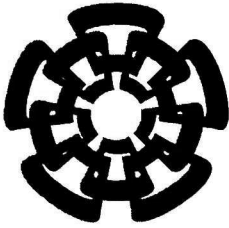


xx(12 9993.7)



CINVESTAV

Centro de Investigación y de Estudios Avanzados del I.P.N.
Unidad Guadalajara

Visión Estereoscópica usando Álgebra Geométrica Conformal

**CINVESTAV
IPN
ADQUISICION
DE LIBROS**

Tesis que presenta:

David Israel González Aguirre

para obtener el grado de:

Maestro en Ciencias

en la especialidad de:

Ingeniería Eléctrica

Director de Tesis

Dr. Eduardo José Bayro Corrochano



CENTRO DE INVESTIGACION DE
ESTUDIOS AVANZADOS DE I.P.N.

SERVICIOS BIBLIOGRAFICOS

Guadalajara, Jalisco, Julio de 2006.

CLASIF.:	TRIGS. G8 G66 2006
ADQUIS:	SSI-402
FECHA:	7-III-2007
PROCED.	Don. 2007
	\$ _____

ID: 129286-1001

Visión Estereoscópica usando Álgebra Geométrica Conformal

**Tesis de Maestría en Ciencias
Ingeniería Eléctrica**

Por:

David Israel González Aguirre
Ingeniero Electrónico en Computación
Centro de Enseñanza Técnica Industrial 1998-2002

Becario de Conacyt, expediente no. 180868

Director de Tesis
Dr. Eduardo José Bayro Corrochano

CINVESTAV del IPN Unidad Guadalajara, Julio de 2006.

Agradecimientos

A Dios por la vida y todo cuanto existe, a mis padres por que todo lo que soy es gracias a ellos, su esfuerzo y esmero. A mis hermanos por vivir, compartir y enseñarme lo que aun no sabemos, de igual forma a mis tíos que son mis segundos padres. A mi novia por su fe y cariño siempre presentes.

Especial agradecimiento a mi asesor el Prof. Dr. Eduardo José Bayro Corrochano por todo el tiempo y conocimientos invertidos. A los compañeros del laboratorio GeovisLab Julio, Rafael, Rubén, Jaime, Miguel-Linux, Miguel-Windows, Falcón, Jorge, Carlos, Nancy, Leo, a mis compañeros de generación. A mis amigos Alfredo, Rubén, Jorge, Juancho, Luís, Roberto, Liz, Alberto, Lissete, Araceli por su amistad, apoyo y paciencia. A todos aquellos que de una u otra forma han estado mi lado. A Cinvestav y en particular a todas las personas de Cinvestav unidad Guadalajara por su esfuerzo continuo, al **Conacyt** por hacer posible la actividad científica en nuestro país y en particular por el apoyo que hizo posibles el presente trabajo mediante la beca numero 180868. Finalmente a todos los mexicanos que gracias a sus impuestos hacen posible la existencia de instituciones como el Cinvestav y el CETI.

Tesis de Maestría en Ciencias en Ingeniería Eléctrica

Presentada por:

Ing. David Israel González Aguirre

Para obtener el título de:

Maestro en Ciencias

Con especialidad en:

Ingeniería Eléctrica

Prof. Dr. Eduardo José Bayro
Corrochano
Director de Tesis

Dr. Mario Angel Siller González
Pico
Sinodal

Dr. Marco Antonio Pérez
Cisneros
Sinodal

4 de Julio de 2006

Índice general

1. Introducción	1
1.1. Motivación .	4
1.2. Estado del Arte	5
1.3. Problema y Objetivos .	5
1.4. Organización y Notación	7
2. Modelo de un Cabezal Binocular	8
2.1. Motivación Biológica	8
2.2. Visión Estereoscópica	10
2.2.1. Correspondencias	11
2.2.2. Reconstrucción del espacio Visual	16
2.3. Modelo conformal del Cabezal Binocular	23
3. Procesamiento de Imágenes Estereoscópicas	27
3.1. Calibración de las Cámaras	27
3.1.1. Calibración del Enfoque	28

<i>ÍNDICE GENERAL</i>	II
3.1.2. Calibración de Distorsión Radial y Tangencial	33
3.1.3. Calibración Geométrica .	33
3.2. Rectificación de Imágenes	33
4. Reconstrucción Activa del Espacio Visual	37
4.1. Cinemática Directa e Inversa del Cabezal	38
4.2. Detección de Primitivas Geométricas	44
4.3. Análisis en Multiresolución	49
4.4. Segmentación en Profundidad	51
5. Aplicaciones	53
5.1. Interfaz Perceptual	53
5.2. Manipulación de Objetos	55
6. Conclusiones	57
A. Álgebra Geométrica	59
A.1. El Producto Exterior	60
A.1.1. Espacio Exterior Nulo	61
A.2. <i>Blades</i>	62
A.2.1. Magnitud de <i>Blades</i>	64
A.3. El Producto Interior	65
A.3.1. Inversa de <i>Blades</i>	68

<i>ÍNDICE GENERAL</i>	III
A.3.2. Espacio Interior Nulo	68
A.4. Dualidad	70
A.5. Intersecciones	71
A.6. Producto Geométrico	74
A.7. Reflexión .	75
A.8. Rotaciones	76
A.9. Espacio Proyectivo	79
A.9.1. Álgebra Geométrica en espacio Proyectivo	81
A.9.2. Espacios Nulos en Espacio Proyectivo $\mathbb{P}E^n$	81
A.9.3. Reflexiones en Espacio Proyectivo	84
A.9.4. Rotaciones en Espacio Proyectivo	85
A.9.5. Modelo de una Cámara	85
A.10. Espacio Conformal	88
A.10.1. Álgebra Geométrica en Espacio Conformal	94
A.10.2. Primitivas Geométricas	95
A.10.3. Puntos	96
A.10.4. Esferas	96
A.10.5. Planos	98
A.10.6. Inversion en Espacio Conformal	100
A.10.7. Rotación en Espacio Conformal	101
A.10.8. Translación en Espacio Conformal	102

<i>ÍNDICE GENERAL</i>	IV
A.10.9. Espacios Euclidiano y Proyectivo en Espacio Conformal	103
Bibliografía	105

Índice de figuras

1.1. Evolución tecnológica.	1
1.2. Revolución cibernética.	2
1.3. Visión computacional.	3
2.1. El sistema visual según lo imaginado por René Descartes (1596-1650). Las dos imágenes retinianas se enfocan sobre la glándula pineal, “El recinto de la imaginación y sentido común”	8
2.2. Los ejes ópticos del sistema de visión estereoscópica en el ser humano.	10
2.3. Configuración epipolar.	12
2.4. a) Ejemplo de oclusión, el sujeto se encuentra frente al sistema a una distancia reducida, la información (regiones de los oídos) de la imagen izquierda no puede ser validada en la derecha y viceversa. b) Región con textura uniforme, los métodos locales son susceptibles a generar correlaciones erróneas debido a la baja varianza en el vecindario. c) Regiones con contenido muy similar separados por segmentos uniformes, son difíciles de correlacionar en forma adecuada sin el uso de restricciones.	13

2.5. Líneas de disparidad, el rango (subespacio) contenido entre cada incremento discreto de disparidad crece en forma cuadrática.	14
2.6. Máscaras de ponderación (aportación) para la región de interés o sub-ventana. a) Función caja. b) Función gaussiana.	15
2.7. Región o vecindario del píxel central, así como la varianza resultado de aplicar el operador C_s .	16
2.8. a) Representación 3D de la intensidad de los píxeles en el vecindario, el plano denota la intensidad del píxel central. b) La señal unidimensional (vector $\Omega_{(x,y)}$) resultado de aplicar el operador C_s .	18
2.9. Función de costo $\tilde{h}(i, j)$, asocia un valor precomputado a la intensidad de dos píxeles para la correlación de regiones.	19
2.10. a) Imagen original. b) Imagen filtrada, los píxeles con una region de soporte con una varianza bajo del umbral U_G no son considerados en etapas posteriores.	21
2.11. Vista en planta del sistema. a) Cálculo de función de correlación y determinación de línea epipolar de disparidad mediante esquema de programación dinámica b) Mapas de disparidad para dos escenas.	22
2.12. a) Mapa de disparidad con falsas correlaciones (<i>Outliers</i>). b) Aplicación del filtro de mediana selectiva para la eliminación de las falsas correlaciones.	23

- 2.13. a) La horóptera y los centros de las esferas divididas por la línea bisectriz cuando los δ de la profundidad se incrementan. b) La horóptera esférica y la esfera unitaria, la horóptera dependiendo del ángulo azimut c) Configuración conformal de la horóptera, el centro de la cámara izquierda L_C , el centro de la cámara derecha R_C y el punto de fijación F_p definen un círculo, el cual mediante la variación del azimut κ define a su vez una familia de esferas o Horoptoesferas. 24
- 2.14. a) Representación polar del espacio geométrico visual. b) Planos de las imágenes y plano de azimut κ (configuración epipolar). 25
- 3.1. a) Región de alta frecuencia limitada mediante el radio de corte R_c y la mitad de la imagen n en su representación de densidad espectral. b) Máxima cantidad de energía $E_R(t)$ en la región de alta frecuencia en el tiempo t_n . 29
- 3.2. Función de ponderación ecuación (3.3). 30
- 3.3. Imágenes izquierda y derecha aislando el contenido previamente acoplado mediante la función Φ . 32
- 3.4. a) Imagen original con distorsión radial. b) Imagen rectificadas. c) Error entre la imagen original y la rectificadas, note que el error crece a razón de un polinomio de segundo grado con argumento radial a partir del píxel central (no necesariamente el centro de la imagen). 34
- 3.5. Resultado de la calibración mediante la implementación Jean-Yves Bouguet. 35
- 3.6. Función de calibración de imagen (*Warp*). 36
- 4.1. Ego-mundo del sistema. 37

4.2. El campo de visión 3D es el espacio contenido en la intersección de los campos visuales de las dos cámaras, el plano de corte en cercanía N y el plano de corte en lejanía F	39
4.3. Modelo 3D del sistema de visión estereoscópica, puntos importantes, ejes de rotación <i>Pan-Tilt</i> .	40
4.4. Algoritmo de cinemática inversa.	42
4.5. Ejemplo de primitiva geométrica en un ambiente semiestructurado.	45
4.6. Segmentación en color de la imagen para la obtención de las regiones con características cromáticas similares. a) Imagen original. b) Imagen con reducción del espacio cromático. c) Aglutinamiento mediante el uso de la variación de la media. d) Fronteras entre las regiones cromáticas.	46
4.7. a) Imagen segmentada para rectificación geométrica. b) Región seleccionada con su centroide. c) Ajuste de plano basado en los vectores propios de la matriz de covarianza.	47
4.8. Implementación del algoritmo de generación de superficies de iso-disparidad en multiresolución, los resultados en los niveles inferiores de la pirámide se propagan a niveles superiores optimizando el proceso.	50

4.9.	1) Imagen del escenario $I(x, y, t_n)$. 2) Imagen calibrada y filtrada $I_c(x, y, t_n)$. 3) Mapa de iso-disparidad $Disp(x, y, t_n)$. 4) Imagen del escenario $I(x, y, t_{n+1})$. 5) Imagen calibrada y filtrada $I_c(x, y, t_{n+1})$. 6) Mapa de iso-disparidad $Disp(x, y, t_{n+1})$. 7) Diferencia en profundidad $D_p(x, y, t_{n+1}) = Disp(x, y, t_{n+1}) -$ $Disp(x, y, t_n) $. 8) Diferencia en profundidad filtrada mediante mediana selectiva $D_{pf}(x, y, t_{n+1})$. 9) Umbralización $M(x, y, t_{n+1})(mascara)$. 10) Resultado de la segmentación en profundidad $I_s(x, y, t_{n+1})($ Segmento de imagen) $Disp_s(x, y, t_{n+1})($ Segmento de disparidad)	52
5.1.	Sistema con ego-visión y usuario en realidad virtual.	55
5.2.	a) Imágenes de entrada para el algoritmo de manipulación. b) Manipu- lación de esfera mediante la aplicación del álgebra geométrica conformal.	56
6.1.	Visualización en realidad virtual del proyecto PAC.	57
A.1.	William Kingdom Clifford (1845-1879).	59
A.2.	Punto a y su respectivo $\text{NO}(a)$ como una línea que pasa por el origen y a .	62
A.3.	El <i>Pseudoescalar</i> o <i>Hipervolumen</i> $I_E = e_1 \wedge e_2 \wedge e_3 = e_{123} \in \mathcal{G}_{3,0}$ resultado del producto exterior de $a = e_1, b = e_2$ y $c = e_3$.	63
A.4.	Vectores $a, b \in \mathbb{R}^3$, representación de la magnitud del <i>bivector</i> $ b \wedge a $.	64
A.5.	Producto Interior de un <i>bivector</i> y un <i>vector</i> .	66
A.6.	Espacio interior nulo de un <i>bivector</i> $\text{NI}(S = a \wedge b)$, espacio exterior nulo $\text{NO}(S = a \wedge b)$.	69
A.7.	Intersección entre el plano $A = e_1 \wedge e_2$ y la línea $B = \alpha e_1 + \beta e_2$.	73
A.8.	Reflexión del <i>vector</i> x respecto al <i>vector</i> y .	77

A.9. Rotación del <i>vector</i> a mediante dos reflexiones consecutivas con los <i>vec-</i> <i>tores</i> unitarios n y m , nótese el eje de rotación es la línea $\text{NI}(n \wedge m)$.	78
A.10. Espacio proyectivo \mathbb{PE}^2 resultado de envolver el espacio Euclidiano \mathbb{E}^2 con el operador ρ .	79
A.11. Línea representada mediante el <i>bivector</i> $\text{NO}_E(A \wedge B)$.	83
A.12. Modelo de una cámara, π_c y π_l denotan los planos que delimitan el campo visual, esto se debe a que fuera de estos planos no sera posible adquirir una imagen con el enfoque correcto.	86
A.13. Intersección del rayo óptico Λ formado por el centro de la cámara C y el punto en el espacio X con el plano de la imagen.	87
A.14. Proyección estereográfica de los puntos x e y en la hiperesfera unitaria.	89
A.15. Proyección estereográfica del círculo y la línea a la hiperesfera unitaria.	90
A.16. Proceso de envolvimiento de un punto $x \in \mathbb{E}^1$, primero proyección es- tereográfica $\kappa(x)$ y posteriormente homogenización $\rho(\kappa(x))$.	91
A.17. Puntos conformales con la componente e_0 normalizada en el hiperplano afino \mathbb{H}_a^1 .	93
A.18. Inversión de un vector $x \in \mathbb{E}^1$ con base en el mapeo estereográfico $\kappa(x)$ y la reflexión sobre el eje e_+ .	100

Capítulo 1

Introducción

La tecnología como extensión de las capacidades del ser humano implica que las máquinas son (en cierta manera) amplificadores de las habilidades, aptitudes y sentidos del hombre. Dichos artefactos proporcionan los instrumentos para modificar y



Figura 1.1: Evolución tecnológica.

manipular el entorno. Con la transformación de materia-energía en términos químicos y mecánicos el género humano ha logrado la subsistencia y crecimiento. Desde la edad de piedra con los primeros artefactos hasta la revolución industrial con las máquinas de vapor, el hombre ha generado mecanismos para producir satisfactores, (en un sentido más amplio) para generar un modo de vida como se conoció hasta principios del siglo pasado. Posteriormente gracias a los avances en física, matemáticas y otras ciencias

afines, así como el desarrollo de dispositivos eléctricos y electrónicos dieron lugar a la aplicación de modelos como el de *Alan Turing* (1912-1954) y la matemática de *George Boole* (1815-1864) para construir las primeras computadoras. Con el incremento de las



Figura 1.2: Revolución cibernética.

capacidades de procesamiento y almacenamiento de información en las computadoras surgieron distintos modos de construir rutinas de procesamiento. Dichas rutinas o programas evolucionaron y con ello no sólo cambió la forma en que se introducían datos y comandos a la computadora sino también la forma en que se presentaban los resultados. La interconexión entre computadoras y dispositivos electrónicos amplificó los campos de acción de dichas máquinas a casi cualquier ámbito del ser humano. Dicha evolución tiene implicaciones científicas, tecnológicas y sociales enormes, tal es la magnitud que algunos historiadores hablan de la revolución cibernética¹ en analogía a la revolución industrial.

El modo de vida contemporáneo está basado en una plataforma tecnológica tan amplia y sutilmente extendida que las máquinas no sólo definen la forma en que se producen los satisfactores, sino también la calidad de vida del individuo, el nivel socioeconómico-cultural, el modo en que se expresan las personas, la forma en que se comunicarán los individuos entre sí, la manera en que piensan, perciben e interactúan

¹ Era de la información.

con su entorno. Es posible decir que las máquinas son ahora extensiones del cuerpo y de la mente de los seres humanos.

Las máquinas que amplifican las capacidades de percepción son aquellas relacionadas con los sentidos. Bajo este marco el sentido de la visión es sin duda uno de los retos más interesantes, dada la basta cantidad de información contenida en el espacio visual, tanto en un sentido semántico como de procesamiento y almacenamiento.

La visión computacional como el proceso adquisición, digitalización, ecualización, filtrado, segmentación, caracterización, interpretación, selección y organización de información obtenida por medio de imágenes es uno de los más amplios esfuerzos realizados por la comunidad científica para dotar a las máquinas de capacidades de percepción visual, lo que implica aumentar las capacidades humanas de percepción y así lograr un manejo objetivo de la información contenida en el espacio visual. La presente tesis es

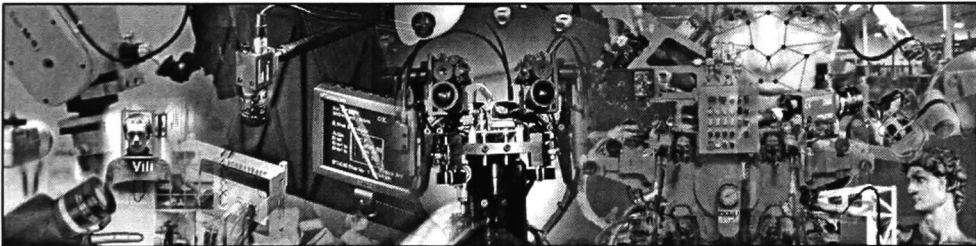


Figura 1.3: Visión computacional.

un esfuerzo en el desarrollo de nuevas técnicas de reconstrucción geométrica del espacio visual haciendo uso del álgebra geométrica conformal como un aparato matemático moderno, sumamente poderoso y prometedor.

Las propuestas principales de este trabajo son; un modelo conformal de la visión humana inspirado en bases biológicas, una técnica para la calibración del enfoque del sistema estereoscópico, un método de obtención y filtrado de superficies de iso-disparidad,

así como la aplicación de técnicas para ajustar planos y operaciones de incidencia con álgebra geométrica conformal. Estos conceptos, incorporando análisis en multiresolución y la técnica de segmentación en profundidad, han sido implementados en un sistema que consta de un cabezal robótico y un sistema de visión estereoscópico. Todo lo anterior puede ser observado y controlado en el puerto visual de realidad virtual del sistema. Se han obtenido buenos resultados en imágenes sintéticas y reales, permitido hacer uso del sistema para etapas posteriores de investigación, así como futuras aplicaciones.

1.1. Motivación

En un futuro cercano la presencia de computadoras estará aun más difundida en la vida cotidiana, la capacidad para procesar información visual en los sistemas será un factor decisivo en la manera que interactuamos con ellos y la forma que ellos interactúen en el mundo. Las interfaces basadas en visión artificial (las cuales cierran el ciclo de generación de imagen sintética en la computadora y el análisis del espacio visual) proveerán nuevos esquemas y analogías tanto físicas como mentales para la interacción hombre-máquina-entorno en diversos campos del hacer humano, desde el trabajo hasta la recreación. Los sistemas con capacidad de procesar información visual en forma adecuada y eficiente tendrán por ende una respuesta adaptativa al medio (más rápida, apropiada y flexible) con lo que se lograrán robots más autónomos, cajeros automáticos más seguros, casa y vehículos más inteligentes, procesos más confiables y económicos, etcétera.

Un lenguaje matemático apropiado, sintético, poderoso y elegante que ofrece entidades geométricas sobre las cuales se opera como variables en un álgebra, donde las rotaciones y translaciones son esencialmente lo mismo, donde la intersección de prim-

itivas geométricas es un operador generalizado, es el álgebra geométrica conformal, la cual alienta a la reflexión sobre el espacio e invita a hacer uso de sus características.

1.2. Estado del Arte

Dado que las áreas comprendidas en la presente tesis abarcan temas como procesamiento de imágenes, análisis estadístico, robótica (las cuales son muy bastas y dado que sólo se usó algunos aspectos en particular) se procederá a citar aquellos cuya importancia lo requiera.

Muchos autores han abordado el tema de la reconstrucción tridimensional partiendo de imágenes tal es el caso de R. Hartley y A. Zisserman[13], M. Pollefeys [22]. En los trabajos anteriores se hace hincapié en la importancia de las esquinas o bordes en las imágenes para obtener la correlación entre imágenes. Existen trabajos que mejoran la calidad de la reconstrucción mediante métodos invasivos, es decir, uso de artefactos como proyectores o luces controladas para mejorar la calidad de la reconstrucción. Por otra parte la determinación de mapas de disparidad densos sin métodos invasivos se encuentran trabajos como el de E. Trucco [7] que incorpora ideas para aceleran el proceso de obtención de disparidades. En lo referente al uso del análisis multiresolución para diversas aplicaciones se denota el trabajo realizado por J. Bouguet[24] en particular la formulación en multiresolución de C. Lucas y T. Kanade [25].

1.3. Problema y Objetivos

En la presente tesis consiste se usará cuello robótico (*Pan-Tilt-Unit*) y dos cámaras montadas sobre dicha unidad (sistema estereoscópico), se pretende obtener una recon-

strucción tridimensional del espacio visual alcanzable con los dos grados de libertad que el mecanismo ofrece, así mismo se detectarán primitivas geométricas para aplicaciones posteriores. Una vez reconstruido el escenario que comprende $\pm 165^\circ \times \pm 65^\circ$ (ver la figura 4.5) se dispondrá en tres resoluciones distintas para lograr la segmentación en profundidad de objetos o personas para aplicaciones de interacción hombre-máquina, manipulación de objetos, etcétera. La intención es que los algoritmos desarrollados sean eficientes para conseguir una frecuencia de trabajo adecuada²

En forma estratificada las tareas a realizar son las siguientes:

- Formalizar el modelo de visión humana.
- Calibrar cámaras.
- Rectificar imágenes.
- Obtener Mapas de iso-disparidad.
- Controlar el cabezal.

Obtener primitivas geométricas.

Segmentar en profundidad.

- Aplicar el sistema en interacción hombre-máquina y manipulación de objetos.

Los objetivos trazados son desarrollar un sistema estable, capaz de hacer reconstrucciones de su entorno en un tiempo adecuado, posteriormente con base en el modelo geométrico del espacio circúndante reconocer la posición relativa de entidades para seguimiento u otro tipo de actividades como suministrar comandos al sistema mediante

² Entendiendo como adecuada la frecuencia de trabajo particular de cada tarea, aproximadamente 15 – 25 Hz.

el posicionamiento de alguna entidad en el espacio (mano u objeto geométrico). La intención es abordar la visión computacional con herramientas del álgebra geométrica, incursionar y proponer técnicas de procesamiento de imágenes para diversas aplicaciones.

1.4. Organización y Notación

En el capítulo 2 se presentan los conceptos del álgebra geométrica partiendo del espacio euclidiano tridimensional \mathbb{E}^3 , pasando por el espacio proyectivo \mathbb{P}^3 y finalizando con el espacio conformal \mathbb{PK}^3 . En el capítulo 3 se plantea y formaliza el modelo de visión humana con el cual se implementa el sistema físico del cabezal robótico. Posteriormente en el capítulo 4 se adentra en las etapas de procesamiento, calibración y rectificación de imágenes propias del sistema. En el capítulo 5 se detalla el método de reconstrucción tridimensional, control del cuello robótico, detección de primitivas geométricas y segmentación en profundidad. Las aplicaciones de dicho modelo serán presentadas junto con sus resultados en el capítulo 6. Las conclusiones se presentan en el capítulo 7.

Los vectores o puntos en R^3 se presentan en *cursivas*, mientras que los puntos conformales se presentaran en **negritas**.

Capítulo 2

Modelo de un Cabezal Binocular

2.1. Motivación Biológica

De manera natural nuestro mecanismo de visión es estéreo, es decir, somos capaces de apreciar, a través de la visión binocular, las diferentes distancias y volúmenes en el entorno que nos rodea. Nuestros ojos (debido a su separación) obtienen dos imágenes con pequeñas diferencias entre ellas, a lo que denominamos disparidad. Nuestro cerebro

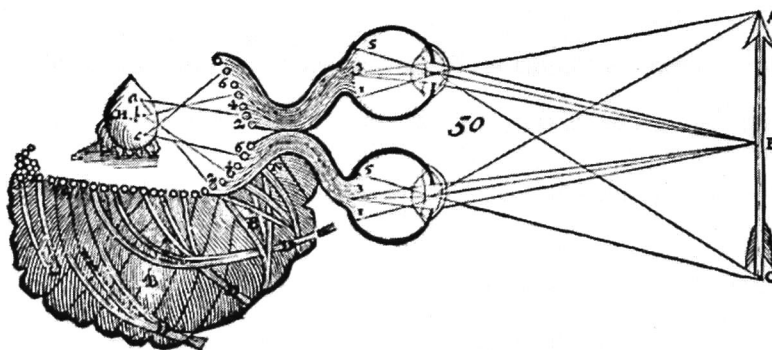


Figura 2.1: El sistema visual según lo imaginado por René Descartes (1596-1650). Las dos imágenes retinianas se enfocan sobre la glándula pineal, “El recinto de la imaginación y sentido común”

procesa las diferencias entre ambas imágenes y las interpreta de forma que percibimos la sensación de profundidad, lejanía o cercanía de los objetos que nos rodean. Este proceso se denomina estereopsis. La distancia interpupilar más habitual es de 65 mm, pero puede variar desde los 45 a los 75 mm. En la estereopsis intervienen diversos mecanismos. Cuando observamos objetos muy lejanos, los ejes ópticos de nuestros ojos son paralelos¹. Cuando observamos un objeto cercano, nuestros ojos giran para que los ejes ópticos estén alineados sobre él, es decir, convergen. A su vez se produce el enfoque para ver nítidamente el objeto. Este proceso conjunto se llama fusión. La agudeza estereoscópica es la capacidad de discernir, mediante la estereopsis, detalles situados en planos diferentes a una distancia mínima. Hay una distancia límite a partir de la cual no somos capaces de apreciar la separación de planos y varía de una persona a otra. Así, la distancia límite a la que dejamos de percibir la sensación estereoscópica puede variar desde 60 metros hasta cientos de metros. Un factor que interviene directamente en esta capacidad es la separación íterocular, a mayor separación entre los ojos, mayor es la distancia a la que apreciamos el efecto de relieve. El efecto obtenido con una separación íterocular mayor que la habitual es que los objetos parecen más pequeños de lo normal (*Liliputismo*), la técnica se denomina *hiperestereoscopia*. El efecto contrario se consigue con la *hipoestereoscopia*, es decir, con la reducción de la distancia íterocular, imprescindible para obtener imágenes estereoscópicas de pequeños objetos (*macrofotografías*). Estas nociones asociadas a los conceptos de la horóptera² reflejan la relación del modelo presentado con el funcionamiento del sistema visual de algunos seres vivos, en particular los primates, es por ello que la distancia interocular y demás conceptos se deben utilizar para obtener un modelo eficiente en términos computacionales e intuitivo debido a la experiencia que los humanos poseen con uno de sus

¹ Los ejes serán paralelos cuando el punto observado sea el punto en el infinito.

² La horóptera es una curva 3D que se define como el conjunto de los puntos para los cuales la luz cae en áreas correspondientes en las dos retinas.

sentidos más evolucionado.

2.2. Visión Estereoscópica

La profundidad de un punto en una escena no puede ser obtenido de una sola imagen, esto se debe a que la proyección de los puntos de la escena a la cámara produce la pérdida de información de profundidad, es decir, con una sola imagen sólo se puede determinar una línea entre el punto que está en la imagen y el centro de la cámara, dicho subespacio posee infinidad de puntos³, por ello es necesario al menos dos imágenes, para obtener la intersección de ambos rayos ópticos y así determinar la profundidad del punto en la escena⁴ (ver la figura 2.2). La visión estéreo o estereoscópica emplea dos cámaras

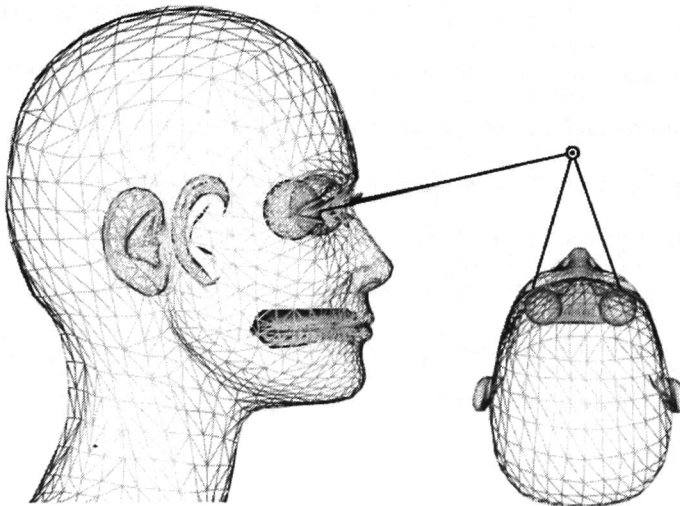


Figura 2.2: Los ejes ópticos del sistema de visión estereoscópica en el ser humano.

³ Todos los puntos contenidos la línea.

⁴ En las imágenes es necesario hacer consideraciones para minimizar los efectos de ruido al momento de la obtención de la profundidad mediante técnicas de retroproyección[14].

para capturar pares de imágenes de la misma escena y así poder inferir un conjunto de puntos en el espacio tridimensional. La forma en que dichas cámaras se describen, así como el espacio visual que puede ser reconstruido se describe en el capítulo 4. En la configuración estéreo la intersección de los campos visuales de ambas cámaras se le llama el campo de visión tridimensional, sólo los puntos contenidos en dicho subespacio podrán ser reconstruidos mediante técnicas de correlación, es decir, técnicas que resuelven el problema de la tridimensionalidad (profundidad), determinando que parejas de píxeles en ambas imágenes corresponden con un mismo punto de la escena (proceso de fusión realizado en la corteza visual). Haciendo uso de los datos obtenidos en la calibración intrínseca y extrínseca de las cámaras es posible generar un modelo tridimensional de la escena, el cual puede ser visualizado mediante técnicas de realidad virtual.

2.2.1. Correspondencias

La búsqueda de parejas de píxeles que correspondan con un mismo punto en la escena tridimensional implica técnicas de correlación, dicho proceso posee una complejidad computacional restrictiva $O(n^4)$, es por ello que algunos trabajos [15][10] han encontrado una serie de restricciones que reducen el espacio de búsqueda y por lo tanto aceleran la tarea, al evitar realizar cómputo innecesario. El primer grupo de restricciones son geométricas en las cámaras, esto se debe a la configuración epipolar[13](ver la figura 2.3). Al momento de adquirir la imagen, cada punto en la escena será mapeado en los planos de las imágenes como un par de rayos y al unir los centros de las cámaras mediante una línea se describe el plano epipolar (restricción epipolar). De esta manera los píxeles que representa dicho punto en la escena se encontrarán en la línea epipolar(producto de la intersección del plano epipolar con el plano de la imagen) en cada cámara. Esta restricción es esencial ya que reduce el espacio de búsqueda a una sola

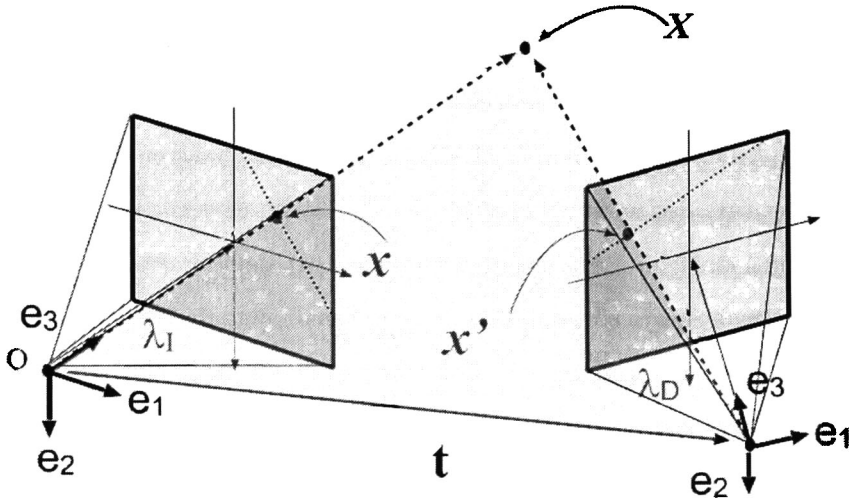


Figura 2.3: Configuración epipolar.

línea⁵ de la imagen para cada píxel de la imagen izquierda en la derecha y viceversa. La siguiente restricción consiste en el orden de las parejas de píxel. Dados dos puntos en la recta epipolar sus puntos correspondientes deben estar ubicados en la recta epipolar homóloga (en la otra imagen) en el mismo orden, esta restricción disminuye notablemente los cálculos, ya que mediante técnicas de programación dinámica que se acelera al aprovechar los cálculos para etapas posteriores de filtrado y discriminación de errores (ver la figura 2.11). La restricción de unicidad define que para cada punto de una imagen existe un único punto correspondiente en su imagen par, esto sucede si dicho punto no se encuentra en una oclusión, en cuyo caso el punto en una imagen no podrá ser encontrado en su imagen par (ver la figura 2.4-a).

El frente de onda o superficie de iso-disparidad (mapa de disparidad)[20] es una superficie continua, es decir, los puntos proyectados sobre la imagen pertenecen a las

⁵ No necesariamente corresponde con las líneas de muestreo.

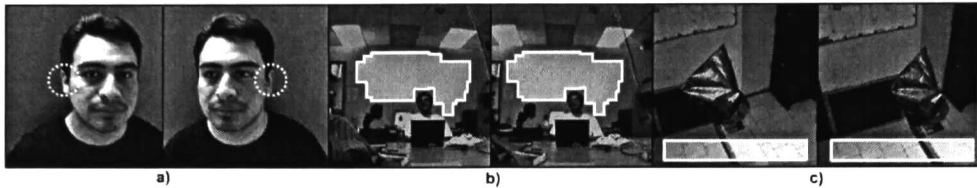


Figura 2.4: a) Ejemplo de oclusión, el sujeto se encuentra frente al sistema a una distancia reducida, la información (regiones de los oídos) de la imagen izquierda no puede ser validada en la derecha y viceversa. b) Región con textura uniforme, los métodos locales son susceptibles a generar correlaciones erróneas debido a la baja varianza en el vecindario. c) Regiones con contenido muy similar separados por segmentos uniformes, son difíciles de correlacionar en forma adecuada sin el uso de restricciones.

superficies de los objetos de la escena que se asumen continuas, excepto en la separación entre los objetos, esta restricción es importante para determinar un máximo de disparidad a computar. La disparidad máxima implica no sólo el grado de discontinuidad entre los objetos, sino también el plano de corte de cercanía (ver la figura 2.5). El límite del gradiente de disparidad es utilizado como restricción que impone un límite en las tangentes de las superficies de los objetos en la escena. El valor angular de dichas tangentes deben ser lo suficientemente pequeñas para que los puntos de la superficie yazcan en un rango factible. Existen varias técnicas de correspondencia con diferentes enfoques, entre las más utilizadas se encuentran:

- Técnicas globales,
- Técnicas locales,
- Técnicas basadas en área,
- Técnicas basadas en características,
- Técnicas jerárquicas,
- Técnicas de programación dinámica,

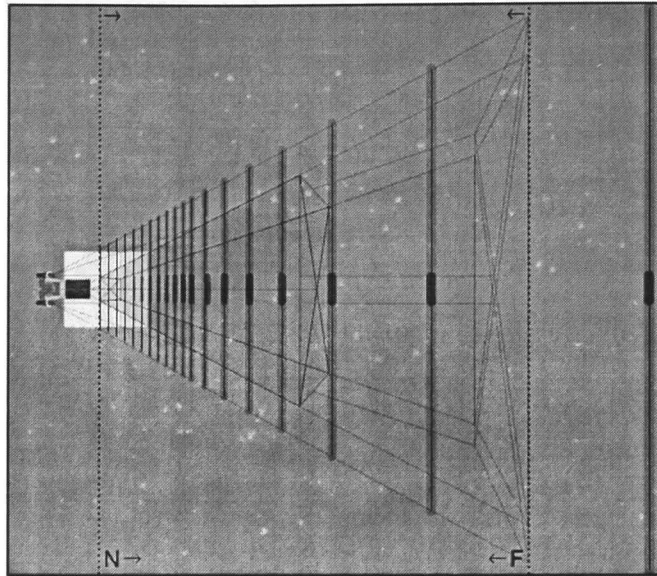


Figura 2.5: Líneas de disparidad, el rango (subespacio) contenido entre cada incremento discreto de disparidad crece en forma cuadrática.

Técnicas basadas en más de una vista.

Los métodos globales se enfocan en determinar la mejor correlación basándose en toda la información de ambas imágenes, esta aproximación es la más fácil de implementar, pero es poco usada debido al costo computacional, dichos métodos son menos sensitivos a las peculiaridades locales puesto que añaden soporte a las regiones de difícil estudio de manera local. Los métodos locales aplican restricciones a un pequeño número de puntos alrededor (vecindario o región de soporte ver las figuras 2.6-2.7) del punto de estudio, son muy eficientes pero sensitivos a las ambigüedades locales de las regiones, (por ejemplo oclusiones, regiones con texturas uniformes o regiones muy similares separadas por variaciones grandes, ver la figura 2.4). En estos métodos se asume que la zona alrededor del píxel de interés no deberá variar en forma radical entre cada par de imágenes debido a las restricciones anteriores. Las regiones de interés pueden ser

definidas como ventanas rectangulares como en la figura 2.6-a o bien como la figura 2.6-b, donde la aportación de los elementos se ponderar por el valor de la región. Los métodos presentados en el capítulo 3 utilizaron las funciones de región de la figura 2.6-a. Los métodos basado en área y en características, así como los basados en la opti-

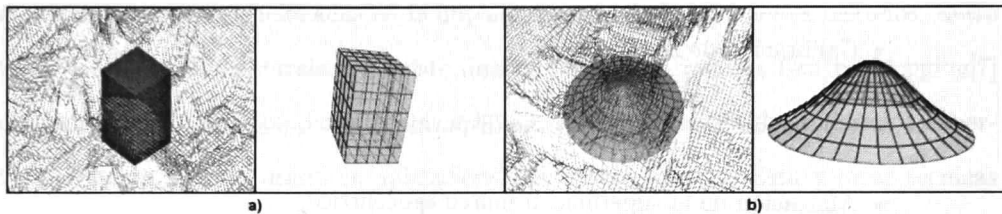


Figura 2.6: Máscaras de ponderación (aportación) para la región de interés o subventana. a) Función caja. b) Función gaussiana.

mización del gradiente son el siguiente escalón en su grado de efectividad ya que hacen uso de la información en el área, es decir, además de hacer uso de la región de interés también analizan las característica de dicho vecindario, si la región es muy continua o bien si la información cromática aporta algún índice de identificación, etcétera. Las técnicas que hacen uso de jerarquías en las peculiaridades de la imagen son excelentes dado su alto desempeño computacional, sin embargo no son usadas ampliamente debido a que la selección y jerarquización de las características puede ser difícil de determinar o ajustar a medios cambiantes. Los métodos de programación dinámica almacenan la información para etapas posteriores de filtrado y eliminación de falsas correlaciones, además evitan cálculos redundantes. Los métodos basados en más de una vista son más robustos en el sentido de la calidad de la reconstrucción. Sin embargo al requerir más imágenes para obtener aquellas correspondencias que no fueron conseguidas con el primer par de imágenes, implica más cálculos y almacenamiento, sin embargo son la mejor forma de verificar los datos obtenidos en etapas anteriores, logrando superficies de reconstrucción más continuas y coherentes con la realidad de la escena.

2.2.2. Reconstrucción del espacio Visual

El proceso de reconstrucción consiste en las siguientes etapas:

- Adquisición sincronizada de imágenes,
- Calibración de imágenes,
- Obtención de superficie de iso-disparidad,
- Alineación de la superficie al marco egocéntrico,
- Rectificación geométrica del espacio.

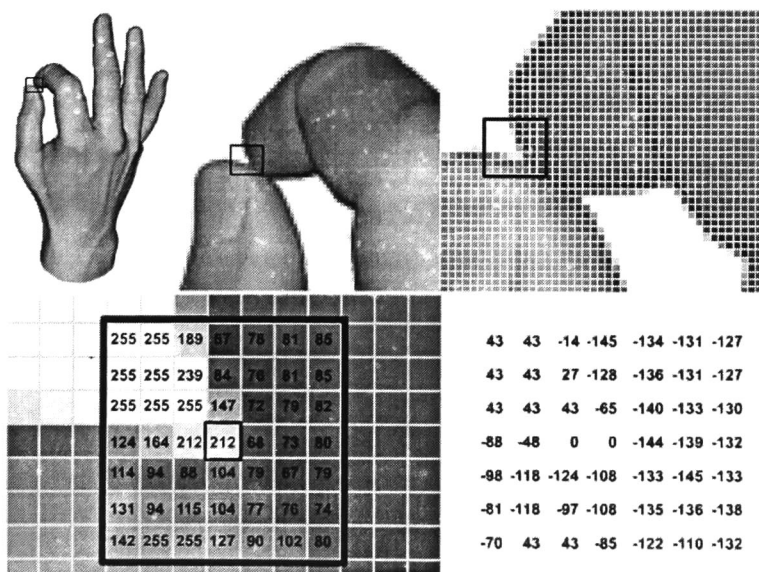


Figura 2.7: Región o vecindario del píxel central, así como la varianza resultado de aplicar el operador C_s .

La adquisición de imágenes se realiza en forma controlada para mantener una correspondencia temporal en ambas imágenes, esto es posible debido a las características de

las cámaras *Flea*⁶ posteriormente se usa la función de calibración previamente computada (ver la sección 3.2) para rectificar las imágenes. La obtención de la superficie de disparidad se realiza en multiresolución para conseguir un proceso más eficiente como se explica en la sección 4.3. La estrategia para lograr calcular la disparidad considera las restricciones antes mencionadas en la implementación de los siguientes métodos. Entre todas las técnicas de correlación existe una en particular llamada *Census Mapping*[7] que por sus ventajas computacionales y la calidad de los resultados obtenidos experimentalmente ha sido usada en aplicaciones de teleconferencia. Ésta y otras técnicas (como la suma del cuadrado de las diferencias de intensidades o el ángulo entre vectores generados mediante el operador *vec*⁷ del vecindario) inspiran la siguiente técnica.

El primer paso consiste en definir una subventana o vecindario de la forma

$$N(x, y, R) = [(x - R) \leq x_i \leq (x + R)] \times [(y - R) \leq y_i \leq (y + R)], \quad (2.1)$$

donde $R \in \mathbb{Z}$ define el radio de la subventana, dicha subventana se transforma en una nueva imagen mediante la siguiente función (ver la figura 2.7)

$$Cs(x, y, R, U) = \begin{cases} N(i, j, R) - N(x, y, R) & \text{si } |N(i, j, R) - N(x, y, R)| > U \\ 0 & \text{si } |N(i, j, R) - N(x, y, R)| \leq U \end{cases}$$

$$N'(x, y, R) = Cs(x, y, R, U)N(x, y, R), \quad (2.2)$$

la transformación realizada al vecindario mediante la función anterior proporciona dos nociones, la primera es generar una relación de todos los píxeles del vecindario con el píxel central, esto es similar al concepto usado en la técnica de correlación cruzada de intensidades [13], no es necesario calcular la media de la intensidad en la región ni la varianza de cada píxel con respecto a dicha media ya que al tomar el píxel central en conjunto con el umbral U se obtiene que aquellos píxeles con intensidades por debajo

⁶ <http://www.ptgrey.com/>

⁷ El operador *vec* crea un vector a partir del apilamiento de columnas de una matriz.

del umbral que no aportarán información en la firma de la región (heurística de rechazo al ruido). Una ventaja más de este paso es que sólo se considera la variación entre las intensidades de los píxeles, obteniendo un grado de insensibilidad sobre la exposición y ganancia de las cámaras⁸ El resultado anterior o firma del vecindario es organizado en forma de vector mediante el operador *vec*, es decir,

$$\Omega_{(x,y)} = \text{vec}(N'(x, y, R)), \quad (2.3)$$

el vector $\Omega_{(x,y)}$ tiene $(2R + 1)^2$ dimensiones (en la practica $2 \leq R \leq 7$), el cual será utilizado para computar la correlación con otros vectores producto del proceso anterior de la imagen derecha. Experimentos con varias funciones de correlación arrojaron resulta-

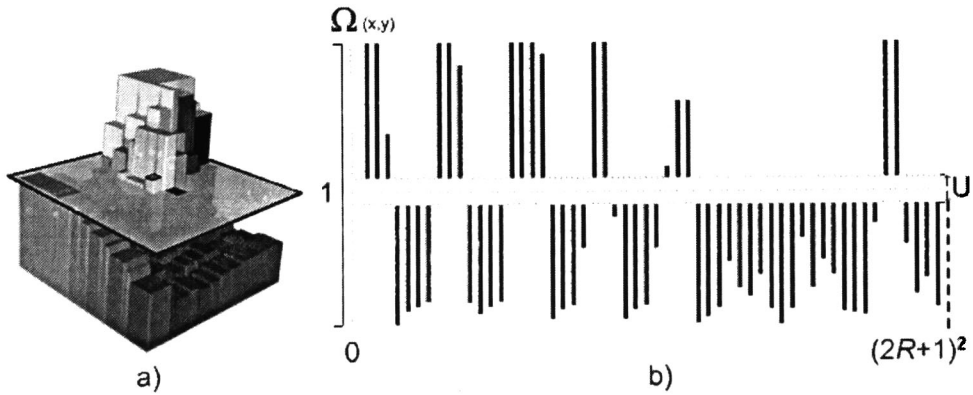


Figura 2.8: a) Representación 3D de la intensidad de los píxeles en el vecindario, el plano denota la intensidad del píxel central. b) La señal unidimensional (vector $\Omega_{(x,y)}$) resultado de aplicar el operador C_s .

dos satisfactorios, entre las cuales se destacan la correlación (\odot) unidimensional de los vectores $\Omega_{(x,y)}$ y $\Omega_{(x+\delta_x,y)}$, la integración de diferencias absolutas $G_{l,r}$ y una función de costo $h(i, j) : (i, j) \in (Z)^2 (\sigma_1, \sigma_2, \sigma_3) \in \mathbb{R}^3$ diseñada mediante criterios heurísticos basados en la similitud de la intensidad de los píxeles en los vecindarios (ver la figura

⁸ Es posible incluso conseguir un grado de inmunidad al cambio de contraste mediante la normalización del vector resultante.

2.9), es decir,

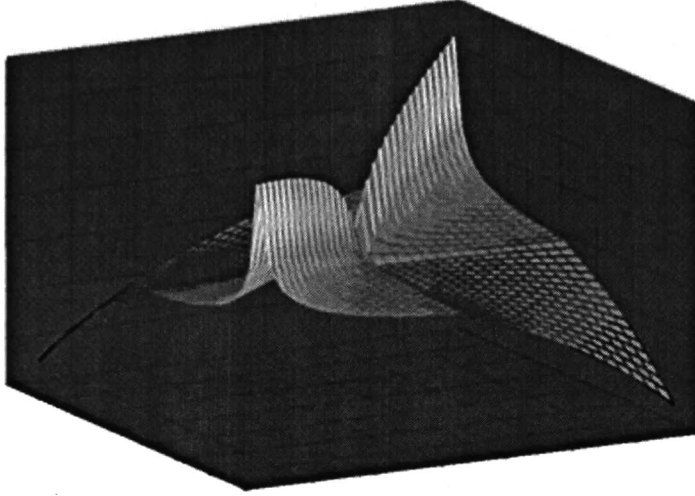


Figura 2.9: Función de costo $\tilde{h}(i, j)$, asocia un valor precomputado a la intensidad de dos píxeles para la correlación de regiones.

$$C_{l,r} = \Omega_{l(x,y)} \odot \Omega_{r(x+\delta_x,y)}, \quad -\delta_{max} \leq \delta_x \leq \delta_{max} \quad (2.4)$$

$$G_{l,r} = \sum_{\delta_x=-\delta_{max}}^{\delta_{max}} |\Omega_{l(x+\delta_x,y)} - \Omega_{r(x+\delta_x,y)}|, \quad (2.5)$$

$$\tilde{h}(i, j) = \begin{cases} \sigma_1(\frac{\pi}{4} - |\arctan(i, j) - \frac{\pi}{4}|) + \sigma_2(i^2 + j^2) & \text{si } i \cdot j > 0 \\ -\sigma_3 \cdot i \cdot j & \text{si } i \cdot j \leq 0 \end{cases} \quad (2.6)$$

la forma en que se aplica cada función es diferente entre si. Mientras que en el caso de $C_{l,r}$ se pretende encontrar la evaluación máxima, de esa manera se encontrará la mejor correlación. Por su parte $G_{l,r}$ pretende lo contrario, es decir, la evaluación mínima será la mejor correlación, ya que este criterio intenta localizar la mínima diferencia entre regiones. Finalmente la función $\tilde{h}(i, j)$ se debe aplicar sobre el vector $\Omega_{(x,y)}$, es

decir, se realiza una integración ponderada discreta de la forma

$$F_{l,r} = \sum_{\delta_x = -\delta_{max}}^{\delta_{max}} \sum_{i=0}^{(2R+1)^2} \hbar(\Omega_{(l,x+\delta_x,y)}(i), \Omega_{(r,x+\delta_x,y)}(i)), \quad (2.7)$$

tal que se localiza en la región donde $F_{l,r}$ sea el máximo, ahí se localiza una correspondencia. Se compararon los resultados contra métodos conocidos como la correlación cruzada de intensidades $CCI_{l,r}$ que se expresa como,

$$\bar{I} = \frac{1}{(2R+1)^2} \sum_{i=-R}^R \sum_{j=-R}^R I(x+i, y+j) \quad (2.8)$$

$$CCI_{l,r} = \sum_{i=-R}^R \sum_{j=-R}^R (I_R(x+i, y+j) - \bar{I}_R) \cdot (I_L(x+i, y+j) - \bar{I}_L), \quad (2.9)$$

la complejidad temporal se incrementa al calcular la intensidad media para cada región, además dicho método ha mostrado ser inapropiado cuando se presentan deformaciones por movimiento o disparidades amplias, lo cual limita la proximidad en que se puede correlacionar los puntos en la escena. Una mejora contundente en la disminución de la complejidad temporal y espacial, consiste en correlacionar sólo los píxeles de las imágenes que contengan suficiente textura o varianza en su región o vecindario, de esta manera las correlaciones serán más acertadas evitando incurrir en falsas correlaciones, que requieren cómputo y espacio para su detección, filtrado y eliminación. Se proponen varias maneras de calcular este criterio, por ejemplo, la varianza, la magnitud del gradiente, entre otros, es decir,

$$T_{var} = \sum_{i=-R}^R \sum_{j=-R}^R (I(x+i, y+j) - \bar{I}), \quad (2.10)$$

$$\vec{\delta}_\alpha(I(x, y)) = \sum_{i=-1}^1 \sum_{j=-1}^1 (I(x+i, y+j) D_\alpha(R+i+1, R+j+1)), \quad (2.11)$$

$$D_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \quad D_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}$$

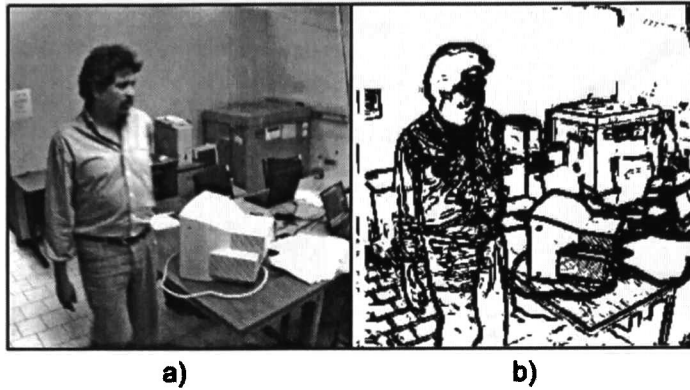


Figura 2.10: a) Imagen original. b) Imagen filtrada, los píxeles con una region de soporte con una varianza bajo del umbral U_G no son considerados en etapas posteriores.

maskas de aproximación a la derivación discreta[16]

$$D_x \rightarrow \frac{\delta I(x,y)}{\delta x} \text{ y } D_y \rightarrow \frac{\delta I(x,y)}{\delta y}$$

$$T_{Grad} = \sqrt{|\vec{\delta}_x(I(x,y))|^2 + |\vec{\delta}_y(I(x,y))|^2}, \quad (2.12)$$

con estos criterios es posible definir un umbral (incluso adaptativo con histéresis) para cada píxel en la imagen, sólo aquellos que estén por encima de dicha frontera serán procesados en las etapas siguientes. Se puede computar una imagen (ver la figura 2.10) de selección de píxeles de la forma

$$I_s(x,y) = \begin{cases} 1 & \text{si } T_{Grad}(I(x,y)) > U_G \\ 0 & \text{si } T_{Grad}(I(x,y)) \leq U_G. \end{cases} \quad (2.13)$$

Ahora con la imagen de selección de píxeles se procede al uso de las técnicas anteriormente descritas para obtener las disparidades, calculando primero a partir de la imagen izquierda a la derecha y viceversa. Una vez terminado el cálculo se procede a eliminar las falsas correspondencias mediante la restricción de unicidad, con tolerancia⁹ D_T píxeles

⁹ Entre 1 y 2 píxeles.

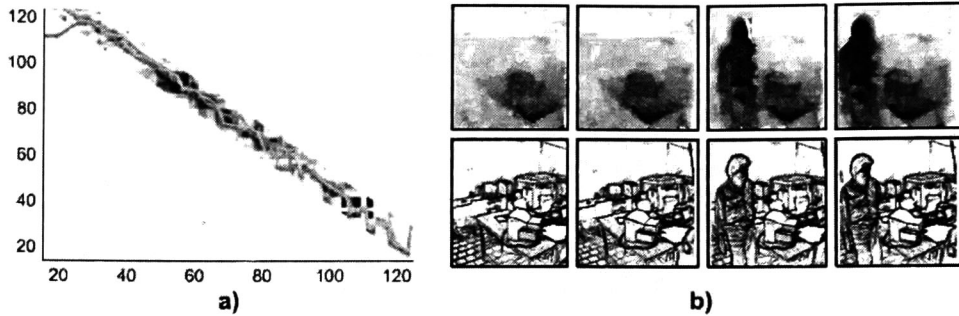


Figura 2.11: Vista en planta del sistema. a) Cálculo de función de correlación y determinación de línea epipolar de disparidad mediante esquema de programación dinámica b) Mapas de disparidad para dos escenas.

para contemplar los fenómenos de discretización, es decir, si $\Delta_{I \rightarrow R}(x, y)$ es el mapa de disparidad de la imagen izquierda hacia derecha ($\Delta_{R \rightarrow I}(x, y)$ es su contraparte),

$$|(\Delta_{I \rightarrow R}(x, y) + \Delta_{R \rightarrow I}((\Delta_{I \rightarrow R}(x, y) + x), y))| \leq D_T. \quad (2.14)$$

Aquellas disparidades que no cumplan con el anterior criterio se considerarán erróneas (*outliers*), por lo tanto se descartan haciendo $I_s(x, y) = 0$, generando un mapa de disparidad disperso¹⁰ en su contenido (ver la figura 2.11-b). Una forma de aproximar la solución es mediante interpolación entre píxeles activos en $I_s(x, y)$, haciendo uso de precisión subpíxel e interpolación con métodos bi-cúbicos¹¹ o aquellos que usan curvas de mayor complejidad¹². Es posible mejorar el mapa de disparidad¹³ aplicando filtros no lineales¹⁴ como la mediana selectiva, que es una aplicación del filtro de la mediana para aquellos píxeles que posean una varianza mayor a cierto umbral $U = \sigma_j \cdot \bar{I}$, tomando en consideración la cantidad de píxeles activos en la región o vecindario (ver la figura 2.12). Por último se ejecuta el paso de la alineación de la malla de iso-disparidad al ego-

¹⁰ Con huecos.

¹¹ <http://www.mathworks.com/access/helpdesk/help/toolbox/vipblks/f13689.html>

¹² <http://bigwww.epfl.ch/thevenaz/interpolation/>

¹³ En el sentido de la certidumbre de los datos.

¹⁴ A costa de perder detalle en la reconstrucción.

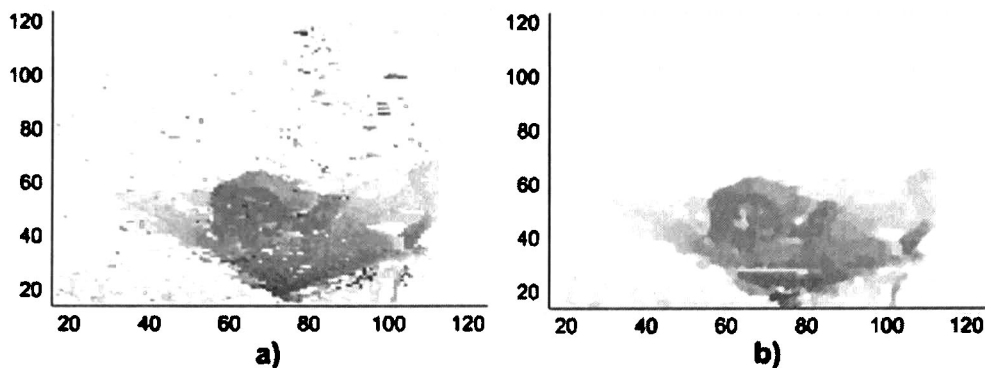


Figura 2.12: a) Mapa de disparidad con falsas correlaciones (*Outliers*). b) Aplicación del filtro de mediana selectiva para la eliminación de las falsas correlaciones.

mundo del sistema, para ello es necesario conocer la orientación del cabezal, al realizar una transformación rígida a los puntos obtenidos. La sección 4.2 explica las medidas tomadas para la rectificación geométrica de la escena, con lo cual se concluye con los pasos para lograr la reconstrucción del espacio visual.

2.3. Modelo conformal del Cabezal Binocular

Ahora se explica el desarrollo de un modelo conformal para representar la información 3D del espacio visual, comenzando por explicar la relación entre la horóptera y el modelo conformal. La horóptera es el lugar geométrico 3D en donde un objeto tiene que ser colocado para estimular exactamente dos puntos en correspondencia en las retinas izquierda y derecha de un sistema biológico de visión binocular. En la figura 2.13.b-c se observa la horóptera depender del ángulo azimut. En otras palabras la horóptera representa un conjunto de puntos que causan disparidad mínima (casi cero) en las retinas. La horóptera se traza mediante un círculo que pasa a través del punto de la fijación y los puntos nodales (L_C y L_R) de las dos retinas, consideran la figura 2.13-a. La horóptera

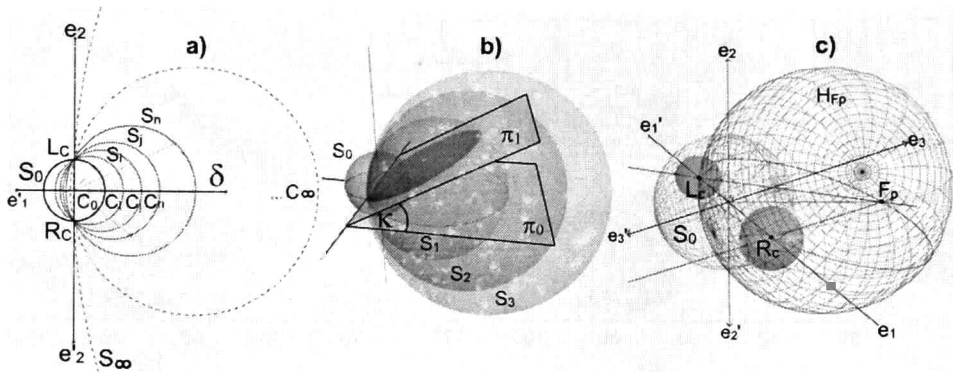


Figura 2.13: a) La horóptera y los centros de las esferas divididas por la línea bisectriz cuando los δ de la profundidad se incrementan. b) La horóptera esférica y la esfera unitaria, la horóptera dependiendo del ángulo azimut κ c) Configuración conformal de la horóptera, el centro de la cámara izquierda L_C , el centro de la cámara derecha R_C y el punto de fijación F_P definen un círculo, el cual mediante la variación del azimut κ define a su vez una familia de esferas o Horoptoesferas.

teórica se conoce como el círculo de *Vieth-Müller*. Observe que cada distancia de fijación tiene su propio círculo de *Vieth-Müller*. Según esta visión teórica, las suposiciones siguientes pueden ser hechas: cada retina describe un círculo perfecto, ambas retinas poseen el mismo tamaño, los puntos correspondientes se correlacionan perfectamente en sus localizaciones de la retina, los puntos en correspondencia se distribuyen uniformemente en el espacio (respecto a la nariz) y temporalmente en los ojos derechos e izquierdos. Si un objeto está situado en cualquier lado de la horóptera, una cantidad pequeña de disparidad es causada por los ojos. El cerebro analiza esta disparidad y computa la distancia relativa del objeto con respecto a la horóptera. En una región estrecha cerca a la horóptera la estereopsis no existe. Eso es debido a las disparidades que son muy pequeñas para estimular la estereopsis. Medidas empíricas de la horóptera (incluso hechas usando el método de *Nonius*¹⁵) no corresponden con el círculo de *Vieth-Müller*. Hay dos razones obvias de esta inconsistencia, las irregularidades en la distribución de

¹⁵ Es el método más fiel para la determinación de la horóptera experimental.

las direcciones visuales en los dos ojos o bien resultado de la distorsión óptica en la imagen en la retina. Existen varias razones fisiológicas por las que la horóptera puede ser deformada, una causa de la distorsión es la distribución asimétrica de las distribuciones visuales oculocéntricas, además de una asimetría regional en muestras locales en algún ojo, la distribución entre los dos ojos puede no ser congruente (problema de la correspondencia), esto puede ser otra causa de la distorsión de la horóptera. El mapeo asimétrico de la retina a la neocorteza en ambos ojos también causa una desviación del círculo de *Vieth-Müller*. La presencia de desviaciones en la horóptera en este sistema artificial dependerá si las cámaras fotográficas digitales satisfacen los requisitos arriba resumidos para obtener los círculos de *Vieth-Müller*. Una contribución importante de este trabajo es demostrar que la horóptera está embebida naturalmente en la esfera unitaria del álgebra geométrica conformal. La simple configuración de la horóptera

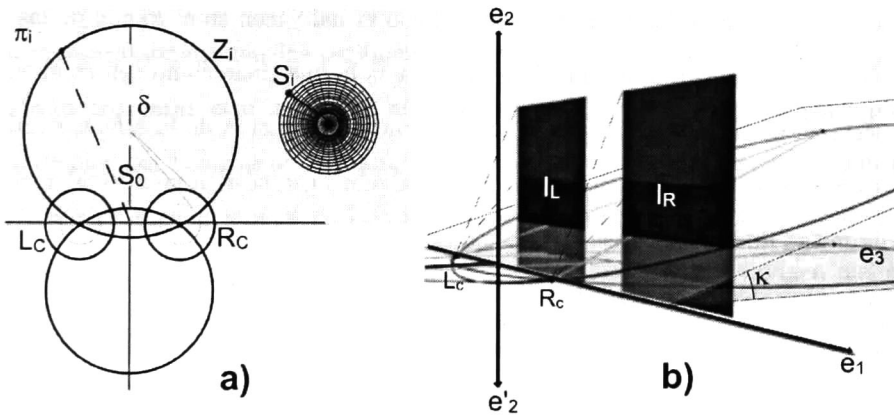


Figura 2.14: a) Representación polar del espacio geométrico visual. b) Planos de las imágenes y plano de azimut κ (configuración epipolar).

mostrada en la figura 2.14-a es una representación usando coordenadas polares del espacio geométrico visual. En contraste usando las herramientas de álgebra geométrica conformal podemos sugerir que la visión binocular se puede reformular con ventajas

usando una retina esférica. Ahora demostramos cómo encontramos la horóptera en la esfera de la geometría conformal. Nos estamos ocupando realmente de un *manejo*¹⁶ de esferas que intersectan los centros de las cámaras fotográficas L_C y R_C (ver la figura 2.14-b), esto es el manejo de esferas en el espacio proyectivo de esferas. Observe que los centros de las cámaras L_C y de R_C son puntos de *Poncelet*. Puesto que un sistema estéreo considera solamente lo que se encuentra delante de él, entonces consideramos las esferas que emergen hacia el frente. Cuando el lugar geométrico de los objetos se amplía, los centros de las esferas se desplazan a lo largo de la línea bisectriz del sistema estéreo, esto sucede cuando la profundidad crece (ver la figura 2.13-a). De aquí en adelante se utilizará la horopteresfera en lugar del círculo de la horóptera, porque cuando varía el azimut del círculo de la horóptera simplemente se está seleccionando diversos círculos de una horopteresfera en particular (ver la figura 2.13), consecuentemente se puede considerar que todos los puntos del espacio visual yacen en el *Pencil* de las horopteresfera. Llamamos la horopteresfera unitaria s_0 la que tiene como centro el egocentro del sistema estéreo (ver la figura 2.13-c), usando la ecuación de la esfera, donde su centro es el origen verdadero del espacio de la álgebra geométrica conformal y cuyo radio es la mitad de línea de base de sistema estéreo, $\rho = \frac{1}{2}|L_C - R_C|$. El centro c_i de cualquier horopteresfera s_i se desplaza hacia el punto en el infinito $c_i = c_i + \frac{1}{2}(c_i^2)e_\infty + e_0$, donde c_i es el vector euclidiano 3D. Así podemos escribir la ecuación de la esfera s_i como $s_i = c_i + \frac{1}{2}(c_i^2 - \rho_i^2)e_\infty$ donde el radio se calcula en los términos de la profundidad de la estereopsis. Se pueden obtener los mismos círculos z_i simplemente intersecando el *Pencil* de planos con las esferas de la tangente como se representa en la figura 2.13-a. Si se toma una inversión de todos los círculos de la horopteresfera con los centros en la línea bisectriz se consigue las líneas radiales del diagrama polar (ver la figura 2.14-a).

¹⁶ En inglés *Pencil*.

Capítulo 3

Procesamiento de Imágenes Estereoscópicas

3.1. Calibración de las Cámaras

Es necesario determinar los parámetros internos y externos de las cámaras, puesto que con dichos parámetros se podrá hacer uso de los puntos en las imágenes para la reconstrucción por técnicas de triangulación o retroproyección. Los parámetros internos de la cámara son conocidos como los parámetros intrínsecos y son básicamente la distancia focal, la resolución horizontal, vertical, el punto central de la imagen y el grado de sesgo (*skew*), mientras que los parámetros externos o extrínsecos se refieren a la posición y orientación de las cámaras respecto al origen universal [13]. En lo referente a la determinación de la distancia focal o enfoque de la cámara, es importante señalar que en la práctica los sistemas de enfoque poseen más de un lente, pero el resultado puede ser visto como un sólo lente, por facilidad y congruencia con el modelo geométrico *Pin-Hole* [13] con una sola distancia focal, aunque dicha situación pueda generar fenómenos como el que se trata en la sección 3.1.1. Por su naturaleza el sistema de lentes puede presentar aberraciones y distorsiones en las imágenes, dichas distorsiones

son modeladas y conocidas como la distorsión radial y tangencial. En la literatura existen métodos y algoritmos para su estimación mediante diversas técnicas, en particular se usó el método implementado en [24]. También existen métodos que determinan los parámetros intrínsecos y extrínsecos, pero debido a su esquema no es posible adaptarlos para realizar la calibración del enfoque en tiempo de ejecución¹. Dadas las hipótesis presentadas en el modelo mostrado en la sección 2.3, es necesario que ambas cámaras estén enfocadas simétricamente, de otra forma la horóptera se deformaría y las hipótesis no podrían ser sustentadas. Se requiere una calibración del enfoque adecuada con simetría en las distancias focales de ambas cámaras.

3.1.1. Calibración del Enfoque

El método de calibración inicia al capturar una imagen en escala de grises $I_l(x, y, t_0)$ en un tiempo t_0 con la cámara izquierda, después se transforma dicha imagen en su representación en el dominio de la frecuencia mediante la transformada rápida de Fourier² [16]

$$F_l(u, v, t_0) = \mathcal{F}\{I_l(x, y, t_0)\}. \quad (3.1)$$

Posteriormente haciendo uso de las propiedades del dominio de la frecuencia se procede a calcular la densidad espectral

$$P_l(u, v, t_0) = \frac{1}{2\pi} F_l(u, v, t_0) \overline{F_l(u, v, t_0)}, \quad (3.2)$$

la cual es usada para obtener la distribución de energía por número de onda. Después se define una región particular a la cual se le llamará la región de altas frecuencias, puede ser concebida como la zona que se extiende más allá de cierto radio de corte R_c ,

¹ Entender aquí tiempo de ejecución como la variación de la distancia focal y adquisición de imágenes en forma simultánea.

² <http://mathworld.wolfram.com/FastFourierTransform.html>

i.e., donde al menos alguna de las variables de frecuencia u o v (o las dos) representen las componentes de alta frecuencia de la imagen (ver la figura 3.1-a). La contribución

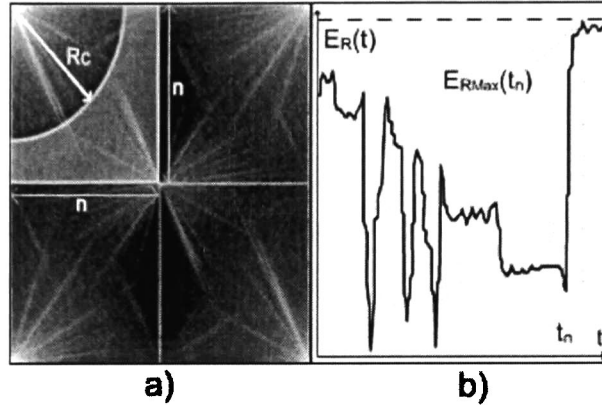


Figura 3.1: a) Región de alta frecuencia limitada mediante el radio de corte R_c y la mitad de la imagen n en su representación de densidad espectral. b) Máxima cantidad de energía $E_R(t)$ en la región de alta frecuencia en el tiempo t_n .

energética de cierta componente $P_i(u, v, t)$ en dicha zona se determina (se pondera para evitar ondulación³) mediante

$$B(i, j, R_c) = 1 - e^{-\sigma(i^2+j^2)R_c/2} \tag{3.3}$$

Esta zona implica todos los componentes de la imagen que constituyen los bordes y esquinas en el dominio del espacio, las cuales son las características a buscar en una cámara con enfoque apropiado; en este sentido es importante denotar que al momento de la calibración se establece un plano o profundidad determinada, esto se logra mediante la selección y aislamiento de la sección en la imagen usada para calibrar. La cantidad de energía en dicha región en un tiempo t se define como $E_R(t)$, se puede visualizar como un indicador del grado de agudeza (*sharpness*) de la imagen, se estima como una

³ *Rippling*

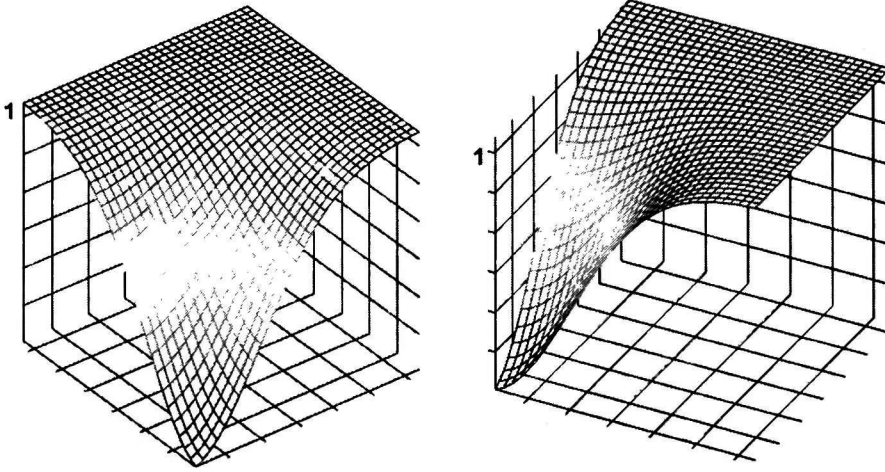


Figura 3.2: Función de ponderación ecuación (3.3).

integración discreta ponderada

$$E_R(t_0) = \sum_{i=1}^n \sum_{j=1}^n B(i, j, R_c) P(i, j, t_0). \quad (3.4)$$

Mientras dicho índice se calcula en forma iterativa, se ajusta físicamente la distancia focal en el lente de la cámara hasta que se logra un valor máximo $E_{R_{Max}}(t_n)$ (ver la figura 3.1-b). Con este método se logra la calibración del enfoque para una cámara, posteriormente se realiza el mismo método en la otra cámara, con ello se calibran ambas cámaras, sin embargo hasta este punto sólo se sabe que ambas cámaras están enfocadas, pero no se tiene la certeza de que ambas cámaras este simétricamente enfocadas. Esta situación es posible si cada cámara esta enfocada a diferente planos de profundidad (en tal caso se debe asegurar la calibración a una misma profundidad), por ello se presenta la siguiente estrategia. Se establece una subventana de la forma

$$W_L(x, y, R_s) = [(x - R_s) \leq x_i \leq (x + R_s)] \times [(y - R_s) \leq y_i \leq (y + R_s)] \quad R_s \in \mathbb{Z} \quad (3.5)$$

en la imagen obtenida mediante la cámara izquierda de tal modo que el segmento de imagen aparezca en el campo visual tridimensional del sistema estereoscopico, así el contenido será localizable en ambas imágenes. Posteriormente se usa dicha imagen como *Kernel* de correlación con la imagen de la cámara derecha, en la práctica se implementó mediante una convolución reflejando vertical y horizontalmente la subventana del lado izquierdo, también se implementó el mejoramiento del *Kernel* (imagen) mediante un filtro detector de bordes[16] con una máscara de 3x3 píxeles en el dominio del espacio,

$$F_R(u, v, t) = \mathcal{F}\{I_R(x, y, t)\}, \quad (3.6)$$

$$F_{Kernel}(u, v, t) = \mathcal{F}\{W_L(R_s - x, R_s - y, R_s)\}, \quad (3.7)$$

$$C_w(i, j) = F_{Kernel}(i, j, t) \cdot F_R(i, j, t), \quad (3.8)$$

$$M_p = \max(\operatorname{Re}(\mathcal{F}^{-1}\{C_w(i, j)\})). \quad (3.9)$$

El punto M_p con máxima amplitud en la escala de grises⁴ será el centro de la máxima correlación, el mejor emparejamiento (*matching*), donde la subventana $W_R(x, y, R_s)$ será establecida en la imagen derecha. De este modo se han determinado dos subventanas una en cada imagen, el contenido de ambas subventanas está en el campo visual tridimensional, ahora se procede a aislar el contenido de las imágenes de la forma (ver la figura 3.3)

$$I_{ins}(x, y, t) = \Phi(x, y, R_s)I(x, y, t), \quad (3.10)$$

$$\Phi(x, y, R_s) = \begin{cases} 1 & (x, y) \in W(x, y, R_s) \\ 0 & (x, y) \notin W(x, y, R_s), \end{cases}$$

haciendo uso del teorema de invarianza de corrimiento (*Shift Invariant*) el contenido energético de ambas señales sería muy similar. Finalmente se procede a llevar ambas

⁴ La función $\operatorname{Re}(x \in \mathbb{C}) \in \mathbb{R}$ selección la parte real de un elemento complejo.

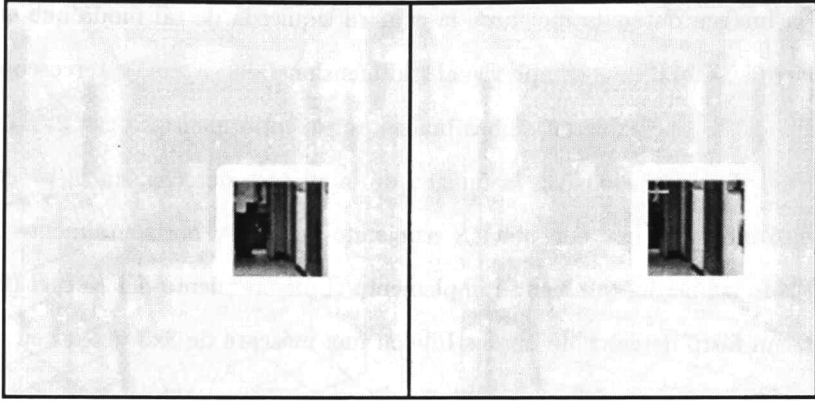


Figura 3.3: Imágenes izquierda y derecha aislando el contenido previamente acoplado mediante la función Φ .

imágenes⁵ a su representación en el dominio de la frecuencia, calcular su índice de energía en la región de alta frecuencia, es decir,

$$F_{Lins}(u, v, t) = \mathcal{F}\{I_{Lins}(x, y, t)\}, \quad (3.11)$$

$$F_{Rins}(u, v, t) = \mathcal{F}\{I_{Rins}(x, y, t)\}, \quad (3.12)$$

$$P_{Lins}(u, v, t) = \frac{1}{2\pi} F_{Lins}(u, v, t) \cdot \overline{F_{Lins}(u, v, t)}, \quad (3.13)$$

$$P_{Rins}(u, v, t) = \frac{1}{2\pi} F_{Rins}(u, v, t) \cdot \overline{F_{Rins}(u, v, t)}, \quad (3.14)$$

$$E_{RL}(t) = \sum_{i=1}^n \sum_{j=1}^n B(i, j, R) P_{Lins}(i, j, t), \quad (3.15)$$

$$E_{RR}(t) = \sum_{i=1}^n \sum_{j=1}^n B(i, j, R) P_{Rins}(i, j, t), \quad (3.16)$$

$$\Delta E_d(t) = E_{RL}(t) - E_{RR}(t). \quad (3.17)$$

El objetivo es ahora minimizar $\Delta E_d(t)$, ésto se logra físicamente mediante la variación de la distancia focal en la cámara cuyo índice $E_R(t)$ sea el menor. Dicho método proporciona resultados adecuados en particular para las etapas posteriores de reconstrucción y manipulación.

⁵ Con los contenidos previamente seleccionados, emparejados y aislados.

3.1.2. Calibración de Distorsión Radial y Tangencial

Las imágenes adquiridas mediante las cámaras pueden estar bien enfocadas, sin embargo la deformación producida por el sistema óptico, conformado por el conjunto de lentes, puede generar una distorsión, la cual puede ser modelada y a su vez compensada mediante diversos esquemas, con dichos modelos es posible precalcular la función inversa a la detectada en las imágenes. Dicha etapa de compensación se le conoce como *warp* (ver la figura 3.6). En la práctica las componentes de distorsión radial son más fuertes que las tangenciales y dado que la magnitud de la distorsión tangencial es insignificante[22], entonces se ignoran al calcular la función de compensación.

3.1.3. Calibración Geométrica

En lo referente a la calibración geométrica se hace uso de la técnica bien conocida en la comunidad de visión artificial publicada por Zhang Z. [26]. La cual no sólo proporciona los parámetros intrínsecos, sino también los extrínsecos basados en la implementación de Jean-Yves Bouguet⁶

3.2. Rectificación de Imágenes

Las etapas de calibración se aplican mediante la función de calibración $\chi(i, j) \in \mathbb{Z}^2$, la cual es un arreglo bidimensional de enteros que contiene las coordenadas donde se deberá mapear los píxeles $P(x, y) \in \mathbb{Z}^2$ de la imagen original para producir la imagen calibrada (ver la figura 3.6), es decir,

$$I_{CL}((x, y), t_n) = I_{UL}(\chi(x, y)^{-1}, t_n) \quad \forall x, y \in I_{CL}. \quad (3.18)$$

⁶ <http://www.vision.caltech.edu/bouguetj/>

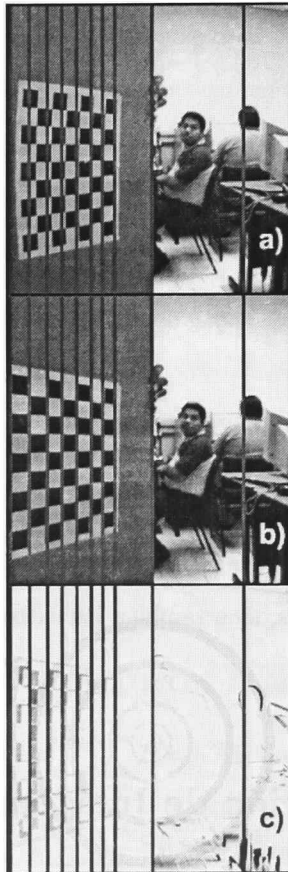


Figura 3.4: a) Imagen original con distorsión radial. b) Imagen rectificada. c) Error entre la imagen original y la rectificada, note que el error crece a razón de un polinomio de segundo grado con argumento radial a partir del píxel central (no necesariamente el centro de la imagen).

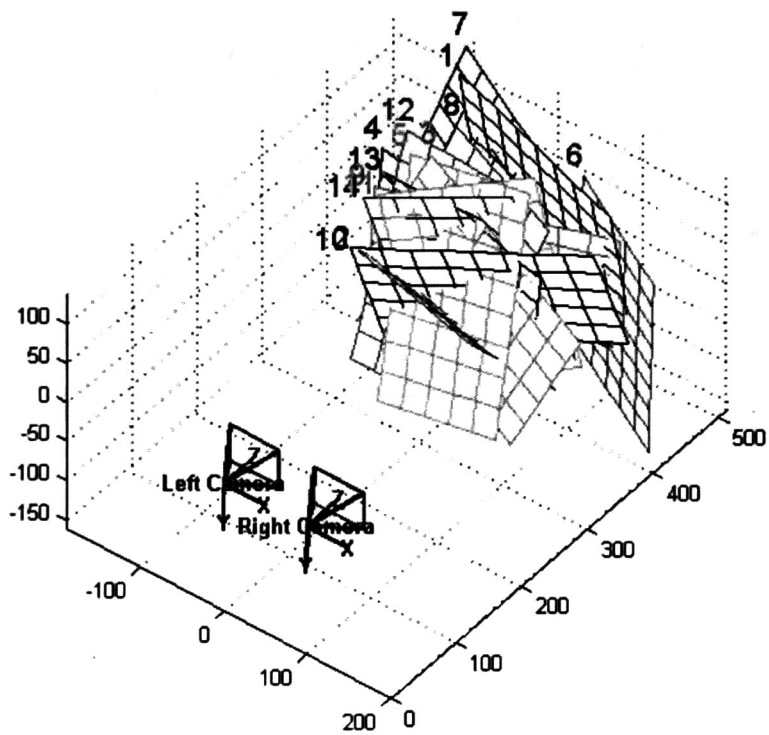


Figura 3.5: Resultado de la calibración mediante la implementación Jean-Yves Bouguet.

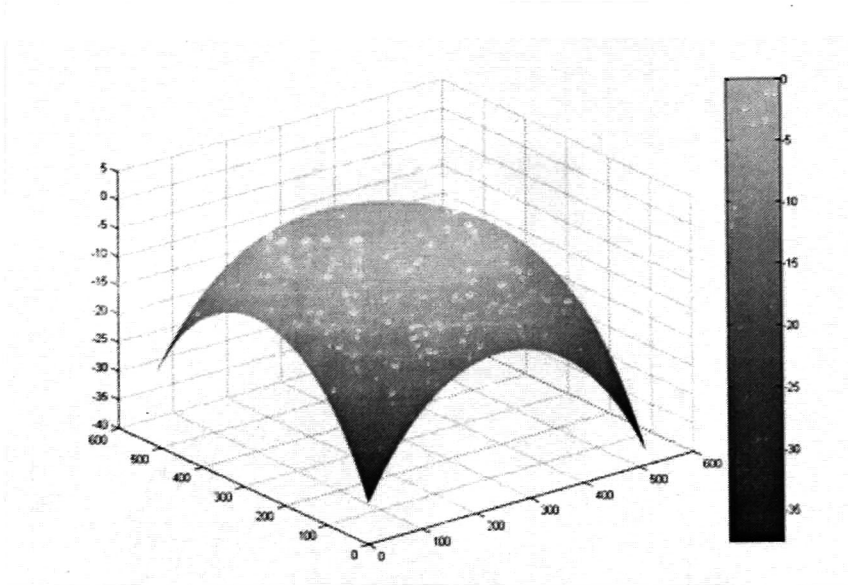


Figura 3.6: Función de calibración de imagen (*Warp*).

Capítulo 4

Reconstrucción Activa del Espacio Visual

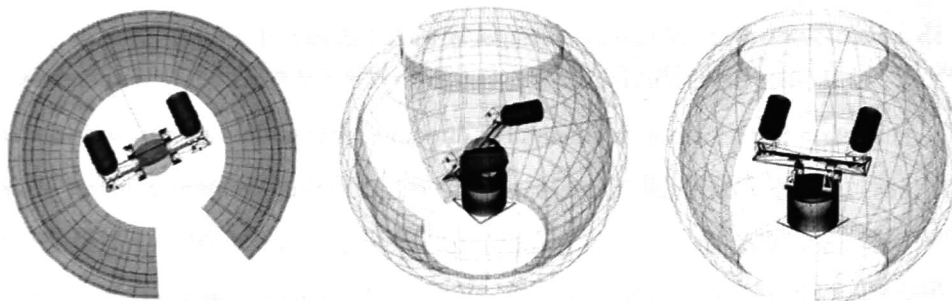


Figura 4.1: Ego-mundo del sistema.

El sistema físico de visión artificial usado en el presente trabajo es parte de un proyecto llamado *PAC*¹ el cual incluye temas relacionados a la investigación sobre algoritmos de percepción y acción, aplicables en áreas tanto industriales como en el ámbito de los humanoides, es por ello que el subsistema estereoscópico se encuentra montado en otro subsistema robótico de dos grados de libertad llamado *Pan-Tilt-Unit* (se le llamará cabezal por su analogía con el ser humano), dicha unidad permite

¹ Perception-Action-Cycle proyecto de investigación del Prof. Dr. Eduardo Bayro Corrochano Cinvestav Unidad Guadalajara 2006 (ver la figura 6.1).

orientar el sistema con un ángulo de rotación horizontal (*Pan* o *PTU*) de $\pm 160^\circ$ (con una precisión de 1.8 arc-min, 33.333 pasos por grado), así mismo posee un eje de rotación vertical (*Tilt*) con $\pm 60^\circ$ (con una precisión de 1.8 arc-min, 33.333 pasos por grado, ver la figuras 4.3 y 4.1). El sistema cuenta con la capacidad de ser controlado por medio del programa desarrollado bajo la presente tesis con diversos perfiles, incluyendo posición, velocidad, aceleración y tirón (*Jerk*). Debido a los objetivos de la presente tesis se hace uso del sistema mediante el perfil de control mediante posición con un algoritmo de lazo cerrado para la verificación de la posición deseada.

4.1. Cinemática Directa e Inversa del Cabezal

El modelo de visión se utiliza en conjunto con el sistema *PTU*, entonces es necesario analizar e implementar la cinemática directa e inversa del cabezal, con la finalidad de posicionar y determinar la orientación del espacio visual comprendido por las imágenes, generando así una vista del entorno del sistema en una forma egocéntrica [23]; considerando el origen de los sucesos alrededor de dicho sistema, este tipo de concepto llamado ego-visión ha sido abordado con variedad de fines tanto en robótica como en áreas de entretenimiento y realidad virtual [23]. La ventaja de obtener un modelo de la cinemática directa e inversa significa que se asociarán posiciones angulares de los ejes de rotación del robot con los parámetros de las reconstrucciones tridimensionales, con el fin de controlar numéricamente el espacio contenido en el ego-mundo del robot (con un grado de precisión adecuado). Ésto puede ayudar en muchas operaciones, como ejemplo la segmentación en profundidad o determinación de movimiento en tres dimensiones, entre otras. En lo referente a la cinemática directa se puede determinar fácilmente mediante el análisis del modelo geométrico del sistema (ver la figura 4.3).

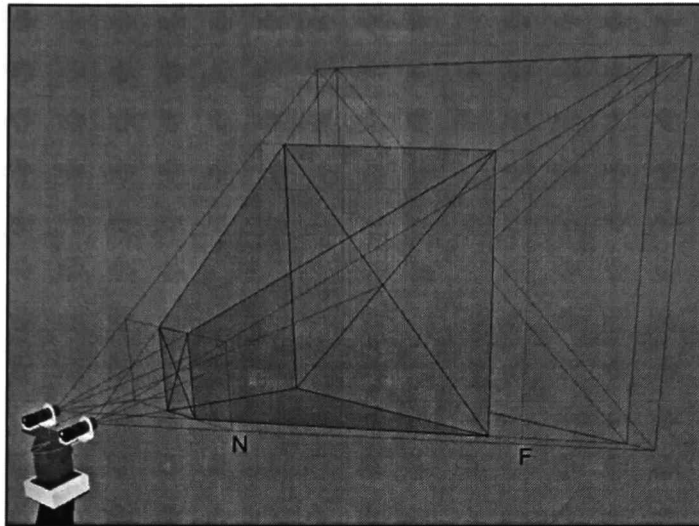


Figura 4.2: El campo de visión 3D es el espacio contenido en la intersección de los campos visuales de las dos cámaras, el plano de corte en cercanía N y el plano de corte en lejanía F .

La línea de base se expresa como $B_l = L_c \wedge R_c \wedge e_\infty$, la cual define la orientación entre los centros de las cámaras y la distancia ínterocular o distancia de línea de base que es de la forma $dist = \sqrt{-\frac{1}{2}(L_c \cdot R_c)}$. El eje de rotación *Pan* E_p se determina mediante el producto exterior de dos puntos marcados en el modelo tridimensional del sistema de la forma $E_p = \frac{X_a \wedge X_b \wedge e_\infty}{|X_a \wedge X_b \wedge e_\infty|}$ (una línea en el espacio Euclidiano). Similarmente, el eje de rotación *Tilt* E_t se construye de forma análoga con los puntos señalados en el modelo 3D del sistema como sigue, $E_t = \frac{Y_a \wedge Y_b \wedge e_\infty}{|Y_a \wedge Y_b \wedge e_\infty|}$, ambas líneas están normalizadas para ser usadas con la definición de los rotores (ver el apéndice A.8). Los centros de las cámaras así como su posición, orientación y su apertura (horizontal y vertical) determinan el campo de visión tridimensional (ver la figura 4.2). El punto central entre los centros de las cámaras C_c , proyectándolo en la dirección perpendicular a la línea de base, conforma el llamado eje óptico de la cámara tridimensional que se obtiene mediante los algorit-

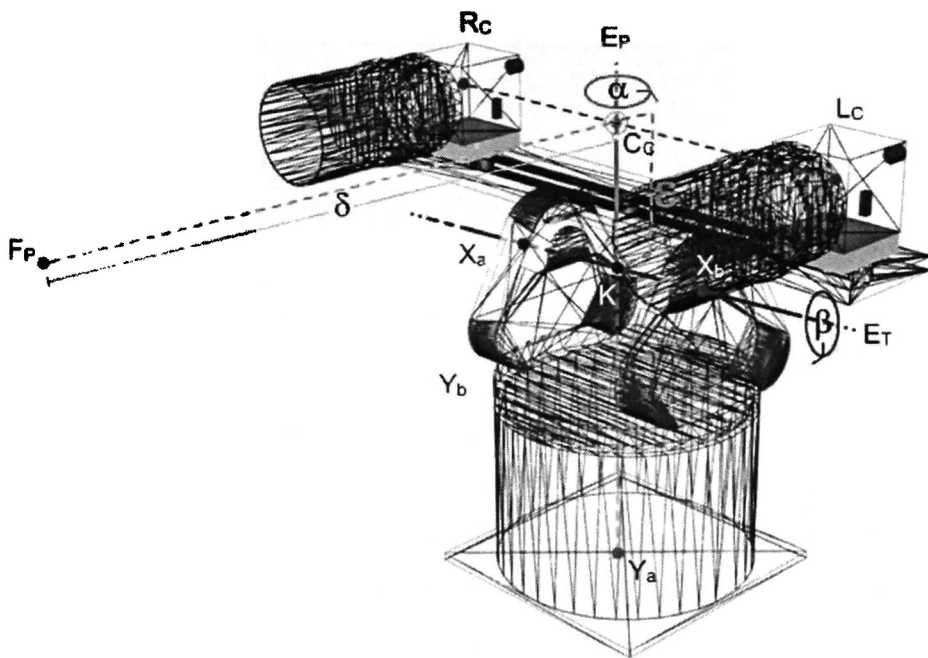


Figura 4.3: Modelo 3D del sistema de visión estereoscópica, puntos importantes, ejes de rotación *Pan-Tilt*.

mos de reconstrucción. La distancia entre el punto de fijación F_p y el punto central se determina como $\delta = \sqrt{-\frac{1}{2}(F_p \cdot C_c)}$. Finalmente se denotan los ángulos α de rotación del eje Pan y el ángulo β de rotación del eje $Tilt$. El eje de rotación Pan esta definido por la línea E_p , el ángulo de rotación para dicho eje es α con estos parámetros es posible definir el rotor

$$R_P = e^{-E_p(\frac{\alpha}{2})}, \quad (4.1)$$

similarmente el rotor en el eje $Tilt$

$$R_T = e^{-E_T(\frac{\beta}{2})}. \quad (4.2)$$

Se puede determinar la posición del punto de fijación F_p mediante una translación de punto C_C (con la magnitud de la profundidad deseada δ), es decir,

$$T_\delta = 1 - \frac{1}{2}\delta e_3 e_\infty. \quad (4.3)$$

La cinemática directa del cabezal se puede expresar como la translación de ese punto a una profundidad, posteriormente un giro con el eje $Tilt$ y por último un giro sobre el eje Pan (cadena cinemática), es decir,

$$F'_p = \{ R_P [R_T (T_\delta F_p \tilde{T}_\delta) \tilde{R}_T] \tilde{R}_P \}. \quad (4.4)$$

La cinemática inversa pretende obtener los ángulos de rotación α y β conociendo la posición del punto de interés F_p , esto es con el fin de mantener separados los parámetros internos del funcionamiento de los subsistemas robóticos (ver la figura 4.4). Las aplicaciones de interacción humano-máquina, inspección y manipulación basadas en visión podrán concentrar el interés en las posiciones absolutas de los objetos o entidades con respecto al ego-centro y así separar las capas de abstracción para los diferentes niveles conceptuales. La esfera S_1 , cuyo centro es el punto de interés F_p , es tal que el punto

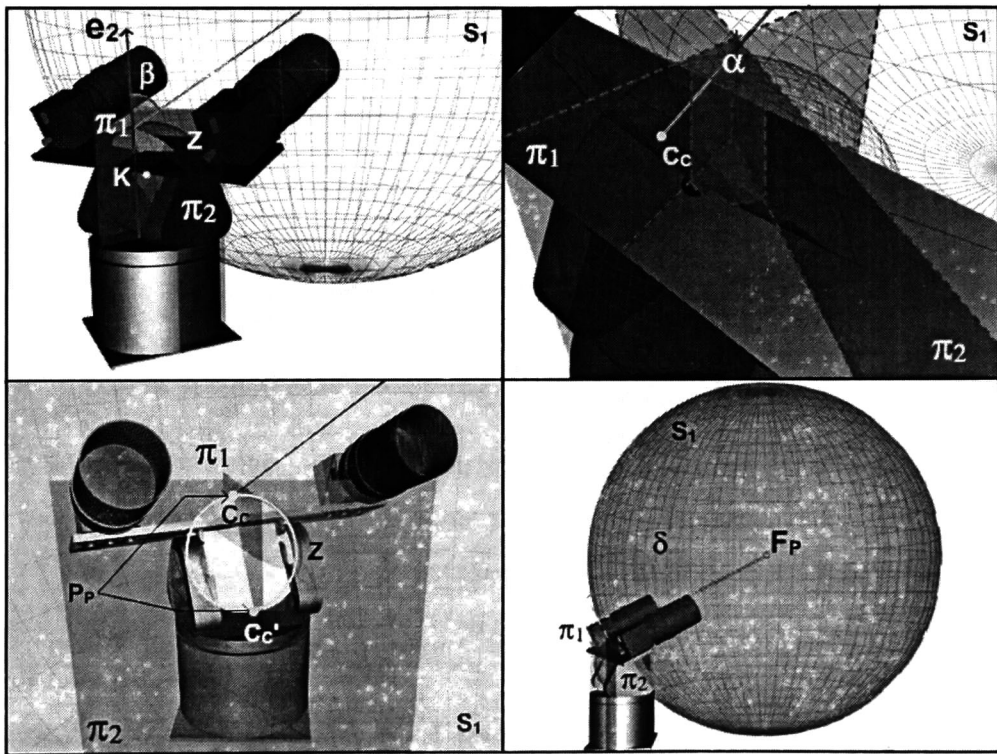


Figura 4.4: Algoritmo de cinemática inversa.

central C_C se encuentra en su superficie, es decir,

$$S_1 = F_P - \frac{1}{2}\delta^2 e_\infty, \quad (4.5)$$

de forma similar la esfera S_2 , cuyo centro K es la intersección ambos ejes de rotación (ver la figura 4.3); pasa por el punto central C_C , es decir,

$$S_2 = K - \frac{1}{2}\epsilon^2 e_\infty. \quad (4.6)$$

El círculo producto de la intersección de ambas esferas,

$$Z = S_1 \wedge S_2, \quad (4.7)$$

en seguida se obtiene el plano π_1 , mediante la línea E_P y el punto F_P , es decir,

$$\pi_1 = E_P \wedge F_P, \quad (4.8)$$

la intersección del círculo Z con el plano π_1 generar un par de puntos P_P ,

$$P_{P1-2} = Z \wedge \pi_1, \quad (4.9)$$

es posible obtener los puntos P_{P1} y P_{P2} por separado mediante la ecuación para el par de puntos (ver [9]). Con estos dos puntos se obtienen las líneas L_{C1} y L_{C2} , es decir,

$$L_{C1} = P_{P1} \wedge K \wedge e_\infty, \quad (4.10)$$

$$L_{C2} = P_{P2} \wedge K \wedge e_\infty, \quad (4.11)$$

las solución matemática para el ángulo β del eje *Tilt* se expresa como,

$$\beta_1 = \arccos \left(\frac{E_P \cdot (L_{C1})}{|E_P| |L_{C1}|} \right) \quad (4.12)$$

$$\beta_2 = \arccos \left(\frac{E_P \cdot (L_{C2})}{|E_P| |L_{C2}|} \right) \quad (4.13)$$

Estas ecuaciones proporcionan dos valores para β (dada la solución dual), sin embargo, se sabe que físicamente el sistema sólo podrá alcanzar uno de los valores angulares por las características del *PTU*. Ahora en lo referente a la solución del ángulo α del eje *Pan* se puede obtener mediante la siguientes afirmaciones. Primero se obtiene el plano $\pi_2 = e_3$ cuya normal es el eje e_3 y pasa por el origen del sistema (ver la figura 4.4), después mediante la intersección de los planos π_1 y π_2 se puede obtener α , es decir,

$$\alpha = \arccos \left(\frac{\pi_1 \cdot \pi_2}{|\pi_1||\pi_2|} \right), \quad (4.14)$$

de esta manera se puede obtener la cinemática inversa del *PTU*.

4.2. Detección de Primitivas Geométricas

La información contenida en la reconstrucción tridimensional del espacio visual es abundante y en muchas ocasiones redundante. Ésto se debe a que en muchos de los medios² en los que la visión artificial se desea aplicar son estructurados o semiestructurados, es decir, poseen características geométricas muy tajantes de tal manera que dichas características pueden ser aprovechadas para hacer las tareas en forma eficiente, como ejemplo vea la figura 4.5, en dichas imágenes la superficie plana del escritorio está almacenada, procesada y concebida en el ego-mundo del sistema como una gran cantidad de puntos en un superficie de iso-disparidad que además de ser redundante es menos susceptible a ser manejada mediante el marco matemático del álgebra geométrica, por ello se sugiere el beneficio de la detección de primitivas geométricas en el espacio visual. Existen una gran variedad de entidades geométricas, las cuales se pueden clasificar como primitivas, sin embargo en el álgebra geométrica son de gran

² Medios circundantes para los sistemas de visión, como interiores de edificios, industrias, vehículos, etcétera.

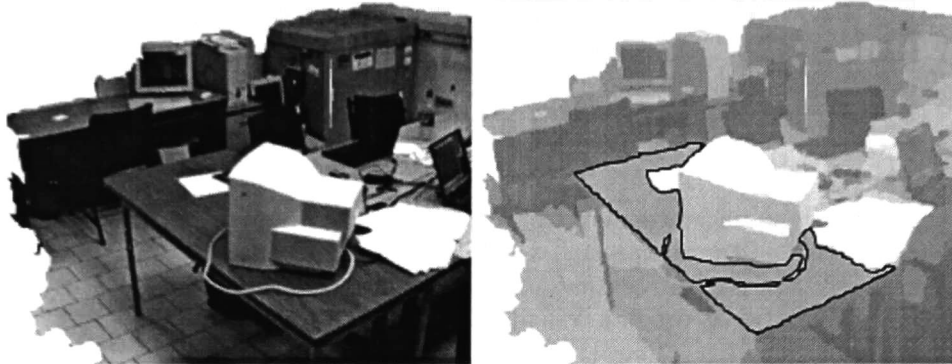


Figura 4.5: Ejemplo de primitiva geométrica en un ambiente semiestructurado.

utilidad los puntos, líneas, círculos, planos y esferas. La detección de puntos y líneas es susceptible a realizarse en el contenido de las imágenes mediante técnicas conocidas como la transformada de *Hough*[6] entre otras, sin embargo, las primitivas como planos y esferas implican un trabajo más amplio por su naturaleza, en este sentido existen varios métodos[17] para el ajuste de planos u otras entidades (superficies). El problema de detección de primitivas en la superficie generada (ver la figuras 4.6 y 4.7) se puede dividir en dos subproblemas, el primero consiste en detectar que región de la superficie posee las características topológicas y cromáticas adecuadas. Manteniendo un criterio o umbral para segmentar dichos grupos es posible generar una representación facetada³ del espacio visual, que es más adecuada para cierta tareas, en particular cuando se hacen algoritmos geométricos para controlar en algún modo el espacio visual. El primer paso es la segmentación en color para agrupar en bloques con contenidos cromáticos similares, esta etapa hace sólo análisis cromático, sin embargo con la etapa posterior se

³ Representación de un politopo n-dimensional, mediante (n-1) simplexes o caras

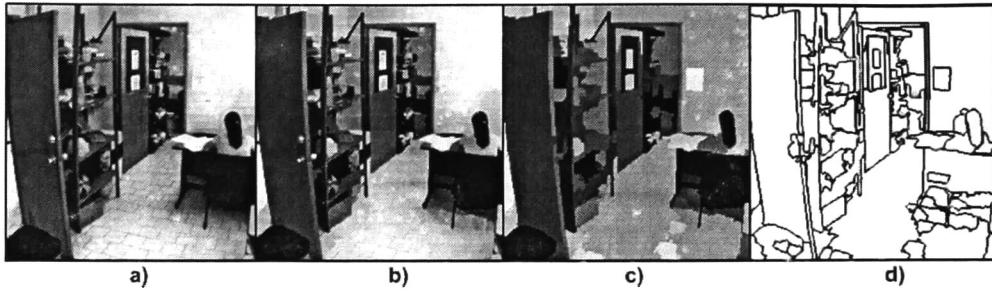


Figura 4.6: Segmentación en color de la imagen para la obtención de las regiones con características cromáticas similares. a) Imagen original. b) Imagen con reducción del espacio cromático. c) Aglutinamiento mediante el uso de la variación de la media. d) Fronteras entre las regiones cromáticas.

podrán detectar bordes o discontinuidades abruptas sobre la superficie de iso-disparidad que implican cambios topológicos substanciales. Existen trabajos relacionados con la segmentación en color, en particular el grupo EDISON (*Edge Detection and Image Segmentation*)[5][19][1] ha publicado una serie de técnicas de segmentación basadas en la variación de la media y la conglomeración mediante técnicas de *clustering*. El proceso consiste en la partición y reducción de resolución del espacio cromático⁴, después se procede a técnicas de suavizado para eliminar las desviaciones de gran magnitud de la media local, posteriormente el aglutinamiento en regiones (*patches*) donde la variación con respecto a la media local es menor a cierto umbral definido como parámetro de la segmentación. Finalmente se detectan los bordes entre las diversas regiones, se etiquetan para obtener un mapa de regiones D_i con características cromáticas similares (ver la figura 4.6). Posteriormente se localizan regiones con características topológicas similares, en particular en el caso de los planos se procede a seleccionar elementos del mapa de regiones y se realiza el método siguiente para la detección de los parámetros de la superficie plana. Para cada región i se tiene una lista de puntos conformales nor-

⁴ En diversos modelos de color según se requiera (RGB, HSB, LUV, etcétera).

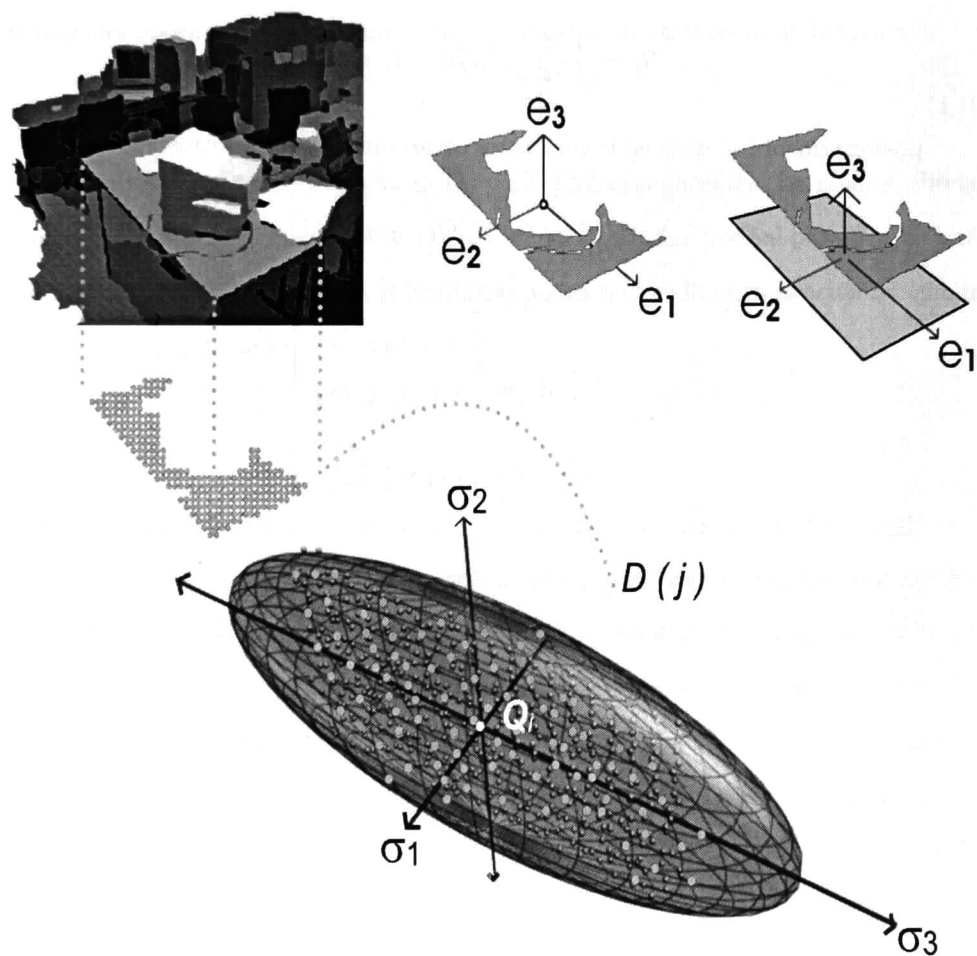


Figura 4.7: a) Imagen segmentada para rectificación geométrica. b) Región seleccionada con su centroide. c) Ajuste de plano basado en los vectores propios de la matriz de covarianza.

malizados⁵ $D_i(j)$ tales que j denota el índice de un punto en particular, el tamaño de la lista se denota como m_i . Ahora se calcula el centroide Q_i de los puntos en su representación euclidiana, es decir,

$$H_i = \frac{1}{m_i} \left[\sum_{j=1}^{m_i} D_i(j) \right], \quad Q_i = \sum_{k=1}^3 (H_i \cdot e_k) e_k, \quad (4.15)$$

posteriormente se obtiene la matriz de covarianza W_i de la forma siguiente

$$\psi_{p,q} = \sum_{j=1}^{m_i} \{ (D_i(j) \cdot e_p) - (Q_i \cdot e_p) \} \{ (D_i(j) \cdot e_q) - (Q_i \cdot e_q) \} \quad (4.16)$$

$$W_i = \begin{bmatrix} \psi_{1,1} & \psi_{1,2} & \psi_{1,3} \\ \psi_{2,1} & \psi_{2,2} & \psi_{2,3} \\ \psi_{3,1} & \psi_{3,2} & \psi_{3,3} \end{bmatrix}$$

dicha matriz W_i tiene valores propios (σ_1 , σ_2 y σ_3) y vectores propios (ν_1 , ν_2 y ν_3), los cuáles determinan los ejes del elipsoide que concuerda con la distribución espacial de los puntos (ver la figura 4.7), dicha noción puede ser utilizada para determinar la orientación de un plano, dado que el vector propio con magnitud menor ν_1 será la normal del plano donde los puntos $D_i(j)$ habitan. En términos del álgebra geométrica conformal se puede expresar como

$$\pi_i = \nu_1 + |Q_i \cdot \nu_1| e_\infty, \quad (4.17)$$

dando que ν_1 representa la normal del plano, $|Q_i \cdot \nu_1|$ es la distancia de *Hesse*. Es posible averiguar cuales puntos en $D_i(j)$ se encuentran alejados de formar un plano, mediante el criterio

$$\iota = |D_i(j) \cdot \nu_1|, \quad (4.18)$$

⁵ Con el objetivo de obtener su representación en Euclidiano mediante la proyección en sus bases e_1 , e_2 y e_3 .

tal que si ι es mayor a cierto umbral ($\iota > \alpha \cdot |Q_i|$), entonces dicho $D_i(j)$ debe considerarse un elemento de otra región y ser tratada en una etapa posterior. Una vez aplicada la técnica sobre toda la superficie de iso-disparidad es posible determinar las fronteras de las regiones ahora planas mediante el uso del álgebra de incidencia de la forma

$$l_{ab} = \pi_a \wedge \pi_b, \quad (4.19)$$

la línea de intersección entre regiones D_a y D_b planas adyacentes. La idea es obtener un modelo paramétrico faceteado de la reconstrucción tridimensional para la aplicación de algoritmos geométricos, dicha información podrá ser usada en aplicaciones como las realizadas en el capítulo 5.

4.3. Análisis en Multiresolución

Debido a la complejidad computacional de los algoritmos, tanto de reconstrucción, como de rectificación geométrica se requiere el mejoramiento del desempeño de la aplicación para fines de carácter práctico, es decir, si se desea hacer uso de los resultados anteriores es necesario que el ciclo completo sea calculado con una frecuencia apropiada para tareas industriales o bien hacer uso de los resultados obtenidos en etapas posteriores. Una solución adecuada para mejorar el desempeño es el análisis en diferentes resoluciones, iniciado con imágenes de baja resolución en las cuales se puede lograr una aproximación tosca de los resultados, una vez concluido el análisis en dicho nivel de resolución se proceden a propagar los resultados al nivel posterior, donde el algoritmo usará la información ahí contenida para computar el mínimo necesario (comprendido en un rango cercano al valor obtenido en el nivel inferior). El resultado del último nivel es la malla de reconstrucción. Este modo de minimizar el cálculo necesario es muy efectiva en particular en las técnicas de correlación y filtrado para la obtención de la dispari-

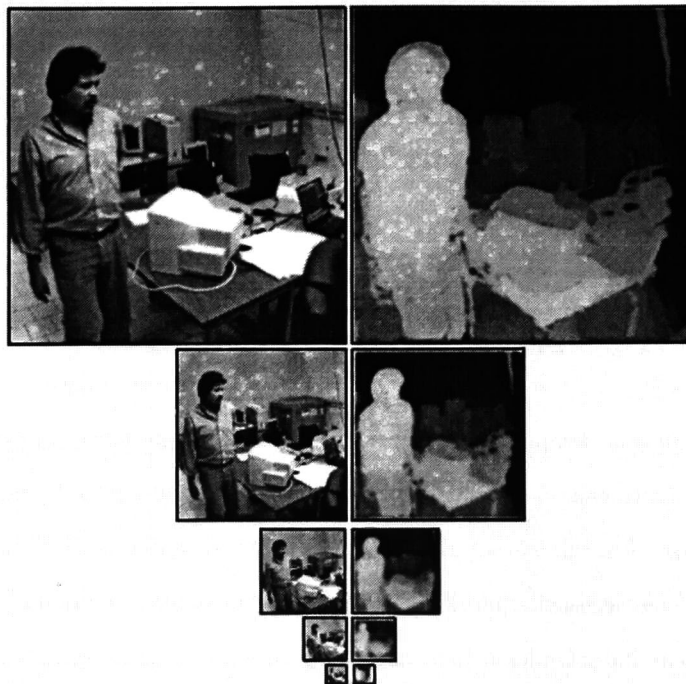


Figura 4.8: Implementación del algoritmo de generación de superficies de iso-disparidad en multiresolución, los resultados en los niveles inferiores de la pirámide se propagan a niveles superiores optimizando el proceso.

dad, dado que el rango de búsqueda entre la correlación de los píxeles queda definida en una región pequeña (aproximadamente entre 3-4 píxeles de radio en su máscara o vecindario), la complejidad-temporal se minimiza, sin embargo la complejidad-espacial se ve incrementada, pero sólo en una razón de $\frac{4}{3}$ el tamaño de la imagen original (ver la figura 4.8).

4.4. Segmentación en Profundidad

Una de las ventajas obtenidas mediante la reconstrucción y el análisis en multiresolución es que se puede almacenar los resultados de la reconstrucción del ego-mundo del sistema, con los cuales se puede hacer una gran variedad de algoritmos y operaciones, entre ellos se destaca la segmentación en profundidad. El modelo ego-céntrico del sistema puede comparar la posición de los datos obtenidos mediante la reconstrucción en un instante de tiempo, con los datos obtenidos anteriormente, de esta manera las variaciones en la profundidad (la posición) de los objetos o personas pueden ser seleccionadas y agrupadas en una colección de puntos o jerarquías U_i las cuales se analizan en forma más exhaustiva con el fin de terminar características tales como posición, velocidad, aceleración, volumen, etcétera, en términos tridimensionales o bien retroproyectando las regiones a las imágenes para la utilización de algoritmos de reconocimiento de rostros u otras posibles aplicaciones (ver figura 4.9).

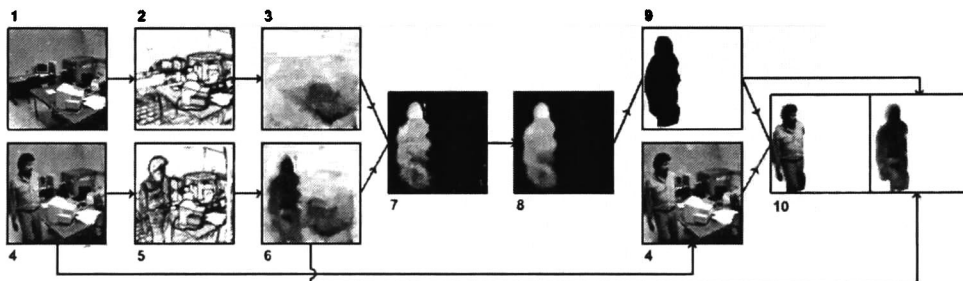


Figura 4.9: 1) Imagen del escenario $I(x, y, t_n)$. 2) Imagen calibrada y filtrada $I_c(x, y, t_n)$. 3) Mapa de iso-disparidad $Disp(x, y, t_n)$. 4) Imagen del escenario $I(x, y, t_{n+1})$. 5) Imagen calibrada y filtrada $I_c(x, y, t_{n+1})$. 6) Mapa de iso-disparidad $Disp(x, y, t_{n+1})$. 7) Diferencia en profundidad $D_p(x, y, t_{n+1}) = |Disp(x, y, t_{n+1}) - Disp(x, y, t_n)|$. 8) Diferencia en profundidad filtrada mediante mediana selectiva $D_{pf}(x, y, t_{n+1})$. 9) Umbralización $M(x, y, t_{n+1})$ (mascara). 10) Resultado de la segmentación en profundidad $I_s(x, y, t_{n+1})$ (Segmento de imagen) $Disp_s(x, y, t_{n+1})$ (Segmento de disparidad)

Capítulo 5

Aplicaciones

5.1. Interfaz Perceptual

La interacción humano máquina basada en esquemas multimodales[18], con varios tipos de Interfaz es una línea de investigación que cambiará en gran medida la forma en los seres humanos se relacionarán con los sistemas computacionales, así se mejorará la forma en que las personas interactúen con los dispositivos sin obstrucciones en un sentido más natural y perceptivo. En esta dirección se encuentran trabajos[4] con excelentes ideas y preguntas abiertas.

La idea fundamental en esta sección es hacer uso del modelo conformal de visión para Interfaz de usuario donde el suministro de información sea en base a interacción o acciones de los usuarios en el espacio visual que controla el sistema, es decir, interacción humano máquina basada en visión artificial mediante la detección de comportamientos o acciones que se deberán interpretar como cambios de modo o comandos emitidos por el usuario. En concreto, se puede establecer un escenario de pruebas donde el usuario haga uso del espacio visual para posicionar alguna de sus extremidades en el espacio delimitado mediante una entidad geométrica, este suceso puede asociar un comando,

así el usuario podrá en una forma sencilla tener el equivalente a un dispositivo de posición y emisión de comandos como el ratón de la computadora, con varias ventajas, la ausencia de artefactos o el suministro de información tridimensional. El grado de interacción entre el usuario y el sistema puede hacer que los objetos virtuales o primitivas tengan posiciones, colores, formas o escalas cambiantes, esto es similar a los botones y las ventanas en los sistemas *WIMP*¹ (*Windows Icons Mouse Pointes*), de esta manera la interacción hace uso de esquemas conocidos, intuitivos y poderosos. La aplicación en concreto de las ideas de interacción humano máquina basadas en visión que aquí se plantean no son limitantes e incluso es sólo una configuración de prueba (*Test-Bed*) para ilustrar los conceptos de interacción que aportan los sistemas de visión tridimensional, en particular el modelo aquí presentado. El esquema (ver la figura 5.1) presenta el sistema con su ego-visión y el usuario como una entidad la cual puede ser segmentada del resto de la escena debido al conocimiento a priori del espacio circundante (ver sección 4.4). Una vez segmentando el *Blob*² que comprende el cuerpo humano, la estrategia consiste en calcular si dicho *Blob* se encuentra en el interior de un subespacio definido por primitivas como esferas o bien si alguna extremidad ha atravesado al otro lado de un plano. El instante en el que se determina que una porción de dicho *Blob* se encuentra en el subespacio determinado, entonces se dispara el evento llamando *OnBlobEntering*, de forma análoga al momento que el *Blob* deja el subespacio se dispara el evento *OnBlobLeaving*, así se puede definir toda una serie de eventos con argumentos como los ya bien definidos en los marcos de trabajo como .Net, MFC de Microsoft³ *Java-EE*⁴ o bien el GTK+ de GNU⁵

¹ [http://en.wikipedia.org/wiki/WIMP_\(computing\)](http://en.wikipedia.org/wiki/WIMP_(computing))

² Superficie procedural.

³ <http://msdn.microsoft.com>

⁴ <http://java.sun.com/>

⁵ <http://www.gtk.org/>

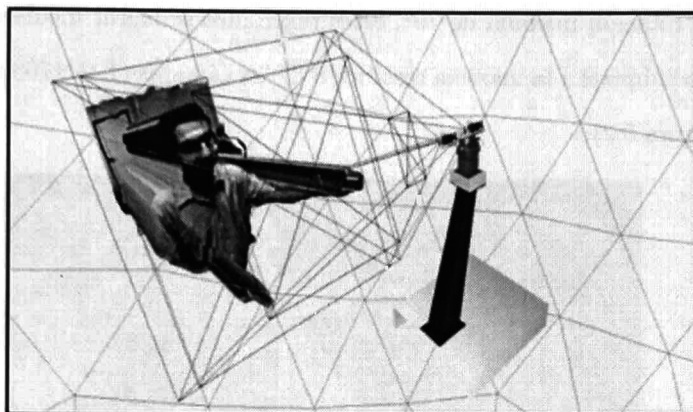


Figura 5.1: Sistema con ego-visión y usuario en realidad virtual.

5.2. Manipulación de Objetos

Entre las aplicaciones importantes de la visión artificial se puede destacar la manipulación de objetos, dado que determinar la posición, orientación y forma de los objetos puede ser obtenida en forma directa mediante las imágenes capturadas. El esquema consiste en tener un sistema donde las partes que lo conforman (en este caso, el cabezal con el sistema estereoscópico y el brazo robótico de forma antropomórfica con el actuador *BarrettHand*⁶ *BH8-260*) sean capaces de obtener información del medio para así poder establecer una trayectoria adecuada para la realización de alguna actividad. En el presente trabajo se hace uso del modelo conformal de visión tridimensional planteado anteriormente para calcular la posición-orientación de objetos mediante un método [2] que hace uso del álgebra geométrica y las técnicas de lógica difusa para realizar tareas de manipulación robótica con fines industriales (ver la figura 5.2). En lo referente a los resultados en métrica euclidiana se logró una precisión de ± 0.75 mm a una distancia de 850 mm, en lo referente a la orientación se logró una precisión de $\pm 0.5^\circ$

⁶ <http://www.barrett.com>

con una rotación máxima de 70° , estos resultados se logran mediante la aplicación de precisión subpíxel a la máxima resolución de las cámaras (1024x768 píxeles por imagen, ver la figura 5.2).

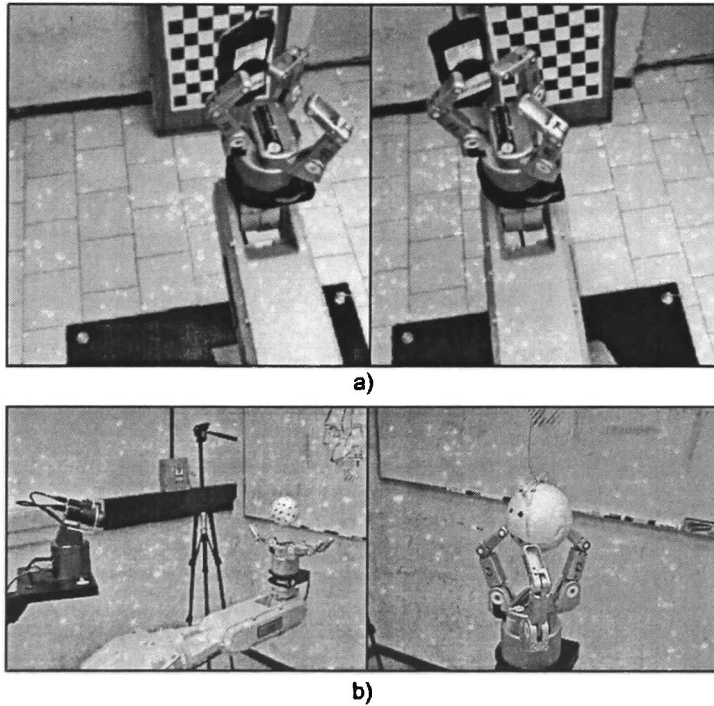


Figura 5.2: a) Imágenes de entrada para el algoritmo de manipulación. b) Manipulación de esfera mediante la aplicación del álgebra geométrica conformal.

Capítulo 6

Conclusiones

El modelo de visión conformal proporciona un esquema de procesamiento geométrico que aporta ventajas, unido a la fusión de técnicas de procesamiento de imágenes para la obtención de superficies de iso-disparidad para la reconstrucción del espacio visual. Sin embargo el volumen de información obtenida por estas etapas es muy superior a la

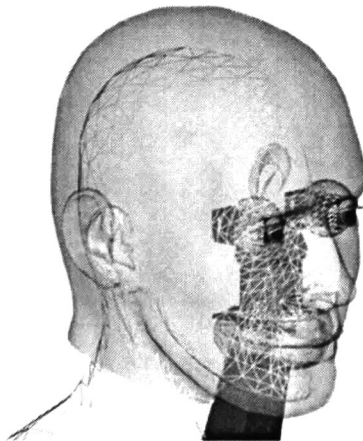


Figura 6.1: Visualización en realidad virtual del proyecto *PAC*.

capacidad de interpretación o manipulación para etapas de reconocimiento o planeación de tareas. Es ahí donde el método de detección de primitivas geométricas resulta útil al

reducir la cantidad de información necesaria para representar regiones de la superficie mediante aproximaciones a planos o alguna otra primitiva que reduce significativamente la complejidad temporal.

En este sentido no sólo se aporta una forma de optimizar la cantidad y calidad de información obtenida para etapas subsecuentes, sino también se aprovecha el lenguaje del álgebra geométrica para hacer uso de las relaciones de incidencia, con las cuales se determina la intersección de subespacios para aplicaciones como manipulación de objetos o interacción humano maquina.

Apéndice A

Álgebra Geométrica

El álgebra geométrica conformal, un caso particular de las álgebras de *Clifford* [21], posee conceptos y propiedades matemáticas de orden superior, tales como geometría proyectiva, álgebras-grupos de Lie, álgebra de incidencias, números complejos, coordenadas de Plücker, cuaternios entre otros. Una descripción formal, amplia y profunda del tema puede ser encontrada en [11][8][3]. A continuación se definen los conceptos funda-

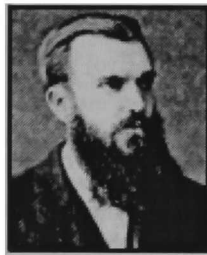


Figura A.1: William Kingdom Clifford (1845-1879).

mentales del álgebra geométrica, posteriormente se presenta una introducción general al álgebra geométrica para modelar los espacios Euclidianos \mathbb{E}^n , Proyectivo \mathbb{P}^n y finalmente el espacio conformal \mathbb{PK}^n , se hace énfasis en los conceptos importantes para el presente trabajo.

En forma general un álgebra geométrica $\mathcal{G}_{p,q}$ es un espacio lineal de 2^n dimensiones, $n = p + q$, donde p es la cantidad de bases cuya *cuadratura*¹ es 1, $e_i^2 = 1 : i \in p$, respectivamente q representa la cantidad de bases que poseen una cuadratura de -1 , $e_j^2 = -1 : j \in q$, de este modo (p, q) es llamada la firma del álgebra², lo cual define las características de las interpretaciones geométricas y algebraicas. En dicho espacio existen estructuras llamadas *Blades*³ que representan los subespacios lineales. Los *Blades* son utilizados para componer entidades algebraicas de grado superior llamadas *Multivectores*⁴

A.1. El Producto Exterior

El producto *wedge* o exterior se define en el álgebra de *Clifford* y se representa por el operador \wedge . Sean a, b y $c \in \mathbb{E}^n$ se tiene que:

- $a \wedge b = -b \wedge a$ (Anticonmutativo)

$$(a \wedge b) \wedge c = a \wedge (b \wedge c) \text{ (Asociativo)}$$

$$a \wedge (b \wedge c) = (a \wedge b) + (a \wedge c) \text{ (Distributivo)}$$

- $a \wedge b = 0 \iff a = \alpha b, \alpha \in \mathbb{R}$ (Nulo si es linealmente dependiente).

Sean $\{a_1, a_2, \dots, a_k\} \subset \mathbb{R}^n : k \leq n$ *vectores* linealmente independientes, entonces se tiene que $(a_1 \wedge a_2 \wedge \dots \wedge a_k) \wedge b = 0$ si b es una combinación lineal de $\{a_1, a_2, \dots, a_k\}$.

¹ Se denota como *cuadratura* de una base vectorial e_i , la operación $e_i^2 = e_i \cdot e_i$, cuyo resultado es un escalar (-1 ó 1).

² La firma de un Álgebra $\mathcal{G}_{p,q}$ define las propiedades del espacio que representa, por ello cada combinación de p y q denotará un espacio particular.

³ Se adopto la palabra inglesa debido que la traducción carece de sentido.

⁴ Combinaciones lineales de bases *Blades*.

El producto exterior de k vectores es llamado un k -Blade, el grado del Blade es el número de vectores que han sido operados con el producto exterior para producirlo, en otros términos el grado de un Blade implica la dimensión del subespacio que representa⁵ es decir,

$$A_{\langle k \rangle} := a_1 \wedge a_2 \wedge \dots \wedge a_k = \wedge_{i=1}^k a_i. \quad (\text{A.1})$$

A.1.1. Espacio Exterior Nulo

El producto exterior posee una interpretación geométrica como subespacio lineal, se puede visualizar mediante la siguiente función. Sea $x \in \mathbb{R}^n$ (1-Blade) y $\mathcal{U}(x) := x \wedge a \in \mathbb{R}^n$. El conjunto de vectores en \mathbb{R}^n que $\mathcal{U}(x)$ evalúa a cero (*Kernel* de \mathcal{U}) es llamado el espacio exterior nulo (*Outer product null space*)⁶ de a , es decir,

$$\text{NO}(a) := \{x \in \mathbb{R}^n : x \wedge a = 0 \in \mathbb{R}^n\}. \quad (\text{A.2})$$

Se sabe que $x \wedge a = 0$ si x es linealmente dependiente de a , es decir, $\text{NO}(a) = \{\alpha a : \alpha \in \mathbb{R}\}$, dicho espacio exterior puede ser visualizado como una línea que pasa por a y el origen (ver la figura A.2). Lo anterior no sólo sucede con los vectores ó 1-Blades, este concepto se puede generalizar a Blades de grado superior. Dado un bivector ó 2-Blade $a \wedge b \in \mathbb{R}^n$ se puede definir similarmente una función $g(x) := x \wedge a \wedge b \in \mathbb{R}^n$, cuyo *Kernel* es igual a $\text{NO}(a \wedge b) = \{\alpha a + \beta b : (\alpha, \beta) \in \mathbb{R}^2\}$, un subespacio bidimensional en \mathbb{R}^n expandido por los vectores a y b , es decir, el plano que pasa por los puntos a , b y el origen. Si se generaliza este concepto a un k -Blade $A_{\langle k \rangle}$ se tiene que el espacio exterior $\text{NO}(A_{\langle k \rangle})$ es un subespacio lineal k -dimensional de \mathbb{R}^n , es decir,

$$\text{NO}(A_{\langle k \rangle}) := \{x \in \mathbb{R}^n : x \wedge A_{\langle k \rangle} = 0\}. \quad (\text{A.3})$$

⁵ $A_{\langle k \rangle}$ denota el Blade A de grado k .

⁶ Esta interpretación se le conoce también como la forma dual o el espacio dual en [12].

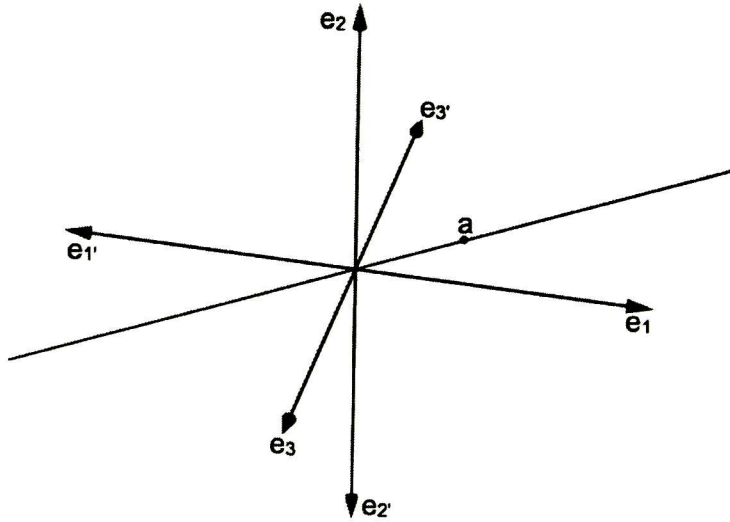


Figura A.2: Punto a y su respectivo $\text{NO}(a)$ como una línea que pasa por el origen y a .

A.2. *Blades*

El producto exterior sobre n vectores en un espacio \mathbb{R}^n genera un n -Blade, tal conjunto de n vectores deberán ser linealmente independientes entre si para que el producto exterior no se anule, por lo tanto dicho conjunto formara una base⁷ en \mathbb{R}^n , además pueden existir distintos vectores que formen diferentes bases para dicho espacio, por lo tanto diferentes conjuntos de vectores (todos de cardinalidad n cuyos elementos son linealmente independientes) operados bajo el producto exterior producirían Blades que variarían únicamente en un factor escalar, por ejemplo dados los vectores linealmente independientes a, b y $c \in \mathbb{R}^3$ entonces $\text{NO}(a \wedge b \wedge c) := \{x \in \mathbb{R}^3 : x \wedge a \wedge b \wedge c = 0 \in \mathbb{R}^3\}$ puesto que $\text{NO}(a \wedge b \wedge c) = \{\alpha a + \beta b + \rho c : (\alpha, \beta, \rho) \in \mathbb{R}^3\}$, por consiguiente, el espacio exterior nulo de $a \wedge b \wedge c$ comprende todo el espacio \mathbb{R}^3 , la única restricción sobre a, b y c es que no sean linealmente dependientes, por ende, los Blades resultantes sólo difieren

⁷ No necesariamente ortonormal.

entre si en un factor escalar, a ese *Blade* singular que comprende todo el espacio se le llama *Hipervolumen* o *Pseudoescalar*. En espacios n -dimensionales no es posible generar

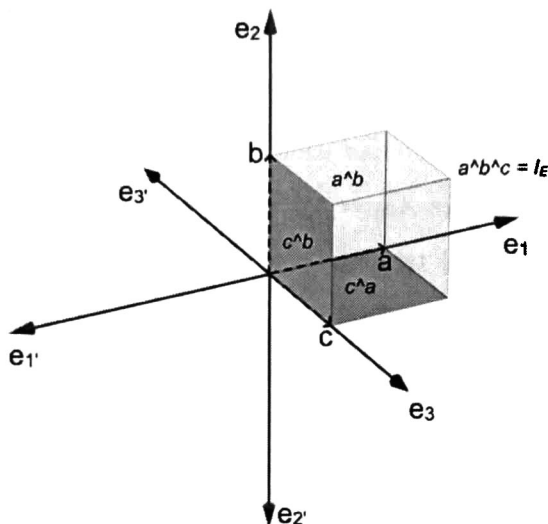


Figura A.3: El *Pseudoescalar* o *Hipervolumen* $I_E = e_1 \wedge e_2 \wedge e_3 = e_{123} \in \mathcal{G}_{3,0}$ resultado del producto exterior de $a = e_1$, $b = e_2$ y $c = e_3$.

Blades de orden superior a n .

El espacio Euclidiano tridimensional \mathbb{E}^3 tiene una álgebra geométrica \mathcal{G}_3 ($n = p = 3$) con tres bases ortonormales $\{e_1, e_2, e_3\}$ ⁸, el álgebra geométrica de dicho espacio consiste de 2^3 *Blades* bases, es decir, $\mathcal{G}_3 = \text{span}^9\{1, e_1, e_2, e_3, e_{23}, e_{31}, e_{12}, e_{123} = I_E\}$, el elemento escalar es el 0-*Blade*, mientras que las bases vectoriales son 1-*Blades*, los elementos $e_{ij} = e_i e_j = e_i \wedge e_j$ son 2-*Blades* unitarios o bases *bivectoriales*, finalmente el e_{123} es un *Trivector* unitario (3-*Blade*), en este caso $n = p = 3$ por ende e_{123} comprende todo el espacio, entonces dicho *Trivector* es el *Hipervolumen* o *Pseudoescalar* cuya cuadratura¹⁰ es -1 (ver la figura A.3.).

⁸ Vectores unitarios \hat{i} , \hat{j} y \hat{k} del cálculo vectorial.

⁹ Del inglés expandir.

¹⁰ $I_E \cdot I_E = (e_1 e_2 e_3) \cdot (e_1 e_2 e_3) = -1$.

A.2.1. Magnitud de *Blades*

Sean a y $b \in \mathbb{R}^3$, tal que $b = b_\alpha + b_\beta$, donde b_α representa la componente de b que es perpendicular a a , b_β representa la componente de b que es paralela a a (ver la figura A.4), entonces $a \wedge b = a \wedge (b_\alpha + b_\beta) = (a \wedge b_\alpha) + (a \wedge b_\beta)$, pero b_β es linealmente dependiente de a entonces $(a \wedge b_\beta) = 0$, por ende $a \wedge b = (a \wedge b_\alpha)$, lo anterior es generalizable a un k -*Blade*, es decir, sea $A_{(k)} = \wedge_{i=1}^k a_i$, es posible obtener un conjunto de k *vectores* ortogonales $\{a'_1, \dots, a'_k\}$ tales que $A_{(k)} = A'_{(k)} = \wedge_{i=1}^k a'_i$. Si el conjunto de *vectores*

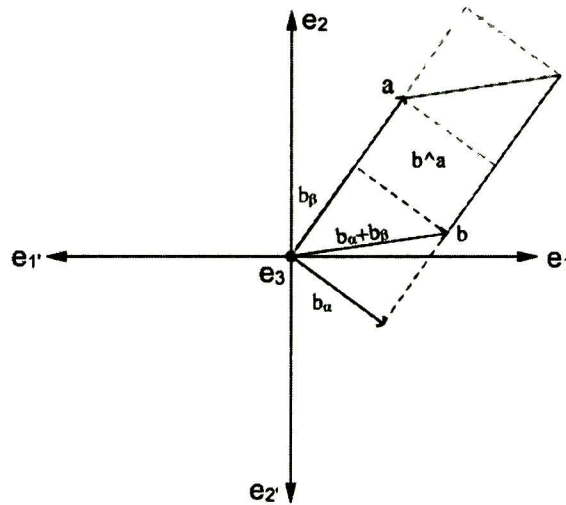


Figura A.4: *Vectores* a $b \in \mathbb{R}^3$, representación de la magnitud del *bivector* $|b \wedge a|$.

$\{a'_1, \dots, a'_k\}$ ortogonales entre si son operados mediante el producto *wedge*, entonces se puede mostrar que la magnitud del k -*Blade* que forman es $|A_{(k)}| = |A'_{(k)}| = \prod_{i=1}^k |a'_i|$. Lo anterior aplica para los k -*Blades* donde $k > 0$. Ejemplo sobre un espacio de tipo $\mathcal{G}_3(n = p = 3)$, el caso $k = 1$ se trata de los *vectores* y la magnitud de tal *Blade* es la norma del mismo. Si $k = 2$ se tiene un *bivector*, su magnitud equivale al area del paralelogramo expandido por los dos (*vectores*) 1-*Blades* que lo forman (ver la figura

A.4), similarmente la magnitud de un 3-*Blade* o *Trivector* puede ser visualizada como el volumen contenido o expandido por el paralelepípedo.

Sean $\{a'_i\}_{i=1}^k \subset \mathbb{R}^n$ un conjunto de k *vectores* en un espacio n -dimensional entonces es posible organizarlos matricialmente como $A = [a_1, \dots, a_k]$ (matriz $n \times k$), ahora se sabe que la norma de dicha matriz es el volumen que expanden los k *vectores* y se calcula mediante el determinante de A , lo cual es análogo a la magnitud del k -*Blade* formado por $|\{a_1 \wedge \dots \wedge a_k\}| = \det([a_1, \dots, a_k])$.

A.3. El Producto Interior

Al igual que el producto exterior el producto interior está definido en el álgebra geométrica para *Blades* de cualquier grado, siendo el caso base los 1-*Blades* o bases vectoriales, en dicho caso el producto interior se reduce al bien conocido producto punto o escalar (denotado por “.”). Si a y $b \in \mathbb{R}^n$ entonces $a \cdot b = s : s \in \mathbb{R}$. El producto interior del *bivector* $b \wedge c$ con un *vector* es de la forma

$$a \cdot (b \wedge c) = (a \cdot b)c - (a \cdot c)b, \tag{A.4}$$

donde $(a \cdot b)$ y $(a \cdot c)$ son escalares, el resultado de aplicar el producto interior entre un *vector* (1-*Blade*) y un *bivector* (2-*Blade*) es un *vector* (1-*Blade*). En forma más amplia si se tiene que $x \cdot A_{(k)} : k \geq 1$ es decir, el producto interior de un 1-*Blade* con un k -*Blade*

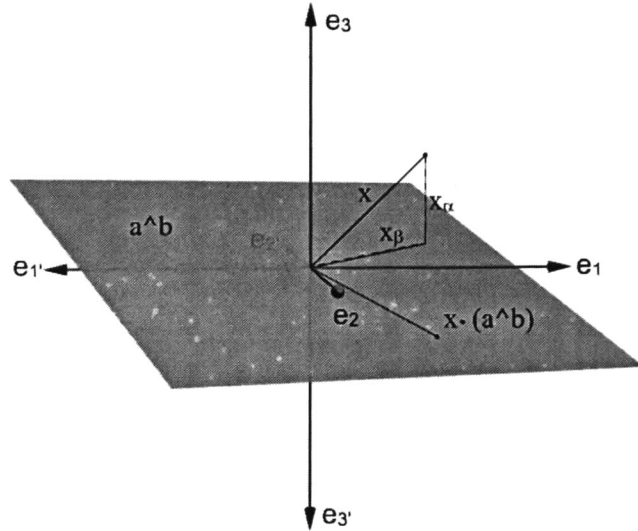


Figura A.5: Producto Interior de un *bivector* y un *vector*.

$$\begin{aligned}
 x \cdot A_{(k)} &= (x \cdot a_1)(a_2 \wedge a_3 \wedge a_4 \wedge \dots \wedge a_k) \\
 &\quad - (x \cdot a_2)(a_1 \wedge a_3 \wedge a_4 \wedge \dots \wedge a_k) \\
 &\quad + (x \cdot a_3)(a_1 \wedge a_2 \wedge a_4 \wedge \dots \wedge a_k) \\
 &\quad - (x \cdot a_4)(a_1 \wedge a_2 \wedge a_3 \wedge \dots \wedge a_k) \dots \\
 &= \sum_{i=1}^k (-1)^{i+1} (x \cdot a_i) [A_{(k)}/a_i],
 \end{aligned}$$

en donde $[A_{(k)}/a_i]$ denota el *Blade* $A_{(k-1)}$ sin el *vector* a_i . El resultado de dicho producto interior es un *Blade* de grado $k-1$, si se ejecuta un producto interior entre un *2-Blade* y un k -*blade* tal que $k \geq 2$ se tiene que

$$(a \wedge b) \cdot A_{(k)} = a \cdot (b \cdot A_{(k)}). \tag{A.5}$$

Generalizando un k -*Blade* y un l -*Blade* el resultado del producto interior es

$$A_{(k)} \cdot B_{(l)} = a_1 \cdot (a_2 \cdot \dots \cdot (a_k \cdot B_{(l)})), \tag{A.6}$$

cuyo resultado es un $(l - k)$ Blade. En este sentido el resultado del producto interior decrementa el grado del Blade resultante mientras que el producto exterior incrementa el grado del Blade, es por ellos que también se denotan como “contracción” y “expansión” al producto interior y exterior respectivamente. Una forma de visualizar las implicaciones del producto interior en un espacio dado es la siguiente. Sean x, y, a y $b \in \mathbb{E}^n$, tal que $y = x \cdot (a \wedge b) = (x \cdot a)b - (x \cdot b)a$ entonces $s = x \cdot y = x \cdot [(x \cdot a)b - (x \cdot b)a] = (x \cdot a)(x \cdot b) - (x \cdot b)(x \cdot a) = 0$, implica que el resultado de $x \cdot (a \wedge b)$ es perpendicular a x , o bien, y es el subespacio resultante de eliminar el subespacio x de $a \wedge b$, por lo tanto y es linealmente independiente de x , es decir, perpendicular a x . El vector $x = x_\alpha + x_\beta$, entonces $y = (x_\alpha + x_\beta) \cdot (a \wedge b) = (x_\alpha \cdot (a \wedge b)) + (x_\beta \cdot (a \wedge b))$, pero como $x_\alpha \cdot (a \wedge b) = 0$ (son perpendiculares) se tiene $y = (x_\alpha + x_\beta) \cdot (a \wedge b) = (x_\beta \cdot (a \wedge b))$ lo cual es el subespacio perpendicular a x_β contenido en $(a \wedge b)$, ver la figura A.5. En forma general el producto interior de dos Blades $A_{(k)}$ y $B_{(l)}$ se define como:

$$\begin{aligned} A_{(k)} &= a_1 \wedge \dots \wedge a_k \\ B_{(l)} &= b_1 \wedge \dots \wedge b_l \\ A_{(k)} B_{(l)} &= (a_1 \wedge \dots \wedge a_k) \cdot (b_1 \wedge \dots \wedge b_l) \\ \text{si } k \geq l &\Rightarrow ((a_1 \wedge \dots \wedge a_k) \cdot b_1) \cdot (b_2 \wedge \dots \wedge b_l) \\ \text{si } k < l &\Rightarrow (a_1 \wedge \dots \wedge a_{k-1}) \cdot (a_k \cdot (b_1 \cdot (b_2 \wedge \dots \wedge b_l))) \end{aligned}$$

donde

$$(a_k \cdot (b_1 \cdot (b_2 \wedge \dots \wedge b_l))) = \sum_{i=1}^l (-1)^{(i-1)} (b_1 \cdot b_2 \wedge \dots \wedge b_{i-1} \wedge (a_k \cdot b_i) \wedge b_{i+1} \wedge \dots \wedge b_l),$$

esta definición recursiva es conocida como el producto interior generalizado [12] y se utilizará posteriormente en conceptos de dualidad e inversa de Blades. Esta manera de interpretar el producto interior así como otras relaciones entre el producto exterior y el Pseudoscalar son presentadas más a detalle en la sección A.4.

A.3.1. Inversa de *Blades*

Sea $A_{\langle k \rangle} : k \leq n$ un k -*Blade*, entonces la inversa se denota como $A_{\langle k \rangle}^{-1}$, deberá satisfacer la relación $A_{\langle k \rangle} A_{\langle k \rangle}^{-1} = A_{\langle k \rangle}^{-1} \cdot A_{\langle k \rangle} = 1$. La inversa de un k -*Blade* es

$$A_{\langle k \rangle}^{-1} = \frac{A_{\langle k \rangle}}{|A_{\langle k \rangle}|^2}, \quad (\text{A.7})$$

sin embargo dado el comportamiento del producto interior es posible que $A_{\langle k \rangle} \cdot A_{\langle k \rangle}^{-1} = -1$ es por ello que $A_{\langle k \rangle}^{-1} = \pm \frac{A_{\langle k \rangle}}{|A_{\langle k \rangle}|^2}$, el signo depende tanto del grado del *Blade* como de la firma del álgebra en cuestión. Es necesario un operador que ajuste el signo, dicho operador se conoce como la reversión. La reversión de un *Blade* $A_{\langle k \rangle}$ se denota como $\tilde{A}_{\langle k \rangle}$ y consiste simplemente en invertir el orden de los *vectores* que conforma el *Blade*, es decir, si $A_{\langle k \rangle} = \wedge_{i=1}^k a_i$ entonces

$$\tilde{A}_{\langle k \rangle} = \wedge_{i=k}^1 a_i = a_k \wedge a_{k-1} \wedge \dots \wedge a_1. \quad (\text{A.8})$$

La anticonmutatividad del producto *wedge* muestra un patrón de comportamiento ya que al realizar una conmutación de la forma $a \wedge b = -b \wedge a$ y $b \in \mathbb{E}^n$, se tiene un cambio de signo, entonces para un k -*Blade* se tendrá

$$\tilde{A}_{\langle k \rangle} = (-1)^{(k(k-1)^2)/2} A_{\langle k \rangle}. \quad (\text{A.9})$$

Es posible definir la inversa para cualquier k -*Blade* en cualquier álgebra como $A_{\langle k \rangle}^{-1} = \frac{\tilde{A}_{\langle k \rangle}}{|A_{\langle k \rangle}|^2}$ siempre que su magnitud sea diferente de cero¹¹

A.3.2. Espacio Interior Nulo

El producto interior define un espacio llamado el espacio interior nulo (*Inner product Null Space*). Sea $l(A_{\langle k \rangle})$ una función definida como $l(A_{\langle k \rangle}) := x \cdot A_{\langle k \rangle}$ donde $x \in \mathbb{R}^n$,

¹¹ Dado que $|A_{\langle k \rangle}| = 0$ es posible en álgebras con $q \neq 0 \rightarrow |A_{\langle k \rangle}| = 0$.

entonces el *Kernel* de dicha función define un espacio interior nulo sobre el *Blade* $A_{(k)}$, lo cual es denotado como $\text{NI}(A_{(k)}) := \{x \in \mathbb{R}^n : l(A_{(k)}(x)) = 0\}$. Ejemplo, considérese el *bivector* $S = a \wedge b$, entonces $\text{NI}(S) = \{x \in \mathbb{E}^3 : x \cdot (a \wedge b) = (x \cdot a)b - (x \cdot b)a = 0\}$ la condición anterior muestra que $(x \cdot a)b = (x \cdot b)a$ dado que a y b deben ser linealmente independientes (de lo contrario $a \wedge b = 0$), entonces tanto $(x \cdot a) = 0$ y $(x \cdot b) = 0$, significa x es perpendicular tanto a a y b , tal *vector* x debe ser perpendicular al plano formado por $a \wedge b$, cuyo NI es la línea formada por los puntos en la dirección normal al plano representado por el espacio exterior nulo del *bivector* S ver la figura A.6. El

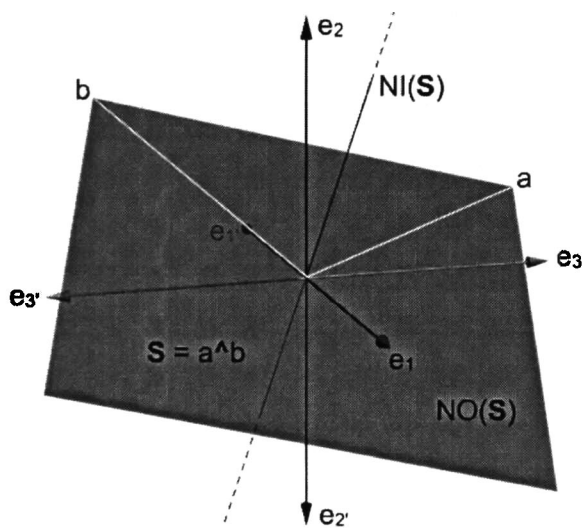


Figura A.6: Espacio interior nulo de un *bivector* $\text{NI}(S = a \wedge b)$, espacio exterior nulo $\text{NO}(S = a \wedge b)$.

espacio exterior nulo de un *vector* (*1-Blade*) representa un línea que pasa por el origen y a su vez el espacio interior de un *bivector* (*2-Blade*) es una línea que pasa por el origen, lo que sugiere la existencia de una relación dual entre ambos espacios nulos.

A.4. Dualidad

Con el objetivo de presentar el operador dual (*) que relaciona los espacios nulos se presentan las siguientes definiciones de operaciones sobre conjuntos de *vectores*. La *suma directa* de conjuntos de *vectores* se denota con \oplus , si se tiene el conjunto $A = \{a_i\}_{i=1}^k \subset \mathbb{R}^n$ y $B = \{b_i\}_{i=1}^l \subset \mathbb{R}^n$ entonces $A \oplus B := \{a_i + b_j \in \mathbb{R}^n : 1 \leq i \leq k, 1 \leq j \leq l\}$, por ejemplo, en el caso de dos conjuntos infinitos $A := \{\alpha \mathbf{a} \in \mathbb{R}^n : \alpha \in \mathbb{R}\}$ y $B := \{\beta \mathbf{b} \in \mathbb{R}^n : \beta \in \mathbb{R}\}$ la suma directa $A \oplus B := \{\alpha \mathbf{a} + \beta \mathbf{b} \in \mathbb{R}^n : (\alpha, \beta) \in \mathbb{R}\}$, es el conjunto de todas las combinaciones lineales. Similarmente se define la *resta directa* entre conjuntos de *vectores* como $A \ominus B := \{x \in A : x \cdot y = 0 \forall y \in B\}$. La *resta directa* no significa eliminar los elemento $b_i \in B$ del conjunto A , sino que implica eliminar la dependencia lineal de todos los elementos de B en los elementos de A . En el espacio Euclidiano tridimensional existen tres *vectores* ortonormales llamados e_1, e_2 y e_3 , haciendo uso de la definición del producto exterior se tiene que $\text{NO}(e_1) = \{\alpha e_1 : \alpha \in \mathbb{R}\}$, visualizado como una línea que pasa por el origen, de este modo $\text{NO}(e_1 \wedge e_2)$ también se puede visualizar con plano que pasa por el origen (ver la figura A.6). Haciendo uso de la *suma directa*

$$\text{NO}(e_1) \oplus \text{NO}(e_2 \wedge e_3) = \{\alpha e_1 + \beta e_2 + \chi e_3 : (\alpha, \beta, \chi) \in \mathbb{R}^3\},$$

es todo el espacio \mathbb{R}^3 . Es posible eliminar la dependencia lineal (*resta directa*) de $\text{NO}(e_1)$ en \mathbb{R}^3 como $\mathbb{R}^3 \ominus \text{NO}(e_1) = \text{NO}(e_2 \wedge e_3)$ en este contexto (espacio particular \mathbb{E}^3) es posible denotar el complemento de $\text{NO}(e_1)$ como $\text{NO}(e_2 \wedge e_3)$, haciendo uso de las operaciones *directas* sobre conjuntos se tiene $\text{NI}(e_1) = \mathbb{R}^3 \ominus \text{NO}(e_1)$. El dual es el operador que transforma un *k-Blade* en su *l-Blade* complementario en un espacio \mathbb{R}^n tal que $l = n - k : k \leq n$, es decir, si $\text{NO}(A_{(k)}) = \mathbb{R}^n \ominus \text{NO}(B_{(l)})$, se representa como $A^* = A \cdot I^{-1}$, donde I^{-1} es el inverso del *Pseudoescalar* unitario. Es necesario hacer uso

del inverso del *Pseudoescalar* unitario dado que de otra forma

$$(A^*)^* = (A \cdot I) \cdot I = A \cdot (I \cdot I) = \pm A, \quad (\text{A.10})$$

donde el cambio de signo dependería de la firma del álgebra específica. Ejemplo, sean las bases ortonormales e_1, e_2 y $e_3 \in \mathbb{R}^3$ entonces el *bivector* $S = e_1 \wedge e_2$, posee un dual de la forma

$$\begin{aligned} S^* &= S \cdot I^{-1} \\ &= (e_1 \wedge e_2) \cdot (e_3 \wedge e_2 \wedge e_1) \\ &= e_1 \cdot (e_2 \cdot (e_3 \wedge e_2 \wedge e_1)) \\ &= e_1 \cdot ((e_2 \cdot e_3)(e_2 \wedge e_1) - (e_2 \cdot e_2)(e_3 \wedge e_1) + (e_2 \cdot e_1)(e_3 \wedge e_2)) \\ &= e_1 \cdot (-(e_3 \wedge e_1)) \\ &= e_1 \cdot (e_1 \wedge e_3) \\ &= (e_1 \cdot e_1)e_3 - (e_1 \cdot e_3)e_1 \\ &= e_3, \end{aligned}$$

de este modo se puede ver que $S \wedge S^* = (e_1 \wedge e_2) \wedge e_3 = e_1 \wedge e_2 \wedge e_3 = I$ dado que el dual de dicho *2-Blade* es su complemento en el espacio, por consiguiente $\text{NO}((e_1 \wedge e_2)) \oplus \text{NO}((e_1 \wedge e_2)^*) = \mathbb{R}^3$, la relación entre el espacio exterior nulo y el espacios nulos interior de un *Blade* es el dual, osea, $\text{NO}((e_1 \wedge e_2)) = \mathbb{R}^3 \ominus \text{NO}(e_3)$ o bien $\text{NO}((e_1 \wedge e_2)) = \text{NI}((e_1 \wedge e_2)^*)$, de modo general se puede decir que $\text{NO}(A_{(k)}) = \text{NI}(A_{(k)}^*)$.

A.5. Intersecciones

El subespacio $\text{NI}(B_{\langle 2 \rangle})$ es el conjunto de puntos contenidos en una línea perpendicular al plano que pasa por el origen, entonces $\text{NI}(a \wedge b) = \text{NI}(a) \cap \text{NI}(b)$, es decir $\text{NI}(a)$ será el plano cuya normal es a y respectivamente $\text{NI}(b)$ será el plano cuya normal es b ,

por lo tanto la intersección de ambos planos en el espacio \mathbb{R}^3 será una línea perpendicular tanto a b como a a , dicha línea a su vez debe ser la norma del plano formado por $a \wedge b$ por consiguiente (en la interpretación del $\mathbb{N}\mathbb{I}$) se puede obtener la intersección de los *Blades* mediante la aplicación del producto exterior, osea $\mathbb{N}\mathbb{I}(a \wedge b) = \mathbb{N}\mathbb{I}(a) \cap \mathbb{N}\mathbb{I}(b)$. Existe una relación entre la interpretación de los espacio nulos y el dual de los mismos para definir un operador de intersección de subespacios.

Supóngase $A = a_1 \wedge a_2$ y $B = b_1 \wedge b_2$ son *bivectores* en \mathbb{R}^3 entonces A y B representan planos en su interpretación de espacio exterior nulo ($\mathbb{N}\mathbb{O}$). Ahora si se tiene $A^* = (a_1 \wedge a_2)^*$ y $B^* = (b_1 \wedge b_2)^*$, entonces $A^* \wedge B^*$ en términos de $\mathbb{N}\mathbb{I}$ representa la intersección de A con B , posteriormente en términos de su espacio exterior nulo está intersección deberá ser de la forma $(A^* \wedge B^*) \cdot I$, obteniendo la expresión¹² para determinar la intersección de dos *Blades*, es decir, $((a_1 \wedge a_2)^* \wedge (b_1 \wedge b_2)^*) \cdot I$. El tener una expresión u operador que obtenga la intersección de subespacios como la anterior debe cumplirse en cualquier combinación de *Blades* lo cual no es posible realizar con la expresión anterior. A modo de contra ejemplo, sean e_1, e_2 y $e_3 \in \mathbb{R}^3$ las bases vectoriales del espacio Euclidiano tridimensional entonces el *bivector* $A = e_1 \wedge e_2$ representa el plano orientado (ver la figura A.7), posteriormente el *vector* $B = \alpha e_1 + \beta e_2 : (\alpha, \beta) \in \mathbb{R}^2$ representa¹³ una línea en el plano A , por lo tanto la intersección de A con B deberá ser B misma, dado que dicha línea yace en el plano A , ahora al calcular la intersección se tiene que

$$\begin{aligned}
 R &= (A^* \wedge B^*) \cdot I \\
 &= (((e_1 \wedge e_2)^* \wedge (\alpha e_1 + \beta e_2)^*)) \cdot I \\
 &= (e_3 \wedge (\alpha(e_3 \wedge e_2) + \beta(e_1 \wedge e_3))) \cdot I \\
 &= (0) \cdot I \\
 &= 0,
 \end{aligned}$$

¹² Esta expresión es conocida también como el producto regresivo.

¹³ En su interpretación $\mathbb{N}\mathbb{O}$.

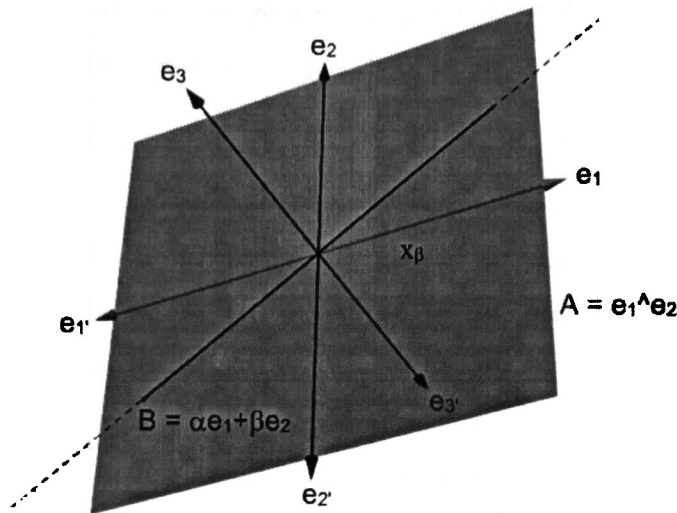


Figura A.7: Intersección entre el plano $A = e_1 \wedge e_2$ y la línea $B = \alpha e_1 + \beta e_2$.

un resultado incorrecto, esto se debe a que ambos *Blades* tanto la línea como el plano habitan en un subespacio bidimensional de \mathbb{R}^3 y al ser dualizados con un *Pseudoescalar*(I^{-1}) que cubre todo el espacio (\mathbb{R}^3) se incluye el subespacio e_3 en la forma dual de A^* y B^* , al aplicar el producto exterior para obtener la intersección se obtendrá siempre cero, es decir, no hay intersección. La idea para solucionar esta situación es que los *Blades* en los que se está evaluando la intersección deberán ser dualizados en el espacio¹⁴ que expandan sus k y l bases. El operador llamado unión o *Join* genera el *Pseudoescalar* adecuado para el subespacio que expanden los dos *Blades* cuya intersección se está evaluando. La unión se denota como $\hat{\wedge}$, por ejemplo en el caso anterior se tendría

$$J = A \hat{\wedge} B = (e_1 \wedge e_2) \hat{\wedge} (\alpha e_1 + \beta e_2) = (e_1 \wedge e_2), \quad (\text{A.11})$$

con lo que es ahora posible obtener la intersección como $R = ((A \cdot J^{-1}) \wedge (B \cdot J^{-1})) \cdot J$.

En el ejemplo anterior se tiene

¹⁴ O subespacio según sea el caso.

$$\begin{aligned}
 R &= (1 \wedge (\beta e_1 - \alpha e_2)) \cdot J \\
 &= (\beta e_1 - \alpha e_2) \cdot J \\
 &= \alpha e_1 + \beta e_2,
 \end{aligned}$$

el resultado esperado, la intersección de la línea con el plano deberá ser la misma línea, nótese el extremo derecho no hace la inversión de la unión (J). Esta expresión de intersección de *Blades* es usualmente conocida como el *Meet* y se denota por \vee , es decir, $A_{(k)} \vee B_{(l)} = [(A_{(k)} \cdot J^{-1}) \wedge (B_{(l)} \cdot J^{-1})] \cdot J$, lo cual implica en la interpretación del producto exterior nulo que $\text{NO}(A_{(k)} \vee B_{(l)}) = \text{NO}(A_{(k)}) \cap \text{NO}(B_{(l)})$, la intersección de los subespacios representados por los *Blades*.

A.6. Producto Geométrico

Un *Multivector* es una combinación lineal de bases *Blades* de diversos grados, por ejemplo un *Multivector* de grado r , es decir, $A_r = \langle A \rangle_0 + \langle A \rangle_1 + \dots + \langle A \rangle_r$, donde $\langle A \rangle_i$ es usado para denotar la parte de grado i de un *Multivector* (operador llamado el *Selector*). Un *Multivector* es homogéneo si consiste sólo de elementos de un sólo grado. En el álgebra geométrica existe el producto geométrico que es denotado por la yuxtaposición de dos *Multivectores* a ser operados, por ejemplo en el caso de los *Multivectores* A_r y B_s se tendría que

$$A_r B_s = \langle AB \rangle_{r+s} + \langle AB \rangle_{r+s-2} + \dots + \langle AB \rangle_{|r-s|}. \quad (\text{A.12})$$

Si a y $b \in \mathbb{R}^n$ son dos *vectores* ó *1-Blades* entonces el producto geométrico $ab = (a \cdot b) + (a \wedge b)$, es decir, un par de *Multivectores* homogéneos como $A_1 = a$ y $B_1 = b$ ambos de grado 1, entonces $A_1 B_1 = \langle AB \rangle_{1+1} + \langle AB \rangle_{|1-1|} = (a \wedge b) + (a \cdot b)$. El producto geométrico es asociativo y distributivo pero no es conmutativo en general, además se

sabe que si la intersección de dos *Blades* no es vacía, entonces su intersección es el producto geométrico que a su vez es igual al producto exterior, es decir,

$$\mathbb{NO}(A_{(k)}) \cap \mathbb{NO}(B_{(l)}) \neq 0 \iff A_{(k)}B_{(l)} = A_{(k)} \wedge B_{(l)}, \quad (\text{A.13})$$

de igual modo si

$$\mathbb{NO}(A_{(k)}) \supseteq \mathbb{NO}(B_{(l)}) \text{ ó } \mathbb{NO}(A_{(k)}) \subseteq \mathbb{NO}(B_{(l)}) \iff A_{(k)}B_{(l)} = A_{(k)} \vee B_{(l)}. \quad (\text{A.14})$$

A.7. Reflexión

Los conceptos anteriores han presentado maquinaria para el manejo de subespacios lineales (*Blades*) sobre un espacio con firma determinada, dicha maquinaria puede construir un subespacio lineal al expandirlo mediante subespacios más simples aplicando el producto exterior, substraer un subespacio de otro mediante el producto interior y obtener la intersección de subespacios lineales mediante la operación de intersección (*Meet*). Es necesario tener mecanismos para operar sobre dicho subespacios lineales con operaciones que mantengan el grado de las entidades intacto, por ejemplo, en el caso de la interpretación geométrica de los *Blades* se puede tener alguna operación que permita girar puntos, líneas o planos en el espacio y el resultado de dicha operación deberá ser de igual forma puntos, líneas o planos. Este tipo de operación o transformación es la reflexión, transformaciones más complicadas pueden ser compuestas por varias reflexiones. Sean $x, y \in \mathbb{R}^n : |y| = 1$ entonces una reflexión de x con respecto a y se define como

$$\begin{aligned}
 x' &= yxy = (yx)y \\
 &= (y \cdot x + y \wedge x)y \\
 &= (y \cdot x)y + (y \wedge x) \cdot y \\
 &= (y \cdot x)y + (x \cdot y)y - (y \cdot y)x \\
 &= 2(y \cdot x)y - x,
 \end{aligned}$$

y como $x = x_\alpha + x_\beta$, tales que x_α y x_β son las partes de x perpendiculares y paralelas a y respectivamente, por lo tanto $y \cdot x = y \cdot x_\alpha = |x_\alpha|$ entonces se tiene que

$$\begin{aligned}
 x' &= 2(y \cdot x)y - x \\
 &= 2(|x_\alpha|)y - x \\
 &= 2x_\alpha - x \\
 &= 2x_\alpha - (x_\alpha + x_\beta) \\
 &= x_\alpha - x_\beta,
 \end{aligned}$$

por consiguiente x' es el *vector* x reflejado respecto a y , es decir, el *vector* cuya parte perpendicular a y ha sido invertida ($-x_\beta$), esto se puede visualizar en la figura A.8. Dicha operación puede reflejar entidades más generales o *Blades* de cualquier grado. Operaciones más complejas se basan en la reflexión. La rotación es la combinación de dos reflexiones, la translación es la rotación con un eje en el infinito y así se expresan sucesivamente las transformaciones geométricas en los espacios Euclidiano, Proyectivo y Conformal. Cualquier transformación que preserve el grado de un *Blade* será una combinación de reflexiones.

A.8. Rotaciones

Al reflejar con respecto a un *Blade* unitario sólo se modifica la dirección, si se ejecutan en forma consecutiva dos reflexiones con respecto a dos distintos *Blades* n y

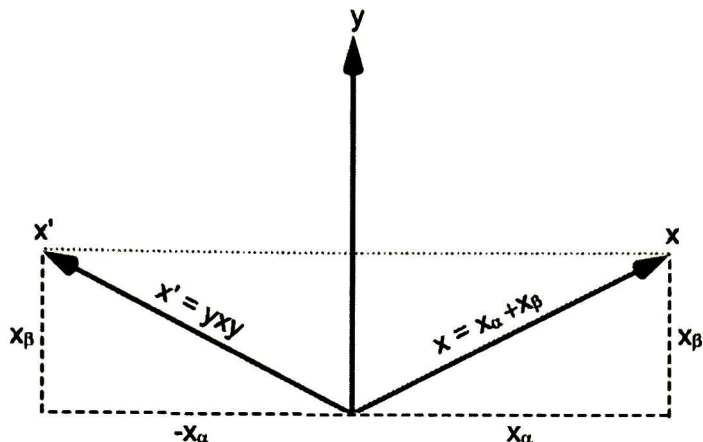


Figura A.8: Reflexión del vector x respecto al vector y .

m ambos unitarios, entonces se puede obtener una rotación que es equivalente a rotar dos veces el ángulo ($\theta = 2\angle_{nm}$) comprendido entre n y m (ver la figura A.9). plano, es decir, (el eje de rotación será la normal del plano) $a' = m(nan)m = mnanm$. Ahora dado que el producto geométrico se define para los *vectores* como $mn = m \cdot n + m \wedge n$, $nm = n \cdot m + n \wedge m$, $n \cdot m = m \cdot n \in \mathbb{R}$ y $m \wedge n = \widetilde{(n \wedge m)}$ y la reversión de cualquier escalar es el escalar en sí, entonces se puede afirmar que $mn = \widetilde{nm}$, dicho producto es conocido como el rotor o $R = mn$ y $\widetilde{R} = nm$. Para que la rotación pueda ser expresada como $a' = Ra\widetilde{R}$ es necesario que $R\widetilde{R} = 1$, es decir, el motor no debe cambiar la escala de la entidad a (sólo rotarla). El rotor puede ser desmembrado de la forma

$$R = mn = m \cdot n + m \wedge n = \cos(\theta) + \psi \sin(\theta) \tag{A.15}$$

y

$$\psi = \frac{(m \wedge n)}{|m \wedge n|}, \tag{A.16}$$

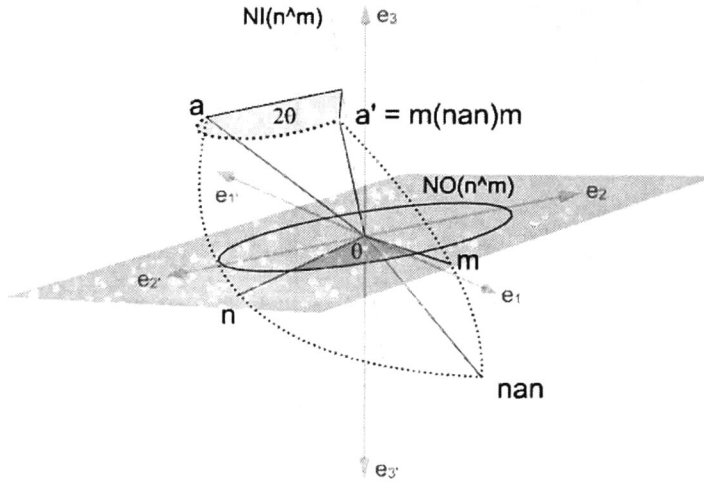


Figura A.9: Rotación del *vector* a mediante dos reflexiones consecutivas con los *vectores* unitarios n y m , nótese el eje de rotación es la línea $\mathbb{N}\mathbb{I}(n \wedge m)$.

es decir, el *bivector* unitario. Dado que

$$\psi \cdot \psi = \frac{(m \wedge n)}{|m \wedge n|} \cdot \frac{(m \wedge n)}{|m \wedge n|} = -1 \quad (\text{A.17})$$

(posee una cuadratura de -1), es posible tener una representación similar a los números complejos en su forma polar

$$z = r(\cos(\theta) + i \sin(\theta)) \approx R = \cos(\theta) + \psi \sin(\theta), \quad (\text{A.18})$$

con el teorema de Euler

$$z = e^{i\theta} \approx R = e^{-\psi(\frac{\theta}{2})}. \quad (\text{A.19})$$

donde sólo rotara la mitad del ángulo (es decir, el ángulo comprendido entre m y n), el cambio de signo tiene el propósito de mantener la convención de la dirección de la rotación. Las rotaciones (al igual que las reflexiones) son generalizables a cualquier dimensión y pueden ser aplicadas a *Blades* de cualquier grado.

A.9. Espacio Projectivo

En las secciones anteriores se presentó brevemente el álgebra geométrica y la forma en que los elementos del álgebra representan entidades geométricas, sin embargo las entidades carecen de generalidad, es decir, todos los planos y líneas pasan por el origen tanto en su representación NO como en NI , es una limitante para describir elementos geométricos. El mecanismo para lograr esta característica y otras más es incluir (envolver) el espacio euclidiano \mathbb{E}^n como un subespacio de otro espacio de más dimensiones. Un aspecto importante es que los *Blades* como subespacios lineales y la reflexión como operador permanecen intactos en el nuevo espacio. En contraparte la interpretación geométrica de los *Blades* y los efectos colaterales de dicho proceso (envolvimiento) cambiarán radicalmente dependiendo del tipo específico de envolvente espacial que se le realice sobre \mathbb{E}^n . El espacio projectivo es un caso particular de inmersión del espacio

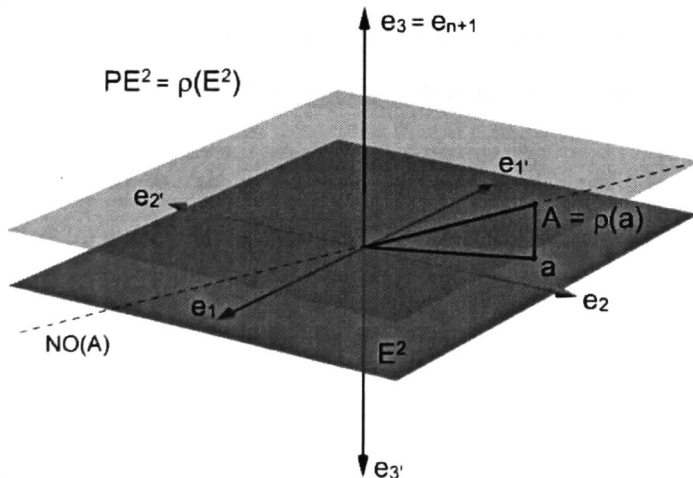


Figura A.10: Espacio projectivo \mathbb{PE}^2 resultado de envolver el espacio Euclidiano \mathbb{E}^2 con el operador ρ .

Euclidiano \mathbb{E}^n denotado por $\mathbb{P}\mathbb{E}^n$. El espacio proyectivo u homogéneo deriva sus características en base a la inmersión que realiza, es decir, el espacio $\mathbb{P}\mathbb{E}^n$ representa $\mathbb{R}^{n+1}/0$, un espacio $n + 1$ -dimensional sin incluir el origen ($/0$). El operador ρ es el encargado de la transformación del espacio Euclidiano al respectivo espacio proyectivo, se define como

$$\rho : A \in \mathbb{E}^n \rightarrow A + e_{n+1} \in \mathbb{P}\mathbb{E}^n \quad (\text{A.20})$$

En otros términos, el espacio Euclidiano \mathbb{E}^n está embebido como un *hiperplano* en $\rho(\mathbb{E}^n)$. Es importante resaltar que el origen Euclidiano no será más un punto singular¹⁵ en dicho espacio ($0 + e_{n+1}$). En la figura A.10 se muestra el caso del espacio $\rho(\mathbb{E}^2)$ dado que dicho espacio proyectivo está en tres dimensiones por su componente homogénea aquí el plano $\rho(\mathbb{E}^2)$ representa todo el espacio Euclidiano bidimensional, en el caso de los espacios Euclidianos de más de dos dimensiones se tendrá el mismo concepto (sólo que ahora dichos planos se convertirán en *hiperplanos*). La transformación inversa ρ^{-1} del espacio proyectivo a su correspondiente espacio Euclidiano se define como

$$\rho^{-1} : A \in \mathbb{P}\mathbb{E} \rightarrow \frac{1}{A \cdot e_{n+1}} \sum_{i=1}^n (A \cdot e_i) e_i \in \mathbb{E}^n, \quad (\text{A.21})$$

es posible ver que la transformación del espacio proyectivo hacia el espacio Euclidiano será válida para todo aquel punto cuya componente homogénea (e_{n+1}) sea diferente de cero. En contraparte aquellos puntos cuya componente homogénea sea cero son transformados al punto en el infinito (*Point at Infinity*),¹⁶. El proceso de mapeo inverso del espacio proyectivo al Euclidiano consiste en normalizar todas las componentes respecto a la componente homogénea, de esa manera la componente homogénea será unitaria, así los primeros n componentes representa el *vector* en el espacio Euclidiano. Todo el

¹⁵ Punto singular es aquel cuyo producto interior con cualquier otro *vector* siempre es 0.

¹⁶ Estos puntos son también llamados *vectores* de dirección.

espacio proyectivo excepto el *hiperplano* $e_{n+1} = 0$ es mapeado inversamente a el espacio Euclidiano, por ello, el proceso de escalamiento isotrópico de un punto proyectivo no tiene ningún efecto en su interpretación Euclidiana, es decir, si un *vector* $a \in \mathbb{E}^n$ y un factor de escalamiento $\alpha \in \mathbb{R}$ entonces $\rho^{-1}(\alpha\rho(a)) = a$ (ver la figura A.10).

A.9.1. Álgebra Geométrica en espacio Proyectivo

Asociar una interpretación geométrica mediante NO o NI implica que los *Blades* representan subespacios lineales, un *bivector* representará un subespacio bidimensional, si este *2-Blade* es un subespacio del espacio proyectivo sucede que dicho *bivector* representará un subespacio distinto en su correspondiente espacio Euclidiano embebido. Por lo tanto la forma en que un *k-Blade* $A_{(k)} \in \mathbb{PE}^n$ se representa en su respectivo \mathbb{E}^n tendrá que ser analizada en relación del grado del *Blade*, la interpretación NO o NI deseada y la transformación ρ .

A.9.2. Espacios Nulos en Espacio Proyectivo \mathbb{PE}^n

Debido a que la interpretación NO o NI aplica sobre el espacio inmediato y no al espacio embebido (el espacio inmediato en este caso es el proyectivo y el embebido es el Euclidiano), es necesario un operador que aplique el concepto de espacio nulo al espacio embebido. El espacio exterior nulo Euclidiano $\text{NO}_{\mathbb{E}}$ se define como

$$\text{NO}_{\mathbb{E}}(A_{(k)} \in \mathbb{PE}^n) := \{x \in \mathbb{E}^n : \rho(x) \wedge A_{(k)} = 0 \in \mathbb{PE}^n\}, \quad (\text{A.22})$$

similarmente el espacio interior nulo Euclidiano como $\text{NI}_{\mathbb{E}}$ de la forma

$$\text{NO}_{\mathbb{E}}(A_{(k)} \in \mathbb{PE}^n) := \{x \in \mathbb{E}^n : \rho(x) \cdot A_{(k)} = 0 \in \mathbb{PE}^n\}. \quad (\text{A.23})$$

La interpretación geométrica del espacio exterior nulo Euclidiano $\text{NO}_{\mathbb{E}}$ es de gran utilidad pues permite representar entidades como puntos, líneas y planos en posición general¹⁷ en el espacio Euclidiano. Sea $a \in \mathbb{E}^n$ cuya representación homogénea será $A = \rho(a) \in \mathbb{P}\mathbb{E}^n$ entonces el $\text{NO}(A) = \{\alpha A \in \mathbb{P}\mathbb{E}^n : \alpha \in \mathbb{R}/0\}$ representará una línea en dicho espacio que pasa por el origen (ver A.10), es decir,

$$\begin{aligned} \text{NO}_{\mathbb{E}}(A) &= \rho^{-1}(\text{NO}(A)) \\ &= \left\{ \frac{1}{(\alpha A) \cdot e_{n+1}} \sum_{i=1}^n ((\alpha A) \cdot e_i) e_i \right\} \\ &= a, \end{aligned}$$

es por ello que $\text{NO}(A \in \mathbb{P}\mathbb{E}^n)$ es una línea, pero en $\text{NO}_{\mathbb{E}}$ es un punto a , es decir, la representación de una entidad como el punto (subespacio cero dimensional). La representación de un *bivector* en el espacio proyectivo respecto al espacio Euclidiano es que mediante el *2-Blade* se representa un subespacio bidimensional que al ser mapeado al espacio Euclidiano mediante la transformación inversa (ρ^{-1}) genera un subespacio unidimensional (línea) (ver la figura A.11). Sea a y $b \in \mathbb{E}^n$ cuya representación homogénea es $A = \rho(a)$ y $B = \rho(b) \in \mathbb{P}\mathbb{E}^n$, entonces el $\text{NO}(A \wedge B)$ representará un plano que pasa por el origen y los puntos homogéneos A y B . Al obtener su interpretación en el espacio Euclidean se tiene

$$\begin{aligned} \text{NO}_{\mathbb{E}}(A \wedge B) &= \rho^{-1}(\text{NO}(A \wedge B)) \\ &= \rho^{-1}\{\alpha A + \beta B \in \mathbb{P}\mathbb{E}^n : (\alpha, \beta) \in \mathbb{R}^2/0\} \\ &= \left\{ \frac{1}{(\alpha A + \beta B) \cdot e_{n+1}} \sum_{i=1}^n ((\alpha A + \beta B) \cdot e_i) e_i \right\} \\ &= \left\{ \frac{\alpha}{(\alpha + \beta)} a + \frac{\beta}{(\alpha + \beta)} b \right\}, \end{aligned}$$

lo cual representa una línea que pasa por los puntos a y b (no necesariamente por el origen). Esto muestra como se puede tratar una línea general en álgebra geométrica

¹⁷ Posición general en este contexto se entiende que la entidad geométrica puede estar localizada y orientada en cualquier posición y dirección.

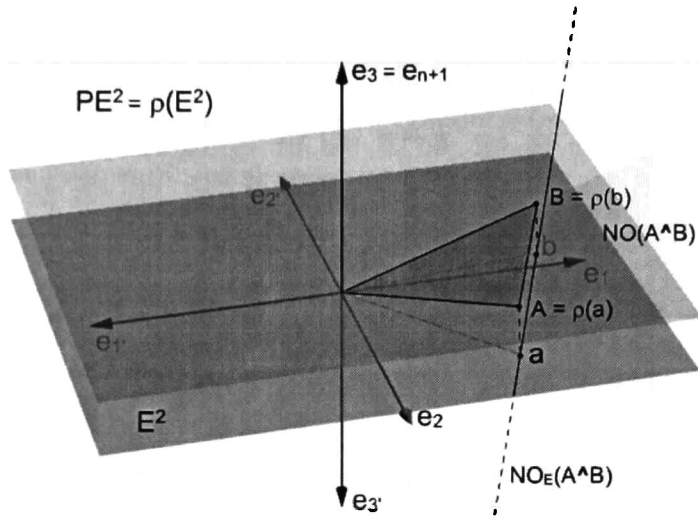


Figura A.11: Línea representada mediante el *bivector* $NO_E(A \wedge B)$.

proyectiva haciendo el producto exterior de los puntos (ver la figura A.11). De igual modo es posible verificar que en el caso del *trivector* formado por los puntos a , b y $c \in \mathbb{E}^n$, se tendrá que su interpretación en el $NO_E(A \wedge B \wedge C)$ es el plano que pasa por los puntos a , b y c . Esto se explica dado que el volumen representado por el producto exterior de los *vectores* homogéneos $A = \rho(a)$, $B = \rho(b)$ y $C = \rho(c)$ al obtener su interpretación geométrica mediante ρ^{-1} al espacio Euclidiano, es equivalente a obtener la intersección de dicho *hipervolumen* con el *hiperplano* e_{n+1} . Al transformar el espacio $\rho(E^3) = \mathbb{P}E^3$ que equivalente a decir $\rho(\mathbb{R}^3) \cong \mathbb{R}^4/0$ (por lo tanto es imposible graficarlo) pero es a análogo a la figura A.10. La dualidad permite interpretaciones en el espacio interior nulo Euclidiano NI_E de los *Blades* tal y como se presentó para el caso del NO_E . Generalizando se tiene:

Punto :	$\text{NO}_{\mathbb{E}}(A)$
Línea :	$\text{NO}_{\mathbb{E}}(A \wedge B)$
Plano :	$\text{NO}_{\mathbb{E}}(A \wedge B \wedge C)$
Plano :	$\text{NI}_{\mathbb{E}}(A)$
Línea :	$\text{NI}_{\mathbb{E}}(A \wedge B)$
Punto :	$\text{NI}_{\mathbb{E}}(A \wedge B \wedge C)$

A.9.3. Reflexiones en Espacio Proyectivo

En el espacio proyectivo se pueden realizar reflexiones al igual que en cualquier otro espacio, sin embargo dado que se agregó una coordenada extra (la componente homogénea) el efecto de la reflexión puede cambiar el concepto representado en el espacio Euclidiano. Lo anterior puede ser ajustado para mantener la misma representación geométrica. Primero se define un operador ψ que proyecta los puntos homogéneos al plano afino¹⁸ es decir,

$$\psi : A \in \mathbb{PE}^n \rightarrow \frac{A}{A \cdot e_{n+1}} \in \mathbb{PE}^n \quad (\text{A.24})$$

Sea $a \in \mathbb{E}^2$ se tiene que $A = \rho(a) = a + e_3 \in \mathbb{PE}^2$, entonces al reflejar A con respecto a e_3 se tendrá que $B = e_3 A e_3 = e_3(a + e_3)e_3 = e_3 a e_3 + e_3 e_3 e_3 = -a + e_3$, implica una reflexión respecto al origen. Ahora sucede que cuando se refleja con respecto a un *vector* dirección¹⁹ unitario (aquellos sin componente homogénea), es decir, sea $m \in \mathbb{PE}^2$ tal que $|m| = 1$ (manteniendo la definición de A), se tiene que $C = m A m = m(a + e_3)m = m a m + m e_3 m = m a m - e_3$, es aquí el uso del operador ψ , puesto que $\psi(m A m) = -(m a m) + e_3 = -(a_\alpha - a_\beta) + e_3 = a_\beta - a_\alpha + e_3$ (donde a_α y a_β son las partes paralelas y perpendiculares de a respecto a m). El resultado anterior muestra

¹⁸ El plano con componente homogénea unitaria.

¹⁹ *Point at Infinity*.

que la reflexión no arrojó el resultado deseado, es decir, reflejó la parte paralela y no la perpendicular. La solución es reflejar previamente con respecto al origen (reflejar con e_3) y posteriormente reflejarlo con m , es decir, $m(e_3 A e_3)m = (m e_3)A(e_3 m) = -mam - e_3$, en el plano afino esto es $\psi(-mam - e_3) = mam + e_3$.

A.9.4. Rotaciones en Espacio Proyectivo

Una rotación se realiza mediante un Rotor, que es la aplicación de dos reflexiones. Sea $a \in \mathbb{E}^2$ de modo que $A = \rho(a) = a + e_3 \in \mathbb{P}\mathbb{E}^2$, dos *vectores* de dirección n y m , entonces como se observó en la sección anterior es posible realizar una reflexión multiplicando con $n e_3$ y $m e_3$ respectivamente, es decir, $B = (m e_3)A(e_3 m)$ y posteriormente $C = (n e_3)B(e_3 n)$ lo cual es equivalente a decir, $C = (n e_3)(m e_3)A(e_3 m)(e_3 n)$ que es análogo a la forma usual de rotación mediante un rotor, es decir, $C = R a \tilde{R}$, $R = (n e_3)(m e_3) = -nm$ $\tilde{R} = (e_3 m)(e_3 n) = -mn$, $C = R A \tilde{R} = (-nm)A(-mn) = (nm)A(mn)$, así el rotor en espacio proyectivo tendrá la misma forma que en espacio Euclidiano (ver la figura A.9).

A.9.5. Modelo de una Cámara

La visión computacional hace uso de cámaras para adquirir imágenes de objetos o escenas con la finalidad de interpretar la información ahí contenida para diversos fines, por consiguiente la cámara se modela matemáticamente como una entidad cuya función principal es convertir los elementos del espacio visual tridimensional (escena) a un espacio bidimensional (imagen). Existen diferentes modelos matemáticos, en este trabajo se usará el modelo llamado *Pin-hole*, donde los puntos en el espacio se proyectan en un plano (ver la figura A.13). En dicho modelo existen puntos destacados, el centro

de la cámara $c \in \mathbb{E}^3$, el punto principal $p \in \mathbb{E}^3$ que definen el eje principal. Dicho eje corresponde a la normal del plano de la imagen. La distancia entre el centro de la cámara c y el punto principal p es conocida como la distancia focal $f \in \mathbb{R}$, la apertura horizontal de la cámara, es decir, el ángulo o campo de visión horizontal se denota como $\theta \in \mathbb{R}$, mientras que la apertura vertical de la cámara o ángulo de visión vertical se denotará como $\zeta \in \mathbb{R}$, finalmente es importante conocer cual de las cuatro aristas del rectángulo de la imagen representa la parte superior de la imagen, esto se puede denotar con un punto u (ver la figura A.12). Ahora en términos de álgebra geométrica

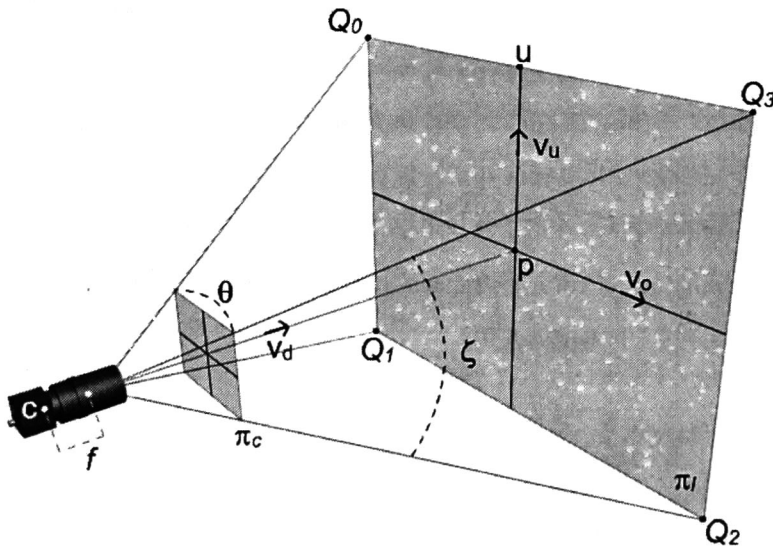


Figura A.12: Modelo de una cámara, π_c y π_l denotan los planos que delimitan el campo visual, esto se debe a que fuera de estos planos no sera posible adquirir una imagen con el enfoque correcto.

proyectiva se puede expresar el centro de la cámara como un punto homogéneo de la forma $C = \rho(c) \in \mathbb{PE}^3$, de igual forma el punto principal y la orientación del rectángulo de la imagen respectivamente $P = \rho(p) \in \mathbb{PE}^3$, $U = \rho(u) \in \mathbb{PE}^3$, se puede calcular los

puntos Q_i en los vértices de la imagen como (ver la figura A.13)

$$v_d = \frac{p - c}{|p - c|}, \quad v_u = \frac{u - p}{|u - p|}, \quad v_o = (v_d \wedge v_u)^* \quad (\text{A.25})$$

$$h = f \tan(\theta), \quad w = f \tan(\zeta), \quad (\text{A.26})$$

$$Q_i = \rho \left(p + (-1)^{i(i+1)/2} w v_o + (-1)^{i(i+3)/2} h v_u \right) \quad i \in \{0, 1, 2, 3\}. \quad (\text{A.27})$$

De este modo el plano de la imagen se puede modelar en su interpretación geométrica

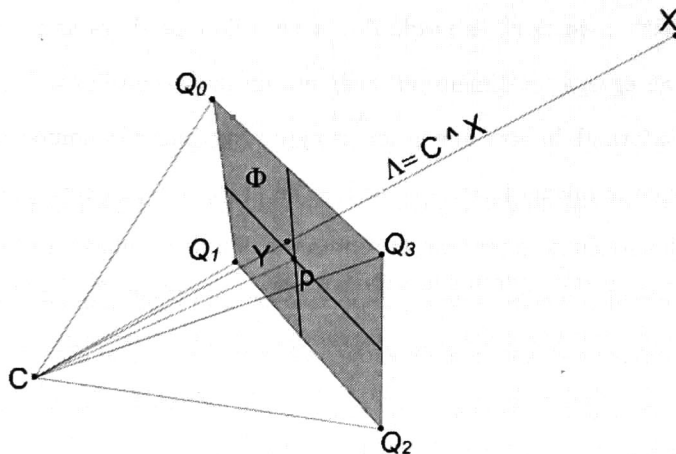


Figura A.13: Intersección del rayo óptico Λ formado por el centro de la cámara C y el punto en el espacio X con el plano de la imagen.

(NO) mediante el producto exterior de tres de los puntos Q_i del rectángulo de la imagen, es decir, $\Phi = Q_1 \wedge Q_2 \wedge Q_3$. El efecto de adquirir una imagen mediante la cámara es equivalente al mapeo de dicho punto en la escena X , al subespacio de la imagen, se modela mediante la intersección del rayo óptico que pasa por el punto X y el centro de la cámara C con el plano de la imagen, dicho rayo óptico puede ser modelado mediante la línea que pasa por dichos puntos, es decir, $\Lambda = X \wedge C$. La intersección de dicho rayo con el plano de la imagen generará el punto Y correspondiente a la proyección de X en

la imagen, es decir, $Y = \Phi \vee \Lambda = (Q_1 \wedge Q_2 \wedge Q_3) \vee (X \wedge C)$.

A.10. Espacio Conformal

Los conceptos subyacentes al álgebra geométrica sobre el espacio conformal homogéneo son similares a los del espacio proyectivo y a la transformación conformal, es decir, la idea de envolver el espacio Euclidiano ($\mathbb{E}^n \cong \mathbb{R}^n$) en un espacio con más dimensiones ($\mathbb{R}^m : m > n$) es usada de nuevo, sólo que ahora se tendrán dos dimensiones extras. Se agrega un dimensión mediante un mapeo no lineal²⁰ produciendo así el espacio conformal, la otra dimensión se agrega mediante la homogenización del resultado anterior en la misma forma que se hizo en el espacio proyectivo. En dicho espacio conformal homogéneo se tendrán la representación de entidades tales como puntos, líneas, planos, círculos y esferas en su respectivo espacio Euclidiano embebido. Las transformaciones conformales son combinaciones de inversiones, con dichas transformaciones serán posible realizar rotaciones, translaciones, dilataciones, etcétera. en el espacio Euclidiano. El primer paso para envolver el espacio Euclidiano \mathbb{E}^n en el espacio Conformal \mathbb{K}^n consiste en realizar una proyección estereográfica mediante la función de mapeo κ definida por

$$\kappa : x \in \mathbb{E}^n \rightarrow \frac{2}{x^2 + 1}x + \frac{x^2 - 1}{x^2 + 1}e_+, \quad (\text{A.28})$$

donde se denota²¹ a $e_+ \equiv e_{n+1}$. Todos los puntos envueltos se encuentran en una *hiperesfera* de radio unitario, por lo tanto $|\kappa(x \in \mathbb{E}^n)| = 1$. La figura A.14 ilustra el envolvimiento de \mathbb{E}^1 , donde el punto $e_+ \in \mathbb{K}^1$ representa $\pm\infty$ y $-e_+$ representa el origen en \mathbb{E}^1 . La figura A.15 muestra el mismo concepto de proyección estereográfica ahora

²⁰ Proyección estereográfica.

²¹ La base e_+ se denota así dado que posee una cuadratura a 1, posteriormente se definirá otra base e_- cuya cuadratura es -1 .

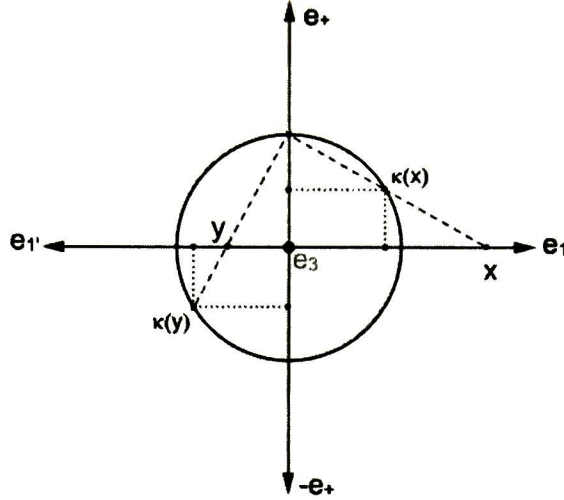


Figura A.14: Proyección estereográfica de los puntos x e y en la hipersfera unitaria.

con líneas y círculos de \mathbb{E}^2 a $\kappa(\mathbb{E}^2)$, esto muestra que ambas ecuaciones se mapean al espacio conformal como círculos, con la diferencia que la línea se mapea como un círculo que pasa por e_+ (la línea se extiende hasta el infinito, o bien la línea es un segmento de un círculo de radio infinito), a su vez el círculo se proyecta estereográficamente como otro círculo que no pasa por e_+ . El operador de envolvimiento conformal κ mapea todo el espacio Euclidiano \mathbb{E}^n a un subespacio n dimensional de \mathbb{K}^n , la transformación inversa κ^{-1} de \mathbb{K}^n a \mathbb{E}^n sólo está definida para aquellos puntos que se encuentren contenidos en dicho subespacio (*Horósfera*). Los puntos Euclidianos x_i transformados al espacio conformal tienen magnitud unitaria $|\kappa(x \in \mathbb{E}^n)| = 1$ (puntos en la *hipersfera*), entonces la transformación inversa está definida sólo para aquellos puntos conformales $y \in \mathbb{K}^n : |y| = 1$, es decir,

$$\kappa^{-1} : y \in \mathbb{K}^n, |y| = 1 \rightarrow \frac{1}{1 - y \cdot e_+} \sum_{i=1}^n (y \cdot e_i) e_i. \quad (\text{A.29})$$

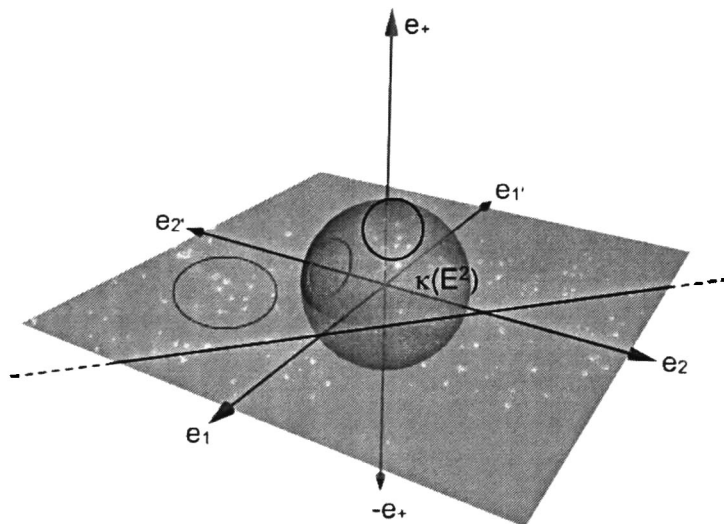


Figura A.15: Proyección estereográfica del círculo y la línea a la hipersfera unitaria.

La homogenización del espacio conformal \mathbb{K}^n se efectúa de la misma forma que el espacio proyectivo, se denota como \mathbb{PK}^n y se representa mediante $\mathbb{R}^{n+1,1}/0$, dicho espacio posee $n + 2$ bases ortonormales de las cuales $n + 1$ tienen una cuadratura igual 1 y la base homogénea²² -1 . El objetivo de estas bases es construir *vectores* nulos (aquellos cuya cuadratura es 0), serán de gran utilidad en interpretaciones geométricas. El operador para realizar este proceso de homogenización del espacio conformal se denota como ρ el cual se define como

$$\rho : x \in \mathbb{K}^n \rightarrow \mathbf{X} = x + e_-, \tag{A.30}$$

donde e_- representa la dimensión homogénea. La figura A.16 se muestra la sucesión de pasos para el envolvimiento desde \mathbb{E}^1 hasta \mathbb{PK}^1 Como consecuencia de homogeneizar los puntos conformales (que representan puntos Euclidianos) la cuadratura de dichos puntos conformales es 0, es decir,

²² e_-

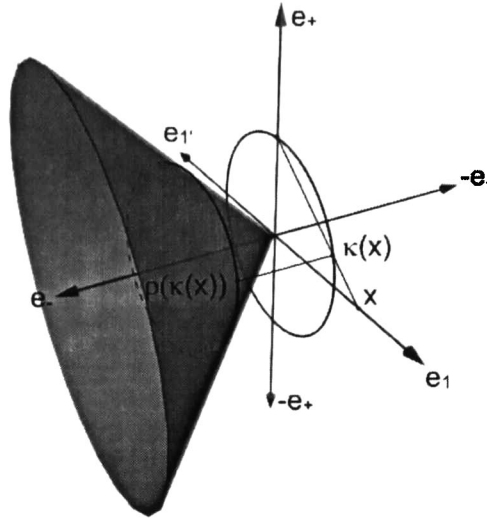


Figura A.16: Proceso de envolvimiento de un punto $x \in \mathbb{E}^1$, primero proyección estereográfica $\kappa(x)$ y posteriormente homogenización $\rho(\kappa(x))$.

$$\begin{aligned}
 & (\alpha(\rho(\kappa(x \in \mathbb{E}^n))))^2 \alpha \in \mathbb{R}^1/0 \\
 & = \alpha^2(\kappa(x \in \mathbb{E}^n) + e_-)^2 \\
 & = \alpha^2(\kappa(x \in \mathbb{E}^n) + e_-) \cdot (\kappa(x \in \mathbb{E}^n) + e_-) \\
 & = \alpha^2(\kappa(x \in \mathbb{E}^n)^2 + e_-^2) \\
 & = \alpha^2(1 - 1) \\
 & = 0,
 \end{aligned}$$

lo cual implica que todos los puntos Euclidianos que están embebidos en el espacio conformal proyectivo tienen cuadratura igual a cero, es decir, son *vectores nulos*, sin importar el escalamiento isotrópico que puedan sufrir en el espacio \mathbb{PK}^n debido a un factor de escalamiento α . En el caso del espacio Euclidiano unidimensional \mathbb{E}^1 (*vectores nulos*) está contenido en el subespacio representado por el cono \mathbb{H}^n de la figura A.17, tal

subespacio²³ (cono nulo) de \mathbb{PK}^n se denota como $\mathbb{H}^n \subset \mathbb{PK}^n$, es decir,

$$\mathbb{H}^n := \{\mathbf{X} \in \mathbb{PK}^n : \mathbf{X}^2 = 0\} = \{\alpha(\rho(\kappa(x \in \mathbb{E}^n))) : \alpha \in \mathbb{R}^1/0\}. \quad (\text{A.31})$$

La transformación inversa ρ^{-1} desde el espacio conformal homogéneo al espacio conformal, está definida para aquellos elementos cuya componente homogénea sea diferente de cero ($\mathbf{X} \cdot e_- \neq 0$), es decir,

$$\rho^{-1} : \mathbf{X} \in \mathbb{PK}^n \rightarrow \frac{1}{\mathbf{X} \cdot e_-} \sum_{i=1}^{n+1} (\mathbf{X} \cdot e_i) e_i \in \mathbb{K}^n \quad (\text{A.32})$$

Un *vector* euclideo $x \in \mathbb{E}^n$ se lleva al espacio conformal homogéneo de la forma $\mathbf{X} = \rho(\kappa(x)) = \frac{2}{x^2+1}x + \frac{x^2-1}{x^2+1}e_+ + e_-$. Se sabe que un escalamiento isotrópico en el espacio proyectivo no afecta la representación en el espacio Euclideo, entonces se puede escalar la expresión sin alterar su representación en Euclideo, en particular si $\frac{1}{2}(x^2 + 1) \neq 0$,

$$\begin{aligned} \frac{1}{2}(x^2 + 1)\rho(\kappa(x)) &= (x^2 + 1)\left\{\frac{2}{x^2+1}x + \frac{x^2-1}{x^2+1}e_+ + e_-\right\} \\ &= x + \frac{1}{2}x^2(e_+ + e_-) + \frac{1}{2}(e_+ - e_-) \end{aligned}$$

de esta manera se puede simplificar el proceso de envolvimiento. Se observa que $(e_+ + e_-)$ representará el infinito en el espacio Euclideo dado que $\kappa^{-1}(\rho^{-1}(e_+ + e_-)) = \infty$, por lo que se define el *vector* $e_\infty := (e_+ + e_-)$, similarmente se puede observar que $\frac{1}{2}(e_+ - e_-)$ representa el origen Euclideo dado que $\kappa^{-1}(\rho^{-1}(\frac{1}{2}(e_+ - e_-))) = 0$, es por ello que se define el *vector* e_0 . Así mismo se usará la definición de $E = e_\infty \wedge e_0$ para secciones posteriores. Los *vectores* definidos anteriormente (que sirven para representar el infinito e_∞ y el origen e_0) poseen algunas características importantes, tales como $e_\infty^2 = e_0^2 = 0$, y también $e_\infty \cdot e_0 = -1$. Estas propiedades serán de gran utilidad posteriormente en el álgebra geométrica. El proceso de envolvimiento se puede reescribir como $(x^2 +$

²³ Llamado horófera.

$\rho(\kappa(x)) = x + \frac{1}{2}x^2 e_\infty + e_0$. Un operador que ejecute la proyección estereográfica, homogenización y escalamiento haciendo uso de los conceptos anteriores, se representa con λ de la forma

$$\lambda : x \in \mathbb{E}^n \rightarrow \frac{1}{2}(x^2 + 1)\rho(\kappa(x)) \in \mathbb{PK}^n, \tag{A.33}$$

la figura A.17 ilustra este proceso de $x \in \mathbb{E}^1$ hasta $\mathbf{X} \in \mathbb{PK}^1$ donde su componente e_0 es unitaria, todos los puntos Euclidianos al ser llevados al espacio conformal homogéneo están contenidos en el subespacio limitado por el cono nulo y el *hiperplano* definido por la base *vectorial* e_0 , es decir, en una parábola denominada \mathbb{H}_a^1 . La función inversa λ^{-1} estaría definida sólo para aquellos puntos que se encuentren en el subespacio del cono nulo (\mathbb{H}^n), es decir,

$$\lambda^{-1} : \mathbf{X} \in \mathbb{H}^n \rightarrow \kappa^{-1}(\rho^{-1}(\mathbf{X})) \in \mathbb{E}^n \tag{A.34}$$

El conjunto de puntos conformales que se encuentran contenidos en el subespacio gener-

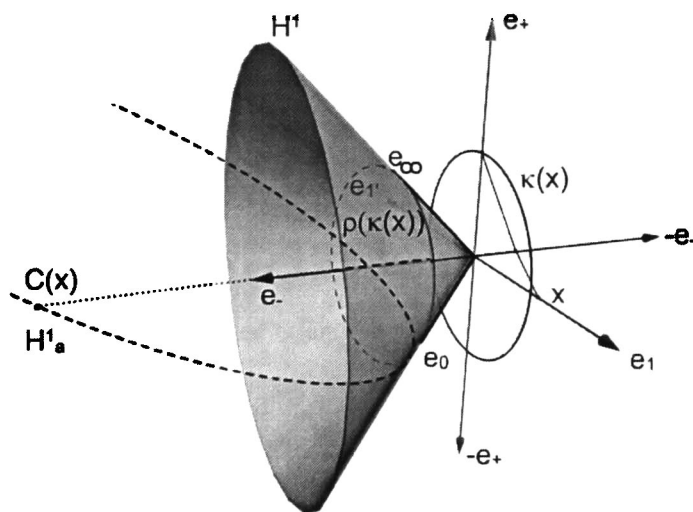


Figura A.17: Puntos conformales con la componente e_0 normalizada en el hiperplano afino \mathbb{H}_a^1 .

ado por la intersección del cono nulo y el *hiperplano* e_0 llamados \mathbb{H}_a son de gran utilidad

en las interpretaciones geométricas. Estos puntos poseen componentes e_0 unitarias similares a los puntos en el espacio proyectivo con componente homogénea unitaria (e_{n+1}) que yacen en el *hiperplano afino*, por ello se le denota a \mathbb{H}_a como el cono nulo afino, haciendo uso de las propiedades anteriores para los *vectores* nulos se reescribe como $\mathbb{H}_a^n := \{\mathbf{X} \in \mathbb{H}^n \subset \mathbb{PK}^n : \mathbf{X} \cdot e_\infty = -1\}$.

A.10.1. Álgebra Geométrica en Espacio Conformal

El espacio conformal homogéneo \mathbb{PK}^n que envuelve el espacio Euclidiano n -dimensional pose $n+2$ dimensiones, entonces el álgebra geométrica de dicho espacio consiste de $2^{(n+2)}$ bases *Blade*. Dichos *Blades* representan subespacios lineales en su interpretación \mathbb{NO} y \mathbb{NI} en \mathbb{PK}^n . El interés de envolver el espacio Euclidiano tiene el propósito de generar propiedades útiles en la interpretación de los espacios nulos Euclidianos, es decir, el interés está en el conjunto de *vectores* Euclidianos que al ser envueltos en el espacio conformal homogéneo se encuentran en el \mathbb{NO} o \mathbb{NI} de algún *Blade* de \mathbb{PK}^n . Sea $\mathbf{A} \in \mathbb{PK}^n$ el espacio interior nulo y el espacio exterior nulo de \mathbf{A} se definen como

$$\mathbb{NI}_E(\mathbf{A} \in \mathbb{PK}^n) := \{x \in \mathbb{E}^n : \lambda(x) \cdot \mathbf{A} = 0\} \quad (\text{A.35})$$

$$= \lambda^{-1}(\{\mathbf{X} \in \mathbb{H}^n : \mathbf{X} \cdot \mathbf{A} = 0\}), \quad (\text{A.36})$$

$$\mathbb{NO}_E(\mathbf{A} \in \mathbb{PK}^n) := \{x \in \mathbb{E}^n : \lambda(x) \wedge \mathbf{A} = 0\} \quad (\text{A.37})$$

$$= \lambda^{-1}(\{\mathbf{X} \in \mathbb{H}^n : \mathbf{X} \wedge \mathbf{A} = 0\}), \quad (\text{A.38})$$

la segunda definición de estos conjuntos es evidente dado que el conjunto de puntos en el cono nulo \mathbb{H}^n pueden ser proyectados al respectivo espacio Euclidiano, de esta manera el producto interior o exterior con dicho punto conformal será equivalente a tener puntos Euclidianos que cumplan con la primera definición del conjunto. Los espacios nulos Euclidianos de los *Blades* serán la intersección de los espacios nulos en \mathbb{PK}^n con el cono

nulo \mathbb{H}^n , es decir, si

$$\begin{aligned} \text{NI}(\mathbf{A} \in \mathbb{PK}^n) &:= \{\mathbf{X} \in \mathbb{PK}^n : \mathbf{X} \wedge \mathbf{A} = 0\}, \\ \text{NO}(\mathbf{A} \in \mathbb{PK}^n) &:= \{\mathbf{X} \in \mathbb{PK}^n : \mathbf{X} \wedge \mathbf{A} = 0\} \end{aligned}$$

, entonces

$$\begin{aligned} \text{NI}(\mathbf{A} \in \mathbb{PK}^n) &:= \text{NI}_E(\mathbf{A} \in \mathbb{PK}^n) := \lambda^{-1}(\{\text{NI}(\mathbf{A}) \cap \mathbb{H}^n\}), \\ &:= \lambda^{-1}(\{\text{NO}(\mathbf{A}) \cap \mathbb{H}^n\}). \end{aligned}$$

El NO del *bivector* formado por $\mathbf{X} \wedge \mathbf{Y} : \mathbf{X}, \mathbf{Y} \in \mathbb{PK}^1$ expande un subespacio bidimensional $\text{NO}(\mathbf{X} \wedge \mathbf{Y})$ y como el espacio exterior nulo Euclidiano son aquellos puntos contenidos en la intersección del espacio exterior conformal con el cono nulo, entonces el $\text{NO}_E(\mathbf{X} \wedge \mathbf{Y}) = \{\lambda^{-1}(\mathbf{X}), \lambda^{-1}(\mathbf{Y})\}$ ²⁴

A.10.2. Primitivas Geométricas

La presente tesis se enfoca en visión estereoscópica usando álgebra geométrica conformal y puesto que los escenarios u objetos existen en el espacio tridimensional, entonces se presenta la descripción de las entidades geométricas básicas en \mathbb{E}^3 haciendo uso del envolvimiento conformal homogéneo \mathbb{PK}^3 e interpretándolo mediante los espacios nulos correspondientes. Antes de iniciar con estas entidades es crucial visualizar una propiedad importante de los puntos conformales. Sea x y $y \in \mathbb{E}^3$ dos puntos Euclidianos, al envolverlos se obtiene que $\mathbf{X} = x + \frac{1}{2}x^2e_\infty + e_0$, $\mathbf{Y} = y + \frac{1}{2}y^2e_\infty + e_0 \in \mathbb{H}^3$. Ahora el producto interior del par de puntos Conformales²⁵ será de la forma

²⁴ Entidad par de puntos.

²⁵ Por brevedad se le denotará como puntos Conformales a los puntos Conformales homogéneos.

$$\begin{aligned}
 \mathbf{X} \cdot \mathbf{Y} &= (x + \frac{1}{2}x^2e_\infty + e_0) \cdot (y + \frac{1}{2}y^2e_\infty + e_0) \\
 &= x \cdot y - \frac{1}{2}x^2 - \frac{1}{2}y^2 \\
 &= -\frac{1}{2}(x - y)^2
 \end{aligned}$$

lo que proporciona una característica muy útil, es decir, el producto interior de dos puntos Conformales proporciona una medida de la distancia Euclidiana entre ambos puntos, es decir, $|x - y| = \sqrt{-2(\mathbf{X} \cdot \mathbf{Y})}$.

A.10.3. Puntos

Sea $x \in \mathbb{E}^3$ un punto en el espacio Euclidiano, entonces se tendrá que $\mathbf{X} = x + \frac{1}{2}x^2e_\infty + e_0 \in \mathbb{H}^3$, entonces su $\text{NI}(\mathbf{X}) = \{\mathbf{X} \cdot \mathbf{Y} = 0\}$, pero como $\mathbf{X} \in \mathbb{H}^3$ entonces \mathbf{X} es un punto sobre el cono nulo, por ende $\text{NI}(\mathbf{X}) = \{\mathbf{X} \cdot \alpha\mathbf{X} = 0 : \alpha \in \mathbb{R} \setminus \{0\}\}$. Esto implica que $\text{NI}_E(\mathbf{X}) = \lambda^{-1}(\{\text{NI}(\mathbf{X})\})$, se obtiene que la entidad (0-dimensional) geométrica *punto* se representa como un subespacio unidimensional $\alpha\mathbf{X}$ en \mathbb{PK}^3 .

A.10.4. Esferas

Las esferas están relacionadas intrínsecamente con el álgebra geométrica conformal, es por ello que cualquier punto conformal que este fuera del cono nulo representará una esfera, ya sea de radio real o imaginario. Sea $\mathbf{A} \in \mathbb{H}_a^n$ un punto conformal en el cono nulo afino y $\mathbf{S} \in \mathbb{PK}^n$ otro punto conformal fuera del cono nulo de la forma $\mathbf{S} = \mathbf{A} - \frac{1}{2}\epsilon^2e_\infty$, $\epsilon \in \mathbb{R}$, entonces el $\text{NO}_E(\mathbf{S}) = \{\mathbf{X} \in \mathbb{H}^n : \mathbf{S} \cdot \mathbf{X} = 0\}$, es decir

$$\begin{aligned}
 \mathbf{S} \cdot \mathbf{X} &= 0 \\
 (\mathbf{A} - \frac{1}{2}\epsilon^2 e_\infty) \cdot \mathbf{X} &= 0 \\
 \mathbf{A} \cdot \mathbf{X} - \frac{1}{2}\epsilon^2 e_\infty \cdot \mathbf{X} &= 0 \\
 -\frac{1}{2}(a - x)^2 + \frac{1}{2}\epsilon^2 &= 0 \\
 (a - x)^2 &= \epsilon^2
 \end{aligned}$$

lo que implica que en el espacio Euclidiano una entidad que posee un punto singular de referencia a y que cualquier otro punto x debe de encontrarse a una distancia fija de ϵ unidades para pertenecer a ese espacio interior nulo. Tal entidad es la esfera, el punto a es su centro, mientras que cualquier otro punto x a una distancia dirigida ϵ (radio) pertenecerá al subespacio que representa la superficie de la esfera. Nótese dos aspectos importantes, primero en el $\text{NO}_E(\mathbf{S})$ se hace uso intencional de $\mathbf{X} \in \mathbb{H}^n$ para tener un versión con componente homogénea unitaria (para evitar el paso de normalización), sin embargo con componente no normalizados el principio se sigue cumpliendo. El segundo aspecto es que \mathbf{S} al ser un elemento de un espacio homogéneo, cualquier escalamiento isotrópico no posee efecto alguno sobre el resultado en su interpretación en el espacio Euclidiano. Si se desea normalizar la componente homogénea de una esfera escalada arbitrariamente \mathbf{S} se tiene que $(\frac{\mathbf{S}}{-\mathbf{S} \cdot e_\infty})^2 = \epsilon^2$. Es posible saber si un punto \mathbf{X} se encuentra en el interior, exterior, o en la superficie de una esfera mediante

$$\begin{aligned}
 \frac{\mathbf{S} \cdot \mathbf{X}}{(\mathbf{S} \cdot e_\infty)(\mathbf{X} \cdot e_\infty)} &= 0 \quad \text{En la superficie de la esfera} \\
 &> 0 \quad \text{Dentro} \\
 &< 0 \quad \text{Fuera.}
 \end{aligned}$$

Ahora si $\mathbf{S} = \mathbf{A} + \frac{1}{2}\epsilon^2 e_\infty \in \mathbb{PK}^n$ es un punto en el interior del cono nulo, se puede verificar que dicho punto representará una esfera de radio imaginario. De esta forma se tiene que los puntos en el espacio proyectivo conformal son esferas de radio mayor a cero (los puntos fuera del cono nulo), esferas de radio cero (puntos contenidos en el

cono nulo), esferas de radio imaginario (puntos contenido en el interior del cono nulo). Es por eso que la entidad primitiva fundamental del espacio conformal homogéneo es la esfera y de ahí se derivan el resto de las entidades geométricas.

A.10.5. Planos

La esfera como entidad geométrica fundamental puede ser usada para aproximarse a un plano, es decir, una esfera con un radio infinito. Recordando que en el espacio proyectivo aquellos *vectores* de \mathbb{P}^n cuya componente homogénea es igual a cero (*vectores* de dirección) se mapean al espacio Euclidiano mediante ρ^{-1} a puntos en el infinito. Es plausible en el espacio conformal homogéneo convertir una esfera en un plano, es decir, sea la esfera normalizada $\mathbf{S} = \mathbf{A} - \frac{1}{2}\epsilon^2 e_\infty$ es posible eliminar su componente homogénea con sólo restarle e_0 , dado que

$$\begin{aligned}
 \mathbf{P} &= \mathbf{S} - e_0 \\
 &= \mathbf{A} - \frac{1}{2}\epsilon^2 e_\infty - e_0 \\
 &= a + \frac{1}{2}a^2 e_\infty + e_0 - \frac{1}{2}\epsilon^2 e_\infty - e_0 \\
 &= a + \frac{1}{2}(a^2 + \epsilon^2)e_\infty.
 \end{aligned}$$

El resultado es un plano \mathbf{P} en su interpretación $\mathbb{N}\mathbb{I}$, puesto que

$$\begin{aligned}
 \mathbf{P} \cdot \mathbf{X} &= 0 \in \mathbb{H}_a^3, \\
 &= (a + \frac{1}{2}(a^2 + \epsilon^2)e_\infty) \cdot (x + \frac{1}{2}x^2 e_\infty + e_0) = 0, \\
 &= a \cdot x - \frac{1}{2}(a^2 + \epsilon^2) = 0, \\
 &= |a||x_\alpha| - \frac{1}{2}(a^2 + \epsilon^2) = 0 : x = x_\alpha + x_\beta, \\
 &= 0 \Leftrightarrow |a||x_\alpha| = \frac{1}{2}(a^2 + \epsilon^2), \\
 &= 0 \Leftrightarrow |x_\alpha| = \frac{a^2 + \epsilon^2}{2|a|}.
 \end{aligned}$$

donde x_a representa la componente paralela de x a lo largo de a . La ecuacion A.10.5 ilustra que todos los puntos que posean una componente paralela a a de magnitud fija se encontrarán en el NI de \mathbf{P} . Esto es análogo al concepto del plano en cálculo vectorial, puesto que un plano se representa como un distancia de Hesse $d \in \mathbb{R}$ y un *vector* unitario normal $n \in \mathbb{R}^3 : |n| = 1$ de la forma que cualquier *vector* $u \in \mathbb{R}^3$ sobre el plano se tendra que $n \cdot u = d$ donde la componente paralela al lo largo de n tendría una magnitud fija d . Finalmente se puede visualizar a la entidad \mathbf{P} como una esfera de radio infinito que ρ^{-1} . En la sección A.5 se observó que en NI el producto exterior de dos *vectores* representa la intersección de sus respectivos NI, lo cual es posible realizar en el espacio conformal, es decir, el NI_E del producto exterior de dos esferas $\mathbf{Z} = \mathbf{S}_1 \wedge \mathbf{S}_2 \in \mathbb{PK}^3$, es la intersección de la esfera, lo cual puede ser un círculo, un punto o en el caso que no se intersecten será un círculo de radio imaginario. Es posible extender dicho producto exterior a tres esferas dando como resultado la intersección un círculo con una esfera que es un *par de puntos*, un sólo punto, o en su caso un par de puntos con un radio imaginario. En este sentido es posible conseguir un producto exterior de cuatro esferas, que en caso de existir la intersección de dichas entidades será un sólo punto. Esto aplica para las esferas que poseen un radio infinito o planos, como se observo anteriormente, en tal caso el círculo será de radio infinito, osea, una línea. El álgebra geométrica conformal hace uso del concepto de dualidad, para encontrar la representación de las entidades en términos de NO , es decir, sean $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W}, \mathbf{V}$ y $\mathbf{U} \in \mathbb{H}_d^e$ *vectores* linealmente independientes entre si tales que $\mathbf{X} = \rho(\kappa(x)), \mathbf{Y} = \rho(\kappa(y))$, etcétera. Entonces se tiene que

- $\text{NO}_E(\mathbf{V}) \rightarrow$ Punto v
- $\text{NO}_E(\mathbf{W} \wedge \mathbf{V}) \rightarrow$ Par de Puntos W y V
- $\text{NO}_E(\mathbf{W} \wedge \mathbf{V} \wedge e_\infty) \rightarrow$ Línea que pasa por W y V
- $\text{NO}_E(\mathbf{W} \wedge \mathbf{V} \wedge \mathbf{Z}) \rightarrow$ Circulo que pasa por W , V y Z
- $\text{NO}_E(\mathbf{W} \wedge \mathbf{V} \wedge \mathbf{Z} \wedge \mathbf{X}) \rightarrow$ Esfera que pasa por W , V , Z y X
- $\text{NO}_E(\mathbf{W} \wedge \mathbf{V} \wedge \mathbf{Z} \wedge e_\infty) \rightarrow$ Plano que pasa por W , V y Z
- $\text{NO}_E(\mathbf{W} \wedge \mathbf{V} \wedge \mathbf{Z} \wedge \mathbf{X} \wedge \mathbf{Y}) \rightarrow$ Todo el espacio Conformal.

A.10.6. Inversion en Espacio Conformal

En el espacio conformal es posible realizar trasformaciones llamadas inversiones, las combinaciones de estas inversiones implican trasformaciones en el espacio Euclidiano envuelto, tales inversiones son herramientas para transformar las entidades en su interpretación en los respectivos espacios nulos. Sea $x \in \mathbb{E}^n$ un punto Euclidiano y el

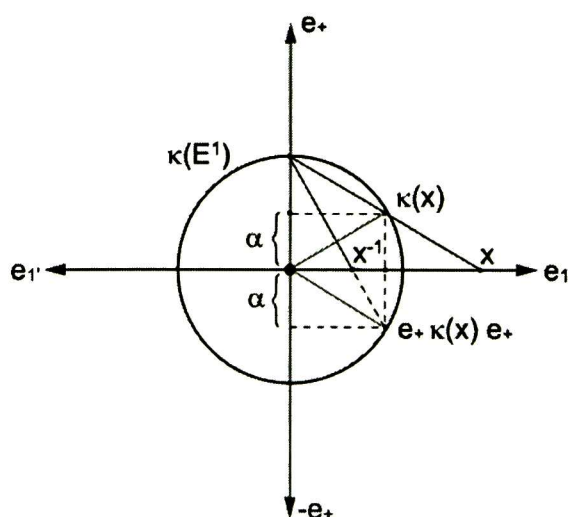


Figura A.18: Inversión de un vector $x \in \mathbb{E}^1$ con base en el mapeo estereográfico $\kappa(x)$ y la reflexión sobre el eje e_+ .

operador de proyección estereográfica $\kappa(x \in \mathbb{E}^n) = \frac{2}{x^2+1}x + \frac{x^2-1}{x^2+1}e_+ \in \mathbb{K}^n$, ahora la inversión de x denotada como $x^{-1} = \frac{x}{x^2}$, es equivalente a la inversión de x con respecto a la *hiperesfera* unitaria centrada en origen. Para verificar como se relaciona x^{-1} en el espacio conformal se proyecta estereográficamente, es decir,

$$\kappa(x^{-1} \in \mathbb{E}^n) = \frac{2}{\left(\frac{x}{x^2}\right)^2+1}\left(\frac{x}{x^2}\right) + \frac{\left(\frac{x}{x^2}\right)^2-1}{\left(\frac{x}{x^2}\right)^2+1}e_+ = \frac{2}{x^2+1}x - \frac{x^2-1}{x^2+1}e_+ \in \mathbb{K}^n \quad (\text{A.39})$$

lo cual muestra que el proceso de invertir un *vector* en \mathbb{E}^n es equivalente en el espacio conformal a invertir el signo de su componente e_+ . Esta característica es muy útil puesto que la inversión de signo respecto a una de la bases (e_+) puedes ser realizada mediante una reflexión respecto al eje (e_1) adecuado (ver la figura A.18).

A.10.7. Rotación en Espacio Conformal

Las transformaciones Euclidianas son un subconjunto de las transformaciones conformales, las transformaciones conformales son combinaciones de inversiones que en álgebra geométrica pueden ser ejecutadas mediante reflexiones, la rotación y la translación se implementan de esa forma, así se definen los rotores y transladores. El álgebra geométrica conformal posee cualidades del espacio proyectivo, entonces es posible expresar planos y líneas (que no necesariamente pasen por el origen), por consiguiente es posible reflejar con respecto a tales planos en posición arbitraria en el espacio (la rotación sería el doble del ángulo comprendido entre el par de planos usados, lo expuesto en la sección A.8). Considerando el ejemplo anterior, es posible observar que el eje de rotación es la línea de intersección de ambos planos. Si los planos efectivamente se intersectan entonces se produce una rotación alrededor de dicho eje, de lo contrario los planos son paralelos y el eje de intersección se encuentra en el infinito, entonces se produce una translación. La línea resultado de la intersección de dos planos o el producto exterior

de dos puntos con e_∞ se representa mediante su \mathbb{NO} , mientras que en su representación \mathbb{NI} la línea se expresa como $L = me_\infty + nE$ donde n representa la dirección y m el momento de la línea (será igual a cero para líneas que pasan por el origen). Una vez normalizada la línea L es posible definir el rotor²⁶ $R = e^{\frac{1}{2}\theta L}$, dicha notación es posible ya que la expansión de la función exponencial en series de Taylor sería de la forma $R = \cos \theta + L \sin \theta$. Rotar un punto en el espacio Euclidiano respecto a una línea arbitraria con un ángulo θ , es decir, si $x_i \in \mathbb{E}^3 \rightarrow \mathbf{X}_i = \rho(\kappa(x_i))$, entonces es posible obtener una línea que pase por cualquier par de puntos $L_{mn} = (\mathbf{X}_m \wedge \mathbf{X}_n \wedge e_\infty)^*$ bivector y normalizando $L'_{mn} = \frac{L_{mn}}{-(L_{mn} \cdot L_{mn})^{\frac{1}{2}}}$, el rotor puede ser construido como $R = e^{\frac{1}{2}\theta L'_{mn}}$, la rotación de cualquier punto $\mathbf{X}_i : i \neq m$ y $i \neq n$ sería de la forma $\mathbf{X}'_i = R\mathbf{X}_i\tilde{R}$, de θ radianes alrededor del eje formado por L_{mn} .

A.10.8. Translación en Espacio Conformal

Es posible expresar una translación en álgebra geométrica mediante dos reflexiones consecutivas con dos líneas paralelas, para lo cual existe un operador llamado el Tradador T , donde el vector $t \in \mathbb{E}^n$ representa la translación en el espacio Euclidiano, es decir,

$$T = 1 - \frac{1}{2}te_\infty : t \in \mathbb{E}^n, \quad (\text{A.40})$$

es también posible expresar el tradador mediante su notación exponencial como

$$T = e^{-\frac{1}{2}te_\infty} \quad (\text{A.41})$$

En el espacio conformal la translación es una rotación con la peculiaridad de que el plano de rotación no se encuentra en el espacio Euclidiano, sino en el espacio \mathbb{PK}^n . Ahora si se aplica un tradador T al origen e_0 , se tendría que $Te_0\tilde{T} = (1 - \frac{1}{2}te_\infty)e_0(1 + \frac{1}{2}te_\infty) =$

²⁶ L es un bivector normalizado.

$\rho(\kappa(t))$, es decir, sería igual al *vector* Euclidiano que constituye a T , lo cual es lo esperado, así mismo si se aplica un translador al punto en el infinito e_∞ se tendría de la forma $Te_\infty\tilde{T} = (1 - \frac{1}{2}t_\infty)e_\infty(1 + \frac{1}{2}te_\infty) = e_\infty$, es decir, el punto permanece ahí, en el infinito. Una de las características interesantes de los rotores y los transladores es que pueden ser aplicados a cualquier *Blade*, el mismo operador puede trasladar puntos, líneas, planos, esferas, círculos, etc. incluso es posible trasladar a un rotor, es decir,

$$M = TR\tilde{T} \tag{A.42}$$

lo cual implica un rotación más general que ahora incluye una translación, tal entidad es llamada motor. Ejemplo, sea $Y \in \mathbb{PK}^3$

$$\begin{aligned} Y' &= MY\tilde{M} \\ &= (TR\tilde{T})Y(TR\tilde{T}) \\ &= \{ T [R (\tilde{T} Y T) \tilde{R}] \tilde{T} \}, \end{aligned}$$

los paréntesis más internos representan una translación con un sentido $-t$, los corchetes muestran una rotación con respecto a R y finalmente las llaves aplican una translación un sentido t .

A.10.9. Espacios Euclidiano y Projectivo en Espacio Conformal

Ahora que se han tratado los espacios Euclidiano \mathbb{E}^3 , Projectivo \mathbb{PE}^3 y Conformal \mathbb{PK}^3 mediante el álgebra geométrica, es posible observar que en cierta forma el espacio conformal homogéneo posee las características de los dos espacios anteriores, codificándolas mediante sus dos bases extras, los *vectores* nulos e_∞ e_0 y la proyección estereográfica. Esta forma de contener los espacios Euclidiano y Projectivo en el espacio

Conformal no es inmediata ni transparente, pero con la ayuda de e_∞ y E es posible encontrar el isomorfismo entre dichos espacios. Sea $\mathbf{X} \in \mathbb{H}_d^n$, entonces es posible definir un operador

$$Q(\mathbf{X} \in \mathbb{H}_d^n) : \mathbf{X} \wedge E, \quad (\text{A.43})$$

de esta manera identificando las bases Euclidianas de la forma

$$\{e_i \wedge E\}_{i=1}^n, \quad (\text{A.44})$$

se visualiza el espacio Euclidiano inmerso en el conformal. Así mismo es posible detectar el espacio proyectivo inmerso, mediante el operador $Z(\mathbf{X} \in \mathbb{H}_d^n) : \mathbf{Z} \wedge e_\infty$, de esta manera identificando las bases proyectivas de la forma $\{e_i \wedge e_\infty\}_{i=1}^n$ y la base $e_0 \wedge e_\infty$ como e_{n+1} (la componente homogénea), con lo que se afirma que el espacio proyectivo se encuentra inmerso en el conformal.

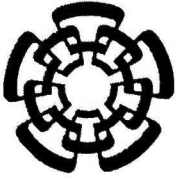
Bibliografía

- [1] B. Georgescu C. Christoudias and P. Meer. Synergism in low-level vision. *Proceedings. 16th International Conference on Pattern Recognition*, pages 150 – 155, 2002.
- [2] Rubén Machucho Cadena. *Métodos Geométricos y Difusos para Manipular Objetos*. PhD thesis, Centro de Investigación y Estudios Avanzados del IPN, Guadalajara, México, 2006.
- [3] Dietmar Hildenbrand Christian B.U. Perwass. Aspects of geometric algebra in euclidean, projective and conformal space. *Interactive Graphics Systems Group, TU Darmstadt, Germany*, September 2003.
- [4] Roberto Cipolla and Alex Pentland. *Computer Vision for Human-Machine Interaction*. Cambridge University Press, 1998.
- [5] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 603–619, 2002.
- [6] R. O. Duda and P. E. Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, pages 11 – 15, January 1972.

- [7] N. Brandenburg P. Kauff M. Karl E. Trucco, K. Plakas and O. Schreer. Real-time disparity maps for immersive 3-d teleconferencing by hybrid recursive matching and census transform. *IEEE Workshop on Video Registration vision List Digest on the International Conference on Computer Vision*, pages 5–9, July 2001.
- [8] Bayro Corrochano Eduardo. *Geometric Computing for Perception Action Systems*. Springer-Verlag, 2001.
- [9] Leo Reyes Lozano Eduardo Bayro Corrochano and Julio Zamora Esquivel. Conformal geometric algebra for robotic vision. *Journal of Mathematical Imaging and Vision*, pages 55 – 81, December 2006.
- [10] L. Falkenhagen. Hierarchical block-based disparity estimation considering neighbourhood constraints. *International workshop on SNHC and 3D Imaging*, pages 5–9, September 1997.
- [11] Sommer G., editor. *Geometric Computing with Clifford Algebras*. Springer-Verlag, Heidelberg, 2001.
- [12] Li H., Hestenes D., and Rockwood A. Generalized homogeneous coordinates for computational geometry. *En [11]*, pages 27–59, 2001.
- [13] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [14] Reyes Lozano Leo Hendrik. *Métodos Geométricos para Visión Robótica*. PhD thesis, Centro de Investigación y Estudios Avanzados del IPN, Guadalajara, México, 2004.
- [15] S. Rao I. Cox, S. Hingorani. A maximum likelihood stereo algorithm. *Computer Vision and Image Understanding*, (3), 1996.

- [16] Bernd Jäne. *Digital Image Processing*. Springer-Verlag, New York, 1993.
- [17] H. Pottmann M. Desbrun. An image processing approach to surface matching. *Eurographics Symposium on Geometry Processing*, 2005.
- [18] Sharon Oviatt. *Multimodal interfaces*. Lawrence Erlbaum Associates, Inc., Mahwah, NJ, USA, 2003.
- [19] B. Georgescu P. Meer. Edge detection with embedded confidence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1351 – 1365, December 2001.
- [20] T. Pajdla and J. Matas. Iso-disparity surfaces for general stereo configurations. *In Proceedings of The European conference on computer vision, ECCV 2004*, pages 509–520, 2004.
- [21] Lounesto Pertti. *Clifford Algebras and Spinors*. Cambridge University Press, 1997.
- [22] Marc Pollefeys. Visual 3d modeling from images. *University of North Carolina, Chapel Hill, USA Tutorial Notes*, 2000.
- [23] K. Kawamura R. A. Peters II, K. Hambuchen and D. M. Wilkes. The sensory ego-sphere as a shortterm memory for humanoids. *In Proceedings of the IEEE-RAS International Conference on Humanoid Robots*, pages 451–459, November 2001.
- [24] Jean-Yves Bouguet Technical report. Pyramidal implementation of the lucas kanade feature tracker description of the algorithm. *Intel Corporation Microprocessor Research Labs*, 2002.
- [25] Carlo Tomasi and Takeo Kanade. Detection and tracking of point features. Technical Report CMU-CS-91-132, Carnegie Mellon University, April 1991.

- [26] Zhengyou Zhang. Flexible camera calibration by viewing a plane from unknown orientations. *Proceedings of the International Conference on Computer Vision*, pages 666–73, 1999.



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL I.P.N. UNIDAD GUADALAJARA

El Jurado designado por la Unidad Guadalajara del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional aprobó la tesis

Visión Estereoscópica usando Álgebra Geométrica Conformal

del (la) C.

David Israel GONZÁLEZ AGUIRRE

el día 13 de Julio de 2006.

Dr. Eduardo José Bayro Corrochano
Investigador CINESTAV 3C
CINESTAV Unidad Guadalajara

Dr. Marco Antonio Perez Cisneros
Profesor Investigador
Centro Universitario de Ciencias
Exactas e Ingenierías

Dr. Mario Angel Siller González
Pico
Investigador CINESTAV
CINESTAV Unidad Guadalajara



CINVESTAV
BIBLIOTECA CENTRAL



SSIT000008567