

X
TJ 211
B37

UNIVERSIDAD DE LOS ANDES
FACULTAD DE INGENIERÍA
CONSEJO DE ESTUDIOS DE POSTGRADO
DIVISIÓN DE POSTGRADO
POSTGRADO EN AUTOMATIZACIÓN E INSTRUMENTACIÓN

**ALGORITMOS DE CONTROL VISUAL
BASADO EN IMAGEN PERMISIBLE A
VARIACIONES DEL ZOOM**

WWW.BDIGITAL.ULA.VE

Por:

Arnaldo Attilio Barrucci Dávila

Tutores: Dr. Ruben Rojas

Dr. Winston Garcia

DONACION

Trabajo de Grado presentado como Requisito Parcial
para obtener el Grado de
MAGÍSTER SCIENTIAE EN AUTOMATIZACIÓN
E INSTRUMENTACIÓN

MÉRIDA, SEPTIEMBRE de 2005

S E R B I U L A
Tulio Febres Cordero

Licencia Creative Commons:
Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

RESUMEN

Este trabajo presenta la propuesta de tres algoritmos de control visual basados en el uso de las características en la imagen, los cuales permiten la variación del zoom de la cámara bajo la condición la convergencia de la distancia focal, al valor utilizado en el momento en que se capturó la imagen de referencia, al final de la tarea de control. Para nuestro conocimiento, a excepción del Control Visual Invariante, estos esquemas son los únicos que permiten la variación del zoom de la cámara mientras se realiza la tarea de control. La Propuesta 1 asume que el valor de la distancia focal de referencia es desconocida, mientras que las Propuestas 2 y 3 asumen lo contrario. Estas dos últimas propuestas se diferencian en el cálculo de la velocidad rotacional respecto al eje óptico: una mediante el uso de las características en la imagen y la otra mediante el uso de puntos invariantes, respectivamente. Las Propuestas 1 y 2 son comparadas con el Control Visual Invariante ante diferentes tareas de control y bajo distintos niveles de ruido en la imagen, encontrándose que la Propuesta 1 presenta un desempeño relativamente menor que el Control Invariante en el caso en que se requiere de una translación y rotación general para el cumplimiento de la tarea. Por su parte la Propuesta 2 presenta un mejor desempeño en el caso en que se conozca la distancia focal usada en la imagen de referencia. Las Propuestas 2 y 3 son evaluadas simulando una situación en que el coeficiente angular s de la cámara (*skew coeficient*) es distinto de cero, encontrándose que ambos esquemas presentan un comportamiento bastante aceptable y muy similar entre si, prefiriéndose el uso de la Propuesta 2 por requerir de menor gasto computacional y ser ligeramente menos susceptible al ruido. Por último, se propuso y evaluó una ley general, aplicable a cualquiera de los esquemas propuestos, para mantener las características de la imagen dentro del campo visual de la cámara, obteniéndose resultados satisfactorios.

Palabras claves: Control visual basado en imagen, zoom variable, robótica, características en la imagen, variación de la distancia focal, ruido en la imagen.

INDICE DE CONTENIDO

CAPÍTULOS

1.- INTRODUCCIÓN.....	1
2.- MARCO TEÓRICO.....	7
2.1.- Modelos de proyección y configuraciones de cámara.....	7
2.1.1.- Proyección perspectiva.....	7
2.1.2.- Modelo de cámara.....	8
2.1.3.- Configuraciones de cámara.....	10
2.2.- Control visual basado en imagen.....	12
2.3.- Esquema dividido de Corke y Hutchinson.....	18
2.4.- Control visual invariante a los parámetros intrínsecos de la cámara.....	22
2.4.1.- Transformación si no se conoce el modelo de objeto.....	22
2.4.2.- Transformación si el modelo es conocido.....	23
2.4.3.- Equivalencia entre las dos transformaciones.....	24
3.- ESQUEMAS DE CONTROL VISUAL PROPUESTOS.....	27
3.1.- Propuesta 1.....	27
3.2.- Propuesta 2.....	36
3.2.1.- Independencia del cálculo de la velocidad de translación y rotación a lo largo del eje óptico de las variaciones del zoom de la cámara.....	39
3.3.- Propuesta 3.....	41
3.4.- Control de la distancia focal para evitar que las características se salgan del campo visual.....	42

4.- DESCRIPCIÓN DE LAS PRUEBAS PARA LA EVALUACIÓN DE LOS ESQUEMAS DE CONTROL VISUAL PROPUESTOS.....	44
4.1.- Simulación del objeto.....	45
4.2.- Descripción de las tareas de control.....	46
4.2.1.- Tarea 1: Rotación con respecto al eje óptico.....	46
4.2.2.- Tarea 2: Translación a lo largo del eje óptico.....	46
4.2.3.- Tarea 3: Rotación respecto a un eje perpendicular al eje óptico.....	47
4.2.4.- Tarea 4: Rotación del objeto.....	47
4.3.- Índices de desempeño.....	47
4.3.1.- Error final en translación (EFT).....	48
4.3.2.- Error final en rotación (EFR).....	48
4.3.3.- Translación máxima inducida (TMI).....	48
4.3.4.- Recorrido realizado por la cámara (RRC).....	49
4.3.5.- Separación máxima de las características del punto principal (SMPP).....	49
4.3.6.- Error final de la distancia focal (EFDF).....	49
4.4.- Condiciones de las Pruebas.....	50
4.4.1.- Prueba 1: Comparación de la Propuesta 1 y 2 con el CV invariante.....	50
4.4.1.1.- Condiciones de ruido.....	50
4.4.1.2.- Condiciones iniciales de los parámetros intrínsecos.....	50
4.4.1.3.- Sintonización de los esquemas de control.....	51
4.4.1.4.- Prueba t-student.....	52
4.4.2.- Prueba 2: Comparación de las Propuestas 2 y 3 para un coeficiente angular de la cámara distinto de cero.....	53
4.4.2.1.- Condiciones de ruido.....	54
4.4.2.2.- Condiciones iniciales de los parámetros intrínsecos.....	54
4.4.2.3.- Sintonización de los esquemas de control.....	55
4.4.2.4.- Prueba t-student.....	55

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

4.4.3.- Prueba 3: Uso del zoom de la cámara para evitar que las características se salgan del campo visual.....	55
4.4.3.1.- Condiciones de ruido.....	56
4.4.3.2.- Condiciones iniciales de los parámetros intrínsecos.....	56
4.4.3.3.- Sintonización de los esquemas de control.....	56
5.- RESULTADOS Y DISCUSIÓN	57
5.1.- Prueba 1: Comparación de la Propuesta 1 y 2 con el CV invariante.....	57
5.1.1.- Tarea 1: Rotación a lo largo del eje óptico con distintos niveles de ruido.....	58
5.1.1.1.- Error final en translación.....	58
5.1.1.2.- Error final en rotación.....	60
5.1.1.3.- Translación máxima inducida durante la tarea.....	62
5.1.1.4.- Separación máxima de las características del punto principal.....	64
5.1.1.5.- Error final de la distancia focal.....	66
5.1.2.- Tarea 2.- Translación a lo largo del eje óptico con distintos niveles de ruido.....	67
5.1.3.- Tarea 3: Rotación respecto a un eje perpendicular al eje óptico con distintos niveles de ruido.....	68
5.1.4.- Tarea 4: Rotación del objeto con diferentes niveles de ruido.....	69
5.2.- Prueba 2: Comparación de las Propuestas 2 y 3 para un coeficiente de s de la cámara distinto de cero.....	69
5.2.1.- Tarea 1: Rotación a lo largo del eje óptico para distintos niveles de ruido.....	69

5.2.2.- Tarea 2: Translación a lo largo del eje óptico con distintos niveles de ruido.....	70
5.2.3.- Tarea 3: Rotación respecto a un eje perpendicular al eje óptico con distintos niveles de ruido.....	71
5.2.4.- Tarea 4: Rotación del objeto con distintos niveles de ruido.....	71
5.3.- Prueba 3: Uso del zoom de la cámara para evitar que las características se salgan del campo visual.....	72
6.- CONCLUSIONES.....	75
REFERENCIAS BIBLIOGRÁFICAS.....	78
APÉNDICE A: Matriz jacobiana en el espacio invariante.....	82
APÉNDICE B: Tabla de la distribución de la t-student.....	87
APÉNDICE C: Detalles Prueba 1 (Tareas 2-4) y Prueba 2 (Tareas 1-4).....	88

INDICE DE FIGURAS

1.1	Control visual basado en posición (3D)	2
1.2	Control visual basado en imagen (2D)	3
2.1	Modelo de cámara.....	9
2.2	Configuraciones de cámara usadas en el control visual de manipuladores.....	11
2.3	Diferencia entre las configuraciones de cámara y su efecto sobre la imagen.....	14
2.4	Comparación del comportamiento entre el CV2D clásico y el esquema dividido, para un error inicial de rotación pura a lo largo del eje óptico de 80°.....	21
3.1	CV2D clásico, objeto no plano y distancia focal inicial distinta al valor con que se capturó la imagen de referencia.....	30
3.2	CV2D 7GDL, objeto no plano y distancia focal inicial distinta al valor con que se capturó la imagen de referencia.....	31
3.3	CV2D clásico, objeto no plano y distancia focal inicial igual al valor con que se capturó la imagen de referencia, error de rotación de 60° en el eje Z.....	32
3.4	CV2D 7GDL, objeto no plano y distancia focal inicial igual al valor con que se capturó la imagen de referencia, error de rotación de 60° en el eje Z.....	33
3.5	CV2D 7GDL, objeto no plano, y distancia focal inicial igual al valor con que se capturó la imagen de referencia, error de rotación de 180° en el eje Z. Distancia focal mínima soportada por la cámara = 300 píxeles.....	34
3.6	Propuesta 1, objeto no plano, y distancia focal inicial igual al valor con que se capturó la imagen de referencia, error de rotación en Z de 180°.....	36
3.7	Diferencia del cálculo de θ en fusión de la orientación del vector B	40
4.1	Posición y orientación de referencia de la cámara con respecto al objeto.....	45
5.1	EFT para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).....	59
5.2	EFR para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).....	61

5.3	TMI para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).....	63
5.4	SMPP para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).....	65
5.5	EFDF para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).....	66
5.6	Desplazamiento de las características en la imagen.....	72
5.7	Variación de la distancia focal.....	73
5.8	Error en translación.....	73
5.9	Error en rotación.....	74
C.1.1	EFT para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	88
C.1.2	EFR para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	89
C.1.3	RRC para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	90
C.1.4	SMPP para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	91
C.1.5	EFDF para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	92
C.2.1	EFT para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	93
C.2.2	EFR para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	94
C.2.3	TMI para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	95
C.2.4	SMPP para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	96
C.2.5	EFDF para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	97
C.3.1	EFT para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	98
C.3.2	EFR para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	99
C.3.3	RRC para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	100
C.3.4	SMPP para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	101
C.3.5	EFDF para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	102
C.4.1	EFT para la tarea 1 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	103
C.4.2	ERT para la tarea 1 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	103
C.4.3	TMI para la tarea 1 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	104
C.4.4	SMPP para la tarea 1 y diferentes niveles de ruido, prueba 2 ($s = 0,5$)....	105

C.5.1	EFT para la tarea 2 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	105
C.5.2	EFR para la tarea 2 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	106
C.5.3	RRC para la tarea 2 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	107
C.5.4	SMPP para la tarea 2 y diferentes niveles de ruido, prueba 2 ($s = 0,5$)...	107
C.6.1	EFT para la tarea 3 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	108
C.6.2	EFR para la tarea 3 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	109
C.6.3	TMI para la tarea 3 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	109
C.6.4	SMPP para la tarea 3 y diferentes niveles de ruido, prueba 2 ($s = 0,5$)....	110
C.7.1	EFT para la tarea 4 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	111
C.7.2	EFR para la tarea 4 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	111
C.7.3	RRC para la tarea 4 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).....	112
C.7.4	SMPP para la tarea 4 y diferentes niveles de ruido, prueba 2 ($s = 0,5$)....	113

WWW.BDIGITAL.ULA.VE

INDICE DE TABLAS

4.1	Parámetros intrínsecos de la cámara para la imagen inicial y la imagen de referencia utilizados en la Prueba 1.....	51
4.2	Ganancias utilizadas por el CV invariante, Propuesta 1 y Propuesta 2 para la prueba 1.....	52
4.3	Parámetros intrínsecos de la cámara para la imagen inicial y la imagen de referencia utilizados en la Prueba 2.....	54
4.4	Parámetros intrínsecos de la cámara para la imagen inicial y la imagen de referencia utilizados en la Prueba 3.....	56
5.1	Ganancias utilizadas por el CV invariante (modificada), Propuesta 1 y Propuesta 2 para la prueba 1.....	57
5.2	EFT promedio por nivel de varianza del ruido para la tarea 1 (desde 30° a 170°), Prueba 1 ($s = 0$).....	60
5.3	EFR promedio por nivel de varianza del ruido para la tarea 1 (desde 30° a 170°), Prueba 1 ($s = 0$).....	61
5.4	TMI promedio por nivel de varianza del ruido para la tarea 1 (desde 30° a 170°), Prueba 1 ($s = 0$).....	63
5.5	SMPP promedio por nivel de varianza del ruido para la tarea 1 (desde 30° a 170°), Prueba 1 ($s = 0$).....	65
5.6	Resumen de la Tarea 2 para los índices EFT, EFR, RRC y SMPP. Prueba 1.....	67
5.7	Resumen de la Tarea 3 para los índices EFT, EFR, RRC y SMPP. Prueba 1.....	68
5.8	Resumen de la Tarea 4 para los índices EFT, EFR, RRC y SMPP. Prueba 1.....	69
5.9	Resumen de la Tarea 1 para los índices EFT, EFR, RRC y SMPP. Prueba 2.....	70
5.10	Resumen de la Tarea 2 para los índices EFT, EFR, RRC y SMPP. Prueba 2.....	70
5.11	Resumen de la Tarea 3 para los índices EFT, EFR, RRC y SMPP. Prueba 2.....	71
5.12	Resumen de la Tarea 4 para los índices EFT, EFR, RRC y SMPP. Prueba 2.....	71
B	Tabla de la distribución de la t -student.....	87

C.1.1	EFT promedio por nivel de varianza del ruido para la tarea 2, Prueba 1 ($s = 0$).....	88
C.1.2	EFR promedio por nivel de varianza del ruido para la tarea 2, Prueba 1 ($s = 0$).....	89
C.1.3	RRC promedio por nivel de varianza del ruido para la tarea 2, Prueba 1 ($s = 0$).....	90
C.1.4	SMPP promedio por nivel de varianza del ruido para la tarea 2, Prueba 1 ($s = 0$).....	91
C.2.1	EFT promedio por nivel de varianza del ruido para la tarea 3, Prueba 1 ($s = 0$).....	93
C.2.2	EFR promedio por nivel de varianza del ruido para la tarea 3, Prueba 1 ($s = 0$).....	94
C.2.3	TMI promedio por nivel de varianza del ruido para la tarea 3, Prueba 1 ($s = 0$).....	95
C.2.4	SMPP promedio por nivel de varianza del ruido para la tarea 3, Prueba 1 ($s = 0$).....	96
C.3.1	EFT promedio por nivel de varianza del ruido para la tarea 4, Prueba 1 ($s = 0$).....	98
C.3.2	EFR para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	99
C.3.3	RRC para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	100
C.3.4	SMPP para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).....	101
C.4.1	EFT promedio por nivel de varianza del ruido para la tarea 1, prueba 2 ($s = 0,5$).....	103
C.4.2	EFR promedio por nivel de varianza del ruido para la tarea 1, prueba 2 ($s = 0,5$).....	104
C.4.3	TMI promedio por nivel de varianza del ruido para la tarea 1, prueba 2 ($s = 0,5$).....	104
C.4.4	SMPP promedio por nivel de varianza del ruido para la tarea 1, prueba 2 ($s = 0,5$).....	105
C.5.1	EFT promedio por nivel de varianza del ruido para la tarea 2, prueba 2 ($s = 0,5$).....	106
C.5.2	EFR promedio por nivel de varianza del ruido para la tarea 2, prueba 2 ($s = 0,5$).....	106

C.5.3	RRC promedio por nivel de varianza del ruido para la tarea 2, prueba 2 ($s = 0,5$).....	107
C.5.4	SMPP promedio por nivel de varianza del ruido para la tarea 2, prueba 2 ($s = 0,5$).....	108
C.6.1	EFT promedio por nivel de varianza del ruido para la tarea 3, prueba 2 ($s = 0,5$).....	108
C.6.2	EFR promedio por nivel de varianza del ruido para la tarea 3, prueba 2 ($s = 0,5$).....	109
C.6.3	RRC promedio por nivel de varianza del ruido para la tarea 3, prueba 2 ($s = 0,5$).....	110
C.6.4	SMPP promedio por nivel de varianza del ruido para la tarea 3, prueba 2 ($s = 0,5$).....	110
C.7.1	EFT promedio por nivel de varianza del ruido para la tarea 4, prueba 2 ($s = 0,5$).....	111
C.7.2	EFR promedio por nivel de varianza del ruido para la tarea 4, prueba 2 ($s = 0,5$).....	112
C.7.3	RRC promedio por nivel de varianza del ruido para la tarea 4, prueba 2 ($s = 0,5$).....	112
C.7.4	SMPP promedio por nivel de varianza del ruido para la tarea 4, prueba 2 ($s = 0,5$).....	113

CAPÍTULO 1

INTRODUCCIÓN

En la actualidad, se considera la robótica como una de las tecnologías con mayor auge e impacto pudiéndose afirmar que los robots desarrollan una tarea muy importante en el área industrial, llegando incluso a considerarse imprescindibles en determinados ámbitos. Sin embargo, el uso de robots para la realización de una tarea en específico depende en gran medida del conocimiento a priori que se tenga sobre el entorno. Debido a esto, gran cantidad de investigaciones están siendo realizadas, teniendo como norte que los robots puedan trabajar en ambientes cada vez menos estructurados, acondicionándolos con sistemas sensoriales que les permita adaptarse mejor a los cambios inesperados. Estas investigaciones, conjuntamente con el alto desarrollo de los microprocesadores y de la instrumentación digital, han provocado un incremento en el estudio del uso de cámaras como sensores de realimentación, ya que éstas proveen de información dinámica acerca del ambiente, además de ser un tipo de información con la cual los humanos están altamente familiarizados. Esta área específica de investigación es conocida en la literatura como control visual de sistemas robóticos (*visual servoing*) y combina principalmente técnicas de visión por computador con estrategias de control.

Este tipo de control consiste en el uso de datos proveniente de una o varias cámaras para controlar un sistema robótico, y gracias al avance tecnológico desarrollado en las últimas décadas en el área computacional y procesamiento de imágenes, es posible llevarlo a cabo en tiempo real. Un sistema robótico podría ser un brazo manipulador, donde un objetivo típico del esquema de control visual sería mantener una posición y orientación relativa (de referencia) a algún objeto específico para así permitirle al robot realizar alguna tarea de interés sobre el mismo. El control visual de brazos manipuladores usualmente utiliza dos tipos de configuraciones de cámara. La

primera, en donde la(s) cámara(s) permanece(n) fija(s) y observa(n) la escena y al robot desde un punto externo [1], siendo esta configuración la típica en los primeros trabajos realizados en control visual de brazos manipuladores. La segunda, donde la(s) cámara(s) tiene(n) libertad de movimiento y se encuentra(n) montada(s) sobre algún eslabón del brazo robot, usualmente sobre el efector final (extremo del robot), denominada configuración de cámara-en-mano (*eye-in-hand configuration*) [2], típica de las investigaciones más recientes. Adicionalmente, existe la posibilidad de combinar ambas configuraciones en un mismo esquema [3]. Por otro lado, independientemente de la configuración utilizada, existe otra clasificación basada en el número de cámaras, que distingue tres tipos de sistemas: los monoculares donde se utiliza solo una cámara, los binoculares o stereos donde se utilizan dos, y los redundantes donde se utilizan más de dos.

Por otro lado, los esquemas de control utilizados se pueden dividir en dos categorías que dependen del espacio en donde se define la función error a minimizar por el controlador [4]. La primera categoría, conocida como control visual basado en posición o control visual 3D, define el error en el espacio cartesiano, y utiliza la información proporcionada por el sistema de visión para realizar una reconstrucción tridimensional (parcial) de la escena [5, 6]. Cada iteración del lazo de control consiste en la captura de la imagen actual, determinación de la posición y orientación actual del robot respecto al objeto (interpretación de la imagen), comparación de esta con la posición y orientación de referencia, cálculo de la señal de control y envío de la misma (ver Fig. 1.1).

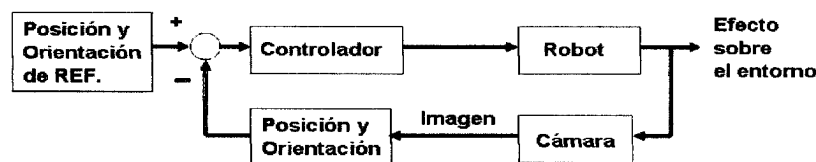


Figura 1.1. Control visual basado en posición (3D)

La segunda categoría, conocida como control visual basado en imagen o control visual 2D, define el error en el espacio bidimensional de la imagen y requiere de una imagen de referencia [7, 8]. Esta imagen es usualmente obtenida a través de la técnica conocida en la literatura como *teach by showing*, la cual consiste en llevar al robot a la posición y orientación deseada respecto al objeto de interés (posición y orientación de referencia) y capturar la imagen. Luego, se mueve el robot y/o el objeto y la tarea de control es iniciada. Comúnmente, de la imagen de referencia se extraen ciertas características del objeto, y en cada iteración del lazo de control las mismas son extraídas de la imagen actual, definiéndose un error entre ellas (ver Fig. 1.2). Este esquema hace uso de una matriz jacobiana, que relaciona la variación de las características en la imagen respecto a la variación de la posición y orientación del robot y/o la cámara, con la intención de generar en cada iteración un vector velocidad de 6 grados de libertad (velocidades de translación y de rotación a lo largo de los ejes X, Y, Z con respecto al sistema de coordenadas de la cámara), que define cómo ha de moverse el robot y/o la cámara para conseguir la configuración deseada en las características de la imagen. Esta matriz jacobiana es conocida en la literatura como *image jacobian*, *feature jacobian*, *feature sensitivity matrix*, o *interaction matrix* [9], y en el presente trabajo será denominada como *matriz jacobiana imagen*, pudiendo ser estimada analíticamente [10] o numéricamente en línea [11].

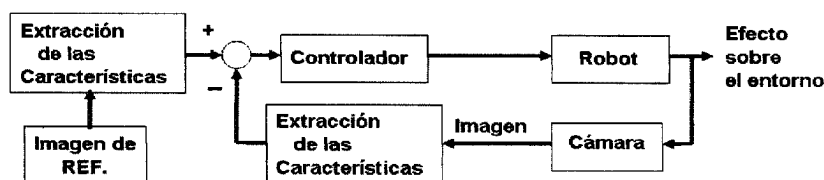


Figura 1.2. Control visual basado en imagen (2D)

Generalmente, la ley de control mas utilizada es la proporcional simple [12], pero también se han registrado en la literatura investigaciones utilizando control predictivo multivariable [13], control basado en asignación de polos, control PID, control

feedforward [14], control lineal cuadrático utilizando un modelo lineal invariante en el tiempo [15], control PI no-lineal [16] y control por modo deslizante [17], entre otros.

Por otro parte, tanto el control basado en imagen como el basado en posición presentan ventajas y desventajas [18-20], las cuales han llevado a la introducción de esquemas divididos, en donde se utiliza el control visual 2D para algunos grados de libertad y otras técnicas para controlar los grados de libertad restantes. Algunas de estas técnicas incluyen el uso de la geometría epipolar calculada en línea (basado en posición), la cual requiere del cálculo de la matriz homográfica estimada de la observación de al menos cuatro puntos en la imagen que pertenezcan a un objeto coplanar [21, 22]. Luego, con el conocimiento de los parámetros intrínsecos de la cámara, esta matriz debe ser descompuesta para extraer la información tridimensional, usando métodos como los presentados en [23-25]. Esta descomposición no es un procedimiento trivial y generalmente no posee una solución única, requiriéndose de un gasto computacional elevado aunque alcanzable en tiempo real. Otro esquema de tipo dividido, es el presentado en [26] donde a pesar que se desacoplan los grados de libertad referentes al eje óptico, estos son controlados utilizando directamente características en la imagen sin requerir el uso de geometría epipolar, controlando los restantes grados con el control visual 2D.

También existen enfoques de control visual no divididos como el control visual dinámico (*dynamic visual servoing*), el cual no utiliza características geométricas de la escena sino características dinámicas [27], y cuyo principal inconveniente es el tiempo de duración de cada iteración debido al gasto computacional, limitando la velocidad del movimiento del robot. Otro esquema no dividido es el control visual invariante [28], siendo una de las tendencias más recientes e interesantes. Este esquema realiza una proyección a un espacio independiente de los parámetros intrínsecos de la cámara (espacio en donde se define la función error), permitiendo la variación del zoom durante la tarea de control.

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

Aunque el control visual 2D puede presentar algunos problemas, este esquema ha mostrado un incremento de popularidad. Principalmente gracias a que cualquier error en la calibración de los sistemas de visión usados en el control visual 3D, conlleva a errores en la reconstrucción tridimensional, y por lo tanto, traerá errores en la ejecución de la tarea de control. Es por esto, que nuestra investigación se centra en el área del control visual 2D. Además, uno de los requerimientos para el uso de este tipo de control es que los parámetros intrínsecos sean iguales a los usados en el momento en que se capturó la imagen de referencia, de hecho si los parámetros intrínsecos usados son diferentes, la posición y orientación del robot al final de la convergencia será completamente distinta a la posición y orientación de referencia, a pesar que se haya minimizado el error de las características en la imagen y su valor tienda a cero.

Por lo tanto, el objetivo principal de esta investigación fue proponer esquemas de control visual, basados en las características de la imagen, que permitan explotar el potencial del zoom de las cámaras modernas (variando la distancia focal, la cual es un parámetro intrínseco), para aumentar el espacio de trabajo del robot y asegurar que las características (del objeto en estudio) se mantengan dentro del campo visual de la cámara durante toda la tarea de control, condición de suma importancia para el éxito de la tarea de control. Estos esquemas están basados principalmente en el control visual 2D clásico y el esquema divididos en donde todos los grados de libertad son calculados utilizando directamente las características en la imagen [26]. Estos conceptos se complementan en el **Capítulo 2**, junto al control visual invariante y las nociones básicas de proyección de imágenes, modelos de cámara y las diferentes configuraciones cámara/robot.

En el **Capítulo 3** se presentan tres propuestas de control visual permisibles a la variación de la distancia focal, de los cuales, la Propuesta 1 considera la distancia focal usada en la captura de la imagen de referencia como desconocida, lo que no

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

hacen las Propuestas 2 y 3. También se plantea una ley de control de la distancia focal para evitar que las características se salgan del campo visual de la cámara, aplicable a cualquiera de los esquemas propuestos. En el **Capítulo 4** se expone la metodología para la evaluación de los esquemas propuestos comparándolos principalmente con el control visual invariante, el cual (a nuestro entender), es el único algoritmo de control visual existente en la literatura que permite la variación de la distancia focal durante la tarea de control. Para ello, se realizan pruebas a través de simulaciones, basadas en tareas de control postuladas en la literatura para la evaluación de esquemas de control visual bajo diferentes condiciones de ruido en la imagen [29]. El **Capítulo 5** muestra los resultados obtenidos y un análisis de los mismos, finalizándose con las conclusiones del trabajo realizado, en el **Capítulo 6**.

WWW.BDIGITAL.ULA.VE

CAPÍTULO 2

MARCO TEÓRICO

En este capítulo se presentan los conceptos básicos relacionados con la proyección de imágenes y modelos de cámara, así como las diferentes configuraciones cámara/robot encontradas para el control visual de sistemas robóticos. También, se examinan los fundamentos del control visual basado en imagen, del esquema dividido propuesto en [26] y del control visual invariante a cambios en los parámetros intrínsecos de la cámara [28].

2.1.- Modelos de proyección y configuraciones de cámara

Cada cámara posee un lente que produce una proyección bidimensional de la escena. Esta proyección causa una pérdida directa de la profundidad debido a que una escena tridimensional se ve proyectada en un plano. Adicionalmente, esta(s) cámara(s) puede(n) estar colocada(s) de diferentes maneras con relación al sistema robótico. A continuación, se describen los modelos de proyección de imagen y de cámara utilizados, así como las distintas configuraciones de cámara/robot encontradas en el control visual.

2.1.1.- Proyección perspectiva [4]

Sea ${}^C \mathbf{p}_i = [x_i \quad y_i \quad z_i]^T$ las coordenadas de un punto P_i , respecto al sistema de coordenadas de la cámara C . La proyección de este punto sobre el plano imagen con coordenadas $\mathbf{p}_i = [u_i \quad v_i]^T$, es dada por la proyección perspectiva:

$$\mathbf{p}_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \frac{f}{z_i} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (\text{donde } z_i \neq 0) \quad (2.1)$$

donde f representa la distancia focal. Esta proyección es una representación no lineal de una escena tridimensional en un plano imagen y es también conocida como modelo *pinhole*. Para lectores interesados en otros tipos de modelos para proyecciones de imágenes como la proyección ortogonal escalada y la proyección *affine* se recomienda el tutorial en control visual de Hutchinson y col. [4].

Es importante señalar que si el punto P_i de la escena se encuentra parametrizado en otro sistema de coordenadas arbitrario α , se debe realizar la transformación necesaria para llevarlo al sistema de coordenadas de la cámara:

$${}^c \mathbf{p}_i = {}^c \mathbf{T}_\alpha {}^\alpha \mathbf{p}_i \quad (2.2)$$

donde ${}^c \mathbf{T}_\alpha$ es la transformación que relaciona al sistema de coordenadas α con el de la cámara, mientras que ${}^\alpha \mathbf{p}_i$ representa las coordenadas del punto \mathbf{p}_i en el sistema de coordenadas de α .

2.1.2.- Modelo de cámara [30]

Un modelo de cámara más completo, pero también de tipo proyectivo, considera que el plano imagen en una cámara real es constituido por una matriz de celdas sensibles a la luz (sensor), conocidas como píxeles. La información dada por la cámara es un punto \mathbf{p}_i que depende de los parámetros intrínsecos de la misma y la posición tridimensional (respecto a la cámara) del punto proyectado (ver Fig. 2.1), tal que:

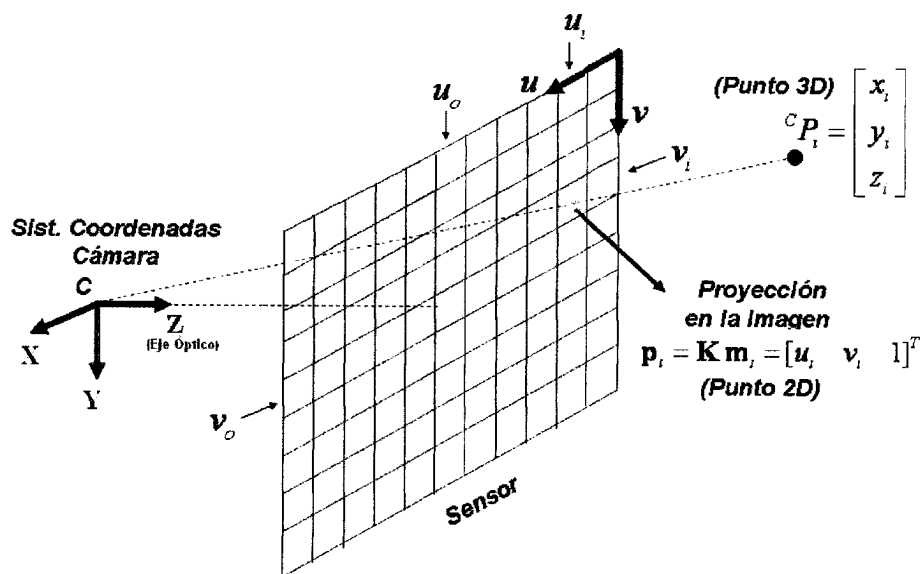


Figura 2.1 Modelo de cámara.

$$\mathbf{p}_i = \mathbf{K} \mathbf{m}_i = [u_i \quad v_i \quad 1]^T \quad (2.3)$$

donde \mathbf{m}_i se denomina punto normalizado y esta definido como:

$$\mathbf{m}_i = \begin{bmatrix} \frac{x_i}{z_i} & \frac{y_i}{z_i} & 1 \end{bmatrix}^T = [\bar{x}_i \quad \bar{y}_i \quad 1]^T \quad (\text{donde } z_i \neq 0) \quad (2.4)$$

y \mathbf{K} es una matriz triangular superior que contiene los parámetros intrínsecos de la cámara [30]:

$$\mathbf{K} = \begin{bmatrix} f & s & u_o \\ 0 & r f & v_o \\ 0 & 0 & 1 \end{bmatrix} \quad (2.5)$$

siendo f la distancia focal (en píxeles), u_0 y v_0 las coordenadas del punto principal (en píxeles) representando la intersección del eje óptico con el plano imagen, s un coeficiente angular dependiente de la forma del píxel (*skew coefficient*) y r la relación dimensional de los lados del píxel conocida como la relación de aspecto.

2.1.3.- Configuraciones de cámara

Las configuraciones de cámara utilizadas en el control visual de manipuladores, están divididas en sistemas de visión monocular, binocular y redundante. Los sistemas de visión monocular por el uso de solo una cámara, minimizan el tiempo de procesamiento necesario para extraer la información visual. Estos sistemas monoculares pueden existir en sistemas cámara-en-mano o con cámara externa al robot. En los sistemas de cámara-en-mano (Fig. 2.2(a)), la cámara esta rígidamente unida al efector final del robot y usualmente la transformación de los sistemas de coordenadas entre ellos (cámara y efector final) es conocida de antemano, obtenida mediante algún método de calibración [31]. En este tipo de configuración, generalmente, el control adopta técnicas basadas en el conocimiento del modelo del objeto para estimar la posición (profundidad) entre la cámara y el objeto [32]. En los sistemas de visión monocular externa al robot (ver Fig. 2.2(b)), la cámara no se encuentra unida directamente a un eslabón móvil del robot y generalmente se obtiene un campo visual mayor que en el caso de sistemas cámara-en-mano. Esta configuración es típica de los primeros sistemas de control visual.

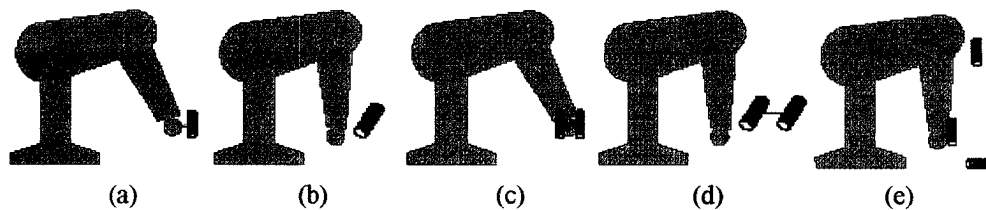


Figura 2.2 Configuraciones de cámara usadas en el control visual de manipuladores. (a) Sistema ojo-en-mano monocular, (b) Sistema con cámara externa monocular, (c) Sistema ojo-en-mano binocular, (d) Sistema con cámara externa binocular, (e) Sistema redundante [12].

Por otro lado, los sistemas de visión binocular, utilizan dos cámaras para generar visión estereo, lo que permite una interpretación tridimensional completa de la escena. Uno de los enfoques mas comunes es la estimación de la disparidad, la cual es usada, para estimar la profundidad. El problema fundamental consiste en determinar la correspondencia entre las características extraídas de las dos imágenes. Aunque, el gasto computacional se duplica, tiene la ventaja de no necesitar explícitamente de un modelo del objeto para obtener información sobre la profundidad. Esta configuración binocular raramente se encuentra en sistemas de cámara-en-mano (Fig. 2.2(c)), ya que a pesar que genera información tridimensional de importancia, la distancia de separación entre las cámaras (*the baseline*), se ve limitada por el espacio disponible, afectando la precisión de la reconstrucción. En cambio, sistemas de visión binocular externa al robot (Fig. 2.2(d)), son usados con mayor frecuencia, debido a que la distancia de separación entre las cámaras puede escogerse lo suficientemente grande como para obtener una reconstrucción adecuada. Esta disposición permite un campo visual bastante amplio que admite observar tanto el objeto como al robot simultáneamente.

Los sistemas de visión redundantes (Fig. 2.2(e)), por contar con mas de dos cámaras pueden proveer información adicional. Estos sistemas son usados muy poco debido a que determinar la correspondencia entre las características extraídas de las distintas

imágenes se convierte en un problema poco trivial, que requiere de un gasto computacional elevado.

El presente trabajo se basa principalmente en el estudio de sistemas de cámara-enmano de tipo monocular, por ser considerado el de mayor uso e interés en las publicaciones mas recientes [12, 13, 29].

2.2.- Control visual basado en imagen

El control visual basado en imagen (CV2D clásico), define su señal de referencia en el espacio bidimensional de la imagen, mediante la extracción de características del objeto observado. Estas características son valores como, el tamaño o coordenadas del centroide de una región, las coordenadas de puntos en la imagen, la longitud de un segmento, entre otros [33].

La variación de estas características, c , utilizadas para el control visual de brazos manipuladores, pueden ser escrita en el caso de una configuración de cámara-enmano como [34]:

$$\dot{c} = \frac{\partial c}{\partial r} \frac{\partial r}{\partial t} + \frac{\partial c}{\partial t} = \mathbf{J} \cdot \mathbf{v}_c + \frac{\partial c}{\partial t} \quad (2.6)$$

donde:

$\mathbf{J} = \frac{\partial c}{\partial r}$ conocida como la *matriz jacobiana imagen* o *matriz interacción*, representa la variación de las características respecto a la variación de la posición y orientación de la cámara.

$\mathbf{v}_c = \frac{\partial \mathbf{r}}{\partial t}$ representa la variación de la posición y orientación de la cámara respecto al tiempo. Siendo un vector de 6x1, tal que:

$$\mathbf{v}_c = [T_x \quad T_y \quad T_z \quad \omega_x \quad \omega_y \quad \omega_z]^T \quad (2.7)$$

donde T_x , T_y , T_z , son velocidades de translación y ω_x , ω_y , ω_z son velocidades angulares referidas a los ejes X, Y, Z del sistema de coordenadas adjunto a la cámara, respectivamente.

$\frac{\partial c}{\partial t}$ representa la variación de las características debido al movimiento potencial del objeto observado. Este término es usualmente considerado igual a cero suponiendo que el objeto es inmóvil.

Análogamente, para el caso de configuración de cámara externa al robot, podemos describir la ecuación (2.6) como:

$$\dot{c} = -\mathbf{J} \cdot \mathbf{v}_c + \frac{\partial c}{\partial t} \quad (2.8)$$

donde el cambio de signo en el término $\mathbf{J} \cdot \mathbf{v}_c$ se debe al cambio de configuración del sensor con respecto a la variables de control (ver Fig. 2.3).

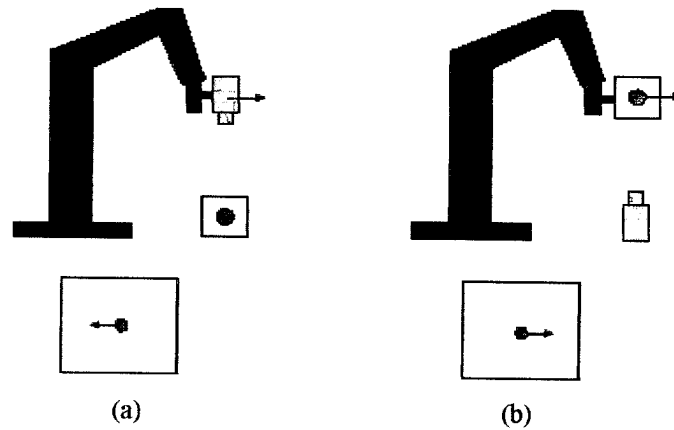


Figura 2.3 Diferencia entre las configuraciones de cámara y su efecto sobre la imagen.
 (a) Configuración cámara-en-mano, (b) Configuración de cámara externa al robot [34].

Por otro lado, una expresión para la *matriz jacobiana imagen* \mathbf{J} , utilizando puntos como características y una transformación ideal de la proyección (ver sección 2.1.1) para un sistema de cámara-en-mano, se expresa como [15]:

$$\mathbf{J} = [\mathbf{J}_1^T \quad \mathbf{J}_2^T \quad \dots \quad \mathbf{J}_n^T]^T \quad (2.9)$$

$$\mathbf{J}_i = \begin{bmatrix} -\frac{f}{z_i} & 0 & \frac{u_i}{z_i} & \frac{u_i v_i}{f} & \frac{-f^2 - u_i^2}{f} & v_i \\ 0 & -\frac{f}{z_i} & \frac{v_i}{z_i} & \frac{f^2 + v_i^2}{f} & -\frac{u_i v_i}{f} & -u_i \end{bmatrix} \quad (2.10)$$

donde f es la distancia focal de la cámara, z_i es la distancia del punto a la cámara a lo largo del eje óptico (profundidad), u_i y v_i son las coordenadas horizontales y verticales del punto en el plano imagen, respectivamente. Se puede apreciar que la *matriz jacobiana imagen* depende de la posición de los puntos en el plano imagen, del

estimado de los parámetros intrínsecos de la cámara, y del estimado de la profundidad de los puntos respecto al sistema de coordenadas de la cámara.

Entonces, considerando que el objeto se mantiene estático, tendremos:

$$\begin{bmatrix} \dot{u}_i \\ \dot{v}_i \end{bmatrix} = \mathbf{J}_i \cdot \mathbf{v}_c \quad (2.11)$$

Existen diferentes expresiones de la *matriz jacobiana imagen* en donde las características extraídas de la imagen son rectas, primitivas planas, círculos, primitivas tridimensionales y esferas [10].

Ahora bien, considerando un modelo de cámara más completo, como el enunciado en la Sección 2.1.2, y utilizando puntos como características en la imagen, obtenemos:

$$\begin{aligned} u_i &= \frac{f}{z_i} x_i + \frac{s}{z_i} y_i + u_o \\ v_i &= \frac{rf}{z_i} y_i + v_o \end{aligned} \quad (2.12)$$

Asumiendo que los parámetros intrínsecos de la cámara permanecen constantes, las derivadas de u_i y v_i con respecto al tiempo, pueden definirse como:

$$\begin{aligned} \dot{u}_i &= \frac{f}{z_i} \dot{x}_i + \frac{s}{z_i} \dot{y}_i - \left(\frac{x_i f + y_i s}{z_i^2} \right) \dot{z}_i \\ \dot{v}_i &= \frac{rf}{z_i} \dot{y}_i - \frac{rf y_i}{z_i^2} \dot{z}_i \end{aligned} \quad (2.13)$$

Por otro lado, la variación respecto al tiempo de las coordenadas de un punto ${}^C \mathbf{p}_i = [x_i \ y_i \ z_i]^T$ referente al sistema de coordenadas de la cámara C , se expresa como [4]:

$$\begin{aligned}\dot{x}_i &= z_i \omega_y - y_i \omega_z + T_x \\ \dot{y}_i &= x_i \omega_z - z_i \omega_x + T_y \\ \dot{z}_i &= y_i \omega_x - x_i \omega_y + T_z\end{aligned}\quad (2.14)$$

Sustituyendo la ecuación (2.14) en la ecuación (2.13) y reagrupando, se obtiene que *matriz jacobiana imagen* para un sistema con cámara externa, puede definirse como:

$$\mathbf{J}_{CE} = [\mathbf{J}_{CE_1}^T \ \mathbf{J}_{CE_2}^T \ \dots \ \mathbf{J}_{CE_n}^T]^T \quad (2.15)$$

$$\mathbf{J}_{CE_i} = \begin{bmatrix} \frac{f}{z_i} & \frac{s}{z_i} & -\left(\frac{fx_i + sy_i}{z_i^2}\right) & -\left(s + \frac{(fx_i + sy_i)y_i}{z_i^2}\right) & \left(f + \frac{(fx_i + sy_i)x_i}{z_i^2}\right) & \left(\frac{sx_i - fy_i}{z_i}\right) \\ 0 & \frac{rf}{z_i} & \frac{ry_i f}{z_i^2} & -\left(rf + \frac{ry_i^2 f}{z_i^2}\right) & \frac{ry_i x_i f}{z_i^2} & \frac{rx_i f}{z_i} \end{bmatrix} \quad (2.16)$$

Mientras que para un sistema de cámara-en-mano, la *matriz jacobiana imagen* se expresa como:

$$\mathbf{J}_C = [\mathbf{J}_{C_1}^T \ \mathbf{J}_{C_2}^T \ \dots \ \mathbf{J}_{C_n}^T]^T \quad (2.17)$$

$$\mathbf{J}_C = \begin{bmatrix} \frac{f}{z_i} & \frac{s}{z_i} & \left(\frac{fx_i + sy_i}{z_i^2}\right) & \left(s + \frac{(fx_i + sy_i)y_i}{z_i^2}\right) & -\left(f + \frac{(fx_i + sy_i)x_i}{z_i^2}\right) & -\left(\frac{sx_i - fy_i}{z_i}\right) \\ 0 & \frac{rf}{z_i} & \frac{ry_i f}{z_i^2} & \left(rf + \frac{ry_i^2 f}{z_i^2}\right) & \frac{ry_i x_i f}{z_i^2} & \frac{rx_i f}{z_i} \end{bmatrix} \quad (2.18)$$

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

donde x_i e y_i pueden despejarse de la ecuación (2.12), tal que:

$$\begin{aligned} x_i &= \frac{(u_i - u_o)z_i}{f} - \frac{s(v_i - v_o)z_i}{f^2 r} \\ y_i &= \frac{(v_i - v_o)z_i}{rf} \end{aligned} \quad (2.19)$$

Sustituyendo las ecuaciones (2.19) en la ecuación (2.18) y considerando valores para $r = 1$, $s = 0$ [30], y el punto principal (u_o, v_o) ubicado en el centro de la imagen [35], como una buena estimación especialmente para cámaras modernas, obtenemos:

$$\mathbf{J}_C = \begin{bmatrix} -\frac{f}{z_i} & 0 & \frac{(u_i - u_o)}{z_i} & \left(\frac{(u_i - u_o)(v_i - v_o)}{f} \right) & -\left(\frac{f^2 + (u_i - u_o)^2}{f} \right) & (v_i - v_o) \\ 0 & -\frac{f}{z_i} & \frac{(v_i - v_o)}{z_i} & \left(\frac{f^2 + (v_i - v_o)^2}{f} \right) & -\frac{(u_i - u_o)(v_i - v_o)}{f} & -(u_i - u_o) \end{bmatrix} \quad (2.20)$$

Observándose que si $u_o = v_o = 0$, se obtiene la misma expresión dada por la ecuación (2.10).

Finalmente, considerando que el objeto observado se mantiene inmóvil, se define una ley de control de tipo proporcional (la cual es la estrategia más comúnmente usada) para calcular el vector velocidad \mathbf{v}_C (señal de control), tal que:

$$\mathbf{v}_C = \mathbf{\Gamma} \cdot \mathbf{J}_C^+ \cdot (\mathbf{c}^* - \mathbf{c}) \quad (2.21)$$

donde \mathbf{J}_C^+ representa la pseudo-inversa de la matriz \mathbf{J}_C , $\mathbf{\Gamma}$ es una matriz diagonal de dimensiones apropiadas, con escalares positivos que sintoniza la tasa de

convergencia. \mathbf{c}^* y \mathbf{c} son vectores columna que contienen las características extraídas de la imagen de referencia y de la imagen actual, respectivamente.

La imagen de referencia es comúnmente adquirida utilizando la técnica conocida como “*teach by showing*” en donde el robot es llevado a la posición y orientación de referencia y la imagen es capturada. Por otro lado, si el objeto no se mantiene inmóvil esto simplemente se verá reflejado como una perturbación en el sistema de control.

2.3.- Esquema dividido de Corke y Hutchinson [26]

Debido a que el control visual basado en imagen es óptimo respecto a la trayectoria de las características en la imagen, éstas tienden a moverse en línea recta hacia su configuración deseada. Esto produce que las rotaciones a lo largo del eje óptico sean acompañadas de una translación a lo largo del mismo, conociéndose este comportamiento como el fenómeno de retracción de cámara [18]. A medida que esta rotación se acerca a los 180° , la retracción tiende al infinito anulándose la componente de ω_z del vector \mathbf{v}_C . Esta componente ω_z es calculada mediante la ecuación (2.21), y puede ser reescrita como:

$$\omega_z = (\mathbf{\Gamma})_{6,6}(\mathbf{J}_C^+)_{6,6}(\mathbf{c}^* - \mathbf{c}) \quad (2.22)$$

en donde $(\mathbf{J}_C^+)_{6,6}$ y $(\mathbf{\Gamma})_{6,6}$ denotan la última fila de la pseudo inversa de \mathbf{J}_C y el valor del escalar ubicado en la última fila y última columna de la matriz de ganancias $\mathbf{\Gamma}$, respectivamente.

En el caso particular de una rotación de 180° , esta operación es nula, inclusive si $(\mathbf{c}^* - \mathbf{c}) \neq 0$, debido a que las diferentes contribuciones de la *matriz jacobiana*

imagen para el cálculo de esta velocidad rotacional se cancelan una a otras [26]. Por lo tanto, el fenómeno de retracción de cámara esta presente ante cualquier rotación alrededor del eje óptico induciendo una translación a lo largo del mismo, la cual es innecesaria desde un punto de vista cartesiano y es crítica para el caso en que esta rotación sea de 180°, en donde la translación tenderá al infinito.

Corke y Hutchinson [26] diseñaron un esquema dividido para evitar estos problemas del CV2D clásico, ante grandes rotaciones respecto al eje óptico. Como solución, este esquema dividido desacopla los componentes relativos a la translación y rotación al eje óptico de la *matriz jacobiana imagen*, y calculan estos grados de libertad usando características simples de la imagen.

Para el cálculo de T_z , se define σ como la raíz cuadrada del área encerrada por el polígono formado por los puntos en la imagen (la cual puede ser calculada usando el método presentado por Wilf y Cunningham [36]), expresando T_z como:

$$T_z = \gamma T_z \cdot (\sigma^* - \sigma) \quad (2.23)$$

o una definición para producir una variación lineal con el movimiento a lo largo del eje óptico [29]:

$$T_z = \gamma T_z \cdot \ln(\sigma^* / \sigma) \quad (2.24)$$

donde σ^* indica el valor de σ en la imagen de referencia y γT_z es una ganancia escalar.

Por otro lado, para ω_z , se define la variable θ , como el ángulo entre el eje horizontal de la imagen y el segmento constituido al unir dos puntos, el cual debe cumplir que $0 \leq \theta \leq 2\pi$, lo que lleva a:

$$\omega_z = \gamma\omega_z \cdot (\theta^* - \theta) \quad (2.25)$$

donde θ^* indica el valor de θ en la imagen de referencia y $\gamma\omega_z$ es una ganancia escalar. Definiendo así un vector \mathbf{v}_{c_z} que contiene la translación y rotación a lo largo de eje óptico, tal que:

$$\mathbf{v}_{c_z} = [T_z \quad \omega_z]^T \quad (2.26)$$

mientras que la expresión que establece los movimientos relativos a los ejes X e Y es:

$$\mathbf{v}_{c_{xy}} = [T_x \quad T_y \quad \omega_x \quad \omega_y]^T = \Gamma \mathbf{J}_{xy}^+ ((\mathbf{c}^* - \mathbf{c}) - \mathbf{J}_z \mathbf{v}_{c_z}) \quad (2.27)$$

donde \mathbf{J}_{xy} y \mathbf{J}_z son matrices construidas con las componentes relativas a la translación y rotación respecto a los ejes X e Y (1^{ra}, 2^{da}, 4^{ta} y 5^{ta} columnas) y las componentes relativas a la translación y rotación al rededor del eje Z (3^{ra} y 6^{ta} columnas) de la *matriz jacobiana imagen*, respectivamente.

En la Fig. 2.4 se observa como para un error inicial de rotación pura de 180° a lo largo del eje óptico, utilizando el CV2D clásico las características se mueven en línea recta tendiendo a intersectarse en el centro de la imagen (Fig. 2.4(a)), mientras se produce una translación a lo largo del eje óptico (alejándose del objeto) que tenderá al infinito (Fig. 2.4(c)), sin modificarse el error de rotación inicial (Fig. 2.4(e)). Por otro lado, utilizado el esquema dividido se observa como las características se mueven en

arco (Fig. 2.4(b)) a la vez que no existe ninguna translación (Fig. 2.4(d)) y el error de rotación disminuye monotónicamente (Fig. 2.4(f)).

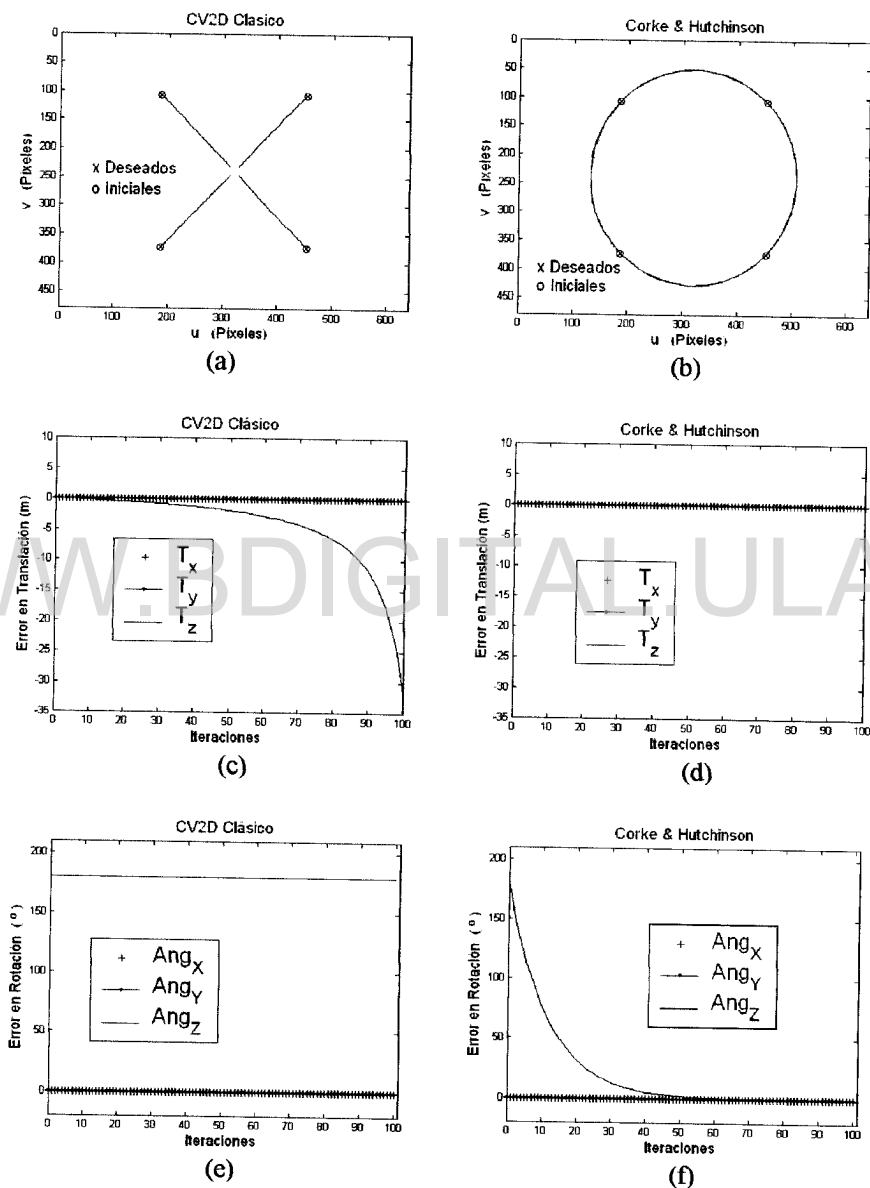


Figura 2.4 Comparación del comportamiento entre el CV2D clásico y el esquema dividido, para un error inicial de rotación pura a lo largo del eje óptico de 180° . Desplazamiento de las características en la imagen: (a) CV2D clásico, (b) Dividido. Error en translación: (c) CV2D clásico, (d) Dividido. Error en rotación: (e) CV2D clásico, (f) Dividido.

2.4.- Control visual invariante a los parámetros intrínsecos de la cámara

Este enfoque permite controlar la posición y orientación de la cámara con respecto a un objeto, de manera independiente de los parámetros intrínsecos de la cámara, ya que la función error es definida en un espacio invariante a los mismos [28], hasta el punto (en el caso de un objeto no plano) que se puede utilizar durante la tarea de control una cámara totalmente distinta a la utilizada para la captura de la imagen de referencia [37]. Quizás una característica de mayor interés práctico, propia de este esquema, es la posibilidad de variar el zoom de la cámara durante la ejecución de la tarea de control, pudiéndose imponer un control paralelo del zoom que evite que las características en la imagen se salgan del campo visual de la cámara. La idea básica de este enfoque es la construcción de una referencia en un espacio de proyección, la cual puede ser determinada si se conoce o no el modelo del objeto.

2.4.1.- Transformación si no se conoce el modelo de objeto

Supóngase que se posee una imagen de referencia capturada con el procedimiento *teach by showing*. Utilizando todos los puntos $\mathbf{p}_i = [u_i, v_i, 1]^T$, se calcula la siguiente matriz simétrica (3x3):

$$\mathbf{S}_p = \frac{1}{n} \sum_{i=1}^n \mathbf{p}_i \mathbf{p}_i^T = \frac{1}{n} \begin{bmatrix} \sum_{i=1}^n u_i^2 & \sum_{i=1}^n u_i v_i & \sum_{i=1}^n u_i \\ \sum_{i=1}^n u_i v_i & \sum_{i=1}^n v_i^2 & \sum_{i=1}^n v_i \\ \sum_{i=1}^n u_i & \sum_{i=1}^n v_i & \sum_{i=1}^n 1 \end{bmatrix} = \begin{bmatrix} \sigma_{11}^p & \sigma_{12}^p & \sigma_{13}^p \\ \sigma_{12}^p & \sigma_{22}^p & \sigma_{23}^p \\ \sigma_{13}^p & \sigma_{23}^p & 1 \end{bmatrix} \quad (2.28)$$

Los valores de esta matriz S_p tienen un significado simétrico simple, de hecho σ_{13}^p y σ_{23}^p representa las coordenadas del centroide del grupo de n puntos, mientras que σ_{11}^p , σ_{12}^p , σ_{22}^p , los momentos de segundo orden. Si los puntos observados son no colineales y el número de puntos (n) es mayor que 3 entonces la matriz S_p es simétrica positiva y puede ser reescrita utilizando la descomposición Cholesky, como [28]:

$$S_p = T_p T_p^T \quad (2.29)$$

donde T_p es una matriz (3x3) triangular superior no singular, definida como:

$$T_p = \begin{bmatrix} \sqrt{\sigma_{11}^p - (\sigma_{13}^p)^2} & \frac{(\sigma_{12}^p - \sigma_{13}^p \sigma_{23}^p)}{\sigma_{22}^p - (\sigma_{23}^p)^2} & \frac{(\sigma_{12}^p - \sigma_{13}^p \sigma_{23}^p)}{\sigma_{22}^p - (\sigma_{23}^p)^2} & \sigma_{13}^p \\ 0 & \sqrt{\sigma_{22}^p - (\sigma_{23}^p)^2} & \sigma_{23}^p & \\ 0 & 0 & 1 & \end{bmatrix} \quad (2.30)$$

Esta matriz define una transformación del espacio de proyección imagen $\mathcal{P} \in \mathbb{P}^2$ a un espacio $Q \in \mathbb{P}^2$. Cada punto $q_i \in Q$ es calculado como:

$$q_i = T_p^{-1} p_i = [a_i \quad b_i \quad 1]^T \quad (2.31)$$

2.4.2.- Transformación si el modelo es conocido

En este caso se conoce el modelo del objeto y para cada punto de interés es posible determinar:

$$m_i = \begin{bmatrix} x_i & y_i & 1 \\ z_i & z_i & 1 \end{bmatrix}^T = [\bar{x}_i \quad \bar{y}_i \quad 1]^T \quad (2.32)$$

Utilizando estos puntos, se puede calcular la siguiente matriz (3x3) simétrica:

$$\mathbf{S}_m = \frac{1}{n} \sum_{i=1}^n \mathbf{m}_i \mathbf{m}_i^T = \frac{1}{n} \begin{bmatrix} \sum_{i=1}^n \bar{x}_i^2 & \sum_{i=1}^n \bar{x}_i \bar{y}_i & \sum_{i=1}^n \bar{x}_i \\ \sum_{i=1}^n \bar{x}_i \bar{y}_i & \sum_{i=1}^n \bar{y}_i^2 & \sum_{i=1}^n \bar{y}_i \\ \sum_{i=1}^n \bar{x}_i & \sum_{i=1}^n \bar{y}_i & \sum_{i=1}^n 1 \end{bmatrix} = \begin{bmatrix} \sigma_{11}^m & \sigma_{12}^m & \sigma_{13}^m \\ \sigma_{12}^m & \sigma_{22}^m & \sigma_{23}^m \\ \sigma_{13}^m & \sigma_{23}^m & 1 \end{bmatrix} \quad (2.33)$$

Si los puntos son no colineales y $n > 3$, entonces la matriz \mathbf{S}_m es simétrica positiva y puede ser reescrita utilizando la descomposición Cholesky como:

$$\mathbf{S}_m = \mathbf{T}_m \mathbf{T}_m^T \quad (2.34)$$

donde \mathbf{T}_m es una matriz triangular superior de la misma forma que \mathbf{T}_p . De manera análoga, la matriz \mathbf{T}_m no singular, puede ser usada para realizar la transformación del espacio $M \in \mathbb{P}^2$ al espacio $Q \in \mathbb{P}^2$:

$$\mathbf{q}_i = \mathbf{T}_m^{-1} \mathbf{m}_i \quad (2.35)$$

Nótese que \mathbf{T}_m no depende de la matriz de parámetros intrínsecos \mathbf{K} , y por lo tanto, los puntos \mathbf{q}_i son independientes de estos.

2.4.3.- Equivalencia entre las dos transformaciones

Siendo $\mathbf{p}_i = \mathbf{K} \mathbf{m}_i$, la matriz \mathbf{S}_p puede ser escrita en función de \mathbf{S}_m y de los parámetros intrínsecos de la cámara \mathbf{K} :

$$\mathbf{S}_p = \frac{1}{n} \sum_{i=1}^n \mathbf{p}_i \mathbf{p}_i^T = \frac{1}{n} \sum_{i=1}^n \mathbf{K} \mathbf{m}_i \mathbf{m}_i^T \mathbf{K}^T = \mathbf{K} \left(\frac{1}{n} \sum_{i=1}^n \mathbf{m}_i \mathbf{m}_i^T \right) \mathbf{K}^T = \mathbf{K} \mathbf{S}_m \mathbf{K}^T \quad (2.36)$$

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

de las ecuaciones (2.29) , (2.34) y (2.36) se obtiene:

$$\mathbf{T}_p = \mathbf{K} \mathbf{T}_m \quad (2.37)$$

la equivalencia entre las ecuaciones (2.31) y (2.35) es demostrada a través:

$$\mathbf{q}_i = \mathbf{T}_p^{-1} \mathbf{p}_i = \mathbf{T}_m^{-1} \mathbf{K}^{-1} \mathbf{p}_i = \mathbf{T}_m^{-1} \mathbf{K}^{-1} \mathbf{K} \mathbf{m}_i = \mathbf{T}_m^{-1} \mathbf{m} \quad (2.38)$$

Por lo tanto, se obtiene los mismos puntos invariantes \mathbf{q}_i , del conocimiento de modelo del objeto o sin él. También se puede obtener la equivalencia en este espacio invariante para curvas no coplanares [28] o líneas rectas [38] como características.

Por otra parte, se puede controlar la localización de la cámara de un sistema cámara-en-mano minimizando una función error definida en el espacio invariante Q , utilizando una ley proporcional simple y un grupo genérico de puntos, tal que:

$$\mathbf{v}_c = \mathbf{\Gamma} \mathbf{J}_q^+ (\mathbf{q}^* - \mathbf{q}) \quad (2.39)$$

en donde \mathbf{q} es un vector $3n \times 1$ que agrupa los puntos actuales en el espacio Q , $\mathbf{q} = [\mathbf{q}_1^T \quad \mathbf{q}_2^T \quad \dots \quad \mathbf{q}_n^T]^T$ y \mathbf{q}^* los puntos de referencia. \mathbf{J}_q^+ es la pseudo-inversa de la *matriz jacobiana invariante* \mathbf{J}_q . Esta matriz depende de los parámetros intrínsecos de la cámara, de los valores de distribución de profundidad z_i , de las coordenadas de los puntos en la imagen \mathbf{p}_i y de los puntos invariantes \mathbf{q}_i [28], cuyo cálculo puede ser consultado en el Apéndice A. $\mathbf{\Gamma}$ es una matriz diagonal de dimensiones apropiadas con escalares positivos que sintoniza la tasa de convergencia. En este caso, $\mathbf{\Gamma} = \text{diag}(k_{Tx} \ k_{Ty} \ k_{Tz} \ k_{\omega_x} \ k_{\omega_y} \ k_{\omega_z})$.

Finalmente, si el objeto no es plano la cámara volverá a la posición y orientación de referencia cuando $\mathbf{q}^* = \mathbf{q}$, mientras el zoom puede ser controlado para evitar que las

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

características se salgan del campo visual de la cámara o para mejorar la resolución de la imagen, sin importar si la cámara usada durante la tarea es la misma que la usada en la captura de la imagen de referencia. En caso que el objeto sea plano, este esquema requiere que la distancia focal converja a la distancia focal usada en la captura de la imagen de referencia [39]. No siendo por tanto, un control invariante respecto a los parámetros intrínsecos de la cámara en todo el sentido de esta afirmación, ya que no podría intercambiarse la cámara por otra sin afectar la tarea, aunque se puede utilizar el zoom de la cámara para evitar que las características se salgan del campo visual (con la premisa que la distancia focal converja a la utilizada en el momento de captura de la imagen de referencia).

WWW.BDIGITAL.ULA.VE

CAPÍTULO 3

ESQUEMAS DE CONTROL VISUAL PROPUESTOS

En esta sección se presentan los esquemas propuestos para el uso de control visual basado en las características de la imagen permisible a variaciones del zoom de la cámara. La Propuesta 1 se enfoca al caso en que la distancia focal con la que se capturó la imagen de referencia sea un parámetro desconocido, mientras que, las Propuestas 2 y 3 la consideran conocida. Finalmente, se presenta una ley de control aplicable al zoom de la cámara, la cual puede ser usada por cualquiera de los esquemas propuestos, para evitar que las características se salgan del campo visual, a la vez que la distancia focal converja al valor usado en la captura de la imagen de referencia.

3.1.- Propuesta 1

El modelado de la variación de las características en la imagen, usualmente no toma en consideración el efecto de la variación del zoom de la cámara sobre las mismas, como se observa en la ecuación (2.6). En este trabajo, se ha tomado en consideración este efecto, introduciendo un término adicional $\left(\frac{\partial c}{\partial f} \frac{\partial f}{\partial t}\right)$ en dicha ecuación, obteniéndose:

$$\dot{c} = \frac{\partial c}{\partial r} \dot{r} + \frac{\partial c}{\partial a} + \frac{\partial c}{\partial f} \frac{\partial f}{\partial t} = \mathbf{J} \cdot \mathbf{v}_c + \frac{\partial c}{\partial a} + \mathbf{j}_f \cdot V_{f_1} \quad (3.1)$$

donde:

$\frac{\partial c}{\partial f} = \mathbf{j}_f$ es la variación de las características respecto a la variación de la distancia

focal f , y $\frac{\partial f}{\partial t} = V_{f_1}$ es la variación de la distancia focal respecto al tiempo.

El efecto de la variación de la distancia focal es independiente de la configuración cámara/robot utilizada, ya que las características tienden a moverse hacia el punto principal si la distancia focal disminuye, y tienden a alejarse radialmente si esta aumenta, indistintamente de cual sea la disposición del sistema de visión.

Así la ecuación (3.1) podría reescribirse considerando este efecto como un grado de libertad adicional de la *matriz jacobiana imagen*, tal que:

$$\dot{c} = \mathbf{J}_n \cdot \mathbf{v}_n + \frac{\partial c}{\partial t} \quad (3.2)$$

donde $\mathbf{v}_n = [Vf_1 \quad \mathbf{v}_c^T]^T$ es un vector 7x1 y $\mathbf{J}_n = [\mathbf{j}_f \quad \mathbf{J}]$.

Utilizado puntos como características y el modelo de cámara dado en la sección 2.1.2, \mathbf{j}_f se puede expresar como:

$$\mathbf{j}_f = [\mathbf{j}_{f_1}^T \quad \mathbf{j}_{f_2}^T \quad \dots \quad \mathbf{j}_{f_n}^T]^T \quad (3.3)$$

$$\mathbf{j}_{f_i}^T = \begin{bmatrix} \frac{x_i}{z_i} \\ r \frac{y_i}{z_i} \end{bmatrix} \quad (3.4)$$

donde considerando $r=1$, $s=0$, como una buena estimación para la mayoría de las cámaras modernas [30], se obtiene:

$$\mathbf{j}_{f_i} = \begin{bmatrix} \frac{u_i - u_0}{f} \\ \frac{v_i - v_0}{f} \end{bmatrix} \quad (3.5)$$

Es importante señalar, que a pesar de que la posición del punto principal puede variar si la distancia focal de la cámara cambia, para el cálculo de \mathbf{j}_{f_i} , este se ha considerado constante y ubicado en el centro de la imagen, lo que es usualmente razonable [35].

Con el uso de esta nueva matriz jacobiana \mathbf{J}_n , podemos definir un esquema de control utilizando una ley de tipo proporcional (y considerando que el objeto se mantiene estático), el cual, se denomina esquema de control visual basado en imagen de 7 grados de libertad (CV2D 7GDL):

$$\mathbf{v}_n = \Gamma \mathbf{J}_n^+ (\mathbf{c}^* - \mathbf{c}) \quad (3.6)$$

donde \mathbf{J}_n^+ representa la pseudo-inversa de la matriz \mathbf{J}_n .

Ahora, a diferencia del caso CV2D clásico, el algoritmo CV2D 7GDL no solo nos permite regular la posición y orientación de la cámara respecto al objeto, sino también regular el zoom de la misma, lo que es especialmente útil para casos en que la imagen inicial (y/o la actual) es captada con una distancia focal distinta a la utilizada en el momento de adquisición de la imagen de referencia, condicionándose la convergencia de este parámetro a este segundo valor, sin un conocimiento a priori del mismo.

En la Fig. 3.1 se observa como para un objeto no plano, sin existir un error de translación o rotación inicial, sino solo una variación de la distancia focal inicial con respecto al valor usado en la captura de la imagen de referencia, el CV2D clásico produce un movimiento en línea recta de las características en la imagen, manteniendo la distancia focal constante al no existir control sobre la misma (sin reducir el error), conservando el error de rotación en cero, pero presentando una

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

translación a lo largo de eje óptico que conlleva a un error de translación en estado estacionario.

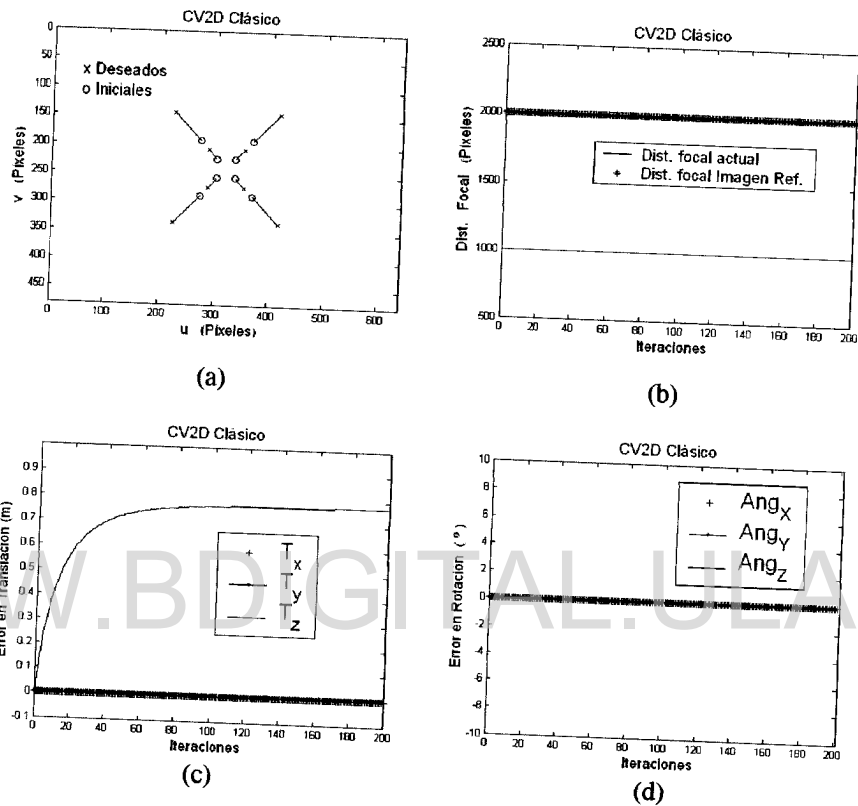


Figura 3.1 CV2D clásico, objeto no plano y distancia focal inicial distinta al valor con que se capturó la imagen de referencia. (a) Desplazamiento de las características en la imagen, (b) Variación de la distancia focal, (c) Error en translación, (d) Error en rotación.

Mientras, que utilizando el esquema de CV2D 7GDL, las características en la imagen se mueven de manera similar, pero la distancia focal converge a la usada durante la captura de la imagen de referencia (sin un conocimiento a priori de la misma), y los errores de rotación y de translación son mantenidos en cero (ver Fig. 3.2).

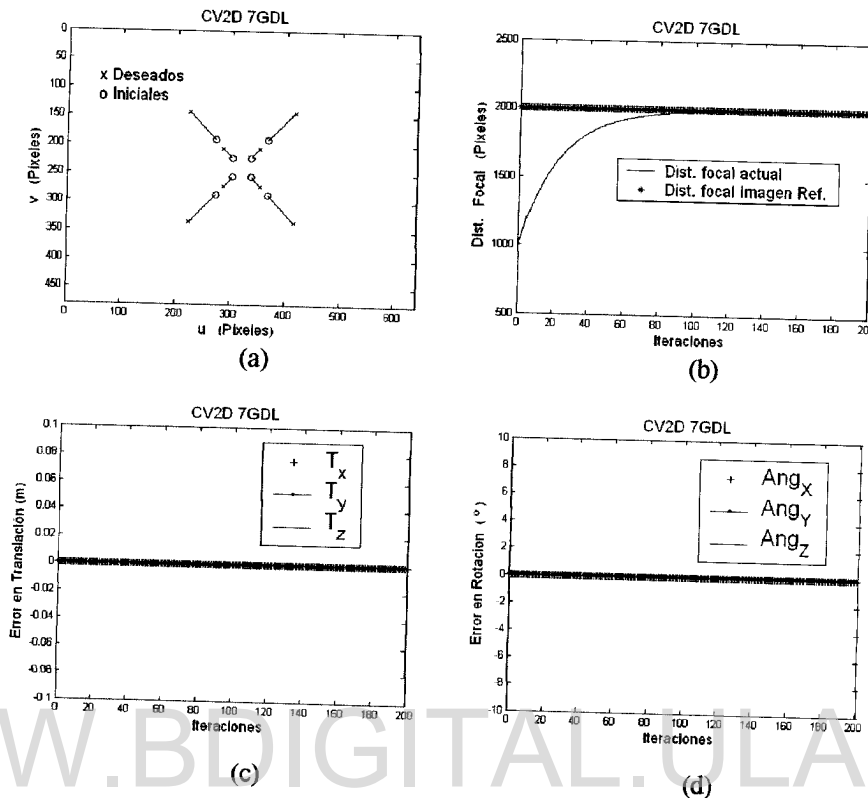


Figura 3.2 CV2D 7GDL, objeto no plano y distancia focal inicial distinta al valor con que se capturó la imagen de referencia. (a) Desplazamiento de las características en la imagen, (b) Variación de la distancia focal, (c) Error en translación, (d) Error en rotación.

Cuando utilizamos el CV2D 7GDL, una condición necesaria (mas no suficiente) para que la distancia focal al final de la tarea converja al valor usado en la captura de la imagen de referencia, sin un conocimiento a priori de la misma, es que la *matriz jacobiana imagen* \mathbf{J}_n sea de rango 7, lo cual se cumple en caso de objetos no coplanares y objetos planos que no se encuentren paralelos al plano imagen. La disminución del rango de la matriz \mathbf{J}_n (rango < 7) cuando el objeto es plano y paralelo al plano imagen se debe a que todos los valores de la profundidad de los puntos involucrados en el calculo de \mathbf{J}_n son iguales ($z_i = \text{constate}$), produciendo que la 1^{ra} columna de la matriz (relativa a las variación de la distancia focal) sea una combinación lineal de la 4^{ta} columna (relativa a las variaciones de translación a lo

largo del eje óptico). Esta combinación lineal entre columnas trae como consecuencia la disminución del rango la matriz J_n a 6. Por lo tanto, se recomienda principalmente el uso de este esquema con objetos no planos, ya que con objetos coplanares podría ocurrir que la cámara se posicione en la configuración anteriormente planteada a lo largo de la ejecución de la tarea de control, inclusive si la orientación inicial y/o de referencia son distintas de esta.

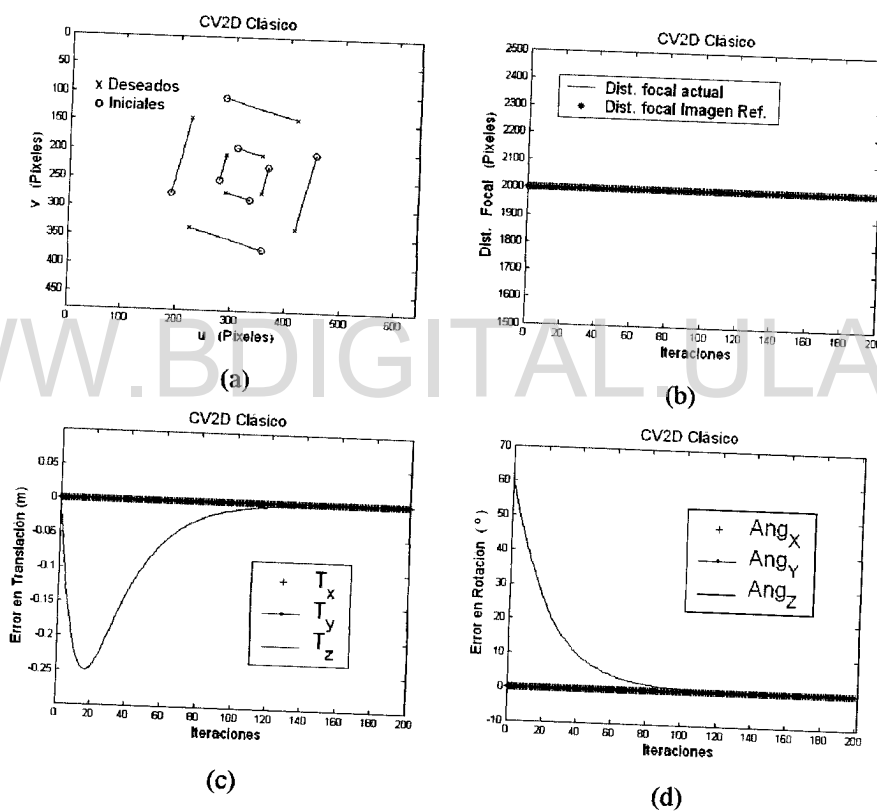


Figura 3.3 CV2D clásico, objeto no plano y distancia focal inicial igual al valor con que se capturó la imagen de referencia, error de rotación de 60° en el eje Z. (a) Desplazamiento de las características, (b) Variación de la distancia focal, (c) Error en translación, (d) Error en rotación.

Por otro lado, el CV2D 7GDL presenta una solución especial al fenómeno de retracción de cámara existente en el CV2D clásico para valores de rotación menores de 180° . En la Fig. 3.3 se observa que para un error inicial de rotación de 60° alrededor del eje óptico, utilizando el CV2D clásico, las características en la imagen

Licencia Creative Commons:

se mueven en línea recta hacia su configuración deseada, a la vez que se minimiza el error de rotación, presentándose una translación a lo largo del eje óptico, innecesaria desde un punto de vista cartesiano (fenómeno de retracción de cámara), sin manifestarse variación sobre la distancia focal.

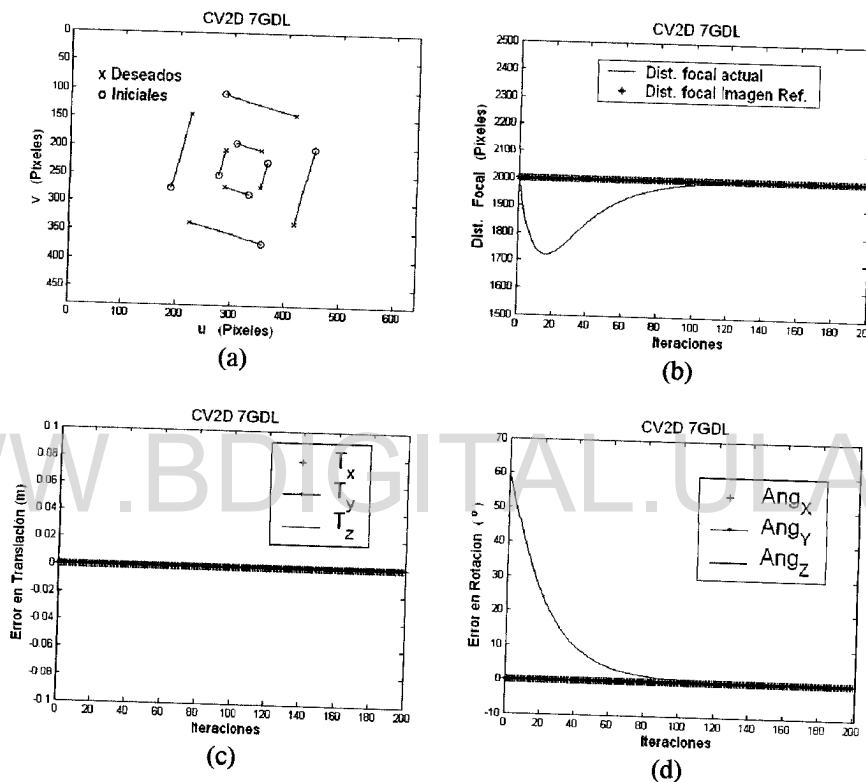


Figura 3.4 CV2D 7GDL, objeto no plano y distancia focal inicial igual al valor con que se capturó la imagen de referencia, error de rotación de 60° en el eje Z. (a) Desplazamiento de las características, (b) Variación de la distancia focal, (c) Error en translación, (d) Error en rotación.

Mientras que utilizando el CV2D 7GDL (Fig. 3.4), las características se mueven en línea recta, el error en rotación se minimiza sin presentarse ninguna translación, existiendo en este caso una retracción del valor de la distancia focal. Esta retracción focal es favorable con respecto a una retracción de cámara debido a que desde un punto de vista de seguridad del robot y su entorno, al no existir translación

(innecesaria), se disminuye la posibilidad que el robot colisione con otros robots u otros elementos que interactúan en su espacio de trabajo.

Sin embargo, para el caso en que el error inicial sea una rotación de 180° a lo largo del eje óptico, se observa en la Fig. 3.5 que utilizando el CV2D 7GDL, las características en la imagen tienden a intersectarse en el punto principal, mientras el error de rotación permanece constante y la distancia focal se satura en el valor mínimo que la cámara pueda proporcionar ($f_{\min.} = 300$ píxeles).

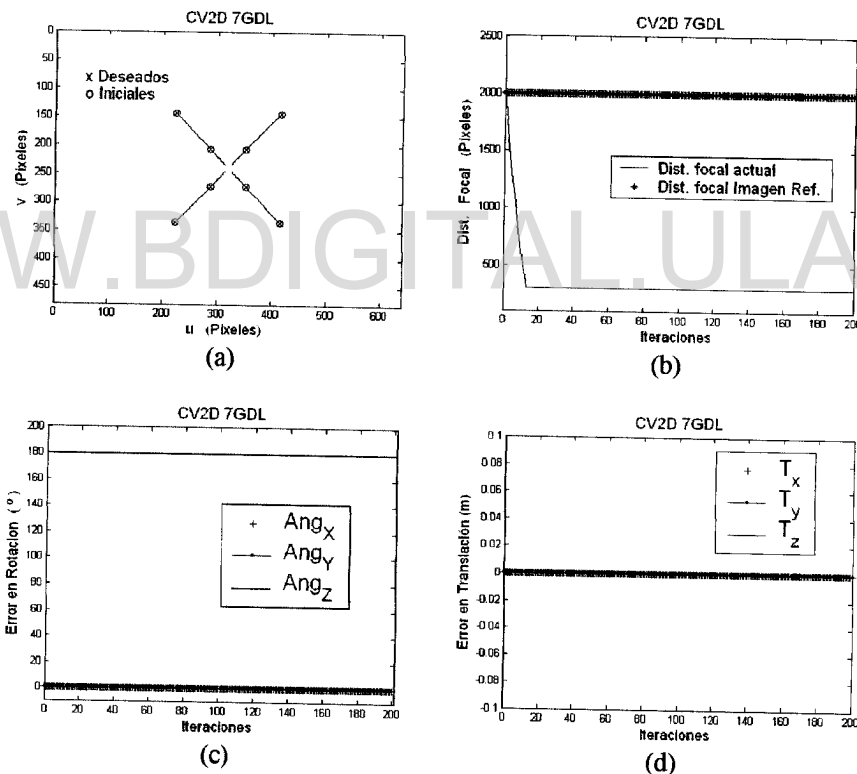


Figura 3.5 CV2D 7GDL, objeto no plano, y distancia focal inicial igual al valor con que se capturó la imagen de referencia, error de rotación de 180° en el eje Z. Distancia focal mínima soportada por la cámara = 300 píxeles. (a) Desplazamiento de las características, (b) Variación de la distancia focal, (c) Error en rotación, (d) Error en traducción.

Debido a este comportamiento, se ha desacoplado el grado de libertad correspondiente a la velocidad de rotación alrededor del eje óptico, denominando a este esquema como Propuesta 1:

$$\mathbf{v}_n' = \Gamma(\mathbf{J}_n')^+((\mathbf{c}^* - \mathbf{c}) - \mathbf{j}_{\omega_z} \omega_z) \quad (3.7)$$

En donde \mathbf{v}_n' es un vector con las mismas componentes que el vector \mathbf{v}_n sin incluir la velocidad de rotación respecto al eje óptico (ω_z). \mathbf{J}_n' es la matriz constituida por las primeras 6 columnas de la matriz \mathbf{J}_n y \mathbf{j}_{ω_z} su última columna (correspondiente a la velocidad rotacional respecto al eje óptico). Γ es una matriz diagonal de dimensiones apropiadas con escalares positivos que sintoniza la tasa de convergencia. En este caso, $\Gamma = \text{diag}(k_{v_{f1}}, k_{T_x}, k_{T_y}, k_{T_z}, k_{\omega_x}, k_{\omega_y})$.

Para el cálculo de ω_z , se define la variable θ como el ángulo entre el eje horizontal de la imagen y el segmento constituido al unir dos puntos, el cual debe cumplir que $0 \leq \theta \leq 2\pi$, lo que lleva a:

$$\omega_z = \gamma \omega_z \cdot (\theta^* - \theta) \quad (3.8)$$

donde θ^* indica el valor de θ en la imagen de referencia y $\gamma \omega_z$ es una ganancia escalar.

En la Fig. 3.6 se observa como utilizando la Propuesta 1, las características ya no se mueven en línea recta y la distancia focal varía convergiendo a su valor de referencia sin producirse ninguna translación, mientras el error de rotación es minimizado. De igual manera que en el caso de CV2D 7GDL, en el uso de la Propuesta 1 conviene

que el objeto observado no sea coplanar, evitándose que la 1^{ra} columna la matriz J_n' se convierta en una combinación lineal de su 4^{ta} columna, y por ende, deje de ser de rango completo (rango < 6 en este caso).

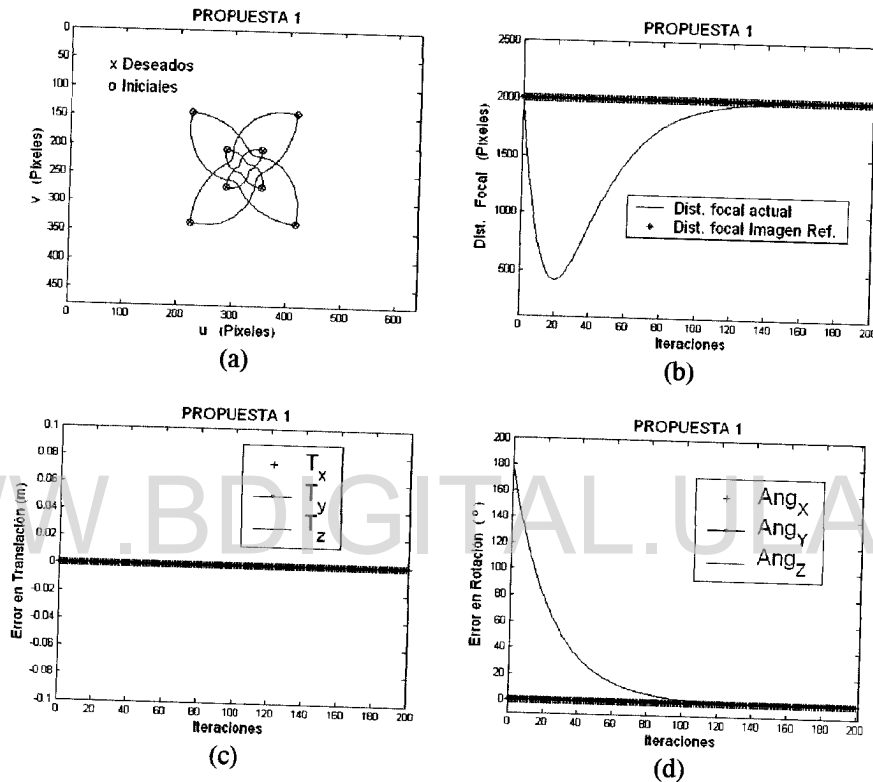


Figura 3.6. Propuesta 1, objeto no plano, y distancia focal inicial igual al valor con que se capturó la imagen de referencia, error de rotación en Z de 180°. (a) Desplazamiento de las características, (b) Variación de la distancia focal, (c) Error en translación, (d) Error en rotación.

3.2.- Propuesta 2

En esta sección se plantea un esquema de control considerando que la distancia focal usada en el momento de captura de la imagen de referencia es conocida a priori. El enfoque principal está basado en la modificación del esquema dividido presentado en la Sección 2.3 para permitirle variaciones del zoom de la cámara. Este nuevo

esquema, el cual denominamos Propuesta 2, se logra desacoplando de la matriz J_n los grados de libertad referidos a la variación de la distancia focal (1^{ra} columna), y los relativos a la translación y rotación respecto al eje óptico (4^{ta} y 7^{ta} columnas), calculado estos grados de libertad de manera independiente.

Para calcular la velocidad de translación a lo largo del eje óptico, T_z , se define σ' como la raíz cuadrada del cociente del doble del área del triángulo delimitado por tres puntos (no colineales) cualesquiera en la imagen y el valor de la distancia focal usada (conocida a priori). Recomendándose el uso de puntos que estén mas distanciados para mejorar la relación señal a ruido. Utilizando conceptos de algebra lineal para determinar el área del triángulo (y considerándose puntos como características en la imagen), obtenemos:

$$2(A_{3P}) = (2)(0.5) \begin{vmatrix} \mathbf{p}_1^T \\ \mathbf{p}_2^T \\ \mathbf{p}_3^T \end{vmatrix} = \begin{vmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_3 & v_3 & 1 \end{vmatrix} \quad (3.9)$$

$$\sigma' = \frac{\sqrt{2(A_{3P})}}{f} \quad (3.10)$$

siendo T_z calculado como:

$$T_z = \gamma T_z \cdot \ln(\sigma'^* / \sigma') \quad (3.11)$$

donde σ'^* se refiere al valor de esta variable con respecto a la imagen de referencia y γT_z es una ganancia escalar positiva que sintoniza la tasa de convergencia.

Mientras, ω_z se define como en la propuesta anterior. Una vez más, se recomienda utilizar los puntos que estén mas distanciados para mejorar la relación señal a ruido.

Adicionalmente, se plantea una ley de control proporcional para imponer que la distancia focal converja al valor usado en la captura de la imagen de referencia, tal que:

$$Vf_1 = \gamma_f \cdot (f^* - f) \quad (3.12)$$

donde f es la distancia focal usada para la captura de la imagen actual y f^* la distancia focal utilizada en el momento de captura de la imagen de referencia (la cual se supone conocida). Vf_1 es la velocidad de la distancia focal a ser aplicada para que f converja al valor de f^* , y γ_f es una ganancia escalar positiva que sintoniza la tasa de convergencia. De esta manera podemos definir el vector \mathbf{v}_{nz} , tal que:

$$\mathbf{v}_{nz} = [Vf_1 \quad T_z \quad \omega_z]^T \quad (3.13)$$

Finalmente, el esquema propuesto queda definido como:

$$\mathbf{v}_{nxt} = [T_x \quad T_y \quad \omega_x \quad \omega_y]^T = \mathbf{\Gamma} \cdot \mathbf{J}_{xy}^+ \cdot ((\mathbf{c}^* - \mathbf{c}) - (\mathbf{J}_{nz})\mathbf{v}_{nz}) \quad (3.14)$$

donde \mathbf{J}_{xy} y \mathbf{J}_{nz} son las matrices construidas con las componentes relativos a la translación y rotación respecto a los ejes X e Y (2^{da}, 3^{ra}, 5^{ta} y 6^{ta} columnas) y los relativos a la variación de la distancia focal, translación y rotación en el eje Z (1^{ra}, 4^{ta} y 7^{ta} columnas) de la matriz \mathbf{J}_n , respectivamente. $\mathbf{\Gamma}$ es una matriz diagonal de

dimensiones apropiadas con escalares positivos que sintoniza la tasa de convergencia.

En este caso, $\Gamma = \text{diag}(k_{Tx}, k_{Ty}, k_{\omega_x}, k_{\omega_y})$.

3.2.1.- Independencia del cálculo de la velocidad de translación y rotación a lo largo del eje óptico de las variaciones del zoom de la cámara

Utilizando el modelo de cámara enunciado en la Sección 2.1.2 y la ecuación (3.9) obtenemos:

$$2(A_{3P}) = \begin{vmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_3 & v_3 & 1 \end{vmatrix} = \begin{vmatrix} \bar{x}_1 f + \bar{y}_1 s + u_o & \bar{y}_1 r f + v_o & 1 \\ \bar{x}_2 f + \bar{y}_2 s + u_o & \bar{y}_2 r f + v_o & 1 \\ \bar{x}_3 f + \bar{y}_3 s + u_o & \bar{y}_3 r f + v_o & 1 \end{vmatrix} \quad (3.15)$$

$$2(A_{3P}) = f^2 r (\bar{x}_1 \bar{y}_2 - \bar{x}_1 \bar{y}_3 - \bar{x}_2 \bar{y}_1 + \bar{x}_2 \bar{y}_3 + \bar{x}_3 \bar{y}_1 - \bar{x}_3 \bar{y}_2) \quad (3.16)$$

$$\sigma' = \frac{\sqrt{2(A_{3P})}}{f} = \sqrt{r (\bar{x}_1 \bar{y}_2 - \bar{x}_1 \bar{y}_3 - \bar{x}_2 \bar{y}_1 + \bar{x}_2 \bar{y}_3 + \bar{x}_3 \bar{y}_1 - \bar{x}_3 \bar{y}_2)} \quad (3.17)$$

Obsérvese que el valor final de σ' no depende del valor de la distancia focal (f) ni de la ubicación del punto principal (u_o, v_o). Por ende, el valor obtenido de T_z a través de la ecuación (3.11), será independiente de la variación del zoom de la cámara.

Con respecto a ω_z , podemos definir el ángulo β , tal que:

$$\beta = \cos^{-1}(\mathbf{a} \bullet \mathbf{b}) / \|\mathbf{B}\| \quad (3.18)$$

donde el operador (\bullet) denota el producto escalar entre dos vectores, \mathbf{a} está definido como vector unitario horizontal $\mathbf{a} = [1 \ 0 \ 0]^T$, y \mathbf{b} como el vector formado por dos puntos en la imagen, tal que:

$$\mathbf{b} = [\mathbf{p}_2 - \mathbf{p}_1] = \begin{bmatrix} u_2 - u_1 \\ v_2 - v_1 \\ 1 - 1 \end{bmatrix} = \begin{bmatrix} u_2 - u_1 \\ v_2 - v_1 \\ 0 \end{bmatrix} \quad (3.19)$$

El valor de la variable β según la ecuación (3.18) estará en el rango $0 \leq \beta \leq \pi$. Para definir a θ dentro del rango $0 \leq \theta \leq 2\pi$, se establece la siguiente regla en función de la segunda componente del vector \mathbf{b} (ver Fig. 3.7):

$$\begin{cases} \theta = \beta & \text{si } (v_2 - v_1) \leq 0 \\ \theta = 2\pi - \beta & \text{si } (v_2 - v_1) > 0 \end{cases} \quad (3.20)$$

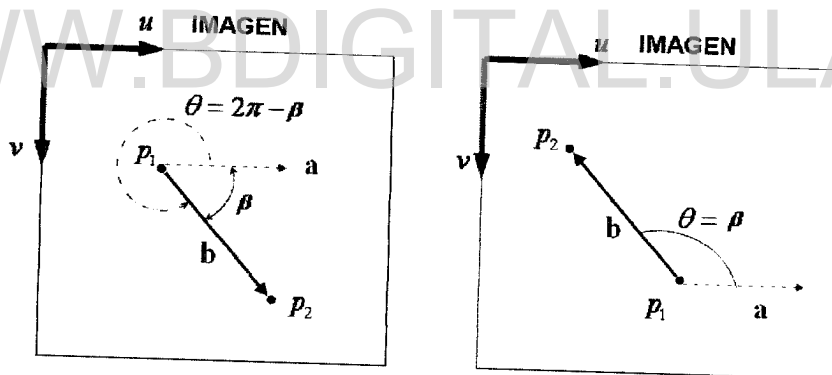


Figura 3.7 Diferencia del cálculo de θ en función de la orientación del vector B .

De nuevo haciendo uso del modelo de cámara presentado en la sección 2.1.2 y sustituyendo las expresiones para \mathbf{a} y \mathbf{b} en la ecuación (3.18), obtenemos:

$$\beta = \cos^{-1} \left(\frac{(u_2 - u_1)}{\sqrt{(u_2 - u_1)^2 + (v_2 - v_1)^2}} \right) \quad (3.21)$$

$$\beta = \cos^{-1} \left(\frac{\bar{x}_2 f + \bar{y}_2 s - \bar{x}_1 f - \bar{y}_1 s}{\sqrt{(\bar{x}_2 f + \bar{y}_2 s - \bar{x}_1 f - \bar{y}_1 s)^2 + (\bar{y}_2 r f - \bar{y}_1 r f)^2}} \right) \quad (3.22)$$

Ahora bien, si $s = 0$, encontramos que:

$$\beta = \cos^{-1} \left(\frac{(\bar{x}_2 - \bar{x}_1)}{\sqrt{(\bar{x}_2 - \bar{x}_1)^2 + (\bar{y}_2 r - \bar{y}_1 r)^2}} \right) \quad (3.23)$$

Se observa como la variable β (y por ende θ) no depende ni de la distancia focal ni de la ubicación del punto principal, para cámaras que puedan ser modeladas con $s = 0$, lo cual es considerado una buena estimación para las mayoría de las cámaras modernas [30]. Por lo tanto, en este caso (para $s = 0$), el cálculo de la velocidad rotacional ω_z es también independiente de las variaciones del zoom de la cámara.

3.3.- Propuesta 3

Como Propuesta 3 se postula una simple variación del Propuesta 2, en donde la velocidad rotacional ω_z no será calculada utilizando puntos en la imagen sino usando puntos invariantes calculados según lo enunciado en Sección 2.4, manteniendo el resto del esquema. Por ende, el cálculo del vector \mathbf{b} dado por la ecuación (3.19), es modificado usando puntos invariantes, tal que:

$$\mathbf{b} = [\mathbf{q}_2 - \mathbf{q}_1] \quad (3.24)$$

Estos puntos invariantes son teóricamente independientes de los parámetros intrínsecos de la cámara, y por lo tanto, la velocidad de rotación alrededor del eje

óptico será independiente de las variaciones del zoom de la cámara (inclusive para valores de $s \neq 0$). Igual que en el la sección anterior se recomienda utilizar los puntos que estén mas distanciados, y el valor de θ debe estar dentro del rango $0 \leq \theta \leq 2\pi$.

3.4.- Control de la distancia focal para evitar que las características se salgan del campo visual

Poder variar el zoom de la cámara durante la tarea de control de manera apropiada, presenta grandes ventajas, como por ejemplo, evitar que alguna(s) característica(s) se salga(n) del campo visual de la cámara. Para ello, se determina (utilizando puntos como características) la distancia del punto más cercano al borde de la imagen, tal que:

$$\delta = \min_i(u_i, v_i, u_{\max} - u_i, v_{\max} - v_i) \quad (3.25)$$

donde u_{\max} y v_{\max} representan las dimensiones de la imagen. Luego, utilizando una ley de control de tipo proporcional, se obtiene:

$$Vf_2 = \lambda_{vf_2} \cdot (\delta - \delta_{REF}) \quad (3.26)$$

donde λ_{vf_2} es una ganancia escalar positiva, Vf_2 la variación de la distancia focal requerida para minimizar la diferencia entre δ y δ_{REF} , siendo δ_{REF} un valor positivo igual o menor a la distancia mas corta que existe entre los puntos observados en la imagen de referencia y el borde de la misma. Nótese que si δ es menor que δ_{REF} (zona de riesgo de que las características se salgan del campo visual), Vf_2 tendrá un valor negativo, sugiriendo una disminución de la distancia focal (*zoom out*), lo cual,

aumentará el campo visual de la cámara y evitará que las características se salgan del mismo.

Finalmente, podemos definir la señal Vf a ser enviada por el controlador para controlar el zoom de la cámara, tal que:

$$Vf = \begin{cases} Vf_1 & \text{si } (\delta - \delta_{REF}) \geq 0 \\ Vf_2 & \text{si } (\delta - \delta_{REF}) < 0 \end{cases} \quad (3.27)$$

Vf asumirá el valor de Vf_1 para buscar que la que la distancia focal converja al valor usado en la captura de la imagen de referencia y adquirirá el valor de Vf_2 cuando exista riesgo de que alguna de las características se salga del campo visual. Es de suma importancia que el valor de δ_{REF} no sea un valor mayor que la distancia mas corta que existente entre los puntos observados en la imagen de referencia y el borde de la misma, condición que es necesaria (mas no suficiente) para que de la distancia focal converja al valor utilizado en la captura de la imagen de referencia.

Esta ley de control del zoom de la cámara puede ser usada con cualquiera de los esquemas control visual propuestos en las secciones anteriores.

CAPÍTULO 4

DESCRIPCIÓN DE LAS PRUEBAS PARA LA EVALUACIÓN DE LOS ESQUEMAS DE CONTROL VISUAL PROPUESTOS

El uso de esquemas de control visual en experimentos reales, se hace necesario resolver el problema de correspondencia entre las características de la imagen actual con la de referencia [40], y en el caso en particular de una cámara con distancia focal variable, entre imágenes de diferente resolución [41]. Las pruebas realizadas a través del uso de simulaciones, están enfocadas a analizar el comportamiento del Control Visual Invariante (CV invariante) junto a las distintas propuestas presentadas en el Capítulo 3 (todos estos esquemas permiten variación del zoom de la cámara), y por ende se ha considerado resuelto el problema de correspondencia de las características entre imágenes.

En todas la pruebas la *matriz jacobiana imagen* y/o la *matriz jacobiana invariante* es actualizada a cada iteración (a menos que se indique lo contrario), y la cámara es simulada según el modelo enunciado en el Capítulo 2, con una velocidad de captura de 30 cuadros por segundo y con capacidad de variar la distancia focal (f entre 300 píxeles y 3800 píxeles). A pesar que la posición del punto principal puede cambiar si la distancia focal varía, éste se ha considerado constante (y ubicado en el centro de la imagen), lo cual, es una estimación razonable [35].

Para la evaluación de las propuesta enunciadas en el Capítulo 3, se han planteado tres pruebas con distintos objetivos. La prueba 1 se concentra en la comparación del CV invariante con las Propuestas 1 y 2, en el caso en que el coeficiente angular de la cámara es nulo ($s = 0$). La prueba 2 se enfoca en comparar el comportamiento de la Propuesta 2 versus la 3, para el caso en que este coeficiente no puede ser modelado como cero ($s = 0,5$). Para estas dos primeras pruebas se hace uso de tareas de control

que ya han sido utilizadas para la evaluación de otros esquemas de control visual [29], siendo estas descritas mas adelante al igual que los índices de desempeño utilizados. Por último, la prueba 3 tiene como objetivo ilustrar el uso de la ley de control de la distancia focal para evitar que las características se salgan del campo visual de la cámara enunciada en el Capítulo anterior.

4.1.- Simulación del objeto

El objeto observado por la cámara ha sido simulado como un conjunto de 8 puntos no coplanares. Cuatro de estos puntos representan las esquinas de un cuadrado con lados de 0,05 m y los restantes cuatro forman otro de 0,15 m, paralelo al anterior y separado por una distancia de 0,05 m. La ubicación del objeto respecto a la posición y orientación de referencia de la cámara es tal, que ambos cuadrados son paralelos al plano imagen y el eje óptico pasa por el centro de los mismos. El cuadrado más pequeño se encuentra a 1,50 m de la cámara (a lo largo del eje óptico) y el mayor a 1,55 m (ver Fig. 4.1). Para esta posición/orientación relativa entre la cámara y el objeto se captura la imagen de referencia. La formación de esta imagen depende de esta posición/orientación relativa y de los parámetros intrínsecos de la cámara.

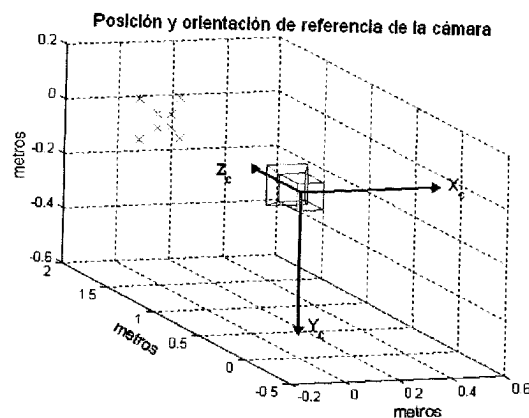


Figure 4.1 Posición y orientación de referencia de la cámara con respecto al objeto

El objeto ha sido configurado de la forma anteriormente descrita debido a que en un experimento real, las esquinas de los cuadrados pueden ser identificadas por algoritmos de detección de esquinas, lo que en las simulaciones es representado por la proyección de los puntos sobre el plano imagen.

4.2.- Descripción de las tareas de control

Para la comparación de los diferentes esquemas de control visual, se han seleccionado cuatro tareas de control, las cuales representan operaciones típicas en control visual y ya han sido utilizadas para evaluación de esquemas de control de este tipo [29]. Todas estas tareas consisten en minimizar un error inicial con respecto a la posición y/o orientación de referencia de la cámara.

WWW.BDIGITAL.ULA.VE

4.2.1.- Tarea 1: Rotación con respecto al eje óptico

La primera tarea de control, corresponde a minimizar un error inicial de rotación pura de la cámara alrededor del eje óptico (eje Z). La dificultad de esta tarea fue notada por primera vez por Chaumette [18]. En esta tarea se evalúa el desempeño de los esquema de control para rotaciones iniciales que van desde 30° a 210°, ejecutándose cada 5°.

4.2.2.- Tarea 2: Translación a lo largo del eje óptico

La segunda tarea de control corresponde a una translación inicial de la cámara a lo largo del eje óptico (eje Z). Los valores de translación están en el rango de -0,9 m (alejándose del objeto) hasta 1,3 m, (acercándose al objeto) con respecto a la posición

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

de referencia de la cámara, evaluándose a intervalos de 0,1 m. La translación a lo largo de un eje paralelo al plano imagen no ha sido considerada debido a que usualmente esta no presenta dificultades para los esquemas de control visual [29].

4.2.3.- Tarea 3: Rotación respecto a un eje perpendicular al eje óptico

En esta tarea se evalúa a los esquemas de control visual, respecto a una rotación inicial de la cámara alrededor del eje Y del sistema de coordenadas de la cámara. Los resultados encontrados son representativos para rotaciones respecto a cualquier eje perpendicular al eje óptico [29]. Es importante señalar, que a pesar que una pequeña rotación inicial sacaría a las características fuera del campo visual de una cámara real, a través de simulaciones, podemos permitir que el plano imagen sea de tamaño infinito, aumentando el rango de estudio. Por ende los valores de rotación se encuentran en el rango de 10° a 80° , incrementándose a intervalos de 5° .

4.2.4.- Tarea 4: Rotación del objeto

La última tarea a evaluar corresponde a una rotación inicial del objeto sobre un eje perpendicular al eje óptico. Para minimizar este error el sistema de control visual debe realizar labores de translación y rotación, evaluándose así un comportamiento general. El objeto es rotado en el rango de 10° a 80° en intervalos de 5° .

4.3.- Índices de desempeño

Se han seleccionado los siguientes índices para medir el desempeño de los distintos esquemas de control visual: error final en translación, error final en rotación, translación máxima inducida, recorrido realizado por la cámara, separación máxima de las características del punto principal y error final de la distancia focal.

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

4.3.1.- Error final en translación (EFT)

Este índice representa el error entre la posición final de la cámara y la posición de referencia. Es medido en metros (m) y como es obvio pensar, entre menor sea su valor, mejor será el desempeño del esquema de control visual. En caso que con el uso de algún esquema de control visual se presenten valores superiores a 1 m, los mismos son acotados a este valor, considerándose que el esquema de control ha fallado por completo para dichos puntos.

4.3.2.- Error final en rotación (EFR)

Con este índice se pretende evaluar cual es el error existente entre la orientación final de la cámara y la orientación referencia. Para ello, se cuantifica en grados (°) el ángulo entre los ejes del sistema de coordenadas de la cámara al final de la tarea y el sistema de coordenadas de la cámara en la orientación de referencia (para los ejes X, Y, Z). Registrándose la suma (del valor absoluto) de los ángulos.

4.3.3.- Translación máxima inducida (TMI)

A cada iteración del lazo de control se evalúa cual es la translación de la cámara con respecto a su posición de referencia, registrándose el valor máximo encontrado durante toda la ejecución de la tarea. Este índice es evaluado solo para las tareas 1 y 3, ya que para la ejecución de las mismas, desde un punto cartesiano, no se requiere de ninguna translación. Es cuantificado en metros (m) y si algún esquema de control visual presenta una translación máxima inducida mayor a 10 m, la misma es acotada a este valor, considerándose que el esquema de control visual ha fallado para dicho punto.

4.3.4.- Recorrido realizado por la cámara (RRC)

Este índice registra cual ha sido la translación necesaria (recorrido) para ejecutar la tarea de control. Solo se evalúa para las tareas 2 y 4, las cuales requieren de una translación para la ejecución de las mismas, pudiéndose entonces, determinar cual esquema de control visual las realiza con un menor recorrido. Si alguno de los esquemas de control visual a evaluar presenta un recorrido mayor a 35 m, el mismo es acotado a este valor, y se considera que el esquema de control ha fallado para dicho punto.

4.3.5.- Separación máxima de las características del punto principal (SMPP)

Este índice mide la tendencia de las características ha salirse del campo visual de la cámara. A cada interacción, la separación (en píxeles) entre la característica más distante al punto principal es medida, registrándose el valor máximo encontrado durante la ejecución de toda la tarea. Valores superiores a los 3000 píxeles se acotan a dicho valor, considerándose que el esquema de control ha fallado para tales puntos.

4.3.6.- Error final de la distancia focal (EFDF)

Este índice registra cual es el error existente (en valor absoluto) entre la distancia focal al final de la tarea de control y la utilizada en la captura de la imagen de referencia. El mismo es medido en píxeles.

4.4.- Condiciones de las Pruebas

4.4.1.- Prueba 1: Comparación de la Propuesta 1 y 2 con el CV Invariante

Esta prueba consiste en realizar una comparación del comportamiento de las Propuesta 1 y 2 versus el CV Invariante ante la ejecución de las distintas tareas de control enunciadas en los apartados anteriores, y bajo diferentes niveles de ruido en la imagen. La tarea de control se realiza durante 300 iteraciones.

4.4.1.1.- Condiciones de ruido

Para cada tarea de control, se evalúan los índices de desempeño bajo diferentes condiciones de ruido en la imagen. Este ruido es simulado adicionándole a las coordenadas de las características en la imagen una variable aleatoria. Esta variable será de distribución gaussiana y de media cero, siendo la varianza un parámetro entre 0 píxel^2 (ausencia de ruido) y 1 píxel^2 , en intervalos de $0,2 \text{ píxel}^2$. Ya que el ruido es un proceso aleatorio, la ejecución de cada tarea por parte de cada esquema de control visual en estudio ante todas las condiciones iniciales de la misma, se realizó 20 veces, para luego obtener índices de desempeño con valores promedio, suavizándose así el efecto de espurios. Es importante señalar, que las características de referencia (extraídas de la imagen de referencia) son también simuladas 20 veces para luego utilizar sus valores promedio.

4.4.1.2.- Condiciones iniciales de los parámetros intrínsecos

La cámara es simulada según el modelo enunciado en el Capítulo 2. La Tabla 4.1 muestra los parámetros intrínsecos utilizados para la imagen inicial y la imagen de referencia.

Tabla 4.1.- Parámetros intrínsecos de la cámara para la imagen inicial y la imagen de referencia utilizados en la Prueba 1.

Parámetros Intrínsecos	Imagen Inicial	Imagen de Referencia
u_o	320 píxeles	320 píxeles
v_o	240 píxeles	240 píxeles
r	1	1
s	0	0
f	500 píxeles	2000 píxeles

Los parámetros intrínsecos de la cámara son mantenidos constantes durante la tarea a excepción de la distancia focal, la cual, para la Propuesta 2 y el CV Invariante variará en función de la ecuación (3.12), y para la Propuesta 1 en función de la ecuación (3.7).

4.4.1.3.- Sintonización de los esquemas de control

Las ganancias de las Propuestas 1, 2 y el CV invariante fueron seleccionadas para minimizar un error relativo al correspondiente grado de libertad en aproximadamente 30 iteraciones, mientras que el movimiento de los restante grados de libertad se mantienen nulos. Esta sintonización es mas agresiva que la típicamente encontrada en esquemas de control visual, representando minimizar el error en 1 s de señal de video (30 cuadros por segundo) [29]. La Tabla 4.2 muestra las diferentes ganancias con las apropiadas unidades usadas con cada uno de los esquemas de control visual a evaluar.

Tabla 4.2.- Ganancias utilizadas por el CV invariante, Propuesta 1 y Propuesta 2 para la prueba 1.

Ganancias	CV invariante	Propuesta 1	Propuesta 2
k_{v/f_i}	N/A	5 s^{-1}	N/A
k_{T_x}	5 s^{-1}	5 s^{-1}	5 s^{-1}
k_{T_y}	5 s^{-1}	5 s^{-1}	5 s^{-1}
k_{T_z}	5 s^{-1}	5 s^{-1}	N/A
k_{ω_x}	5 rad/s	5 rad/s	5 rad/s
k_{ω_y}	5 rad/s	5 rad/s	5 rad/s
k_{ω_z}	5 rad/s	N/A	N/A
γ_f	5 s^{-1}	N/A	5 s^{-1}
γT_z	N/A	N/A	$7,5 \text{ m/s}$
$\gamma \omega_z$	N/A	5 s^{-1}	5 s^{-1}

4.4.1.4.- Prueba t-student

La prueba t-student (comparación de pares combinados) es aplicada a las diferencias de los índices de desempeño promedio por nivel de varianza del ruido entre el CV invariante y las Propuesta 1 y 2, denominando estas diferencias, Diferencias 1 y 2, respectivamente. El estimador \hat{t} es calculado para cada diferencia, con el fin de determinar si la mismas son estadísticamente significativas. Este estimador \hat{t} se define como:

$$\hat{t} = \frac{|\bar{d}\sqrt{n}|}{std} \quad (4.1)$$

donde, \bar{d} y std representan la media y la desviación estandar de la diferencia en cuestión (Diferencia 1 ó 2), y n , el tamaño de la misma (cantidad de niveles de ruido evaluados, $n = 6$).

Luego el estimador \hat{t} es comparado con un valor extraído de la tabla t -student, (ver apéndice B), evaluándose la siguiente expresión [42]:

$$\hat{t} > t_{n-1; \alpha/2} \quad (4.2)$$

Entonces, si se satisface (no satisface) la desigualdad dada por la ecuación (4.2), se dice que los dos métodos comparados en la diferencia, difirieren (no difieren) significativamente a un nivel de confianza de $100(1-\alpha)\%$.

Se ha seleccionado para α un valor de 0.05 (es decir, un nivel de confianza de 95%), lo que trae que:

$$t_{n-1; \alpha/2} = t_{5; 0,025} = 2,5710 \quad (4.3)$$

el cual es valor a compararse con \hat{t} .

4.4.2.- Prueba 2: Comparación de las Propuestas 2 y 3 para un coeficiente angular de la cámara distinto de cero

En el Capítulo 3, se observó que el cálculo de la velocidad rotacional ω_z para la Propuesta 2 es independiente de las variaciones del zoom, si el coeficiente angular, s , puede ser modelado como cero. El principal objetivo de esta prueba es evaluar el comportamiento de la Propuesta 2, cuando $s \neq 0$, determinándose a la vez si el uso de

puntos invariantes para el cálculo de la velocidad rotacional respecto al eje óptico (Propuesta 3) representa una mejora o no para el esquema. Para ello, se analiza el desempeño de las Propuestas 2 y 3 ante las tareas de control descritas en la Sección 4.2 (Tareas 1-4), ejecutándose bajo diferentes niveles de ruido y utilizándose para la simulación de la cámara un valor de s igual a 0,5.

4.4.2.1.- Condiciones de ruido

Las tareas de control son ejecutadas bajo diferentes niveles de ruido, el cual es simulado de manera análoga a la Prueba 1, repitiéndose también, todas las tareas y la captura de la imagen de referencia 20 veces para obtener valores promedio.

4.4.2.2.- Condiciones iniciales de los parámetros intrínsecos

La cámara será simulada según el modelo enunciado en el Capítulo 2. La Tabla 4.3 muestra los parámetros intrínsecos utilizados para la imagen inicial y la imagen de referencia.

Tabla 4.3.- Parámetros intrínsecos de la cámara para la imagen inicial y la imagen de referencia utilizados en la Prueba 2.

<i>Parámetros Intrínsecos</i>	<i>Imagen Inicial</i>	<i>Imagen de Referencia</i>
u_o	320 píxeles	320 píxeles
v_o	240 píxeles	240 píxeles
r	1	1
s	0.5	0.5
f	500 píxeles	2000 píxeles

Estos parámetros se mantienen constantes a excepción de la distancia focal, la cual varía en función de la ecuación (3.12) para ambos esquemas de control visual.

4.4.2.3.- Sintonización de los esquemas de control

Las ganancias usadas para las Propuestas 2 y 3 en esta prueba, son las mismas que las utilizadas por la Propuesta 2 en la prueba anterior (Prueba 1).

4.4.2.4.- Prueba t-student

Para esta prueba, el método de comparación de pares combinados es aplicado a la diferencia de los índices de desempeño promedio por nivel de ruido entre la Propuesta 2 y 3, (Diferencias 3), siguiendo el procedimiento explicado en la sección 4.4.1.3, seleccionando también un $\alpha = 0.05$.

WWW.BDIGITAL.ULA.VE

4.4.3.- Prueba 3: Uso del zoom de la cámara para evitar que las características se salgan del campo visual

Esta prueba estudia la ley de control de la distancia focal (enunciada en la sección 3.4) para evitar que las características se salgan del campo visual. Se seleccionó la Propuesta 2 como esquema de control visual a usar para comparar el comportamiento con y sin el uso esta ley, mientras la *matriz jacobiana imagen* es mantenida constante (calculada con los valores relativos a la imagen de referencia). La labor de control corresponde a minimizar un error inicial dado por una rotación de 25° con respecto al eje Y de la cámara (como en la Tarea 3).

4.4.3.1.- Condiciones de ruido

El sistema es simulado sin ruido en la imagen, con la intención de poder distinguir solamente el efecto de usar o no la ley de control de la distancia focal de la cámara para evitar que las características se salgan del campo visual.

4.4.3.2.- Condiciones iniciales de los parámetros intrínsecos

La cámara es simulada según el modelo enunciado en el Capítulo 2. La Tabla 4.4 muestra los parámetros intrínsecos utilizados para la imagen inicial y la imagen de referencia.

Tabla 4.4.- Parámetros intrínsecos de la cámara para la imagen inicial y la imagen de referencia utilizados en la Prueba 3.

<i>Parámetros Intrínsecos</i>	<i>Imagen Inicial</i>	<i>Imagen de Referencia</i>
u_o	320 píxeles	320 píxeles
v_o	240 píxeles	240 píxeles
r	1	1
s	0	0
f	500 píxeles	2000 píxeles

Estos parámetros son mantenidos constantes a excepción de la distancia focal, la cual variará en función de la ecuación (3.12) o (3.27) según el caso.

4.4.3.3.- Sintonización de los esquemas de control

Las ganancias en este caso, son las mismas utilizadas en la prueba 1. Adicionalmente, los valores de δ_{REF} , λ_{vf_2} requeridos para el uso de la ley de control se establecieron en 117 píxeles y $0,5 \text{ s}^{-1}$, respectivamente.

CAPÍTULO 5

RESULTADOS Y DISCUSIÓN

En este Capítulo se presentan y analizan los diferentes resultados encontrados para las pruebas enunciadas en el Capítulo anterior.

5.1.- Prueba 1: Comparación de la Propuesta 1 y 2 con el CV invariante

Al comenzar a simular las diferentes tareas, se observó que la sintonización seleccionada para la Propuesta 1 resultó ser muy agresiva y se modificó para obtener un mejor comportamiento. La Tabla 5.1 muestra las ganancias definitivas con las unidades apropiadas usadas para cada uno de los esquemas de control visual en evaluación.

Tabla 5.1.- Ganancias utilizadas por el CV invariante (modificada), Propuesta 1 y Propuesta 2 para la prueba 1.

Ganancias	CV invariante	Propuesta 1	Propuesta 2
k_{Vf_i}	N/A	1 s ⁻¹	N/A
k_{Tx}	5 s ⁻¹	1 s ⁻¹	5 s ⁻¹
k_{Ty}	5 s ⁻¹	1 s ⁻¹	5 s ⁻¹
k_{Tz}	5 s ⁻¹	1 s ⁻¹	N/A
k_{ω_x}	5 rad/s	5 rad/s	5 rad/s
k_{ω_y}	5 rad/s	5 rad/s	5 rad/s
k_{ω_z}	5 rad/s	N/A	N/A
γ_f	5 s ⁻¹	N/A	5 s ⁻¹
γT_z	N/A	N/A	7,5 m/s
$\gamma \omega_z$	N/A	5 s ⁻¹	5 s ⁻¹

5.1.1.- Tarea 1: Rotación a lo largo del eje óptico con distintos niveles de ruido

En esta tarea se hace uso de gráficas 3D y de tablas para mostrar el comportamiento de los índices de desempeño para cada uno de los esquemas de control visual estudiados. Dichas graficas muestran los resultados encontrados para todo el rango de rotaciones iniciales (30° a 210°), mientras que las tablas están referidas a rotaciones iniciales que van desde 30° a 170° . Esto se debe a que el CV invariante no esta diseñado para rotaciones iniciales de 180° respecto al eje óptico, y por ende, se ha obviado esta zona de operación para las evaluaciones cuantitativas (usando las tablas), realizándose una comparación mas justa.

5.1.1.1- Error final en translación

La Figura 5.1 muestra que en ausencia de ruido en la imagen (Varianza = 0), el error final de translación (EFT) para todos los esquemas de control visual estudiados, es nulo indistintamente del valor de rotación inicial. Concluyéndose que estos esquemas de control poseen un buen desempeño (con respecto al EFT y esta tarea de control), si no existe ruido en la imagen. Por otro lado, el CV invariante, Fig. 5.1(a), presenta valores de 1 m para rotaciones cercanas a los 180° y una varianza del ruido distinta de cero. También, a partir de una varianza de $0,6 \text{ píxel}^2$ (y rotación distinta de 180°), muchos puntos se encuentran con este mismo valor, lo que indica que para todos estos casos este esquema de control falla (acorde con lo mencionado en el Capítulo anterior para este índice).

La Propuesta 1, no presenta valores de EFT iguales a 1 m, alcanzando un valor máximo de $0,3272 \text{ m}$. Se aprecia además, que existe una tendencia de aumento de este índice a medida que el nivel de varianza de ruido aumenta (ver Fig. 5.1(b)). Esto puede corroborarse en la Tabla 5.2. Así mismo, esta tabla muestra que la Diferencia 1 es positiva para todos los niveles de varianza distintos de cero, indicando un mejor

comportamiento de la Propuesta 1 sobre el CV invariante (en el caso de que exista ruido en la imagen), aunque la prueba t para la comparación de estos dos métodos no muestra una diferencia estadísticamente significativa (pero si en un 90% de confianza $\hat{t} = 2,1619 > t_{5; 0,05} = 2,0150$).

Por otro lado, la Propuesta 2 posee un mejor comportamiento, mostrando un EFT máximo de 0,0158 m, a la vez, que se observa que este índice aumenta a medida que el nivel de varianza de ruido aumenta (ver Fig. 5.1(c)). La Tabla 5.2 también muestra que la Diferencia 2 es positiva y estadísticamente significativa, indicando que la Propuesta 2 posee un mejor desempeño que el CV invariante ($\hat{t} = 2,8729 > t_{5; 0,025} = 2,5710$).

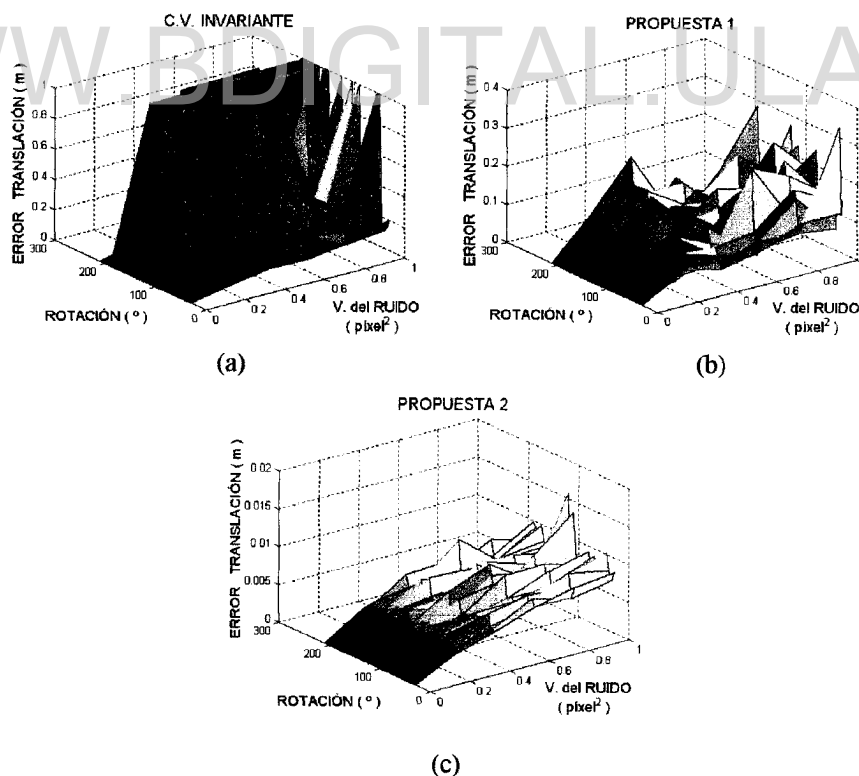


Figura 5.1 EFT para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta2.

Tabla 5.2 EFT promedio por nivel de varianza del ruido para la tarea 1 (desde 30° a 170°), Prueba 1 ($s = 0$).

Varianza del ruido (píxel ²)	CV Invariante (m)	Propuesta 1 (m)	Propuesta 2 (m)	Diferencia 1 (m)	Diferencia 2 (m)
0,0	0,0000	0,0000	0,0000	0,0000	0,0000
0,2	0,0636	0,0521	0,0037	0,0115	0,0599
0,4	0,1310	0,0697	0,0050	0,0613	0,1260
0,6	0,3167	0,0905	0,0063	0,2262	0,3104
0,8	0,1993	0,1186	0,0070	0,0807	0,1923
1,0	0,4270	0,1226	0,0081	0,3044	0,4189
MEDIA	0,1896	0,0756	0,0050	0,1140	0,1846
			\hat{t}	2,2619	2,8729

5.1.1.2.- Error final en rotación

Para un nivel de varianza de ruido igual a cero, todos los esquemas de control visual estudiados presentaron un error final en rotación (EFR) nulo indistintamente del valor de rotación inicial, excepto para el CV invariante usado con una rotación inicial de 180° (Fig. 5.2). Para esta rotación inicial, el CV invariante (Fig. 5.2(a)) mantiene el EFR en 180° debido a que las distintas componentes del cálculo de la velocidad de rotación respecto al eje óptico se anulan entre si. En cambio, si la varianza del ruido es distinta de cero (para esta misma rotación inicial), el EFR tiende a disminuir gracias a que el ruido en la imagen produce que las diferentes componentes ya no se anulen. Adicionalmente, se observa que para cualquier intensidad de ruido, el EFR es máximo para rotaciones iniciales cercanas a 180°. Por otro lado, para rotaciones iniciales menores de 170°, la Tabla 5.3, muestra una tendencia general a que el promedio de este EFR por nivel de ruido sea mayor a medida que se incrementa, excepto para una varianza de 0,8 píxel², la cual posee un promedio menor que el correspondiente a una varianza de 0,6 píxel².

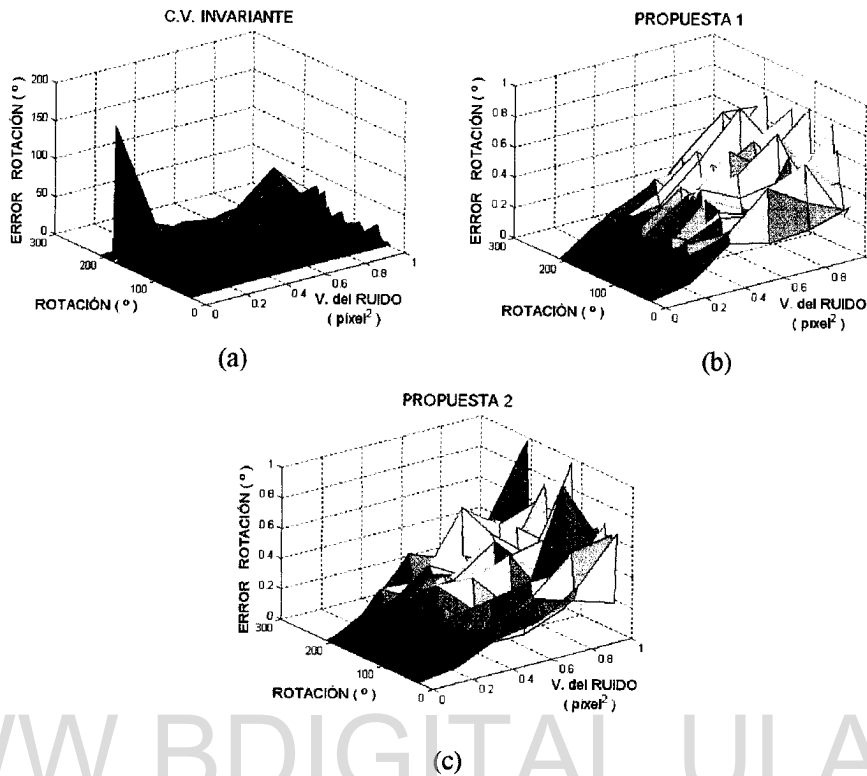


Figura 5.2 EFR para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla 5.3 EFR promedio por nivel de varianza del ruido para la tarea 1 (desde 30° a 170°), Prueba 1 ($s = 0$).

Varianza del ruido (píxel ²)	CV Invariante (°)	Propuesta 1 (°)	Propuesta 2 (°)	Diferencia 1 (°)	Diferencia 2 (°)
0,0	0,0000	0,0000	0,0000	0,0000	0,0000
0,2	0,9598	0,1907	0,2272	0,7691	0,7326
0,4	1,8029	0,2984	0,2692	1,5045	1,5337
0,6	3,8472	0,3215	0,3361	3,5257	3,5111
0,8	3,4199	0,4198	0,3536	3,0001	3,0663
1,0	6,4791	0,4011	0,4684	6,0780	6,0107
MEDIA	2,7515	0,2719	0,2758	2,4796	2,4757
			\hat{i}	2,7529	2,7696

La Fig. 5.2(b) muestra que la Propuesta 1 presenta valores de EFR menores a 1°, inclusive para una rotación inicial de 180°. Además, el EFR tiende a aumentar a medida que el nivel de ruido aumenta. Por otra parte, la Diferencia 1 (Tabla 5.3) es

positiva y estadísticamente significativa, indicando un mejor desempeño de la Propuesta 1 sobre el CV invariante ($\hat{t} = 2,7529 > t_{5; 0,025} = 2,5710$).

La Propuesta 2 presenta un comportamiento muy similar al de la Propuesta 1 (ver Fig. 5.2(c)). Por otro parte, la Tabla 5.3 muestra que el EFR promedio para cada nivel de varianza del ruido, aumenta a medida que el nivel de ruido se incrementa, así como también, que la Diferencia 2 es positiva y estadísticamente significativa, sugiriendo un desempeño superior de la Propuesta 2 sobre el CV invariante ($\hat{t} = 2,7696 > t_{5; 0,025} = 2,5710$).

5.1.1.3.- Translación máxima inducida durante la tarea

Para un nivel de varianza de ruido igual a cero (y para cualquier valor de rotación inicial), los tres esquemas de control visual estudiados presentaron una translación máxima inducida (TMI) nula (Fig. 5.3).

La Figura 5.3(a), muestra que para rotaciones cercanas a 180° y una varianza del ruido distinta de cero, la TMI adquiere el valor de 10 m (valor de saturación), sugiriendo falla del esquema de CV invariante, para dichas condiciones. Adicionalmente, a partir de una varianza de $0,6 \text{ píxel}^2$ (y valores de rotación distintos a 180°), muchos puntos presentan también el mismo comportamiento.

La Fig. 5.3(b), presenta de manera general un aumento del TMI a medida que el nivel de ruido se incrementa, aunque el valor máximo (2,2271 m) es alcanzado para una rotación de 135° y una varianza del ruido de $0,8 \text{ píxel}^2$. Por otra parte, para valores de rotación menores a 170° , la Diferencia 1 (Tabla 5.4) indica un mejor desempeño promedio por parte de la Propuesta 1 sobre el CV invariante, sin ser estadísticamente

significativo (aunque si lo es para un 90% de confianza, $\hat{t} = 2,0445 > t_{5; 0,05} = 2,0150$).

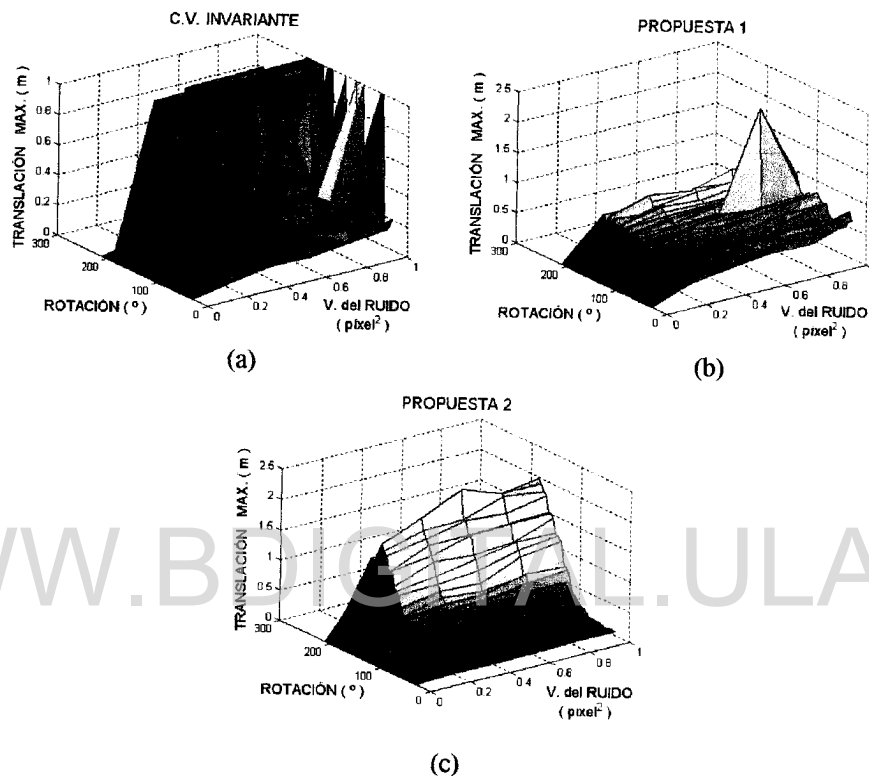


Figura 5.3 TMI para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla 5.4 TMI promedio por nivel de varianza del ruido para la tarea 1 (desde 30° a 170°), Prueba 1 ($s = 0$).

Varianza del ruido (pixel ²)	CV Invariante (m)	Propuesta 1 (m)	Propuesta 2 (m)	Diferencia 1 (m)	Diferencia 2 (m)
0,0	0,0000	0,0000	0,0000	0,0000	0,0000
0,2	0,3585	0,4225	0,3570	-0,0640	0,0015
0,4	0,8841	0,5040	0,4262	0,3801	0,4579
0,6	2,5907	0,5518	0,4550	2,0389	2,1357
0,8	1,5549	0,6596	0,5052	0,8953	1,0497
1,0	3,8067	0,6792	0,5310	3,1275	3,2757
MEDIA	1,5325	0,4695	0,3791	1,0630	1,1534
			\hat{t}	2,0445	2,1517

La Propuesta 2 presenta un aumento de la TMI a medida que el valor de rotación inicial se acerca a 180° y el nivel de la varianza del ruido aumenta (ver Fig. 5.3(c)). Adicionalmente, la Tabla 5.4 muestra un mejor desempeño de la Propuesta 2 ante el CV invariante (Diferencia 2 positiva) a pesar de no ser estadísticamente significativo (pero si a 90 % confianza $\hat{t} = 2,1517 > t_{5; 0,05} = 2,0150$).

5.1.1.4.- Separación máxima de las características del punto principal

En la Figura 5.4(a) se observa que el CV invariante mantiene la SMPP en un valor constante para el caso en que no existe ruido en la imagen. Cuando la varianza del ruido es distinta de cero, este índice aumenta a medida que la rotación inicial se acerca a 180° y el nivel del ruido aumenta, presentándose una gran cantidad de puntos en que el valor SMPP es igual a 3000 píxeles (valor de saturación), condición considerada como falla del esquema de control.

La Propuesta 1 no presenta ninguna SMPP equivalente a 3000 píxeles (ver Fig. 5.4(b)) y según se observa en la Tabla 5.5 el promedio de este índice aumenta a medida que el nivel de ruido aumenta. Adicionalmente, la Diferencia 1 por ser estadísticamente significativa (y positiva) sugiere que la Propuesta 1 posee mayor tendencia a mantener las características dentro del campo visual de la cámara que el CV invariante ($\hat{t} = 3,02 > t_{5; 0,025} = 2,5710$).

Por otra parte, la Fig. 5.4(c) muestra que para la Propuesta 2 la SMPP tiende a aumentar a medida que la rotación inicial se acerca a 180° , alcanzado el valor de 3000 píxeles para valores cercanos a esta rotación inicial y varianza del ruido igual o mayor a $0,6 \text{ píxel}^2$. En la Tabla 5.5 se observa que la Diferencia 2 es positiva y estadísticamente significativa, indicando que la Propuesta 2 posee un mejor desempeño que el CV invariante ($\hat{t} = 3,32 > t_{5; 0,025} = 2,5710$).

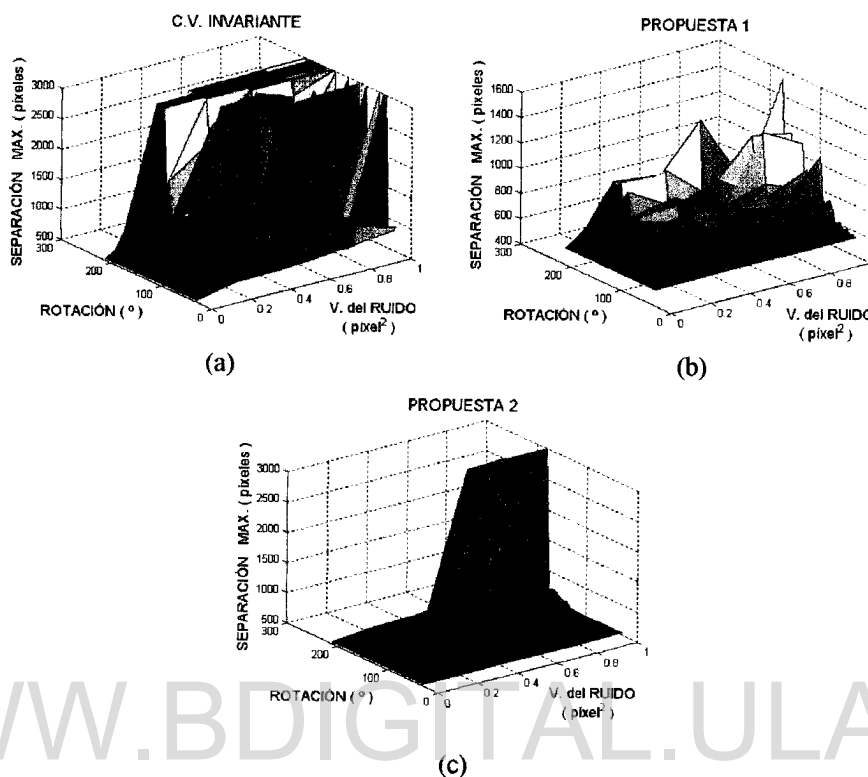


Figura 5.4 SMPP para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla 5.5 SMPP promedio por nivel de varianza del ruido para la tarea 1 (desde 30° a 170°), Prueba 1 ($s = 0$).

Varianza del ruido (píxe ²)	CV Invariante (píxel)	Propuesta 1 (píxel)	Propuesta 2 (píxel)	Diferencia 1 (píxel)	Diferencia 2 (píxel)
0,0	535,89	535,83	535,83	0,06	0,06
0,2	811,71	674,85	561,55	136,86	250,16
0,4	1038,61	688,21	566,07	350,40	472,54
0,6	1274,53	693,36	567,74	581,17	706,79
0,8	1379,09	744,52	570,96	634,57	808,13
1,0	1758,46	754,75	576,48	1003,71	1181,98
MEDIA	1133,05	681,92	563,11	451,13	569,94
			\hat{i}	3,02	3,32

5.1.1.5.- Error final de la distancia focal

En las Figuras 5.5(a) y 5.5(c) se observa que para el CV invariante y la Propuesta 2, el error final de la distancia focal (EFDF) con respecto a la utilizada para la captura de la imagen de referencia es nulo para cualquier valor de rotación inicial y nivel de ruido. Esto era de esperarse, ya que la ley control utilizada para la variación de la distancia focal para estos esquemas (ecuación (3.12)) no depende de las características de la imagen y por ende no es susceptible al ruido en ella. Por otro lado, la Propuesta 1, la cual fue diseñada para minimizar el error entre la distancia focal actual y la usada en la captura de la imagen de referencia (sin un conocimiento a priori de la misma) utilizando las características en la imagen, muestra que el EFDF es afectado por el ruido en la imagen, llegando a adquirir valores de 419,15 píxeles para altos niveles de ruido (ver Fig. 5.5(b)).

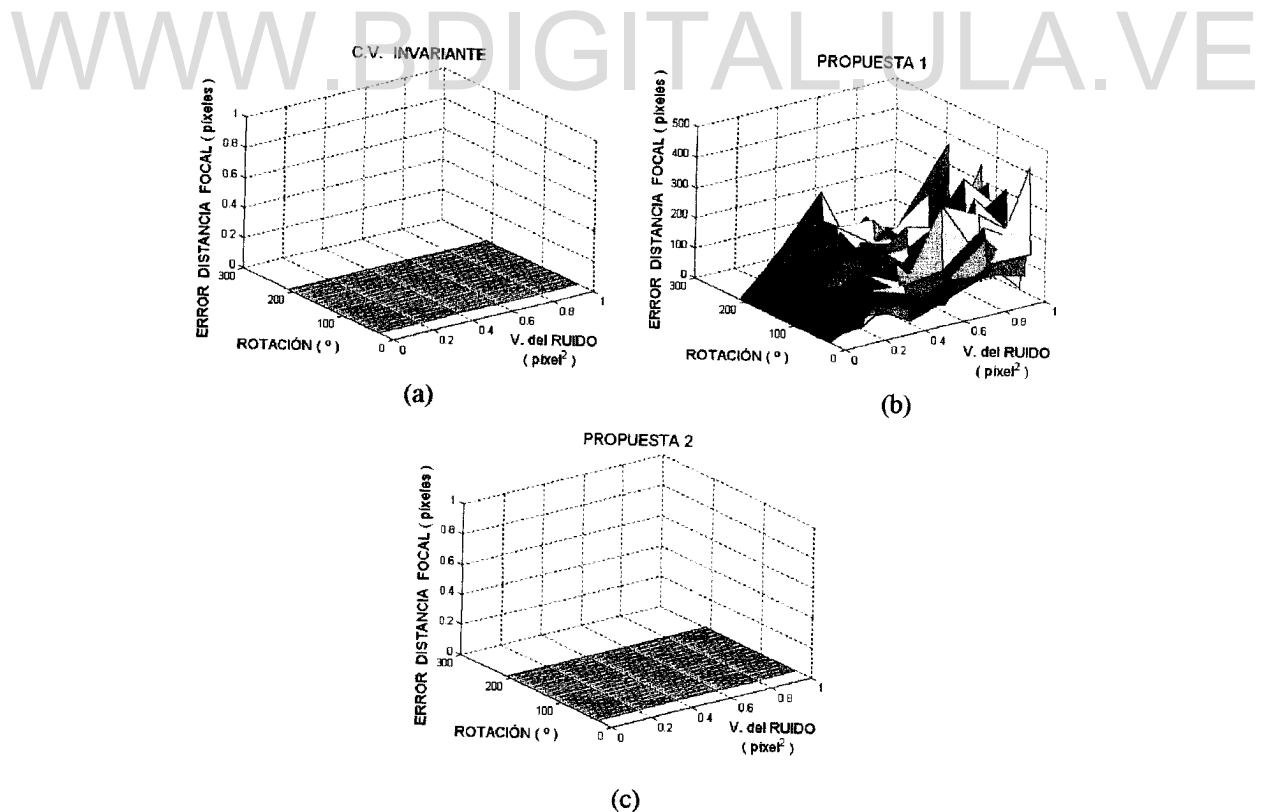


Figura 5.5 EFDF para la tarea 1 (de 30° a 210°) y diferentes niveles de ruido, prueba 1 ($s = 0$).
(a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

5.1.2.- Tarea 2.- Translación a lo largo del eje óptico con distintos niveles de ruido

Esta tarea fue analizada de manera similar a la tarea anterior. La Tabla 5.6, presenta un resumen del desempeño, basado en las medias de la Diferencia 1 y 2 y el estimador \hat{i} respectivo.

Tabla 5.6 Resumen de la Tarea 2 para los índices EFT, EFR, RRC y SMPP. Prueba 1.

Índices de Desempeño	Diferencia 1		Diferencia 2	
	Media	\hat{i}	Media	\hat{i}
EFT (m)	0,123	2,5102	0,2526	2,8570
EFR (°)	2,7083	2,7183	3,5386	2,7631
RRC (m)	12,5365	3,9686	15,8340	3,9023
SMPP (píxeles)	368,47	2,93	596,63	2,79

Se observa que la media de la Diferencia es positiva para todos los índices, lo que revela un mejor desempeño de la Propuesta 1 sobre el CV invariante. Esta diferencia es estadísticamente significativa, excepto para el EFT (aunque sigue siendo significativa con un 90% de confianza, $\hat{i} = 2,5102 > t_{5, 0,05} = 2,0150$). La media de la Diferencia 2 mostró ser estadísticamente significativa para todos los índices, manifestando la superioridad de la Propuesta 2 sobre el desempeño del algoritmo CV invariante. Finalmente, el EFDF aumenta para la Propuesta 1 a medida que el nivel de ruido lo hace, llegando a adquirir de 414,35 píxeles para los niveles más alto de ruido, mientras que para el CV invariante y la Propuesta 2, este índice es nulo independiente de la rotación inicial o nivel de ruido en la imagen. Esto demuestra que la efectividad de la ley control utilizada por estos esquemas para la variación de la distancia focal y su independencia de las características de la imagen y/o la translación inicial. Detalles de estos resultados pueden observarse en las Figuras y Tablas del Apéndice C.1.

5.1.3.- Tarea 3: Rotación respecto a un eje perpendicular al eje óptico con distintos niveles de ruido

Análogamente a la tarea anterior, la Tabla 5.7, muestra un resumen del desempeño de los esquemas de control estudiados, basado en las medias de la Diferencia 1 y 2 y el estimador \hat{t} para cada una de ellas.

Tabla 5.7 Resumen de la Tarea 3 para los índices EFT, EFR, RRC y SMPP. Prueba 1.

Índices de Desempeño	Diferencia 1		Diferencia 2	
	Media	\hat{t}	Media	\hat{t}
EFT (m)	0,0878	2,4749	0,1536	3,1540
EFR (°)	2,0788	2,9857	2,1068	3,0255
TMI (m)	0,7760	2,1364	0,5840	1,5431
SMPP (píxeles)	418,45	6,90	444,08	6,62

De nuevo, la media de la Diferencia 1 es positiva para todos los índices mostrados, lo que indica un mejor desempeño de la Propuesta 1 sobre el CV invariante. Esta diferencia es estadísticamente significativa para todos los índices, excepto para la TMI (aunque si lo es si se considera un 90% de confianza, $\hat{t} = 2,1364 > t_{5; 0,05} = 2,0150$). Por otro lado, la media de la Diferencia 2 es positiva y estadísticamente significativa para todos los índices, indicando un mejor desempeño de Propuesta 2 sobre CV invariante.

El EFDF se encontró nulo ante cualquier rotación inicial y cualquier nivel de ruido cuando se utilizó el CV invariante y la Propuesta 2. Esto se debe a que la ley de control usada por estos dos esquemas para la variación de la distancia focal, no depende de las características en la imagen como ya se ha mencionado con anterioridad. Mientras que utilizando la Propuesta 1, el EFDF llega a alcanzar 371,83 píxeles (para máximo nivel de ruido). Detalles de estos resultados pueden observarse en las Figuras y Tablas del Apéndice C.2.

5.1.4.- Tarea 4: Rotación del objeto con diferentes niveles de ruido

La Tabla 5.8 muestra que la media de la Diferencia 1 es negativa para todos los índices, sugiriendo un comportamiento superior por parte del CV invariante sobre la Propuesta 1. Mientras, que la media de la Diferencia 2 sugiere un mejor desempeño por parte de la Propuesta 2 (sobre el CV invariante) al ser positiva. Ambas diferencias son estadísticamente significativas ($\hat{t} > t_{5; 0,025} = 2,5710$).

Tabla 5.8 Resumen de la Tarea 4 para los índices EFT, EFR, RRC y SMPP. Prueba 1.

Índices de Desempeño	Diferencia 1		Diferencia 2	
	Media	\hat{t}	Media	\hat{t}
EFT (m)	-0,4721	6,4281	0,1662	3,0828
EFR (°)	-29,9363	5,7939	1,8744	2,7571
RRC (m)	-11,7240	4,4536	13,0322	4,6512
SMPP (píxeles)	-1236,08	7,99	552,06	3,20

Por otro lado, el EFDF es nulo para el CV invariante y la Propuesta 2 indistintamente del valor de la rotación inicial o del nivel de ruido como era de esperarse. Mientras que utilizando la Propuesta 1, el valor medio de este índice por nivel de intensidad de ruido fluctúa entre 205,43 píxeles y 280.03 píxeles, lo que manifiesta una respuesta poco satisfactoria por parte de este esquema para minimizar este error. Detalles de estos resultados pueden observarse en las Figuras y Tablas del Apéndice C.3.

5.2.- Prueba 2: Comparación de las Propuestas 2 y 3 para un coeficiente angular de la cámara distinto de cero.

5.2.1.- Tarea 1: Rotación a lo largo del eje óptico para distintos niveles de ruido

La Tabla 5.9 muestra que la media de la Diferencia 3 es negativa para todos los índices, y a pesar de ser estadísticamente significativa solo para la SMPP (aunque lo

es también para el resto de los índices si se considera un 90% de confianza, $t_{5; 0,05} = 2,0150$), esta indica un mejor desempeño de la Propuesta 2 sobre la 3. Detalles de estos resultados pueden ser observados en las Figuras y Tablas del Apéndice C.4.

Tabla 5.9 Resumen de la Tarea 1 para los índices EFT, EFR, RRC y SMPP. Prueba 2.

Índices de Desempeño	Diferencia 3	
	Media	\hat{t}
EFT (m)	-0,0248	2,2303
EFR (°)	-0,3120	2,2129
TMI (m)	-0,2131	2,2765
SMPP (píxeles)	-93,06	3,11

5.2.2.- Tarea 2: Translación a lo largo del eje óptico con distintos niveles de ruido

Al igual que en la tarea anterior, la media de la Diferencia 3 es negativa, pero estadísticamente significativa para todos los índices, excepto el EFR (Tabla 5.10). Sugiriendo que la Propuesta 2 presenta un mejor desempeño sobre la Propuesta 3 para esta tarea. Detalles de estos resultados pueden consultarse en las Figuras y Tablas del Apéndice C.5.

Tabla 5.10 Resumen de la Tarea 2 para los índices EFT, EFR, RRC y SMPP. Prueba 2.

Índices de Desempeño	Diferencia 3	
	Media	\hat{t}
EFT (m)	-0,0019	3,9428
EFR (°)	-0,0109	0,7890
RRC (m)	-0,2380	4,1837
SMPP (píxeles)	-0,25	3,70

5.2.3.- Tarea 3: Rotación respecto a un eje perpendicular al eje óptico con distintos niveles de ruido

En esta tarea, la media de la Diferencia 3 adquiere valores negativos y positivos (Tabla 5.11), indicando un mejor desempeño de la Propuesta 2 con respecto al EFT y la SMPP, y uno mejor para el EFR y el TMI por parte de la Propuesta 3. Siendo estas diferencias, estadísticamente significativas solo para el EFT y el TMI. Detalles sobre estos resultados pueden ser consultados en el Apéndice C.6.

Tabla 5.11 Resumen de la Tarea 3 para los índices EFT, EFR, RRC y SMPP. Prueba 2.

Índices de Desempeño	Diferencia 3	
	Media	\hat{t}
EFT (m)	-0,0015	3,7970
EFR (°)	0,0052	0,1588
TMI (m)	0,0097	56,4047
SMPP (píxeles)	-0,07	1,60

5.2.4.- Tarea 4: Rotación del objeto con distintos niveles de ruido

Como se observa en la Tabla 5.12, la media de la Diferencia 3 es negativa y estadísticamente significativa para todos los índices, excepto para el RRC, siendo positiva y no estadísticamente significativa. A pesar de ello se puede considerar un mejor desempeño general por parte de la Propuesta 2 sobre la 3. Detalles sobre estos resultados pueden ser consultados en el Apéndice C.7.

Tabla 5.12 Resumen de la Tarea 4 para los índices EFT, EFR, RRC y SMPP. Prueba 2.

Índices de Desempeño	Diferencia 3	
	Media	\hat{t}
EFT (m)	-0,0018	4,1498
EFR (°)	0,0108	0,2708
RRC (m)	-0,2345	4,1435
SMPP (píxeles)	-0,29	4,46

Licencia Creative Commons:

5.3.- Prueba 3: Uso del zoom de la cámara para evitar que las características se salgan del campo visual

La Figura 5.6 muestra los resultados para los casos en que no se utiliza (a), o se utiliza (b) la ley de control para la distancia focal diseñada para prevenir que las características de la imagen se salgan del campo visual de la cámara (representado por el polígono delimitado con líneas segmentadas). La Fig. 5.6(a) muestra como las trayectorias se salen de este campo visual cuando esta estrategia no se usa, mientras que la Fig. 5.6(b) demuestra la efectividad de la misma para mantener estas trayectorias dentro del campo visual de la cámara.

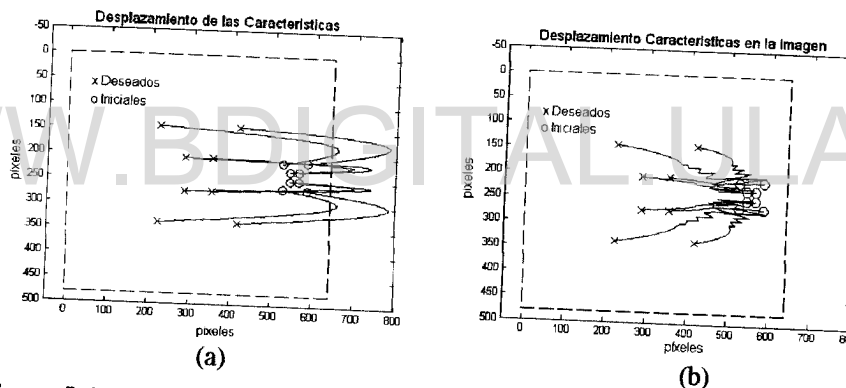


Figura 5.6 Desplazamiento de las características en la imagen. (a) Sin utilizar la ley control de la distancia focal para evitar que las características se salgan del campo visual, (b) Utilizando dicha la ley de control.

La Figura 5.7 muestra como en ambos casos la distancia focal converge al valor usado en la captura de la imagen de referencia. Es de hacer notar que para mantener las características dentro del campo visual la ley de control fuerza en ciertos intervalos a la distancia focal a mantenerse constante mientras que en otros disminuye (ver, Fig. 5.7(b)), aumentándose ligeramente las cantidad de iteraciones necesarias para converger al valor de la distancia focal usada en la captura de la imagen de referencia.

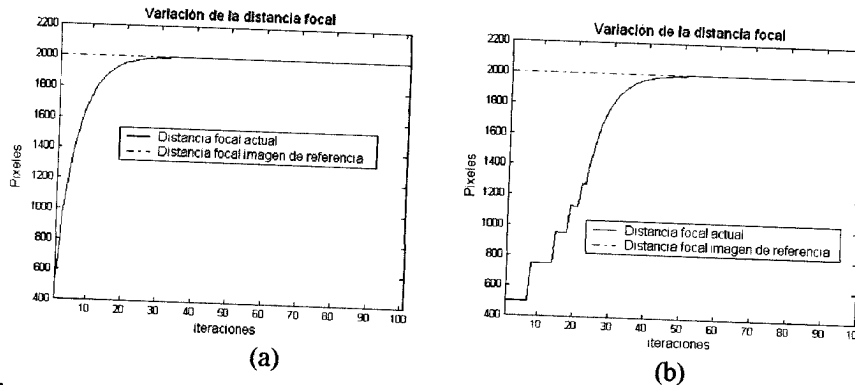


Figura 5.7 Variación de la distancia focal. (a) Sin utilizar la ley control de la distancia focal para evitar que las características se salgan del campo, (b) Utilizando dicha ley de control.

En las Figuras 5.8 y 5.9 se muestran como los errores en translación y en rotación convergen a cero para ambos casos con el evidente retardo en convergencia para cuando se regula la distancia focal para evitar que las características se salgan del campo (ver Fig. 5.8(b) y 5.9(b)). En particular, la Fig. 5.8 muestra además, que el error de translación se induce durante la tarea de control, ya que este es nulo al inicio de la misma.

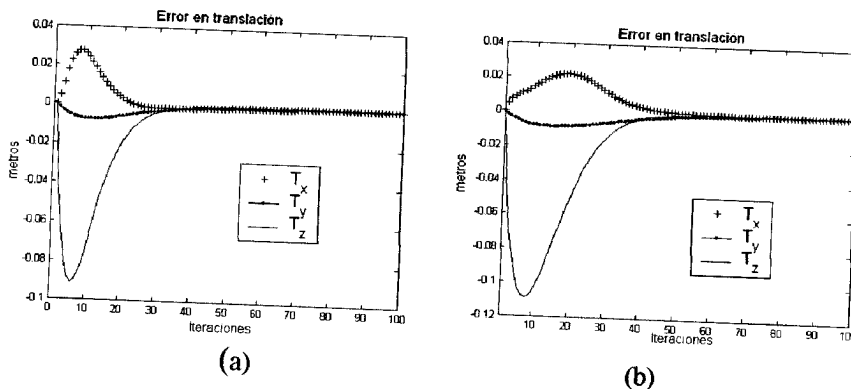


Figura 5.8 Error en translación. (a) Sin utilizar la ley control de la distancia focal para evitar que las características se salgan del campo visual (b) Utilizando dicha ley de control.

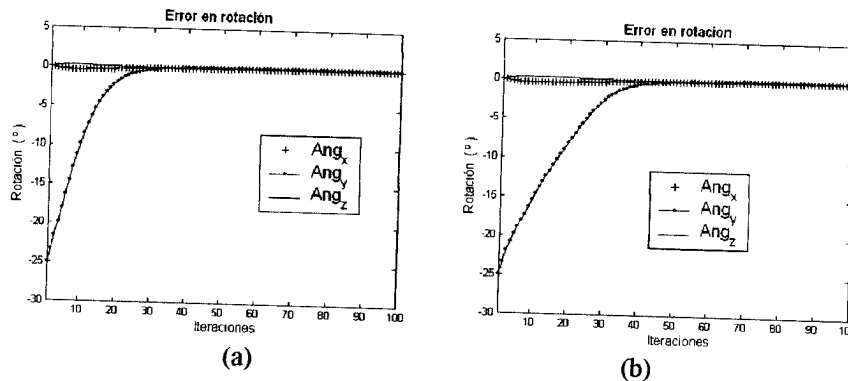


Figura 5.9 Error en rotación. (a) Sin utilizar la ley control de la distancia focal para evitar que las características se salgan del campo visual (b) Utilizando dicha ley de control.

Cabe destacar que en la ejecución de la tarea de control durante esta prueba, las trayectorias, para el caso en que no se utiliza la regulación para prevenir que se salgan las características del campo visual de la cámara (Fig. 5.6(a)), no se ven afectadas por la salida de estas del marco de referencia. Esto se debe a que se consideró para efecto de cálculos y simulación, que el plano imagen es de tamaño infinito; ya que solo se deseaba determinar si las características se mantenían o no dentro del campo visual. Los efectos que podría traer el hecho que una o más características se salgan del campo visual han sido estudiados por García [43], e incluso algunos autores consideran que si esto llegara a ocurrir, como medida de seguridad debería detenerse la tarea de control [21], lo que resalta la importancia del uso de ley de control de la distancia focal para evitar que las características se salgan del campo visual.

CAPÍTULO 6

CONCLUSIONES

Este trabajo muestra que los tres algoritmos propuestos para el control visual basado en el uso de las características en la imagen, realmente permiten la variación del zoom de la cámara, lo que hasta el momento según nuestro conocimiento a excepción del Control Visual Invariante no había sido reportado en la literatura. Estos algoritmos tienen como condición que al final de la tarea de control la distancia focal converge al valor utilizado en el momento en que se capturó la imagen de referencia.

La Propuesta 1 mostró un desempeño similar al Control Visual Invariante en ausencia de ruido en la imagen, pero a medida que el nivel de ruido aumenta, la Propuesta 1 presenta un mejor desempeño excepto para una rotación inicial del objeto observado (Tarea 4), mostrando un mal comportamiento, en especial para rotaciones del objeto mayores a 25° , y no logra que la distancia focal converja al valor usado en la captura de la imagen de referencia. Esto lo descalifica para controlar robots que requieran realizar tareas que involucren una translación y rotación general (necesarias para minimizar el error producido por una rotación inicial del objeto observado).

La Propuesta 2 mostró también un desempeño similar al Control Visual Invariante en ausencia de ruido, pero uno superior en el caso contrario. Esto se debe a que el Control Invariante requiere de cálculos de proyección de las características de la imagen sobre el espacio invariante, que además de consumir cierto tiempo se pueden ver afectados por la presencia del ruido. La Propuesta 2 utiliza las características directamente, evitándose la propagación de error y disminuyendo el costo computacional. Así, si la distancia focal utilizada en la imagen de referencia es conocida, la Propuesta 2 presentará un mejor desempeño. Cabe destacar, que con objetos no planos, teóricamente el Control Visual Invariante puede llevar a cabo la

tarea de control, incluso si la cámara utilizada durante la tarea es distinta a la usada en la captura de la imagen de referencia. Esto puede ser de poco interés práctico a menos que se piense en el envejecimiento del lente de la cámara (cambios de los parámetros intrínsecos) o en el reemplazo de la misma por otra de tipo distinto debido a rotura. En caso de un reemplazo de este tipo para sistemas controlados mediante la Propuesta 2, se requeriría, simplemente, de una nueva captura de la imagen de referencia. Sin embargo, con objetos coplanares, esta pequeña ventaja se desvanece y el Control Visual Invariante necesita que la distancia focal al final de la tarea de control converja a la usada en la imagen de referencia y que ambas cámaras sean la misma (la usada en la captura de la imagen de referencia y durante la tarea), esperándose análogamente (ya que las pruebas realizadas fueron simuladas para objetos no planos), que este esquema control visual sea más susceptible al ruido que la Propuesta 2.

Adicionalmente, la Propuesta 2 aunque no mostró un desempeño ampliamente mejor que la Propuesta 3, ante situaciones en que el coeficiente angular s de la cámara (*skew coefficient*) no pueda ser considerado nulo (prueba 2), ella es preferible ya que requiere menor gasto computacional y es ligeramente menos susceptible al ruido. De esta manera, la Propuesta 2 presenta un buen desempeño a pesar que el cálculo de la velocidad rotacional respecto al eje óptico se hace dependiente del valor de la distancia focal cuando el coeficiente angular s es distinto de cero. Esto puede deberse a que usando este algoritmo eventualmente la distancia focal converge al valor utilizado para capturar la imagen de referencia.

Finalmente, la ley propuesta para mantener las características dentro del campo visual de la cámara, mostró ser efectiva (prueba 3), además de ser aplicable en cualquiera de los algoritmos planteados. Este resultado es de gran importancia ya que el mantener las características dentro del campo visual de la cámara pronostica el éxito de la tarea de control.

En conclusión los resultados obtenidos para los algoritmos propuestos mediante este estudio basado en simulaciones son prometedores, por lo que recomendamos para futuras investigaciones se estudie de manera experimental el desempeño de los mismos. Además, para fortalecer estos resultados exhortamos: a estudiar estos esquemas para el caso en que las características en la imagen sean distintas a puntos, y a evaluar su desempeño utilizando leyes de control avanzado.

WWW.BDIGITAL.ULA.VE

REFERENCIAS BIBLIOGRÁFICAS

- [1] Kelly, R., Shirkey, P. y Spong, M.W. "Fixed Camera Visual Servo Control for Planar Robots". *IEEE Intl. Conf. on Robotics and Automation*, Minneapolis, MN, April, 1996.
- [2] Fujimoto, H. "Visual Servoing of 6 DOF Manipulator by Multirate Control with Depth Identification". *Proceedings of the 42nd IEEE Conference on Decision and Control*, Maui, Hawaii USA, December 2003.
- [3] Allen, P., Timcenko, A., Yoshimi, B. y Mitchelman, P. "Automated tracking and grasping of a moving object with a robot hand eye system". *IEEE Transactions on Robotics and Automation* 9(2), 125-165, 1993.
- [4] Hutchinson, S., Hager, G. y Corke, P. "A tutorial on visual servo control". *IEEE Transactions on Robotics and Automation* 12(5), 651-670, 1996.
- [5] Wilson, W., Hulls, W. y Bell, G., S. "Relative end-effector control using cartesian position-based visual servoing". *IEEE Trans. on Robotics and Automation*, 12(5), 684-696, October 1996.
- [6] Gangloff, J. y De Mathelin, M. "Visual servoing of a 6 DOF manipulator for unknown 3D profile following". *IEEE Trans. Robotics and Automation*, vol. 18, no 4, pages 511-520, August 2002.
- [7] Hashimoto, K., Kimoto, T., Ebine, T. y Kimura, H. "Manipulator Control with Image-Based Visual Servo". *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*.
- [8] Tsay, T. I., Lee, M. K., G. L. Wang, Hsu, M. S. y Lai, C. H. "Robotic Eye/Arm Coordination via Visual Servoing". *Proceedings of the 2002 IEEE International Conference on Control Applications*. September 18-20, 2002. Glasglow, Scotland.
- [9] Corke, Peter. "Visual control of robot manipulators: A review". 1994.
- [10] Espiau, B., Chaumette, F. y Rives, P. "A new approach to visual servoing in robotics". *IEEE Trans. On Robotics and Automation*, 8 (3) pp 313-326, June 1992.

- [11] Miura, K., Hashimoto, K., Gangloff, J. y De Mathelin, M. "Visual Servoing without Jacobian Using Modified Simplex Optimization". *IEEE International Conference on Robotics and Automation*, Barcelona, Spain, April 2005.
- [12] Kragic, D. y Christensen, L. S. "Survey on Visual Servoing for Manipulation" *Tech. Rep. ISRN KTH/NA/P--02/01--SE*, Jan. 2002., CVAP259.
- [13] Gangloff J. A. y de Mathelin, M. F. "High speed visual servoing of a 6 DOF manipulator using multivariable predictive control". *Advanced Robotics*, Vol. 17, no 10, pp 993-1021, 20 December 2003.
- [14] Corke, P. y Good, M. "Dynamics effects in visual closed-loop systems". *IEEE Transactions on Robotics and Automation* 12(5), 671-696, 1996
- [15] Hashimoto, K., Ebine, T. y Kimura, H. "Visual Servoing with Hand-Eye Manipulator-Optimal Control Approach". *IEEE Transactions on Robotics and Automation*, 12(5), 766-774, 1996.
- [16] Astolfi, A., Hsu, L., Netto, M. y Ortega, R. "Two Solutions to the Adaptive Visual Servoing Problem". *IEEE Transactions on Robotics and Automation*, 18(3), 387-392, 2002.
- [17] Zanne, P., Morel, G., y Plestaqn, F. "Robust vision based 3D trajectory tracking using sliding model control". *IEEE Int. Conf. Robotics and Automation*, pp. 2088-2093, San Francisco, CA, 2000.
- [18] Chaumette, F. "Potential problems of stability and convergence in image-based and position-based visual servoing". In D. Kriegman, G. Hager, and A. Morse, eds., *The Confluence of Vision and Control , Lecture Notes in Control and Information Sciences* Vol. 237. Springer Verlag, Berlin, pp. 66-78, 1998.
- [19] Martinet, P. "Comparison of visual servoing techniques: Experimental results". *In Proceedings of the European Control Conference, ECC'99*, August 1999.
- [20] Chesi, G., Hashimoto, K. "Static-eye against hand-eye visual servoing". *1st IEEE Conf. on Decision and Control* (Las Vegas, 2002.12.10-13)/pp.2854-2859.
- [21] Malis, E., Chaumette, F. y Boudet S. "2-1/2-D Visual Servoing", *IEEE Transactions on Robotics and Automation* 15(2), 238-250, 1999.

- [22] Deguchi, K. "Optimal motion control for image-based visual servoing by coupling translation and rotation". In *Proc. Int. Conf. Intelligent Robots and Systems*, Oct. 1998, pp. 705-711.
- [23] Zhang Z. y Hanson, A.R. "3D reconstruction based on homography mapping". In *ARPA Image Understanding Workshop*, Palm Springs, CA, 1996.
- [24] Faugeras, O. y Lustman, F. "Motion and structure from motion in a piecewise planar environment". *Int. Journal of Pattern Recognition and Artificial Intelligence*, 2(3): 485-508, 1988.
- [25] Ma, Y., Soatto, S., Kosecka, J. y Sastry, S. (MASKS). An Introduction to 3D Vision. Capitulo 5.
- [26] Corke, P. y Hutchison, S. "A new partitioned approach to image-based visual servo control". *IEEE Trans. Robot. Automat.*, vol 17, pp. 507-515, Aug 2001.
- [27] Cretual, A y Chaumette, F. "Positionning a camera parallel to a plane using dynamic visual servoing". In *IEEE Int. Conf. on Intelligent Robots and Systems*, volumen 1, pages 43-48, Grenoble, France, September 1997.
- [28] Malis, E. "An unified approach to model-based and model-free visual servoing". In *European Conference of Computer Vision*, Copenhagen, Denmark, May 2002.
- [29] Gans, N., Hutchinson, S. y Corke, P. "Performance Tests for Visual Servo Control Systems, with Application to Partitioned Approaches to Visual Servo Control". *The International Journal of Robotics Research*, Vol. 22, No. 10-11, 955-981, 2003.
- [30] Malis, E. & Bartoli, A. "Euclidean Reconstruction Independent on Camera Intrinsic Parameters". *IEEE/RSJ International Conference on Intelligent Robots Systems*, Sendai, Japan, October 2004.
- [31] Lenz, R. y Tsai, R. "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration". *IEEE Transactions on Robotics and Automation*, vol. 5, no. 3, pp. 345-358, 1989.
- [32] Dementhon, D. y Davis, L. S. "Model-based object pose in 25 lines of code. *International journal of Computer Vision*", 15(1/2):123-141, June 1995.

- [33] Hashimoto, K. "A review on vision-based control of robot manipulators". *Advance Robotics*, vol. 17, No 10, pp 969-991, 2003.
- [34] Chaumette, F. y Marchand, E. "Recent results in visual servoing for robotics applications". In 8th ESA Workshop on Advanced Space Technologies for Robotics and Automation, ASTRA 2004, Pages 471-478, Noordwijk, The Netherlands, November.
- [35] Hartley, R. "Extraction of focal lengths from the fundamental matrix". Technical Report, G.E. CRD, Schenectady, NY, 1993.
- [36] Wilf, J. y Cunningham R. "Computing Region Moments From Boundary Representations", NASA JPL, JPL 79-45, 1979.
- [37] Malis, E. "Vision-based control using different cameras for learning the reference image and for servoing". *IEEE/RSJ International Conference on Intelligent Robots Systems*, volume 3, pages 1428-1433, Hawaii, USA, November 2001.
- [38] E. Malis, J. Borrelly, P. Rives. "Intrinsics-free visual servoing with respect to straight lines". *IEEE/RSJ International Conference on Intelligent Robots Systems*, Lausanne, Switzerland, October 2002.
- [39] Benhimane, S. y Malis, E. "Vision-based control with respect to planar and non-planar objects using a zooming camera". In *IEEE International Conference on Advanced Robotics*, volume 2, pages 991-996, Coimbra, Portugal 2003.
- [40] Zhang, Z., Deriche, R., Faugeras, O. y Luong Quang-Tuan. "A robust technique for matching two uncalibrated images through the recovery of the unakwon epipolar geometry". Technical Report 2273, INRIA, May 1994.
- [41] Dufournaud, Y., Shcmid, C. y Horaud, R. "Matching images with different resolutions". In *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, South California, USA, June 2000.
- [42] Guttman, I., Wilks, S. y Hunter, J. Introductory engineering statistics. 3rd Edition, John Wiley & Sons, 1982, New York.
- [43] Garcia, N. "Técnicas avanzadas para el control visual de robots basadas en características de la imagen. Aplicación a la solución de posibles singularidades". Tesis Doctoral. Universidad Miguel Hernández, 2004.

APÉNDICE A: Matriz jacobiana en el espacio invariante

La matriz jacobiana invariante \mathbf{J}_q se obtiene agrupando todas las matrices \mathbf{J}_{q_i} relativas a los puntos invariantes \mathbf{q}_i , tal que:

$$\mathbf{J}_q = [\mathbf{J}_{q_1}^T \quad \mathbf{J}_{q_2}^T \quad \dots \quad \mathbf{J}_{q_n}^T]^T \quad (\text{A.1})$$

Derivando la ecuación (2.31), se obtiene:

$$\dot{\mathbf{q}}_i = \dot{\mathbf{T}}_p^{-1} \mathbf{p}_i + \mathbf{T}_p^{-1} \dot{\mathbf{p}}_i \quad (\text{A.2})$$

Sabiendo que $\dot{\mathbf{T}}_p^{-1} = -(\mathbf{T}_p^{-1})(\dot{\mathbf{T}}_p)(\mathbf{T}_p^{-1})$, nos queda:

$$\dot{\mathbf{q}}_i = \mathbf{T}_p^{-1}(\dot{\mathbf{p}}_i - \mathbf{A}\mathbf{p}_i) \quad (\text{A.3})$$

donde \mathbf{A} es definida como:

$$\mathbf{A} = \dot{\mathbf{T}}_p \mathbf{T}_p^{-1} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & 0 \end{bmatrix} \quad (\text{A.4})$$

Si se agrupan los elementos de \mathbf{A} en un vector $\mathbf{a} = [a_{11} \quad a_{12} \quad a_{13} \quad a_{21} \quad a_{22}]^T$, la ecuación (A.3) puede ser reescrita como:

$$\dot{\mathbf{q}}_i = \mathbf{T}_p^{-1}(\dot{\mathbf{p}}_i - \mathbf{C}_i \mathbf{a}) \quad (\text{A.5})$$

donde:

$$\mathbf{C}_i = \begin{bmatrix} u_i & v_i & 1 & 0 & 0 \\ 0 & 0 & 0 & v_i & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{A.6})$$

Derivando la expresión de \mathbf{S}_p dada por la ecuación (2.29):

$$\begin{aligned} \dot{\mathbf{S}}_p &= \dot{\mathbf{T}}_p \mathbf{T}_p^T + \mathbf{T}_p \dot{\mathbf{T}}_p^T = \dot{\mathbf{T}}_p \mathbf{T}_p^{-1} \mathbf{T}_p \mathbf{T}_p^T + \mathbf{T}_p \mathbf{T}_p^T \mathbf{T}_p^{-T} \dot{\mathbf{T}}_p^T \\ &= \mathbf{A} \mathbf{S}_p + \mathbf{S}_p \mathbf{A}^T \end{aligned} \quad (\text{A.7})$$

Al reescribir la ecuación anterior en forma matricial nos queda:

$$\begin{bmatrix} \dot{\sigma}_{11}^p & \dot{\sigma}_{12}^p & \dot{\sigma}_{13}^p \\ \dot{\sigma}_{12}^p & \dot{\sigma}_{22}^p & \dot{\sigma}_{23}^p \\ \dot{\sigma}_{13}^p & \dot{\sigma}_{23}^p & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{11}^p & \sigma_{12}^p & \sigma_{13}^p \\ \sigma_{12}^p & \sigma_{22}^p & \sigma_{23}^p \\ \sigma_{13}^p & \sigma_{23}^p & 1 \end{bmatrix} + \begin{bmatrix} \sigma_{11}^p & \sigma_{12}^p & \sigma_{13}^p \\ \sigma_{12}^p & \sigma_{22}^p & \sigma_{23}^p \\ \sigma_{13}^p & \sigma_{23}^p & 1 \end{bmatrix} \begin{bmatrix} a_{11} & 0 & 0 \\ a_{12} & a_{22} & 0 \\ a_{13} & a_{23} & 0 \end{bmatrix} \quad (\text{A.8})$$

Se obtiene:

$$\begin{aligned} \dot{\sigma}_{11}^p &= 2\sigma_{11}^p a_{11} + 2\sigma_{12}^p a_{12} + 2\sigma_{13}^p a_{13} \\ \dot{\sigma}_{12}^p &= \sigma_{12}^p a_{11} + \sigma_{22}^p a_{12} + \sigma_{23}^p a_{13} + \sigma_{12}^p a_{22} + \sigma_{13}^p a_{23} \\ \dot{\sigma}_{13}^p &= \sigma_{13}^p a_{11} + \sigma_{23}^p a_{12} + a_{13} \\ \dot{\sigma}_{22}^p &= 2\sigma_{22}^p a_{22} + 2\sigma_{23}^p a_{23} \\ \dot{\sigma}_{23}^p &= \sigma_{13}^p a_{22} + a_{23} \end{aligned} \quad (\text{A.9})$$

En forma matricial nos queda:

$$\begin{bmatrix} \dot{\sigma}_{11}^p \\ \dot{\sigma}_{12}^p \\ \dot{\sigma}_{13}^p \\ \dot{\sigma}_{22}^p \\ \dot{\sigma}_{23}^p \end{bmatrix} = \begin{bmatrix} 2\sigma_{11}^p & 2\sigma_{12}^p & 2\sigma_{13}^p & 0 & 0 \\ \sigma_{12}^p & \sigma_{22}^p & \sigma_{23}^p & \sigma_{12}^p & \sigma_{13}^p \\ \sigma_{13}^p & \sigma_{23}^p & 1 & 0 & 0 \\ 0 & 0 & 0 & 2\sigma_{22}^p & 2\sigma_{23}^p \\ 0 & 0 & 0 & \sigma_{13}^p & 1 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{22} \\ a_{23} \end{bmatrix} = \mathbf{B}(\sigma^p) \mathbf{a} \quad (\text{A.10})$$

Despajando:

$$\mathbf{a} = \mathbf{B}^{-1}(\sigma^p) \begin{bmatrix} \dot{\sigma}_{11}^p \\ \dot{\sigma}_{12}^p \\ \dot{\sigma}_{13}^p \\ \dot{\sigma}_{22}^p \\ \dot{\sigma}_{23}^p \end{bmatrix} \quad (\text{A.12})$$

WWW.BDIGITAL.ULA.VE

Para obtener los elementos $\dot{\sigma}_{kk}^p$, se deriva la expresión de \mathbf{S}_p (ecuación 2.28), tal que:

$$\begin{aligned} \dot{\sigma}_{11}^p &= \frac{2}{n} \sum_{i=1}^n u_i \dot{u}_i = \left(\frac{2}{n} \sum_{i=1}^n u_i \mathbf{j}_{u_i} \right) \mathbf{v}_C \\ \dot{\sigma}_{12}^p &= \frac{1}{n} \sum_{i=1}^n u_i \dot{v}_i + \frac{1}{n} \sum_{i=1}^n v_i \dot{u}_i = \left(\frac{1}{n} \sum_{i=1}^n u_i \mathbf{j}_{v_i} + \frac{1}{n} \sum_{i=1}^n v_i \mathbf{j}_{u_i} \right) \mathbf{v}_C \\ \dot{\sigma}_{13}^p &= \frac{1}{n} \sum_{i=1}^n \dot{u}_i = \left(\frac{1}{n} \sum_{i=1}^n \mathbf{j}_{u_i} \right) \mathbf{v}_C \\ \dot{\sigma}_{22}^p &= \frac{2}{n} \sum_{i=1}^n v_i \dot{v}_i = \left(\frac{2}{n} \sum_{i=1}^n v_i \mathbf{j}_{v_i} \right) \mathbf{v}_C \\ \dot{\sigma}_{23}^p &= \frac{1}{n} \sum_{i=1}^n \dot{v}_i = \left(\frac{1}{n} \sum_{i=1}^n \mathbf{j}_{v_i} \right) \mathbf{v}_C \end{aligned} \quad (\text{A.13})$$

donde \mathbf{j}_{u_i} y \mathbf{j}_{v_i} son la primera y segunda fila de la matriz \mathbf{J}_i (matriz jacobiana imagen), respectivamente.

La ecuación (A.13) puede ser expresada de manera matricial, tal que:

$$\begin{bmatrix} \dot{\sigma}_{11}^p \\ \dot{\sigma}_{12}^p \\ \dot{\sigma}_{13}^p \\ \dot{\sigma}_{22}^p \\ \dot{\sigma}_{23}^p \end{bmatrix} = \begin{bmatrix} \left(\frac{2}{n} \sum_{i=1}^n u_i \mathbf{j}_{u_i} \right) \\ \left(\frac{1}{n} \sum_{i=1}^n u_i \mathbf{j}_{v_i} + \frac{1}{n} \sum_{i=1}^n v_i \mathbf{j}_{u_i} \right) \\ \left(\frac{1}{n} \sum_{i=1}^n \mathbf{j}_{u_i} \right) \\ \left(\frac{2}{n} \sum_{i=1}^n v_i \mathbf{j}_{v_i} \right) \\ \left(\frac{1}{n} \sum_{i=1}^n \mathbf{j}_{v_i} \right) \end{bmatrix} \mathbf{v}_C = \mathbf{F} \mathbf{v}_C \quad (\text{A.14})$$

Luego los elementos de la matriz \mathbf{A} pueden ser expresados como:

$$\mathbf{a} = \mathbf{B}^{-1}(\sigma^p) \mathbf{F} \mathbf{v}_C \quad (\text{A.15})$$

Sustituyendo en la ecuación (A.5) se obtiene:

$$\dot{\mathbf{q}}_i = \mathbf{T}_p^{-1} \cdot (\dot{\mathbf{p}}_i - \mathbf{C}_i \cdot \mathbf{B}^{-1}(\sigma^p) \cdot \mathbf{F} \cdot \mathbf{v}_C) \quad (\text{A.16})$$

Definiendo $\dot{\mathbf{p}}_i$ como:

$$\dot{\mathbf{p}}_i = \begin{bmatrix} \dot{u}_i \\ v_i \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{j}_{u_i} \\ \mathbf{j}_{v_i} \\ 0 \end{bmatrix} \mathbf{v}_C = \mathbf{J}_i \mathbf{v}_C \quad (\text{A.17})$$

Entonces,

$$\dot{\mathbf{q}}_i = \mathbf{T}_p^{-1} \cdot (\mathbf{J}_i - \mathbf{C}_i \cdot \mathbf{B}^{-1}(\sigma^p) \cdot \mathbf{F}) \cdot \mathbf{v}_c \quad (\text{A.18})$$

Y por lo tanto, la matriz jacobiana invariante \mathbf{J}_{q_i} queda definida como:

$$\mathbf{J}_{q_i} = \mathbf{T}_p^{-1} (\mathbf{J}_i - \mathbf{C}_i \cdot \mathbf{B}^{-1}(\sigma^p) \cdot \mathbf{F}) \quad (\text{A.19})$$

WWW.BDIGITAL.ULA.VE

APÉNDICE B: Tabla de la distribución de la t -student, $t_{m,p}$ [42].

$m \backslash p$	0,25	0,1	0,05	0,025	0,01	0,005
1	1,000	3,078	6,314	12,706	31,821	63,657
2	0,816	1,886	2,920	4,303	6,965	9,925
3	0,765	1,638	2,353	3,182	4,541	5,841
4	0,741	1,533	2,132	2,776	3,747	4,604
5	0,727	1,476	2,015	2,571	3,365	4,032
6	0,718	1,440	1,943	2,447	3,143	3,707
7	0,711	1,415	1,895	2,365	2,998	3,499
8	0,706	1,397	1,860	2,306	2,896	3,355
9	0,703	1,383	1,833	2,262	2,821	3,250
10	0,700	1,372	1,812	2,228	2,764	3,169
11	0,697	1,363	1,796	2,201	2,718	3,106
12	0,695	1,356	1,782	2,179	2,681	3,055
13	0,694	1,350	1,771	2,160	2,650	3,012
14	0,692	1,345	1,761	2,145	2,624	2,977
15	0,691	1,341	1,753	2,131	2,602	2,947
16	0,690	1,337	1,746	2,120	2,583	2,921
17	0,689	1,333	1,740	2,110	2,567	2,898
18	0,688	1,330	1,734	2,101	2,552	2,878
19	0,688	1,328	1,729	2,093	2,539	2,861
20	0,687	1,325	1,725	2,086	2,528	2,845
21	0,686	1,323	1,721	2,080	2,518	2,831
22	0,686	1,321	1,717	2,074	2,508	2,819
23	0,685	1,319	1,714	2,069	2,500	2,807
24	0,685	1,318	1,711	2,064	2,492	2,797
25	0,684	1,316	1,708	2,060	2,485	2,787
26	0,684	1,315	1,706	2,056	2,479	2,779
27	0,684	1,314	1,703	2,052	2,473	2,771
28	0,683	1,313	1,701	2,048	2,467	2,763
29	0,683	1,311	1,699	2,045	2,462	2,756
30	0,683	1,310	1,697	2,042	2,457	2,750
40	0,681	1,303	1,984	2,021	2,423	2,704
60	0,679	1,296	1,671	2,000	2,390	2,660
120	0,677	1,289	1,658	1,980	2,358	2,617
∞	0,674	1,282	1,645	1,960	2,326	2,576

APÉNDICE C: Detalles Prueba 1 (Tareas 2-4) y Prueba 2 (Tareas 1-4).

C.1.- Prueba 1, Tarea 2

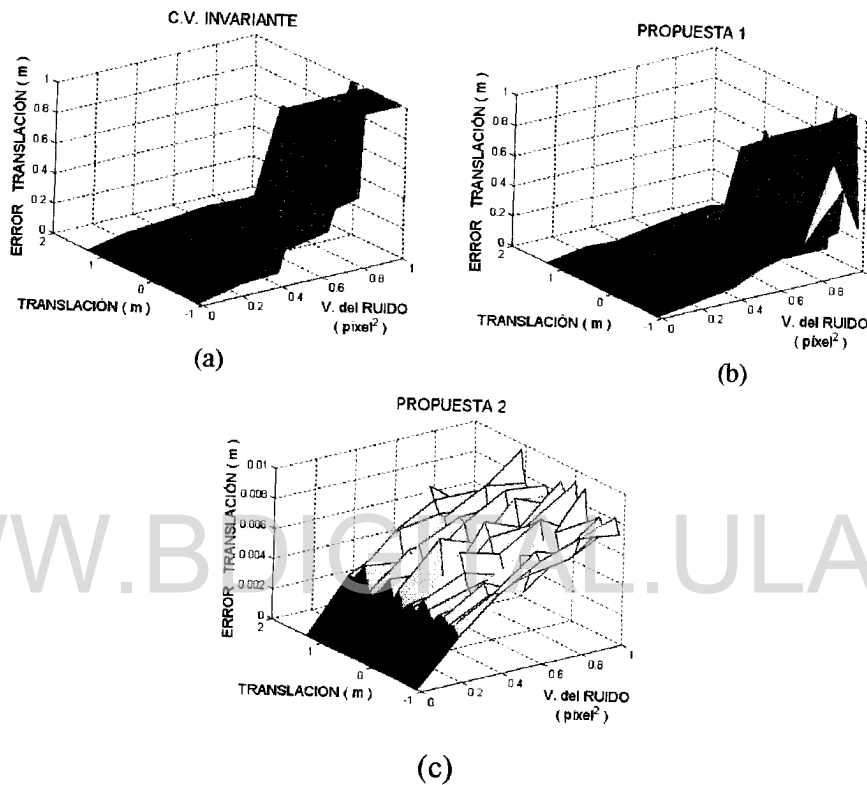


Figura C.1.1 EFT para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.1.1 EFT promedio por nivel de varianza del ruido para la tarea 2, Prueba 1 ($s = 0$).

Varianza del ruido (píxel^2)	CV Invariante (m)	Propuesta 1 (m)	Propuesta 2 (m)	Diferencia 1 (m)	Diferencia 2 (m)
0,0	0,0000	0,0001	0,0000	-0,0001	0,0000
0,2	0,0639	0,0443	0,0035	0,0196	0,0604
0,4	0,1335	0,1019	0,0049	0,0316	0,1286
0,6	0,4143	0,2059	0,0060	0,2084	0,4083
0,8	0,3992	0,2019	0,0067	0,1973	0,3925
1,0	0,5330	0,2519	0,0075	0,2811	0,5255
MEDIA	0,2573	0,1343	0,0048	0,1230	0,2526
			\hat{i}	2,5102	2,8570

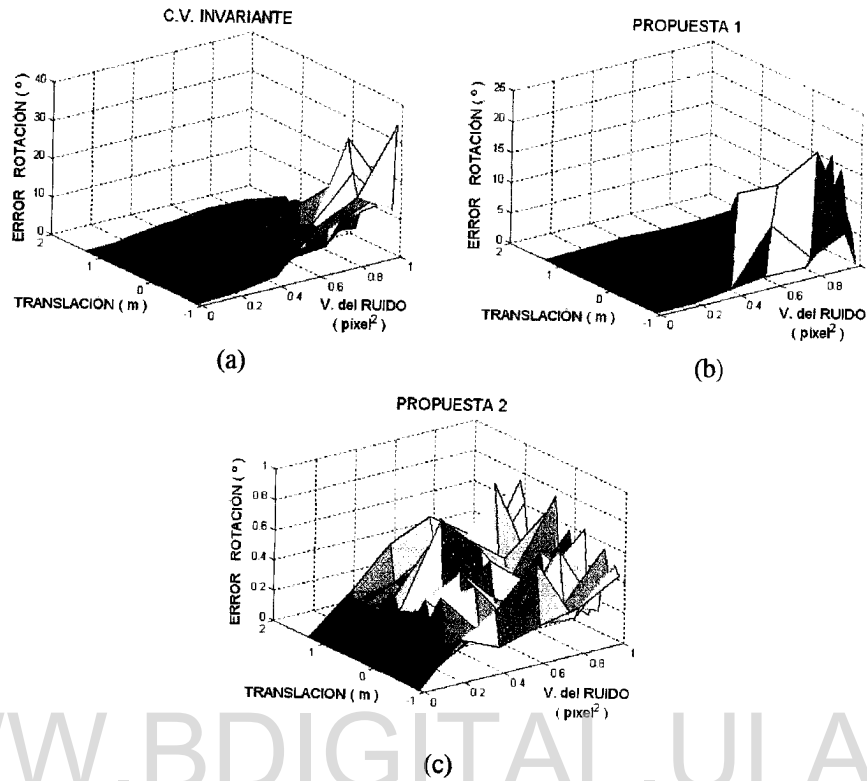


Figura C.1.2 EFR para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.1.2 EFR promedio por nivel de varianza del ruido para la tarea 2, Prueba 1 ($s = 0$).

Varianza del ruido (pixel ²)	CV Invariante (°)	Propuesta 1 (°)	Propuesta 2 (°)	Diferencia 1 (°)	Diferencia 2 (°)
0,0	0,0000	0,0000	0,0000	0,0000	0,0000
0,2	1,2574	0,1844	0,2336	1,0730	1,0238
0,4	1,9966	0,9841	0,3023	1,0125	1,6943
0,6	4,6241	1,6666	0,3350	2,9575	4,2891
0,8	6,7877	1,3718	0,3281	5,4159	6,4596
1,0	8,1697	2,3787	0,4049	5,7910	7,7648
MEDIA	3,8059	1,0976	0,2673	2,7083	3,5386
			\hat{i}	2,7183	2,7631

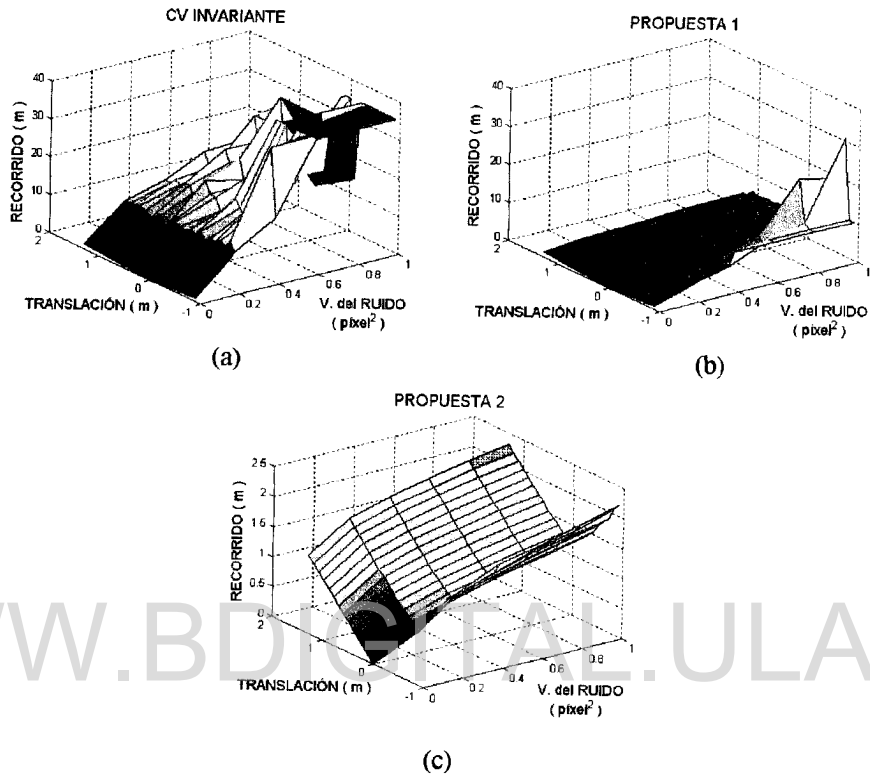


Figura C.1.3 RRC para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.1.3 RRC promedio por nivel de varianza del ruido para la tarea 2, Prueba 1 ($s = 0$).

Varianza del ruido (pixel ²)	CV Invariante (m)	Propuesta 1 (m)	Propuesta 2 (m)	Diferencia 1 (m)	Diferencia 2 (m)
0,0	0,5913	0,5913	0,5913	0,0000	0,0000
0,2	10,3623	2,8456	1,0691	7,5167	9,2932
0,4	15,6734	4,0876	1,2790	11,5858	14,3944
0,6	23,3584	5,2666	1,4415	18,0918	21,9169
0,8	24,8391	6,6541	1,5786	18,1850	23,2605
1,0	27,8425	8,0028	1,7034	19,8397	26,1391
MEDIA	17,1112	4,5747	1,2772	12,5365	15,8340
			\hat{i}	3,9686	3,9023

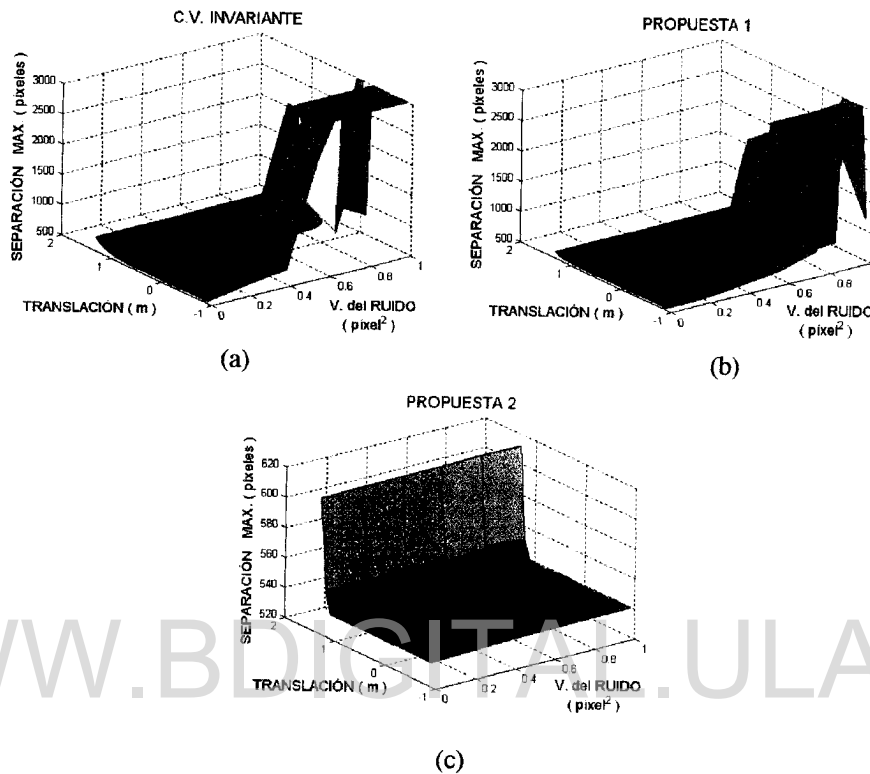


Figura C.1.4 SMPP para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.1.4 SMPP promedio por nivel de varianza del ruido para la tarea 2, Prueba 1 ($s = 0$).

Varianza del ruido (píxel ²)	CV Invariante (píxel)	Propuesta 1 (píxel)	Propuesta 2 (píxel)	Diferencia 1 (píxel)	Diferencia 2 (píxel)
0,0	555,71	539,72	539,72	15,99	15,99
0,2	656,47	546,37	541,25	110,10	115,22
0,4	809,42	659,60	541,89	149,82	267,53
0,6	1560,41	874,70	542,29	685,71	1018,12
0,8	1469,65	883,34	542,83	586,31	926,82
1,0	1779,18	1116,29	543,06	662,89	1236,12
MEDIA	1138,47	770,00	541,84	368,47	596,63
			\hat{i}	2,93	2,79

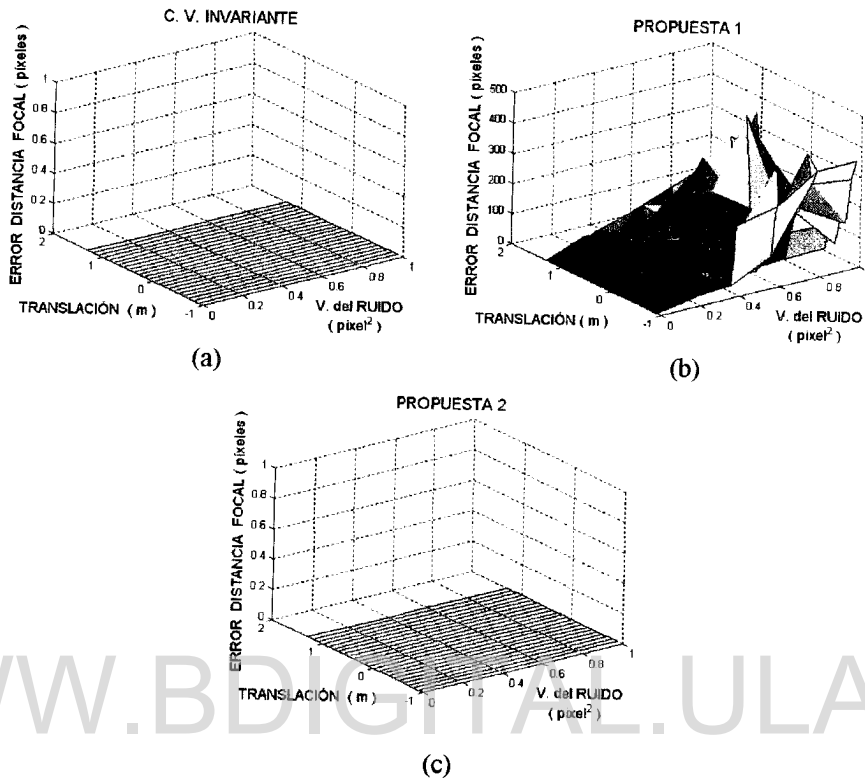


Figura C.1.5 EFDF para la tarea 2 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

C.2.- Prueba 1, Tarea 3

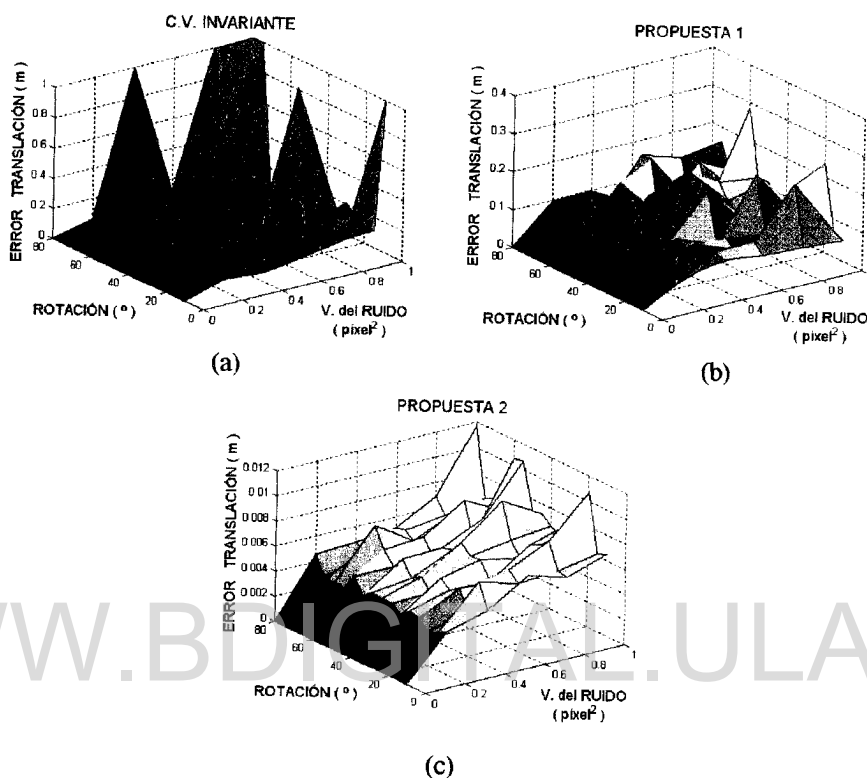


Figura C.2.1 EFT para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.2.1 EFT promedio por nivel de varianza del ruido para la tarea 3, Prueba 1 ($s = 0$).

Varianza del ruido (pixel ²)	CV Invariante (m)	Propuesta 1 (m)	Propuesta 2 (m)	Diferencia 1 (m)	Diferencia 2 (m)
0,0	0,0000	0,0000	0,0000	0,0000	0,0000
0,2	0,0717	0,0456	0,0034	0,0261	0,0683
0,4	0,1647	0,0755	0,0048	0,0892	0,1599
0,6	0,1156	0,0907	0,0054	0,0249	0,1102
0,8	0,2862	0,0977	0,0065	0,1885	0,2797
1,0	0,3112	0,1134	0,0078	0,1978	0,3034
MEDIA	0,1582	0,0705	0,0047	0,0878	0,1536
			\hat{i}	2,4749	3,1540

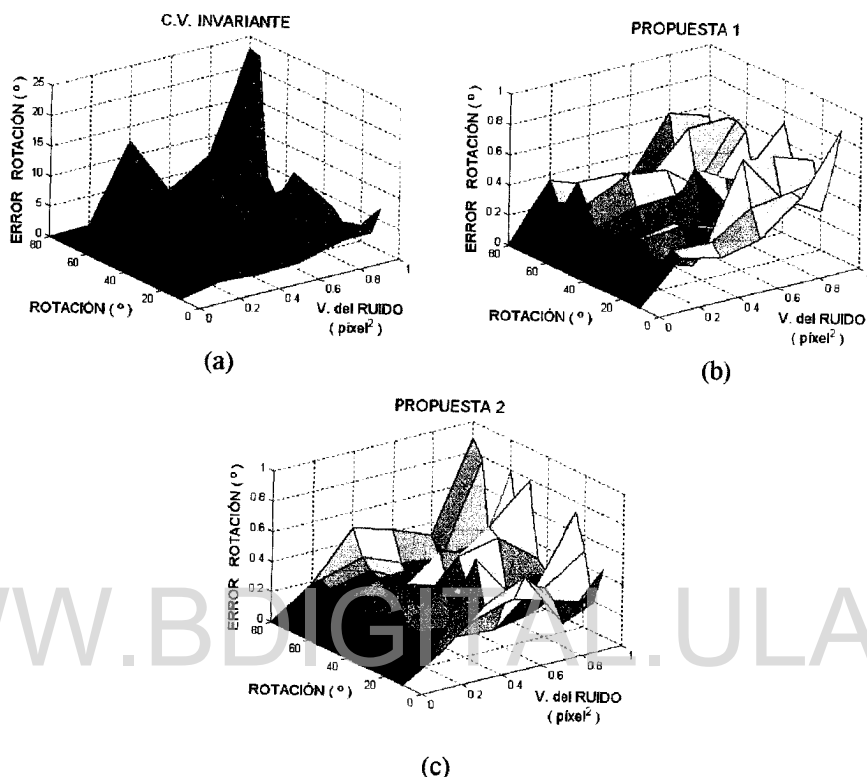


Figura C.2.2 EFR para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.2.2 EFR promedio por nivel de varianza del ruido para la tarea 3, Prueba 1 ($s = 0$).

Varianza del ruido (píxel^2)	CV Invariante ($^\circ$)	Propuesta 1 ($^\circ$)	Propuesta 2 ($^\circ$)	Diferencia 1 ($^\circ$)	Diferencia 2 ($^\circ$)
0,0	0,0000	0,0000	0,0000	0,0000	0,0000
0,2	1,1806	0,2270	0,1896	0,9536	0,9910
0,4	2,2621	0,2387	0,2910	2,0234	1,9711
0,6	2,2380	0,3397	0,2748	1,8983	1,9632
0,8	2,9562	0,3829	0,2586	2,5733	2,6976
1,0	5,5125	0,4882	0,4945	5,0243	5,0180
MEDIA	2,3582	0,2794	0,2514	2,0788	2,1068
			\hat{t}	2,9857	3,0255

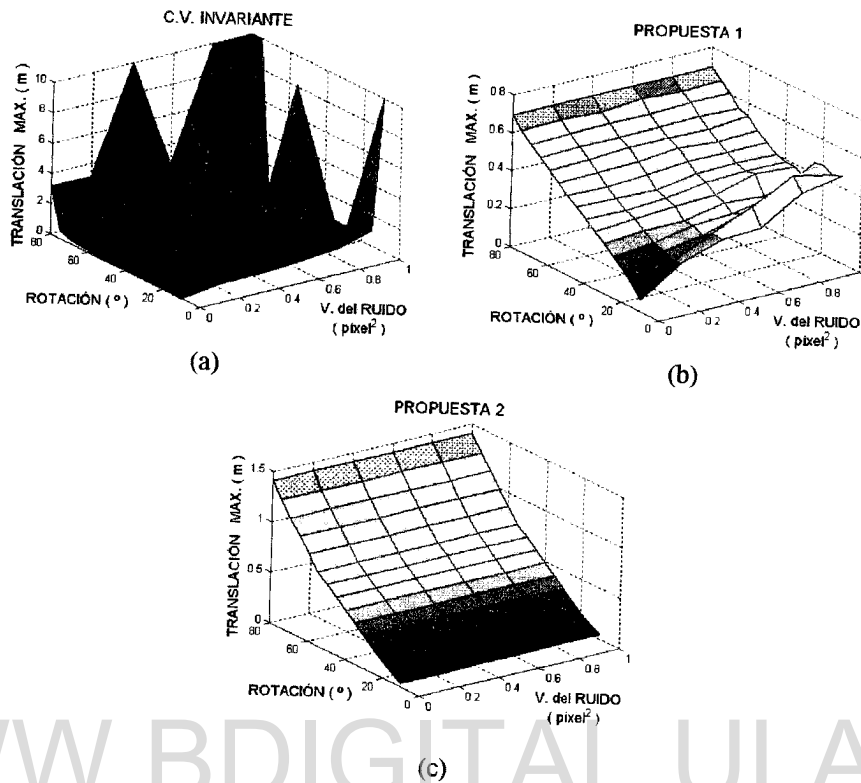


Figura C.2.3 TMI para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.2.3 TMI promedio por nivel de varianza del ruido para la tarea 3, Prueba 1 ($s = 0$).

Varianza del ruido (pixel ²)	CV Invariante (m)	Propuesta 1 (m)	Propuesta 2 (m)	Diferencia 1 (m)	Diferencia 2 (m)
0,0	0,2939	0,3776	0,6238	-0,0837	-0,3299
0,2	0,4762	0,4004	0,6244	0,0758	-0,1482
0,4	1,1121	0,4203	0,6247	0,6918	0,4874
0,6	0,7021	0,4442	0,6260	0,2579	0,0761
0,8	2,0838	0,4672	0,6253	1,6166	1,4585
1,0	2,5858	0,4882	0,6258	2,0976	1,9600
MEDIA	1,2090	0,4330	0,6250	0,7760	0,5840
			\hat{i}	2,1364	1,5431

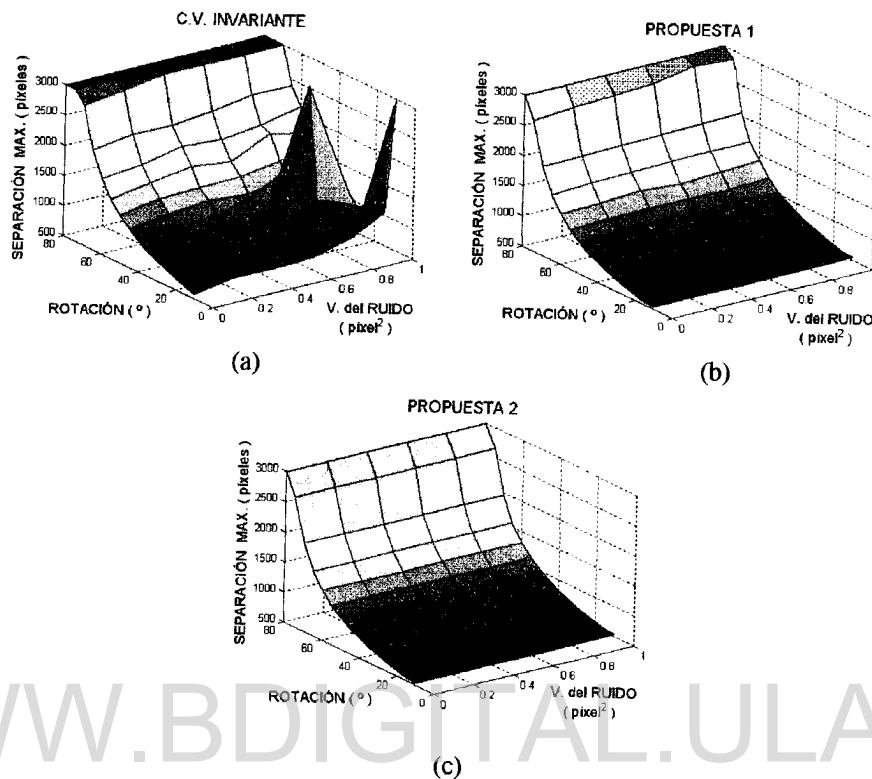


Figura C.2.4 SMPP para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.2.4 SMPP promedio por nivel de varianza del ruido para la tarea 3, Prueba 1 ($s = 0$).

Varianza del ruido (píxel ²)	CV Invariante (píxel)	Propuesta 1 (píxel)	Propuesta 2 (píxel)	Diferencia 1 (píxel)	Diferencia 2 (píxel)
0,0	1465,48	1212,52	1212,52	252,96	252,96
0,2	1546,21	1228,13	1213,25	318,08	332,96
0,4	1595,58	1238,69	1213,46	356,89	382,12
0,6	1619,48	1241,24	1213,50	378,24	405,98
0,8	1857,83	1255,48	1213,98	602,35	643,85
1,0	1860,60	1258,42	1214,02	602,18	646,58
MEDIA	1657,53	1239,08	1213,46	418,45	444,08
			\hat{i}	6,90	6,62

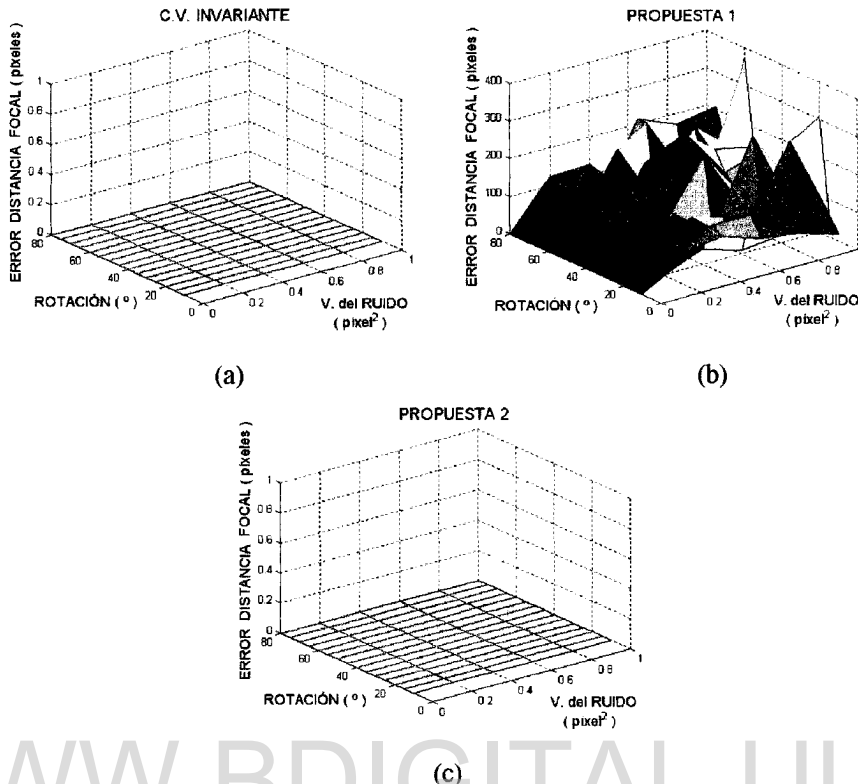


Figura C.2.5 EFDF para la tarea 3 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

C.3.- Prueba 1, Tarea 4

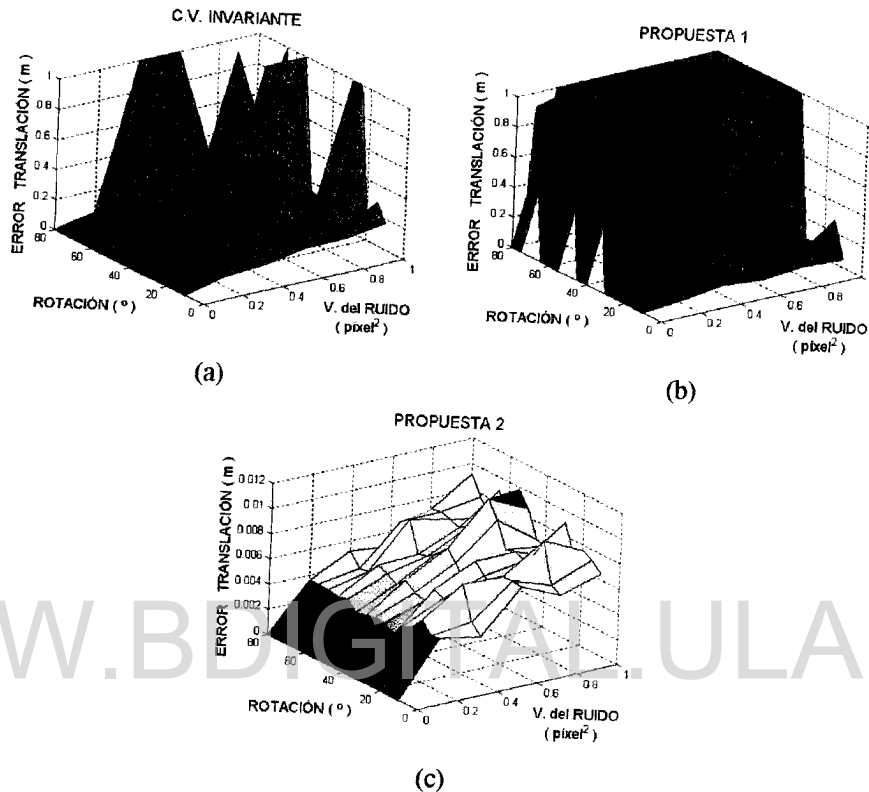


Figura C.3.1 EFT para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.3.1 EFT promedio por nivel de varianza del ruido para la tarea 4, Prueba 1 ($s = 0$).

Varianza del ruido (pixel ²)	CV Invariante (m)	Propuesta 1 (m)	Propuesta 2 (m)	Diferencia 1 (m)	Diferencia 2 (m)
0,0	0,0000	0,2001	0,0000	-0,2001	0,0000
0,2	0,0678	0,7455	0,0036	-0,6777	0,0642
0,4	0,1620	0,7676	0,0048	-0,6056	0,1572
0,6	0,1682	0,6906	0,0063	-0,5224	0,1619
0,8	0,2457	0,7548	0,0073	-0,5091	0,2384
1,0	0,3832	0,7009	0,0075	-0,3177	0,3757
MEDIA	0,1712	0,6433	0,0049	-0,4721	0,1662
			\hat{i}	6,4281	3,0828

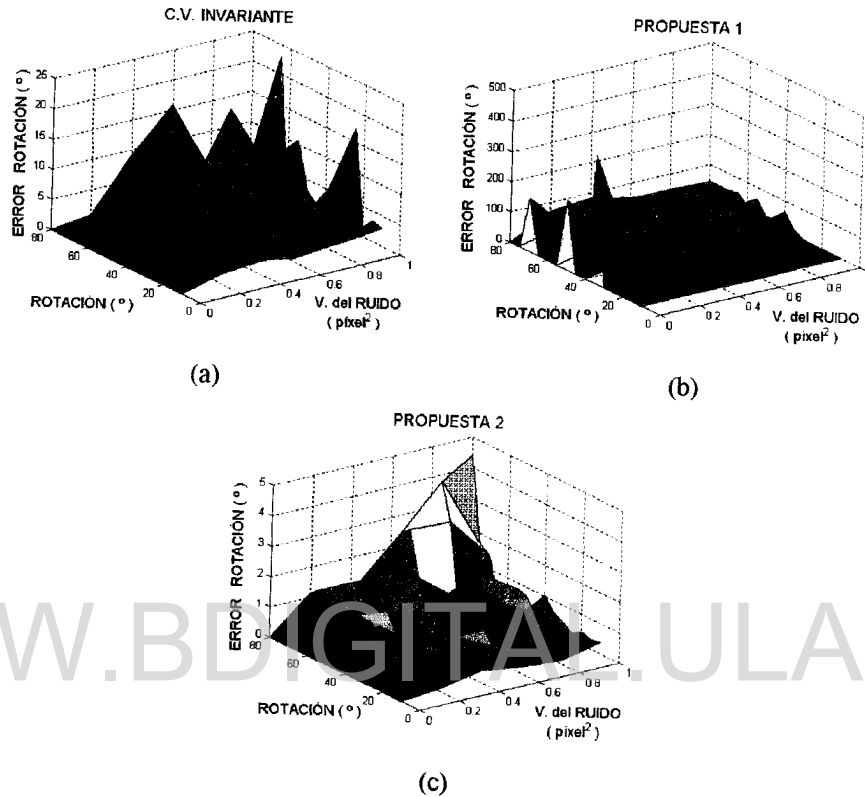


Figura C.3.2 EFR para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.3.2.- EFR promedio por nivel de varianza del ruido para la tarea 4, Prueba 1 ($s = 0$).

Varianza del ruido (píxel ²)	CV Invariante (°)	Propuesta 1 (°)	Propuesta 2 (°)	Diferencia 1 (°)	Diferencia 2 (°)
0,0	0,0000	55,3793	0,0000	-55,3793	0,0000
0,2	1,1452	28,4326	0,4655	-27,2874	0,6797
0,4	2,1714	26,9958	0,6459	-24,8244	1,5255
0,6	2,8943	29,3943	0,8343	-26,5000	2,0600
0,8	3,0943	27,7955	0,9350	-24,7012	2,1593
1,0	5,7304	26,6558	0,9085	-20,9254	4,8219
MEDIA	2,5059	32,4422	0,6315	-29,9363	1,8744
			\hat{i}	5,7939	2,7571

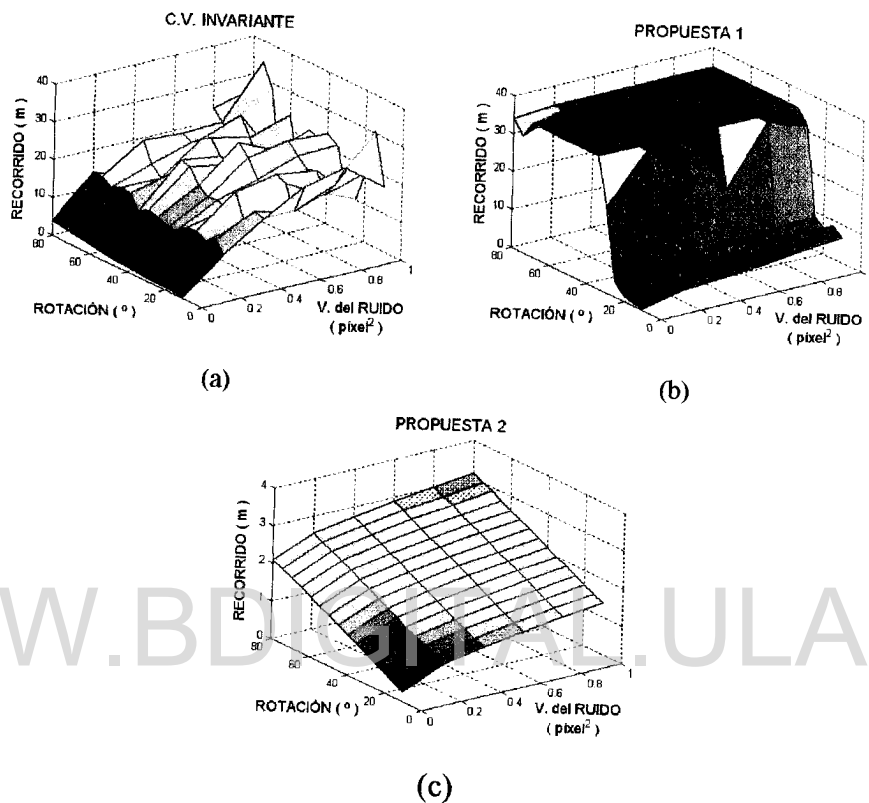


Figura C.3.3 RRC para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.3.3 RRC promedio por nivel de varianza del ruido para la tarea 4, Prueba 1 ($s = 0$).

Varianza del ruido (píxel ²)	CV Invariante (m)	Propuesta 1 (m)	Propuesta 2 (m)	Diferencia 1 (m)	Diferencia 2 (m)
0,0	1,7446	11,5779	1,1957	-23,1667	0,5489
0,2	11,5660	8,3075	1,6578	-15,0748	9,9082
0,4	16,2511	8,6804	1,8556	-10,7626	14,3955
0,6	19,6254	9,4875	2,0105	-6,5288	17,6149
0,8	19,5547	9,1384	2,1399	-7,9171	17,4148
1,0	20,5669	10,7943	2,2562	-6,8941	18,3107
MEDIA	14,8848	9,6643	1,8526	-11,7240	13,0322
			\hat{i}	4,4536	4,6512

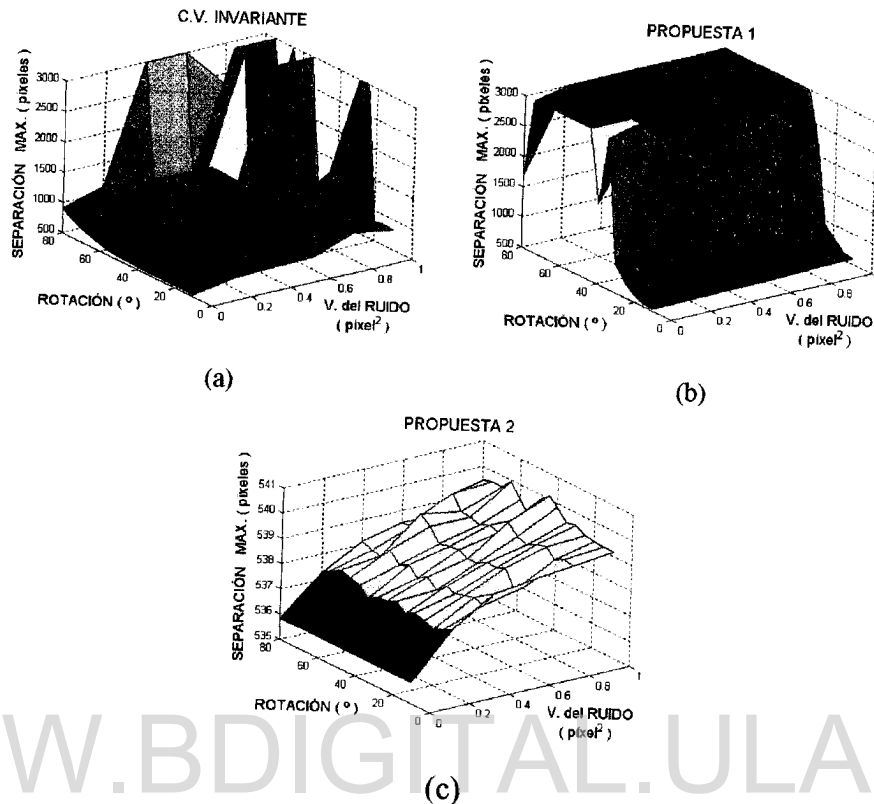


Figura C.3.4 SMPP para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

Tabla C.3.4 SMPP promedio por nivel de varianza del ruido para la tarea 4, Prueba 1 ($s = 0$).

Varianza del ruido (píxe²)	CV Invariante (píxel)	Propuesta 1 (píxel)	Propuesta 2 (píxel)	Diferencia 1 (píxel)	Diferencia 2 (píxel)
0,0	610,27	2062,86	535,83	-1452,59	74,44
0,2	746,85	2376,74	537,45	-1629,89	209,40
0,4	931,16	2386,30	538,05	-1455,14	393,11
0,6	1129,05	2397,13	538,57	-1268,08	590,48
0,8	1363,89	2387,62	539,00	-1023,73	824,89
1,0	1759,52	2346,57	539,49	-587,05	1220,03
MEDIA	1090,12	2326,20	538,07	-1236,08	552,06
			\hat{i}	7,99	3,20

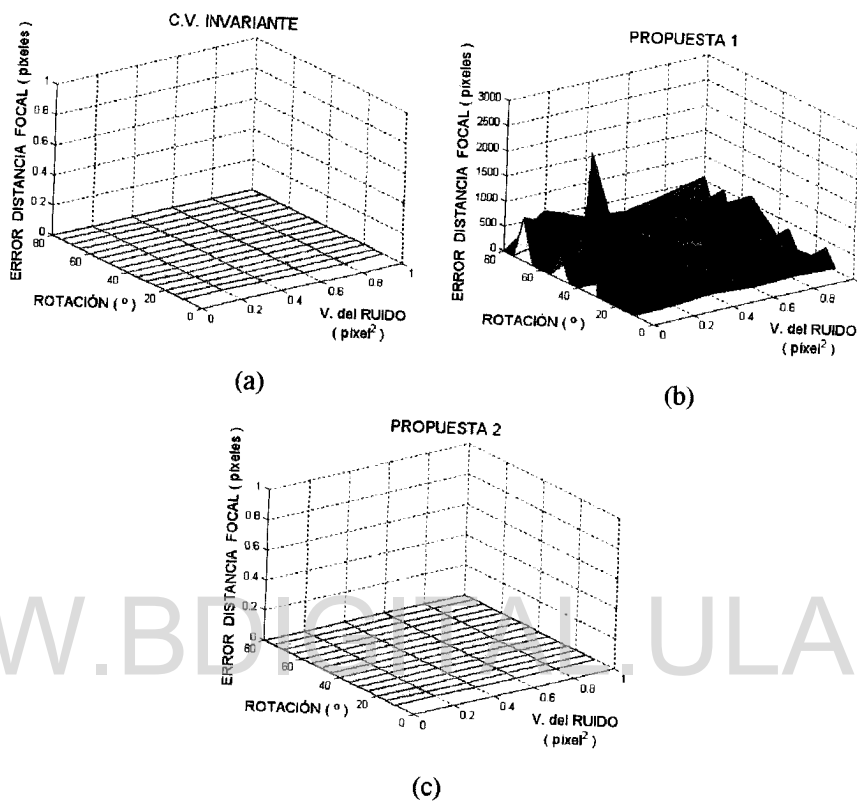


Figura C.3.5 EFDF para la tarea 4 y diferentes niveles de ruido, prueba 1 ($s = 0$).
 (a) CV invariante, (b) Propuesta 1 y (c) Propuesta 2.

C.4.- Prueba 2, Tarea 1

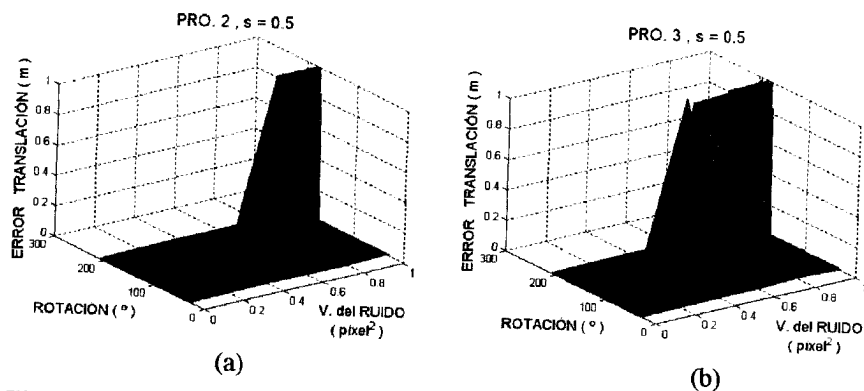


Figura C.4.1 EFT para la tarea 1 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.4.1.- EFT promedio por nivel de varianza del ruido para la tarea 1, prueba 2 ($s = 0,5$).

Varianza del ruido (pixel ²)	Propuesta 2 (m)	Propuesta 3 (m)	Diferencia 3 (m)
0,0	0,0000	0,0000	0,0000
0,2	0,0036	0,0050	-0,0014
0,4	0,0050	0,0068	-0,0018
0,6	0,0061	0,0616	-0,0555
0,8	0,0337	0,0674	-0,0337
1,0	0,0617	0,1179	-0,0562
MEDIA	0,0184	0,0431	- 0,0248
		\hat{i}	2,2303

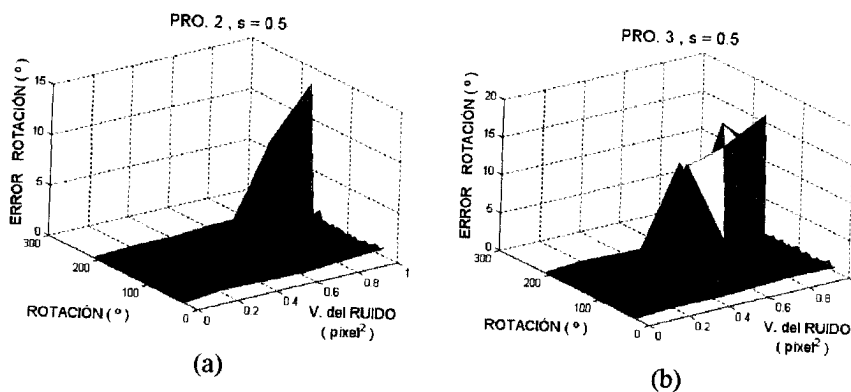
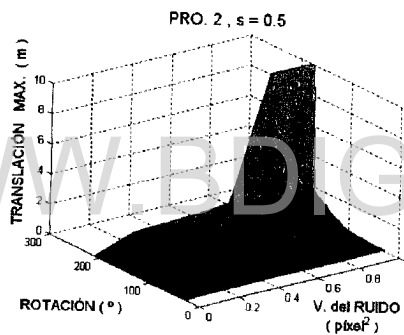


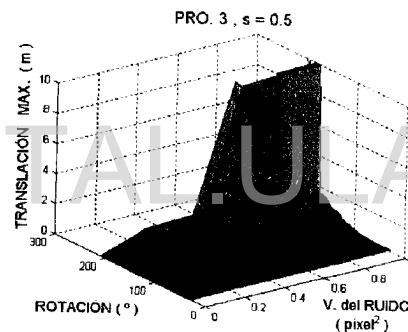
Figura C.4.2 ERT para la tarea 1 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.4.2.- EFR promedio por nivel de varianza del ruido para la tarea 1, prueba 2 ($s = 0,5$).

Varianza del ruido (pixel ²)	Propuesta 2 (°)	Propuesta 3 (°)	Diferencia 3 (°)
0,0	0,0000	0,0000	0,0000
0,2	0,2217	0,2138	0,0079
0,4	0,2920	0,2943	-0,0023
0,6	0,3701	0,9374	-0,5673
0,8	0,5823	1,2320	-0,6497
1,0	0,9622	1,6229	-0,6607
MEDIA	0,4047	0,7167	-0,3120
		\hat{t}	2,2129



(a)



(b)

Figura C.4.3.- TMI para la tarea 1 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.4.3 TMI promedio por nivel de varianza del ruido para la tarea 1, prueba 2 ($s = 0,5$).

Varianza del ruido (pixel ²)	Propuesta 2 (m)	Propuesta 3 (m)	Diferencia 3 (m)
0,0	0,0000	0,0000	0,0000
0,2	0,5597	0,5616	-0,0019
0,4	0,6478	0,6640	-0,0162
0,6	0,7114	1,1589	-0,4475
0,8	0,9574	1,3172	-0,3598
1,0	1,1981	1,6511	-0,4530
MEDIA	0,6791	0,8921	-0,2131
		\hat{t}	2,2765

Licencia Creative Commons:

Atribución - No Comercial - Compartir Igual 3.0 Venezuela
(CC BY-NC-SA 3.0 VE)

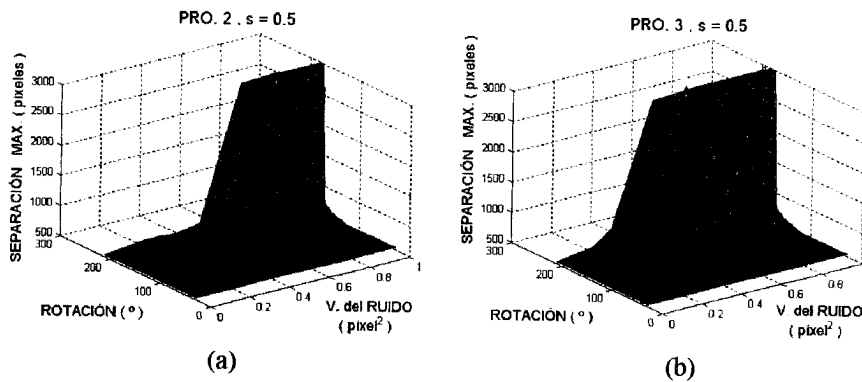


Figura C.4.4 SMPP para la tarea 1 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
 (a) Propuesta 2, (b) Propuesta 3.

Tabla C.4.4 SMPP promedio por nivel de varianza del ruido para la tarea 1, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (píxel)	Propuesta 3 (píxel)	Diferencia 3 (píxel)
0,0	535,85	535,85	0,00
0,2	582,90	595,46	-12,56
0,4	590,19	713,81	-123,62
0,6	719,21	910,74	-191,53
0,8	780,40	898,90	-118,50
1,0	742,06	854,20	-112,14
MEDIA	658,44	751,49	-93,06
		\hat{i}	3,11

C.5.- Prueba 2, Tarea 2

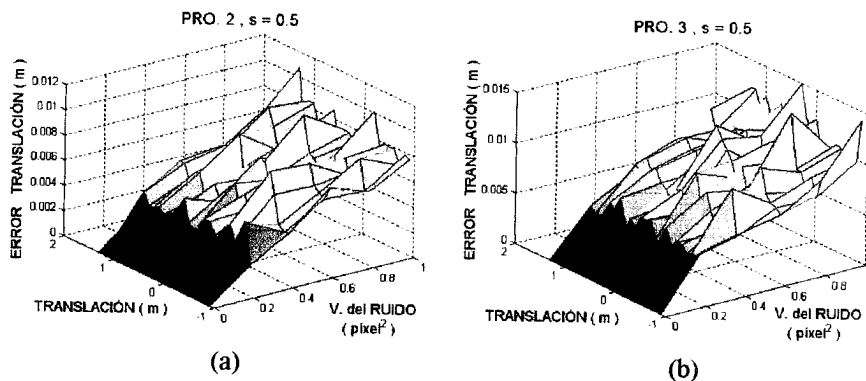


Figura C.5.1 EFT para la tarea 2 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
 (a) Propuesta 2, (b) Propuesta 3.

Tabla C.5.1 EFT promedio por nivel de varianza del ruido para la tarea 2, prueba 2 ($s = 0,5$).

Varianza del ruido (pixel ²)	Propuesta 2 (m)	Propuesta 3 (m)	Diferencia 3 (m)
0,0	0,0000	0,0000	0,0000
0,2	0,0037	0,0048	-0,0011
0,4	0,0050	0,0070	-0,0020
0,6	0,0062	0,0087	-0,0025
0,8	0,0067	0,0094	-0,0027
1,0	0,0076	0,0109	-0,0033
MEDIA	0,0049	0,0068	-0,0019
		\hat{t}	3,9428

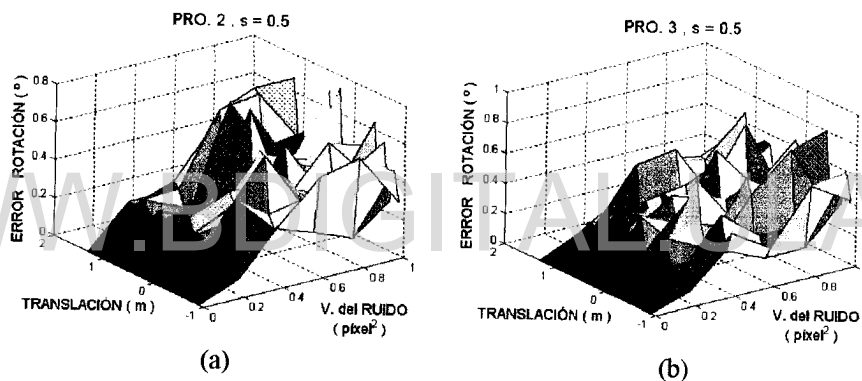


Figura C.5.2 EFR para la tarea 2 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.5.2 EFR promedio por nivel de varianza del ruido para la tarea 2, prueba 2 ($s = 0,5$).

Varianza del ruido (pixel ²)	Propuesta 2 (°)	Propuesta 3 (°)	Diferencia 3 (°)
0,0	0,0000	0,0000	0,0000
0,2	0,2116	0,1781	0,0335
0,4	0,2636	0,3234	-0,0598
0,6	0,3761	0,3605	0,0156
0,8	0,3438	0,3720	-0,0282
1,0	0,3900	0,4165	-0,0265
MEDIA	0,2642	0,2751	-0,0109
		\hat{t}	0,7890

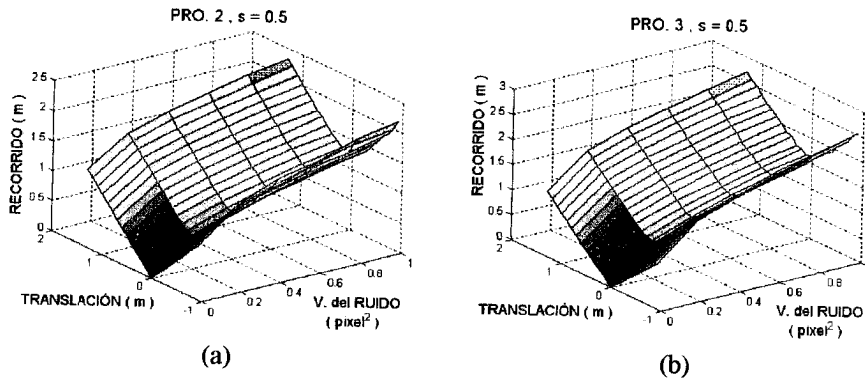


Figura C.5.3 RRC para la tarea 2 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.5.3 RRC promedio por nivel de varianza del ruido para la tarea 2, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (m)	Propuesta 3 (m)	Diferencia 3 (m)
0,0	0,5913	0,5913	0,0000
0,2	1,0679	1,2376	-0,1697
0,4	1,2788	1,5140	-0,2352
0,6	1,4392	1,7363	-0,2971
0,8	1,5784	1,9186	-0,3402
1,0	1,6988	2,0845	-0,3857
MEDIA	1,2757	1,5137	-0,2380
		\hat{i}	4,1837

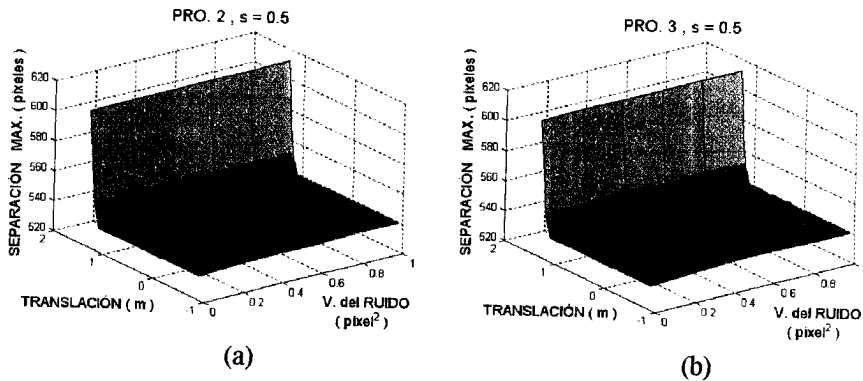


Figura C.5.4 SMPP para la tarea 2 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.5.4 SMPP promedio por nivel de varianza del ruido para la tarea 2, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (píxel)	Propuesta 3 (píxel)	Diferencia 3 (píxel)
0,0	539,75	539,75	0,00
0,2	541,22	541,39	-0,17
0,4	541,89	542,11	-0,22
0,6	542,31	542,61	-0,30
0,8	542,74	543,23	-0,49
1,0	543,06	543,35	-0,29
MEDIA	541,83	542,07	-0,25
		\hat{i}	3,70

C.6.- Prueba 2, Tarea 3

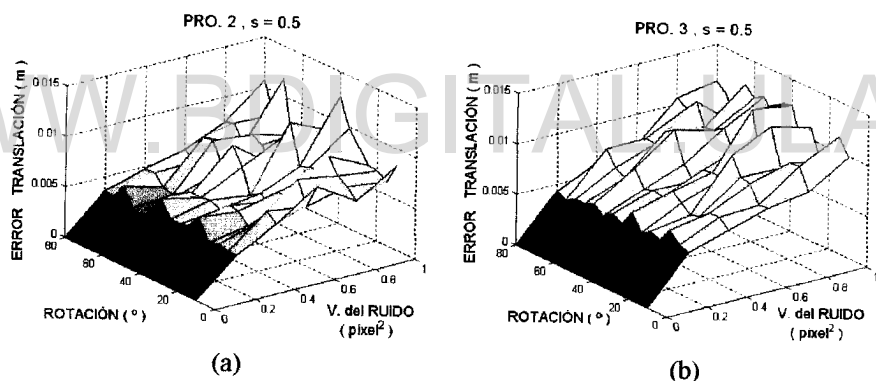


Figura C.6.1 EFT para la tarea 3 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.6.1 EFT promedio por nivel de varianza del ruido para la tarea 3, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (m)	Propuesta 3 (m)	Diferencia 3 (m)
0,0	0,0000	0,0000	0,0000
0,2	0,0039	0,0046	-0,0007
0,4	0,0047	0,0063	-0,0016
0,6	0,0063	0,0080	-0,0017
0,8	0,0073	0,0097	-0,0024
1,0	0,0083	0,0106	-0,0023
MEDIA	0,0051	0,0065	-0,0015
		\hat{i}	3,7970

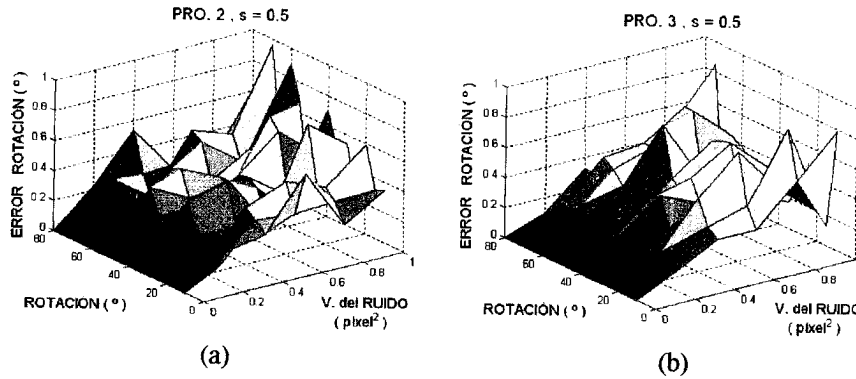


Figura C.6.2 EFR para la tarea 3 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.6.2 EFR promedio por nivel de varianza del ruido para la tarea 3, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (°)	Propuesta 3 (°)	Diferencia 3 (°)
0,00	0,0000	0,0000	0,0000
0,20	0,2046	0,1743	0,0303
0,40	0,2955	0,3060	-0,0105
0,60	0,3485	0,3646	-0,0161
0,80	0,3495	0,4596	-0,1101
1,00	0,5026	0,3650	0,1376
MEDIA	0,2835	0,2783	0,0052
		\hat{i}	0.1588

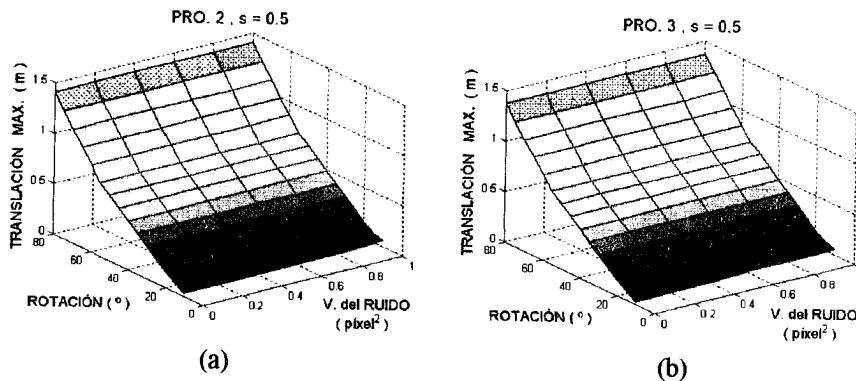


Figura C.6.3 TMI para la tarea 3 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.6.3 RRC promedio por nivel de varianza del ruido para la tarea 3, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (m)	Propuesta 3 (m)	Diferencia 3 (m)
0,0	0,6239	0,6142	0,0097
0,2	0,6242	0,6141	0,0101
0,4	0,6245	0,6151	0,0094
0,6	0,6252	0,6161	0,0091
0,8	0,6250	0,6151	0,0099
1,0	0,6261	0,6159	0,0102
MEDIA	0,6248	0,6151	0,0097
		\hat{t}	56,4047

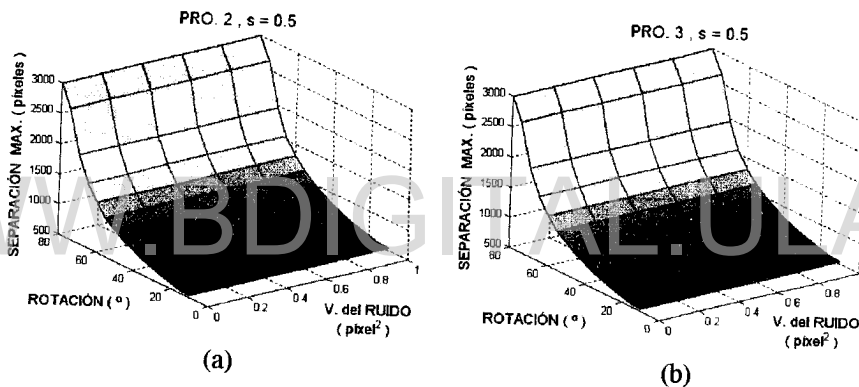


Figura C.6.4 SMPP para la tarea 3 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.6.4 SMPP promedio por nivel de varianza del ruido para la tarea 3, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (píxel)	Propuesta 3 (píxel)	Diferencia 3 (píxel)
0,0	1212,56	1212,56	0,00
0,2	1213,26	1213,22	0,04
0,4	1213,53	1213,56	-0,03
0,6	1213,61	1213,81	-0,20
0,8	1213,82	1214,01	-0,19
1,0	1214,05	1214,07	-0,02
MEDIA	1213,47	1213,54	-0,07
		\hat{t}	1,60

C.7.- Prueba 2, Tarea 4

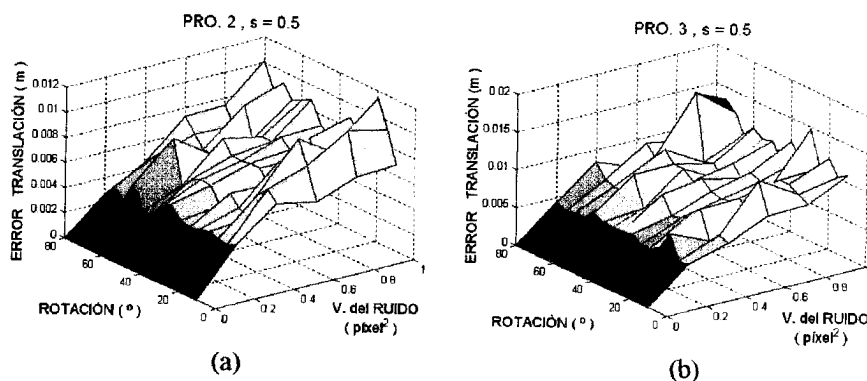


Figura C.7.1 EFT para la tarea 4 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.7.1 EFT promedio por nivel de varianza del ruido para la tarea 4, prueba 2 ($s = 0,5$).

Varianza del ruido (pixel^2)	Propuesta 2 (m)	Propuesta 3 (m)	Diferencia 3 (m)
0,0	0,0000	0,0000	0,0000
0,2	0,0034	0,0047	-0,0013
0,4	0,0051	0,0068	-0,0017
0,6	0,0065	0,0088	-0,0023
0,8	0,0069	0,0091	-0,0022
1,0	0,0083	0,0113	-0,0030
MEDIA	0,0050	0,0068	-0,0018
		\hat{i}	4,1498

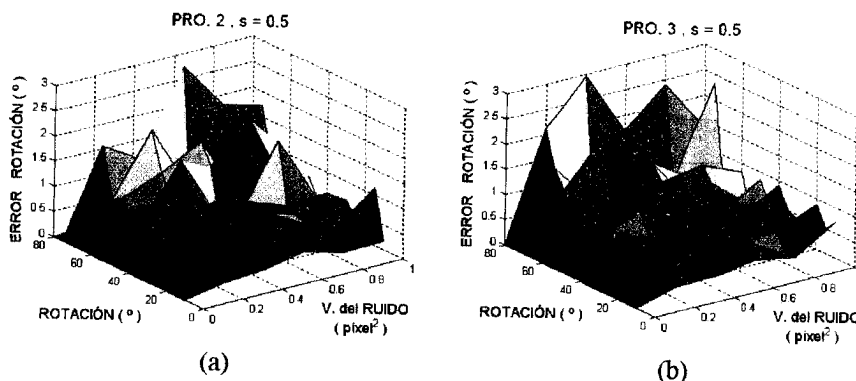
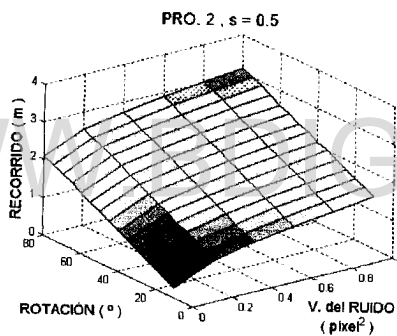


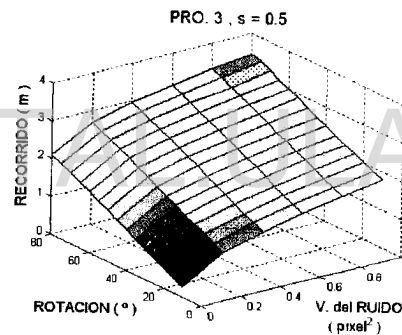
Figura C.7.2 EFR para la tarea 4 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.7.2 EFR promedio por nivel de varianza del ruido para la tarea 4, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (°)	Propuesta 3 (°)	Diferencia 3 (°)
0,0	0,0000	0,0000	0,0000
0,2	0,4572	0,5455	-0,0883
0,4	0,6667	0,6031	0,0636
0,6	0,7777	0,6646	0,1131
0,8	0,7325	0,6363	0,0962
1,0	0,6678	0,7877	-0,1199
MEDIA	0,5503	0,5395	0,0108
		\hat{i}	0,2708



(a)



(b)

Figura C.7.3 RRC para la tarea 4 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.7.3 RRC promedio por nivel de varianza del ruido para la tarea 4, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (m)	Propuesta 3 (m)	Diferencia 3 (m)
0,0	1,1957	1,1910	0,0047
0,2	1,6582	1,8246	-0,1664
0,4	1,8564	2,0968	-0,2404
0,6	2,0117	2,3011	-0,2894
0,8	2,1400	2,4777	-0,3377
1,0	2,2548	2,6326	-0,3778
MEDIA	1,8528	2,0873	-0,2345
		\hat{i}	4,1435

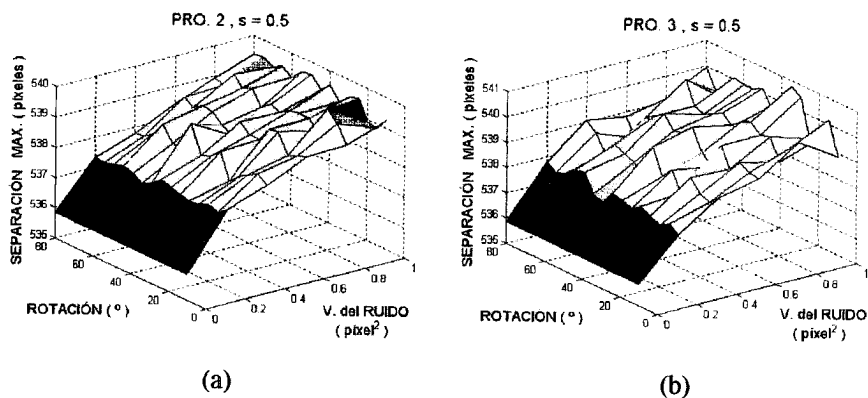


Figura C.7.4 SMPP para la tarea 4 y diferentes niveles de ruido, prueba 2 ($s = 0,5$).
(a) Propuesta 2, (b) Propuesta 3.

Tabla C.7.4 SMPP promedio por nivel de varianza del ruido para la tarea 4, prueba 2 ($s = 0,5$).

Varianza del ruido (píxel ²)	Propuesta 2 (píxel)	Propuesta 3 (píxel)	Diferencia 3 (píxel)
0,0	535,85	535,85	0,00
0,2	537,42	537,66	-0,24
0,4	538,15	538,45	-0,30
0,6	538,67	539,01	-0,34
0,8	539,04	539,42	-0,38
1,0	539,32	539,77	-0,45
MEDIA	538,08	538,36	-0,29
		\hat{t}	4,46