



xx (104959.1)



# CINVESTAV

Centro de Investigación y de Estudios Avanzados del IPN  
Unidad Guadalajara

---

**CINVESTAV I.P.N.  
SECCION DE INFORMACION  
Y DOCUMENTACION**

## **MÉTODOS GEOMÉTRICOS PARA LA MANIPULACIÓN DE OBJETOS**

**TESIS QUE PRESENTA  
JAIME SILVERIO ORTEGÓN AGUILAR**

**PARA OBTENER EL GRADO DE  
MAESTRO EN CIENCIAS**

**EN LA ESPECIALIDAD DE  
INGENIERÍA ELÉCTRICA**

Guadalajara, Jal. Agosto de 2002

**C I N V E S T A V  
I P N  
A D Q U I S I C I O N  
D E L I B R O S**

CLASIF.	
ADQUIS.	SSI-220
FECHA:	27 NOV-2002
PROCED.	TESIS-021
	\$

# **MÉTODOS GEOMÉTRICOS PARA LA MANIPULACIÓN DE OBJETOS**

**Tesis de Maestría en Ciencias  
Ingeniería Eléctrica**

**Por:**

**Jaime Silverio Ortegón Aguilar**

**Lic. en Ciencias de la Computación  
Universidad Autónoma de Yucatán**

**Becario de CONACYT, expediente No. 157868**

**Director de Tesis:  
Dr. Eduardo José Bayro Corrochano.**

**CINVESTAV del IPN Unidad Guadalajara, Junio del 2001**

# ***Agradecimientos***

*Agradezco a Dios, a mis padres, a mi asesor, a mis amigos,*

*pero sobre todo a mi esposa Amy.*

*Un agradecimiento especial al Conacyt.*

# Índice General

<b>Agradecimientos.</b>	<b>i</b>
<b>Índice General.</b>	<b>ii</b>
<b>Tabla de Simbolos.</b>	<b>v</b>
<b>Índice de Figuras.</b>	<b>vi</b>
<b>1 Introducción.</b>	<b>1</b>
1.1 Motivación.	1
1.2 Estado del arte.	2
1.3 Objetivo.	3
1.4 Organización y notación.	3
<b>2 Geometría Proyectiva.</b>	<b>4</b>
2.1 Espacio proyectivo.	4
2.1.1 El plano proyectivo 2D.	4
2.1.2 Geometría proyectiva 3D.	6
2.1.3 Generalización a $P^n$ .	8
2.2 Modelos de cámaras	8
2.2.1 El modelo básico: la cámara 'Pinhole'.	8
2.2.2 Calibración.	10
2.2.3 Proyección y retroproyección.	11
2.3 Sistemas estereoscópicos	12
2.3.1 Triangulación.	12

<b>3</b>	<b>Álgebra de Incidencia y Distancia Dirigida.</b>	<b>14</b>
3.1	Álgebra geométrica.	14
3.1.1	Definiciones básicas.	14
3.1.2	Geometría del espacio n-dimensional.	15
3.1.3	Rotores.	16
3.2	Los conos nulos, la horófera y el plano afino n-dimensional.	17
3.2.1	La horófera y el plano afino n-dimensional.	18
3.3	Distancia dirigida, álgebra de incidencia y movimiento rígido.	20
3.3.1	Distancia dirigida de objetos paralelos.	20
3.3.2	Álgebra de incidencia.	21
3.3.3	Relaciones de incidencia en el plano afino 3D.	22
3.3.4	Indicadores geométricos.	23
3.3.5	Movimiento rígido en el plano afino. .	23
<b>4</b>	<b>Votación Tensorial .</b>	<b>25</b>
4.1	Introducción.	25
4.2	Tensores.	25
4.2.1	Interpretación geométrica.	26
4.3	Votación tensorial	27
4.3.1	El proceso de votación.	27
4.3.2	El campo fundamental 2D.	28
4.3.3	Campos de votación 3D.	29
4.3.4	Ilustración	30
4.4	Ejemplos de aplicaciones.	31
4.4.1	Extracción de características sobresalientes.	31
4.4.2	Posibles aplicaciones.	32
<b>5</b>	<b>Reconstrucción y Planeación usando n Vistas.</b>	<b>35</b>
5.1	Reconstrucción de vistas.	36
5.1.1	Obtención de las vistas.	36
5.1.2	Retroproyección.	36



5.2	Fusión de vistas. .	37
5.2.1	Estimación de movimiento rígido.	37
5.2.2	Fusión de vistas.	39
5.3	Planeación.	39
5.3.1	Obtención de primitivas geométricas.	40
5.3.2	Reglas de decisión.	41
5.3.3	Ejecución.	42
6	Conclusiones.	45
6.1	Resumen.	45
6.2	Trabajo futuro.	46
	<b>Bibliografía.</b>	<b>47</b>

Símbolo	Interpretación
$x$	punto en 2D
$X$	punto en 3D
$R$	matriz de rotación o rotor
$t$	matriz de traslación o trasladador
$H$	homografía
$P$	matriz de cámara
$l_{\infty}$	línea en el infinito
$\pi_{\infty}$	plano en el infinito
$\Omega_{\infty}$	cónica en el plano en el infinito
$\omega_{\infty}$	imagen de la cónica en el plano en el infinito
$M$	motor
$\tilde{M}$	motor conjugado l
$\sigma_i$	vector base de alguna álgebra l
$\eta_i, \sigma_j$	bases del plano afino l l
$\mathcal{G}_{4,1,0}$	álgebra geométrica del plano afino
$I$	pseudoescalar del álgebra
$\bar{I}$	pseudoescalar recíproco del álgebra
$\cup$	operador unión "join"
$\cap$	operador intersección "meet"
$V_F$	campo de votación fundamental 2D
$V_S$	campo de votación del tensor $S$ en 3D
$V_P$	campo de votación del tensor $P$ en 3D
$V_B$	campo de votación del tensor $B$ en 3D

Tabla 1: Símbolos que se utilizan en esta tesis.

# Índice de Figuras

2.1	Una línea se puede representar como la unión de dos puntos o la intersección de dos planos.	7
2.2	Los puntos del mundo se proyectan al plano de la imagen.	9
2.3	Un cabezal binocular.	13
2.4	Punto 3D estimado a partir de puntos en las imágenes .	13
3.1	a) El área dirigida o bivector. b) El volumen dirigido o trivector.	15
3.2	La horósfera de $\mathbf{R}^2$ con triángulos proyectados en el plano afino 2D.	19
3.3	a) Línea en es espacio afino 2D. b) Plano en el espacio afino 2D.	21
4.1	a) Representación Geométrica de un tensor en 2D como una elipse. b) Representación en 3D como un elipsoide.	26
4.2	a) Tensor $B$ . b) Tensor $P$ . c) Tensor $S$ .	27
4.3	Ilustración del proceso de votación (en 2D). Primeramente se alinean los campos de votación y despues se suman las contribuciones.	28
4.4	El campo fundamental 2D. a) Versión tangente. b) Versión normal.	29
4.5	Campos de Votación 3D. a) Tensor $V_S$ , muestra del proceso de obtención y corte mostrando su dirección. b) Tensor $V_P$ , corte del tensor con la dirección $\hat{e}_3$ normal a la página.	30
4.6	Resultados de la extracción de características sobresalientes. a) Puntos de entrada. b) Superficie generada.	33
4.7	Datos a fusionar con la votación tensorial. a) Datos obtenidos mediante una unidad Laser. b) Datos obtenidos con sonares.	34
5.1	Cámara binocular montada sobre una unidad pan-tilt.	35
5.2	Imágenes obtenidas con las cámaras. a) Imagen izquierda. b) Imagen derecha.	36
5.3	Imágenes con las esquinas detectadas. a) Imagen izquierda. b) Imagen derecha.	37

<b>5.4</b>	<b>Reconstrucción de las esquinas del cubo de la Figura 5.3.</b>	<b>38</b>
<b>5.5</b>	<b>Correspondencias de puntos entre vistas. a) Imagen izquierda. b) Imagen derecha.</b>	<b>40</b>
<b>5.6</b>	<b>Correspondencias de puntos entre vistas. a) Segunda imagen izquierda. b) Segunda imagen derecha.</b>	<b>40</b>
<b>5.7</b>	<b>Puntos reconstruidos de diferentes vistas.</b>	<b>41</b>
<b>5.8</b>	<b>Reconstrucción con puntos de diferentes vistas ya fusionados.</b>	<b>41</b>
<b>5.9</b>	<b>Robot GEOMETER sujetando un cubo.</b>	<b>42</b>
<b>5.10</b>	<b>Reconstrucción de la pinza y el cubo (ya fusionado).</b>	<b>43</b>
<b>5.11</b>	<b>Planos paralelos encontrados con la regla <math>\phi_i \cap \phi_j \approx l_\infty</math>.</b>	<b>43</b>
<b>5.12</b>	<b>El plano de la pinza se alineará con el plano del cubo.</b>	<b>43</b>
<b>5.13</b>	<b>Trayectoria que deberá seguir la pinza para sujetar el cubo.</b>	<b>44</b>

# Capítulo 1

## Introducción.

Un sistema de visión computacional debe generar una descripción compacta útil de una escena a partir de una o más vistas. Debido a la naturaleza proyectiva, al obtener las imágenes del espacio visual, este problema no tiene restricciones muy bien delimitadas. Los investigadores en visión computacional se enfrentan al problema de obtener y/o generar suficiente información (descripción de la escena), con información pobre y de poca calidad en comparación con la realidad (vistas, imágenes). Además este proceso se ve complicado debido al ruido del que pueden ser objeto las lecturas de la escena. En esencia, la meta de la investigación en visión computacional es formular un método computacional que obtenga una descripción de la escena en términos de geometría y movimiento con información de una o más observaciones de una escena (imágenes obtenidas con cámaras, laser, etc.).

Esta tesis se basa en los conceptos relacionados con la geometría proyectiva, el álgebra geométrica y la votación tensorial. Estas áreas al ser combinadas proveen un marco computacional realmente poderoso para lidiar con transformaciones proyectivas, movimiento rígido, álgebra de incidencia, distancia dirigida y razonamiento en el sentido geométrico. El enfoque que se presenta parece prometedor para el desarrollo de una maquinaria para razonamiento geométrico útil para sistemas que operan en el ciclo de percepción-acción.

La organización de este capítulo es como sigue: la primera sección explica la motivación de este trabajo. Seguidamente se presenta el estado del arte del problema que aquí se trata. En la tercera sección se presenta una descripción del problema y los objetivos de este trabajo. Finalmente se presenta una descripción de este documento y de la notación que se usará en el mismo.

### 1.1 Motivación.

En la actualidad existen diversos métodos para hacer la reconstrucción de datos 3D, la mayoría de los cuales son demandantes de muchos recursos computacionales. La mayoría de los métodos de reconstrucción lo hacen con datos densos, o no reconstruyen en términos de entidades geométricas, como líneas y planos, a partir de datos reales. Al no reconstruir en términos de entidades geométricas, es difícil que se pueda reconocer la estructura de la escena y es aun más difícil describir el movimiento de objetos en la escena. A pesar de que existen muchos trabajos referentes a

la fusión de datos provenientes de diferentes observaciones de la escena, estos métodos no tienen una representación simple o es difícil expresar las transformaciones como trayectorias suaves.

Los investigadores que utilizan el álgebra geométrica normalmente lo hacen para trabajar en problemas de matemáticas o física. Existen algunos trabajos en visión computacional, por ejemplo [Bayro96, Lasenby99], sin embargo estos no usan el plano afino que da un mayor poder para el trabajo con entidades geométricas y sus relaciones.

La votación tensorial ha sido utilizada en visión computacional para la segmentación de imágenes, estimación de flujo óptico, sin embargo es posible utilizar el concepto para la fusión de datos provenientes de varias fuentes, así como para la validación de restricciones geométricas de objetos en la reconstrucción de la escena. De esta forma la votación tensorial nos permite tener un enfoque que relaciona geometría e incertidumbre cuando estamos abocados a reconstruir e interpretar el mundo visual.

## 1.2 Estado del arte.

Desde mediados de los años 80 la mayoría de la literatura de visión computacional discute la geometría y los invariantes ha usado el lenguaje de la geometría proyectiva. Durante la pasada década ha habido un gran avance en el entendimiento y modelado de la geometría de múltiples vistas en visión computacional. La teoría y práctica ha alcanzado un nivel de madurez con la que se pueden conseguir excelentes resultados para problemas que no tenían solución hace varios años, o que incluso se pensaban irresolubles. Los avances en la geometría de múltiples vistas han sido posibles gracias al desarrollo en el entendimiento de los problemas y las mejoras en la estimación matemática de los objetos en las imágenes. Entre los problemas que se han resuelto se encuentran:

- Calcular la geometría epipolar de un cabezal estéreo, la geometría trifocal de un arreglo trifocal, sin ningún objeto de calibración.
- Calcular la calibración interna de la cámara con una secuencia de imágenes.
- La estimación de los tensores multifocales con correspondencias entre puntos de las imágenes, particularmente la matriz fundamental, y el tensor trifocal.
- La extracción de las matrices de las cámaras a partir de los tensores, y subsecuente reconstrucción con 2, 3 o más vistas.

El álgebra geométrica ha sido aplicada con éxito en diferentes campos de la matemática, física, e ingeniería. En la matemática y física se ha aplicado en campos como la relatividad general, física de electrones, partículas fundamentales; gravedad e hipergravedad. En visión computacional ha sido aplicada con éxito en la búsqueda y representación de propiedades geométricas de un objeto que permanecen invariantes bajo cualquier cambio en los parámetros de observación, tales como calibración de la cámara o posición de la misma. Entre dichas invariantes se encuentran el "cross-ratio". En algunas otras áreas como la robótica, ampliamente relacionada con la visión computacional, se ha aplicado en el estudio del movimiento de objetos rígidos. En años más recientes los investigadores han utilizado el álgebra geométrica para la calibración mano-ojo (hand-eye calibration) y para la expresión de restricciones para la navegación.

Guy y Medioni [Guy96] propusieron una metodología llamada *votación vectorial* para atacar el problema de la inferencia en 2D y 3D. Este método hace uso de restricciones perceptuales para la segmentación de imágenes y reconstrucción de superficies. Esto se logra con un proceso parecido a la convolución, donde cada entrada es alineada con una máscara vectorial predefinida. El concepto base de la votación tensorial fue presentado primeramente por Westin [Westin94] y recientemente extendido por Lee y Medioni [Lee98] al combinarlo con la metodología de Guy en lo que llamaron *votación tensorial*. Este trabajo combina el poder de la representación tensorial (introducido por Knutsson [Knutsson95] al contexto de la visión computacional) y la eficiencia computacional de la votación. Entre los últimos avances en la votación tensorial se encuentra su extensión a un marco n-dimensional y su utilización para la estimación de la geometría epipolar.

### 1.3 Objetivo.

El problema que se considerará en este trabajo es el reconstruir entidades geométricas usando datos reales obtenidos principalmente con una cámara binocular, reconocer estructura y cinemática de dichas entidades, y finalmente poder controlar un manipulador robótico montado en un robot móvil para la navegación en un ambiente o ejecución de tareas simples. Los objetivos de esta tesis son:

- La obtención de imágenes mediante una cámara binocular,
- La reconstrucción de entidades geométricas, y su representación con álgebras geométricas,
- La estimación de la cinemática de dichas entidades, mediante el uso del álgebra geométrica,
- La planeación y ejecución de tareas con un robot.

### 1.4 Organización y notación.

Esta tesis está organizada de la siguiente manera: El presente capítulo presentó la motivación y los objetivos que persigue nuestro trabajo. En el capítulo 2, se revisará la geometría proyectiva, así como los modelos de las cámaras. Las álgebras geométricas, además de los conceptos del plano afino y la horófera, se describirán en el capítulo 3. En el capítulo 4 se presentará el formalismo de la votación tensorial. Finalmente, en el capítulo 5 se tratará el problema de la reconstrucción y la planeación en 3D. Las conclusiones de nuestro trabajo y el trabajo futuro que se pretende llevar a cabo se presenta en el capítulo 6.

En esta tesis, los escalares se denotarán por letras itálicas, por ejemplo  $I$ , los tensores y vectores en 3D por letras mayúsculas en negritas, por ejemplo  $\mathbf{X}$ . Los vectores en 2D se denotarán con letras minúsculas en negritas, por ejemplo  $\mathbf{x}$ , a excepción de las bases vectoriales que se denotarán con letras minúsculas solamente, por ejemplo  $\sigma$ . El vector unitario de  $e$ , se denotará como  $\hat{e}$ .

## Capítulo 2

# Geometría Proyectiva.

Este capítulo introduce algunos conceptos básicos de la geometría proyectiva que se usarán en esta tesis. Esta introducción se pretende que sea breve pero informativa. Para un estudio exhaustivo refiérase a [Hartley00], [Faugeras9] o el conocido libro sobre geometría proyectiva [Semple79]. La organización del capítulo es como sigue: la primera sección presenta el espacio proyectivo, la segunda inicia con la descripción de las cámaras proyectivas, se presenta la cámara básica (pinhole camera) y luego se describe la calibración de las cámaras proyectivas. Finalmente se presenta una sección que trata las propiedades de un sistema estéreo.

### 2.1 Espacio proyectivo.

Es conveniente iniciar presentando la geometría proyectiva del *plano proyectivo 2D*, dado que así es más fácil visualizar las ideas que se presentan. Luego se presenta la geometría del espacio 3D que es sólo una generalización del caso plano, concluyendo con la generalización para el caso n-dimensional.

#### 2.1.1 El plano proyectivo 2D.

Como es bien sabido, un punto en el plano puede ser representado por un par de coordenadas  $(x,y)$  en  $\mathbf{R}^2$ . Entonces es fácil identificar un plano con  $\mathbf{R}^2$ ; ahora, si consideramos  $\mathbf{R}^2$  como un espacio vectorial, el par  $(x,y)$  es un vector, el cual diremos que está en coordenadas inhomogéneas. Además haremos la distinción entre vectores columna y vectores fila, ya que una matriz puede multiplicarse por la izquierda con un vector fila, y multiplicarse por la derecha con un vector columna. En caso de no mencionarse lo contrario se entenderá por vector un vector columna.



## Puntos, líneas y cónicas.

Una línea se expresa algebraicamente con una ecuación de la forma  $ax + by + c = 0$ , donde  $a, b, c \in \mathbb{R}$ . Eligiendo diferentes valores para  $a, b$  y  $c$  se representa todas las líneas posibles en el plano. Entonces se puede representar una línea como el vector  $(a, b, c)^T$ . Dada la forma de la ecuación es fácil ver que la línea representada por  $(a, b, c)^T$  es la misma que  $k(a, b, c)^T$  para cualquier  $k \neq 0$ . La clase de equivalencia definida por esta propiedad se conoce como *vector homogéneo*. El conjunto de clases de equivalencia de los vectores en  $\mathbb{R}^3 - (0, 0, 0)^T$  forma el espacio proyectivo  $\mathbb{P}^2$ .

Un punto  $x = (x, y)^T$  estará sobre la línea  $l = (a, b, c)^T$ , si y solo si,  $ax + by + c = 0$ , esto se puede escribir en términos del producto interior de vectores como  $(x, y, 1)(a, b, c)^T = (x, y, 1)l = 0$ , es decir, el punto  $(x, y)^T \in \mathbb{R}^2$  se representa como un 3-vector si se añade una coordenada final igual a 1. Debemos notar que la ecuación  $(kx, ky, k)l = 0$  también es válida para cualquier valor  $k \neq 0$ , por lo que se puede considerar que el conjunto de vectores  $k(x, y, 1)^T$  con  $k \neq 0$  representa al punto  $(x, y)^T \in \mathbb{R}^2$ . En general un vector de la forma  $(x_1, x_2, x_3)^T \in \mathbb{P}^2$  representa al punto  $(x_1/x_3, x_2/x_3)^T \in \mathbb{R}^2$ . Si el vector es de la forma  $(x, y, 0)^T$  al intentar encontrar las coordenadas inhomogéneas obtendremos  $(x/0, y/0)^T$ , es decir, no corresponde a un punto finito en  $\mathbb{R}^2$ . Los puntos con la última coordenada  $x_3 = 0$  se conocen como *puntos en el infinito*, o puntos ideales. Nótese que todos estos puntos yacen sobre una misma línea, la *línea en el infinito*, denotada por  $l_\infty = (0, 0, 1)^T$ . La línea en el infinito se puede entender como la línea que contiene todas las direcciones posibles en el plano.

La ecuación de una cónica en coordenadas inhomogéneas es  $ax^2 + bxy + cy^2 + dx + ey + f = 0$ , haciendo el cambio a coordenadas homogéneas se obtiene  $ax_1^2 + bx_1x_2 + cx_2^2 + dx_1x_3 + ex_2x_3 + fx_3^2 = 0$  o en forma matricial  $x^T Cx = 0$ , donde  $x \in \mathbb{P}^2$  y la matriz de coeficientes  $C$  esta dada por

$$C = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}. \quad (2.1)$$

Una línea  $l$  tangente a la cónica  $C$  satisface la ecuación  $l^T C^* l = 0$ . La notación  $C^*$  indica que se trata de la matriz adjunta de  $C$ . A la matriz  $C^*$  se le conoce como la cónica dual.

## Transformaciones proyectivas.

Una proyectividad es un mapeo invertible  $h : \mathbb{P}^2 \rightarrow \mathbb{P}^2$  tal que 3 puntos  $x_1, x_2$  y  $x_3$  están sobre la misma línea (son colineales), si y solo si,  $h(x_1), h(x_2)$  y  $h(x_3)$  lo están también. A las proyectividades también se les conoce como *Colineaciones*, *Transformaciones Proyectivas*, u *homografías*. Una homografía en  $\mathbb{P}^2$  se puede representar como una matriz de  $3 \times 3$  tal que  $h(x) = Hx$ , es decir,

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (2.2)$$

o más brevemente como  $\mathbf{x}' = \mathbf{H}\mathbf{x}$ .

Existen un tipo especial de homografías, las similaridades, que mantienen la forma del objeto cambiando tan sólo su orientación, posición y tamaño. Las similaridades son importantes por que existen dos puntos sobre  $\mathbf{L}_\infty$  que son invariantes bajo cualquier similaridad, estos puntos son llamados *puntos circulares* o puntos absolutos,  $\mathbf{I}$ ,  $\mathbf{J}$ , con coordenadas canónicas  $\mathbf{I} = (1, i, 0)^T$  y  $\mathbf{J} = (1, -i, 0)^T$ . Una similaridad  $\mathbf{H}$  tiene la siguiente representación matricial:

$$\begin{aligned} \mathbf{x}' &= \mathbf{H}\mathbf{x} \\ \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} &= \begin{bmatrix} s \cos \theta & -s \sin \theta & t_x \\ s \sin \theta & s \cos \theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned} \quad (2.3)$$

Una homografía  $\mathbf{H}$  es una matriz homogénea, o lo que es lo mismo, tiene tan solo 8 grados de libertad. Bajo una homografía de puntos  $\mathbf{x}' = \mathbf{H}\mathbf{x}$ , una línea se transforma como  $\mathbf{l}' = \mathbf{H}^{-T}\mathbf{l}$ , o de manera equivalente,  $\mathbf{l}'^T = \mathbf{l}^T \mathbf{H}^{-1}$ . De manera similar una cónica se transforma como  $\mathbf{C}' = \mathbf{H}^{-T}\mathbf{C}\mathbf{H}^{-1}$  y la cónica dual como  $\mathbf{C}^{*'} = \mathbf{H}\mathbf{C}^*\mathbf{H}^T$

## 2.1.2 Geometría proyectiva 3D.

La mayoría de las entidades geométricas del espacio proyectivo 3D o  $\mathbf{P}^3$ , y sus propiedades, son generalizaciones de las pertenecientes al plano proyectivo 2D. Un punto  $X$  en 3D se representa en coordenadas homogéneas como un 4-vector, es decir, el vector  $X = (x_1, x_2, x_3, x_4)^T$  con  $x_4 \neq 0$  representa al punto  $(X, Y, Z)^T \in \mathbf{R}^3$  con coordenadas inhomogéneas  $X = x_1/x_4$ ,  $Y = x_2/x_4$ ,  $Z = x_3/x_4$ .

### Planos.

En el espacio proyectivo 3D es posible representar planos. En el espacio euclidiano  $\mathbf{R}^3$  un plano  $\pi$  se representa con la siguiente ecuación  $\pi_1 X + \pi_2 Y + \pi_3 Z + \pi_4 = 0$ , en el espacio proyectivo se hace con el 4-vector  $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)^T$ . Los primeros tres elementos representan la normal del plano, mientras que el cuarto representa la distancia del plano al origen, o *distancia de Hesse*. Haciendo un razonamiento parecido al hecho en el plano proyectivo, podemos decir que un punto, en coordenadas homogéneas, se encuentra en un plano si cumple  $\pi_1 x_1 + \pi_2 x_2 + \pi_3 x_3 + \pi_4 x_4 = 0$ , o  $\pi^T X = 0$ . Así como en el plano proyectivo  $\mathbf{P}^2$  existe la línea en el infinito, en  $\mathbf{P}^3$  existe el *plano en el infinito*, que es el plano  $\pi_\infty$  con coordenadas canónicas

$$\pi_\infty = (0, 0, 0, 1)^T \quad (2.4)$$

## Líneas.

En el plano proyectivo las líneas tienen una representación natural, en el espacio proyectivo no sucede de la misma manera, ya que en él, las líneas se representan como la unión de dos puntos o como la intersección de dos planos, ver figura 2.1. La representación más usual de las líneas es mediante las coordenadas de Plücker, definidas a partir de una matriz homogénea antisimétrica de 4x4 elementos. La línea que une los puntos  $A$  y  $B$  se representa como la matriz  $L$  con elementos  $l_{ij} = A_i B_j - B_i A_j$  o en notación vectorial  $L = AB^T - BA^T$ . Aquí los elementos  $l_{12}, l_{13}, l_{23}$  representan la dirección de la línea, mientras que los elementos  $l_{14}, l_{42}, l_{34}$  representan el momento, es decir el plano donde yace la línea. Nótese que se usa el elemento  $l_{42}$  ya que esto elimina elementos negativos de las ecuaciones relacionadas con el elemento  $l_{24}$ . La línea formada por la intersección de dos planos  $P$  y  $Q$  se representa como la matriz  $L^*$  con elementos  $l_{ij}^* = P_i Q_j - Q_i P_j$ , o en notación vectorial  $L^* = PQ^T - QP^T$ . La matriz  $L^*$  se puede obtener directamente de la matriz  $L$  con la siguiente regla de reescritura  $l_{12} : l_{13} : l_{14} : l_{23} : l_{42} : l_{34} = l_{34}^* : l_{42}^* : l_{23}^* : l_{14}^* : l_{13}^* : l_{12}^*$ . La ecuación para comprobar que un punto está sobre una línea es  $L^* X = 0$ .

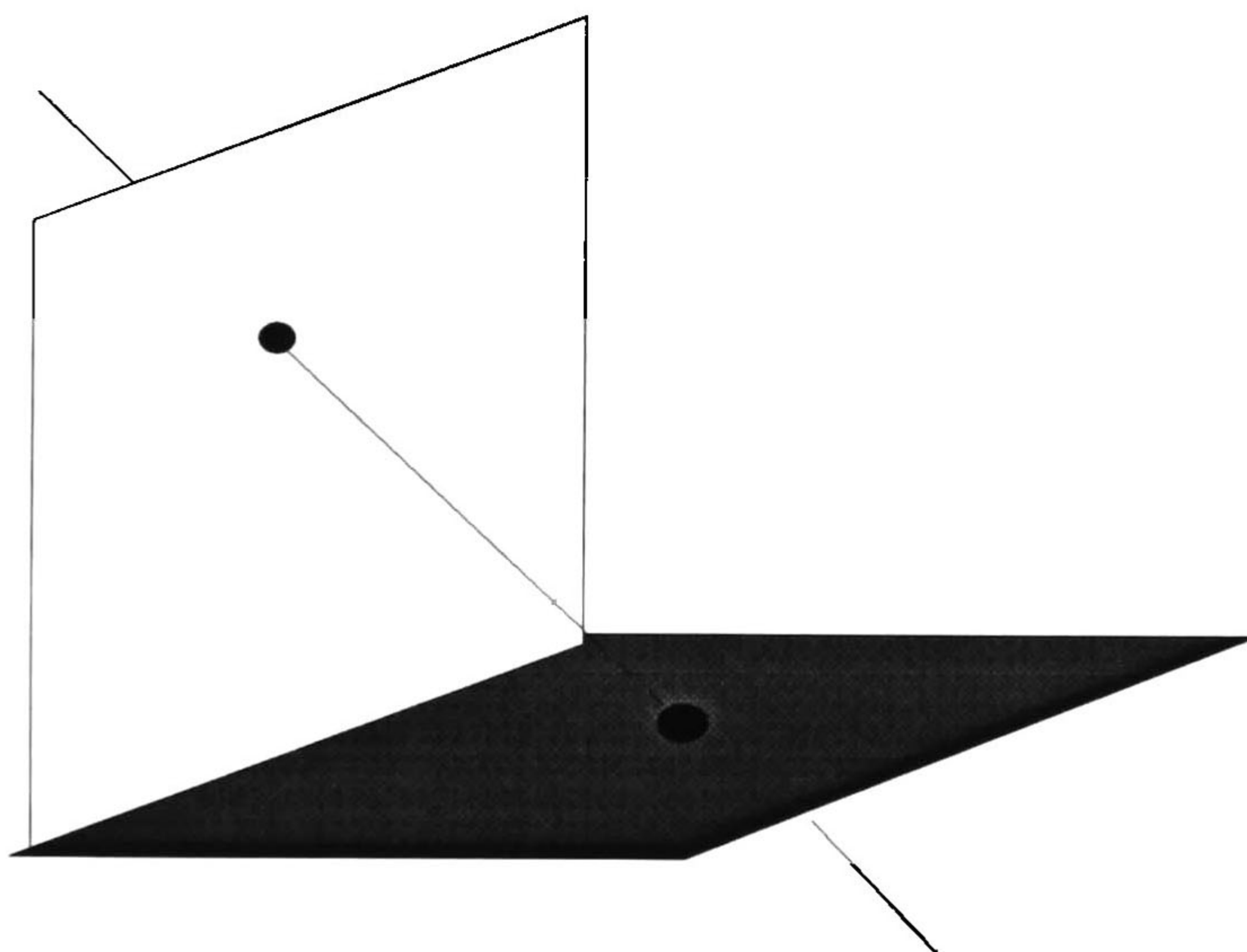


Figura 2.1: Una línea se puede representar como la unión de dos puntos o la intersección de dos planos.

## La Cónica absoluta.

La cónica absoluta,  $\Omega_{\infty}$ , es una cónica sobre  $\pi_{\infty}$ . La cónica absoluta tiene una matriz de coeficientes  $C=I$ , por lo que es una cónica de puntos imaginarios solamente. Algunas propiedades que posee esta cónica son: todos los círculos intersectan  $\Omega_{\infty}$  en dos puntos, y todas las esferas intersectan  $\pi_{\infty}$  en  $\Omega_{\infty}$ .

## Transformaciones proyectivas.

En el espacio proyectivo, un punto  $X$  se transforma con una homografía, una matriz homogénea  $H$  de 4x4 elementos, como  $X' = HX$ . Bajo una homografía  $H$  de transformación de puntos, los planos se transforman como  $\pi' = H^{-T}\pi$ . Finalmente las líneas se transforman como  $L = HLH^T$  y  $L^* = H^{-T}LH^{-1}$ .

En el espacio proyectivo, al igual que en el plano proyectivo, existen *puntos en el infinito*, representados con un 4-vector con el cuarto elemento igual a cero,  $X = (x_1, x_2, x_3, 0)^T$ . Dada la ecuación  $\pi^T X = 0$ , podemos deducir que el *plano en el infinito* se representa con un 4-vector de la forma  $(0, 0, 0, \pi_4)^T$  como se menciona en la ecuación 2.4. Igualmente se puede deducir que una *línea en el infinito* se representa con una matriz de coordenadas de Plücker con los elementos de la dirección igual a cero.

### 2.1.3 Generalización a $P^n$ .

Un punto en un *espacio proyectivo*  $P^n$  puede ser representado con vectores columna de  $n+1$  componentes

$$X = (x_1, \dots, x_{n+1})^T \in R^{n+1} - (0, \dots, 0)^T \quad (2.5)$$

, como en los anteriores casos los vectores  $(x_1, \dots, x_{n+1})^T$  y  $\lambda(x_1, \dots, x_{n+1})^T$  representan el mismo punto de  $P^n$  para toda  $\lambda \neq 0$ . Las  $x_i$ 's son llamadas *las coordenadas homogéneas* del punto proyectivo. En este espacio existe lo que es llamado el *hiperplano* que es equivalente al plano del espacio proyectivo 3D, que se representa con el vector  $A = (a_1, \dots, a_{n+1})^T$ . Los puntos  $X$  que se encuentran sobre el hiperplano  $A$  cumplen con la ecuación  $x_1 a_1 + \dots x_{n+1} a_{n+1} = 0$ , o en forma matricial  $A^T X = 0$ .

## 2.2 Modelos de cámaras

### 2.2.1 El modelo básico: la cámara 'Pinhole'.

Este modelo es el más sencillo de las cámaras proyectivas. Una cámara se puede entender como un mapeo del mundo real en 3D a un plano 2D, el plano de la imagen. Se puede tomar el origen como el centro de proyección y el plano  $Z=f$  como el plano de proyección, el plano de la imagen, ver figura 2.2. El centro de proyección también es conocido como *centro de la cámara*, o centro óptico. La línea perpendicular al plano de la imagen que pasa por el centro de la cámara es conocida como el *eje principal* de la cámara, el punto donde el eje principal interseca el plano de la imagen se conoce como *punto principal* y el plano paralelo al plano de la imagen que pasa por el centro de la cámara se conoce como el *plano principal* de la cámara.

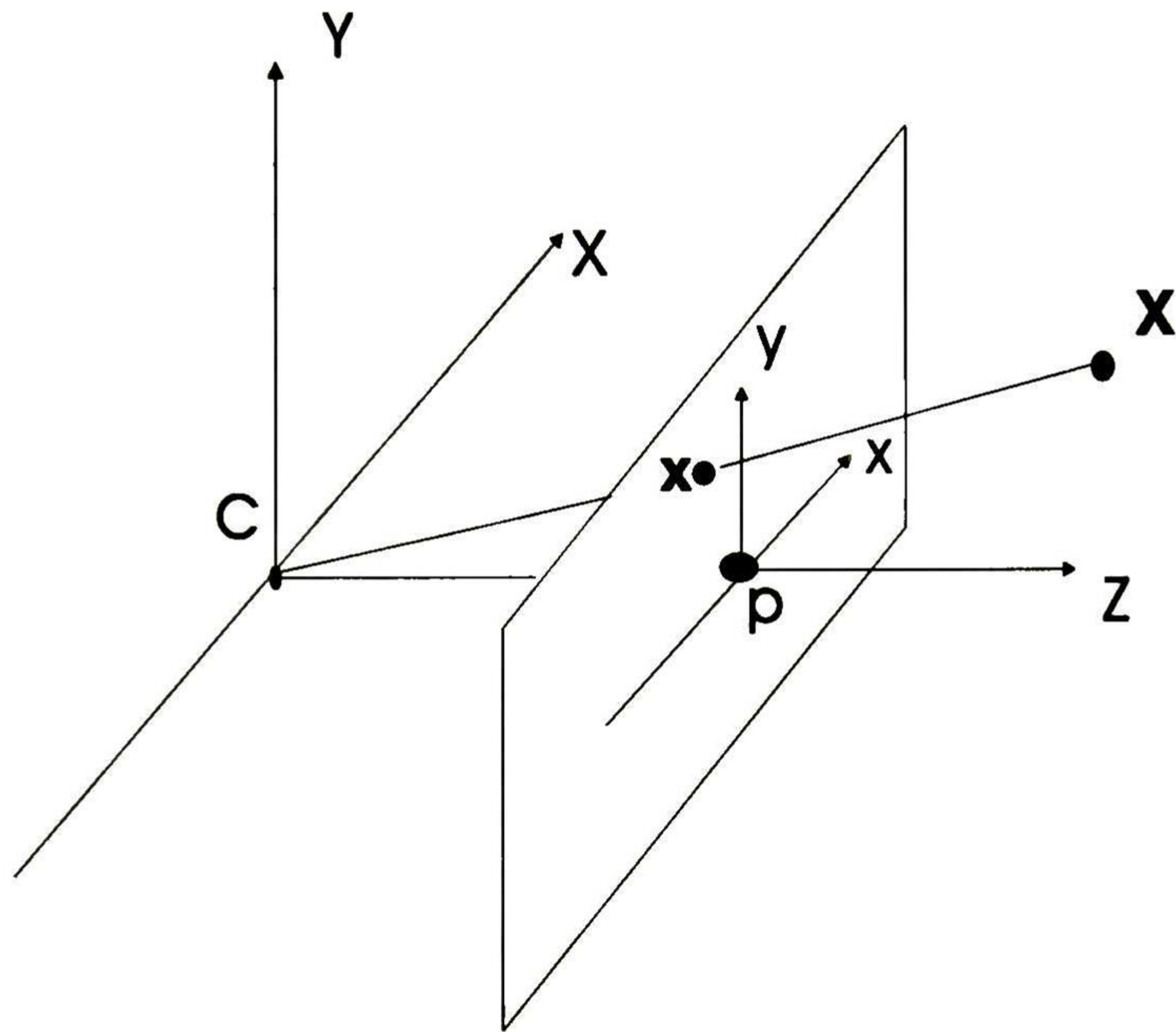


Figura 2.2: Los puntos del mundo se proyectan al plano de la imagen.

En este sencillo caso el punto 3D  $(X, Y, Z, 1)^T$  se mapearía al punto  $(fX/Z, fY/Z, 1)^T$ , este mapeo se representa con la transformación

$$\begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.6)$$

El caso ideal de que el centro de la cámara se encuentre en el origen del sistema de coordenadas es realmente difícil de encontrar en la práctica, por lo que es necesario introducir algunos elementos extras a la transformación o proyección. Estos elementos extras se introducirán como la multiplicación de dos matrices  $P = K[R|t]$ , la matriz  $K$  se conoce como la matriz de *parámetros internos* de la cámara. Esta matriz representa la transformación que llevará el punto principal al origen del plano de la imagen,  $x_0$  y  $y_0$ . También aquella que tomará en cuenta la distancia focal,  $f$ , y la distorsión generada por la relación entre la altura y el ancho de los pixeles de la imagen,  $m_x$  y  $m_y$ . Finalmente además tendrá en consideración la inclinación que pudiera tener el plano de la imagen con respecto al sistema de coordenadas,  $s$ . Esta matriz  $K$  es de la forma

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2.7)$$

donde  $\alpha_x = fm_x$  y  $\alpha_y = fm_y$ .

La matriz  $[R|t]$  es conocida como la matriz de *parámetros externos* de la cámara ya que estos parámetros moverán el centro de la cámara  $\tilde{C}$ , a  $t = R\tilde{C}$ , y alinearán el plano de la imagen para que sea perpendicular al eje z mediante una rotación  $R$ . Entonces la cámara queda completamente descrita con la siguiente matriz

$$P = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} & | & t_1 \\ R_{21} & R_{22} & R_{23} & | & t_2 \\ R_{31} & R_{32} & R_{33} & | & t_3 \end{bmatrix} = K[R|t], \quad (2.8)$$

donde  $t = -R\tilde{C}$ .

## 2.2.2 Calibración.

El proceso de determinar la matriz  $K$  se conoce como proceso de calibración. Una de las herramientas que se usan para la calibración es la imagen de la cónica absoluta  $\Omega_\infty$ ,  $\omega$ . Esto es porque dicha imagen es  $\omega = (KK^T)^{-1} = K^{-T}K^{-1}$ . Al igual que  $\Omega_\infty$ , la cónica  $\omega$  es una cónica de puntos imaginarios, sin ningún punto real. Un ejemplo simple de calibración es el siguiente: La imagen de tres cuadrados (en planos no paralelos) da suficiente información para determinar  $K$ . Siguiendo los siguientes pasos para cada uno de los cuadrados se puede obtener la matriz de calibración  $K$ .

- Considere uno de los cuadrados, las correspondencias entre las cuatro esquinas definen una homografía  $H$  entre el plano en cuestión,  $\pi$ , y el plano de la imagen.
- Aplicando esta homografía a los puntos circulares en  $\pi$  se determinan sus imágenes como  $H(1, \pm i, 0)^T$ , estos dos puntos se encuentran en  $\omega$ .

Después de aplicar a los tres cuadros los pasos anteriores se obtienen 6 puntos sobre  $\omega$ , suficientes para determinarla. Finalmente se obtiene  $K$  de  $\omega = (KK^T)^{-1}$  mediante una factorización de Cholesky.

En la práctica no sólo se determina la matriz  $K$ , sino se determina la matriz  $P$  completa, es decir, se calculan los parámetros internos y externos. Esto se puede hacer conociendo cierto número  $n$  de correspondencias  $X_i \leftrightarrow x_i$  entre puntos en 3D  $X_i$  y puntos en la imagen  $x_i$ . A partir de cada correspondencia se obtiene un sistema de ecuaciones

$$\begin{bmatrix} 0^T & -w_i X_i^T & y_i X_i^T \\ w_i X_i^T & 0^T & -x_i X_i^T \\ -y_i X_i^T & x_i X_i^T & 0^T \end{bmatrix} \begin{pmatrix} P^1 \\ P^2 \\ P^3 \end{pmatrix} = 0, \quad (2.9)$$

donde cada  $P^{iT}$  es un 4-vector, la  $i$ -ésima fila de  $P$ , y  $x_i = (x_i, y_i, w_i)^T$ . En la práctica tan sólo se usan las dos primeras ecuaciones ya que las tres son linealmente dependientes. Entonces, de un conjunto de  $n$  correspondencias obtenemos una matriz  $A$  de  $2n \times 12$ . La matriz  $P$  se obtiene resolviendo el sistema  $Ap=0$ , donde  $p$  es el vector con las entradas de  $P$ .

### 2.2.3 Proyección y retroproyección.

La proyección de un punto en 3D es el mapeo mediante la matriz  $P$  del punto 3D en cuestión a un punto sobre el plano de la imagen. La retroproyección de un punto en la imagen es la determinación del conjunto de puntos en 3D que se mapean a dicho punto en la imagen. También es posible hablar de la proyección y la retroproyección de líneas y cónicas.

#### Puntos.

Un punto  $X$  en 3D se proyecta con la cámara  $P$  al plano de la imagen como

$$\mathbf{x} = P\mathbf{X}. \quad (2.10)$$

La retroproyección de un punto es un rayo en el espacio que pasa a través del centro de la imagen. Existen varias formas de representar dicho rayo, aquí se usa la representación mediante la suma de dos puntos. Nótese que se conocen dos puntos sobre el rayo, el centro de la cámara  $C$  (el punto tal que  $PC=0$ ) y el punto  $P^+\mathbf{x}$ , donde  $P^+$  es la matriz pseudoinversa de  $P$ . La pseudoinversa de  $P$  es la matriz

$$P^+ = P^T (PP^T)^{-1}, \quad (2.11)$$

para la cual  $PP^+ = I$ . Es fácil verificar que el punto  $P^+\mathbf{x}$  cae sobre el rayo,  $P(P^+\mathbf{x}) = I\mathbf{x} = \mathbf{x}$ . Entonces el rayo retroproyectado es

$$X(\lambda) = P^+\mathbf{x} + \lambda C. \quad (2.12)$$

En el caso de cámaras finitas podemos deducir otra expresión ya que el rayo corta al plano en el infinito. Si la cámara tiene una matriz  $P = [M|p_4]$ , el centro de la cámara está dado por

$$\tilde{C} = -M^{-1}p_4 \quad (2.13)$$

y el rayo retroproyectado desde un punto  $\mathbf{x}$  en la imagen intersecará al plano en el infinito en el punto

$$D = ((M^{-1}\mathbf{x})^T, 0)^T \quad (2.14)$$

Entonces el rayo será

$$\begin{aligned} X(\lambda) &= \lambda \begin{pmatrix} M^{-1}\mathbf{x} \\ 0 \end{pmatrix} + \begin{pmatrix} -M^{-1}p_4 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} M^{-1}(\lambda\mathbf{x} - p_4) \\ 1 \end{pmatrix}. \end{aligned} \quad (2.15)$$

## Líneas.

Con la cámara  $P$ , una línea  $L$  en 3D se proyecta al plano de la imagen como

$$[l]_x = PLP^T \quad (2.16)$$

Esto se puede ver claramente si consideramos dos puntos en la imagen  $a = PA$  y  $b = PB$ . Como se presentó anteriormente, la línea que pasa por los puntos  $A$ ,  $B$ , en el espacio 3D es  $L = AB^T - BA^T$ . Entonces la matriz  $E = PLP^T = ab^T - ba^T$  es una matriz antisimétrica, con espacio nulo  $a \times b$ . Por lo tanto  $E = [a \times b]_x$ , y en 2D la línea  $l$  que pasa por los puntos,  $a$  y  $b$ , es  $l = a \times b$ .

La retroproyección de una línea  $l$ , con la cámara  $P$ , es  $P^T l$ . Un punto  $x$  está sobre la línea  $l$ , si y solo si,  $x^T l = 0$ . Un punto  $X$  en el espacio se mapea al punto  $PX$ , que estará en la línea  $l$ , si y solo si,  $X^T P^T l = 0$ . Si consideramos  $P^T l$  como un plano, un punto  $X$ , estará sobre el plano, si y solo si,  $x$  está sobre la línea  $l$ ; en otras palabras la retroproyección de  $l$  es

$$\pi_l = P^T l. \quad (2.17)$$

## Cónicas.

La proyección de una cónica se puede ver como el mapeo de un plano a otro, por lo que es tan solo aplicar la homografía que mapea el plano donde yace la cónica al plano de la imagen. Haciendo un razonamiento similar al de la retroproyección de una línea, es fácil ver que la retroproyección de una cónica, con la cámara  $P$ , es el cono

$$Q_\infty = P^T CP. \quad (2.18)$$

## 2.3 Sistemas estereoscópicos .

Se habla de un sistema estereoscópico cuando se tiene dos cámaras alineadas paralelamente y fijas en un soporte. En el caso de nuestro laboratorio, se cuenta con un cabezal binocular, dos cámaras alineadas horizontalmente, ver la figura 2.3. Las cámaras están calibradas individualmente y la distancia entre ellas es conocida . De los sistemas estereoscópicos se puede mencionar que dos vistas es el número mínimo para hacer una reconstrucción 3D a partir de imágenes. En este trabajo se utilizó las matrices de las cámaras para calcular los rayos que se proyectan a un punto en la imagen, teniendo los rayos se calcularon los puntos más cercanos entre las líneas y se tomó al promedio de ellos como el punto reconstruido en 3D. Los cálculos anteriores se hicieron utilizando el álgebra geométrica de Clifford que se explicará en detalle en el siguiente capítulo.

### 2.3.1 Triangulación.

Un punto 3D se proyectará a puntos correspondientes en ambas imágenes. Dados puntos en dos imágenes, correspondientes entre sí, deberían intersectarse en el espacio los rayos retroproyectados a partir de dichos puntos; sin embargo en la práctica, esto rara vez sucede debido al ruido. A la búsqueda del punto 3D donde se intersectan los rayos se





Figura 2.3: Un cabezal binocular.

le conoce como triangulación. Existen varios métodos para corregir el problema de la intersección, uno de ellos es encontrar nuevos puntos en las imágenes cuyas retroproyecciones si se cruzan en el espacio, esto se hace tratando de minimizar la distancia (error) entre los puntos conocidos y los puntos estimados. En nuestro caso, el punto en el espacio es el punto medio de la línea más corta entre los rayos retroproyectados, ver la figura 2.4. Ver detalles en [Hartley00].

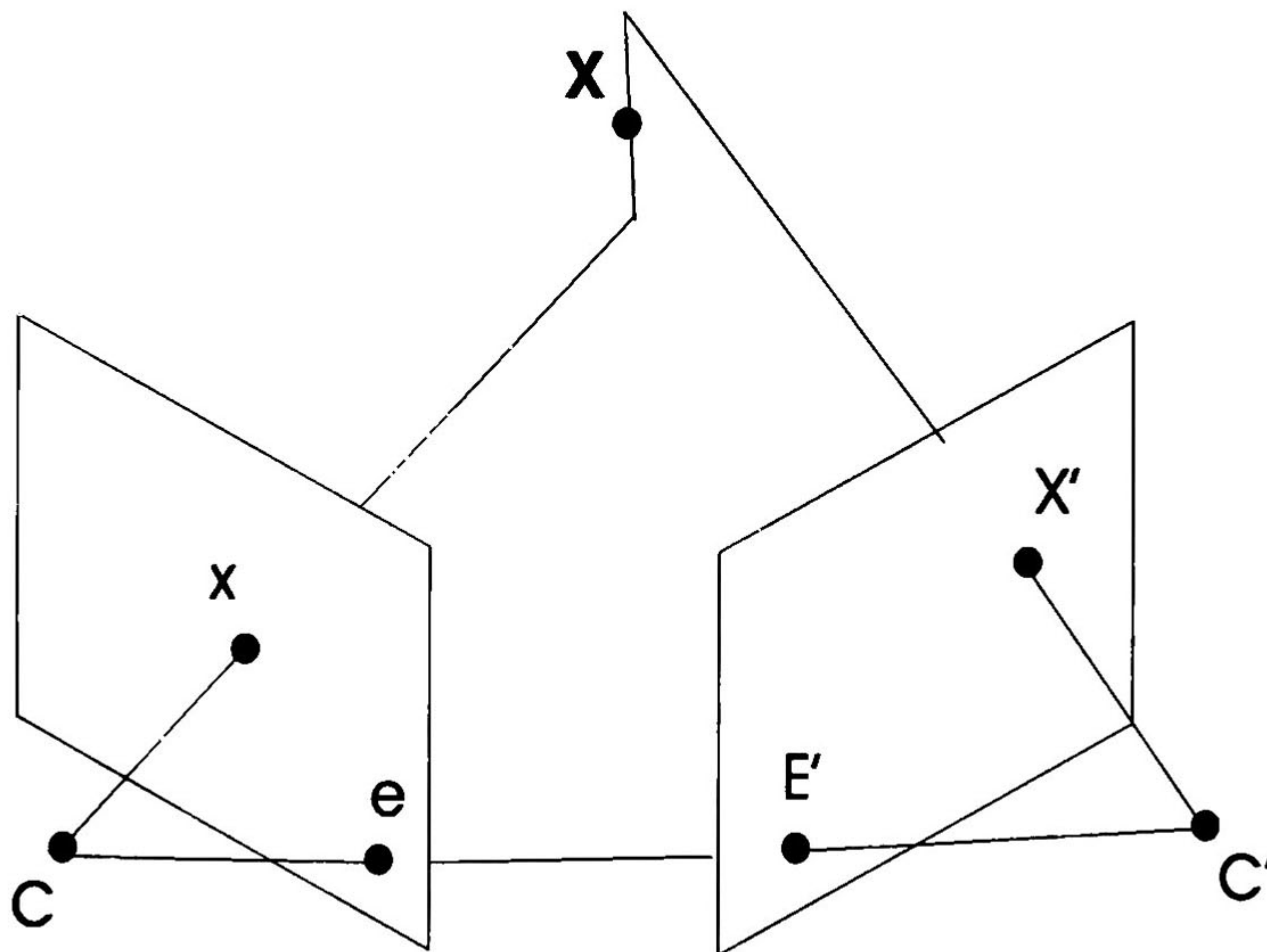


Figura 2.4: Punto 3D estimado a partir de puntos en las imágenes .

Los conceptos presentados en este capítulo se utilizarán en el capítulo 6, donde se presentará un ejemplo práctico de como se pueden combinar esto con el Álgebra Geométrica, que se expondrá en el siguiente capítulo.

## Capítulo 3

# Álgebra de Incidencia y Distancia Dirigida.

En este capítulo se dará una introducción al marco de trabajo llamado el plano afino n-dimensional, el cual permite el uso de las operaciones del álgebra de incidencia y la distancia directa, junto con transformaciones euclidianas representadas como espinores. Se presentará una introducción a: el álgebra geométrica, el álgebra de incidencia y la distancia dirigida; el cálculo geométrico. Finalmente se hará una presentación de como se puede aplicar este razonamiento a problemas prácticos.

### 3.1 Álgebra geométrica.

Las *Álgebras Geométricas* (de Clifford) extienden el sistema de números reales para incluir vectores  $u, v, w, \dots$ , y sus productos  $uv, uvw, \dots$ . Estas álgebras son útiles para modelar geometría, y sus productos vectoriales, representando superficies y objetos de mayores dimensiones, permiten hacer descripciones computacionalmente poderosas de rotaciones, traslaciones y otras transformaciones geométricas. El nombre "álgebra de Clifford" es en honor del matemático inglés William Kingdon Clifford (1845-1879), quien reconoció la importancia de las ideas introducidas por el alemán Hermann Günther Grassmann (1809-1877) y Sir William Rowan Hamilton (1805-1865). Clifford fusionó las ideas de Grassmann con las de Hamilton. Durante la década de 1960, David Hestenes desarrolló un sistema matemático basado en las álgebras de Clifford, el cual es llamado *álgebras geométricas* [Hestenes66]. Los números complejos y los cuaterniones forman dos álgebras geométricas de Clifford particularmente simples. Para mayor información sobre álgebras geométricas de Clifford refiérase a [Lounesto97] y sobre las álgebras geométricas a [Hestenes84]. La mayoría del trabajo aquí presentado es basado en las álgebras geométricas de Clifford.

#### 3.1.1 Definiciones básicas.

Se Denota  $G_n$  al álgebra geométrica de n dimensiones.  $G_n$  es un espacio lineal. Además de la suma de vectores y la multiplicación escalar tiene un producto anticonmutativo que es asociativo y distributivo sobre la adición, el *producto Clifford* o *geométrico*. Una característica especial del álgebra es que cualquier vector al cuadrado es un escalar. El

producto geométrico de dos vectores  $\mathbf{a}$  y  $\mathbf{b}$  es escrito como  $\mathbf{ab}$  y puede ser expresado como la suma de sus partes simétricas y antisimétricas como sigue,

$$\mathbf{ab} = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b}. \tag{3.1}$$

Usando esta definición se puede expresar el producto interior  $\mathbf{a} \cdot \mathbf{b}$  y el producto exterior como  $\mathbf{a} \wedge \mathbf{b}$  en términos de productos geométricos no conmutativos

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= \frac{1}{2}(\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b} + \mathbf{a} \cdot \mathbf{b} - \mathbf{a} \wedge \mathbf{b}) = \frac{1}{2}(\mathbf{ab} + \mathbf{ba}) \\ \mathbf{a} \wedge \mathbf{b} &= \frac{1}{2}(\mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b} - \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b}) = \frac{1}{2}(\mathbf{ab} - \mathbf{ba}) \end{aligned} \tag{3.2}$$

El producto interno de dos vectores es el producto *escalar* o *punto* estándar y da como resultado un escalar. El producto externo o *wedge* de dos vectores es una nueva cantidad llamada *bivector*. Se puede entender un bivector como un área orientada en el plano que contiene  $\mathbf{a}$  y  $\mathbf{b}$ , formado deslizando  $\mathbf{a}$  a lo largo de  $\mathbf{b}$ , ver figura 3.1. Entonces,  $\mathbf{b} \wedge \mathbf{a}$  tendrá la orientación opuesta, mostrando así la propiedad conmutativa que se dió en la ecuación 3.2. El producto externo es fácilmente generalizable a mayores dimensiones; por ejemplo,  $(\mathbf{a} \wedge \mathbf{b}) \wedge \mathbf{c}$  es un trivector que se interpreta como un volumen con orientación: el obtenido deslizando el area  $\mathbf{a} \wedge \mathbf{b}$  a lo largo de  $\mathbf{c}$ . El producto externo de  $k$  vectores es un  $k$ -vector, y se dice que tiene grado  $k$ . Un *multivector* (la combinación lineal de objetos de diferente tipo) es homogéneo si contiene términos de un solo grado. El álgebra geométrica provee una manera de manipular multivectores que permite el seguimiento de objetos de diferente grado simultáneamente.

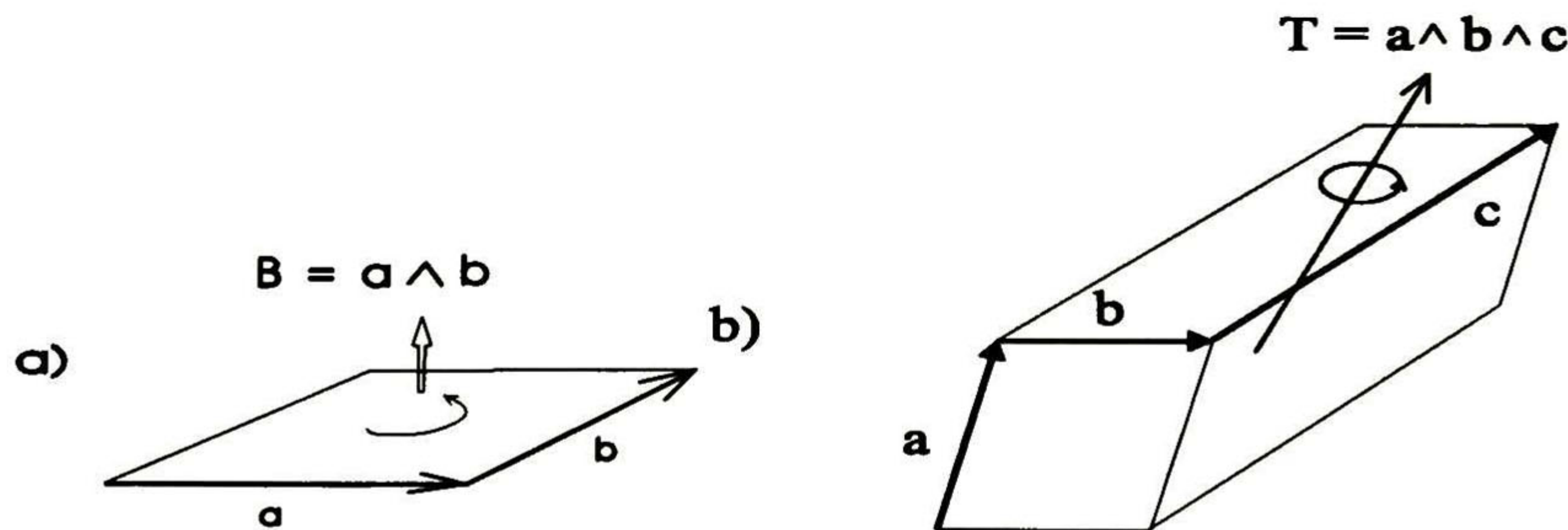


Figura 3.1: a) El área dirigida o bivector. b) El volumen dirigido o trivector.

En un espacio de tres dimensiones se puede construir un trivector  $\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c}$ , pero no 4-vectores ya que éstos no existen (no es posible deslizar el volumen  $\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c}$  en la cuarta dimensión). El elemento vectorial de mayor grado en un espacio es llamado el *pseudoescalar*. El pseudoescalar se denota  $I$  y es crucial cuando se discute la *dualidad*.

### 3.1.2 Geometría del espacio n-dimensional.

En un espacio n-dimensional se pueden introducir bases vectoriales ortonormales  $\{\sigma_i\}$ ,  $i=1, \dots, n$ , tal que  $\sigma_i \cdot \sigma_j = \delta_{ij}$ . Esto nos lleva a las bases para el álgebra completa:

$$1, \{\sigma_i\}, \{\sigma_i \wedge \sigma_j\}, \{\sigma_i \wedge \sigma_j \wedge \sigma_k\}, \dots, \sigma_1 \wedge \sigma_2 \wedge \dots \wedge \sigma_n. \quad (3.3)$$

Cualquier multivector puede ser expresado en términos de estas bases. En este documento se especificará un álgebra geométrica  $\mathcal{G}_n$  del espacio n-dimensional como  $\mathcal{G}_{p,q,r}$ , donde p, q, r son el número de bases vectoriales que elevadas al cuadrado dan como resultado 1, -1 y 0 respectivamente y satisfacen  $n=p+q+r$ . Si una subálgebra es generada tan sólo por las bases de grado par será llamada una subálgebra par y denotada por  $\mathcal{G}_{p,q,r}^+$ .

Considere dos multivectores  $A_r$  y  $B_s$  de grado r y s respectivamente. El producto geométrico de  $A_r$  y  $B_s$  puede ser escrito como

$$A_r B_s = \langle AB \rangle_{r+s} + \langle AB \rangle_{r+s-2} + \dots + \langle AB \rangle_{|r-s|} \quad (3.4)$$

donde  $\langle M \rangle_t$  es usado para denotar la parte de grado t del multivector M.

A continuación un ejemplo simple del producto geométrico. Sea  $A = 2\sigma_1 + 5\sigma_1\sigma_2$  y  $B = 1\sigma_1 + 4\sigma_3$  en  $\mathcal{G}_{3,0,0}$  el producto geométrico será

$$\begin{aligned} AB &= 2\sigma_1\sigma_1 + 8\sigma_1\sigma_3 + 5\sigma_1\sigma_1\sigma_2 + 20\sigma_1\sigma_2\sigma_3 \\ &= 2 + 5\sigma_2 + 8\sigma_1\sigma_3 + 20\sigma_1\sigma_2\sigma_3 \end{aligned} \quad (3.5)$$

note que  $\sigma_i\sigma_i = (\sigma_i)^2 = \sigma_i \cdot \sigma_i = 1$  y  $\sigma_i\sigma_j = \sigma_i \wedge \sigma_j$ , es decir, el producto geométrico de bases iguales es igual a 1 y de bases diferentes es igual a su producto exterior (wedge), el cual para simplificar la notación es omitido. El lector puede realizar esta y muchas más operaciones en álgebra de Clifford usando CLICAL [Lounesto87].

### Álgebra geométrica del espacio 3-D.

El álgebra geométrica que representa el espacio euclideo 3-D es  $\mathcal{G}_{3,0,0}$  y tiene  $2^3 = 8$  elementos base que son:

$$\left\{ \underbrace{1}_{\text{escalar}}, \underbrace{\sigma_1, \sigma_2, \sigma_3}_{\text{vectores}}, \underbrace{\sigma_1\sigma_2, \sigma_2\sigma_3, \sigma_3\sigma_1}_{\text{bivectores}}, \underbrace{\sigma_1\sigma_2\sigma_3 \equiv I}_{\text{trivector}} \right\} \quad (3.6)$$

Es importante mencionar que al elemento de mayor grado del álgebra es llamado pseudoescalar y generalmente denotado como  $I$ .

### 3.1.3 Rotores.

La multiplicación de las tres bases vectoriales  $\sigma_1, \sigma_2$  y  $\sigma_3$  por  $I$  da como resultado las tres bases bivectoriales  $\sigma_1\sigma_2 = I\sigma_3, \sigma_2\sigma_3 = I\sigma_1$  y  $\sigma_3\sigma_1 = I\sigma_2$ . Estos bivectores rotan los vectores en su propio plano en  $90^\circ$ , por ejemplo,  $(\sigma_1\sigma_2)\sigma_2 =$

$\sigma_1, (\sigma_2\sigma_3)\sigma_2 = -\sigma_3$ , etc. Si se identifica el  $i, j, k$  del álgebra de cuaternios con  $I\sigma_1, -I\sigma_2, I\sigma_3$ , las famosas ecuaciones de Hamilton,  $i^2 = j^2 = k^2 = -1$ , se recuperan. Dado que  $i, j, k$  son bivectores, es fácil esperar que representen rotaciones de  $90^\circ$  en direcciones ortogonales y provean un sistema sólido para la representación de rotaciones en 3D generales.

En el álgebra geométrica un *rotor* (del inglés "rotator"),  $\mathbf{R}$ , es un elemento de grado par del álgebra que satisface  $\mathbf{R}\tilde{\mathbf{R}} = 1$ , donde  $\tilde{\mathbf{R}}$  es el conjugado de  $\mathbf{R}$ .

Sea  $\mathbf{A} = \{a_0, a_1, a_2, a_3\} \in \mathcal{G}_{3,0,0}$  un cuaternion unitario, entonces el rotor que describe la misma rotación será

$$\begin{aligned} \mathbf{R} &= \underbrace{a_0}_{\text{escalar}} + \underbrace{a_1(I\sigma_1) - a_2(I\sigma_2) + a_3(I\sigma_3)}_{\text{bivectores}} \\ &= a_0 + a_1\sigma_2\sigma_3 - a_2\sigma_3\sigma_1 + a_3\sigma_3\sigma_1. \end{aligned} \quad (3.7)$$

El álgebra de cuaterniones es entonces un subconjunto del álgebra geométrica del espacio 3D.

El rotor conjugado se calcula como sigue

$$\tilde{\mathbf{R}} = a_0 - a_1(I\sigma_1) + a_2(I\sigma_2) - a_3(I\sigma_3). \quad (3.8)$$

La transformación en términos de un rotor  $\mathbf{R}$  se aplica a un vector  $\mathbf{a}$  como  $\mathbf{a} \mapsto \mathbf{R}\mathbf{a}\tilde{\mathbf{R}} = \mathbf{b}$ . Un rotor funciona de la misma manera con multivectores de cualquier grado y en espacios de cualquier dimensión. Además los rotores se combinan de una manera realmente simple, por ejemplo, un rotor  $\mathbf{R}_1$  seguido de un rotor  $\mathbf{R}_2$  es equivalente a un rotor total  $\mathbf{R}$ , donde  $\mathbf{R} = \mathbf{R}_2\mathbf{R}_1$ .

## 3.2 Los conos nulos, la horósfera y el plano afino n-dimensional.

Los vectores base de los *conos nulos* recíprocos  $N$  y  $\bar{N}$  generan el álgebra  $2^n$ -dimensional  $\mathcal{G}_N = \text{gen}\{e_1, \dots, e_n\}$  y el álgebra  $2^n$ -dimensional recíproca  $\mathcal{G}_{\bar{N}} = \text{gen}\{\bar{e}_1, \dots, \bar{e}_n\}$ , que tienen la estructura de las álgebras de Grassmann. Cualquier elemento de las *bases vectoriales ortogonales nulas*  $\{e_i, \bar{e}_i\}$  del álgebra geométrica  $\mathcal{G}_{n,n}$  elevada al cuadrado da como resultado cero, esta es la razón por la cual se les conoce como *vectores nulos*, además se debe mencionar que el vector  $\bar{e}_i$  es el dual de  $e_i$ .

El álgebra geométrica  $\mathcal{G}_{n,n}$  se construye mediante el producto directo de las dos subálgebras de Grassmann:

$$\mathcal{G}_{n,n} = \mathcal{G}_N \otimes \mathcal{G}_{\bar{N}} = \text{gen}\{e_1, \dots, e_n, \bar{e}_1, \dots, \bar{e}_n\}. \quad (3.9)$$

Cuando  $n$  es infinito contable llamamos  $\mathcal{G}_{\infty,\infty}$  el *álgebra geométrica universal*. El álgebra geométrica universal  $\mathcal{G}_{\infty,\infty}$  contiene todas las álgebras geométricas  $\mathcal{G}_{n,n}$  como subálgebras [Sobczyk01].

Las bases recíprocas  $\{e_i\} \in N$  y  $\{\bar{e}_i\} \in \bar{N}$  son llamadas *duales*, esto es por que cumplen la ecuación  $e_i \cdot \bar{e}_i = 1$ . Estas bases generan las bases k-vectoriales de  $\mathcal{G}_{n,n}$ ,

$$\{1, \{e_i\}, \{\bar{e}_i\}, \{e_i e_j\}, \{\bar{e}_i \bar{e}_j\}, \{e_i \bar{e}_j\}, \dots, \{e_{j_1} \dots e_{j_i} \bar{e}_{j_k} \dots \bar{e}_{j_l}\}, I, \bar{I}\}, \quad (3.10)$$

consistente de escalares, vectores, bivectores, trivectores, ..., y los pseudoescalares  $I = e_1 \wedge e_2 \wedge e_3 \wedge \dots \wedge e_n$  y  $\bar{I} = \bar{e}_1 \wedge \bar{e}_2 \wedge \bar{e}_3 \wedge \dots \wedge \bar{e}_n$ , que satisfacen  $I\bar{I} = 1$ .

Usando estas bases ortogonales nulas, se puede construir unas *nuevas* bases ortonormales  $\{\sigma, \eta\}$  de  $\mathcal{G}_{n,n}$ ,

$$\sigma_i = \frac{1}{\sqrt{2}}(e_i + \bar{e}_i) \quad \eta_i = \frac{1}{\sqrt{2}}(e_i - \bar{e}_i) \quad (3.11)$$

para  $i = 1, 2, \dots, n$ . Estas bases vectoriales para  $i \neq j$  satisfacen las siguientes ecuaciones,

$$\begin{aligned} \sigma_i^2 &= 1, & \sigma_i \cdot \sigma_j &= \delta_{ij}, & \sigma_i \sigma_j &= -\sigma_j \sigma_i \\ \eta_i^2 &= -1, & \eta_i \cdot \eta_j &= -\delta_{ij}, & \eta_i \eta_j &= -\eta_j \eta_i \\ \sigma_i \eta_j &= -\sigma_j \eta_i, & \eta_i \cdot \sigma_j &= 0. \end{aligned} \quad (3.12)$$

Estas bases  $\{\sigma_i\}$  expanden un espacio vectorial euclideo real  $\mathbf{R}^n$  y genera el álgebra geométrica  $\mathcal{G}_{n,0}$ , mientras que las bases  $\{\eta_i\}$  expanden un espacio anti-euclideo  $\mathbf{R}^{0,n}$  y genera el álgebra  $\mathcal{G}_{0,n}$ . A partir de estas dos últimas álgebras se crea el álgebra geométrica  $\mathcal{G}_{n,n}$  como el producto directo de ellas:

$$[\mathcal{G}_{n,n} = \mathcal{G}_{n,0} \otimes \mathcal{G}_{0,n} = \text{gen}\{\sigma_1, \dots, \sigma_n, \eta_1, \dots, \eta_n\}. \quad (3.13)$$

El pseudoescalar y el pseudoescalar dual está dado por  $I = \sigma_1 \wedge \sigma_2 \wedge \dots \wedge \sigma_n$  y  $\bar{I} = \eta_1 \wedge \eta_2 \wedge \dots \wedge \eta_n$ , que satisface  $I\bar{I} = 1$ .

### 3.2.1 La horósfera y el plano afino n-dimensional.

El *Plano Afino n-Dimensional*  $\mathcal{A}_e(\mathbf{R}^{p,q})$  es una representación homogénea de los puntos  $\mathbf{x} \in \mathbf{R}^{p,q}$ . El plano afino extiende  $\mathbf{R}^{p,q}$  al espacio proyectivo  $\mathbf{R}^{p,q,1}$  usando un vector nulo  $e$  como sigue:

$$\mathbf{x}_h = \mathbf{x} + e \in \mathcal{A}_e(\mathbf{R}^{p,q}). \quad (3.14)$$

La  $(p,q)$ -horósfera normalmente es definida como:

$$\mathcal{H}_e^{p,q}(\mathbf{R}^{p+1,q+1}) = \left\{ \frac{1}{2} \mathbf{x}_h \bar{e} \mathbf{x}_h \mid \mathbf{x}_h \in \mathcal{A}_e(\mathbf{R}^{p,q}) \right\} \subset \mathbf{R}^{p+1,q+1}, \quad (3.15)$$

donde el espacio  $\mathbf{R}^{p,q}$  ha sido extendido a  $\mathbf{R}^{p+1,q+1}$  para poder tener disponibles dos vectores nullos,  $e = e_{n+1}$  y  $\bar{e} = \bar{e}_{n+1}$ .

Tanto el plano afino n-dimensional como la horósfera son útiles cuando se trabaja con transformaciones conformales [Pozo01]. Este tipo de transformaciones conserva los ángulos entre los vectores tangentes en cada punto. Por lo que

puede ser útil expresar un punto  $\mathbf{x} \in \mathbf{R}^{p,q}$  y un punto  $\mathbf{x}_h \in \mathcal{A}_e(\mathbf{R}^{p,q})$  con su representación conformal  $\mathbf{x}_c$  tal como sigue,

$$\begin{aligned} \mathbf{x}_c &= \frac{1}{2} \mathbf{x}_h \bar{e} \mathbf{x}_h = \frac{1}{2} [(\mathbf{x}_h \cdot \bar{e}) \mathbf{x}_h + (\mathbf{x}_h \wedge \bar{e}) \mathbf{x}_h] \\ &= \mathbf{x}_h - \frac{1}{2} \mathbf{x}_h^2 \bar{e} = \sigma \mathbf{x} - \frac{1}{2} \mathbf{x}^2 \bar{e} + e \\ &= \exp\left(\frac{1}{2} \mathbf{x} \bar{e}\right) e \exp\left(-\frac{1}{2} \mathbf{x} \bar{e}\right). \end{aligned} \quad (3.16)$$

Esta ecuación implica que todos los puntos en  $\mathcal{H}_e^{p,q}$  pueden ser obtenidos con una rotación de  $e$  con respecto al plano indicado por el bivector  $\mathbf{x} \bar{e}$ .

Los puntos de la horófera pueden ser proyectados al plano afino aplicando la siguiente fórmula,

$$\begin{aligned} \mathbf{x}_h &= P_A(\mathbf{x}) \\ &= (\mathbf{x}_c \wedge \bar{e}) \cdot e, \end{aligned} \quad (3.17)$$

y al espacio  $\mathbf{R}^{p,q}$  usando

$$\mathbf{x} = (\mathbf{x}_c \wedge \bar{e} \wedge e) \cdot (\bar{e} \wedge e). \quad (3.18)$$

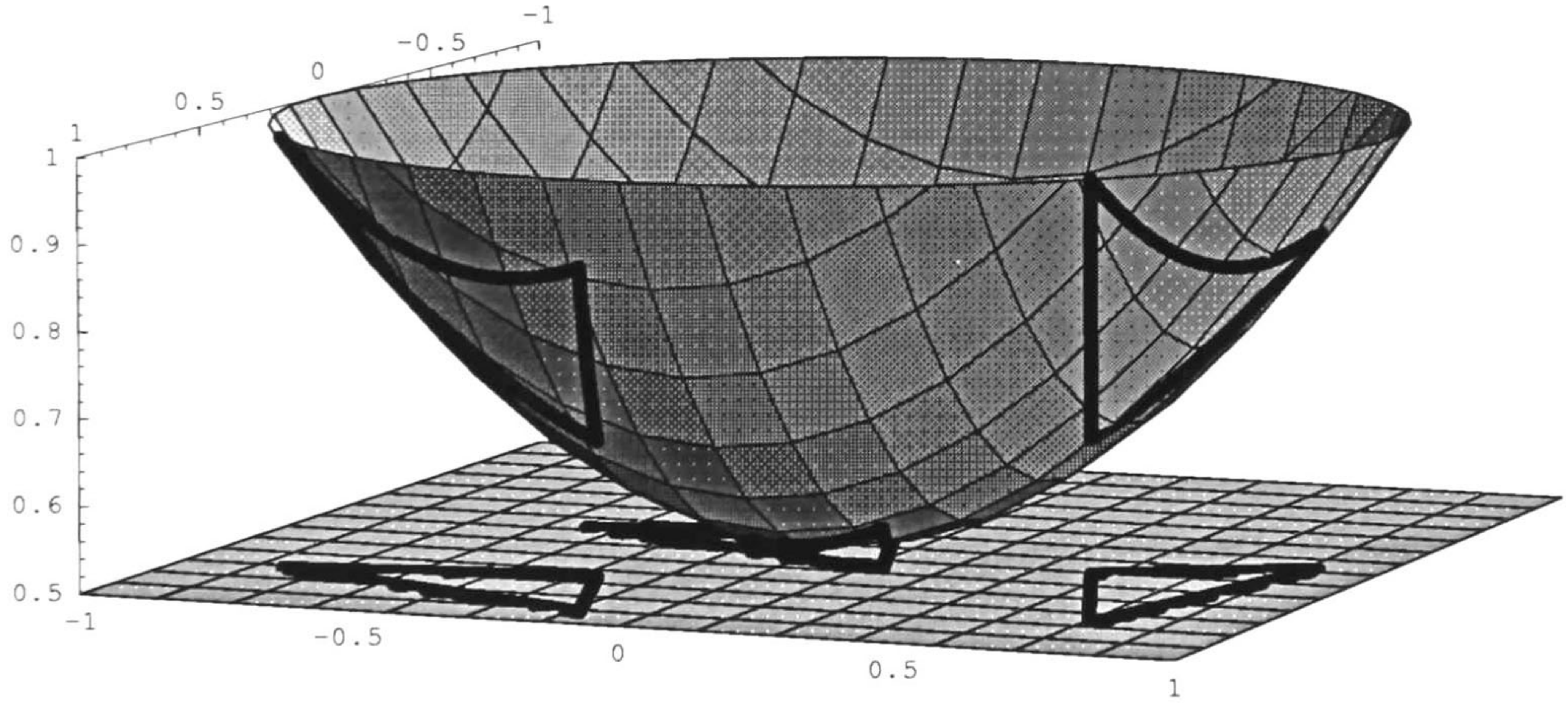


Figura 3.2: La horófera de  $\mathbf{R}^2$  con triángulos proyectados en el plano afino 2D.

En la figura 3.2 se muestra la horófera y triángulos proyectados en el plano afino. En dicha figura se presentan el plano afino 2D

$$\mathcal{A}_e^2 = \{\mathbf{x}_h | \mathbf{x}_h = \mathbf{x} + e, \mathbf{x} \in \mathbf{R}^2\}, \quad (3.19)$$

y la horófera

$$\mathcal{H}_e^2 = \{\mathbf{x}_c = \frac{1}{2} \mathbf{x}_h \bar{e} \mathbf{x}_h | \mathbf{x}_h \in \mathcal{A}_e^2\}. \quad (3.20)$$

Dado que  $\mathbf{x} \in \mathbf{R}^2 = \text{gen}\{\sigma_1, \sigma_2\}$ ,  $e = \frac{1}{\sqrt{\eta^2}}(\sigma + \eta)$ , y  $\bar{e} = \frac{1}{\sqrt{\eta^2}}(\sigma - \eta)$ , cualquier punto sobre la horósfera en estos términos está dado por

$$\begin{aligned} \mathbf{x}_c &= \mathbf{x} - \frac{1}{2}\mathbf{x}^2\bar{e} + e \\ &= \mathbf{x} - \frac{1}{2}(\mathbf{x}^2 - 1)\sigma + \frac{1}{2}(\mathbf{x}^2 + 1)\eta \\ &= \mathbf{x} + x_3\sigma + x_4\eta. \end{aligned} \tag{3.21}$$

Para poder dibujar la horósfera en 3D, en la figura 3.2 se ignoró la coordenada  $x_3$ , considerando en su lugar la condición de que  $\eta$  es ortogonal a  $\sigma_1, \sigma_2$ .

### 3.3 Distancia dirigida, álgebra de incidencia y movimiento rígido.

En esta sección se utilizará el álgebra geométrica  $G_{4,1,0}$  para describir el plano afino tridimensional,  $\mathcal{A}_h^3$ . En las siguientes subsecciones se describirán las reglas para establecer relaciones entre objetos geométricos del espacio 3D: puntos, líneas y planos. Para una introducción más detallada del álgebra de incidencia y el plano afino refiérase a [Bayro01].

#### 3.3.1 Distancia dirigida de objetos paralelos.

Para derivar las ecuaciones se usarán la línea y el plano mostrados en la figura 3.3. En general un  $k$ -plano  $A^h$  consiste de un momento de grado  $k$  y una dirección de grado  $k-1$ . Por lo cual, si se toma el producto interno entre la dirección unitaria y el momento se obtendrá la distancia dirigida como el vector  $p_h$ . Si se multiplica con el producto punto  $\mathcal{A}^h$  con  $\bar{e}$  y se divide este resultado por su norma se tendrá la dirección unitaria  $D_u$

$$\mathbf{D} = \bar{e} \cdot A^h \longrightarrow D_u = \frac{\mathbf{D}}{|\mathbf{D}|}, \tag{3.22}$$

la cual se usa para calcular la distancia al origen de  $A^h$  como sigue

$$p_h = D_u \cdot A^h. \tag{3.23}$$

En la práctica se suele utilizar  $\mathbf{D}$  en vez de  $D_u$ . El punto  $p_h$  está referido al origen y toca de manera ortogonal al  $k$ -plano  $A^h$ . Algo que se debe mencionar al respecto de  $p_h$  es que su norma es igual a la distancia de Hesse.

El punto  $p_h$  se usa para calcular la distancia dirigida del  $k$ -plano  $A^h$  hasta el objeto paralelo  $B^h$  de la siguiente manera

$$d[A^h, B^h] = d[\mathbf{D} \cdot A^h, B^h] = d[(\bar{e} \cdot A^h) \cdot A^h, B^h]. \tag{3.24}$$

La distancia de un punto  $b^h$  a la línea  $L^h = a_1^h \wedge a_2^h$  es la *magnitud* o la *norma* de su distancia dirigida

$$|d| = \left| \left[ \{\bar{e} \cdot (a_1^h \wedge a_2^h)\} \wedge \{\bar{e} \cdot (b^h)\} \right]^{-1} \left[ \bar{e} \cdot (a_1^h \wedge a_2^h \wedge b^h) \right] \right|. \tag{3.25}$$



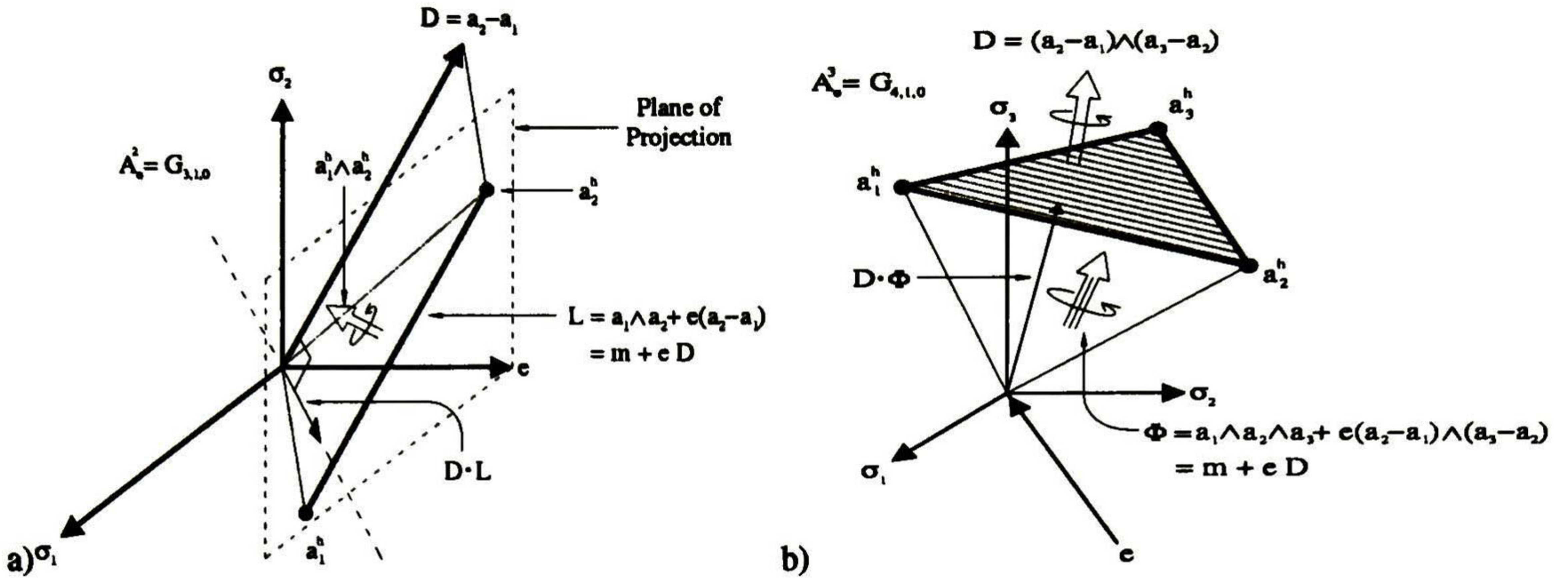


Figura 3.3: a) Línea en es espacio afino 2D. b) Plano en el espacio afino 2D.

La distancia de un punto  $b^h$  a un plano  $A^h = a_1^h \wedge a_2^h \wedge a_3^h$  es

$$|d| = \left| \left[ \{ \bar{e} \cdot (a_1^h \wedge a_2^h \wedge a_3^h) \} \wedge \{ (\bar{e} \cdot (b^h)) \} \right]^{-1} \left[ \bar{e} \cdot (a_1^h \wedge a_2^h \wedge a_3^h \wedge b^h) \right] \right|. \quad (3.26)$$

### 3.3.2 Álgebra de incidencia.

Primeramente se definirán las operaciones *Meet*,  $\cap$ , y *Join*,  $\cup$ , que se utilizarán a lo largo de ésta y la siguiente sección y el resto de la tesis. Suponga que  $r$  puntos  $a_1, a_2, \dots, a_r \in \Pi^n$ , donde  $\Pi^n$  es el hiperplano de grado  $n$ , y los puntos  $a_i$  están en posición general (son linealmente independientes). Entonces un  $(r-1)$ -plano en  $\Pi^n$  está especificad por

$$A_r = a_1 \wedge a_2 \wedge \dots \wedge a_r \neq 0. \quad (3.27)$$

Similarmente, un  $(s-1)$ -plano en  $\Pi^n$  está especificado por

$$B_s = b_1 \wedge b_2 \wedge \dots \wedge b_s \neq 0 \quad (3.28)$$

determinado por  $s$  puntos en posición general en  $\Pi^n$ . Considerando que los puntos  $a_i$  y  $b_j$  son las bases de los respectivos subespacio  $\mathcal{A}^r$  y  $\mathcal{B}^s$  estos pueden ser ordenados para definir su unión como el máximo espacio generado por todos los elementos linealmente independientes, es decir,

$$\mathcal{A}^r \cup \mathcal{B}^s = \text{gen}\{a_1, a_2, \dots, a_r, b_{\lambda_1}, \dots, b_{\lambda_k}\}. \quad (3.29)$$

Entonces si  $B_s = b_{\lambda_1} \wedge \dots \wedge b_{\lambda_k} \wedge b_{\alpha_1} \wedge \dots \wedge b_{\alpha_{s-k}}$ , el operador unión o más conocido como join es

$$A_r \cup B_s = A_r \wedge b_{\lambda_1} \wedge \dots \wedge b_{\lambda_k} \quad (3.30)$$

y la intersección de los subespacios será

$$\mathcal{A}^r \cap \mathcal{B}^s = \text{gen}\{b_{\alpha_1}, \dots, b_{\alpha_{s-k}}\}. \quad (3.31)$$

Definiendo el elemento pseudoescalar recíproco  $\bar{I}_{A \cup B}$  de la unión  $\mathcal{A}^r \cup \mathcal{B}^s$  como

$$\bar{I}_{A \cup B} = \bar{a}_1 \wedge \dots \wedge \bar{a}_r \wedge \bar{b}_{\lambda_1} \wedge \bar{b}_{\lambda_k}, \quad (3.32)$$

su operador intersección o más conocido como meet es

$$A_r \cap B_s = A_r \cdot (B_s \cdot \bar{I}_{A \cup B}). \quad (3.33)$$

### 3.3.3 Relaciones de incidencia en el plano afino 3D.

La relación de incidencia entre las líneas  $L_1^h = \mathbf{a}_1^h \wedge \mathbf{a}_2^h$  y  $L_2^h = \mathbf{b}_1^h \wedge \mathbf{b}_2^h$  está completamente definida por su unión  $I_{L_1^h \cup L_2^h} = L_1^h \cup L_2^h$ . Si  $I_{L_1^h \cup L_2^h}$  es un bivector, las líneas coinciden y  $L_1^h = tL_2^h$  para alguna  $t \in \mathbf{R}$ . Si  $I_{L_1^h \cup L_2^h}$  es un trivector, las líneas o son paralelas o se intersectan en un punto. En este último caso la operación *meet*

$$\mathbf{p} = L_1^h \cap L_2^h = L_1^h \cdot [(L_2^h \cdot \bar{I}_{L_1^h \cup L_2^h})]. \quad (3.34)$$

Si  $\bar{\mathbf{e}} \cdot \mathbf{p} = 0$  las líneas son paralelas, en caso contrario se intersectan en el punto  $\mathbf{p}_h = \frac{\mathbf{p}}{\bar{\mathbf{e}} \cdot \mathbf{p}}$  en el espacio afino 3D  $\mathcal{A}_e^3$ . Finalmente, si  $I_{L_1^h \cup L_2^h}$  es un 4-vector, las líneas no se intersectan y tienen direcciones diferentes.

La relación de incidencia entre una línea  $L^h = \mathbf{a}_1^h \wedge \mathbf{a}_2^h$  y un plano  $B^h = \mathbf{b}_1^h \wedge \mathbf{b}_2^h \wedge \mathbf{b}_3^h$  está también determinada por su unión  $L^h \cup B^h$ . Es claro que si la unión es un trivector, la línea  $L^h$  yace sobre el plano  $B^h$ . La otra posibilidad es que la unión sea el pseudoescalar  $I = \sigma_{123}e$ . En este caso, se calcula el meet

$$\mathbf{p} = L^h \cap B^h = L^h \cdot [(B^h \cdot \bar{I})]. \quad (3.35)$$

Si  $\bar{\mathbf{e}} \cdot \mathbf{p} = 0$ , la línea es paralela al plano con la distancia dirigida determinada por la ecuación 3.22. En caso contrario,  $\mathbf{p}_h = \frac{\mathbf{p}}{\bar{\mathbf{e}} \cdot \mathbf{p}}$  es el punto de intersección en el plano afino.

Dos planos  $A^h = \mathbf{a}_1^h \wedge \mathbf{a}_2^h \wedge \mathbf{a}_3^h$  y  $B^h = \mathbf{b}_1^h \wedge \mathbf{b}_2^h \wedge \mathbf{b}_3^h$  en el plano afino  $\mathcal{A}_e^3$  pueden ser paralelos o se intersectan en una línea, o coinciden. Si su unión es un trivector, entonces  $A^h = tB^h$  para algún  $t \in \mathbf{R}$ , o en otras palabras, los planos coinciden. Si los planos no coinciden, entonces su unión es el pseudoescalar  $I = \sigma_{123}e$ . En este caso, se calcula el meet según la ecuación 3.33 como

$$L = A^h \cap B^h = [A^h \cdot B^h] \cdot (\bar{I}_{A \cup B}). \quad (3.36)$$

Si  $\bar{\mathbf{e}} \cdot L = 0$  los planos son paralelos con distancia dirigida determinada por la ecuación 3.22. En otro casa,  $L$  representa la línea de intersección en el plano afino con dirección  $\bar{\mathbf{e}} \cdot L$ .

### 3.3.4 Indicadores geométricos.

A menudo se requiere verificar la configuración durante un movimiento rígido en el espacio Euclidiano. Las relaciones de incidencia geométricas pueden ser utilizadas para este proposito. Por ejemplo, un punto  $p$  yace sobre una línea  $L$ , si y sólo si,

$$p \wedge L = 0.$$

Similarmente, un punto  $p$  yace sobre un plano  $A$  si

$$p \wedge A = 0. \tag{3.37}$$

Una línea  $L$  estará sobre un plano  $A$  si

$$L \cap A = L. \tag{3.38}$$

Para verificar si una línea  $L$  interseca un plano  $A$  en un sólo punto  $p$ ,

$$L \cap A = p, \tag{3.39}$$

o si la línea  $L$  es paralela al plano  $A$ ,  $L \cap A$  dará un punto en el infinito.

### 3.3.5 Movimiento rígido en el plano afino.

Una rotación en el plano afino 3D  $\mathcal{A}_e^3 = \mathcal{A}_e(\mathbf{R}^3)$  es similar a la rotación en el espacio euclidiano  $\mathbf{R}^3$ , entonces el objeto encargado de la rotación, el rotor, está definido como

$$\mathbf{R} = \mathbf{m}\mathbf{n} = e^{\frac{\theta}{2}\mathbf{B}} = \cos\left(\frac{\theta}{2}\right) + \mathbf{B}\sin\left(\frac{\theta}{2}\right), \tag{3.40}$$

donde  $\mathbf{B}$  es el bivector unitario que define el plano de rotación.

Una traslación del vector  $\mathbf{x}_h \in \mathcal{A}_e^3$  a lo largo del vector  $\mathbf{t} \in \mathbf{R}^3$ , al vector  $\mathbf{x}'_h = \mathbf{x}_h + \mathbf{t} \in \mathcal{A}_e^3$ , es realizado por el objeto geométrico que llamaremos aquí trasladador,

$$\mathbf{T} = \exp\left(\frac{1}{2}\mathbf{t}\bar{e}\right) = 1 + \frac{1}{2}\mathbf{t}\bar{e}, \tag{3.41}$$

cuando esta transformación es seguida por la proyección  $P_A(\mathbf{x}') \equiv (\mathbf{x}' \wedge \bar{e}) \cdot e$ . Entonces para  $\mathbf{x}_h \in \mathcal{A}_e^3$ , se tendrá

$$\begin{aligned} \mathbf{x}' &= \mathbf{T}\mathbf{x}\mathbf{T}^{-1} \\ &= \exp\left(\frac{1}{2}\mathbf{t}\bar{e}\right)\mathbf{x}_h\exp\left(-\frac{1}{2}\mathbf{t}\bar{e}\right) \\ &= \mathbf{x}_h + \mathbf{t} - (\mathbf{t} \cdot \mathbf{x}_h + \frac{1}{2}\mathbf{t}^2)\bar{e}. \end{aligned} \tag{3.42}$$

Aplicando  $P_A$  a este resultado, se tendrá el vector trasladador esperado

$$\begin{aligned} \mathbf{x}'_h &= P_A(\mathbf{x}') \\ &= P_A[\mathbf{x}_h + \mathbf{t} - (\mathbf{t} \cdot \mathbf{x}_h + \frac{1}{2} \mathbf{t}^2) \bar{e}] \\ &= \mathbf{x}_h + \mathbf{t}. \end{aligned} \tag{3.43}$$

El material presentado en este capítulo se utilizará en secciones posteriores de este documento para validar restricciones geométricas en la configuración de objetos en el espacio 3D, así como para la estimación de transformaciones rígidas de los mismos y estimaciones de trayectorias. El lector interesado puede profundizar más en este tema consultando [Bayro01, Lounesto97, Sobczyk01].

## Capítulo 4

# Votación Tensorial .

En este capítulo se explicará la metodología de la votación tensorial (Tensor Voting) la cual es una herramienta útil en la extracción de características sobresalientes en las imágenes y en segmentación. Además se discutirá su posible utilidad en la validación de restricciones presente en la configuración de objetos geométricos.

### 4.1 Introducción.

En esta metodología se hace uso del cálculo de tensores para la representación de los datos, posiblemente ruidosos, y de una votación tensorial para la comunicación entre ellos. La votación de tensores se puede resumir como sigue: cada lugar (tensor) de entrada da un voto a sus vecinos y recolecta los votos que éstos le entregan para refinar la información que contenía, si es que esta fue proporcionada, dando resultados robustos ante el ruido y altamente prometedores para los investigadores e ingenieros que trabajan en los problemas del área de visión computacional. Para conocer más a fondo el uso de los tensores para la extracción de características sobresalientes en las imágenes refiérase a [Westin94]o a [Tang99].

### 4.2 Tensores.

Los tensores son conceptos matemáticos que tienen ciertas leyes específicas de transformación, cuando existe un cambio en el sistema de coordenadas. Existen vectores de diferentes órdenes, pero en el presente trabajo, solo se manejarán tensores de primer orden y de segundo orden. Para mayor simplicidad se puede considerar a los tensores de primer orden como vectores y a los de segundo como matrices. Para obtener mayor información sobre los tensores refiérase a [Hay5].

En el cálculo tensorial, los puntos se pueden representar por sus coordenadas.  $\mathbf{x} = (x_1, x_2, x_3)^T$ , las curvas se representan por las coordenadas de un punto, descripción local de primer orden, y su tangente asociada:  $\mathbf{x} = (x_1, x_2, x_3)^T$  y  $\mathbf{t} = (t_1, t_2, t_3)^T$ . Finalmente, las superficies se representan por las coordenadas de un punto, descripción local de primer orden, y su normal asociada:  $\mathbf{x} = (x_1, x_2, x_3)^T$  y  $\mathbf{n} = (n_1, n_2, n_3)^T$ .

### 4.2.1 Interpretación geométrica.

Un tensor simétrico de segundo orden (tensor en 2D) tiene la propiedad de poder representar los diferentes tipos de objetos geométricos mencionados anteriormente. Este tipo particular de tensor puede ser visualizado geoméricamente en 2D como una elipse donde los ejes principales corresponden a las direcciones del sistema de eigenvectores, ver la figura 4.1. Entonces un tensor simétrico en 2D se puede escribir

$$\mathbf{T} = \lambda_1 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \lambda_2 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T \quad (4.1)$$

, donde  $\lambda_n, \hat{\mathbf{e}}_n$  denotan los eigenvalores y eigenvectores de  $\mathbf{T}$  respectivamente, o en matrices

$$\mathbf{T} = (\hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2) \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{pmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \end{pmatrix}. \quad (4.2)$$

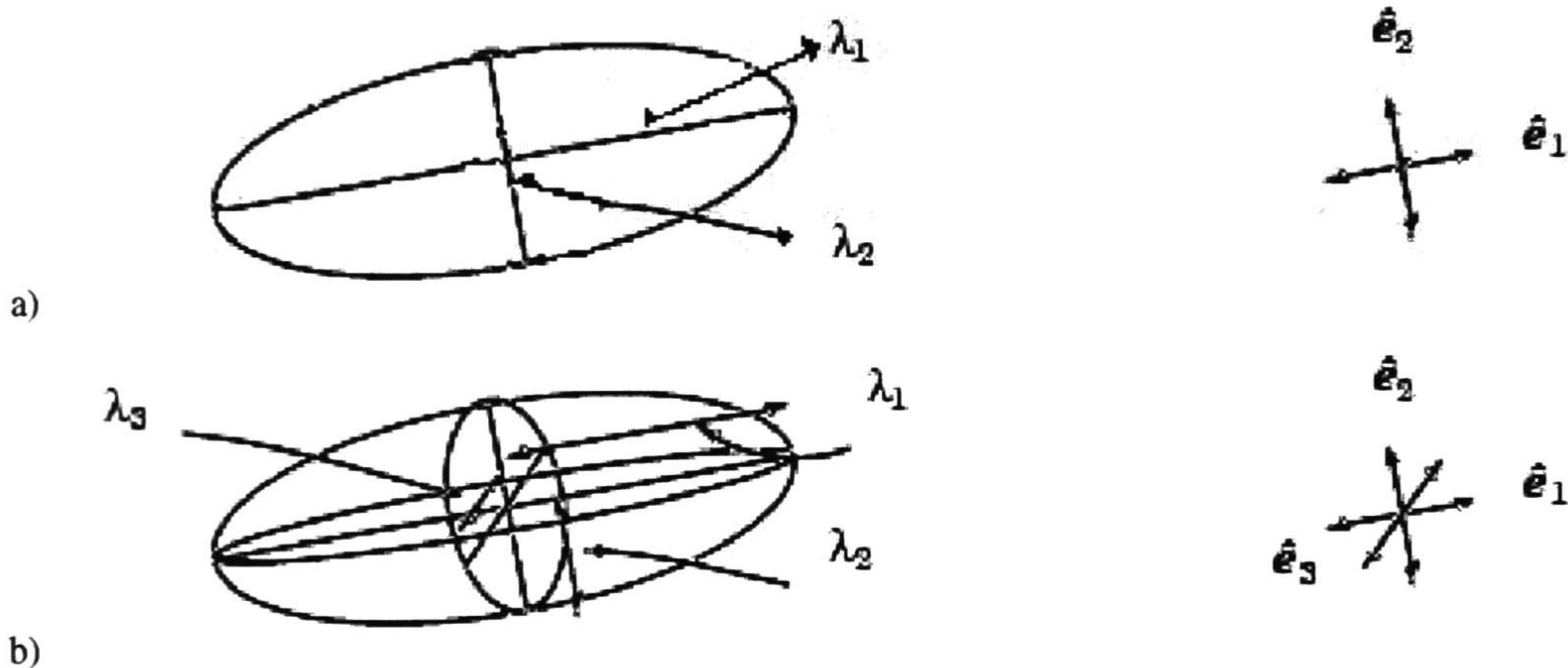


Figura 4.1: a) Representación Geométrica de un tensor en 2D como una elipse. b) Representación en 3D como un elipsoide.

En 3D, la visualización del tensor puede ser hecha similarmente pero con un elipsoide, ver la figura 4.1. El tensor se puede escribir como

$$\mathbf{T} = \lambda_1 \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \lambda_2 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T + \lambda_3 \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_3^T, \quad (4.3)$$

donde  $\lambda_n, \hat{\mathbf{e}}_n$  denotan los eigenvalores y eigenvectores de  $\mathbf{T}$  respectivamente. A partir de este punto cuando se hable de tensores, nos referiremos a tensores en 3D. El eigensistema se puede reescribir como

$$\mathbf{T} = (\lambda_1 - \lambda_2)S + (\lambda_2 - \lambda_3)P + \lambda_3 B, \quad (4.4)$$

donde  $S = \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T$ ,  $P = \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T$  y  $B = \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T + \hat{\mathbf{e}}_3 \hat{\mathbf{e}}_3^T$ , de tal forma que el tensor  $\mathbf{T}$  puede ser representado como una combinación lineal de  $S$ ,  $P$  y  $B$ , ver la figura 4.2. Si se visualiza geoméricamente el tensor  $S$  se podrá observar que describe una línea ya que los dos eigenvalores más pequeños son cero, de esta manera se puede decir que el tensor tiene una alta certeza sobre su dirección. De manera similar el tensor  $P$  describe un plano y el tensor  $B$  una esfera, de

ambos se puede decir que tienen incertidumbre acerca de su dirección en un plano y en el espacio respectivamente. Knutsson estableció que cualquier tensor en 3D puede ser representado como una combinación lineal de estos tensores básicos [Knutsson95].

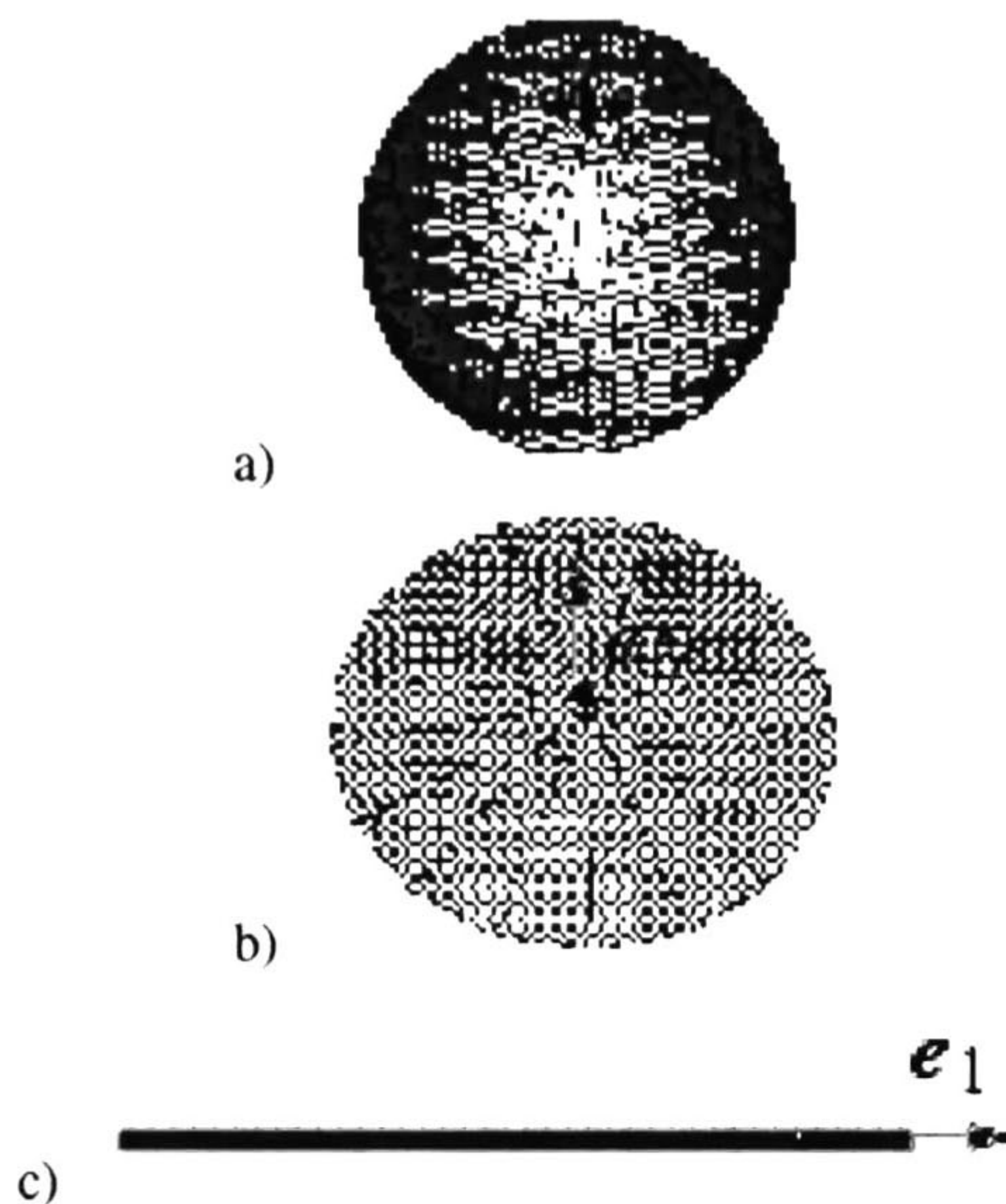


Figura 4.2: a) Tensor  $B$ . b) Tensor  $P$ . c) Tensor  $S$ .

### 4.3 Votación tensorial

Antes de poder hacer uso de la votación tensorial como medio de comunicación entre los tensores, debemos codificar cada punto de entrada como un tensor y éste a la vez como una combinación lineal de los tres tensores básicos tratados anteriormente. Un punto es codificado como un tensor  $B$  ( $\lambda_1 = \lambda_2 = \lambda_3 = 1$ ), dado que no se conoce ninguna dirección en particular para él. Una curva se codifica como un tensor  $P$  ( $\lambda_1 = \lambda_2 = 1, \lambda_3 = 0$ ) debido a su incertidumbre sobre la dirección de la tangente a la curva. Finalmente una superficie se codifica como un tensor  $S$  ( $\lambda_1 = 1, \lambda_2 = \lambda_3 = 0$ ) por la certeza que existe acerca de la dirección de su normal.

El objetivo de la comunicación de los tensores se puede resumir como sigue: derivar la orientación preferida, o refinar la orientación inicial si se dió alguna, para cada uno de los elementos; extrapolar la información antes inferida a todas las localidades del dominio, para los objetivos que se desee, por ejemplo extracción de características sobresalientes o la comprobación de restricciones en la configuración geométrica de los objetos geométricos obtenidos a partir de los datos de entrada.

#### 4.3.1 El proceso de votación.

La comunicación puede ser implementada como un proceso de votación, lo que implica una alineación con núcleos (campos) de votación predefinidos. Esta alineación no es más que una traslación seguida de una rotación. Los núcleos

de votación codifican los tensores básicos. El proceso de votación es similar a una convolución excepto que tanto las entradas como la salida del proceso son tensores (no escalares), debido a esto es llamado *Convolución Tensorial*, ver la figura 4.3. Esta convolución es posible ya que las operaciones necesarias para una convolución escalar también se encuentran definidas para los tensores.

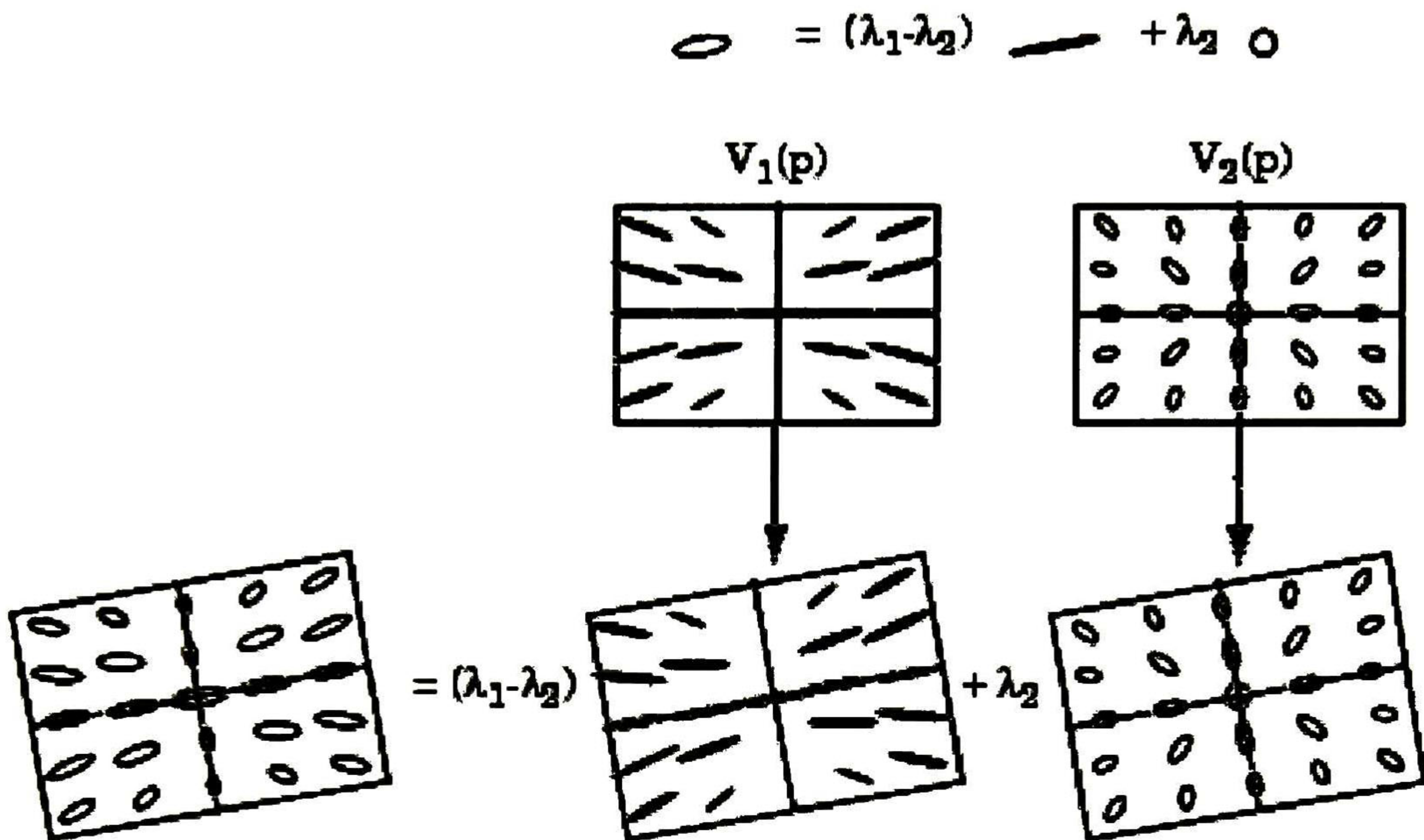


Figura 4.3: Ilustración del proceso de votación (en 2D). Primeramente se alinean los campos de votación y después se suman las contribuciones.

A continuación se describe el diseño del *Campo Fundamental 2D* y la derivación de los campos de votación 3D haciendo rotar e integrando el campo fundamental.

### 4.3.2 El campo fundamental 2D.

Los campos de votación de cualquier dimensión pueden ser derivados a partir del campo lineal 2D, motivo por el cual es llamado campo fundamental 2D, ver figura 4.4. Nótese que una dirección se puede expresar como un vector tangente o normal, que son ortogonales entre sí. Entonces es posible definir dos versiones del campo fundamental 2D, dependiendo entre si se asigna un vector tangente o normal a cada objeto de entrada. Aquí se presenta la versión normal del campo fundamental 2D.

En coordenadas esféricas, el campo fundamental 2D se puede expresar como

$$\overline{DF}(r, \varphi, \sigma) = e^{-\left(\frac{r^2 + \varphi^2}{\sigma^2}\right)} \quad (4.5)$$



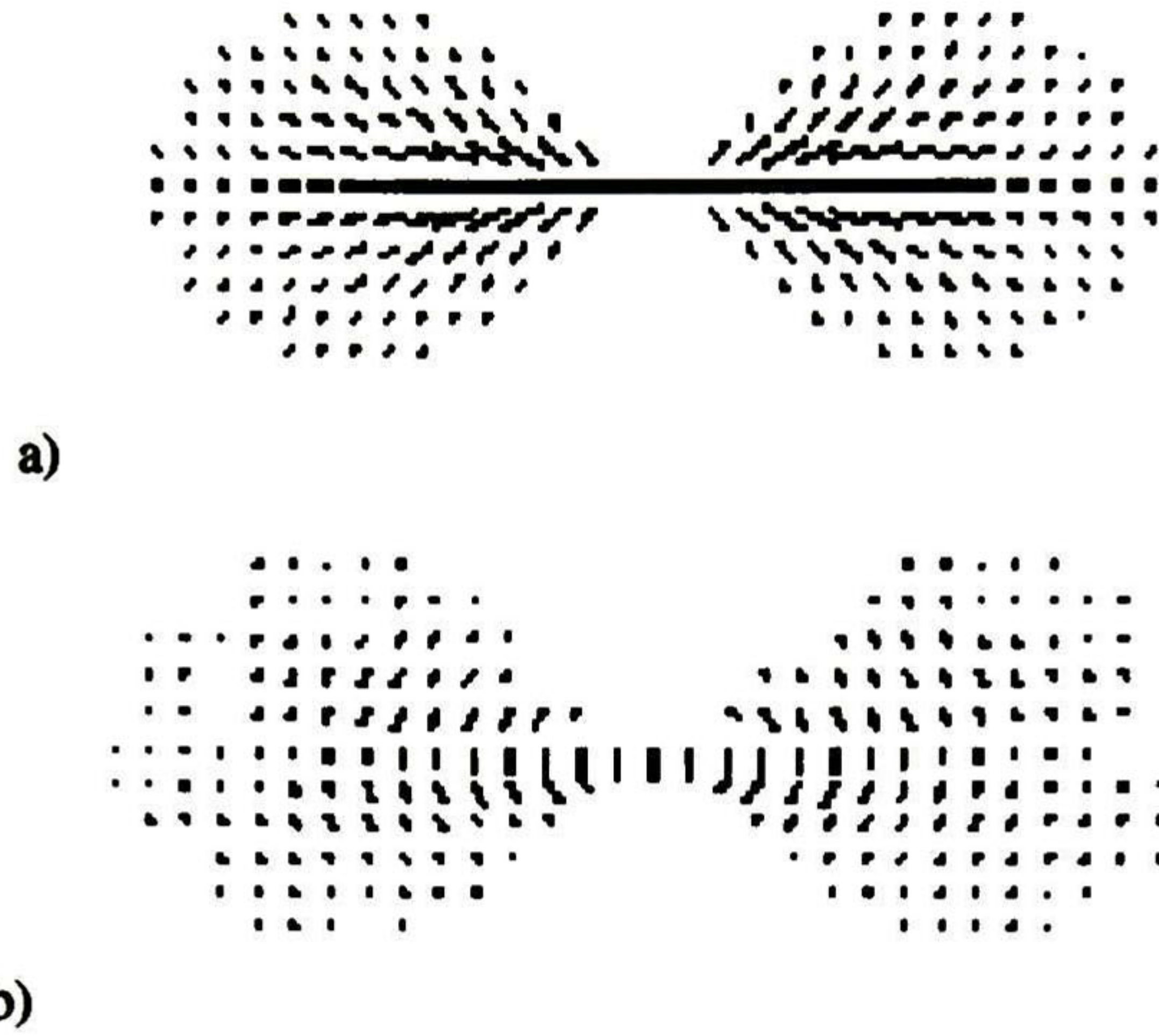


Figura 4.4: El campo fundamental 2D. a) Versión tangente. b) Versión normal.

donde  $r$  es la longitud del arco,  $\varphi$  es la curvatura, y  $\sigma$  es la escala del análisis, el único parámetro libre del formalismo. Para mayor información sobre esta función refiérase a [Guy96].

### 4.3.3 Campos de votación 3D.

Denotemos  $V_F$  al campo fundamental 2D en su forma normal. En 3D se necesitan tres campos de votación,  $V_S$ ,  $V_P$ ,  $V_B$ , que corresponden a los tensores básicos  $S$ ,  $P$  y  $B$  respectivamente. Rotando  $V_F$   $90^\circ$  en torno al eje  $z$  obtendremos un núcleo, al que llamaremos  $V'_F$ . El campo de votación  $V_S$  se obtiene integrando las contribuciones obtenidas rotando  $V'_F$  en torno al eje  $x$ , el cual describe la orientación  $(1,0,0)^T$ . Entonces el campo  $V_S$  es definido matemáticamente como

$$V_S = \int_0^\pi V'_F d\alpha |_{\beta=0, \gamma=0} \quad (4.6)$$

donde  $\alpha$ ,  $\beta$ ,  $\gamma$  son los ángulos de rotación en torno al eje  $x$ ,  $y$ ,  $z$  respectivamente.

Para obtener el campo de votación  $V_P$ , se debe rotar  $V_S$  en torno al eje  $z$  e integrar las contribuciones obtenidas, este campo describe el plano con la orientación normal  $(0,0,1)^T$ . Entonces el campo  $V_P$  es definido matemáticamente como

$$V_P = \int_0^\pi V_S d\gamma |_{\alpha=0, \beta=0} \quad (4.7)$$

Finalmente, para obtener el campo de votación  $V_B$ , se debe rotar  $V_S$  en torno a los ejes  $z$  e  $y$  e integrar las contribuciones obtenidas. Entonces el campo  $V_B$  es definido matemáticamente como

$$V_B = \int_0^\pi V_S d\beta d\gamma |_{\alpha=0}. \quad (4.8)$$

Las representaciones geométricas de estos campos se presentan en la figura 4.5.

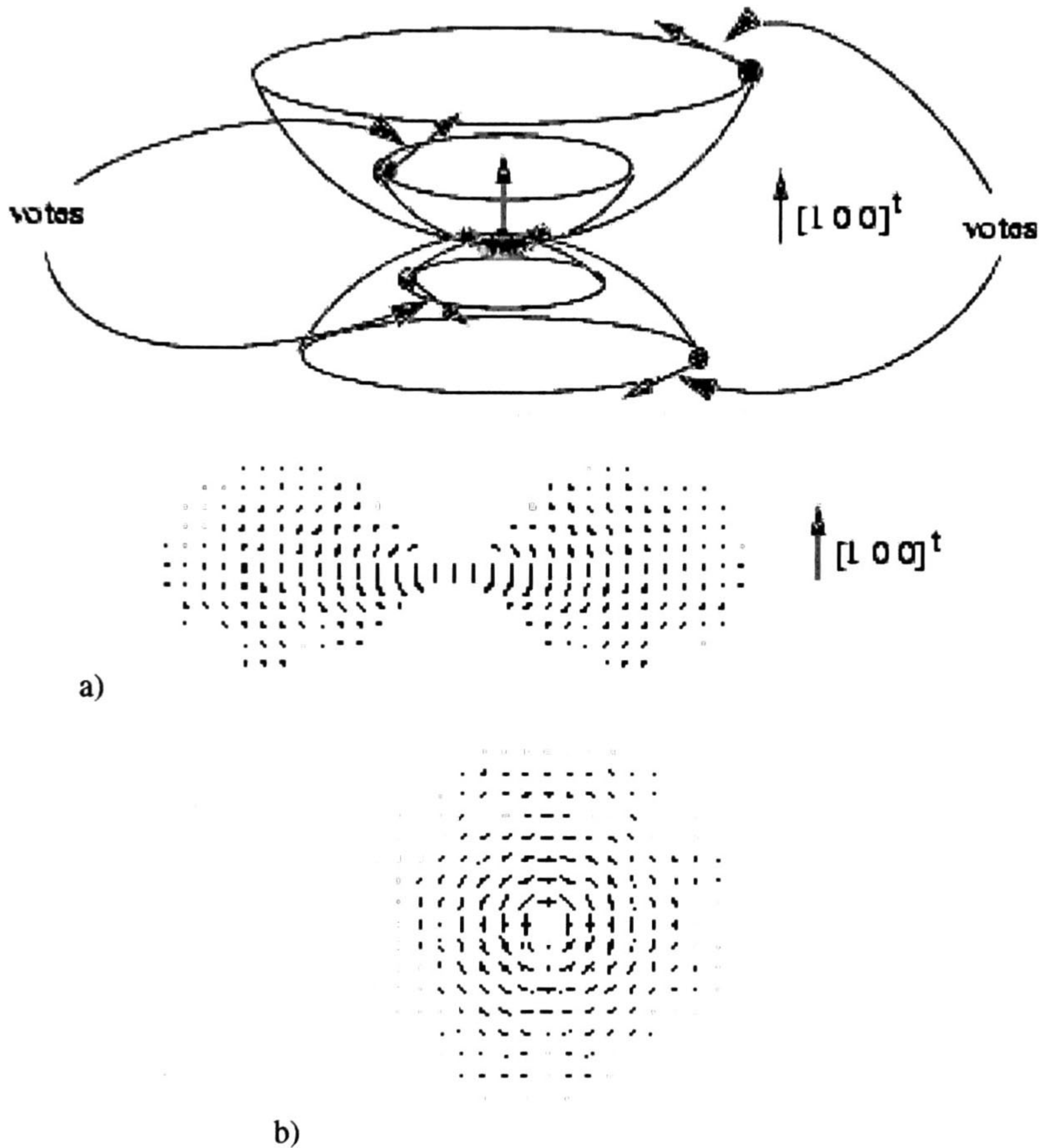


Figura 4.5: Campos de Votación 3D. a) Tensor  $V_S$ , muestra del proceso de obtención y corte mostrando su dirección. b) Tensor  $V_P$ , corte del tensor con la dirección  $\hat{e}_3$  normal a la página.

#### 4.3.4 Ilustración

Para simplificar la comprensión de la votación tensorial se presenta una ilustración del proceso con tensores en 2-D. Sean dos tensores  $T_{0,1}$  y  $T_{0,2}$ , primero se codifican como sigue: si son puntos como un tensor  $B$  y si es una tangente como un tensor  $S$ .

- $S. \lambda_1 = 1, \lambda_2 = 0$  y  $\hat{e}_1 = (t_x, t_y)^T, \hat{e}_2 = (t_y, t_x)^T$

- $B. \lambda_1 = \lambda_2 = 1$  y  $\hat{\mathbf{e}}_1 = (1,0)^T$ ,  $\hat{\mathbf{e}}_2 = (0,1)^T$

Descomponemos cada tensor en sus componentes  $S$  y  $B$ .

$$\begin{aligned}
\mathbf{T}_{0,j} &= (\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2) \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{pmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \end{pmatrix} \\
&= (\lambda_1 - \lambda_2) \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \lambda_2 (\hat{\mathbf{e}}_1 \hat{\mathbf{e}}_1^T + \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_2^T) \\
&= \mathbf{T}_{0,j}^S + \mathbf{T}_{0,j}^B
\end{aligned} \tag{4.9}$$

con  $j=1,2$ .

Sea  $v$  el voto  $V_S$  recolectado en la posición  $j$  entonces el nuevo valor del tensor será:

$$\mathbf{T}_{1,j}^S = \mathbf{T}_{0,j}^S + (\lambda_1 - \lambda_2) \begin{bmatrix} v_x^2 & v_x v_y \\ v_y v_x & v_y^2 \end{bmatrix} \tag{4.10}$$

con  $j=1,2$ .

Sea  $\mathbf{T}_B$  el voto  $V_B$  en la posición  $j$  entonces el nuevo tensor será:

$$\mathbf{T}_{1,j}^B = \mathbf{T}_{0,j}^B + \lambda_2 \mathbf{T}_B \tag{4.11}$$

con  $j=1,2$ .

## 4.4 Ejemplos de aplicaciones.

En esta sección se presentará una aplicación existente de la votación tensorial, si se desea obtener más información sobre las aplicaciones de la votación tensorial refiérase a [Medioni99]. Además se sugieren algunas otras aplicaciones en las cuales la aplicación de la votación tensorial parece prometedora. Esta es una de las aplicaciones de la votación tensorial.

### 4.4.1 Extracción de características sobresalientes.

Una vez finalizado el proceso de votación, se produce un mapa tensorial denso, el cual es descompuesto en tres mapas vectoriales para el caso 3-D. Estos mapas están compuestos de cubos en el espacio (voxels) que tienen una tupla  $(s,d)$ , donde  $s$  es un escalar que representa la magnitud, y  $d$  la dirección.

- Mapa de superficies:  $s = \lambda_1 - \lambda_2$  y  $d = \hat{\mathbf{e}}_1$  la dirección normal.
- Mapa de curvas:  $s = \lambda_2 - \lambda_3$  y  $d = \hat{\mathbf{e}}_3$  la dirección tangencial.

- Mapa de uniones:  $s = \lambda_3$  y  $d$  arbitrario.

Cuando se buscan los extremos de la función de superficie (máximos o mínimos) se hace de igual manera que con una función real, pero aplicando la derivada correspondiente, es decir, buscamos:  $\frac{ds}{d\hat{n}} = 0$ , donde  $s$  es la magnitud de la superficie y  $\hat{n}$  es la dirección normal. En la práctica estos extremos se obtienen buscando los cruces por cero usando el vector gradiente  $\bar{g}$  definido como sigue:

$$\bar{g} = \nabla S = \left[ \frac{\partial s}{\partial x} \quad \frac{\partial s}{\partial y} \quad \frac{\partial s}{\partial z} \right]^T$$

y se proyecta  $\bar{g}$  sobre  $\hat{n}$ , es decir,  $q = \hat{n} \cdot \bar{g}$ , entonces los extremos de la superficie serán los puntos donde  $q = 0$ .

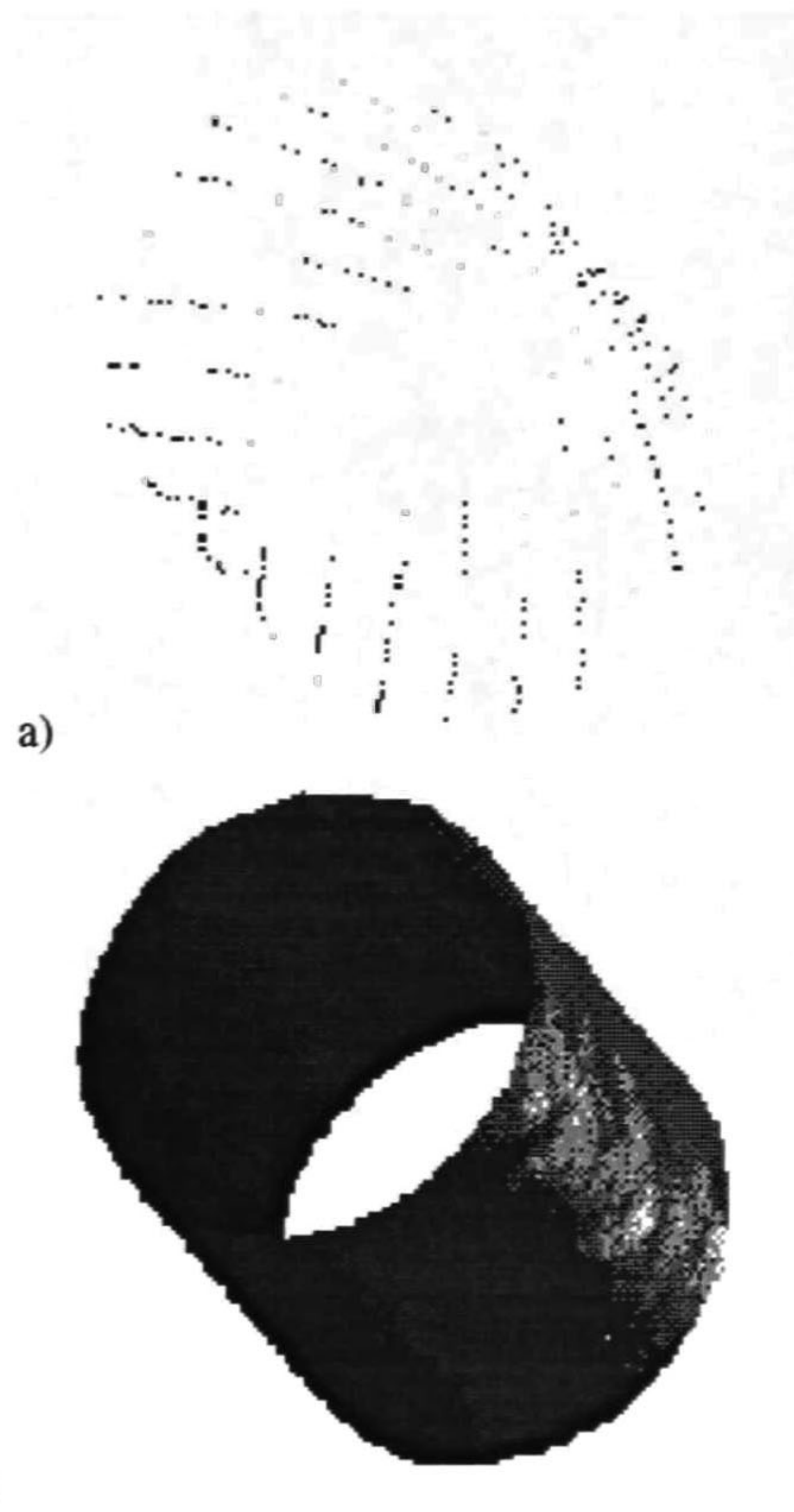
Para los extremos de las curvas definimos el plano formado por dos vectores  $\hat{u}$  y  $\hat{v}$  ortogonales entre sí, que es normal a la dirección tangencial  $\hat{t}$  de la curva, y tendremos un extremo cuando:  $\frac{ds}{d\hat{u}} = \frac{ds}{d\hat{v}} = 0$ . Definiéndolo en términos de un cruce por cero  $q = \mathbf{R}(\hat{t} \times \bar{g})$ , donde  $\mathbf{R}$  define una rotación que alinea el marco con el plano  $\hat{u}$ - $\hat{v}$ . Los extremos estarán en los puntos tales que  $q = 0$ .

Para poder obtener dichos extremos ya sean de curvas o superficies se utiliza una versión modificada del algoritmo Marching Cubes desarrollado por Lorensen y Cline en 1987[Lorensen87]. Se presentan algunos de los resultados obtenidos en esta aplicación en la Figura 4.6.

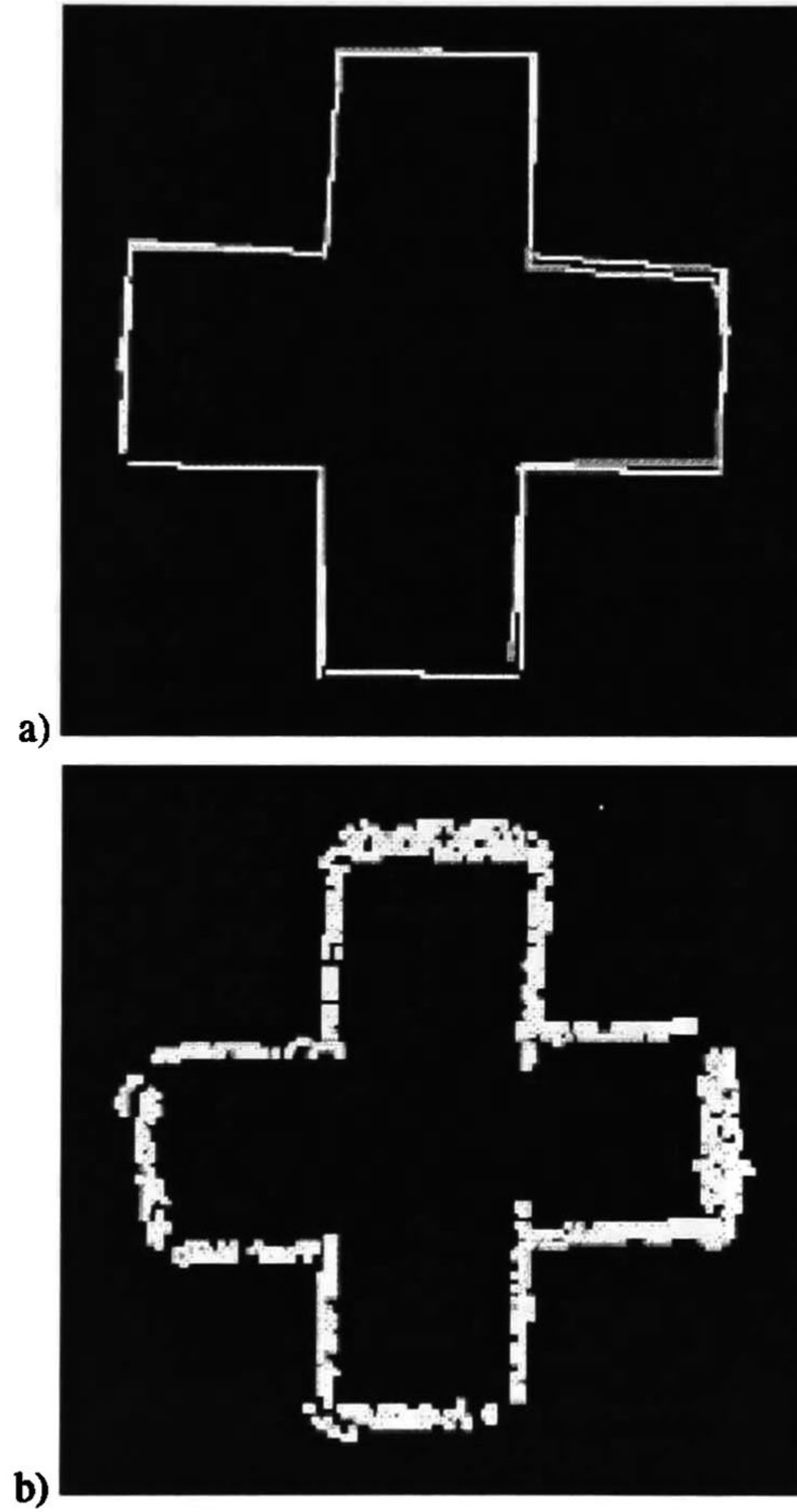
#### 4.4.2 Posibles aplicaciones.

La expresión de restricciones en la configuración de objetos en el mundo, y la validación de las mismas mediante la votación tensorial. Esto se pretende mediante la comprobación de que el producto tensorial exterior (en forma matricial  $\hat{e}_1 \hat{e}_1^T$ ) es equivalente al producto Clifford por el pseudoescalar ( $\sigma_1 I$ ), donde  $I$  es el pseudoescalar en el álgebra en cuestión,  $(\mathcal{G}_{3,0,0})$ .

También se pretende utilizar esta metodología no sólo para la extracción de características sobresalientes, sino también para la fusión de datos en 3-D provenientes de varias fuentes (laser, cámaras, sonar, odometría), ver figura 4.7.



**Figura 4.6: Resultados de la extracción de características sobresalientes. a) Puntos de entrada. b) Superficie generada.**



**Figura 4.7:** Datos a fusionar con la votación tensorial. a) Datos obtenidos mediante una unidad Laser. b) Datos obtenidos con sonares.

## Capítulo 5

# Reconstrucción y Planeación usando $n$ Vistas.

En este capítulo se presenta la metodología utilizada para la reconstrucción de objetos utilizando  $n$  pares de vistas. En el trabajo presentado en este capítulo se utilizan pares de vistas obtenidos con la cámara binocular mostrada en la Figura 2.3.



Figura 5.1: Cámara binocular montada sobre una unidad pan-tilt.

## 5.1 Reconstrucción de vistas.

### 5.1.1 Obtención de las vistas.

Antes de poder trabajar con las imágenes obtenidas con el cabezal binocular, debemos calibrar las cámaras individualmente, la calibración de las cámaras se hace mediante el algoritmo presentado por Tsai [Tsai87]. Una vez calibradas las cámaras se obtienen las matrices que describen cada una de las cámaras, PLeft y PRight. Las matrices PLeft y PRight se pueden expresar como  $P = [M|p_4]$ , claro cada una por separado y con matrices M y p diferentes para cada una.

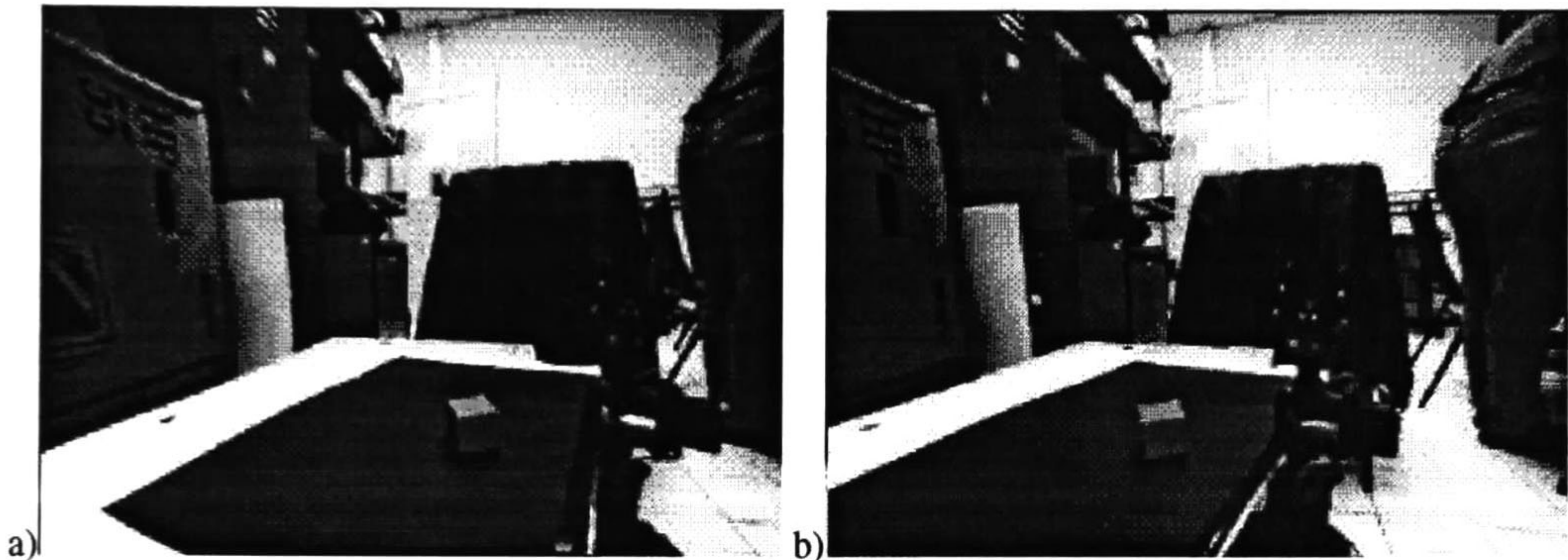


Figura 5.2: Imágenes obtenidas con las cámaras. a) Imagen izquierda. b) Imagen derecha.

En la Figura 5.2 se presentan un par de vistas obtenidas con el cabezal binocular.

### 5.1.2 Retroproyección.

Una vez obtenidas las imágenes, el siguiente paso es la detección de esquinas en cada una de las imágenes mediante el algoritmo de Harris [Harris88], ver figura 5.3. Esto se hace con motivo de obtener puntos relevantes en cada imagen para tener correspondencias entre puntos de las imágenes para la retroproyección y subsecuente triangulación y obtención del punto 3D.

Después de detectadas las esquinas, se hace la selección de puntos en correspondencia, los cuales son retroproyectados siguiendo los siguientes pasos:

- Se obtiene el centro de cada una de las cámaras mediante la fórmula

$$C = \begin{pmatrix} -M^{-1}p_4 \\ 1 \end{pmatrix} \quad (5.1)$$

- Se obtiene otro punto en el espacio que se proyecte al mismo punto en la imagen, esto se hace con la fórmula

$$X_1 = \begin{pmatrix} M^{-1}(\lambda x - p_4) \\ 1 \end{pmatrix}, \quad (5.2)$$



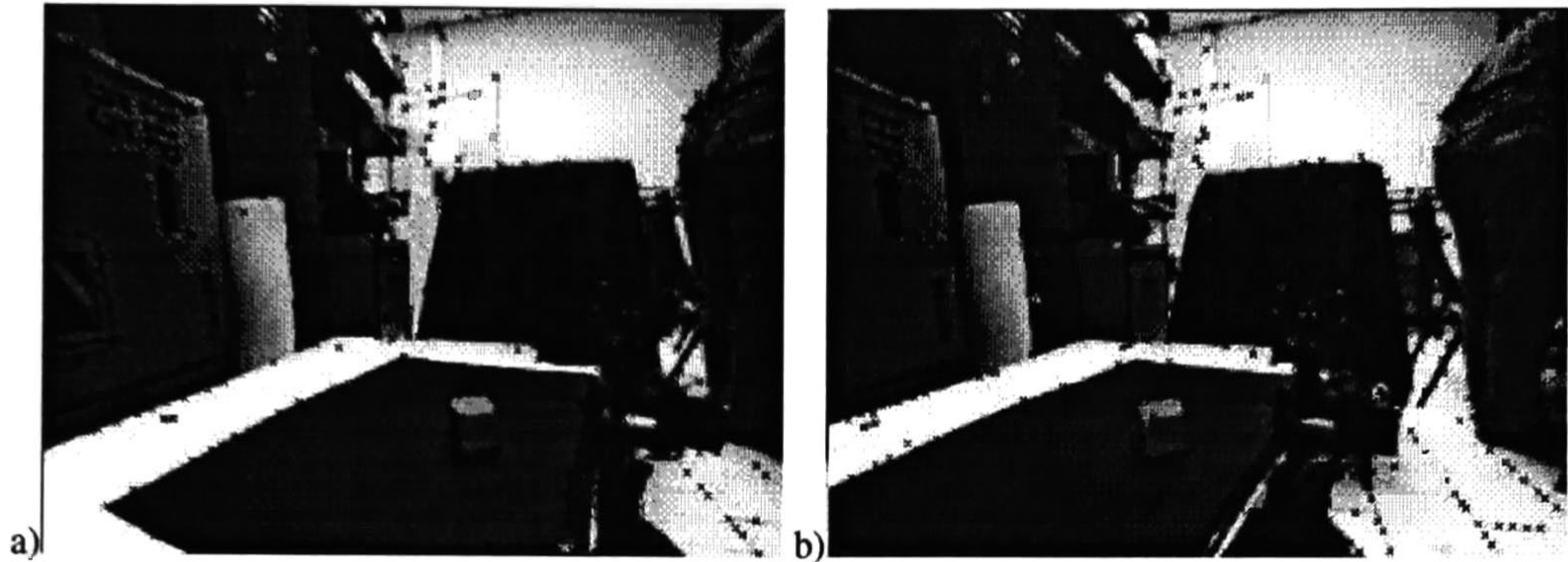


Figura 5.3: Imágenes con las esquinas detectadas. a) Imagen izquierda. b) Imagen derecha.

esto se hace para cualquier valor de  $\lambda$ , en nuestro caso usamos  $\lambda = 1000$ .

- Se hace la codificación de  $C$  y  $X_1$  en términos del álgebra geométrica  $\mathcal{G}_{4,1,0}$  mediante la fórmula

$$\begin{aligned} C_g &= C + e \\ X_g &= X_1 + e. \end{aligned} \tag{5.3}$$

- Se obtiene la línea que representa el rayo retroproyectado mediante el producto exterior (wedge) como

$$L = C_g \wedge X_g. \tag{5.4}$$

Ya que se han retroproyectado los puntos en las imágenes, lo que sigue es la triangulación con los rayos, lo cual se realiza calculando la distancia dirigida entre los rayos, como se explicó en el capítulo 3, y tomando el punto medio del vector que representa la distancia más corta entre ellos.

En la figura 5.4 se muestra la reconstrucción de los puntos que generan el cubo en la Figura 5.3. Note que las líneas que unen los puntos son únicamente con el objetivo de hacer más entendible la representación.

## 5.2 Fusión de vistas.

Después de que se han reconstruido los puntos de cada par de imágenes, se deben fusionar en el espacio 3D. Esta fusión se realiza calculando la transformación rígida (rotación y traslación) entre los puntos en 3D, dicha transformación se encuentra mediante el algoritmo presentado por Lasenby [Lasenby95]. Aquí se presenta una descripción del mismo mismo.

### 5.2.1 Estimación de movimiento rígido.

Dadas las coordenadas en 3D  $\{X_1\}$  y  $\{X'_1\}$  de dos reconstrucciones expresadas en el álgebra geométrica del espacio euclidiano tridimensional,  $\mathcal{G}_{3,0,0}$ . Entonces para  $n$  correspondencias en los dos marcos se quiere encontrar la rotación,

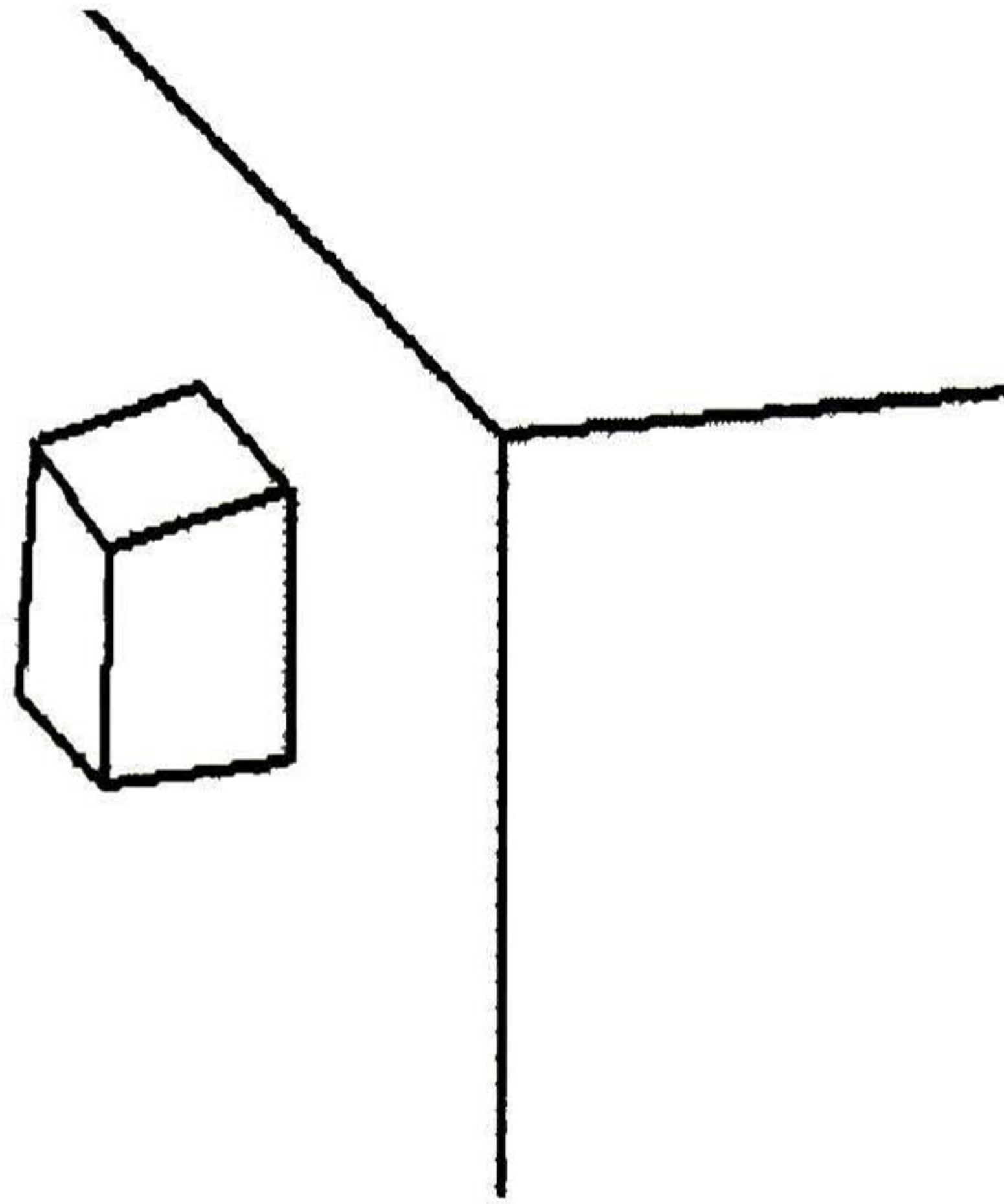


Figura 5.4: Reconstrucción de las esquinas del cubo de la Figura 5.3.

expresada como un rotor  $R$ , y la traslación,  $t$ , que minimize

$$S = \sum_{i=1}^n [\mathbf{X}'_i - R(\mathbf{X}_i - t)\tilde{R}]^2 \quad (5.5)$$

Luego dicha minimización implica la diferenciación de  $S$  con respecto a  $R$  y  $t$ . La diferenciación con respecto a  $t$  es sencilla y está dada por

$$\partial_t S = \sum_{i=1}^n [\mathbf{X}'_i - R(\mathbf{X}_i - t)\tilde{R}] \partial_t (Rt\tilde{R}) = 0. \quad (5.6)$$

La cual se resuelve para  $t$ , dando como resultado

$$t = \frac{1}{n} \sum_{i=1}^n [\mathbf{X}_i - \tilde{R}\mathbf{X}'_i R], \quad (5.7)$$

esto puede ser reescrito como

$$t = \bar{\mathbf{X}} - \tilde{R}\bar{\mathbf{X}}'R, \quad (5.8)$$

donde  $\bar{\mathbf{X}}$  y  $\bar{\mathbf{X}}'$  son los centroides de los puntos en las dos reconstrucciones.

La diferenciación con respecto al rotor  $R$  es mucho más complicada y excede los objetivos de este documento. Por este motivo aquí se presenta sólo la ecuación final, para mayores detalles sobre el proceso de diferenciación y solución de la expresión vea [Lasenby95]. El resultado de la diferenciación es

$$\sum_{i=1}^n \mathbf{X}'_i \wedge R(\mathbf{X}_i - t)\tilde{R} = 0. \quad (5.9)$$

Si se sustituye el valor obtenido para  $t$  en la ecuación anterior 5.9 se tendrá

$$\sum_{i=1}^n \mathbf{X}'_i \wedge R(\mathbf{X}_i - \bar{\mathbf{X}} - \tilde{R}\bar{\mathbf{X}}'R)\tilde{R} = 0. \quad (5.10)$$

Nótese que el término  $\sum_{i=1}^n \mathbf{X}'_i \wedge \bar{\mathbf{X}}'$  que se obtendrá de la ecuación 5.10, será igual a cero. Entonces la ecuación se reduce a

$$\sum_{i=1}^n w_i \wedge R u_i \tilde{R} = 0, \quad (5.11)$$

con

$$\begin{aligned} u_i &= \mathbf{X}_i - \bar{\mathbf{X}} \\ w_i &= \mathbf{X}'_i. \end{aligned} \quad (5.12)$$

Con estos nuevos valores se define una matriz  $F$  como sigue

$$F_{\alpha\beta} = \sum_{i=1}^n (\sigma_\alpha \cdot u_i)(\sigma_\beta \cdot w_i). \quad (5.13)$$

Obteniendo el SVD de la matriz  $F$  da  $F = USV^T$  y la rotación  $R$  estará dada por

$$R = VU^T \quad (5.14)$$

Una vez que se ha obtenido el rotor  $R$ , el trasladador  $t$  se calcula con la ecuación 5.8.

### 5.2.2 Fusión de vistas.

Para hacer la fusión de las vistas lo primero que se debe hacer es obtener una correspondencia de puntos entre los pares de imágenes como las que se muestran en la figura 5.5 y la figura 5.6. Luego se reconstruyen los puntos de las imágenes para obtener puntos en 3D, ver figura 5.7. Con los puntos en 3D reconstruidos, sabiendo las correspondencias, se procede a calcular la rotación y traslación que relaciona dichos puntos. Ya conocidas las transformaciones se deben aplicar para obtener una reconstrucción del escenario pero con la información recopilada por más de una vista, ver figura 5.8.

## 5.3 Planeación.

Este capítulo pretende presentar al lector las reglas necesarias para hacer un razonamiento, tomar una decisión y ejecutar alguna acción en caso de ser necesario. En la primera sección se plantea un método para obtener *primitivas geométricas*, objetos tales como puntos, líneas, planos y volúmenes. En la segunda sección se presentan algunas de las reglas de decisión. Finalmente en la tercera sección se presenta la ejecución de una tarea elegida con las reglas de

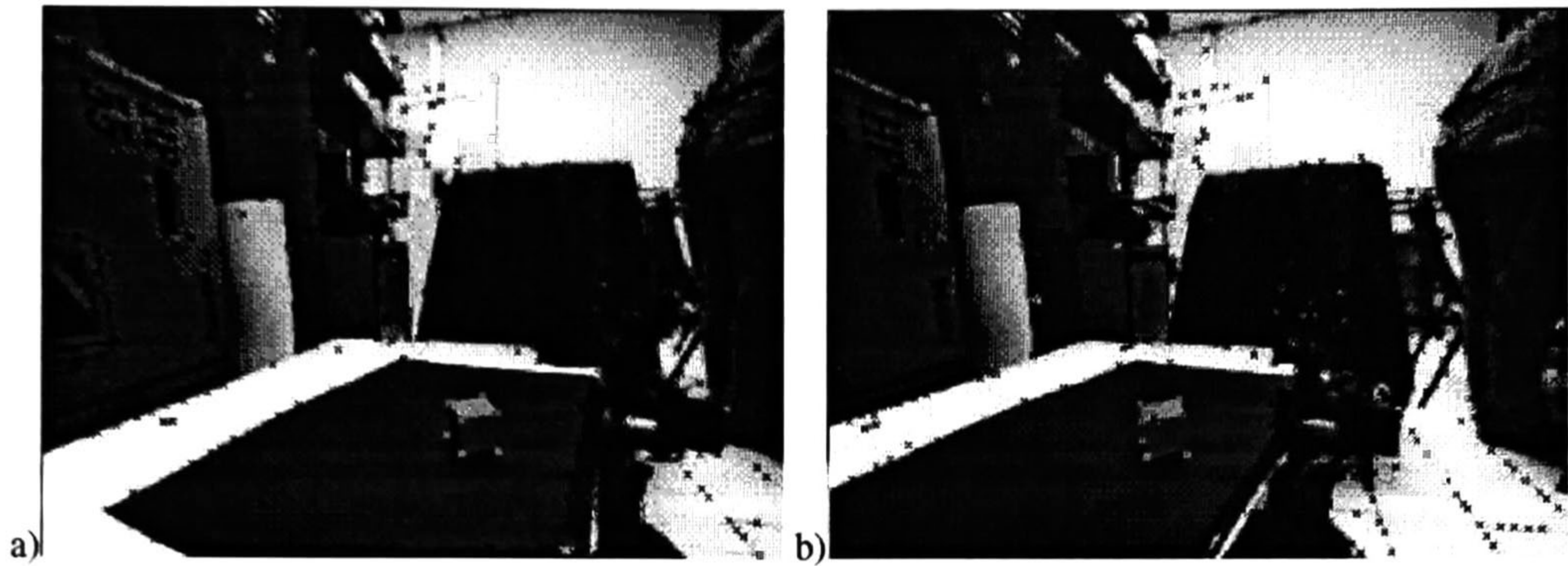


Figura 5.5: Correspondencias de puntos entre vistas. a) Imagen izquierda. b) Imagen derecha.

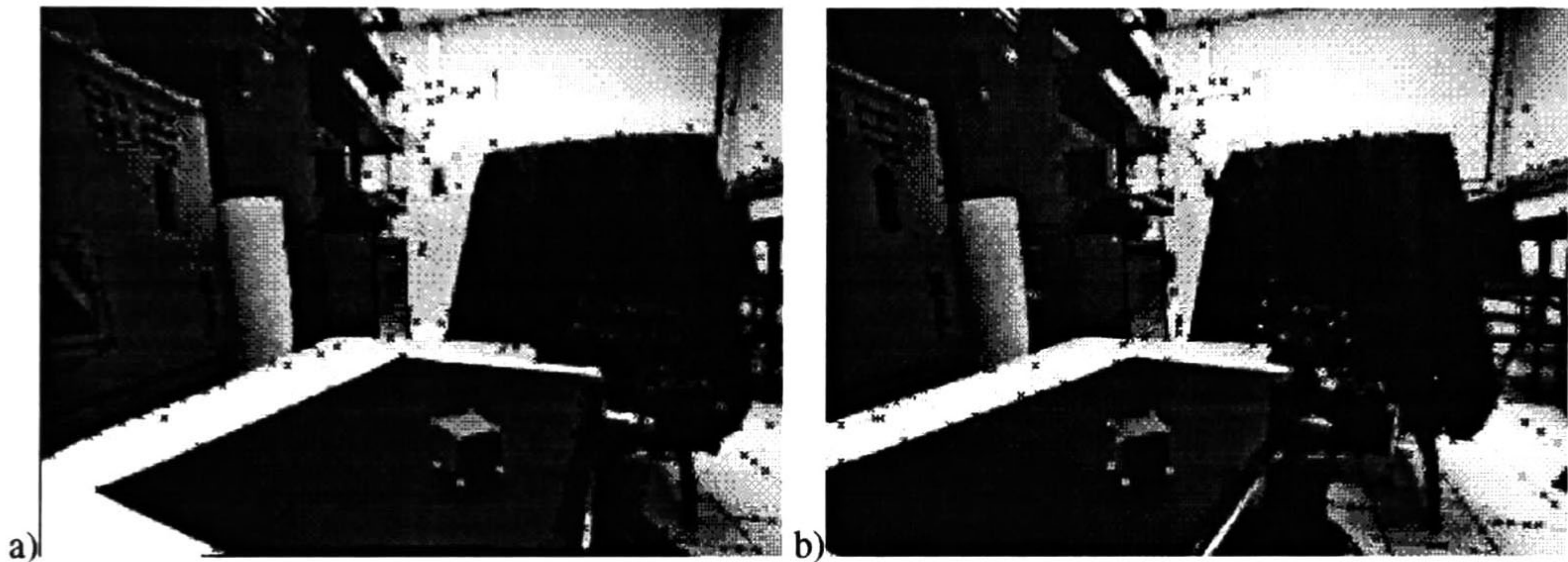


Figura 5.6: Correspondencias de puntos entre vistas. a) Segunda imagen izquierda. b) Segunda imagen derecha.

decisión. Toda esta planeación se puede ver como un ejemplo de un sistema de *Percepción-Acción*, siendo la obtención de primitivas la percepción y la ejecución de la tarea la acción.

### 5.3.1 Obtención de primitivas geométricas.

La obtención de las primitivas geométricas se hace a partir de los puntos reconstruidos. Los puntos reconstruidos representados en el álgebra geométrica  $G_{4,1,0}$  son primitivas por si solos. Las líneas se obtienen haciendo la unión de dos puntos diferentes, con el operador unión, ecuación 3.30. Los planos se obtienen con la unión de tres puntos en posición general, o de una línea y un punto fuera de ella. Similarmente los volúmenes se obtienen mediante la unión de cuatro puntos en posición general, una línea y dos puntos en posición general, o dos líneas que no se corten en el espacio. En general de un conjunto de puntos se pueden obtener muchas primitivas geométricas, incluso varias irrelevantes o redundantes. Es por este motivo que a partir de todas las primitivas iniciales se deben obtener las de relevancia, esto mediante la aplicación de algunas reglas como las que se presentan en la siguiente sección.

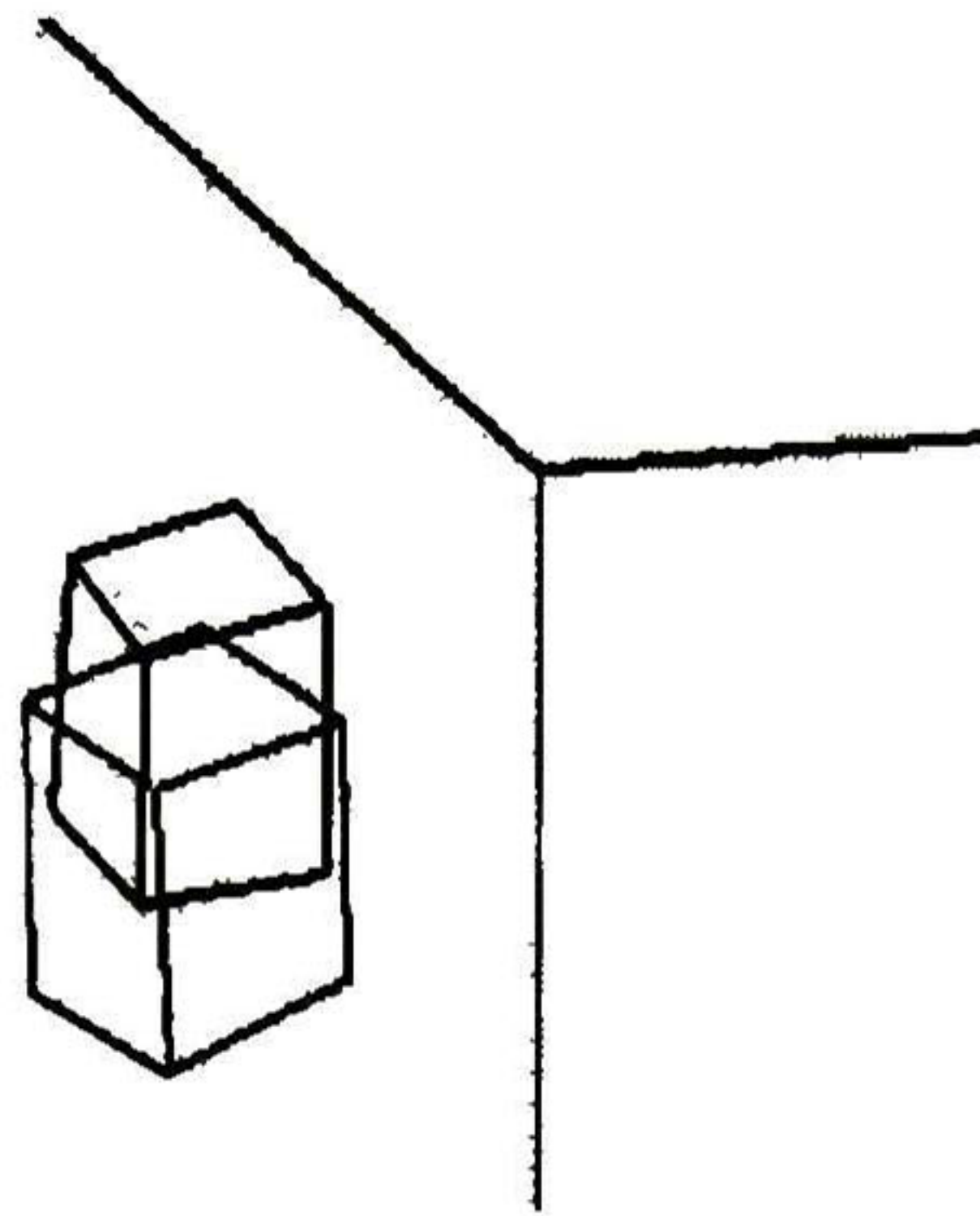


Figura 5.7: Puntos reconstruidos de diferentes vistas.

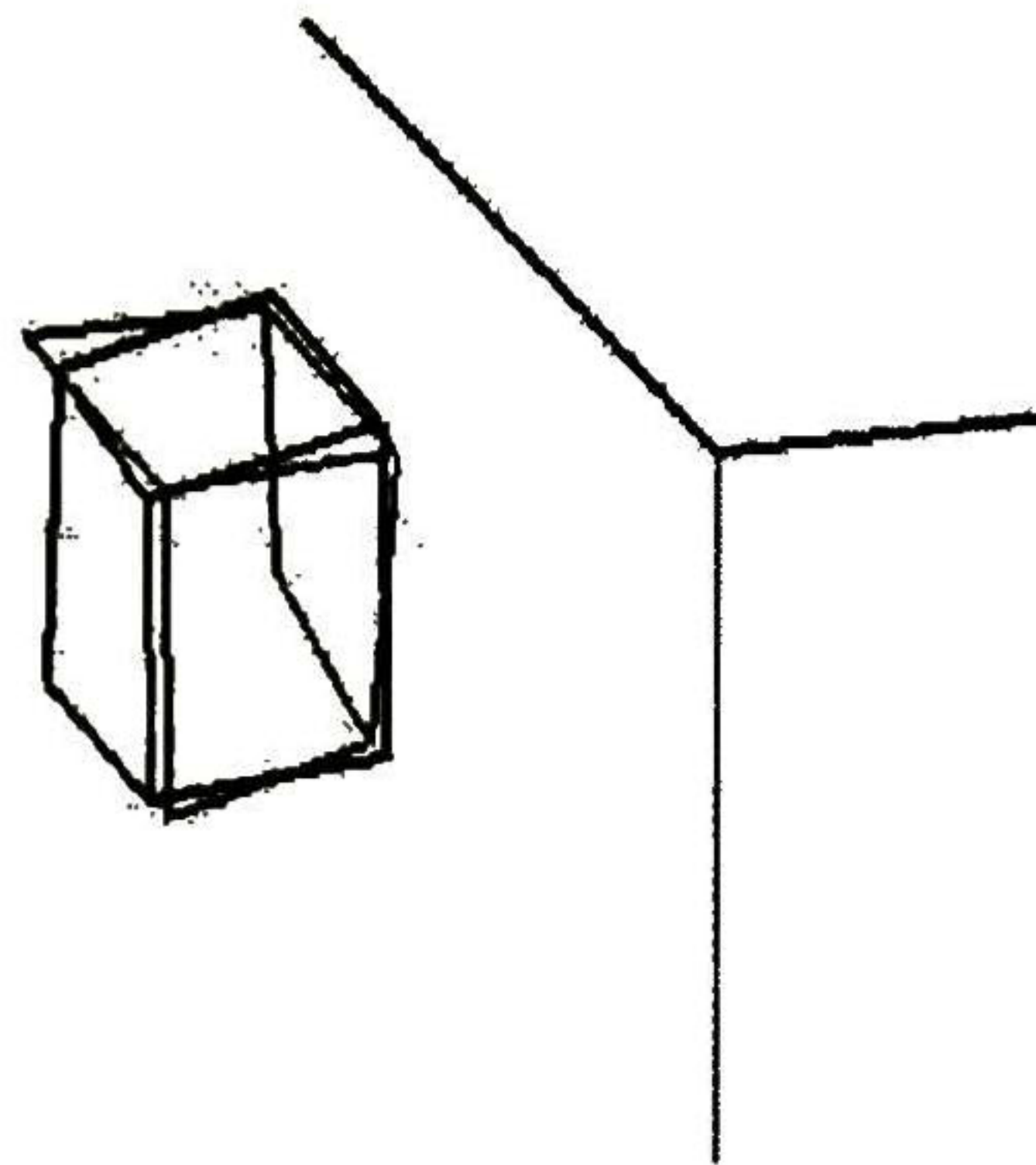


Figura 5.8: Reconstrucción con puntos de diferentes vistas ya fusionados.

### 5.3.2 Reglas de decisión.

Existen diversas reglas de decisión que están basadas en las relaciones de incidencia y los indicadores geométricos que se presentaron en el capítulo 3. En nuestro caso después de la fusión de vistas del cubo que se presentó en la sección anterior, se pretende que sea sujetado por una pinza que posee el robot como se presenta en la figura 5.9.

Para realizar la tarea de sujeción se reconstruyó la pinza del robot con la información de la última vista, ya que ésta es la última posición conocida de la pinza. La reconstrucción de la pinza y el cubo se presenta en la figura 5.10.

Ya con el escenario completo, la pinza y el cubo, el siguiente paso es decidir los puntos de sujeción. Esto se hace decidiendo primeramente los planos de sujeción, para esto se calculan todos los planos posibles del cubo. Seguidamente se calcula cuáles, de entre todos los planos, son los más paralelos. En el capítulo 3 se mencionó que si dos planos son paralelos su distancia dirigida será diferente de cero, sin embargo en la presencia de ruido, los planos se intersecarán y por lo tanto su distancia dirigida será cero. Entonces, en vez de usar la distancia dirigida, se decidió usar el operador "meet" ya que éste da como resultado la línea de intersección dos planos. En el caso de planos paralelos, la línea de intersección es una línea en el infinito. Al no ser los planos exactamente paralelos, la línea no estará en el infinito. Sin embargo los planos más cercanos a ser paralelos tendrán la línea de intersección muy cercana al plano en el infinito.

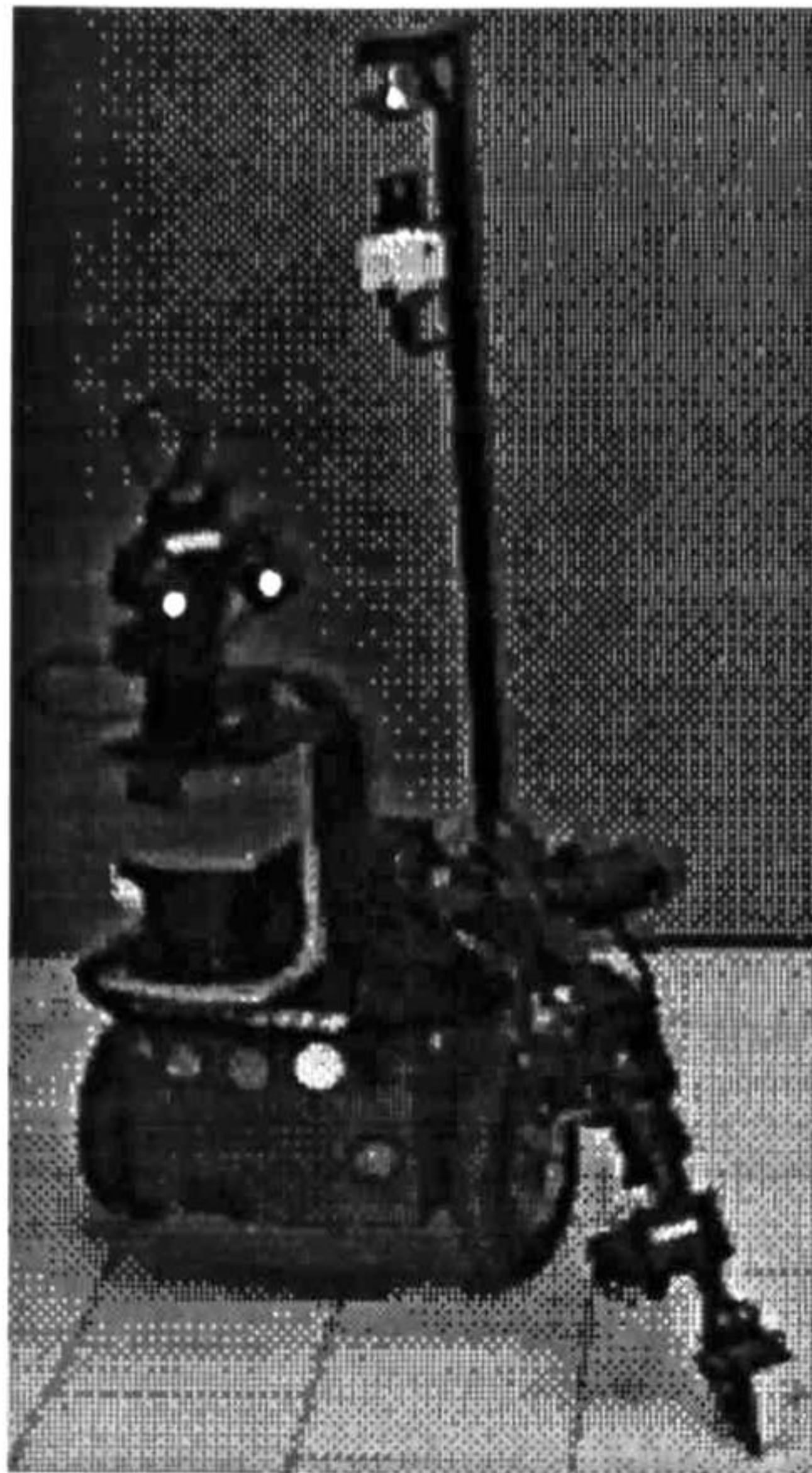


Figura 5.9: Robot GEOMETER sujetando un cubo.

Entonces la regla de decisión para elegir los planos paralelos será

$$\phi_i \cap \phi_j \approx l_{\infty}.$$

Esta es la expresión algebraica de la regla, pero en la práctica se hace con la componente dirigida de la línea. Una línea en el infinito tiene componente dirigida igual a cero, por lo que se establece un umbral que definirá si la componente dirigida de la línea es lo suficientemente pequeña. Los resultados obtenidos aplicando esta regla se presentan en la figura 5.11.

### 5.3.3 Ejecución.

Ya que se han decidido los planos de sujeción, cualesquiera dos puntos en estos planos son candidatos para ser los puntos de sujeción de la pinza. Se eligen los centroides de los planos de sujeción, y con un punto extra se genera un plano con el que se alinearé el plano que representa la pinza, estos dos planos se muestran en la figura 5.12.

Para alinear los planos del cubo y de la pinza se generaron tres pares de puntos correspondientes en los planos. A estos puntos se les aplicará el método de Lasenby[Lasenby95] para encontrar la transformación que alinearé los planos. Además de encontrar la transformación total, se interpoló la transformación, de tal forma que la pinza tenga un trayectoria suave, dicho trayectoria se muestra en la figura 5.13.

En este capítulo se presento un ejemplo de como se puede aplicar un enfoque geométrico a la manipulación de objetos. Aquí tan sólo se presento la sujeción, pero se puede aplicar un método similar, para trasladar los objetos a un lugar diferente de donde fueron tomados.

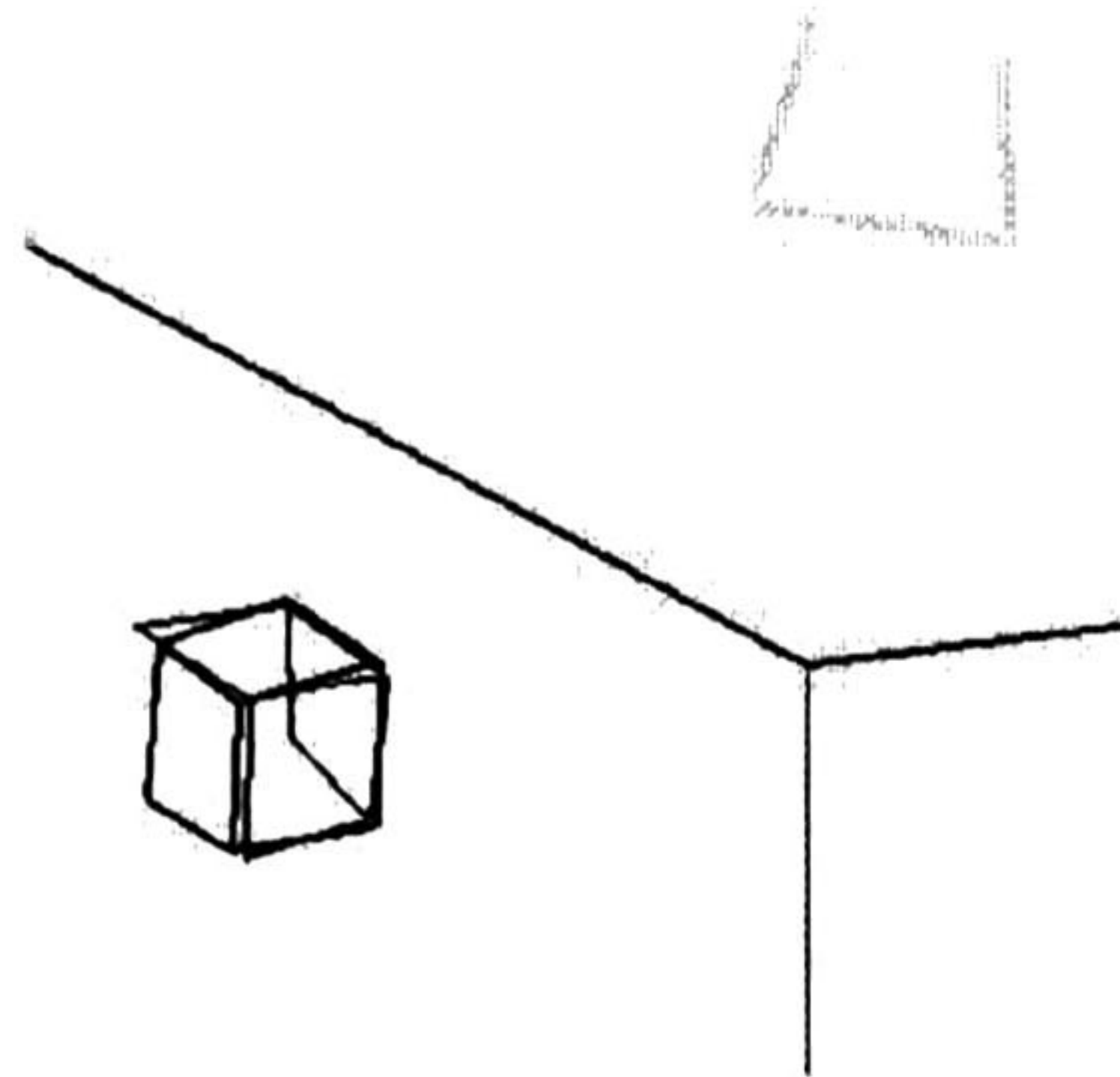


Figura 5.10: Reconstrucción de la pinza y el cubo (ya fusionado).

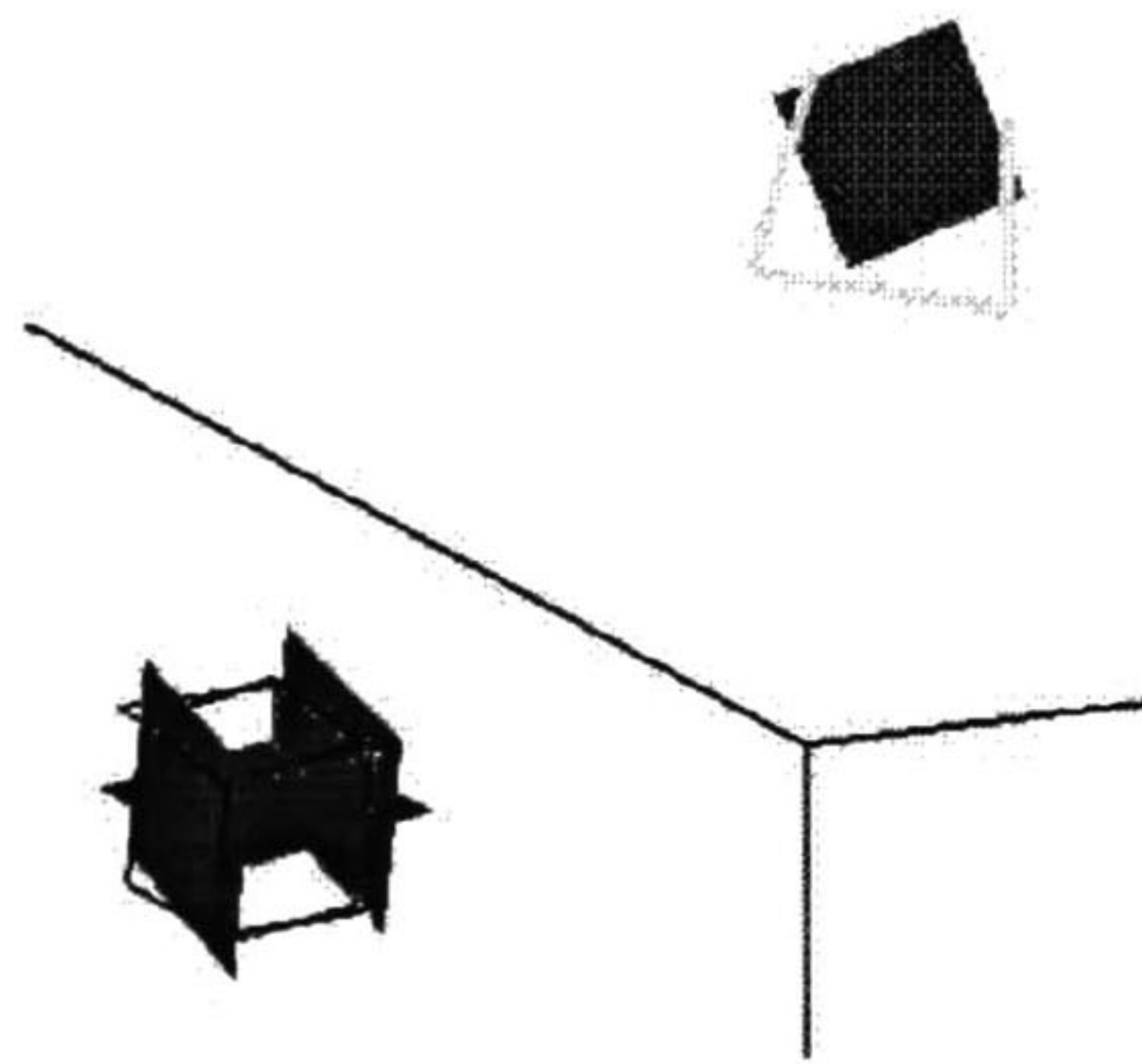


Figura 5.11: Planos paralelos encontrados con la regla  $\phi_i \cap \phi_j \approx \mathbf{l}_\infty$ .

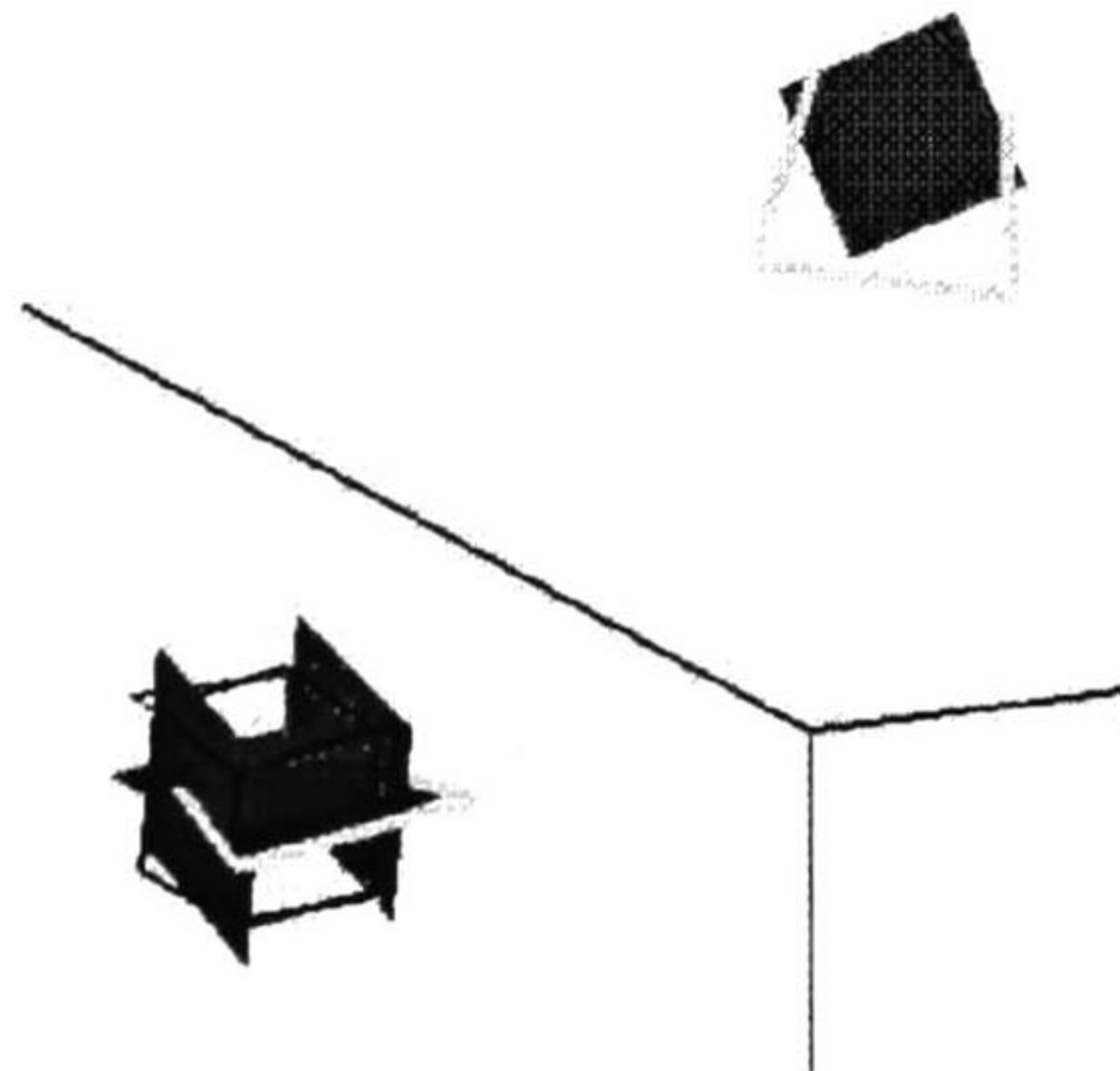


Figura 5.12: El plano de la pinza se alineará con el plano del cubo.



Figura 5.13: Trayectoria que deberá seguir la pinza para sujetar el cubo.



# Capítulo 6

## Conclusiones.

### 6.1 Resumen.

En este documento se presentó una manera de combinar diferentes áreas de investigación para un sólo objetivo. Las áreas que se combinaron son la geometría proyectiva y el álgebra geométrica, con posibilidad de relacionar la votación tensorial.

El objetivo del trabajo de tesis fue la planeación de una tarea a realizar con un brazo robótico, para ello fue necesario

- Captar varias imágenes del escenario con un cabezal binocular.
- Reconstruir puntos relevantes de las imágenes, para ellos se utiliza la geometría proyectiva y un poco de álgebra geométrica al expresar los puntos en 3D para su uso futuro.
- Fusionar los puntos en 3D en un solo ambiente, mediante una transformación rígida que se encuentra después de la reconstrucción.
- Con los datos fusionados, realizar la planeación de una trayectoria para poder sujetar un objeto, esto basándose en reglas de decisión expresadas en álgebra geométrica, específicamente en el álgebra  $G_{4,1,0}$ .
- Realizar una interpolación de la trayectoria, expresando el motor que representa la transformación como una secuencia de motores.
- Simular la variación continua del motor para mover una pinza y sujetar un objeto.

Además de todo lo anterior se demostró la efectividad de las álgebras geométricas en el área de la visión computacional como una herramienta de gran poder para la expresión de restricciones geométricas y reglas de decisión.

## 6.2 Trabajo futuro.

El trabajo futuro será dedicado a relacionar los sistemas matemáticos considerados y resolver varios problemas técnicos. A continuación enumeramos algunas tareas que merecerán nuestra atención:

- La formalización de la votación tensorial en tales términos que se pueda relacionar con el álgebra geométrica.
- La expresión de restricciones en la configuración de objetos en el espacio.
- La validación de restricciones mediante la votación tensorial. Tal vez sea posible expresar el producto tensorial exterior (en forma matricial  $\hat{e}_1 \hat{e}_1^T$ ) equivalentemente como el producto Clifford por el pseudoescalar ( $\sigma_1 I$ ), donde  $I$  es el pseudoescalar en el álgebra en cuestión,  $(\mathcal{G}_{3,0,0})$ .
- La utilización de la votación tensorial no sólo para la extracción de características sobresalientes, sino también para la fusión de datos en 3-D provenientes de varias fuentes (laser, cámaras, sonar, odometría).
- La extensión de la manipulación objetos a no solo sujetar como es el caso del ejemplo que se presento en ésta tesis, sino también al traslado de objetos. Además se piensa en generar una máquina de inferencia para ejecutar la toma de decisiones en el ciclo de percepción-acción.
- La automatización de la fusión de vistas, para que los datos obtenidos de una secuencia de vistas se agreguen al ambiente automática y sucesivamente, para tener una descripción aún más completa de la escena. Para esto se necesita la implementación de algún algoritmo de búsqueda de correspondencias en tiempo real para aplicarlo a la secuencia de vistas y poder aplicar el método de búsqueda de transformaciones rígidas.

# Bibliografía

- [Hartley00] Hartley, R. and Zisserman, A., *Multiple View Geometry in Computer Vision*. Cambridge University Press, UK 2000.
- [Faugeras9] Faugeras, O.D., *Three Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, New York 1993.
- [Knutsson95] H. Knutsson and G.H. Granlund, *Signal Processing for Computer Vision*, Kluwer Academic Publishers 1995.
- [Westin94] C.-F. Westin., *A Tensor Framework for Multidimensional Signal Processing*, Ph.D. Thesis, Linköping University, Sweden 1994.
- [Tang99] Tang, C.-K., *Tensor Voting in Computer Vision, Visualizaion, and Higher Dimensional Inferences*, Ph.D. Thesis, University of Southern California, 1999.
- [Lee98] Lee, M.-S., *Tensor Voting for Salient Feature Inference in Computer Vision*, Ph.D. Thesis, University of Southern California, 1998.
- [Guy96] Guy, G., *Inference of Multiple Curves and Surfaces from Sparse Data*, Ph.D. Thesis, University of Southern California, 1996.
- [Medioni99] G. Medioni, M.-S. Lee, C.-K- Tang, *A computational Framework for Feature Extraction and Segmentation*, Elsevier 1999.
- [Hay5] Hay, G.E., *Vector and Tensor Analysis*, Dover Publications Inc., New York 1953.
- [Lorensen87] Lorensen, W.E., Cline, H.E., *Marching Cubes: A High Resolution 3D Surface Reconstruction Algorithm*, Computer Graphics, 21(3), pp.163-169, 1987.
- [Lounesto87] P. Lounesto, R. Mikkola, and V. Vierros, *CLICAL User Manual*. Helsinki University of Technology, Institute of Mathematics, Research Reports A248, Helsinki 1987.
- [Lounesto97] Lousnesto, P., *Clifford Algebras and Spinors*. Cambridge University Press 1997.
- [Sobczyk01] Sobczyk G. *Universal Geometric Algebra*. In advances in Geometric Algebra with Aplications in Science and Engineering. Bayro-Corrochano E. and Sobczyk G. (eds.), Chap. 2. Birkhäuser, New York 2001, pp. 18-41.

- [Poza01] Pozo J.M. and Sobczyk G. *Realizations of the Conformal Group*. In advances in Geometric Algebra with Applications in Science and Engineering. Bayro-Corrochano E. and Sobczyk G. (eds.), Chap. 3. Birkhäuser, New York 2001, pp. 42-60.
- [Bayro01] Bayro-Corrochano E. 2001. *Geometric Computing for Perception Action Systems*, Springer Verlag, Boston 2001.
- [Tsai87] Tsai, Roger Y. *A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using off-The Shelf TV Cameras and Lenses*, IEEE Journal of Robotics and Automation, vol. RA-3, no. 4, August 1987, pp. 323-344.
- [Harris88] Harris. C.J.Harris and M.Stephens. *A combined corner and edge detector*. Proc. 4th Alvey Vision Conf. Manchester, pp. 147-151, 1988.
- [Lasenby95] Lasenby, J., *Geometric Algebra: Applications in Engineering*. Clifford (Geometric) Algebras, with Applications in Physics, Mathematics, and Engineering. Baylis W.E. (ed.), Chapter 29. Birkhäuser, Boston 1995, pp. .
- [Bayro96] Bayro-Corrochano, E., Lasenby, J. and Sommer, G. 1996. *Geometric Algebra: a framework for computing point and line correspondences and projective structure using n-uncalibrated cameras*. Proceedings of the International Conference on Pattern Recognition (ICPR'96), Vienna, August 1996., Vol.I, pp. 393-397.
- [Lasenby99] Lasenby, J., Bayro-Corrochano, E. 1999. *Analysis and Computation of Projective Invariants from Multiple Views in the Geometric Algebra Framework*. In Special Issue on Invariants for Pattern Recognition and Classification, ed. M.A. Rodrigues. Int. Journal of Pattern Recognition and Artificial Intelligence, Vol13, No. 8, December 1999, pp. 1105-1121.
- [Semple79] Semple, J.G. and Kneebone, G.T. *Algebraic Projective Geometry*. Oxford University Press, 1979.
- [Hestenes66] Hestenes D. *Space-Time Algebra*. Gordon and Breach, 1966.
- [Hestenes84] Hestenes D. and Sobczyk G. *Clifford Algebra to Geometric Calculus: A Unified Language for Mathematics and Physics*, 1984. D. Reidel, Dordrecht.



**Centro de Investigación y de Estudios Avanzados del  
IPN**

**Unidad Guadalajara**

El Jurado designado por la Unidad Guadalajara del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, aprobó la tesis: MÉTODOS GEOMÉTRICOS PARA LA MANIPULACIÓN VISUAL DE OBJETOS del(a) C. Jaime Silverio ORTEGÓN AGUILAR el día 7 de Agosto de 2002 .

**Dr. Valeri KORJIK**  
Profesor Investigador 3C  
CINVESTAV GDL  
Guadalajara

**Dr. Eduardo Jose BAYRO  
CORROCHANO**  
Investigador Cinvestav 3B  
CINVESTAV GDL  
Guadalajara

**Dr. Félix Francisco RAMOS  
CORCHADO**  
Investigador Cinvestav 2A  
CINVESTAV GDL  
Guadalajara



CINVESTAV  
BIBLIOTECA CENTRAL



SSIT000004445