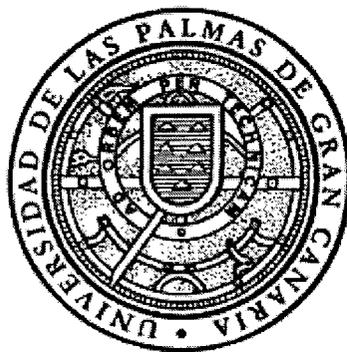


UNIVERSIDAD

DE

LAS PALMAS DE GRAN CANARIA

Departamento de Informática y Sistemas



TESIS DOCTORAL

**Estudio experimental de métodos de calibración y
autocalibración de cámaras**

José Isern González

Julio 2003

**Don Francisco J. Santana Pérez SECRETARIO DEL
DEPARTAMENTO DE Informática y Sistemas DE LA
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA,**

CERTIFICA,

Que el Consejo de Doctores del Departamento, en su sesión de fecha 17 de julio, tomó el acuerdo de dar el consentimiento para su tramitación a la tesis doctoral titulada "*Estudio experimental de métodos de calibración y autocalibración de cámaras.*" presentada por el doctorando Don José Isern González y dirigida por el Doctor Don Jorge Cabrera Gámez.

Y para que así conste, y a efectos de lo previsto en el Artº 73.2 del Reglamento de Estudios de Doctorado de esta Universidad, firmo la presente en Las Palmas de Gran Canaria, a veintitrés de julio de dos mil tres.



Francisco J. Santana Pérez

Anexo II

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Departamento: Informática y Sistemas

Programa de Doctorado: Tecnología de Visión por Computadora

Título de la Tesis

Estudio experimental de métodos de calibración y autocalibración de cámaras.

Tesis Doctoral presentada por D. José Isern González

Dirigida por el Dr. D. Jorge Cabrera Gámez

Codirigida por el Dr/a. D/D^a. _____

El/la Director/a,

El/la Codirector/a

El/la Doctorando/a,

(firma)

(firma)

(firma)



Las Palmas de Gran Canaria, a 8 de julio de 2003

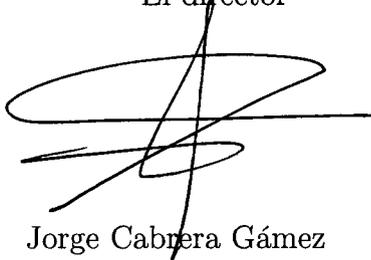
A handwritten signature in black ink, appearing to read 'José Isern González'.

UNIVERSIDAD
DE
LAS PALMAS DE GRAN CANARIA
Departamento de Informática y Sistemas

Tesis Titulada Estudio experimental de métodos de calibración y autocalibración de cámaras, que presenta D. José Isern González, realizada bajo la dirección del Doctor D. Jorge Cabrera Gámez.

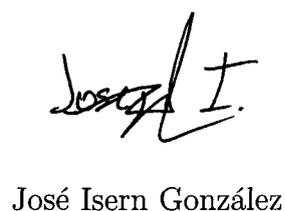
Las Palmas de Gran Canaria, Julio de 2003

El director



Jorge Cabrera Gámez

El doctorando



José Isern González

A Mónica.

Agradecimientos

Es difícil sentarse a escribir los agradecimientos, porque son tantas las personas que me han prestado su ayuda y colaboración en estos años que he tardado en realizar esta tesis que no quisiera dejarme a nadie atrás.

En primer lugar, quiero agradecer el trabajo del Dr. Jorge Cabrera Gámez en la dirección de esta tesis. No sólo por su trabajo propio de dirección de la investigación, sino también por sus consejos y su apoyo en los momentos de desánimo.

También quisiera dar mi agradecimiento al Grupo de Inteligencia Artificial y Sistemas creado por los doctores Antonio Falcón Martel, Francisco Mario Hernández Tejera y Juan Méndez Rodríguez. Ya que me han proporcionado un ambiente de aprendizaje y la posibilidad de trabajar en distintos proyectos que han propiciado el desarrollo de este documento. En particular quiero agradecer al resto de los miembros del grupo: Javier, Daniel, Modesto, Cayetano, Antonio Carlos, Irael y Óscar por su colaboración en este proceso de aprendizaje y su ayuda inestimada tanto en este trabajo como en el desarrollo de los proyectos que hemos llevado a cabo en común.

Por último, quiero agradecer a mi familia que han hecho posible con su amor y esfuerzo que pueda haber dedicado mi vida al estudio. Especialmente, quiero agradecer a mi esposa Mónica, que ha compartido conmigo el esfuerzo que ha supuesto sacar este trabajo adelante, el apoyo que me ha prestado en todo momento.

Gracias a todos.

Índice general

1. Introducción	1
1.1. Soluciones al problema de la calibración	2
1.1.1. Contribuciones	4
1.2. Organización del trabajo	6
2. Métodos de calibración clásicos	9
2.1. Modelo de cámara	14
2.1.1. Modelo Pin-hole	14
2.1.2. Modelo de Gauss	19
2.2. Descripción de los métodos	20
2.2.1. Método de Tsai [Tsai, 1987]	20
2.2.2. Método lineal	23
2.2.3. Método de Faugeras [Faugeras, 1993]	24
2.2.4. Método de Batista [Batista et al., 1998][Batista, 1996]	25
2.2.5. Método de Zhang [Zhang, 1998][Zhang, 2000]	30
2.2.6. Método de Heikkilä [Heikkilä and Silvén, 1997]	32
2.2.7. Método de Ahmed [Ahmed et al., 1999]	33
2.3. Experimentos realizados en el simulador	35
2.3.1. Inyección de ruido en los datos de entrada	39
2.3.2. Variación en la configuración de la cámara	49
2.3.3. Variación de la colocación del patrón	66
2.3.4. Comparación de tiempos de ejecución	85
2.3.5. Otros resultados	86
2.4. Conclusiones de los experimentos sobre el simulador	86
3. Métodos clásicos: experimentos con datos reales	89
3.1. Descripción de la toma de datos y resultados de la calibración	91
3.2. Precisión de los métodos	101
3.3. Estabilidad de los métodos	103

3.3.1.	Influencia de la configuración de la cámara en los parámetros extrínsecos	103
3.3.2.	Desplazamiento horizontal del patrón	104
3.4.	Variación en la configuración del patrón	107
3.5.	Variación del zoom	109
3.6.	Influencia del enfoque	112
3.7.	Desplazamiento de la distancia del patrón respecto a la cámara y zoom .	116
3.8.	Conclusiones del estudio sobre datos reales	120
3.9.	Experimentos en el simulador en las mismas circunstancias que los datos reales	122
3.9.1.	Variación de la distancia focal	123
3.9.2.	Influencia de la distorsión	126
3.9.3.	Influencia de la ortogonalidad	129
3.9.4.	Reproducción de los experimentos reales con condiciones aproximadas	130
3.10.	Comparación de los resultados a partir de datos reales con los de datos simulados	132
3.11.	Estudios realizados por otros autores	137
3.12.	Conclusiones	141
4.	Métodos de autocalibración	145
4.1.	Introducción	145
4.2.	Revisión bibliográfica	148
4.2.1.	Movimiento en general	148
4.2.2.	Desplazamientos en el mismo plano	151
4.2.3.	Movimientos específicos	152
4.2.4.	Sólo rotaciones	153
4.2.5.	Otros trabajos relacionados	155
4.3.	Métodos de autocalibración	157
4.3.1.	Método de McLauchlan [McLauchlan and Murray, 1996]	157
4.3.2.	Método Lineal de Agapito [de Agapito et al., 1999]	159
4.3.3.	Método iterativo de Agapito [de Agapito et al., 1998]	161
4.3.4.	Método de Seo-Hong [Seo and Hong, 1998]	161
4.4.	Experimentos en el simulador	163
4.4.1.	Análisis del coste computacional	166
4.4.2.	Estabilidad de los métodos	173
4.4.3.	Parámetros intrínsecos variables	174

4.4.4.	Influencia del nivel de ruido en las coordenadas 2D de la imagen	182
4.4.5.	Parámetros intrínsecos no modelados por los métodos	188
4.4.6.	Conclusiones	204
5.	Métodos de autocalibración: experimentos con datos reales	207
5.1.	Influencia de la longitud de la secuencia	208
5.2.	Estabilidad de los resultados	212
5.3.	Influencia del zoom	215
5.4.	Reproducción de las condiciones de los experimentos reales en el simulador	218
5.5.	Comparativa de resultados con los métodos clásicos de calibración	221
5.5.1.	Experimentos en el simulador	222
5.5.2.	Experimentos con una cámara real	223
5.6.	Conclusiones	231
6.	Conclusiones	233
6.1.	Futuras investigaciones	237
A.	Conceptos básicos	239
A.1.	Descomposición en Valores Singulares (SVD - Singular Value Decomposition)	239
A.2.	Normalización Isotrópica	240
A.3.	Homografía	240
A.4.	Medidas de error	241

Capítulo 1

Introducción

La calibración de una cámara es el proceso que permite, dentro del campo de la visión artificial, la obtención de los parámetros que definen las condiciones de formación de la imagen, incluyendo la geometría interna y la óptica de la cámara, así como su posición y orientación respecto a un objeto de referencia o patrón de calibración. En definitiva, la calibración es un procedimiento que trata de conocer cómo una cámara proyecta un objeto 3D en el plano imagen para así poder extraer información métrica a partir de las imágenes.

El uso de cámaras calibradas puede ser útil para resolver una serie de problemas relacionados con la obtención de la posición 3D de los objetos en el espacio a partir de sus imágenes o para su reconstrucción tridimensional. Esto puede permitir, entre otras tareas, la realización de mapas del entorno de la cámara, el seguimiento de un objeto específico o la obtención de la posición de la cámara respecto a objetos que la rodeen. También puede facilitar la navegación por su entorno de un robot móvil, permitiendo evitar obstáculos, dirigirse a objetos determinados o facilitar la definición de la trayectoria más adecuada para alcanzar su destino.

La calibración de una cámara es un problema complejo ya que muchos son los parámetros a resolver y muchos son los factores que influyen sobre los resultados. En parte, esta complejidad es acotada por los métodos de calibración empleando modelos de cámara que son, en realidad, modelos ideales o simplificados de sus equivalentes físicos. Con estas simplificaciones se consiguen resultados aproximados (aunque no exactos), pero provoca que no se parametrizen muchos factores. El modelo de cámara de agujero o "pin-hole", que es el utilizado por la mayoría de los métodos de calibración, no parametriza aspectos ópticos tales como la distancia de enfoque, la profundidad de campo, la apertura o la posible desalineación entre el plano imagen y la lente. Incluso, un efecto tan importante como la distorsión de las imágenes, se modela en la mayoría de los casos

de forma muy simplificada y no en todos los casos se parametriza.

Otros factores importantes son las posibles fuentes de ruido que influyen en el proceso de formación de la imagen, ya que la imagen que obtiene una cámara incorpora errores de cuantización, o de la precisión en la ubicación real de los elementos (puntos del patrón) utilizados para realizar la calibración. Debido a estas cuestiones el proceso de calibración de una cámara puede tener múltiples soluciones, donde los errores se compensan entre los grados de libertad (los parámetros del modelo), pudiendo obtenerse como resultado de la calibración una parametrización que, si bien minimiza una función objetivo final, es irreal desde el punto de vista físico.

1.1. Soluciones al problema de la calibración

A lo largo de los años se han un cierto número de soluciones para resolver el problema de la calibración. Las primeras vienen dadas por los métodos clásicos de calibración, los cuales se basan en el uso de patrones de puntos conocidos. Entre este tipo de métodos se han desarrollado varias tendencias. Por un lado, hay métodos que obtienen de forma explícita los parámetros físicos de la cámara [Heikkilä and Silvén, 1996] [Tsai, 1987], y por otro, métodos que los obtienen de forma implícita o agrupada [Ahmed et al., 1999] [Faugeras, 1993]. También se han desarrollado métodos que realizan la calibración utilizando patrones con los puntos distribuidos en uno o varios planos independientes [Zhang, 2000] [Batista et al., 1998] y en otros casos se utilizan patrones con los puntos distribuidos de forma que no estén en el mismo plano, habitualmente en forma de diedro [Ahmed et al., 1999] [Faugeras, 1993]. En algunos se han desarrollado métodos especializados en la calibración de cámaras con lentes que producen gran distorsión en las imágenes [Shah and Aggarwal, 1996] [Basu and Licardie, 1993]. En cualquier caso, un problema común a todos los métodos de calibración clásicos es que necesitan un patrón como referencia, lo que implica un proceso manual de selección de puntos y la dependencia de la cámara a un objeto externo a ella.

En los últimos años ha aparecido una nueva tipología de métodos que no necesitan un patrón como referente, denominados métodos de autocalibración [Maybank and Faugeras, 1992] [Luong et al., 1992] [Hartley, 1994b]. Estos métodos necesitan obtener una secuencia de imágenes y realizan la calibración en base a la correspondencia de los puntos entre las distintas vistas. Los primeros métodos de autocalibración imponían la restricción de que la configuración de la cámara no podía variar entre las vistas de la secuencia. Actualmente, ya se han desarrollado métodos, como los incluidos en [Heyden and Åström, 1997] [Pollefeys and Gool, 1999] [Seo and Hong, 2001] [de Agapito et al., 1999], donde se evita esta restricción. El problema que presentan este

tipo de métodos es que suelen utilizar modelos más simples de cámara, dejando sin parametrizar, en la mayoría de los casos, factores como la distorsión. Otro problema que se presenta es la precisión con que se realiza la correspondencia de puntos entre imágenes. La aparición de falsos positivos en este proceso implica que procesos de bajo nivel, los cuáles no son analizados en este trabajo, obtengan resultados deficientes. En estos caso se hace necesario, por ejemplo, el uso de algoritmos más robustos para obtener elementos como la homografía (ej: algoritmo de RANSAC [Fischler and Bolles, 1981]).

Ante la amplísima variedad de métodos de calibración, la primera cuestión que se plantea una persona cuando necesita calibrar una determinada cámara es cuál de los métodos debe utilizar. Es decir, qué método realizará mejor la calibración de su cámara. El problema que se presenta es que, aunque son muchos los métodos que se han desarrollado a lo largo de los años, son muy pocos los estudios realizados en los que se comparen diferentes métodos bajo las mismas condiciones experimentales e indiquen cuál es el mejor o el más apropiado.

Los pocos trabajos existentes [Tsai, 1989] [Armangué et al., 2000] [Salvi et al., 2002] [Grattoni et al., 1991] que analizan de forma comparativa métodos de calibración presentan estudios incompletos o parciales de éstos. En unos casos sólo se analiza el problema desde el plano teórico [Tsai, 1989]. En otros, se restringe el análisis a un único conjunto de datos reales [Armangué et al., 2000] [Salvi et al., 2002]. Pero, en ningún caso se hace un estudio global de varios métodos que permita concluir sobre la precisión de los métodos o su estabilidad frente a ruidos o efectos no modelados, tanto en condiciones simuladas como reales. Además, en la mayoría de los casos los métodos comparados son relativamente antiguos [Martins et al., 1981] [Hall et al., 1982] [Weng et al., 1992]. Por otro lado, hay trabajos que se han centrado en estudiar de forma detallada la influencia de algún determinado elemento en el proceso de calibración como [Lai, 1993] [Wang et al., 2001] [Hayman and Murray, 2002], pero estos trabajos son muy escasos.

Con frecuencia la calibración de una cámara en visión por computador se plantea en el contexto de un problema más amplio. Como "subproblema", el objetivo es resolverlo de forma rápida y eficaz, y el planteamiento natural es intentar aplicar el "mejor" método de calibración, es decir, aquel que garantice el mejor resultado posible. De un análisis exhaustivo de la bibliografía parece desprenderse que hoy por hoy no existen resultados publicados que permitan establecer que cierto método de calibración sea netamente superior a otros. En estas circunstancias, la cuestión original, cuál es el mejor método de calibración, se debe transformar en qué método puede ser aplicado con garantías en el contexto del problema particular a resolver, y cuál es el grado de precisión esperable. Esta cuestión plantea un problema adicional a la hora de realizar la calibración de una cámara real, y es cómo identificar en qué circunstancias se está realizando el proceso de

calibración para saber qué método a aplicar.

Por tanto, esta tesis surge ante la necesidad de dar respuesta a todas estas cuestiones. Con el objetivo de responder a la pregunta de qué método es el más adecuado, se ha investigado la precisión, robustez, autonomía y coste computacional de diversos métodos. La precisión es una medida de la corrección con que se estiman cada uno de los parámetros que modela un método. La robustez es un indicativo de lo bien que funciona un método en todas las situaciones. La autonomía expresa el nivel de automatización del proceso de calibración, es decir, de lo necesaria que puede ser un persona para guiar dicho proceso. Por último, el coste computacional mide el tiempo que necesita un método para completar el procedimiento y la cantidad de memoria que requiere.

1.1.1. Contribuciones

Esta tesis presenta, en primer lugar, un revisión bibliográfica extensa sobre los métodos de calibración y autocalibración de cámaras. Además, los algoritmos de los métodos que se han implementado para la realización son descritos ampliamente, presentándolos con una notación unificada. Este hecho es resaltable debido a que en la bibliografía se utilizan diferentes notaciones para referirse a los mismos elementos, cuestión que supone para los lectores de este tipo de trabajo un esfuerzo de comprensión adicional, y que en algún caso puede dar lugar a malas interpretaciones.

El estudio realizado establece cuáles son los condicionantes o elementos que influyen en la fiabilidad de cada método. Además, se indica qué procesos se pueden realizar para identificar qué factores están influyendo en un determinado proceso de calibración de una cámara real. También se analizan las ventajas y desventajas de los métodos de calibración que usan un patrón como referente frente a los métodos de autocalibración.

Para realizar el estudio se ha empleado tanto un simulador como una cámara real. La ventaja de emplear datos simulados es que se tiene un control sobre la configuración de la cámara y del entorno, lo cual, permite medir la precisión con que se obtienen cada uno de los parámetros del modelo de cámara utilizado. La desventaja es que se está utilizando un modelo de la realidad, el cuál no es más que una simplificación de la misma, por tanto, puede no ser fiel en todos los aspectos que influyen en el proceso de formación de las imágenes, aunque sí en los más importantes.

Para realizar los experimentos con datos simulados se ha construido un simulador en Matlab. Esta herramienta reproduce el proceso de formación de las imágenes utilizando como base el modelo pin-hole, e incluye todos aquellos aspectos que tienen en cuenta todos los métodos utilizados en la comparativa. Básicamente, permite definir la configuración de la cámara y la disposición de los puntos tridimensionales a procesar en

qué parámetros del modelo se han obtenido con más precisión y qué método ha sido el más fiable. En última instancia se ha desarrollado una guía de pasos o criterios a seguir que permiten realizar la identificación de los factores que inciden en mayor medida en los resultados. En los resultados obtenidos de los experimentos realizados para este estudio se demuestra que no existe un método que puede ser calificado como "el mejor", por lo menos en todas las condiciones.

Tras la realización de los experimentos se puede decir que los métodos de calibración analizados son fiables en el sentido de que, salvo en determinadas circunstancias, el error global que se alcanza en el proceso de calibración es pequeño; es decir, se puede realizar la reconstrucción 3D de objetos con bastante precisión (errores de milímetros). Pero, por el contrario, se muestra que los resultados de la calibración varían dependiendo de las condiciones en las que se realicen. Además, se demuestra que aunque el error en la reconstrucción de las coordenadas de los puntos es pequeño, en la mayor parte de los casos, los parámetros del modelo no se ajustan a la realidad. Se muestra que la desviación en la estimación de un parámetro se ve compensada por la desviación de otro, consiguiendo, no obstante, que el error global de reconstrucción sea mínimo. La manifestación más importante de este problema se presenta cuando la compensación del error en un parámetro intrínseco se produce por un parámetro extrínseco, lo que da como resultado "una buena calibración", pero que sólo es válida para el conjunto de datos que se ha utilizado para realizar el proceso. Esto se ve reflejado en experimentos realizados en los que variando sólo la posición de los puntos (del patrón o de la escena) respecto a la cámara, es decir, variando los parámetros extrínsecos, la estimación de los parámetros intrínsecos cambia de valor, y eso a pesar de que la configuración de la cámara permanece constante.

1.2. Organización del trabajo

Esta tesis se organiza como sigue: los dos siguientes capítulos muestran el estudio realizado sobre los métodos de calibración que utilizan un patrón como referente. En el primero de ellos se describe el modelo de cámara del que parte cada método para resolver la calibración, y los algoritmos de cada uno se los métodos empleados en la comparativa. Finalmente, se realiza el análisis de los experimentos realizados en el simulador. En el segundo de estos capítulos se analizan los resultados obtenidos de la calibración de una cámara real, y se muestra el proceso de identificación de los factores que han influido en los resultados de la calibración con datos reales, para lo cual se reproducen en el simulador algunos de los experimentos realizados con la cámara real.

El capítulo 4 está dedicado a los métodos de autocalibración. En primer lugar, se

relación a ésta y, en función de estos datos, obtener las coordenadas de imagen de estos puntos.

El estudio se ha restringido a la calibración de cámaras estáticas, es decir, sin movimientos de traslación, lo cual ha permitido realizar el análisis de dos tipos de métodos. Por un lado, aquellos que utilizan un patrón como referente, también denominados clásicos, en los cuales las coordenadas 3D de los puntos del patrón es conocida. Y por otro, los métodos de autocalibración basados en movimientos de rotación de la cámara. Estos dos tipos de métodos de calibración se han analizado primero por separado, y luego, se han comparado sus resultados.

Para la realización del estudio comparativo de los métodos de calibración clásicos se han seleccionado ocho métodos que fueran representativos de las diversas estrategias de resolución existentes para este problema. Se han seleccionado los métodos de Tsai [Tsai, 1987] (se incluyen dos versiones), Faugeras [Faugeras, 1993], Batista [Batista et al., 1998], Heikkilä [Heikkilä and Silvén, 1996], Ahmed [Ahmed et al., 1999], Zhang [Zhang, 2000], y por último, un método lineal, variante del método de Faugeras. Estos métodos son descritos en el capítulo 2 usando una notación unificada.

En el caso del estudio de los métodos de autocalibración se han seleccionado cuatro métodos: el método de McLauchlan [McLauchlan and Murray, 1996], el método de Seo [Seo and Hong, 1998] y el método de Agapito, del que se han implementado dos versiones, una lineal [de Agapito et al., 1999] y otra iterativa [de Agapito et al., 1998]. En este caso también, los diferentes métodos son descritos empleando una notación unificada.

La metodología seguida para la realización de la tesis se ha estructurado en varias fases. En primer lugar, se ha realizado un análisis de los factores que a priori pueden tener influencia sobre la calibración. A continuación, se ha estudiado, en un entorno controlado como es un simulador, la incidencia de estos factores sobre cada uno de los parámetros que modelan la cámara en los distintos métodos implementados. En este punto se ha analizado el efecto de factores como el ruido en los datos de entrada, variación en la configuración de la cámara, y la configuración y disposición del patrón en el caso de los métodos clásicos.

El marco experimental se ha trasladado posteriormente a un entorno real corroborar los resultados obtenidos en la fase de estudio en el simulador. El problema que se planteó en esta última fase era saber identificar qué factores están incidiendo en un determinado proceso de calibración y, en la medida de lo posible, cuantificarlos. La comparación de los resultados de la calibración a partir datos reales con aquellos obtenidos en los experimentos en el simulador ha permitido identificar los factores con mayor influencia en los resultados de la calibración con datos reales. La reproducción, en las mismas circunstancias en las que se realizó la calibración de datos reales, ha permitido conocer

realiza un recorrido por la bibliografía del tema, catalogando los diversos métodos de autocalibración en función de la variabilidad de los parámetros intrínsecos y del tipo de movimiento que realiza la cámara durante la calibración. Se describen con mayor detalle los algoritmos de los métodos empleados en la comparativa. Por último, se analizan los resultados obtenidos en los experimentos realizados en el simulador. En el capítulo 5, se presenta el análisis de los resultados obtenidos en las pruebas realizadas con los métodos de autocalibración sobre una cámara real. Al igual que en los métodos clásicos se reproduce en el simulador las condiciones de los experimentos con datos reales. Por último, los resultados de los experimentos de los métodos de autocalibración son comparados con los obtenidos por los métodos clásicos.

Finalmente, el último capítulo recoge las conclusiones del trabajo y se plantean futuras líneas de trabajo. Además, el apéndice A contiene algunas definiciones de conceptos básicos y se describen algoritmos de cálculo que son ampliamente empleados en los métodos de calibración de cámaras.

Capítulo 2

Métodos de calibración clásicos

En este capítulo se realiza un estudio de los métodos que usan un patrón como referente. Se ha agrupado estos métodos bajo la denominación genérica de "métodos clásicos" para distinguirlos de métodos más recientes de autocalibración.

Los métodos clásicos parten del conocimiento de la colocación de los puntos respecto a un sistema de coordenadas exterior a la cámara (situado en el patrón), es decir, de las coordenadas 3D de los puntos que integran el patrón, para reconstruir el proceso de formación de la imagen. Existen muchos factores que generan incertidumbre sobre los resultados de la calibración. Algunos de ellos están relacionados con el modelo de cámara utilizado, y otros con el proceso de generación de las imágenes.

El modelo de cámara que intentan resolver la mayoría de los métodos de calibración es el modelo de cámara de agujero o pin-hole (se realiza una completa descripción de este modelo en la sección 2.1). Esto es debido a que es el más simple y que requiere un menor número de parámetros. Pero este modelo presenta deficiencias en la explicación de la formación de las imágenes. Por ejemplo, el modelo no parametriza aspectos ópticos tales como la distancia de enfoque, la profundidad de campo o la apertura. No se modela o se realiza de forma muy simplificada el hecho de que el plano imagen y la lente puedan no encontrarse exactamente paralelos (Figura 2.1) en una cámara real. Esto produce que si hay un desplazamiento de la lente (zoom) o del plano imagen (enfoque), exista un desplazamiento del centro óptico, ya que el eje óptico variará el punto de incidencia sobre el plano imagen ([Willson, 1994]). En la mayoría de los casos se modela la distorsión de forma muy simplificada, y no en todos los casos se parametriza. Habitualmente, sólo se modela la distorsión radial y con un único coeficiente.

Por otro lado, el proceso físico de calibración es críticamente dependiente de los datos de entrada, o sea, las coordenadas tridimensionales de los puntos en el patrón y las coordenadas de sus proyecciones en el plano imagen. Entre estos problemas destaca:

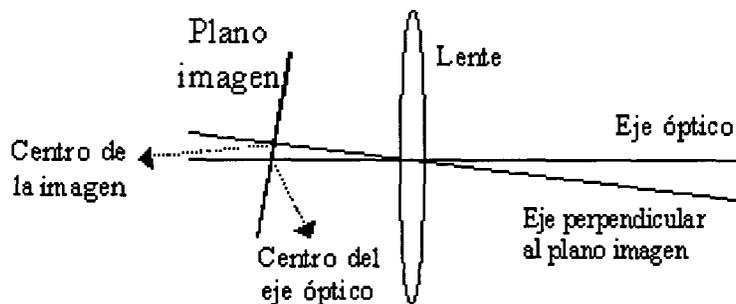


Figura 2.1: Desalineamiento del plano imagen respecto a la lente.

- El hecho de que las imágenes resultantes estén discretizadas introduce un ruido por el propio proceso de cuantización. Aunque, a priori, el ruido producido debe estar entorno al medio pixel, se han desarrollado métodos de procesamiento que permiten obtener una mayor precisión (precisión subpixel). Esto se puede conseguir, por ejemplo, utilizando como patrón un círculo y obteniendo el centroide de su proyección en la imagen como coordenadas 2D del patrón en la imagen. La iluminación de la escena, el tamaño de los puntos del patrón o la distancia a la que se encuentre el mismo son factores que influyen en el grado de precisión con que se obtiene el punto en la imagen. Esto es debido a que influyen, entre otras cosas, sobre el tamaño de los puntos proyectados en la imagen, y se supone que cuanto mayor sea el tamaño del punto en la imagen mayor será la precisión conseguida en la estimación del centroide del punto. También el número de puntos procesados en la imagen, es decir, el número de puntos que integren el patrón, puede ser un factor determinante, ya que, debido al carácter aleatorio del ruido se espera que la calidad de la calibración aumente con el número de puntos de referencia.
- No se puede conocer con exactitud la posición real de los puntos del patrón usado como referente. A priori, la precisión conseguida no debe ser superior a medio milímetro debido a los instrumentos de medición utilizados. En algunos casos, también pueden influir otros factores como, por ejemplo, que el ángulo que forman los dos planos del diedro, usado como patrón de referencia en la mayoría de los métodos clásicos, no sea exactamente de 90° , o que no se conozca con precisión la ubicación del origen de coordenadas común a dichos planos.

Los métodos de calibración clásicos se pueden agrupar en función de diferentes criterios:

- Computación lineal vs no lineal

Lineal. Usan técnicas de resolución de sistemas de ecuaciones lineales. Son muy simples de implementar y muy rápidos ([Abdel-Aziz and Karara, 1971] [Faugeras, 1993] [Hall et al., 1982]).

No lineal. Se basan en el uso de métodos iterativos. Generalmente, se requiere una buena aproximación inicial obtenida de un método lineal. Son mucho más lentos, pero permiten resolver modelos de cámara más complejos (v.g. modelar la distorsión) que incluyen un mayor número de parámetros ([Faugeras, 1993] [Heikkilä and Silvén, 1996] [Tsai, 1987]).

- Calibración explícita vs. implícita

Explícita. Se obtienen los valores de cada uno de los parámetros que forman el modelo ([Batista et al., 1998] [Heikkilä and Silvén, 1996] [Tsai, 1987]).

Implícita. Se obtienen, generalmente, matrices de transformación que contienen el conjunto de todos los parámetros. No se puede conocer el valor exacto de algunos parámetros ([Ahmed et al., 1999] [Faugeras, 1993] [Zhang, 1998]).

- Calibración en un paso vs. multipasos

Un solo paso. En cada ciclo del proceso de resolución se actualizan todos los parámetros a la vez ([Ahmed et al., 1999] [Faugeras, 1993]).

Multipaso. En cada fase se obtiene un conjunto distinto de parámetros, usándose aproximaciones en los primeros pasos para aquellos que aún no se hayan calculado y aplicándose los resultados que se van obteniendo en los siguientes pasos ([Batista et al., 1998] [Tsai, 1987] [Weng et al., 1992]).

- Patrón en un plano vs. múltiples planos

Un plano. Todos los puntos del patrón están en el mismo plano. Por tanto, tienen la ventaja de reducir el ruido en las coordenadas del patrón, ya que una de las coordenadas 3D es nula ([Batista et al., 1998] [Tsai, 1987]).

Múltiples planos. Dentro de este grupo se pueden distinguir dos tipos. Por un lado, aquellos que necesitan conocer la relación entre los planos, generalmente, se opta por que formen un diedro, es decir, dos planos que forman un ángulo de 90° (en este caso se introduce un elemento que puede generar errores como es el ángulo entre los dos planos [Ahmed et al., 1999] [Faugeras, 1993] [Tsai, 1987] [Heikkilä and Silvén, 1996]). Por otro lado, aquellos en los que

no es necesario conocer la relación entre las posiciones de los planos. Generalmente, pueden ser adquisiciones sobre el mismo patrón variando la colocación del mismo o realizando movimientos de la cámara ([Zhang, 1998] [Yu et al., 1998]).

El propósito de este estudio es realizar una comparación objetiva entre distintos métodos de calibración de una cámara que usan un patrón conocido como referente. Este trabajo surge ante la necesidad de seleccionar un método adecuado para realizar la calibración de una cámara, y, aunque son muchos los métodos que se han desarrollado a lo largo de los años, son muy pocos los estudios realizados en los que se comparen diferentes métodos bajo las mismas condiciones experimentales.

Previamente se han realizado otros trabajos en los que se ha efectuado una comparativa entre varios métodos de calibración. Tsai [Tsai, 1989] da unos criterios (velocidad, precisión y autonomía) para realizar la comparativa. Realiza una comparación a nivel teórico de cuatro métodos, todos ellos propuestos en 1987, y además cataloga otros más antiguos.

La diferencia de notación es un problema cuando se busca el método apropiado. En [Armangué et al., 2000] y [Salvi et al., 2002] se describen en detalle, usando una notación unificada, cuatro métodos diferentes y variantes de algunos de ellos (Tsai y Faugeras). En el estudio experimental de estos métodos no se analiza la influencia del ruido ni la precisión en la estimación de los parámetros del modelo de cámara empleado. Sólo se estudia la distancia entre los puntos reales obtenidos (tanto en 2D como en 3D) y la reconstrucción de éstos usando los valores de la calibración, pero no cómo se ve afectado cada uno de los parámetros que resulta de la calibración. Es infrecuente que las comparativas analicen la estabilidad de los parámetros de calibración ante cambios en las condiciones experimentales (configuración y colocación del patrón, configuración de la cámara, etc.), limitándose, por lo general, a cuantificar el error en la reconstrucción de las coordenadas 2D o 3D de los puntos (ec. A.5 y A.6). Una excepción es el trabajo de Lai [Lai, 1993] quien estudia, tanto desde el punto de vista teórico como experimental, cómo los parámetros extrínsecos se ven afectados por errores sobre algunos de los parámetros intrínsecos (centro del eje óptico, la distorsión y el factor de proporción). Este autor muestra que la distorsión de la lente y el centro del eje óptico tienen un efecto pequeño sobre la estimación de la posición y orientación del patrón, aunque el centro del eje óptico tiene mayor incidencia que la distorsión; por otro lado, el factor de proporción tiene una gran influencia sobre el cálculo de la posición del patrón.

Finalmente, Grattoni et al. [Grattoni et al., 1991] en un artículo en el que se comparan cinco métodos, todos ellos propuestos antes de 1991, realizan experimentos sobre

imágenes reales obtenidas de un patrón colocado sobre una plataforma móvil para controlar los desplazamientos. En este trabajo no se puede evaluar la precisión conseguida en los parámetros intrínsecos por cada método porque los experimentos fueron llevados a cabo con imágenes adquiridas con cámaras cuya configuración exacta no era conocida. Sólo tres de los métodos estiman los parámetros extrínsecos que son los que permiten una comparación de la precisión en términos de error en la reconstrucción de las coordenadas 3D de los puntos.

En este capítulo se comparan ocho métodos distintos. En unos casos los métodos se han seleccionado por ser clásicos dentro de la literatura como pueden ser el caso del método de Tsai [Tsai, 1987] o el de Faugeras [Faugeras, 1993]. Además de la versión clásica, se ha utilizado en el presente estudio una variante optimizada del método de Tsai, ya que la clásica no permite obtener todos los parámetros del modelo de cámara. En otros casos se ha buscado un método que cumpliera alguna característica en especial. Así, se ha buscado un método que fuera lo más simple posible como es el Lineal; otro que descomponga el problema en múltiples pasos como es el método de Batista [Batista et al., 1998]; uno especializado en focales cortas como es el método de Heikkilä [Heikkilä and Silvén, 1996]. También se ha incluido un método basado en redes neuronales, el método de Ahmed [Ahmed et al., 1999], por utilizar una aproximación completamente diferente al resto de los métodos. Por último, se ha incluido el método de Zhang [Zhang, 2000], aparte de por ser reciente, por ser un método que mezcla el uso de criterios de autocalibración con la utilización de patrones conocidos. En resumen, se ha intentado seleccionar métodos que fueran representativos de las diferentes aproximaciones que existen hoy en día para la resolución del problema de la calibración de cámaras mediante patrones.

El resto del capítulo se organiza en varias secciones. En primer lugar, se describe el modelo de cámara que utilizan como base los métodos utilizados. En segundo lugar, se describe cada uno de los métodos, mediante una nomenclatura unificada para permitir un mejor seguimiento de los mismos y facilitar la comparación de resultados. A continuación, sobre todos ellos se han efectuado en un simulador una serie de experimentos en los que se han variado las condiciones de la cámara y del patrón, y también se ha estudiado la influencia de ruidos que puedan afectar a la calibración. Los resultados de cada método se han comparado entre sí, observando tanto la precisión a la hora de reconstruir las coordenadas del patrón, como la precisión de los resultados de cada parámetro del correspondiente modelo respecto a la configuración establecida a priori. También se ha analizado el coste computacional que presenta cada uno de ellos. Finalmente, en la última sección se presentan las conclusiones del estudio.

2.1. Modelo de cámara

En este capítulo se hace una descripción de algunas formas de modelar el proceso de formación de imágenes en las cámaras. Se entiende por modelo de cámara aquel que explica cómo un punto 3D se proyecta en el plano imagen (2D) de una cámara. Los modelos de cámara, que se intentan sintetizar y resolver, son simplificaciones de la realidad. Esto se debe a la cantidad de tiempo y al coste computacional que supondría calibrar un modelo real, y a que con estas simplificaciones se consiguen resultados bastante aceptables (aunque no exactos). En trabajos como [Willson, 1994] [Jiménez, 2000] [Hartley and Zisserman, 2000] se puede encontrar una descripción más completa de la formación de las imágenes en una cámara y de modelos que la describen.

A continuación, se describe ampliamente el modelo de cámara que emplean los métodos utilizados en este estudio, denominado de agujero o pin-hole, y más brevemente el modelo de Gauss.

El último apartado de esta sección está dedicado a localizar las posibles fuentes de error en los resultados de la calibración, analizando tanto el proceso de formación de las imágenes como el modelo de cámara que utilizan los distintos métodos. Este análisis sirve como punto de partida a los experimentos realizados y que son analizados en las siguientes secciones.

2.1.1. Modelo Pin-hole

La gran mayoría de los procedimientos de calibración se basan en el modelo de cámara de **agujero** o **pin-hole**, que es el modelo más simple que se puede obtener de una cámara. De hecho, todos los métodos utilizados en este trabajo realizan la proyección en imagen siguiendo este modelo. Es el modelo es el más simple y, por tanto, él que necesita un menor número de parámetros para ser representado. Se basa en que la proyección de un punto de la escena se obtiene de la intersección de una línea que pasa por este punto y el centro de proyección (foco) con el plano imagen (Figura 2.2).

Básicamente, este modelo aplica una matriz de proyección para transformar las coordenadas 3D de los puntos del objeto en coordenadas 2D de la imagen:

$$m \doteq P \cdot M \quad (2.1)$$

donde $M = [X_w, Y_w, Z_w, 1]^t$ es el vector que contiene las coordenadas del punto en el sistema de referencia exterior a la cámara, P es una matriz de 3x4 denominada **matriz de proyección**, $m = [u, v, 1]^t$ es el vector de las coordenadas del punto proyectado en la imagen y transformado a pixels. El símbolo "≐" indica que los elementos de ambos

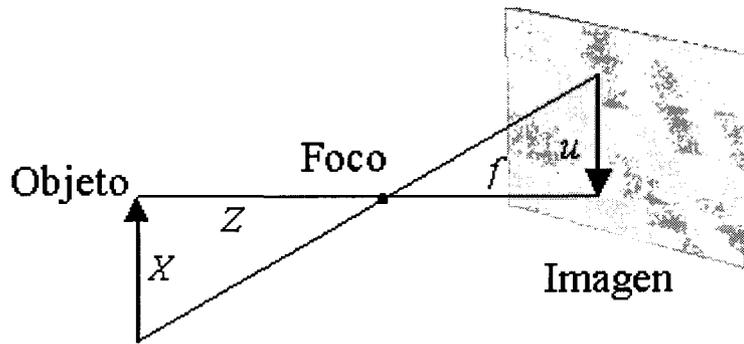


Figura 2.2: Modelo pin-hole.

lados son equivalentes excepto por un factor de proporcionalidad, es decir, $\lambda m = P \cdot M$, siendo λ el factor de escala.

Los parámetros que caracterizan una matriz de proyección se clasifican en dos tipos:

Parámetros intrínsecos. Son aquellos que definen el modelo de cámara utilizado, estableciendo sus características geométricas y ópticas. Los parámetros intrínsecos básicos son:

- Centro del eje óptico o punto principal (u_0, v_0) . Define el punto donde el eje óptico atraviesa el plano imagen. Las coordenadas de este punto vienen dadas en pixels.
- Factores de escalado (k_x, k_y) . Es un parámetro que indica la proporción de tamaño de un objeto visto en la realidad respecto a su proyección en el plano imagen. Esta proporción puede ser distinta en cada eje. Frecuentemente este parámetro se descompone a su vez en:
 - Factores de conversión pixel-milímetros (d_x, d_y) . Indica el número de pixels por milímetro que usa la cámara. Esta relación se obtiene dividiendo la dimensión en pixels de la imagen por el tamaño en mm. del CCD.
 - Distancia focal (f) . Distancia entre el foco y el plano imagen, viene dada en milímetros.
 - Factor de proporción (s_x) . Indica la relación de tamaño entre la dimensión horizontal y vertical de un pixel.

La relación entre ellos (cuando no existe distorsión) viene dada por: $k_x = s_x d_x f$ y $k_y = d_y f$.

- Ortogonalidad de los ejes del plano imagen o asimetría (Ω_1, Ω_2) . Geométricamente es un parámetro que mide el ángulo que forman los dos ejes de la imagen (Figura 2.3), es decir, intenta medir el grado de ortogonalidad de los

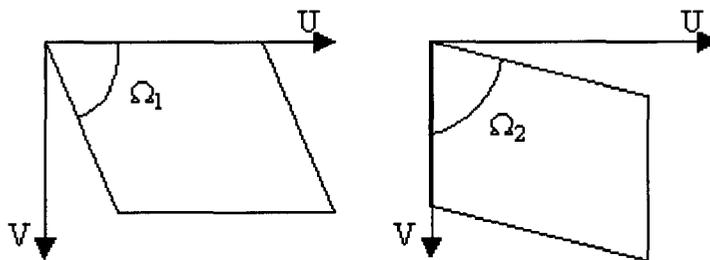


Figura 2.3: Representación de los ángulos de la asimetría.

ejes del plano imagen. En el caso ideal este ángulo debe ser de 90° , es decir, estos ejes deben ser perpendiculares. En la mayoría de los casos sólo se utiliza uno de los ángulos ($\Omega = \Omega_1$). En situaciones reales la desviación respecto a la ortogonalidad se puede producir cuando la lente de la cámara no es paralela al plano imagen [Hartley and Zisserman, 2000].

- **Distorsión.** El efecto de la distorsión en imágenes se modela mediante dos componentes: una radial (k_1, k_2, \dots) y otra tangencial (p_1, p_2, \dots). Cada una de ellas se desarrolla como una serie infinita de términos, aunque en la mayoría de los casos sólo se utiliza la distorsión radial, la cuál se modela sólo con uno o dos de esos términos.

Parámetros extrínsecos. Son los que definen la relación (traslación y rotación) entre un sistema de coordenadas absoluto (exterior a la cámara) y otro ligado a la cámara. Se incluyen 6 parámetros: tres para la traslación (T_x, T_y, T_z) y tres para los ángulos rotados sobre cada uno de los ejes (α, β, γ).

Proyección de un punto 3D en una imagen 2D

La transformación de las coordenadas 3D de un punto en el sistema de referencia situada en el mundo a uno situado en la cámara y su proyección al plano imagen de ésta (coordenadas 2D) se puede observar en la figura 2.4 y se desarrolla en una secuencia de cuatro pasos:

1. Transformación del punto en coordenadas 3D del mundo a coordenadas 3D de la cámara. Primero se realiza una rotación alrededor de los tres ejes:

$$\begin{bmatrix} X_r \\ Y_r \\ Z_r \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} \quad (2.2)$$

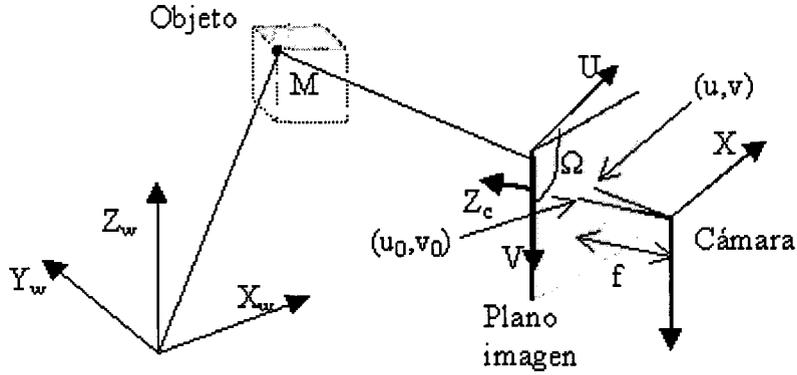


Figura 2.4: Modelo de cámara. Proyección de puntos en la imagen.

seguida de una traslación:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} X_r \\ Y_r \\ Z_r \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \quad (2.3)$$

Expresado en forma matricial se tiene:

$$M_c = R \cdot M + T \quad (2.4)$$

donde M es el vector que contiene las coordenadas del punto en sistema de coordenadas del mundo, T es el vector de traslación y R es la matriz de rotación. En todos los métodos, la matriz de rotación es el resultado de aplicar sucesivamente giros alrededor de cada uno de los ejes, por lo que $R = R_z R_y R_x$, quedando como:

$$R = \begin{bmatrix} C_\beta C_\gamma & S_\alpha S_\beta C_\gamma - C_\alpha S_\gamma & C_\alpha S_\beta C_\gamma + S_\alpha S_\gamma \\ C_\beta S_\gamma & S_\alpha S_\beta S_\gamma + C_\alpha C_\gamma & C_\alpha S_\beta S_\gamma - S_\alpha C_\gamma \\ -S_\beta & S_\alpha C_\beta & C_\alpha C_\beta \end{bmatrix} \quad (2.5)$$

donde S representa el seno y C el coseno del ángulo correspondiente, y α es el giro alrededor del eje X_w , β sobre el eje Y_w y γ sobre Z_w . La matriz de rotación debe ser una matriz ortonormal.

2. Proyección sobre el plano imagen aplicando el modelo pin-hole.

$$u_p = f \frac{X_c}{Z_c} \quad v_p = f \frac{Y_c}{Z_c} \quad (2.6)$$

donde f es la distancia focal.

3. Distorsión de la lente. La distorsión varía en función de la distancia de cada punto al centro del eje óptico.

$$r^2 = u_p^2 + v_p^2 \quad (2.7)$$

Distorsión radial:

$$\begin{aligned} u_r &= u_p(k_1 r^2 + k_2 r^4 + \dots) \\ v_r &= v_p(k_1 r^2 + k_2 r^4 + \dots) \end{aligned} \quad (2.8)$$

donde k_i son los coeficientes de distorsión radial.

Distorsión tangencial:

$$\begin{aligned} u_t &= 2p_1 u_p v_p + p_2 (r^2 + 2u_p^2) \\ v_t &= p_1 (r^2 + 2v_p^2) + 2p_2 u_p v_p \end{aligned} \quad (2.9)$$

donde p_i son los coeficientes de distorsión tangencial. Los resultados de modelar cada tipo de distorsión se suman a las coordenadas sin distorsión, es decir:

$$\begin{aligned} u_d &= u_p + u_r + u_t \\ v_d &= v_p + v_r + v_t \end{aligned} \quad (2.10)$$

4. Conversión de unidades (paso de milímetros a pixels) y desplazamiento del origen al centro del eje óptico.

$$\begin{aligned} u &= d_x s_x u_d + u_0 \\ v &= d_y v_d + v_0 \end{aligned} \quad (2.11)$$

donde d_x y d_y son los factores de conversión, el resultado del cociente entre el tamaño de la imagen (pixels) y el tamaño del CCD (milímetros); s_x es un factor de proporción del pixel; y por último, u_0 y v_0 son las coordenadas del centro del eje óptico. Hay métodos como [Faugeras, 1993] que incluyen un factor que mide la ortogonalidad del plano imagen (Ω) transformándose la formula anterior en la siguiente:

$$\begin{aligned} u &= d_x s_x (u_d - \cot g(\Omega) v_d) + u_0 \\ v &= d_y \frac{v_d}{\text{sen}(\Omega)} + v_0 \end{aligned} \quad (2.12)$$

y también hay métodos como [Batista et al., 1998] que miden este factor mediante dos ángulos (Ω_1, Ω_2):

$$\begin{aligned} u &= \frac{d_x s_x \text{sen} \Omega_1}{\cos(\Omega_1 - \Omega_2)} u_d - \frac{d_x s_x \cos \Omega_1}{\cos(\Omega_1 - \Omega_2)} v_d + u_0 \\ v &= \frac{d_y \cos \Omega_2}{\cos(\Omega_1 - \Omega_2)} u_d + \frac{d_y \text{sen} \Omega_2}{\cos(\Omega_1 - \Omega_2)} v_d + v_0 \end{aligned} \quad (2.13)$$

En resumen, en caso de no tener en cuenta la distorsión (cosa que ocurre frecuentemente y en especial en los métodos de autocalibración), la matriz de proyección se

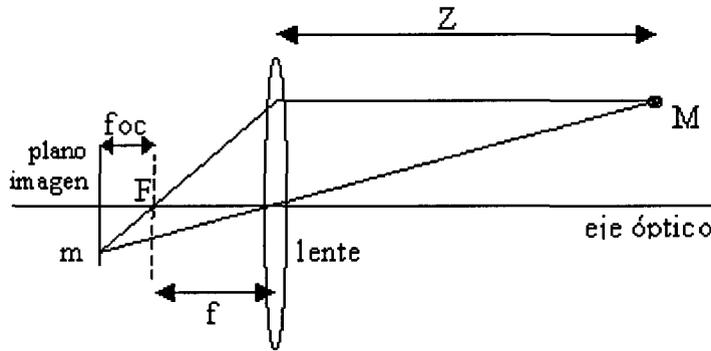


Figura 2.5: Modelo de Gauss.

puede descomponer en matrices, una para los parámetros intrínsecos (A) y otra para los extrínsecos (D):

$$P = AD \quad (2.14)$$

donde la matriz de parámetros intrínsecos es:

$$A = \begin{bmatrix} k_x & -k_x \cot \Omega & u_0 & 0 \\ 0 & k_y / \sin \Omega & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.15)$$

siendo $k_x = s_x d_x f$ y $k_y = d_y f$; y la matriz de parámetros extrínsecos está formada por:

$$D = [RT] = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.16)$$

siendo R la matriz de rotación y T el vector de traslación.

2.1.2. Modelo de Gauss

El modelo de **lente delgada** o de **Gauss** [Hecht and Zajac, 1979] intenta resolver algunos aspectos de una cámara real que el modelo pin-hole no incorpora. Este último se limita a reproducir únicamente los efectos del zoom a través de la variación de la distancia focal. El modelo de lente delgada permite parametrizar el zoom (distancia focal), pero, además, modela la distancia de enfoque, la profundidad de campo, y la apertura.

La formación de la imagen en este modelo se basa en el hecho de que todos los rayos de luz paralelos al eje óptico convergen en el foco. La proyección de un punto en el plano

imagen se obtiene como la intersección del rayo paralelo al eje óptico que pasa por dicho punto y del rayo que pasa a través del centro óptico de la lente (Figura 2.5).

Este modelo plantea la siguiente relación entre la distancia de focal nominal (f) y la distancia de enfoque (foc):

$$\frac{1}{f} = \frac{1}{f + foc} + \frac{1}{Z} \quad (2.17)$$

donde Z es la distancia del objeto M a la cámara, y $f + foc$ es la distancia focal efectiva.

Como distancia focal nominal suele utilizarse la aportada por el fabricante, en cambio, la distancia focal efectiva depende de la distancia de enfoque que haya en cada instante. En el modelo de pin-hole estas dos distancias focales se parametrizan como una sola ya que no se contempla el enfoque.

2.2. Descripción de los métodos

En esta sección se describen los algoritmos de cada uno de los métodos de calibración de una cámara que se ha utilizado en la comparativa.

2.2.1. Método de Tsai [Tsai, 1987]

Es el método más referenciado en la literatura y probablemente el primero en incluir la distorsión. Se basa en el modelo pin-hole y para corregir la distorsión utiliza un único coeficiente, que sólo considera la distorsión radial. El sistema que plantea tiene 9 incógnitas que se obtienen de forma explícita: 6 extrínsecas (rotación y traslación del patrón) y 3 intrínsecas (distancia focal, coeficiente de distorsión, factor de escala). El método tiene dos fases, usando computación lineal en el primer paso y en el segundo un método de optimización iterativo.

Las dos fases del método se descomponen en las siguientes etapas:

1. **Cálculo de la orientación del patrón, la traslación (T_x, T_y), y el factor de proporción (s_x).** Se realiza la conversión de pixels a milímetros en función de los valores aportados por el fabricante de la cámara y situando el centro del eje óptico en el centro de la imagen se obtienen u_d y v_d . Suponiendo nula la distorsión se plantea la siguiente equivalencia $\frac{u_d}{v_d} = \frac{X_c}{Y_c}$ (la relación entre las coordenadas transformadas del plano imagen es igual a la relación de las coordenadas del objeto en el sistema de coordenadas de la cámara), obteniéndose la siguiente ecuación:

$$\begin{bmatrix} v_d M^t & v_d & -u_d M^t \end{bmatrix} a = u_d \quad (2.18)$$

siendo M las coordenadas del objeto en el sistema de coordenadas del mundo, mientras que u_d y v_d se obtienen despejando de la ecuación 2.11 y donde

$$a = \begin{bmatrix} s_x r_{11}/T_y \\ s_x r_{12}/T_y \\ s_x r_{13}/T_y \\ s_x T_x/T_y \\ s_x r_{21}/T_y \\ s_x r_{22}/T_y \\ s_x r_{23}/T_y \end{bmatrix} \quad (2.19)$$

siendo r_{ij} los elementos de la matriz de rotación (ec. 2.2). Resolviendo esta ecuación por mínimos cuadrados se puede extraer las dos primeras filas de dicha matriz, mientras que la tercera se obtiene del producto vectorial de las otras dos. Los valores de T_x, T_y, s_x se pueden extraer de la siguiente forma:

$$|T_y| = \frac{1}{\sqrt{a_5^2 + a_6^2 + a_7^2}} \quad (2.20)$$

Para conocer el signo de T_y se selecciona un punto cuyas coordenadas se encuentren lejos del centro del eje óptico, se transforman sus coordenadas del mundo a coordenadas de la cámara usando T_y con signo positivo, si tanto el signo de X_c y de Y_c obtenidos coinciden, respectivamente, con los de u_d y v_d de ese punto. Entonces el signo de T_y es positivo, en otro caso es negativo. Por otro lado, el valor de s_x se obtiene de:

$$s_x = \sqrt{a_1^2 + a_2^2 + a_3^2} \cdot |T_y| \quad (2.21)$$

Las dos primeras filas de la matriz de rotación y la traslación en el eje X se obtienen de:

$$\begin{aligned} r_{11} &= a_1 T_y / s_x \\ r_{12} &= a_2 T_y / s_x \\ r_{13} &= a_3 T_y / s_x \\ r_{21} &= a_4 T_y \\ r_{22} &= a_4 T_y \\ r_{23} &= a_4 T_y \\ T_x &= a_4 T_y \end{aligned} \quad (2.22)$$

2. **Cálculo de la distancia focal (f), del coeficiente de distorsión (k_1) y la traslación en Z (T_z).** Para obtener la focal y la traslación se sigue suponiendo

nula la distorsión. Partiendo de $v_d = f \frac{Y_c}{Z_r + T_z}$ (es la ecuación 2.6 reescrita) se obtiene la siguiente ecuación:

$$\begin{bmatrix} Y_c & -v_d \end{bmatrix} \begin{bmatrix} f \\ T_z \end{bmatrix} = Z_r v_d \quad (2.23)$$

Al resolver esta ecuación, se obtienen los valores que sirven de aproximación inicial a un método de optimización, el cuál refinará estos valores junto con el coeficiente de distorsión. Para ello, se intenta minimizar mediante un proceso iterativo la siguiente ecuación:

$$\frac{v_f - v_0}{d_y} (1 + k_1 r) - f \frac{Y_c}{Z_r + T_z} = 0 \quad (2.24)$$

donde

$$r = \left(\frac{u_f - u_0}{s_x d_x} \right)^2 + \left(\frac{v_f - v_0}{d_y} \right)^2 \quad (2.25)$$

El método permite utilizar como patrón un diedro o un plano, pero en el segundo caso debe existir una inclinación de más de 30 grados respecto al plano imagen de la cámara. En este estudio se han realizado los experimentos utilizando un diedro como patrón.

Uno de los problemas que plantea este método es que da por hecho que el centro del eje óptico se encuentra en el centro de la imagen. Otro inconveniente es que sólo se refinan los parámetros obtenidos en la segunda fase del método, mientras el resto (s_x, T_x, T_y, R) absorben el error que se comete al estimarlos suponiendo la distorsión nula.

En este trabajo, aparte del método original, también se ha utilizado una versión optimizada de este método, la cuál se basa en añadir dos pasos adicionales a los descritos aquí. El primero consiste en utilizar un método iterativo para optimizar todos los parámetros a la vez, si bien se sigue suponiendo el centro del eje óptico en el centro de la imagen. Para este proceso de optimización se toman como aproximación inicial los valores obtenidos en las dos fases anteriores ya descritas del algoritmo original. El último paso consiste en repetir el proceso de optimización, pero esta vez incluyendo también el centro del eje óptico. En ambos casos se utiliza el método de Levenberg-Marquardt [More, 1977] para realizar la optimización. Para la versión optimizada se ha utilizado el código generado por X. Armangue and J. Salvi para la realización los distintos experimentos del presente trabajo, el cual se puede obtener en: "<http://eia.udg.es/%7Earmangue/research/>". En las gráficas, esta versión aparece como "Tsai-LM".

2.2.2. Método lineal

Es un método clásico y bien conocido de resolución de la calibración usando computación lineal. El método es bastante similar al método lineal de Faugeras [Faugeras and Toscani, 1986]. Frecuentemente se usa como aproximación inicial para obtener la matriz de transformación (P), como ocurre en [Zhang, 2000]. Una vez obtenida la matriz, se ha utilizado la descomposición descrita en [Faugeras, 1993] para obtener los distintos parámetros.

Este método resuelve la calibración mediante la descomposición en valores singulares de una matriz que forma el sistema de ecuaciones que proyecta un punto del espacio en el plano imagen de la cámara. Lo que se obtiene es una matriz de proyección a partir de la cual se extraen todos los parámetros, tanto los extrínsecos (traslación y rotación del patrón respecto a la cámara) como los intrínsecos (centro del eje óptico, factores de conversión y grado de ortogonalidad del plano imagen o asimetría).

Partiendo de $m = P \cdot M$ (ecuación 2.1), y redefiniendo P (es una matriz de 3×4) como:

$$P = \begin{bmatrix} q_1 & q_{14} \\ q_2 & q_{24} \\ q_3 & q_{34} \end{bmatrix} \quad (2.26)$$

El objetivo del método es minimizar el error que se produce entre las coordenadas reales del objeto en la imagen y las obtenidas como consecuencia de la proyección de los puntos del objeto usando la matriz de transformación. Para ello, hay que resolver el siguiente sistema de ecuaciones:

$$\begin{bmatrix} M^t & 0^t & -u \cdot M^t \\ 0^t & M^t & -v \cdot M^t \end{bmatrix} x = 0 \quad (2.27)$$

donde $x = [q_1, q_{14}, q_2, q_{24}, q_3, q_{34}]^t$. Con lo que queda una ecuación matricial del tipo $L \cdot x = 0$. La solución a esta ecuación viene dada por el autovector de $L^t \cdot L$ asociado al autovalor más pequeño.

La matriz resultante se normaliza, es decir, se dividen todos los elementos de la matriz por $\|q_3\|$, que se identifica con la norma de los tres primeros elementos de la última fila de P , correspondientes con la última fila de la matriz de rotación. En caso de que el elemento q_{34} sea negativo hay que cambiar de signo toda la matriz ¹. A partir de esta matriz se pueden obtener cada uno de los parámetros siguiendo la descomposición realizada en [Faugeras, 1993], proceso que aparece descrito en el siguiente apartado (ecuación 2.30).

La principal ventaja de este método es que es muy rápido. Como desventaja pre-

¹Este elemento se corresponde con T_z y por tanto siempre debe ser positivo ya que el objeto se encuentra delante de la cámara.

senta que no incluye la distorsión. En muchos métodos iterativos se pueden utilizar los resultados de este método como aproximación inicial.

2.2.3. Método de Faugeras [Faugeras, 1993]

Es otro de los métodos que con más frecuencia es referenciado en la literatura sobre este tema. Se basa en la resolución de un sistema de ecuaciones en forma matricial. Los parámetros se obtienen de forma implícita, ya que el resultado de la calibración es una matriz de transformación de 3x4 en la que se incluyen todos los parámetros, y en algunos casos no se puede extraer un parámetro en concreto sino una combinación de los mismos. Sólo necesita una imagen del patrón y no modela la distorsión. Presenta dos alternativas para la resolución del problema, una lineal y la otra no lineal, aunque demuestra que esta última es mucho más estable ante la presencia de ruido y es la que se ha utilizado para realizar el presente estudio.

En este método se obtiene primero una aproximación mediante un método lineal (en este trabajo se ha utilizado el método descrito en el punto anterior), y luego mediante un proceso iterativo se refinan los valores de la matriz de proyección perspectiva. A partir de ella se obtienen los parámetros intrínsecos y extrínsecos de la cámara. Los parámetros que se modelan son los seis extrínsecos (tres de desplazamiento y tres de orientación) y cinco intrínsecos (coordenadas del centro del eje óptico (u_0, v_0) , factores de escalado (k_x, k_y) y ortogonalidad de los ejes de la imagen o asimetría (Ω)).

El método trata de minimizar el error en la estimación de los N puntos del patrón mediante la siguiente función:

$$Error = \sum_{i=1}^N \left\| \frac{q_1 M_i + q_{14}}{q_3 M_i + q_{34}} - u_i \right\|^2 + \left\| \frac{q_2 M_i + q_{24}}{q_3 M_i + q_{34}} - v_i \right\|^2 \quad (2.28)$$

donde la matriz de proyección es:

$$P = \begin{bmatrix} q_1 & q_{14} \\ q_2 & q_{24} \\ q_3 & q_{34} \end{bmatrix} \quad (2.29)$$

Esta optimización esta sujeta a que $\|q_3\|^2 = 1$, lo que indica que la matriz de rotación, implícita en esta matriz de proyección, debe ser ortogonal.

Para obtener los valores de los parámetros de la calibración, la matriz resultante se

puede descomponer siguiendo los siguientes pasos: ¹

$$\begin{aligned}
T_z &= q_{34} \\
r_3 &= q_3 \\
u_0 &= q_1 q_3^t \\
v_0 &= q_2 q_3^t \\
c_1 &= q_1^t \otimes q_3^t \\
c_2 &= q_2^t \otimes q_3^t \\
\Omega &= \arccos\left(-\frac{c_1 \cdot c_2}{\|c_1\| \|c_2\|}\right) \\
k_x &= \|c_1\| \operatorname{sen} \Omega \\
k_y &= \|c_2\| \operatorname{sen} \Omega \\
r_2 &= q_3^t \otimes \frac{c_2}{\|c_2\|} \\
r_1 &= \frac{1}{\operatorname{sen} \Omega} \left(q_3^t \otimes \frac{c_1}{\|c_1\|} + q_3^t \otimes \frac{c_2}{\|c_2\|} \cos \Omega \right) \\
T_y &= (q_{24} - v_0 q_{34}) \frac{\operatorname{sen} \Omega}{k_x} \\
T_x &= (q_{14} + (q_{24} - v_0 q_{34}) \frac{k_x}{k_y} \cos \Omega - u_0 q_{34}) \frac{1}{k_x}
\end{aligned} \tag{2.30}$$

2.2.4. Método de Batista [Batista et al., 1998][Batista, 1996]

Es un método de calibración basado en la detección de puntos en un plano y sólo necesita la adquisición de una imagen. Los parámetros intrínsecos y extrínsecos se obtienen de forma explícita. El método es iterativo y se resuelve en varias fases usando computación lineal. Es necesaria una aproximación inicial para la mayoría de los pasos. Por ejemplo, al inicio se supone que la distorsión es nula, la ortogonalidad perfecta, el factor de proporción igual a uno, el centro óptico coincide con el centro de la imagen y los factores de conversión son obtenidos de los datos ofrecidos por el fabricante. La distorsión se modela con un coeficiente (k_1) y la asimetría con dos (Ω_1 y Ω_2).

El modelo de cámara en el que se basa el método de Batista difiere en pequeños detalles respecto al utilizado por el resto de los métodos que se incluyen en este trabajo. Por un lado, para obtener la matriz de rotación, utilizada para alinear el sistema de coordenadas del objeto con el de la cámara, se siguen los siguientes pasos: primero se realiza un giro alrededor del eje Z_w (ϕ), luego alrededor del eje Y_w (θ), quedando el objeto paralelo al plano imagen con el eje Z_w apuntando hacia la cámara, finalmente se rota de nuevo sobre el eje Z_w (ψ) de tal forma que los ejes X e Y coincidan con los ejes U y V de la imagen (Figura 2.6). En la siguiente ecuación se puede ver la matriz de rotación empleada en este método:

¹Nota: El símbolo \otimes denota el producto vectorial.

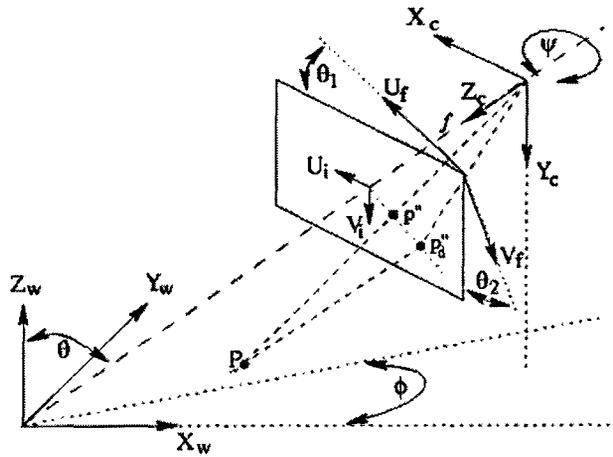


Figura 2.6: Modelo aplicado en el método de Batista [Batista, 1996].

$$R = \begin{bmatrix} -C_\psi S_\phi + S_\psi C_\phi C_\theta & C_\psi C_\phi + S_\psi S_\phi C_\theta & -S_\psi S_\theta \\ S_\psi S_\phi + C_\psi C_\phi C_\theta & -S_\psi C_\phi + C_\psi S_\phi C_\theta & -C_\psi S_\theta \\ -S_\theta C_\phi & -S_\theta S_\phi & -C_\theta \end{bmatrix} \quad (2.31)$$

Por otro lado, se usa la siguiente transformación para modelar la distorsión (sólo se modela la radial):

$$\begin{aligned} u_d &= \frac{2u_p}{1 + \sqrt{1 - 4k_1 r}} \\ v_d &= \frac{2v_p}{1 + \sqrt{1 - 4k_1 r}} \end{aligned} \quad (2.32)$$

El proceso de calibración se divide en varias etapas, cada una de ellas dedicadas a la obtención de ciertos parámetros. En estas etapas los parámetros son obtenidos mediante ecuaciones lineales. Se utiliza un plano como patrón y sólo se necesitan las cuatro esquinas de un rectángulo para obtener la orientación del patrón respecto a la cámara. La calibración se descompone en los siguientes cuatro pasos:

1. Obtención de los ángulos de rotación.
2. Traslación en X e Y, y obtención del factor de escala horizontal (S_x).
3. Cálculo de la focal, coeficiente de distorsión radial y traslación en Z.
4. Estimación del factor de escala, centro de la imagen y desviación de la ortogonalidad del eje de coordenadas de la imagen.

El procedimiento itera en dos puntos concretos: el primero finaliza cuando el factor de escala horizontal (S_x) converge a uno, y el segundo cuando la media de la distancia euclídea entre las coordenadas cada punto proyectado usando el modelo y las coordenadas reales de cada punto en la imagen (disparidad de la imagen) converge a un mínimo.

En cada bucle se repiten los cuatro pasos descritos anteriormente, aunque en el primero no se actualizan la mayoría de los parámetros intrínsecos (factor de conversión vertical, centro del eje óptico y ángulos de asimetría).

Para obtener los ángulos de rotación se supone que el origen del sistema de coordenadas del mundo coincide con el de la cámara, es decir, la traslación es nula. En este paso, los factores de conversión y el centro de la imagen se suponen conocidos (información aportada por el fabricante). Aunque el método es iterativo y se van actualizando todos los parámetros, la obtención de los ángulos de rotación se sigue realizando sobre estos valores predefinidos, lo que confiere estabilidad al método, ya que no permite que diverjan los valores de ninguno de los parámetros. En la primera iteración también se supone que el factor de proporción es 1, la distorsión es nula (en las siguientes iteraciones se utilizan las coordenadas de imagen con la distorsión corregida para realizar este proceso) y la distancia focal viene dada por el fabricante. Para calcular estos ángulos sólo se utilizan cuatro puntos que deben formar un rectángulo en el patrón; es importante que estos puntos se encuentren lo más separados entre sí que sea posible ya que esto disminuirá la incidencia del ruido que puedan tener los puntos seleccionados en la calibración.

Partiendo de la matriz de rotación que se usa en el modelo planteado en este método (ecuación 2.31) se obtienen los ángulos de la siguiente forma:

$$\psi = \arctan \frac{A(v_{p4} - v_{p3}) + B(v_{p1} - v_{p2}) + C(v_{p2} - v_{p4}) + D(v_{p3} - v_{p1})}{A(u_{p3} - u_{p4}) + B(u_{p2} - u_{p1}) + C(u_{p4} - u_{p2}) + D(u_{p1} - u_{p3})} \quad (2.33)$$

donde

$$A = \frac{u_{p1}v_{p2} - u_{p2}v_{p1}}{E} \quad B = \frac{u_{p3}v_{p4} - u_{p4}v_{p3}}{E} \quad C = \frac{u_{p3}v_{p1} - u_{p1}v_{p3}}{F} \quad D = \frac{u_{p4}v_{p2} - u_{p2}v_{p4}}{F} \quad (2.34)$$

$$E = f \cdot [(u_{p3} - u_{p4})(v_{p1} - v_{p2}) - (u_{p1} - u_{p2})(v_{p3} - v_{p4})]$$

$$F = f \cdot [(u_{p4} - u_{p2})(v_{p1} - v_{p3}) - (u_{p3} - u_{p1})(v_{p2} - v_{p4})]$$

siendo $u_{p,i=1:4}$ y $v_{p,i=1:4}$ las coordenadas de imagen transformadas de las esquinas de un rectángulo. Para obtener el valor de θ se puede utilizar cualquiera de las siguientes dos fórmulas:

$$\theta = \arctan \frac{f[(u'_3 - u'_4)(v'_1 - v'_2) - (u'_1 - u'_2)(v'_3 - v'_4)]}{(u'_1v'_2 - u'_2v'_1)(v'_3 - v'_4) - (u'_3v'_4 - u'_4v'_3)(v'_1 - v'_2)} \quad (2.35)$$

$$\theta = \arctan \frac{f[(u'_4 - u'_2)(v'_1 - v'_3) - (u'_3 - u'_1)(v'_2 - v'_4)]}{(u'_3v'_1 - u'_1v'_3)(v'_2 - v'_4) - (u'_4v'_2 - u'_2v'_4)(v'_1 - v'_3)} \quad (2.36)$$

donde u' y v' se corresponden con las coordenadas X e Y de los puntos del rectángulo una vez rotado sobre Z_w e Y_w (ángulos ϕ y θ), o visto de otra forma, son las coordenadas ideales del plano imagen (sin distorsión) a las que se ha aplicado una rotación sobre el

eje Z de ψ grados, por tanto, se obtienen de:

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} \cos\psi & -\operatorname{sen}\psi \\ \operatorname{sen}\psi & \cos\psi \end{bmatrix} \begin{bmatrix} u_p \\ v_p \end{bmatrix} \quad (2.37)$$

El último ángulo se obtiene como sigue:

$$\phi = \arctan \frac{\operatorname{sen}\theta[(u'_3v'_4 - u'_4v'_3)(u'_1 - u'_2) - (u'_1v'_2 - u'_2v'_1)(u'_3 - u'_4)]}{f[(v'_1 - v'_2)(u'_3 - u'_4)(u'_1 - u'_2)]} \quad (2.38)$$

Los valores de ψ y ϕ tienen un rango de 2π , mientras que los valores de θ tienen un rango de $\pi/2$. Si el valor de θ se encuentra en el rango $[-\pi/2, 0]$ entonces hay que obtener el simétrico de θ y rotar π radianes el ángulo ψ .

La traslación (T_x, T_y) y el factor de escala horizontal (S_x) se obtienen al resolver por mínimos cuadrados una ecuación lineal. En ella se relacionan las coordenadas de imagen de todos los puntos, transformadas a milímetros pero suponiendo la distorsión nula, con los puntos del patrón rotado, los cuáles se obtienen aplicando la matriz de rotación obtenida en el paso anterior a las coordenadas 3D. Una vez calculado S_x se puede actualizar el factor de conversión horizontal d_x .

$$\begin{bmatrix} d_x(v_f - v_0) & -d_y(u_f - u_0) & d_x(v_f - v_0)X_r \end{bmatrix} \begin{bmatrix} S_x T_x \\ T_y \\ S_x \end{bmatrix} = \begin{bmatrix} d_y(u_f - u_0)Y_r \end{bmatrix} \quad (2.39)$$

Para obtener la profundidad del objeto (T_z) y el coeficiente de distorsión radial se resuelve por mínimos cuadrados una ecuación lineal. En esta ecuación se relacionan las coordenadas de imagen, transformadas a milímetros mediante los parámetros actualizados, con las coordenadas del objeto respecto a la cámara, también utilizando los valores actualizados.

$$\begin{bmatrix} X_c f R_d^2 & -u_d \\ Y_c f R_d^2 & -v_d \end{bmatrix} \begin{bmatrix} k_1 \\ T_z \end{bmatrix} = \begin{bmatrix} u_d Z_r - X_c f \\ v_d Z_r - Y_c f \end{bmatrix} \quad (2.40)$$

donde $R_d^2 = u_d^2 + v_d^2$.

El último parámetro en actualizarse es la distancia focal efectiva, para lo que utiliza el modelo de lente de Gauss. Según este modelo la distancia focal se descompone en dos: distancia focal de la lente, que coincide con la proporcionada por el fabricante y es fija, y la distancia de enfoque. Esto proporciona dos ventajas, por un lado, permite modelar cámaras con zoom variable, y por otro lado, da estabilidad al método ya que los valores de la focal serán siempre próximos al proporcionado por el fabricante y se evita que el valor de la focal caiga a valores próximos a cero. Hay

que tener en cuenta que en muchos métodos de calibración existe el problema de que al minimizar el error, la relación entre focal y profundidad del objeto (f/T_z) tiende a cero. Se resuelve por mínimos cuadrados la ecuación lineal que consta de dos componentes, uno proporcionado por el modelo pin-hole y el otro proporcionado por el modelo de lente de Gauss:

$$\begin{bmatrix} \sqrt{X_c^2 + Y_c^2} \\ Z_c - F \end{bmatrix} f = \begin{bmatrix} Z_c \sqrt{u_p^2 + v_p^2} \\ Z_c F \end{bmatrix} \quad (2.41)$$

donde $f = F + foc$, es decir la distancia focal efectiva (f) es igual a la suma de la distancia focal de la lente (F), que la suministra el fabricante de la misma, y distancia de enfoque (foc).

Estos pasos se repiten dentro del bucle interno hasta que el factor de escala horizontal haya convergido, pasándose entonces al bucle externo. En el segundo bucle se repite el mismo proceso, pero los ángulos de rotación se siguen calculando en función de los valores de los parámetros iniciales, excepto para la focal, que utiliza su nuevo valor. Las ecuaciones empleadas en esta fase son las mismas que las anteriores, salvo (2.39) que se sustituye por:

$$\begin{bmatrix} u_p & -v_p \end{bmatrix} \begin{bmatrix} T_x \\ T_y \end{bmatrix} = [u_p Y_r - v_p X_r] \quad (2.42)$$

Además, y a diferencia de lo que ocurría en el bucle anterior, en este segundo proceso se actualizan los valores de los parámetros intrínsecos. Este paso es el último dentro de este proceso iterativo y para calcular los nuevos valores de estos parámetros se resuelve el siguiente sistema lineal:

$$\begin{bmatrix} u_d & 0 & v_d & 0 & 1 & 0 \\ 0 & u_d & 0 & v_d & 0 & 1 \end{bmatrix} a = \begin{bmatrix} u_f \\ v_f \end{bmatrix} \quad (2.43)$$

donde $a = \left[\frac{d_x \text{sen} \Omega_1}{\cos(\Omega_1 - \Omega_2)}, \frac{d_y \text{cos} \Omega_2}{\cos(\Omega_1 - \Omega_2)}, \frac{-d_x \text{cos} \Omega_1}{\cos(\Omega_1 - \Omega_2)}, \frac{d_y \text{sen} \Omega_2}{\cos(\Omega_1 - \Omega_2)}, u_0, v_0 \right]^t$.

En definitiva, éste es un método iterativo en el que se resuelven en distintas etapas y de forma lineal cada uno de los parámetros. Tiene como problema la dependencia inicial que existe de los parámetros proporcionados por el fabricante, y que para alguno de ellos no se utilizan valores actualizados, sino que siempre se utilizan los iniciales. Si estos valores iniciales están muy lejos de la realidad el error cometido puede ser grande. Por contra, esto genera estabilidad ante la presencia de ruidos, ya que no permite que parámetros interdependientes se alejen mucho de la realidad al irse compensando los errores sobre los valores de los diferentes parámetros.

2.2.5. Método de Zhang [Zhang, 1998][Zhang, 2000]

Tras obtener una aproximación inicial mediante un método lineal, se realiza un refinamiento iterativo empleando un criterio de máxima probabilidad. La calibración se realiza sobre un plano de puntos, y aunque no es necesario conocer el desplazamiento que se realiza entre tomas de imágenes, para que el método funcione correctamente se necesitan al menos tres imágenes con el patrón tomado en distintas orientaciones. Este número de imágenes puede ser inferior si se fijan los valores de algunos parámetros intrínsecos. Por ejemplo, si no se calcula la ortogonalidad del plano imagen, sólo son necesarias dos imágenes. Los parámetros se obtienen implícitamente, es decir, lo que se obtiene es una matriz cuyos elementos son función de los parámetros intrínsecos. La distorsión que se incluye es de tipo radial y se modela con dos coeficientes.

El método sigue los siguientes pasos:

1. Se calcula la matriz de proyección (P) de cada imagen en dos pasos. Primero se obtiene una aproximación inicial de la matriz de proyección mediante un método lineal. El método utilizado es el mismo que se ha descrito en este trabajo como Método Lineal (2.27), salvo que se elimina la componente Z_w ya que al trabajar sobre planos este valor siempre va a ser nulo. La matriz resultante es, por tanto, de 3×3 , a diferencia de los otros métodos en los que esta matriz es de 3×4 ya que trabajan con diedros. Esta matriz se obtiene resolviendo:

$$\begin{bmatrix} X_w & Y_w & 1 & 0 & 0 & 0 & -u \cdot X_w & -u \cdot Y_w & -u \\ 0 & 0 & 0 & X_w & Y_w & 1 & -v \cdot X_w & -v \cdot Y_w & -v \end{bmatrix} x = 0 \quad (2.44)$$

donde $x = [p_1, p_2, p_3]^t$ siendo p_i la i -ésima fila de la matriz P . De esta forma se consigue una aproximación inicial que luego se intenta refinar mediante un método iterativo. En éste se trata de minimizar el error cometido al proyectar la imagen, es decir, minimizar:

$$\sum \|m_i - \hat{m}_i\|^2 \quad (2.45)$$

donde m_i son las coordenadas reales de cada punto del patrón en la imagen obtenida, mientras que \hat{m}_i es la proyección en la imagen del patrón aplicando los valores calibrados y se obtiene de:

$$\hat{m}_i = \frac{1}{p_3 M_i} \begin{bmatrix} p_1 M_i \\ p_2 M_i \end{bmatrix} \quad (2.46)$$

2. Se obtiene la matriz que define la 'Cónica Absoluta' (B) a partir de la cual se obtiene la matriz de parámetros intrínsecos A (ec. 2.15). Siendo la relación entre

ambas matrices: $B = A^{-t}A^{-1}$. Como B es una matriz simétrica sólo hay que obtener seis de sus elementos (triangular superior), para eso se convierte esta matriz en un vector: $b = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]$. Debido a que en la matriz de rotación las columnas son ortonormales entre si, se presentan dos limitaciones, por un lado $p_1^t A^{-t} A^{-1} p_2 = 0$ y por otro $p_1^t A^{-t} A^{-1} p_1 = p_2^t A^{-t} A^{-1} p_2$, donde p_i es la i -ésima columna de la homografía. Uniendo estas limitaciones a las ecuaciones anteriores se tiene que para obtener los valores del vector b hay que resolver el siguiente sistema:

$$\begin{bmatrix} v_{12}^t \\ (v_{11} - v_{22})^t \end{bmatrix} b = 0 \quad (2.47)$$

siendo

$$v_{ij} = [p_{1i}p_{1j}, p_{1i}p_{2j} + p_{2i}p_{1j}, p_{2i}p_{2j}, p_{3i}p_{1j} + p_{1i}p_{3j}, p_{2i}p_{3j} + p_{3i}p_{2j}, p_{3i}p_{3j}]^t \quad (2.48)$$

Esta ecuación es del tipo $Vb = 0$, por tanto se pueden resolver calculando el autovector de V^tV asociado a su autovalor más pequeño (o lo que es lo mismo, el vector singular derecho de V asociado a su valor singular más pequeño). Una vez se conoce el vector b se pueden obtener los valores de la matriz de parámetros intrínsecos:

$$\begin{aligned} v_0 &= (b_2b_4 - b_1b_5)/(b_1b_3 - b_2^2) \\ \lambda &= b_6 - [b_4^2 + v_0(b_2b_4 - b_1b_5)]/b_1 \\ \alpha &= \sqrt{\lambda/b_1} \\ \beta &= \sqrt{\lambda b_1/(b_1b_3 - b_2^2)} \\ \gamma &= -b_2\alpha^2\beta/\lambda \\ u_0 &= \gamma v_0/\alpha - b_4\alpha^2/\lambda \end{aligned} \quad (2.49)$$

3. Los parámetros extrínsecos se obtienen despejando de $P = A \cdot \begin{bmatrix} r'_1 & r'_2 & T \end{bmatrix}$, donde r'_i son las columnas de la matriz de rotación R y T es el vector de traslación.

$$\begin{aligned} r'_1 &= \lambda A^{-1} p'_1 \\ r'_2 &= \lambda A^{-1} p'_2 \\ r'_3 &= r'_1 \wedge r'_2 \\ T &= \lambda A^{-1} p'_3 \end{aligned} \quad (2.50)$$

con $\lambda = 1/\|A^{-1}p'_1\|$. Por supuesto este resultado es distinto para cada una de las imágenes empleadas.

4. Se estiman los coeficientes de distorsión radial por mínimos cuadrados, resolviendo

las siguientes ecuaciones para cada punto de cada imagen:

$$\begin{bmatrix} (u_f - u_0)r & (u_f - u_0)r^2 \\ (v_f - v_0)r & (v_f - v_0)r^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} u - u_f \\ v - v_f \end{bmatrix} \quad (2.51)$$

donde $r = (X_c^2 + Y_c^2)/Z_c^2$. Hay que hacer notar que el modelo de cámara en que se basa este método supone la corrección de la distorsión entre la aplicación de la matriz de parámetros extrínsecos y la matriz de parámetros intrínsecos 2.14. Es decir, primero se transforma del sistema de coordenadas del mundo al de la cámara (matriz de parámetros extrínsecos), luego se realiza la corrección de la distorsión, y por último, se obtienen las coordenadas 2D de la imagen (matriz de parámetros intrínsecos). La diferencia con otros métodos estriba en que se corrige la distorsión antes de realizar la proyección de los puntos en el plano imagen, y por tanto las coordenadas no están afectas por la distancia focal. Por ese motivo, los valores de los coeficientes de distorsión no son comparables al del resto de los métodos que obtienen este parámetro.

5. Por último, se refinan todos los parámetros mediante un método iterativo que minimiza el criterio de máxima probabilidad. Es decir, se trata de minimizar la siguiente función:

$$\sum \| \begin{bmatrix} u & v \end{bmatrix} - \begin{bmatrix} u_f & v_f \end{bmatrix} \|^2 \quad (2.52)$$

es decir, la diferencia entre las coordenadas de cada punto de cada una de las imágenes reales con respecto a la proyección realizada con los valores de todos los parámetros (incluidos los coeficientes de distorsión) obtenidos en la calibración de forma que todos ellos se puedan actualizar.

2.2.6. Método de Heikkilä [Heikkilä and Silvén, 1997]

Es un método orientado a cámaras con distancia focal muy pequeña y en los que, por tanto, el tipo de lente introduce una importante distorsión en la imagen. Corrigen la distorsión radial y tangencial con polinomios de segundo orden.

El método consta de cuatro pasos, aunque en el estudio realizado en el simulador sólo se reflejan los dos primeros ya que los otros afectan a la extracción de los puntos de las imágenes. Se utiliza un método lineal como aproximación inicial a los parámetros extrínsecos, a la distancia focal y al centro del eje óptico; y a continuación usa un método iterativo para ajustar el valor de estos parámetros y calcular el resto de los parámetros (factor de proporción, coeficientes de distorsión radial y tangencial). En el tercer paso realizan una corrección sobre la imagen de puntos del patrón, ya que estos

no se proyectan de forma homogénea sobre la imagen al no estar el patrón paralelo al plano imagen [Heikkilä and Silvén, 1996] y en el cuarto controlan la iluminación.

En el primer paso se estiman los parámetros aplicando un método lineal, el método de Transformación Lineal Directa (DLT) desarrollado por [Abdel-Aziz and Karara, 1971], en el que se plantea una ecuación igual a (2.27), aunque para extraer los parámetros usa el método propuesto en [Melem, 1994] basado en descomposición RQ, donde dada la matriz P obtenida de aplicar el algoritmo DLT, se descompone en: $A = \lambda V^{-1} B^{-1} F R T$ siendo R y T las matrices de rotación y traslación, y el resto contiene los parámetros intrínsecos:

$$V = \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 0 & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 + k_1 & k_2 & 0 \\ k_2 & 1 - k_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad F = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.53)$$

En el segundo paso, se minimiza el siguiente funcional mediante un método iterativo de optimización (método de Levenberg-Marquardt [More, 1977]):

$$\sum (u - \hat{u})^2 + (v - \hat{v})^2 \quad (2.54)$$

es decir, la suma de los cuadrados de las diferencias entre el punto real (u, v) y el estimado aplicando el modelo (\hat{u}, \hat{v}) .

En este trabajo se ha utilizado el código generado por Janne Heikkilä (Calibration toolbox v2.1b) para realizar los distintos experimentos. Este software puede ser obtenido en la siguiente dirección: "<http://www.ee.oulu.fi/~jth/calibr/>".

2.2.7. Método de Ahmed [Ahmed et al., 1999]

Es un método de calibración basado en redes neuronales. A priori no se considera la distorsión, aunque se indica que es factible añadir algún coeficiente que la corrija o utilizar algún método que realice la corrección sobre la imagen capturada antes del proceso de calibración.

Tiene como ventaja que no necesita una aproximación inicial a los parámetros, aunque en los experimentos realizados en este trabajo se ha utilizado una aproximación inicial (Método Lineal) para acelerar el proceso de calibración. Como desventaja presenta la lentitud característica del entrenamiento de una red neuronal. Es, además, muy dependiente de los factores de aprendizaje (α_1 y α_2).

El método entrena una red neuronal que intenta obtener la matriz de proyección que transforma los puntos 3D del patrón en coordenadas en pixels de la imagen. La red neu-

ronal utilizada es de dos niveles con propagación hacia adelante, donde las coordenadas 3D de los puntos del patrón son la entrada y la coordenadas 2D de imagen son la salida. Existen dos matrices de pesos asociadas a cada nivel (V, W), relacionándose una con la matriz de parámetros extrínsecos (primer nivel o nivel oculto), y la otra con la matriz de parámetros intrínsecos (segundo nivel o nivel de salida).

La salida del nivel oculto viene dada por:

$$Y = VM_i \quad (2.55)$$

Mientras que la salida del segundo nivel viene dada por:

$$\hat{m}_i = \gamma_i WY \quad (2.56)$$

donde $M_i = [X_w, Y_w, Z_w, 1]^t$ son las coordenadas homogéneas del i -ésimo punto del patrón y $\hat{m}_i = [\hat{u}, \hat{v}, \hat{w}]$ son las coordenadas ¹ de su proyección en la imagen. Existe un factor de escalado (γ_i) asociado a los puntos del patrón, teniendo un valor distinto para cada uno de ellos. Este proceso se repite para cada uno de los puntos del patrón hasta que la medida de error (2.57) converge a un mínimo.

El entrenamiento de la matriz asociada al nivel oculto (V) viene restringido a que la submatriz que corresponde con la matriz de rotación sea ortogonal. Por lo que la medida del error tiene en cuenta esta limitación:

$$E_{tot} = E_i + \beta E_{ort} \quad (2.57)$$

donde E_i mide la disparidad de cada punto de la imagen:

$$E_i = \frac{1}{2} d_i^2 \quad (2.58)$$

donde $d_i = (\hat{u} - u)^2 + (\hat{v} - v)^2 + (\hat{w} - 1)^2$ mide la distancia entre las coordenadas reales y las estimadas. β es un factor de ponderación, su valor es pequeño y positivo. Por último, E_{ort} mide la no ortogonalidad de la matriz de rotación:

$$\begin{aligned} E_{ort} = & \sum_{l=1}^3 (V_{1l}^2 + V_{2l}^2 + V_{3l}^2 - a)^2 \\ & + (\sum_{c=1}^3 V_{c1} V_{c2})^2 \\ & + (\sum_{c=1}^3 V_{c1} V_{c3})^2 \\ & + (\sum_{c=1}^3 V_{c2} V_{c3})^2 \end{aligned} \quad (2.59)$$

siendo a una constante que debe ser igual a 1, salvo que la matriz V esté escalada.

¹ \hat{w} es un factor de escala de las coordenadas de imagen.

Cada vez que se procesa un punto del patrón se actualizan las matrices de ponderación siguiendo la regla del gradiente descendiente. En la actualización se utilizan unas constantes de aprendizaje (α_1, α_2) que se definen positivas. La actualización se obtiene de:

$$\Delta W = \alpha_1(m_i - \hat{m}_i)\gamma_i Y \quad (2.60)$$

Para actualizar la matriz V se secciona la matriz en dos, de tal forma que las columnas de la submatriz $V_{1:3,1:3}$, que se corresponde con la matriz de rotación, se actualizan de la siguiente forma:

$$\Delta V_l = \alpha_1 \gamma_i d_i W M_i - 4\alpha_1 \beta (V_{1l}^2 + V_{2l}^2 + V_{3l}^2 - a)V - 2\alpha_1 \beta \delta v_l \quad (2.61)$$

donde

$$\begin{aligned} \delta v_{j1} &= (V_{j2} + V_{j3})[(\sum_{c=1}^3 V_{c1}V_{c2}) + (\sum_{c=1}^3 V_{c1}V_{c3})] \\ \delta v_{j2} &= (V_{j1} + V_{j3})[(\sum_{c=1}^3 V_{c1}V_{c2}) + (\sum_{c=1}^3 V_{c2}V_{c3})] \\ \delta v_{j3} &= (V_{j1} + V_{j2})[(\sum_{c=1}^3 V_{c1}V_{c3}) + (\sum_{c=1}^3 V_{c2}V_{c3})] \end{aligned} \quad (2.62)$$

Para obtener la última columna de V (que se corresponde con el vector de traslación), se utiliza:

$$\Delta V_4 = \alpha_1 \gamma_i d W M_i \quad (2.63)$$

Por último, para actualizar el factor de escalado (γ_i) se utiliza:

$$\Delta \gamma_i = \alpha_2 d(WY) \quad (2.64)$$

Los valores de entrada y salida de la red son normalizados (S_1, S_2), es decir, se divide las coordenadas 2D y 3D de los puntos por su norma (s_1, s_2 , respectivamente). Por tanto, la matrices de normalización son matrices diagonales de tal forma que: $S_1 = \text{diag}(s_1, s_1, 1)$ y $S_2 = \text{diag}(1/s_2, 1/s_2, 1/s_2, 1)$. Cuando el método converge se obtiene la matriz de proyección como:

$$P = S_1 W V S_2 \quad (2.65)$$

La matriz de proyección se resuelve usando, por ejemplo, la descomposición que realiza [Faugeras, 1993] (ya descrita en la sección dedicada a este método), obteniéndose así los valores de los distintos parámetros.

2.3. Experimentos realizados en el simulador

Se han realizado diversos experimentos con el objetivo de conocer la estabilidad y robustez de cada uno de los métodos descritos, así como la fiabilidad del modelo en el

que se basan.

Los experimentos realizados se pueden clasificar en 3 tipos. Por un lado, los que intentan comprobar como se comportan los distintos métodos cuando los datos de entrada están afectados por ruido. Por otro lado, los relacionados con la variación de la configuración de la cámara. Y por último, los relacionados con la colocación del patrón respecto a la cámara.

Para simular los problemas derivados del desconocimiento de la ubicación exacta de los datos de entrada al sistema (coordenadas 2D de la imagen o 3D de los puntos en el espacio), se ha inyectado ruido gaussiano de media nula y desviación típica σ , tanto sobre las coordenadas de imagen como sobre las coordenadas 3D de los puntos del patrón. En el primer caso, se intenta simular los efectos de digitalización sobre la imagen, y en el segundo caso, la mala estimación que se puede producir en la posición real de los puntos del patrón, un aspecto siempre sujeto a un cierto nivel de error.

El ruido introducido en los experimentos descritos es aleatorio en cada punto según la curva de probabilidad que forma la gaussiana, por lo que para asegurar que el promedio de los valores introducidos se aproxime a la desviación típica propuesta se ha repetido cada experimento al menos 50 veces.

Otro tipo de experimento ha sido la inclusión, en el modelo simulado, de parámetros que no se tienen en cuenta en la resolución de alguno de los métodos. Este es el caso del coeficiente de distorsión radial que no se obtiene en los métodos de Faugeras, Ahmed y Lineal; la desviación respecto a la ortogonalidad del plano imagen, que no es estimada por los métodos de Tsai, Tsai-optimizado y Heikkilä; y el centro del eje óptico es fijado en el centro de la imagen en el método de Tsai.

También se han realizado experimentos que varían otros aspectos que a priori podrían afectar al resultado de la calibración de alguno de los métodos empleados, como por ejemplo, la posición en la que se encuentra el patrón, la orientación del mismo o la distancia focal. Se han estudiado estos parámetros ya que afectan al tamaño y colocación del patrón en la imagen.

Para realizar los experimentos, se ha simulado en Matlab cómo una cámara proyecta los puntos de un objeto en coordenadas 3D en el plano imagen. Para realizar la proyección se ha usado el modelo pin-hole, que es común a todos los métodos. Se han realizado estas pruebas en un simulador ya que este permite realizar una comprobación exhaustiva de la fiabilidad de cada método; es decir, se pueden comparar los resultados obtenidos en cada uno de los parámetros con la configuración real de la cámara. Esto no sería posible en un experimento realizado con imágenes reales, ya que se desconoce la configuración de la cámara. En la medida de lo posible se ha utilizado el código proporcionado por los autores u otro obtenido en Internet, previamente usado en otros análisis.

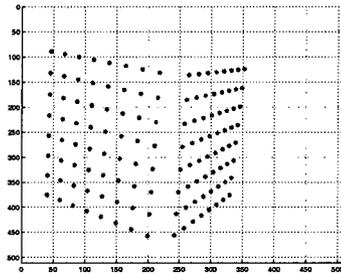


Figura 2.7: Proyección de la distribución de puntos del diedro en el plano imagen.

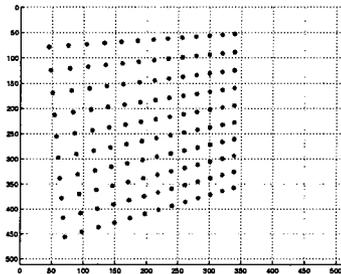


Figura 2.8: Proyección en la imagen de la distribución de puntos del plano usado para el método de Batista.

En todos los experimentos realizados, y para todos los métodos utilizados, se ha usado la misma configuración de cámara para proyectar el patrón de calibración. El plano imagen es de 512x512 pixels situando el centro del eje óptico ligeramente desplazado respecto al centro de la imagen (255,255). La distancia focal es de 11.25 mm y los ejes del plano imagen son ortogonales. El tamaño del CCD es de 8.5 mm x 5.8 mm y el factor de escala es de 0.994. No se ha introducido inicialmente ninguna distorsión.

En cuanto a los parámetros extrínsecos, la forma y colocación del patrón han variado, ya que hay métodos que utilizan un diedro mientras otros necesitan uno o varios planos. Por ello ha sido necesario usar distintos patrones, aunque se ha procurado que en todos los casos se encontraran aproximadamente a la misma distancia de la cámara (entre 40 y 45 cm), que aproximadamente tuvieran la misma cantidad de puntos (entre 126 y 130 puntos) y que ocuparan un área similar en la imagen.

Para realizar la calibración con aquellos métodos que necesitan que los puntos no se encuentren en el mismo plano (Lineal, Tsai, Tsai-optimizado, Faugeras, Heikkilä y Ahmed), se ha utilizado una distribución de puntos en forma de diedro con 8x8 puntos en cada cara y separados entre si 2 cm (Figura 2.7). Las coordenadas del centro del patrón respecto a la cámara son: $x = -20mm$, $y = 0mm$, $z = 420mm$; con una rotación de 11.5° en el eje X_c , -60° en el Y_c , y -8.5° en el eje Z_c .

Para el método de Batista, se necesita que todos los puntos estén en el mismo plano. Se ha utilizado un plano de 13x10 puntos separados entre si 1.75 cm (Figura 2.8). Las

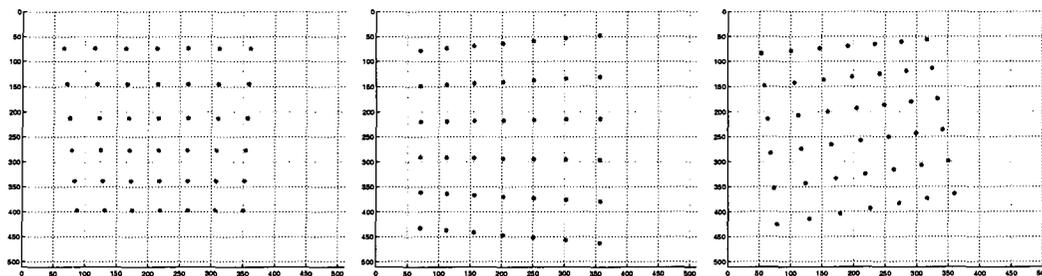


Figura 2.9: Proyecciones en la imagen de la distribución de puntos del plano usado para el método de Zhang.

coordenadas del centro del patrón respecto a la cámara son: $x = -27mm$, $y = 8mm$, $z = 453mm$; con una rotación de 11.5° en el eje X_c , -30° en el Y_c , y -8.5° en el eje Z_c .

Para el método de Zhang, se necesitan tres planos de puntos no paralelos entre sí. Se ha utilizado un plano de 7x6 puntos separados entre si 3 cm (Figura 2.9). El mismo patrón se ha proyectado en la imagen con tres orientaciones (R_x, R_y, R_z) distintas: 1) $20^\circ, 0^\circ, 0^\circ$; 2) $0^\circ, 20^\circ, 0^\circ$; 3) $-15^\circ, -13.5^\circ, -6.5^\circ$. La distancia del plano varió ligeramente en los tres casos (410 mm, 420 mm, 435 mm, respectivamente).

Para cada uno de los métodos se ha evaluado en cada experimento el comportamiento de cada parámetro (tanto intrínsecos como extrínsecos) incluido en el modelo que utiliza y se han comparado los resultados de cada método. Además, se han usado dos medidas de error para realizar la comparación de los distintos métodos implementados. Una obtiene la disparidad de imagen o error 2D (A.5), que es el error cometido al proyectar sobre la imagen un punto de coordenadas conocidas. Este error se mide como la distancia en pixels entre el punto real y la proyección usando los valores resultantes de la calibración. La otra medida calcula el error cometido en la reconstrucción de la posición del punto en el espacio 3D (A.6) calculando también la distancia, esta vez en milímetros, entre el punto real y la reconstrucción realizada usando los valores obtenidos de la calibración. En ambos casos se usa el promedio de las distancias obtenidas en todos los puntos del patrón como medida de error.

Hay que resaltar que cada método genera un conjunto de parámetros distinto como resultado de la calibración. Por tanto, hay gráficas en las que no aparecen los resultados de alguno de los métodos, ya que no calculan el parámetro correspondiente:

- El método de Tsai no aparece en las gráficas que muestran el error en la obtención del centro del eje óptico.
- En las gráficas del error en la estimación de la ortogonalidad del plano imagen no aparecen los métodos de Tsai, Tsai-optimizado y Heikkilä.

- En las gráficas que muestran el error en el cálculo del primer coeficiente de la distorsión radial no aparecen los métodos de Ahmed, Lineal y Faugeras.

Para poder comparar los resultados de todos los métodos se han tenido que realizar operaciones sobre alguno de los resultados. Así, en los métodos que resuelven estos parámetros explícitamente, se han combinado los resultados de cada uno de ellos para obtener un parámetro equivalente a los que lo resuelven implícitamente:

- El parámetro k_x se obtiene como producto de s_x , d_x y f en los métodos de Tsai, Tsai-optimizado, Batista y Heikkilä. El valor de este parámetro en la configuración utilizada de la cámara es de 673.6 píxeles.
- El parámetro k_y se obtiene como producto de d_y y f en los métodos de Tsai, Tsai-optimizado, Batista y Heikkilä. El valor de este parámetro en la configuración empleada es de 993.1 píxeles.
- El parámetro k_1 en el método de Zhang se divide por f^2 para que su magnitud sea comparable a la del resto. Hay que recordar que en el modelo usado por este método se corrige la distorsión antes de realizar las transformaciones de todos los parámetros intrínsecos, mientras que el resto lo hace cuando las coordenadas de los puntos de la cámara han sido proyectadas sobre el plano imagen, por tanto son multiplicadas por la distancia focal antes de corregir la distorsión (ec. 2.6).

2.3.1. Inyección de ruido en los datos de entrada

Los métodos que utilizan un patrón como referente necesitan para realizar el proceso de calibración conocer las coordenadas tridimensionales de los puntos incluidos en el patrón y su proyección 2D en el plano imagen. El problema que se presenta es que estos datos están contaminados con ruido por la influencia de diversos factores, como puede ser el proceso de cuantificación al digitalizar las imágenes o la mala estimación de las coordenadas 3D de los puntos.

En este apartado se analiza la sensibilidad de los distintos métodos a que los datos de entrada presenten estas pequeñas perturbaciones. A priori se supone que aquellos métodos que propongan un modelo con un mayor número de parámetros deben ser menos precisos. Aunque esto depende de la exactitud con que los métodos, que modelan el proceso con menos parámetros, fijen los parámetros no incluidos en dicho proceso. En este caso se hablará de proximidad de la configuración de la cámara utilizada a la "configuración ideal", que en el presente estudio, consiste en que el centro óptico coincida con el centro de la imagen, el factor de proporción sea uno, la ortogonalidad de la imagen

sea perfecta (90°) y la distorsión nula. En la siguiente tabla se muestra el número de parámetros que resuelve cada uno de los métodos aquí analizados.

	intrínsecos	extrínsecos
Lineal	5	6
Tsai	3	6
Tsai-LM	5	6
Faugeras	5	6
Zhang	7	18 (6 por imagen)
Batista	7	6
Heikkilä	8	6
Ahmed	5	6

Tabla 2.1: Número de parámetros que incluye el modelo utilizado por cada método.

Por otro lado, se estudia cuál de los datos de entrada influye más sobre los resultados finales. Hay que decir que en gran parte de los trabajos sobre calibración sólo se incluye un estudio sobre la influencia del ruido en las coordenadas de 2D de la imagen, pero se deja de analizar la incidencia del ruido proveniente de la otra entrada de datos (las coordenadas 3D de los puntos del patrón).

Por último, se intenta observar qué parámetros son más sensibles a las perturbaciones de los datos de entrada y en qué medida afecta esto al resultado global de la calibración.

Variación del nivel de ruido en las coordenadas 2D de la imagen del patrón

El ruido de digitalización aparece como consecuencia de la resolución espacial del sensor CCD, esto es, del número de pixels por unidad de longitud tanto en su dimensión vertical como horizontal.

Para simular el nivel de ruido en la digitalización de la imagen se ha inyectado a las coordenadas en imagen del objeto un ruido gaussiano de media 0 y desviación típica σ . El valor de σ ha variado entre 0 (condiciones libres de ruidos) y 1.5 pixels. El procedimiento de calibración se ha repetido 100 veces para cada nivel de ruido.

Los resultados de este experimento se muestran en las gráficas de las figuras 2.10 - 2.12. Todos los métodos tienen un incremento lineal del error en todos los parámetros a medida que aumenta el nivel de ruido en las coordenadas 2D. La única excepción la presenta el método de Tsai-LM que en muchos casos presenta un incremento exponencial del error.

Destacan como peores métodos los de Heikkilä y Tsai-LM en casi todas las gráficas. El primero de ellos es el peor en todos los parámetros, siendo del orden de 3 veces peor que la mayoría de los otros métodos; excepto en los que interviene la distancia focal (k_x y k_y) y la distancia del patrón (T_z), donde es el método optimizado de Tsai es el que resulta más afectado por el ruido. En ambos casos se cumple la hipótesis planteada inicialmente en la que se indicaba que el método más sensible al ruido sería el que incluyese un mayor

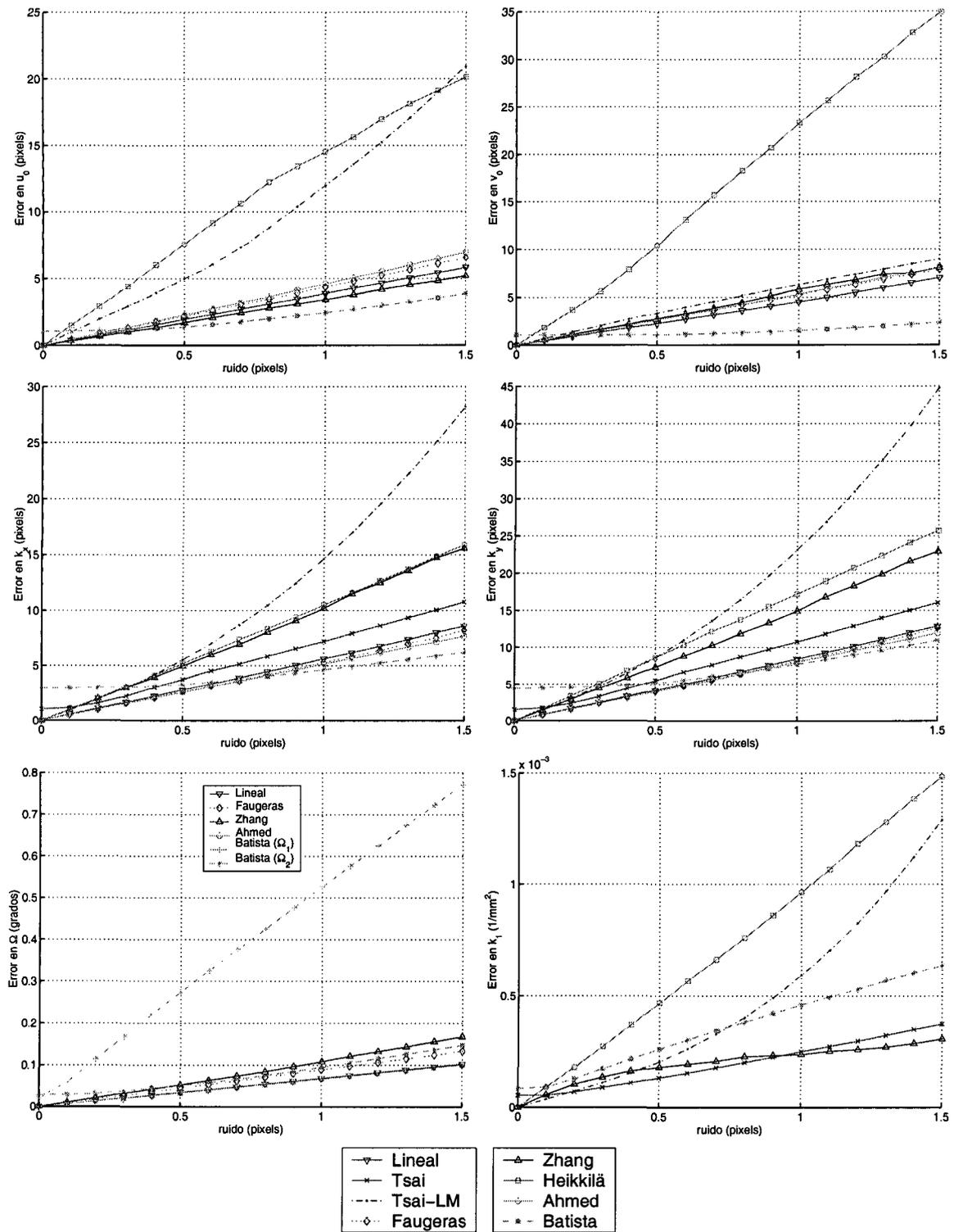


Figura 2.10: Error en el cálculo de los parámetros intrínsecos ante la variación del nivel de ruido en la digitalización de la imagen.

número de parámetros. El método de Heikkilä parametriza 8 elementos, y es el método que mayor número de parámetros intrínsecos modela, seguido del método de Tsai-LM con siete. La excepción a esta tónica son los métodos de Batista (8 parámetros) y el método de Zhang (7 parámetros). En el caso del método de Batista puede influir el hecho de que muchos parámetros permanecen muy próximos a la aproximaciones iniciales introducidas al método. En el caso del método de Zhang influir el hecho de utilizar varias vistas, lo que permite reducir la interacción entre los parámetros extrínsecos y los intrínsecos.

En el lado contrario, y aparentemente, el método de Batista presenta un comportamiento muy bueno ante la presencia de ruido en las coordenadas de imagen. Este análisis se realiza en base a que estima bastante bien la mayoría de los parámetros. En este método el parámetro que ofrece mayor sensibilidad es la asimetría, probablemente debido al hecho de que es el único método que la modela con dos ángulos. También se ven afectados k_x y la orientación sobre los ejes X e Y. En el primer caso se debe a que el factor de proporción se calcula en una primera fase y para su estimación se usan valores predefinidos de algunos parámetros (centro óptico, ortogonalidad, factor de escalado vertical) y, aunque en la segunda fase se ajusta su valor, se ve que el error acumulado no permite converger a un resultado más próximo al real. En el caso de la orientación llama la atención que tenga muy buen comportamiento en la estimación del giro alrededor del eje Z de la cámara, todo lo contrario que en los otros dos giros; esto puede ser debido a que en la ecuación que obtiene este parámetro no interviene la distancia focal, en cambio en los otros dos ángulos sí interviene; además es el primer ángulo que se calcula, por lo que los otros dependen del valor obtenido para este parámetro y, por tanto, pequeños errores en su estimación se propagaran a los otros dos ángulos.

El método de Tsai también tiene muy buen comportamiento, excepto en la estimación de la distancia focal y T_z . En este experimento también se puede observar las diferencias entre las dos versiones del método de Tsai. La versión clásica, al ser muy dependiente de los factores iniciales (y mantener constante alguno de ellos), se ve muy poco influenciado por el ruido, mientras que la versión optimizada al no tener fijados esos parámetros (principalmente, el centro del eje óptico), los valores de todos ellos se ven muy afectados por el ruido. En estos dos métodos (Batista y Tsai), la estabilidad que tienen viene dada por la proximidad de la configuración de la cámara respecto a la aproximación inicial introducida en cada método, es decir, el centro del eje óptico sólo dista 1 pixel respecto al centro de la imagen, el factor de escala es próximo a 1 ($s_x = 0,994$), el plano imagen es perfectamente ortogonal y la distorsión es nula. Pero pequeñas diferencias respecto a esta configuración ideal producen, que en condiciones libres de ruido en las coordenadas 2D, estos métodos tengan un pequeño error. En posteriores experimentos se demostrará

que este buen comportamiento de los métodos de Batista y Tsai es ficticio, ya que se observará la gran dependencia que tienen estos dos métodos respecto a la fiabilidad de las aproximaciones iniciales.

Los métodos de Ahmed, Faugeras y Lineal tienen unos resultados muy parecidos en casi todos los parámetros. Si tenemos en cuenta que se ha utilizado el método Lineal como aproximación inicial de los otros dos, y se observa que la mejora que aportan es mínima en contraposición al coste computacional que supone el proceso iterativo por un lado y el entrenamiento de la red neuronal por otro, se puede concluir que es más rentable usar el método Lineal que los otros dos en condiciones similares a las de estos experimentos.

Por último, se puede decir que el método Zhang es el que mejor reconstruye la posición de los puntos, destacando ligeramente sobre el resto en la reconstrucción tridimensional. También cabe indicar que es el menos sensible en la estimación de la distorsión y de la orientación sobre el eje Z. En cambio, en los parámetros dependientes de la distancia focal y en la distancia del patrón el nivel de error es próximo al método de Heikkilä. En el resto de los parámetros mejora ligeramente los resultados de los tres métodos comentados en el párrafo anterior.

Si se analiza la forma de cada una de las gráficas se puede observar como hay parámetros que están directamente relacionados entre sí y entre los que existe una compensación de los errores. Como se puede ver en la figura 2.12, el error promedio es muy parecido en todos los métodos. En concreto, se observa la íntima relación que existe, por un lado, entre los errores cometidos en el cálculo del centro del eje óptico y de la traslación en X e Y del patrón; y entre los parámetros que incluyen la distancia focal y la distancia a la que se encuentra el objeto (T_z) por otro. Esto es explicable por el modelo de cámara utilizado (pin-hole).

	u_0	v_0	k_x	k_y	Ω	k_1
Real	255	255	673.58	993.1	90	0
Lineal	254.8(5.1)	256.1(5.7)	674.4(7)	994.41(10.5)	90(0.1)	-
Tsai	-	-	672.9(9)	992.7(13.5)	-	$-1,5 \cdot 10^{-5}(3,3 \cdot 10^{-4})$
Tsai-LM	245.3(8.8)	253.4(5.5)	662.7(11.1)	975.8(17.2)	-	$-5,1 \cdot 10^{-4}(5,1 \cdot 10^{-4})$
Faugeras	254.9(5.6)	255.2(6.7)	674.3(6.7)	994.3(10.4)	90(0.06)	-
Zhang	253.3(4.1)	256.9(7.3)	672.3(12.3)	991.2(18.1)	90(0.1)	$-6,8 \cdot 10^{-4}(0,04)$
Batista	256.3(3.7)	255.9(2.0)	680.4(5.5)	997.7(10.5)	90.1(0.7)	$-8,5 \cdot 10^{-5}(5,5 \cdot 10^{-4})$
Heikkilä	253.5(18.4)	253.5(29.9)	672.7(13.2)	992.0(22.0)	-	$3,2 \cdot 10^{-5}(0,0012)$
Ahmed	255.9(5.8)	256.7(6.6)	674.1(6.5)	994.3(10)	90(0.08)	-

Tabla 2.2: Comparación de los resultados de los parámetros intrínsecos de cada método respecto a los datos reales cuando el nivel de ruido en las coordenadas 2D de la imagen es de un pixel. Valor medio obtenido y, entre paréntesis, la desviación típica.

Finalmente, en la tabla 2.2 se pueden observar los valores promedio obtenido por cada método y su desviación típica, que aparece entre paréntesis en la tabla, para cada

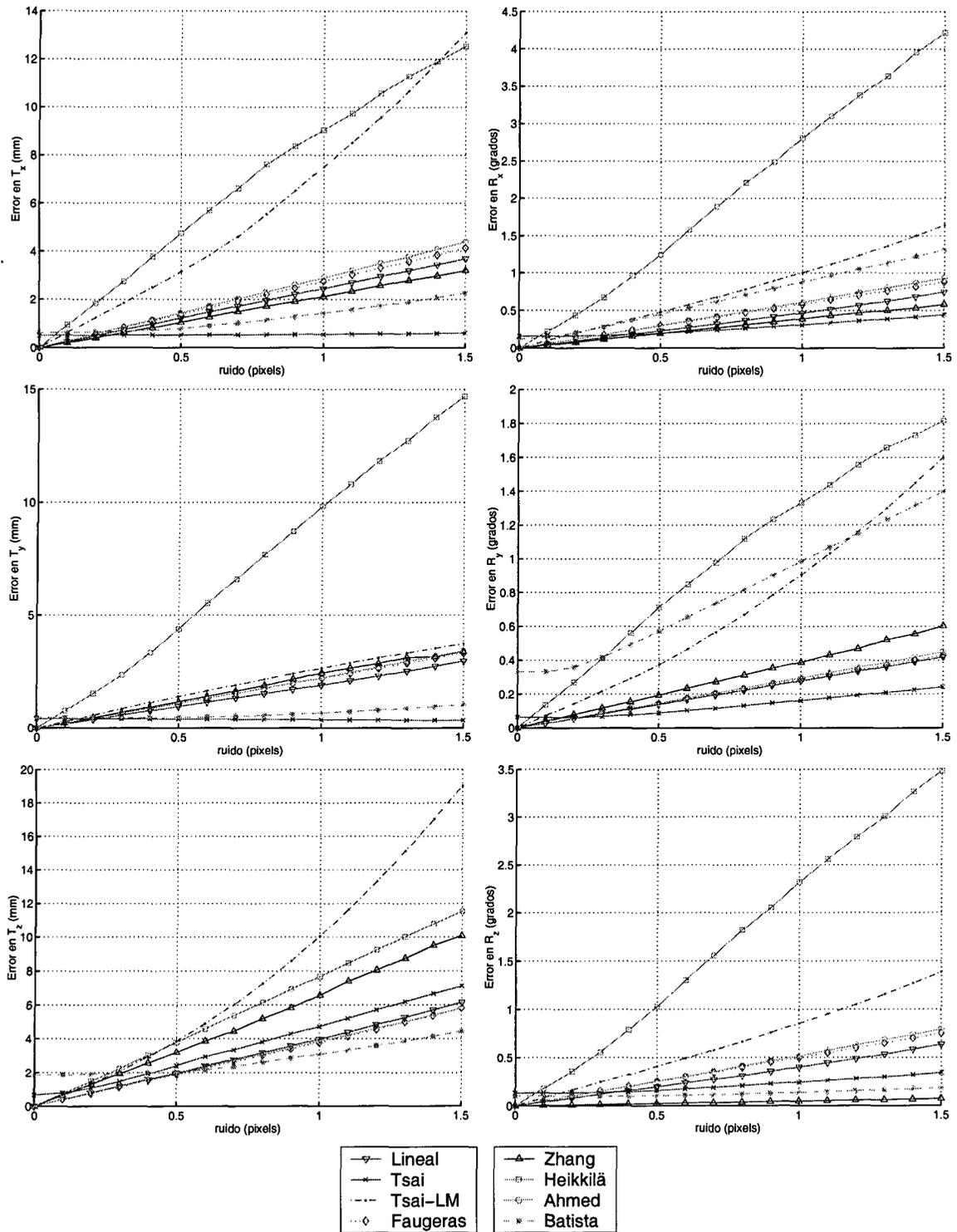


Figura 2.11: Error en el cálculo de los parámetros extrínsecos ante la variación del nivel de ruido en la digitalización de la imagen.

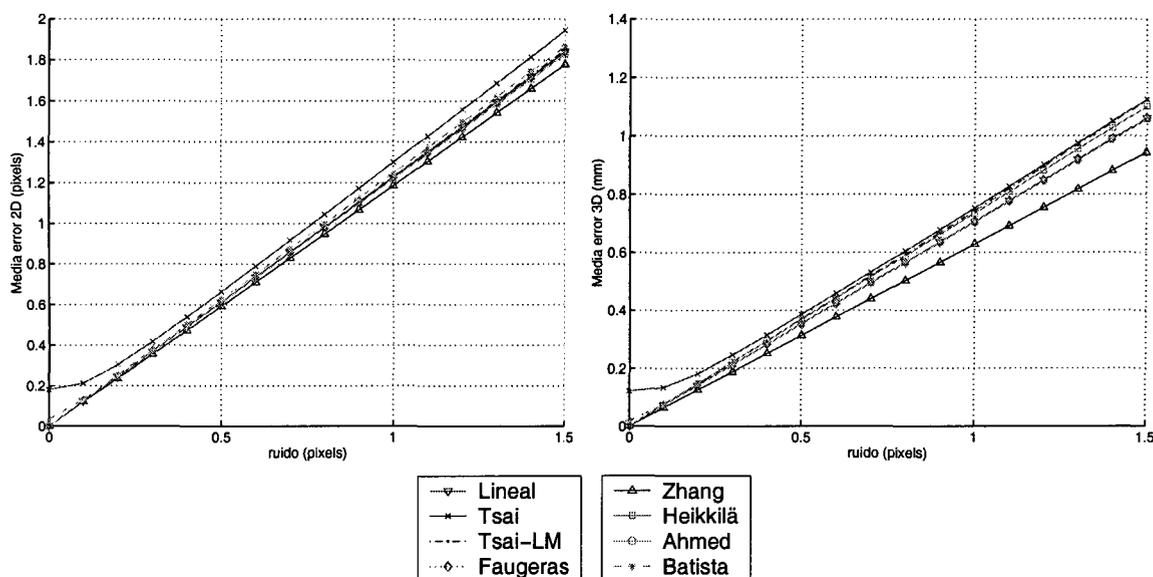


Figura 2.12: Error promedio ante la variación del nivel de ruido en la digitalización de la imagen.

uno de los parámetros intrínsecos cuando existe un nivel de ruido de 1 píxel en las coordenadas 2D de la imagen.

Variación del nivel de ruido en las coordenadas 3D de los puntos en el patrón

En la estimación de las coordenadas de los puntos en el patrón también pueden existir errores. Para simular esta situación se ha inyectado a las coordenadas 3D de cada punto del patrón un ruido gaussiano de media 0 y desviación típica σ , variando el valor de ésta entre 0 y 2 milímetros. El procedimiento de calibración se ha repetido 50 veces para cada nivel de ruido.

Los resultados de este experimento se muestran en las gráficas de las figuras 2.13 - 2.15. En general, se observa que la forma de las curvas en las gráficas es bastante similar a las obtenidas en el experimento anterior. No así la magnitud del error.

En general, el método que mejor comportamiento tiene ante la presencia de ruido en las puntos del patrón es el de Batista, teniendo como en el caso anterior un error inicial que permanece prácticamente constante ante la variación del ruido en casi todos los parámetros. Este método es el único que disminuye la magnitud del error respecto al experimento anterior. En especial, cabe destacar la mejora que se produce en la obtención de los ángulos que indican la orientación del patrón (R_x y R_y). Esta mejora se debe al hecho de que este método obtiene los ángulos de la orientación directamente de cuatro puntos que forman un rectángulo en la imagen, si estos puntos están afectados por ruido en las coordenadas de imagen el resultado será un alto nivel de error en los

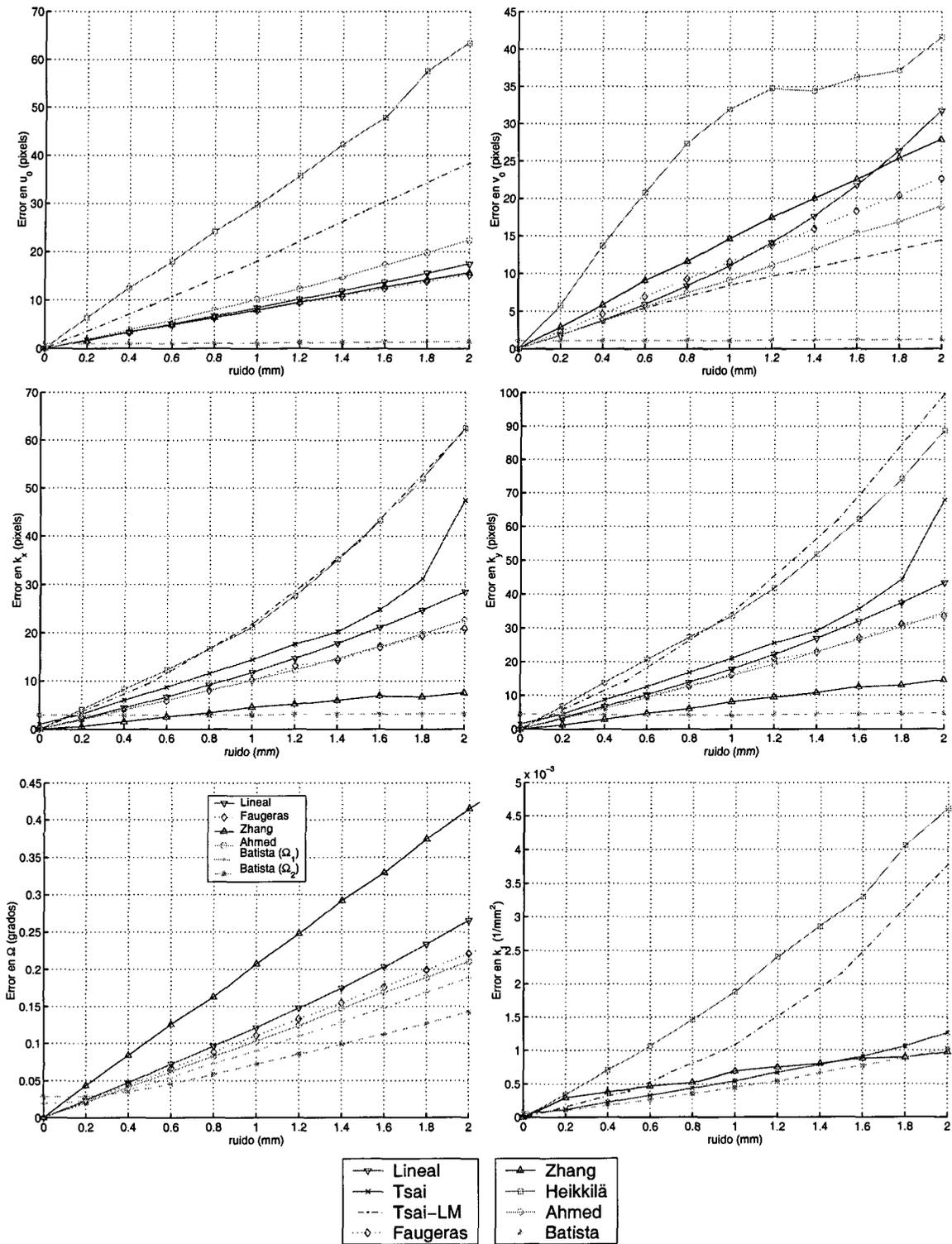


Figura 2.13: Error en el cálculo de los parámetros intrínsecos ante la variación del nivel de ruido en la posición de los puntos en el patrón.

parámetros. Ya que los ángulos de la orientación del patrón son los primeros parámetros que se estiman entonces el resto de los parámetros absorben el residuo de este error.

En casi todos los casos el método de Heikkilä sigue siendo el menos estable, salvo en los parámetros que dependen de la distancia focal y la distancia del patrón en los que el método de Tsai-LM es el que se ve más afectado por este ruido, sobretodo cuando el nivel de ruido es alto. En estos parámetros también se observa un comportamiento exponencial en el error en el método de Tsai cuando el nivel de ruido es próximo a los 2 mm.

A diferencia del experimento anterior, a medida que aumenta el nivel de ruido se observa como los métodos de Ahmed y Faugeras mejoran en casi todos los casos los resultados obtenidos por el método Lineal.

El método de Zhang mejora sensiblemente el rendimiento respecto al caso anterior en la estimación de los factores de escalado y la distancia a la que se encuentra el patrón, siendo junto con el método de Batista el que menor error produce en este parámetro. Sin embargo, en la estimación de la ortogonalidad del plano imagen el método de Zhang es el que peor resultado ofrece.

Al igual que antes, tras la observación de las gráficas de error es evidente la íntima relación que existe, por un lado, entre los errores cometidos en el cálculo del centro del eje óptico y de la traslación en X e Y del patrón; y por otro lado, entre los parámetros que incluyen la distancia focal y la distancia a la que se encuentra el objeto (T_z).

Por último, si se compara la magnitud del error obtenido en los dos experimentos se observa que, en general, un milímetro de ruido en las coordenadas 3D produce casi el doble de error que un pixel de ruido en las coordenadas 2D de la imagen. Este resultado es lógico, ya que a 42 cm de distancia que se encuentra el patrón, y con la configuración de cámara utilizada, un milímetro es equivalente a 1.6 pixels en el eje horizontal de la imagen y a 2.3 pixels en el vertical, situándose ambos por encima del máximo nivel (1.5 pixels) empleado en el experimento anterior.

Variación del nivel de ruido en las coordenadas 3D de los puntos en el patrón usando imágenes con ruido gaussiano en las coordenadas 2D de la imagen

Con el fin de observar si la unión de ambas fuentes de ruido produce un incremento del error, e intentado presentar un experimento más próximo a la realidad, se ha repetido el experimento anterior incluyendo además un ruido gaussiano de media 0 y desviación típica 0.5 en las coordenadas 2D para imitar la digitalización de una imagen. Para simular esta situación se ha inyectado a las coordenadas 3D de cada punto del patrón un ruido gaussiano de media 0 y desviación típica σ , variando el valor de ésta entre 0 y 2

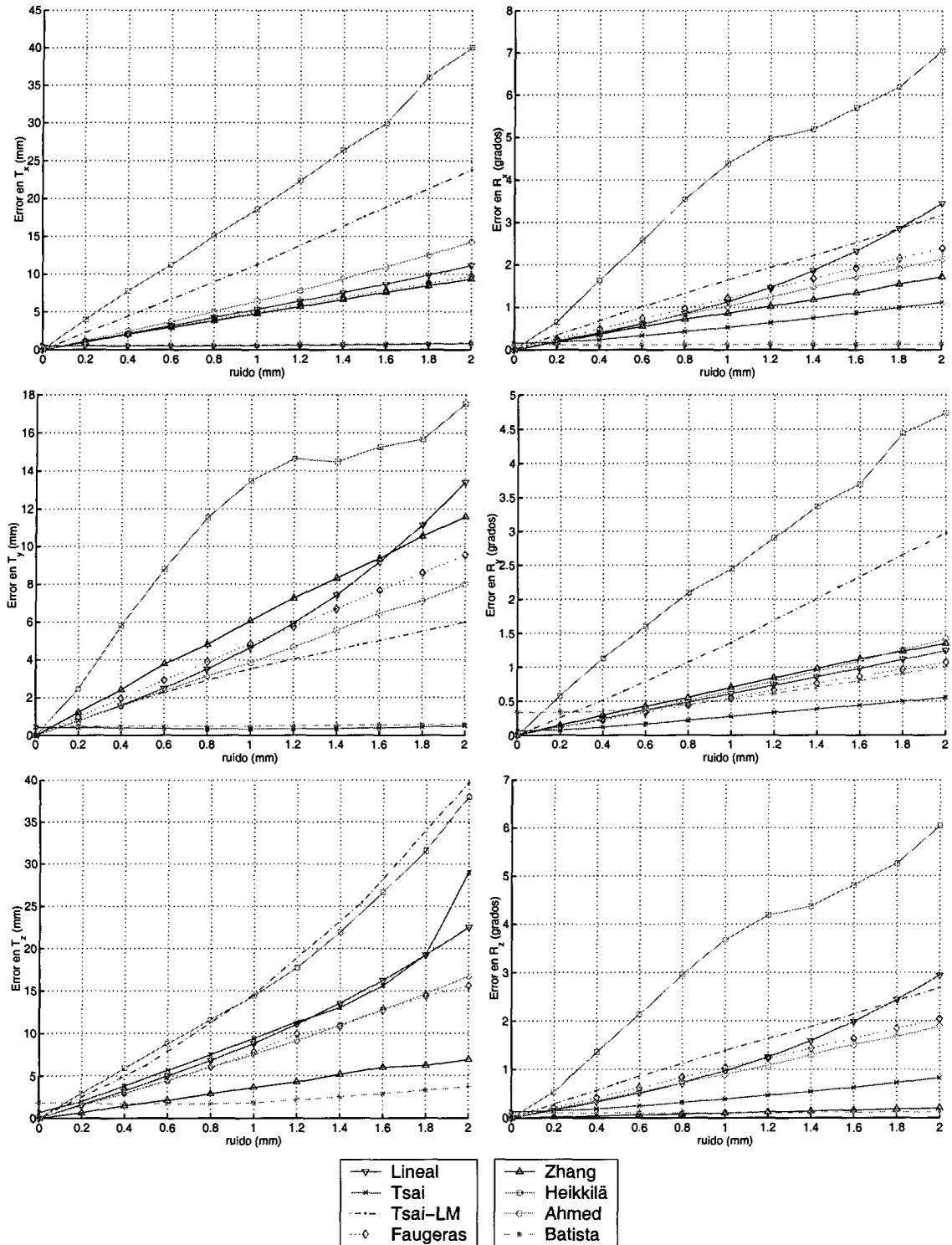


Figura 2.14: Error en el cálculo de los parámetros extrínsecos ante la variación del nivel de ruido en la posición de los puntos en el patrón.

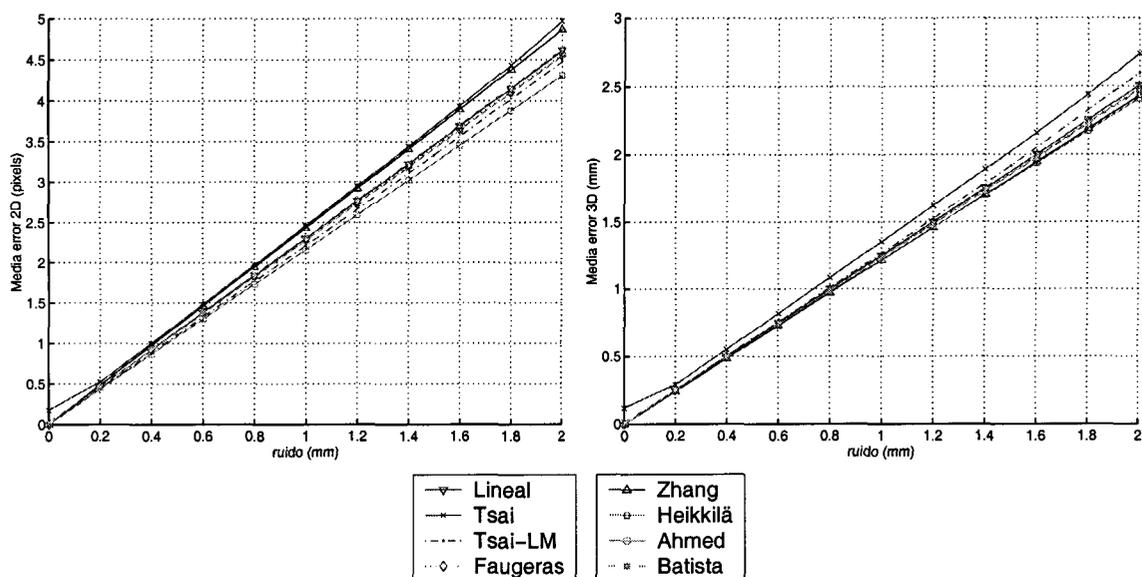


Figura 2.15: Error promedio ante la variación del nivel de ruido en la posición de los puntos en el patrón.

milímetros. El procedimiento de calibración se ha repetido 50 veces para cada nivel de ruido.

Los resultados obtenidos (Figuras 2.16 a 2.18) muestran que las diferencias existentes respecto al experimento anterior son mínimas. En general, se produce un incremento del error que ronda el 10 %, aunque, excepcionalmente, en el método de Batista se produce un incremento que ronda el 50 %. En este caso particular el motivo de este gran aumento en el error se debe a que el método de Batista, al contrario, que el resto de los métodos es mucho más sensible al ruido en las coordenadas 2D que al ruido en las coordenadas 3D, aún así, sigue siendo el método que mejor estima la mayoría de los parámetros en las condiciones que se han realizado estos experimentos.

Se puede concluir que el error que se produce esta motivado por el ruido que tenga una mayor incidencia. Debido a esto, cuando el nivel de ruido en las coordenadas 3D es muy bajo el error obtenido es producto del ruido en las coordenadas 2D y viceversa. También hay que indicar que la unión de los ruidos no produce una suma en la magnitud del error, aunque sí un ligero incremento.

2.3.2. Variación en la configuración de la cámara

En este apartado se estudia la influencia de los distintos parámetros que configuran la cámara sobre el resultado de la calibración. A priori se producirá una alteración del resultado en aquellos casos en los que un parámetro no es modelado o se supone conocido, y en aquellos casos en los que, aunque sea modelado, el método resuelve el problema

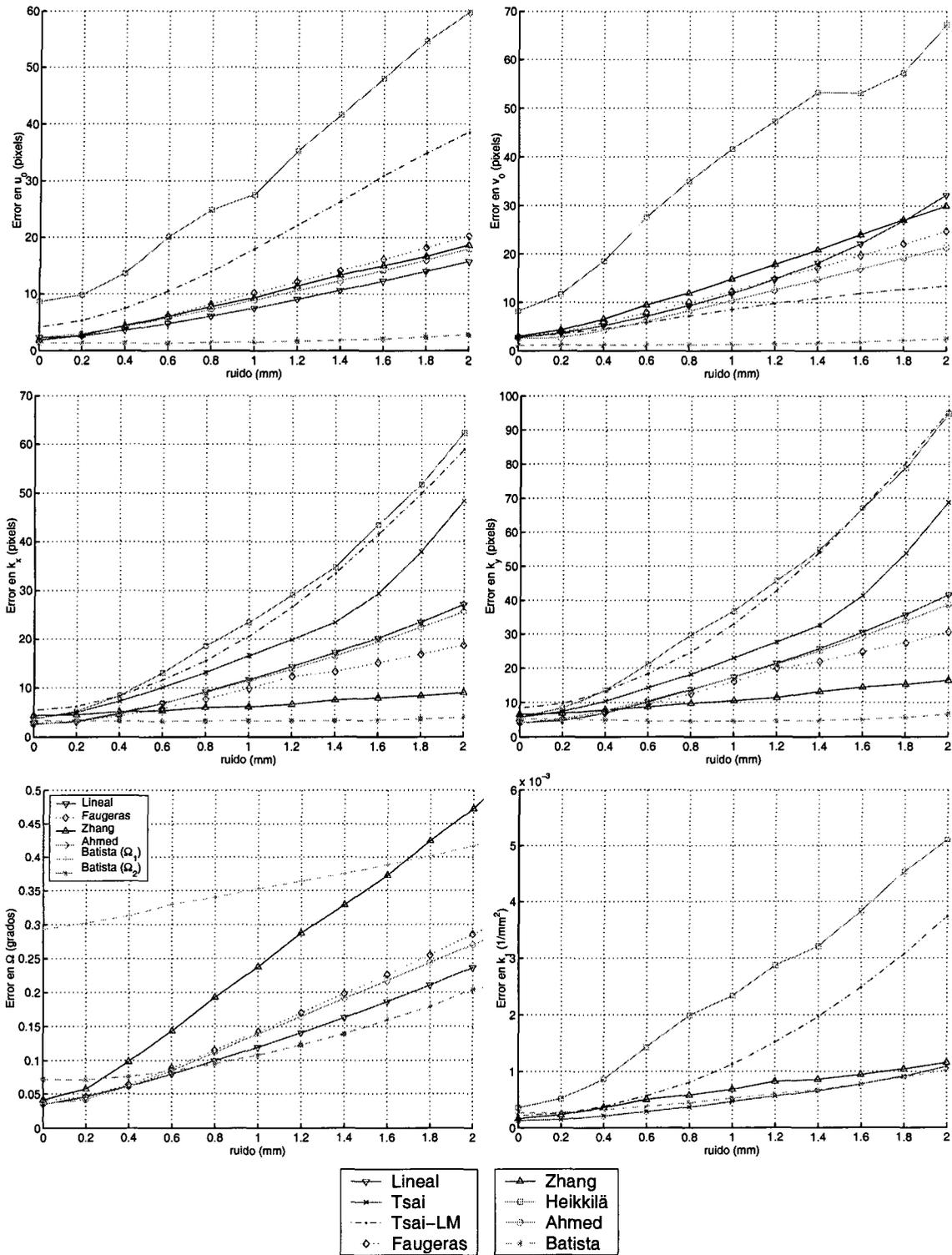


Figura 2.16: Error en el cálculo de los parámetros intrínsecos ante la variación del nivel de ruido en la posición de los puntos en el patrón usando imágenes con ruido de 0.5 pixels en las coordenadas 2D.

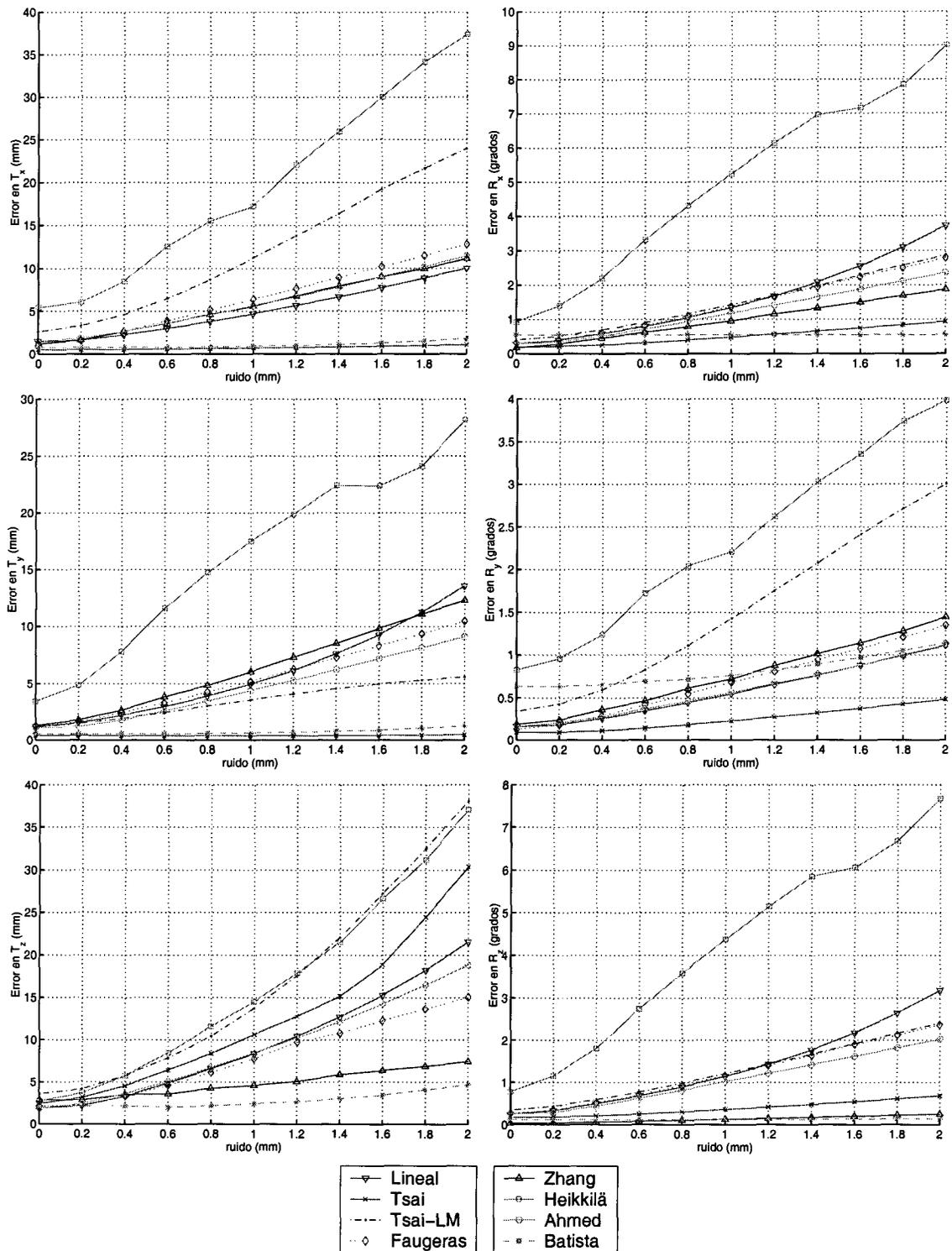


Figura 2.17: Error en el cálculo de los parámetros extrínsecos ante la variación del nivel de ruido en la posición de los puntos en el patrón usando imágenes con ruido de 0.5 pixels en las coordenadas 2D.

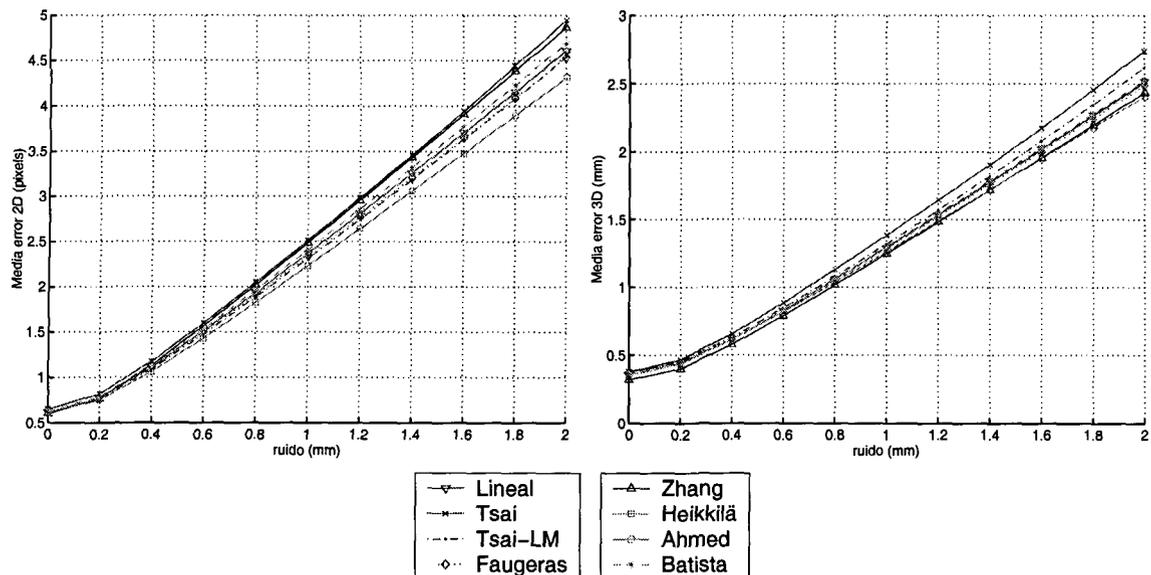


Figura 2.18: Error promedio ante la variación del nivel de ruido en la posición de los puntos en el patrón usando imágenes con ruido de 0.5 pixels en las coordenadas 2D.

de la calibración en varias fases, necesitando para las primeras fases una aproximación inicial. La siguiente tabla muestra qué factores se necesita conocer previamente en cada uno de los métodos o aquellos que el modelo no parametriza.

Método	No modela	Aproximación inicial
Lineal	distorsión	
Tsai	$\Omega, d_x, d_y, u_0, v_0$	k_1
Tsai-LM	Ω, d_x, d_y	k_1, u_0, v_0
Faugeras	distorsión	
Zhang		distorsión
Batista		intrínsecos
Heikkilä	Ω, d_x, d_y	$s_x, \text{distorsión}$
Ahmed	distorsión	

Tabla 2.3: Parámetros que no modela o para los que necesita una aproximación inicial cada uno de los métodos.

Variación del primer coeficiente de distorsión radial (k_1)

Un factor que se debe estudiar es la distorsión, ya que muchos métodos no lo incluyen dentro de su procedimiento (Lineal, Faugeras, Ahmed). En este experimento se ha variado el primer coeficiente de distorsión radial, sin que los resultados sean perturbados por ningún tipo de ruido y en las condiciones descritas en la introducción de esta sección. El valor de k_1 se ha variado entre -0.02 y -10^{-6} . Nótese que estos valores son negativos y por tanto en las gráficas la magnitud de este parámetro decrece. En la figura 2.19 se puede observar el efecto de la distorsión sobre uno de los patrones. Por último, las gráficas que muestran los resultados se presentan en escala logarítmica para una mejor

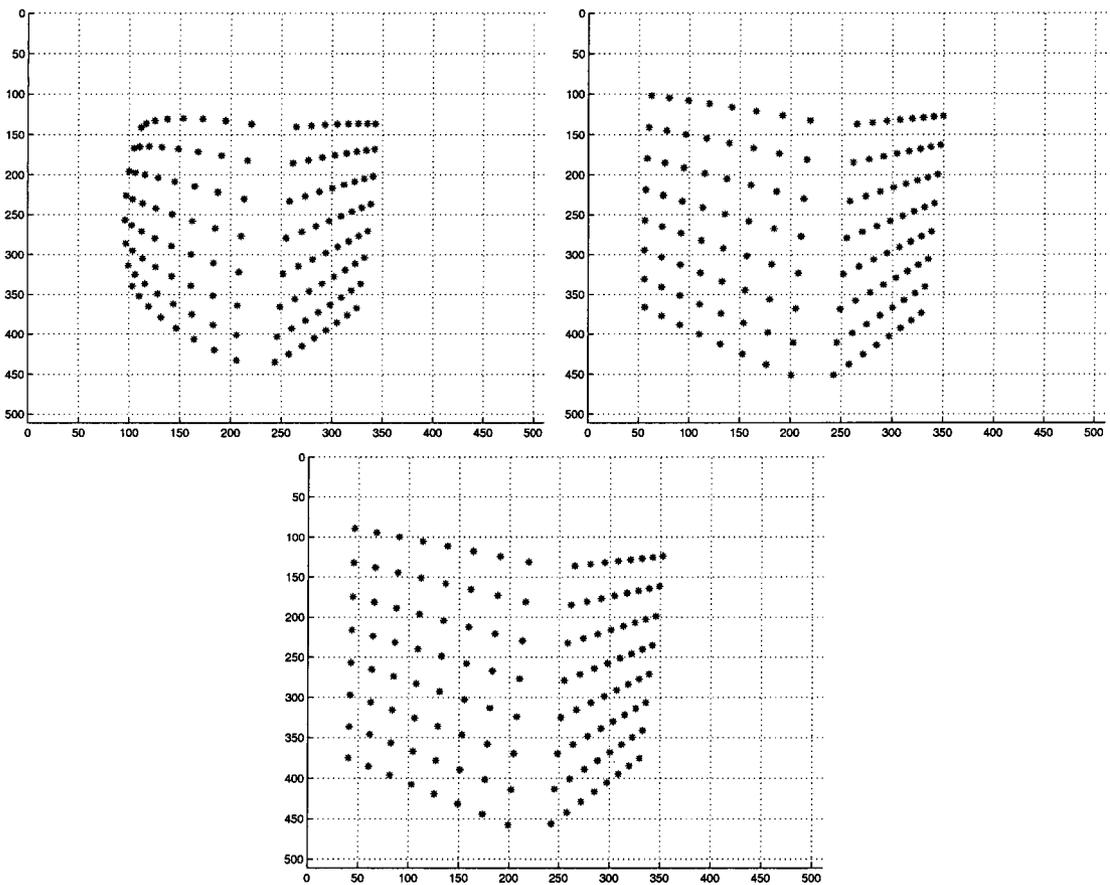


Figura 2.19: Proyección del diedro en la imagen. a) $k_1 = -0,02$. b) $k_1 = -0,005$. c) $k_1 = -10^{-6}$.

lectura de los mismos, ya que el rango de variación del parámetro de distorsión es muy amplio.

Como se puede observar en las figuras 2.20 y 2.21, el método de Heikkilä es el método que mejor modela la distorsión ya que tiene un error prácticamente nulo. Esto es lógico ya que es el método, dentro de los aquí descritos, que tiene una orientación más clara a trabajar con cámaras con focal pequeña donde la distorsión es elevada, y por ello modela con cuatro parámetros este factor. Además, hay que tener en cuenta que las imágenes se encuentran libres de ruido, que es el punto débil de este método.

Destaca también el comportamiento de los métodos de Tsai y Batista, los cuales, aunque tienen un buen comportamiento con valores altos de distorsión, luego estabilizan el error muy por encima de los demás métodos, incluso, cuando la distorsión es prácticamente nula. Esto es debido a que ambos métodos dependen mucho de las aproximaciones iniciales que se les introduzca. Por ejemplo, el método de Tsai presupone que el centro del eje óptico es el centro de la imagen, y como en la configuración de las cámaras usadas éste se encuentra desplazado en un pixel, este error es asumido por el

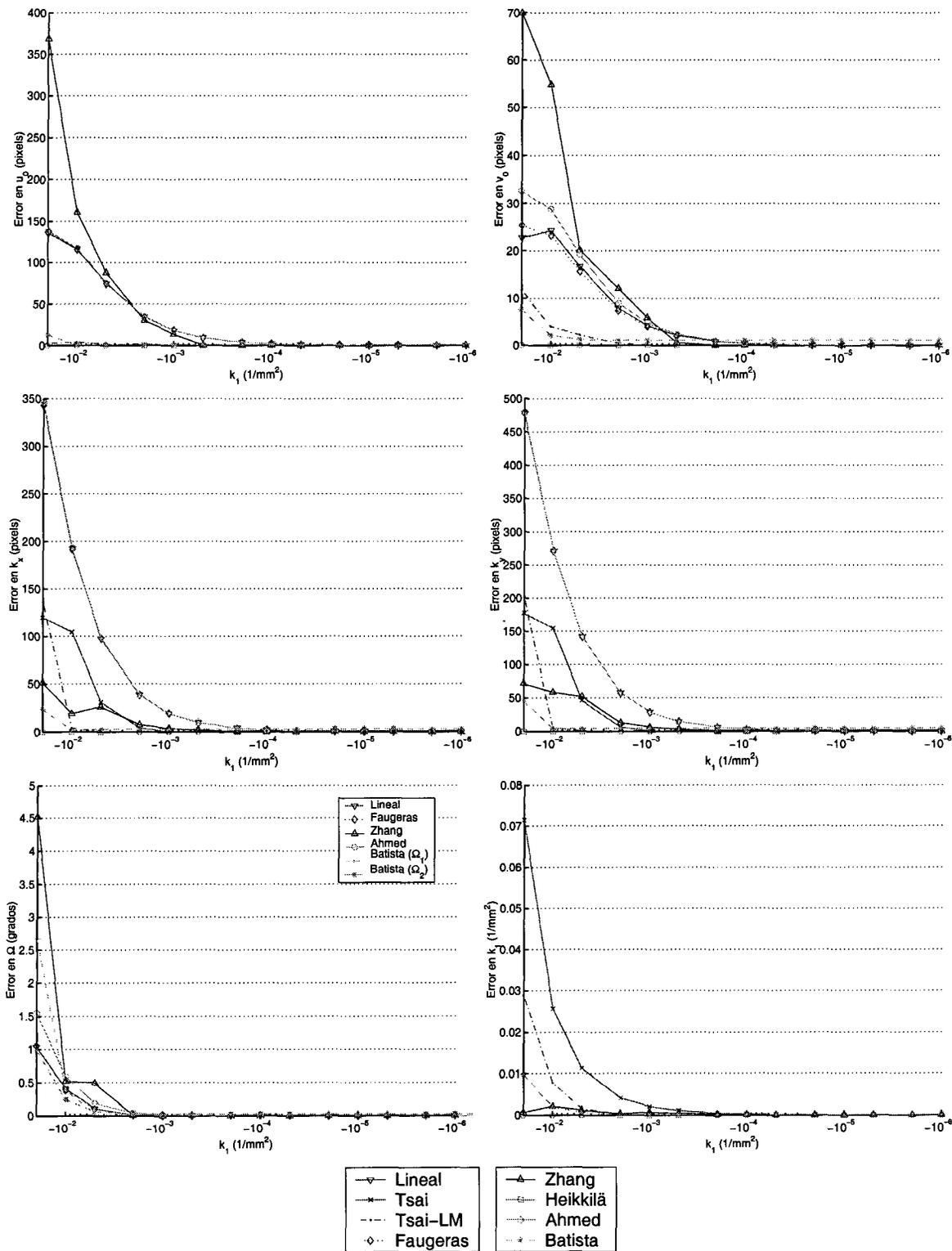


Figura 2.20: Error en el cálculo de los parámetros intrínsecos ante la variación de la distorsión radial (k_1).

resto de los parámetros. Una cosa parecida ocurre en el de Batista, ya que en su primer bucle iterativo se realiza la misma suposición, teniendo la misma consecuencia. Y aunque existe un segundo bucle iterativo que sí actualiza este parámetro, esto no es suficiente para anular el error en todos los parámetros. También realizan suposiciones sobre otros parámetros (factor de escala, coeficiente de distorsión, focal, etc.) que pueden afectar los resultados finales.

Se observa que el resultado de estos dos métodos son semejantes a los resultados de los experimentos de inyección de ruido. En el caso de Tsai los parámetros más sensibles son los que dependen de la distancia focal y la distancia del patrón. En el método de Batista ocurre algo parecido, aunque la magnitud del error es mucho menor en esos casos. Sin embargo, en este método el parámetro más sensible es R_y , posiblemente motivado por su dependencia a la distancia focal en las ecuaciones utilizadas para su obtención.

El método de Zhang funciona bastante mal cuando el grado de distorsión es muy alto (del orden de -10^{-2}), en especial en parámetros como el centro óptico (y por extensión T_x y T_y) y la rotación sobre el eje Y. El error tiende a converger a partir de $-5 \cdot 10^{-4}$. Llama la atención el mal comportamiento del método de Zhang porque modela la distorsión con dos coeficientes, pero a la hora de realizar la calibración trata este parámetro como un residuo del error cometido en la aproximación al resto, y aunque hay proceso de optimización final, el resultado parece quedar estancado en un mínimo local.

El método de Tsai-optimizado tiene un comportamiento parecido al de método Tsai, salvo en el caso de la reconstrucción de los puntos, tanto en 2D como en 3D, donde es el método que presenta mayor nivel de error. Hay que indicar que las dos variantes del método de Tsai son los que peor estiman el coeficiente de distorsión cuando el valor de éste es elevado. En este caso la magnitud del error supera la del propio valor del parámetro.

El resto de los métodos tienen decrecimiento continuo, necesitando que k_1 tenga una magnitud muy pequeña (aproximadamente $-5 \cdot 10^{-5}$) para tener un error aceptable, teniendo todos ellos un comportamiento idéntico entre sí. Esto es debido a que no modelan la distorsión y, por tanto, el error será proporcional al grado de distorsión incluido.

Por último, si se analizan las gráficas del error promedio (figura 2.22) se puede observar que el método de Heikkilä presenta un error prácticamente nulo. Después de este método el mejor es el método de Batista que en el peor de los casos observados no llega a superar los dos pixels de error en la reconstrucción de las coordenadas 2D y los 2 mm en las coordenadas 3D de los puntos. En el lado contrario está el método de Tsai-LM que supera los 25 pixels en las coordenadas 2D y los 15 mm en las coordenadas 3D en la misma situación. A este método le sigue el de Zhang, aunque el error no llega

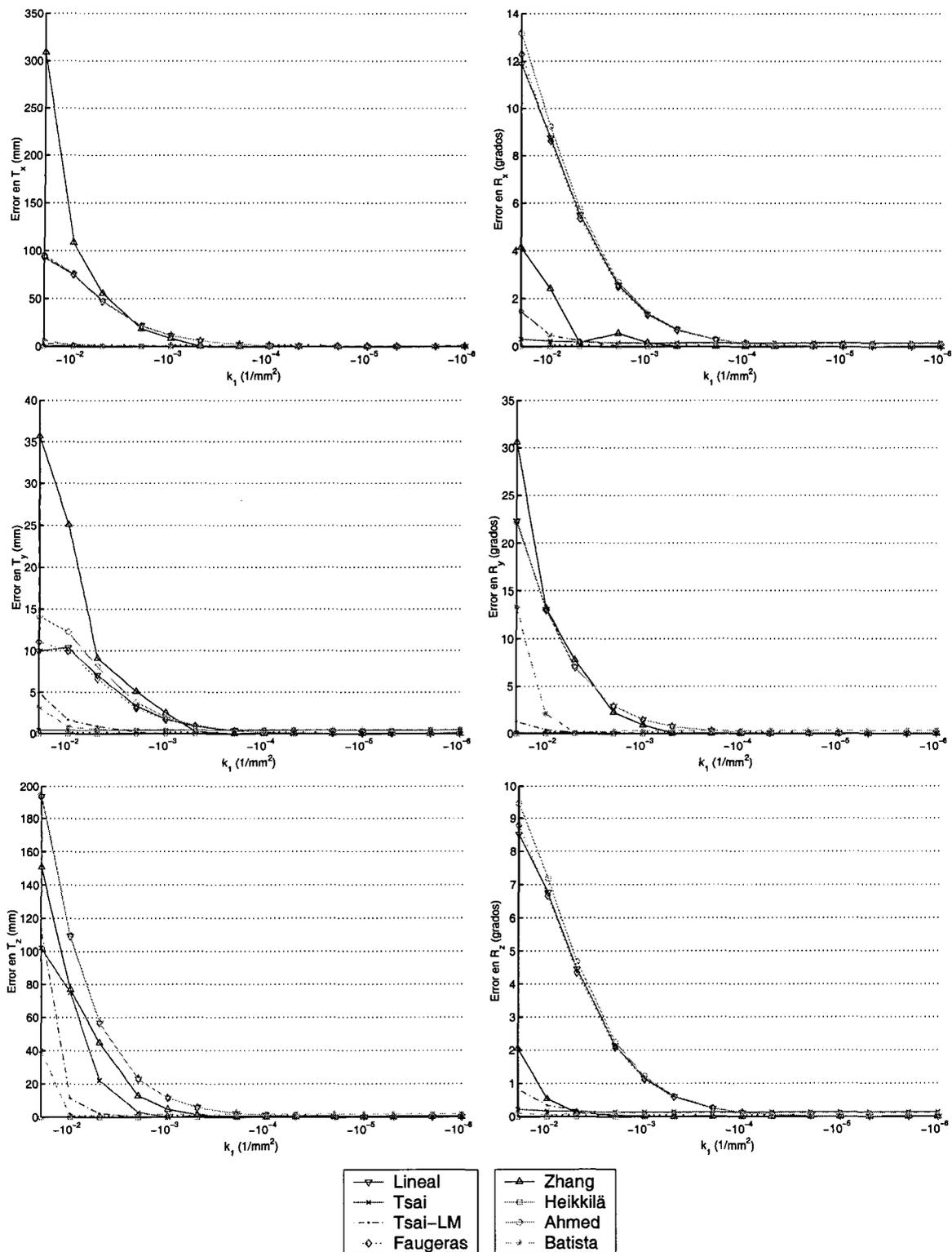


Figura 2.21: Error en el cálculo de los parámetros extrínsecos ante la variación de la distorsión radial (k_1).

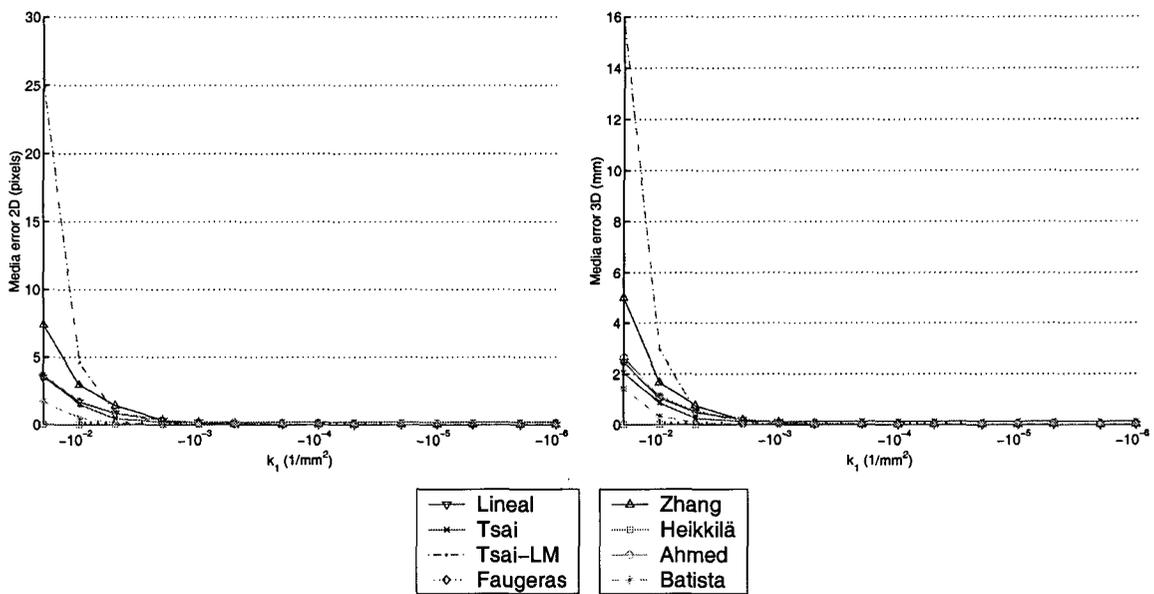


Figura 2.22: Error promedio ante la variación de la distorsión radial (k_1).

a superar los 8 pixels y los 5 mm respectivamente.

Variación del centro del eje óptico

Este experimento se ha planteado debido a que hay métodos que asumen que el centro del eje óptico se encuentra en el centro de la imagen (Tsai) o utilizan este dato como aproximación inicial (Tsai-LM y Batista). Por tanto, el objetivo es comprobar cual es la magnitud del error que resulta de separar estos dos centros. El experimento plantea mover el centro del eje óptico hasta 40 pixels respecto al centro imagen (hay que recordar que el tamaño de la imagen es de 512x512, por tanto supone en torno a un 15 % de la misma). El rango de variación va desde 216 a 296 pixels. Las imágenes se sintetizan en ausencia de cualquier tipo de ruido. Sólo se ha modelado la variación en el eje horizontal del plano imagen (u_0) ya que los resultados se pueden extender al eje vertical.

Como resultado del experimento se observa que los únicos métodos afectados por este desplazamiento son los de Batista y Tsai y para aportar mayor claridad a las gráficas (figuras 2.23 - 2.25) sólo se muestran las curvas de estos métodos y el de Tsai-LM. Como se puede observar, el método optimizado de Tsai, aunque usa como aproximación inicial al centro óptico el centro de la imagen, converge a la solución exacta en ausencia de otros ruidos. Debido a esto no se aprecia en las gráficas, ya que se confunde con el eje de abscisas.

Cuando se desplaza u_0 , en el método de Batista, el error cometido en el cálculo de

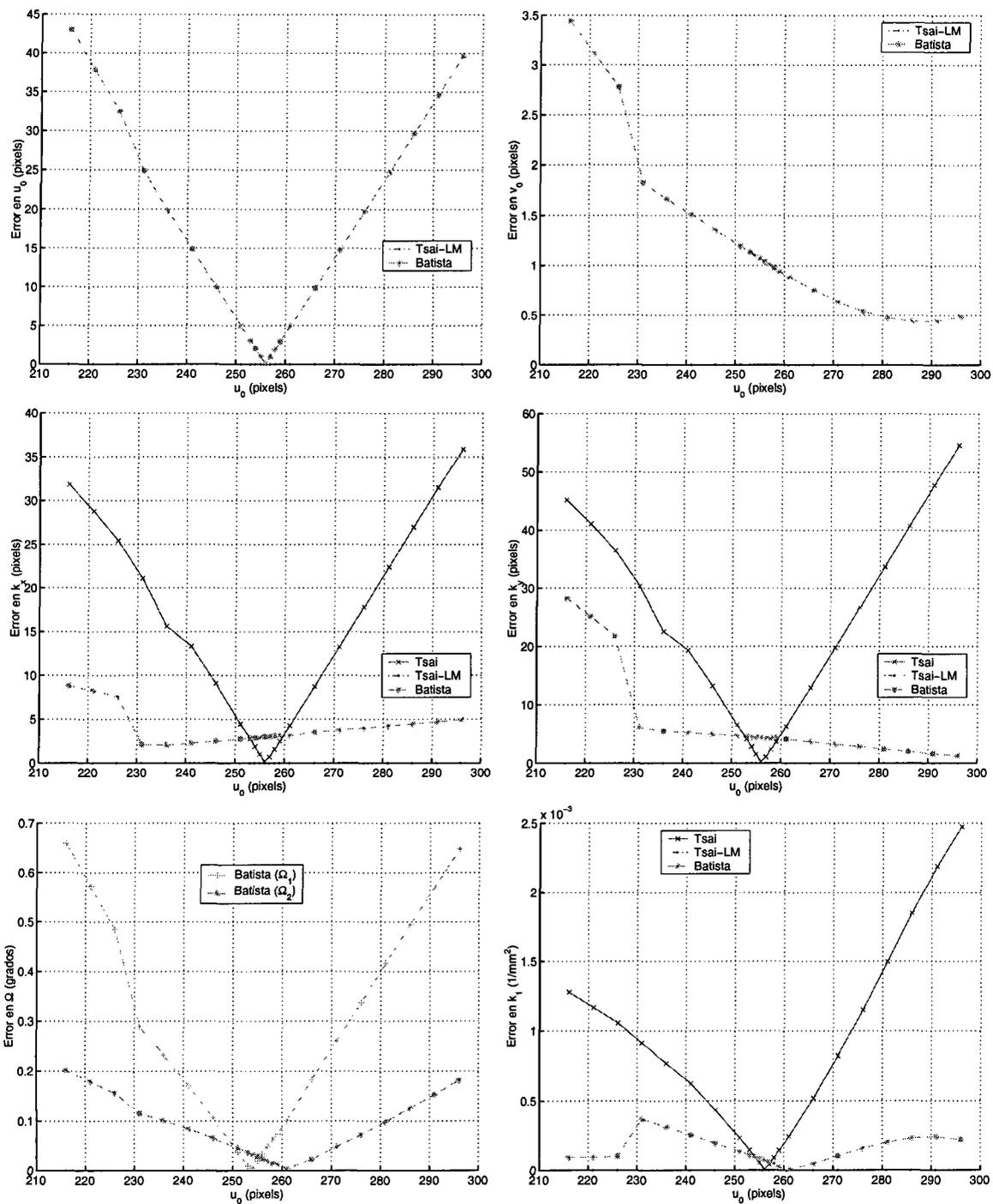


Figura 2.23: Error en el cálculo de los parámetros intrínsecos ante la variación de u_0 .

este parámetro es igual a la separación respecto al centro de la imagen; es decir, este método da un resultado prácticamente igual a la aproximación inicial de este parámetro (el centro del eje óptico). En la gráfica se puede comprobar este efecto ya que se genera un mínimo en ese punto. En la estimación de v_0 se observa que, a medida que aumenta el valor de u_0 , el error disminuye (en este caso no aparece ningún mínimo como ocurre para u_0). Esto se debe a que los puntos del patrón (figura 2.8) se encuentran desplazados en la imagen hacia el lado izquierdo, por tanto, al desplazar a la izquierda el centro del eje óptico, los puntos se encontrarán cada vez más cerca del borde de la imagen, donde el efecto de la distorsión es mucho mayor. En cambio, al desplazarlo hacia la derecha se acerca al centro de la imagen, por lo que el efecto de la distorsión es menor. Aunque en la imagen sintetizada usada en este experimento la distorsión es nula, el método sí que obtiene un pequeño valor para el coeficiente de distorsión, y este error afecta al resto de los parámetros. Por este mismo motivo, si se observa la gráfica de u_0 , también se puede ver que existe un poco más de error en el lado izquierdo de la curva que en el derecho, siendo valores que se encuentran a la misma distancia del centro de la imagen (256).

En el caso del factor de escalado horizontal (k_x), el método de Batista tiene un error bastante inferior al método de Tsai. Pero, mientras el método de Tsai tiene el mínimo del error de k_x en el centro de la imagen (256), en el método de Batista este valor está desplazado a la izquierda (231), esto es debido a que k_x es una combinación de tres parámetros ($k_x = s_x d_x f$) y como este método supone inicialmente que $s_x = 1$, el error se transmite a k_x . Como en la configuración de la cámara usada s_x es menor que uno, el mínimo se encuentra desplazado a la izquierda del centro; mientras que si fuera mayor que uno se encontraría desplazado a la derecha. Si se observan las gráficas, el método de Tsai genera una curva en forma de 'V', donde el lado derecho es mayor que el izquierdo, lo que se debe a que el patrón en la imagen se encuentra desplazado hacia un lado de la misma.

Las curvas que se obtienen de Ω_1 y Ω_2 para el método de Batista tienen forma de 'V' y se observa que Ω_1 es mucho más sensible al desplazamiento horizontal del centro óptico. En el caso del coeficiente de distorsión, el método de Batista presenta un mínimo que está desplazado respecto al centro de la imagen debido a la configuración de s_x , y el método de Tsai obtiene diferente tamaño en los lados de la "V" que forma la curva debido a la posición de los puntos del patrón en la imagen.

Como en los anteriores experimentos mostrados, existe una coincidencia entre los resultados de T_x con u_0 y de T_y con v_0 . Sólo destacar la forma de la curva del método de Tsai, cuyo comportamiento es lógico ya que, al no estimar el centro del eje óptico y dada la relación entre ambos parámetros (centro del eje óptico con traslación en X e Y), el error será proporcional a la diferencia entre el centro del eje óptico y el centro de la imagen.

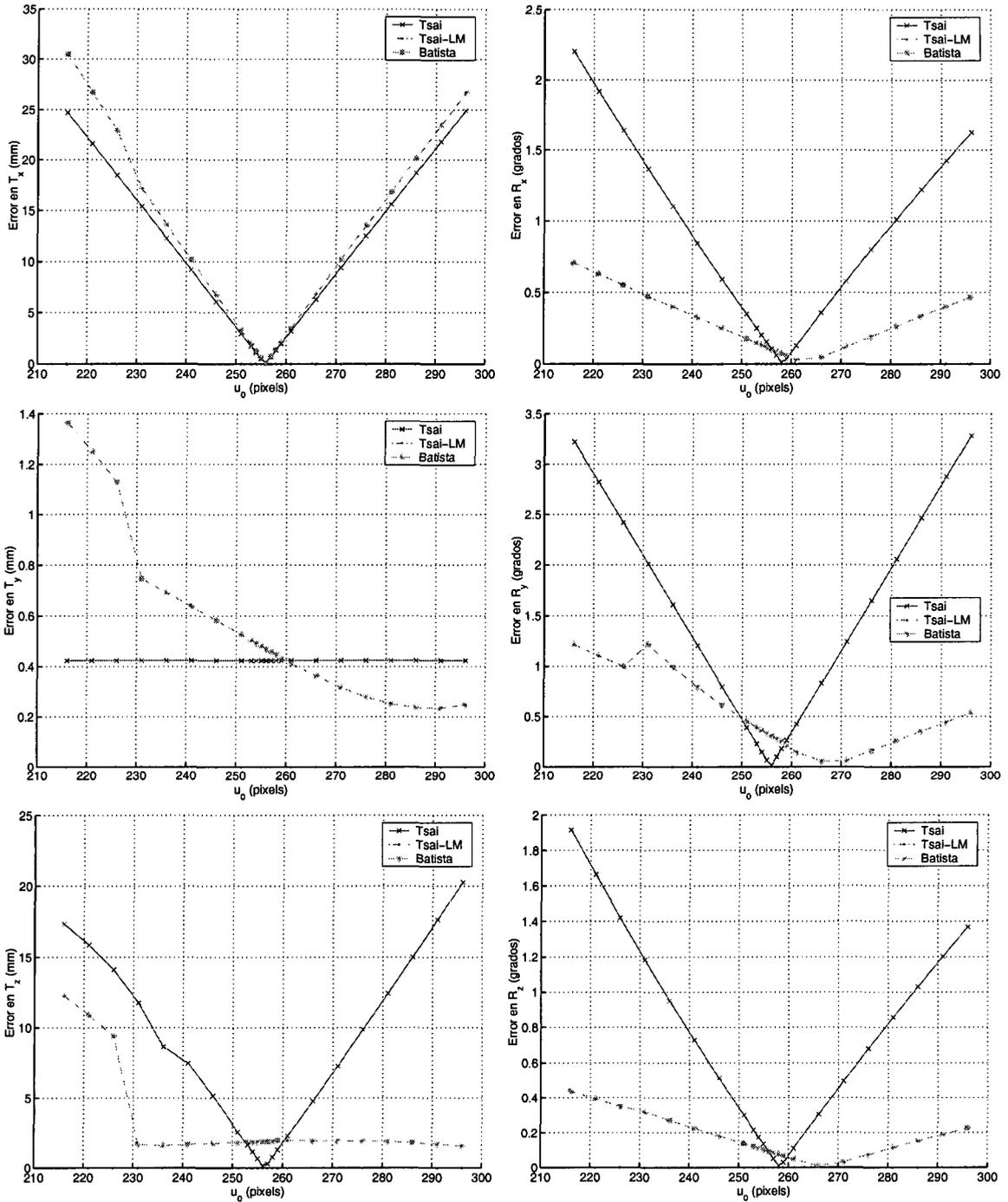


Figura 2.24: Error en el cálculo de los parámetros extrínsecos ante la variación de u_0 .

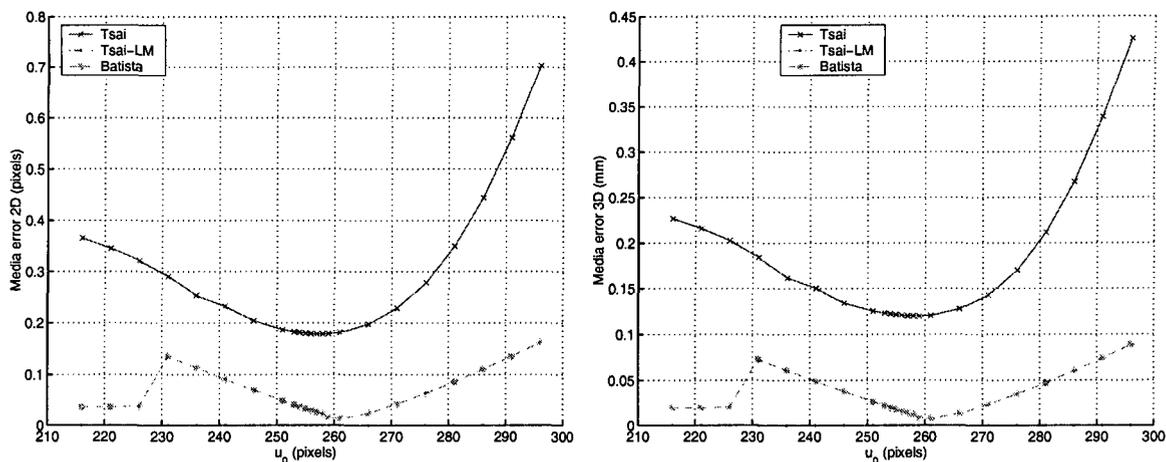


Figura 2.25: Error promedio ante la variación de u_0 .

En el caso de T_z coinciden los resultados con los de los factores de escalado (en especial con k_y). La única diferencia estriba en que en el método de Batista el desplazamiento del mínimo de la curva es hacia la derecha en vez de a la izquierda, aunque la causa que motiva el desplazamiento es la misma que se ha mencionado anteriormente.

En las gráficas relativas a los ángulos de la orientación, el mínimo de las curvas correspondientes al método de Tsai aparece ligeramente desplazado respecto al centro de la imagen en algunas gráficas debido a la influencia del valor de s_x en la configuración de la cámara. El mismo efecto ocurre en el caso del método de Batista (aunque de forma más atenuada). La excepción es el caso de R_z en el que, aparte del factor mencionado, también influye la configuración de v_0 (en la cámara el valor es de 255, mientras que el centro de la imagen es 256).

En la figura 2.25 se puede observar, aparte de lo ya mencionado en lo referente al desplazamiento del mínimo de las curvas y la inclinación de la misma, que el error en el método de Tsai es muy superior al del método de Batista. Llama la atención el hecho de que, aunque haya parámetros que se ven muy afectados por la separación entre el centro óptico y el centro de la imagen, el error en la reconstrucción de los puntos es mínimo. En especial, en el caso del método de Batista donde no se superan los 0.2 pixels de error en las coordenadas 2D y los 0.1 mm en las coordenadas 3D.

Finalmente, como resumen, se puede indicar que en este experimento se muestra cómo los métodos que resuelven la calibración presuponiendo unos valores (centro del eje óptico en el centro de la imagen) para la estimación de otros parámetros se ven afectados por un error generalizado en todos los parámetros. De los métodos implementados, sólo los métodos de Tsai y Batista son sensibles a este problema. En los resultados obtenidos también se observa la interdependencia de algunos parámetros, por ejemplo, como afecta la configuración de s_x para obtener el mínimo de las curvas, y también como influye en

los resultados la situación dentro de la imagen de los puntos del patrón.

Variación de la desviación respecto a la ortogonalidad del plano imagen

Hay métodos que no contemplan este parámetro en el modelo de cámara que resuelven. Tres de los métodos utilizados en este estudio no obtienen este parámetro: Tsai, Tsai-optimizado y Heikkilä; además, está el método de Batista que lo modela con dos parámetros y en las primeras fases del método los aproxima a los 90° . Con este experimento se intenta mostrar cómo afecta la desviación de la ortogonalidad del plano imagen en el resultado del resto de los parámetros en estos métodos. Para ello, se ha variado un grado, en ambos sentidos, respecto a los 90° que forman los dos ejes del plano imagen en condiciones ideales.

Como era de esperar, en los resultados del experimento se observa que los únicos métodos afectados por esta inclinación de los ejes son los de Batista, Tsai, Tsai-LM y Heikkilä. Para aportar mayor claridad a las gráficas (figuras 2.26 - 2.28) sólo se muestran las curvas de estos cuatro métodos.

Se puede observar que en casi todos los casos el método de Heikkilä es el que peor estimación hace de los parámetros cuando Ω dista de los 90° . La curva que genera tiene forma de 'V' y debido a que el patrón se encuentra desplazado del centro de la imagen, los dos lados de la curva tienen distinta forma.

En la estimación del centro óptico se puede observar que este método llega a alcanzar errores superiores a los 30 pixels en u_0 y superiores a 60 pixels en v_0 cuando la desviación es de 1 grado. En cambio, el error en los factores de escalado no es tan desproporcionado, aunque sí elevado en comparación al resto.

El método de Tsai-LM también presenta altos niveles de error en casi todos los parámetros, aunque un poco más atenuados que los del método de Heikkilä. Al igual que en el otro método, la curva tiene forma de 'V', y uno de los lados tiene mayor envergadura que el otro, debido a que el patrón no se proyecta sobre el centro de la imagen. En ambos casos el mínimo se encuentra en los 90° donde el error es nulo. En la estimación del centro óptico se puede observar que este método ronda los 30 pixels de error con un grado de desviación respecto a la ortogonalidad. En este caso, también el error en la estimación de los factores de escalado es bastante inferior al obtenido en el centro óptico. Este error es de alrededor de 2 pixels en k_x y de 6 pixels en k_y .

En las curvas correspondientes al método de Batista se puede observar que para u_0 el error aumenta ligeramente a medida que aumenta el ángulo entre los ejes del plano imagen, mientras que para v_0 ocurre lo contrario. Si se aumentara el grado de desviación en el experimento se podría ver que este error alcanza un mínimo (error nulo). Esta

desviación del mínimo de la curva respecto a los 90° viene dado por el hecho de que el centro del eje óptico este separado del centro de la imagen y que s_x no sea igual a 1 en la configuración de la cámara que se ha utilizado. Aún así, el error cometido es muy pequeño en comparación con el de los otros dos.

En los parámetros extrínsecos el comportamiento es semejante al descrito hasta ahora en este experimento. El método de Tsai parece no verse afectado en la mayoría de los casos por la desviación respecto a la ortogonalidad de los ejes del plano imagen. Sólo parece verse afectado mínimamente en unos pocos parámetros: R_y , R_z y T_x , aunque sólo en una décima de grado los ángulos y en un par de milímetros la posición. La desviación del mínimo de las curvas respecto a los 90° es producto de la configuración del centro del eje óptico.

Como se puede ver en las gráficas de la figura 2.28, el método de Tsai es el más sensible a la variación de Ω . También el método optimizado de Tsai-LM tiene un elevado error, aunque de menor magnitud que el obtenido por la versión clásica del método. El comportamiento del método de Tsai es destacable ya que no hay muchos parámetros afectados por la asimetría del plano imagen y, sin embargo, tanto al reconstruir los puntos en la imagen como la pose del patrón, se ven afectados por esta desviación. En cambio, en el método de Heikkilä se produce el efecto contrario, es decir, los distintos parámetros se ven altamente afectados por la ortogonalidad de los ejes del plano imagen, pero no así la reconstrucción de los puntos. Finalmente, se observa que el parámetro más sensible en esta situación son el centro óptico, y los parámetros extrínsecos.

En resumen, como ocurría en el experimento anterior, se puede observar la interdependencia de los distintos parámetros y también como afecta la ausencia de un parámetro en el modelo a la resolución del problema de la calibración. En general, los que parecen más afectados en todos los parámetros son los métodos de Heikkilä y Tsai-optimizado.

Conclusión

Los resultados obtenidos en esta sección muestran la fuerte dependencia que tienen los distintos métodos con las aproximaciones iniciales o con los parámetros que no modelan. La magnitud del error cuando las condiciones distan de las ideales superan a los errores producidos por el ruido en las coordenadas de los puntos (tanto 2D como 3D), para los niveles de ruido y las desviaciones de los parámetros consideradas en estos experimentos. Se puede observar que todos los métodos presentan deficiencias. Por un lado, los métodos de Zhang, Tsai-LM, Faugeras, Lineal y Ahmed, son muy sensibles a la distorsión; por otro lado, los métodos de Tsai y Batista son muy sensibles al valor del

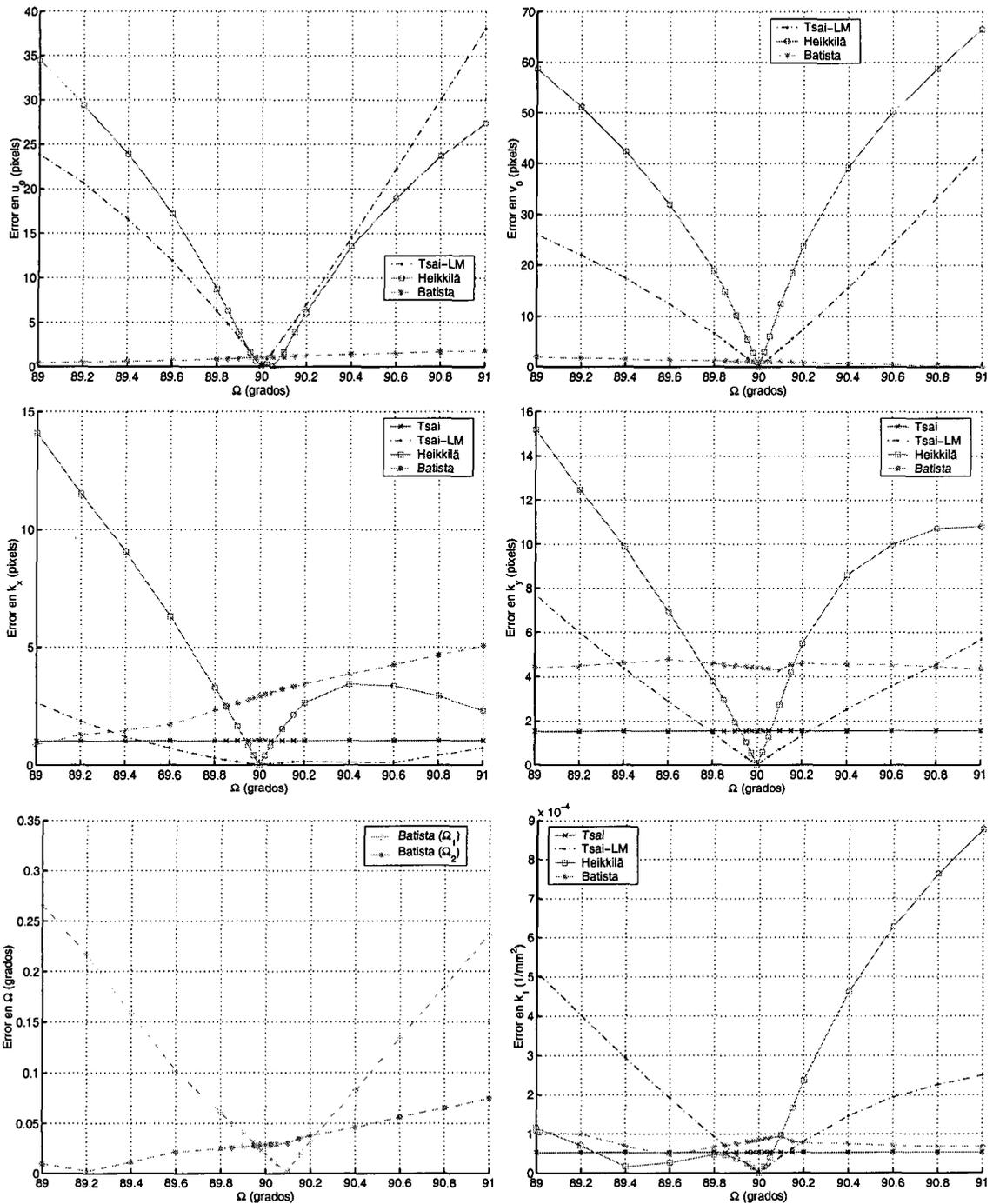


Figura 2.26: Error en el cálculo de los parámetros intrínsecos ante la variación de la desviación respecto a la ortogonalidad de los ejes del plano imagen.

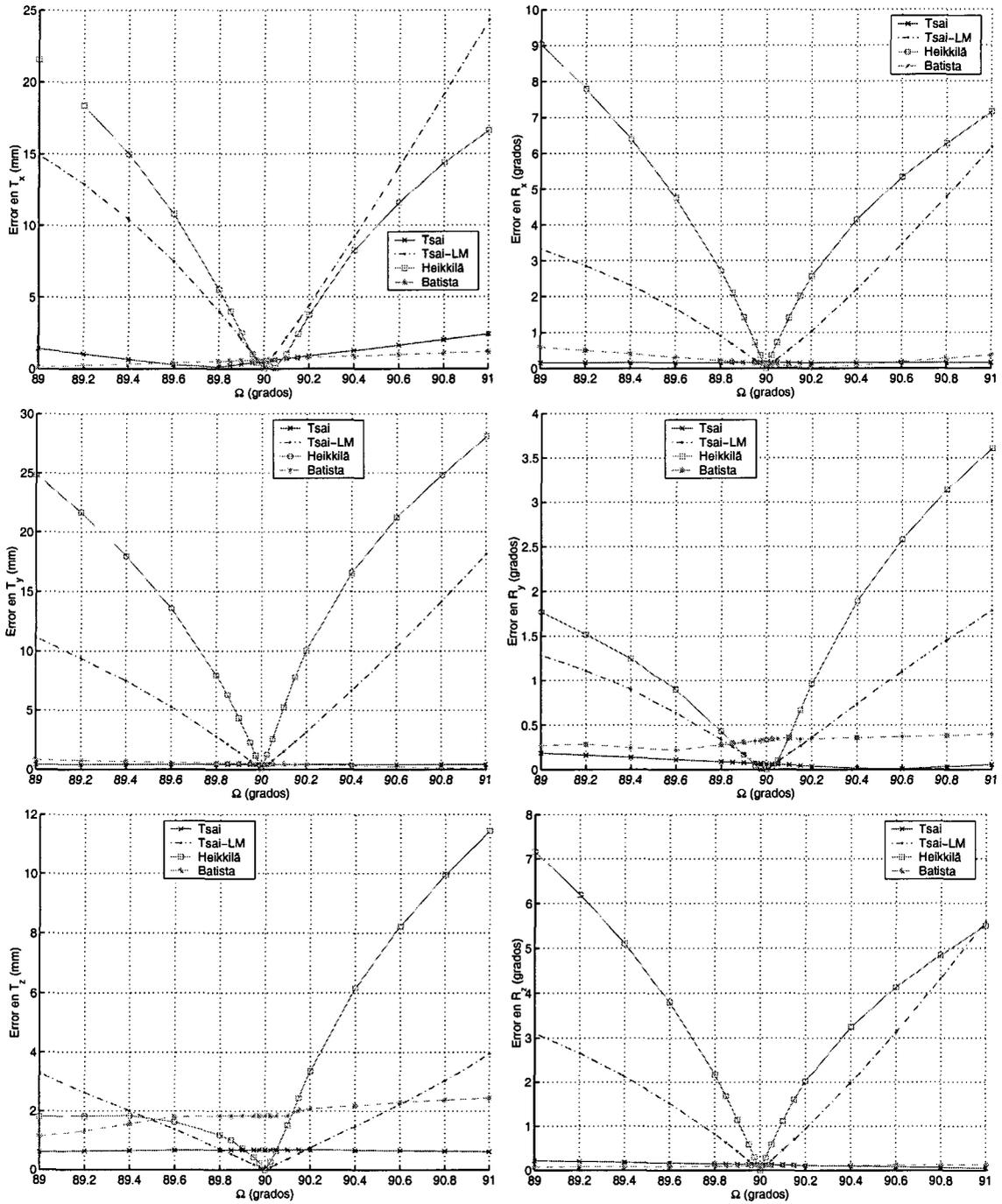


Figura 2.27: Error en el cálculo de los parámetros extrínsecos ante la variación de la desviación respecto a la ortogonalidad de los ejes del plano imagen.

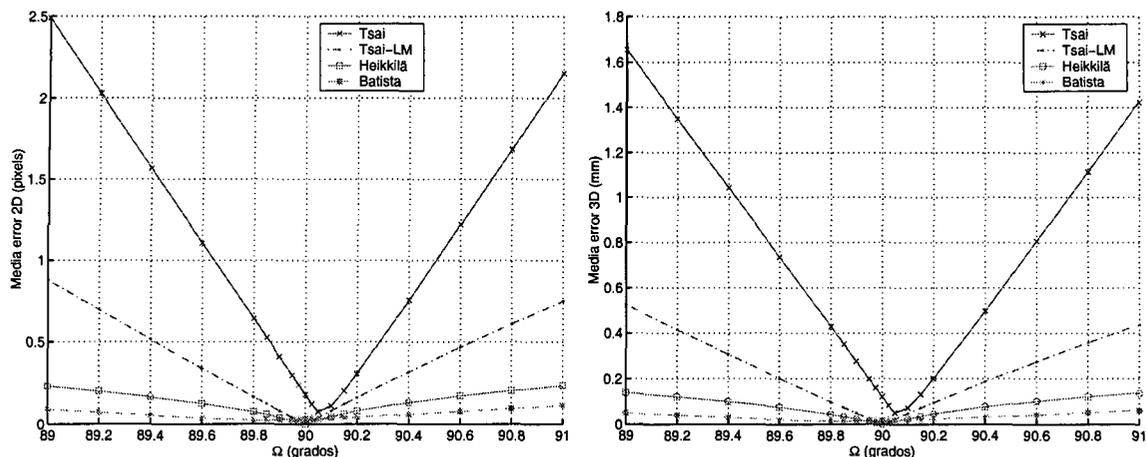


Figura 2.28: Error promedio ante la variación de la desviación respecto a la ortogonalidad de los ejes del plano imagen.

centro óptico y en menor medida a la ortogonalidad de los ejes del plano imagen; factor este último que también influye sobremanera en los métodos de Tsai-LM y Heikkilä.

2.3.3. Variación de la colocación del patrón

En esta sección se analiza si la posición del patrón influye a la hora de obtener buenos resultados en la calibración. El objetivo de los siguientes experimentos será determinar qué factores influyen en los resultados y qué parámetros son los que se ven más afectados. También será importante obtener cuál es la posición más adecuada que deben tener los patrones respecto a la cámara para conseguir los mejores resultados. A priori parece que la posición más lógica sea aquella que presente una perfecta simetría del patrón respecto al centro de la imagen.

El estudio se ha centrado en analizar la influencia de la distancia a la que se encuentra el patrón ante diversas circunstancias.

Variación de la distancia del patrón (T_z)

La distancia a la que se encuentra el patrón de calibración influye en el tamaño que tiene la proyección del mismo en el plano imagen. Luego, lo que se evalúa con este experimento es en qué medida afecta la porción de imagen que ocupa el patrón en los resultados de la calibración. Debido a que cada patrón se debe colocar a diferente distancia para que ocupe la misma porción de imagen, en las gráficas se muestra la variación respecto a la posición inicial utilizada en el resto de los experimentos (sección 2.3). Es decir, ΔT_z , respecto a la posición inicial, ha variado desde acercar el patrón 60 cm hasta alejarlo 197 cm. Por tanto, el plano utilizado para la calibración del método

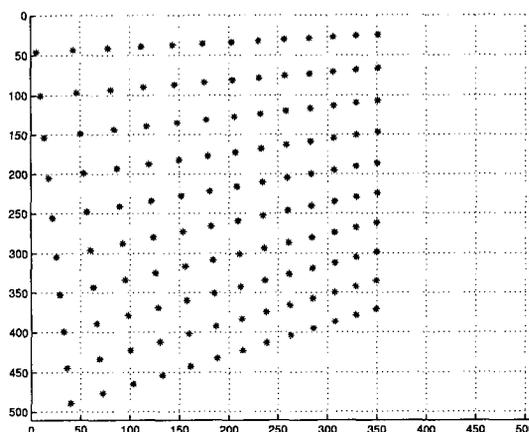


Figura 2.29: Patrón usado en el método de Batista situado a 39.3 cm y a 238.3 cm.

de Batista ha variado entre 39.3 y 238.3 cm (Figura 2.29). El plano utilizado en la calibración del método de Zhang ha variado entre 35 y 234 cm en la primera vista, entre 36 y 235 cm en la segunda, y entre 37.5 y 236.5 cm en la tercera. El diedro utilizado en el resto de los métodos ha variado su distancia respecto a la cámara entre 36 cm y 235 cm. Las imágenes están libres de ruidos y la distorsión es nula.

En los resultados, que se muestran en las figuras 2.30 - 2.32, se observa que, en condiciones libres de ruido y distorsión, los únicos métodos que se ven afectados por el desplazamiento del patrón son los métodos de Batista, Tsai, y, en menor medida, el método de Heikkilä. En las gráficas sólo se muestran estos métodos para una mejor lectura de los resultados.

En las gráficas relativas al error en la estimación del centro óptico se observa que el método de Heikkilä se ve afectado a partir de distancias superiores a metro y medio (el error aumenta, pero con valores inferiores a una décima de pixel). En cambio, el método de Batista es más sensible a la distancia, en concreto, para v_0 tiene un error menor cuando el patrón está cerca, a menos de 65 cm, luego se estabiliza la diferencia de error (respecto al valor mínimo) en 0.1 pixel. En el caso de u_0 se produce un efecto contrario, ya que disminuye. Este efecto, al igual que el alto error que mantiene en todos los valores está motivado por la diferencia entre el centro del eje óptico y el centro de la imagen en la configuración de cámara usada para los experimentos.

En las gráficas de los factores de escalado se puede ver que el método de Tsai disminuye muy levemente el error a medida que se va alejando el patrón de la cámara. El método de Batista ofrece un mínimo en valores intermedios entorno a los 50-55 cm., en distancias más cortas aumenta un poco el error, pero aumenta mucho más en largas distancias. Por encima de un metro el error supera los 8 pixels. Este comportamiento se debe al error en la distancia focal; el valor de este parámetro varía ligeramente en

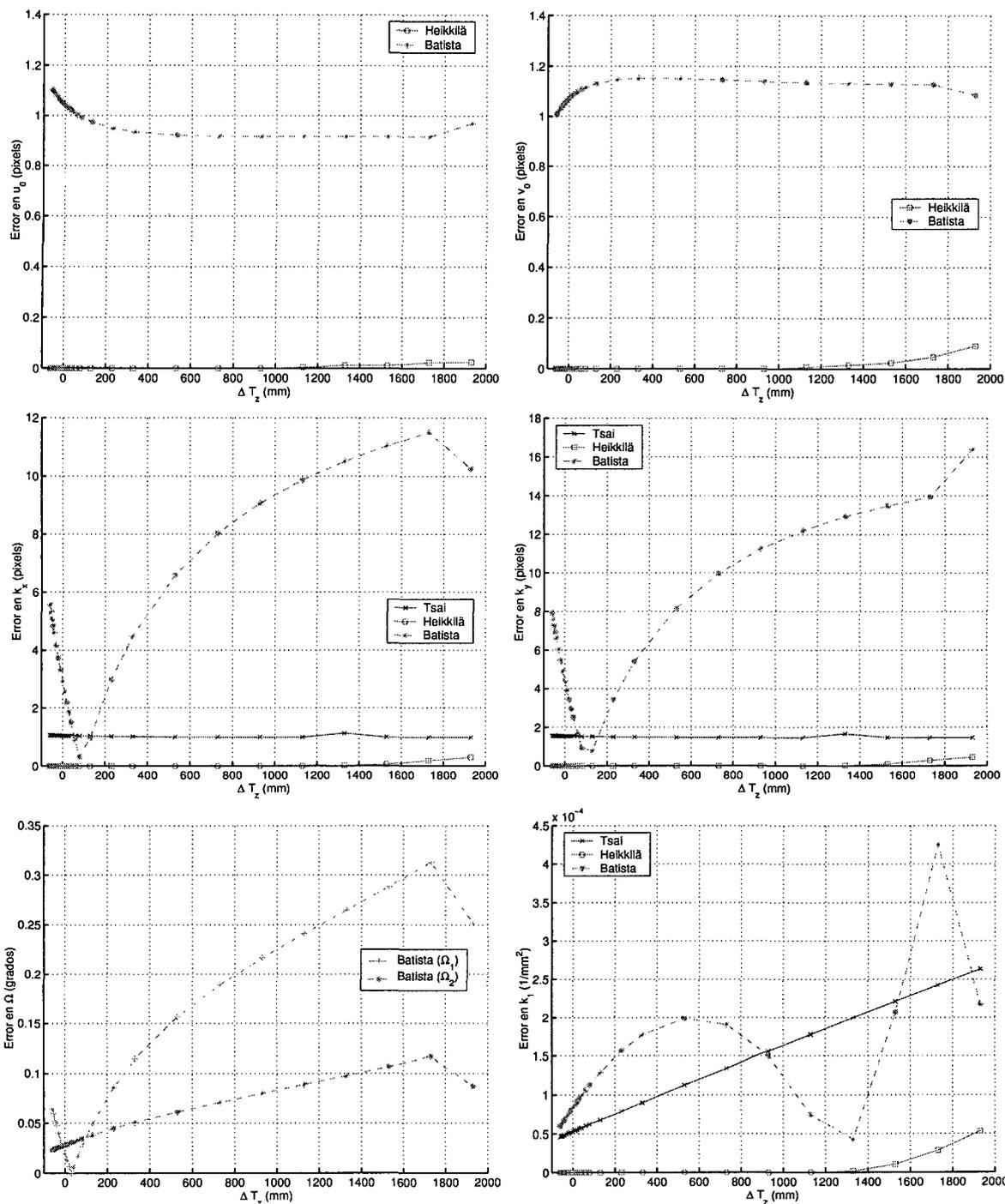


Figura 2.30: Error en el cálculo de los parámetros intrínsecos ante la variación de la distancia entre el patrón y la cámara.

función de la distancia y el error cometido se traslada al resto de los parámetros. El método de Heikkilä tiene un comportamiento muy parecido al resultante de la estimación del centro óptico.

El método de Batista tiene un aumento progresivo del error en la estimación de la ortogonalidad a medida que es mayor la distancia entre la cámara y el patrón. En cuanto al coeficiente de distorsión este método presenta un comportamiento bastante irregular. El método de Tsai tiene un crecimiento constante del error según va aumentando la distancia. Por último, el método de Heikkilä muestra un comportamiento similar al descrito en los anteriores parámetros de este experimento.

Los métodos de Tsai y de Batista son muy parecidos, tanto en la estimación de T_x como de T_y , en ambos caso el error es proporcional a la distancia entre el patrón y la cámara, aunque en el caso de Tsai el incremento es un poco más moderado. Si bien el método de Tsai mantiene este mismo comportamiento para la estimación de T_z , el método de Batista presenta un error mucho más elevado, siendo la curva parecida a la obtenida en el cálculo de los factores de escalado. El método de Heikkilä muestra un ligero error a partir de unos 180 cm de separación.

El método de Tsai no se ve afectado en la estimación de la orientación del patrón ante la variación de la distancia del mismo respecto de la cámara. El método de Batista tiene un comportamiento peculiar, mientras para R_x y R_y las gráficas son semejantes a la de los factores de escalado (hay que recordar que en el cálculo de dos de los ángulos interviene la distancia focal (2.35) y (2.38)), en cambio, para R_z disminuye constantemente el error a medida que aumenta la distancia. El método de Heikkilä mantiene el mismo comportamiento que en los otros parámetros.

En la gráfica que muestra el error promedio en la imagen (Figura 2.32) se observa que en el método de Tsai desciende el error a medida que el patrón se va alejando de la cámara. Esto parece lógico ya que al aumentar la distancia del patrón respecto a la cámara, el tamaño del mismo en la imagen es cada vez más pequeño. Por tanto, las imprecisiones que presente el método se verán atenuadas ya que las distancias entre puntos en el plano imagen serán mucho menores. El efecto contrario se produce en la reconstrucción de las coordenadas en el mundo del patrón a partir de su imagen, ya que, pequeños errores en la estimación de los valores de los parámetros, especialmente de los intrínsecos, se verán aumentados a medida que aumente la distancia. El otro método que se ve sensiblemente afectado es el de Batista, el cuál tiene un comportamiento parecido en las dos gráficas. En ambas aparece un mínimo en torno a los 170 cm. Este comportamiento se debe a la influencia del factor de proporción (s_x), ya que si este valor en la configuración de la cámara fuera uno, el error promedio en 2D sería prácticamente nulo y en 3D aumentaría ligeramente al aumentar la distancia.

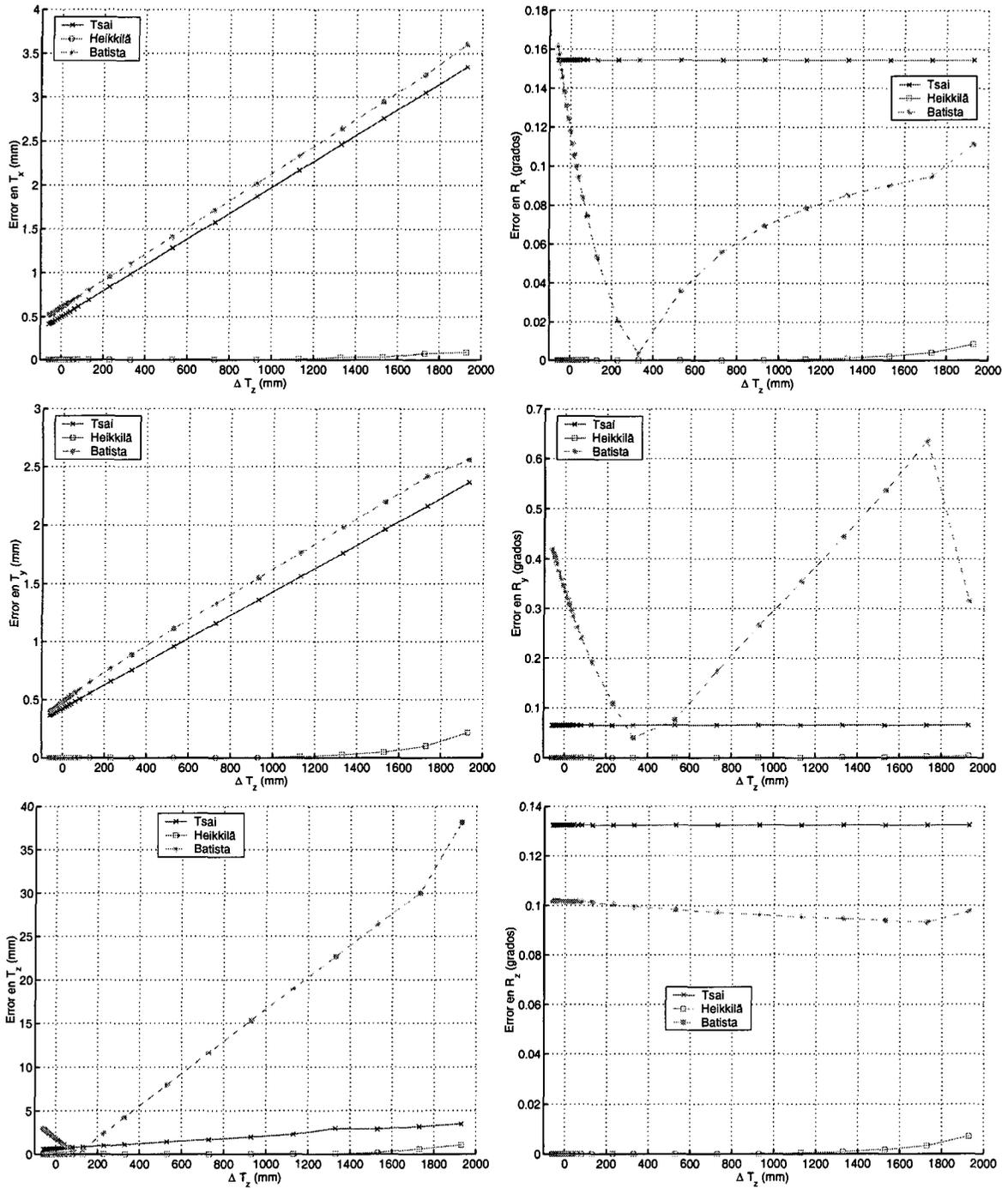


Figura 2.31: Error en el cálculo de los parámetros extrínsecos ante la variación de la distancia entre el patrón y la cámara.

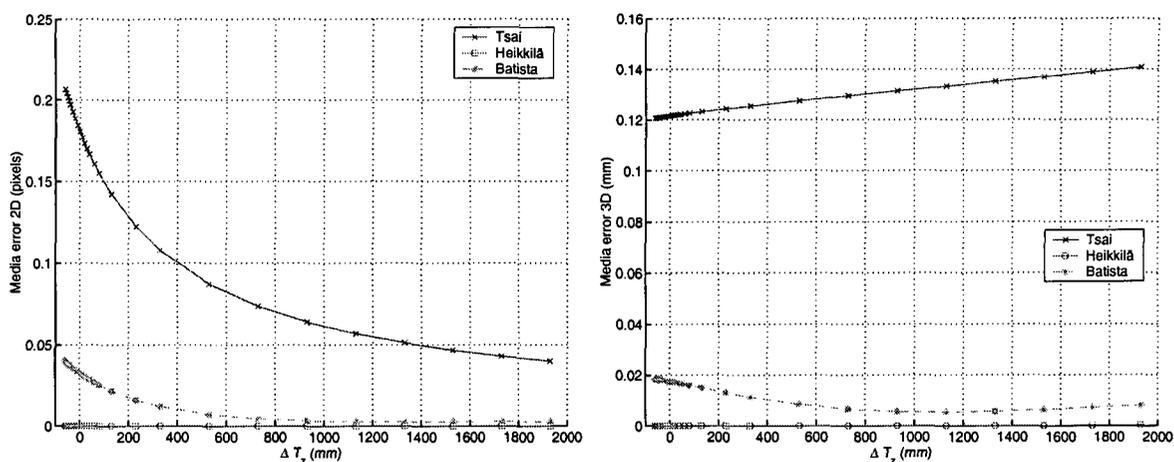


Figura 2.32: Error promedio ante la variación de la distancia entre el patrón y la cámara.

Como conclusión del experimento se pueden destacar un par de observaciones. Por un lado, sólo tres de los métodos se ven afectados por la distancia a la que se encuentra el patrón respecto a la cámara: Batista, Tsai, y en menor medida, Heikkilä. En este caso coincide que estos métodos presuponen inicialmente conocidos algunos de los parámetros (centro del eje óptico, factor de proporción, etc.) lo que provoca errores en el cálculo del resto de los parámetros, como ya se ha comentado en otros experimentos anteriores. Por otro lado, como en todos los métodos se intenta minimizar el error entre la imagen obtenida y la proyección aplicando el modelo, al aumentar la distancia a la que se encuentra el patrón este error será mucho más pequeño ya que la distancia entre puntos es mucho menor (como refleja la gráfica del error promedio en 2D), aún cuando la estimación de cada parámetro sea peor. El resto de métodos no se ve afectado, ya que en condiciones ideales, es decir, en ausencia de ruidos, el error de estos métodos es nulo.

Variación de la distancia del patrón (T_z) en imágenes distorsionadas

Se ha repetido el experimento anterior, pero esta vez se han utilizado imágenes con distorsión ($k_1 = -0,005$). En este experimento se ha utilizado el mismo rango de variación de la distancia que en el anterior caso. Por tanto, el plano utilizado para la calibración del método de Batista ha variado entre 39.3 y 238.3 cm (Figura 2.33). El plano utilizado en la calibración del método de Zhang ha variado entre 35 y 234 cm en la primera vista, entre 36 y 235 cm en la segunda, y entre 37.5 y 236.5 cm en la tercera. El diedro utilizado en el resto de los métodos ha variado su distancia respecto a la cámara entre 36 cm y 235 cm. Las imágenes están libres de ruidos. La importancia del tamaño del patrón viene dada por el hecho que cuanto más porción de imagen es ocupada, los puntos más se acercan a los bordes de la imagen, zona donde es mayor el

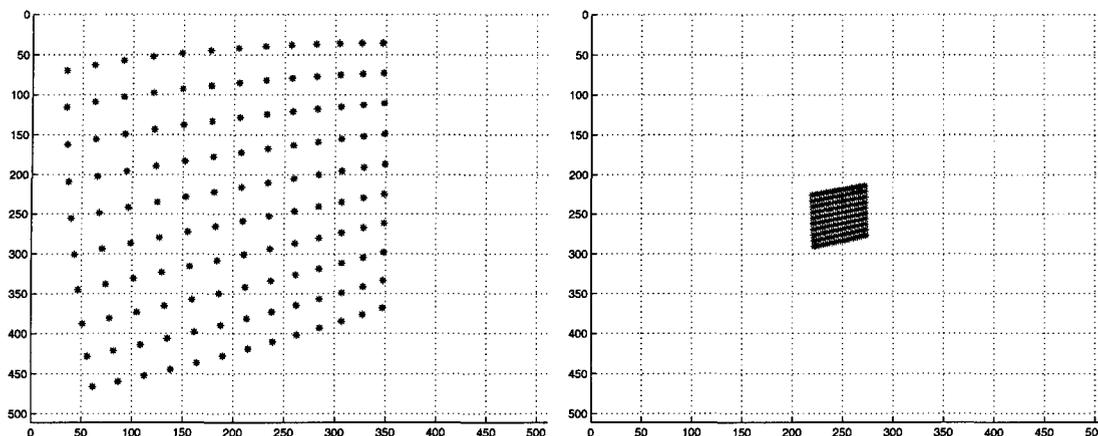


Figura 2.33: Patrón usado en el método de Batista situado a 39.3 cm y a 238.3 cm con distorsión en la imagen ($k_1 = -0,005$).

efecto de la distorsión radial. Por tanto, teóricamente, los errores deben ser de mayor magnitud cuando el patrón está más cerca de la cámara (salvo aquellos métodos que modelen muy bien la distorsión). En las gráficas de las figuras 2.34 y 2.35 se pueden observar los resultados de este experimento para cada uno de los parámetros.

Como era de esperar (ya que son los métodos más sensibles a valores altos de distorsión), los métodos que se ven más afectados son los de Zhang, Lineal, Faugeras y Ahmed. Como es lógico, a medida que se aleja el patrón de la cámara (y es más pequeño en la imagen), la distorsión tiene un menor efecto (el patrón ocupa los puntos centrales de la imagen) y por tanto el error se reduce. El método de Batista obtiene un resultado muy parecido al experimento anterior. Ya que como se puede observar, el error cometido por la influencia del valor de la distorsión es mínimo. Al contrario que el resto de los métodos, el método de Heikkilä sólo presenta errores apreciables cuando el patrón se encuentra alejado de la cámara al igual que ocurría en el experimento anterior. Como en el caso del método Batista, este método modela adecuadamente la distorsión y, por tanto, no se ve afectado por la presencia de ésta en la imagen. En el caso del método de Tsai-optimizado se observa un leve error si la distancia es corta, pero para valores mayores el error es inapreciable. Esto se debe a que este método también modela bastante bien la distorsión como se pudo ver en la figura 2.20.

También hay que hacer notar que el error para u_0 es mucho mayor que para v_0 . Esto se debe a que como se observa en la figura 2.33 los patrones se encuentran desplazados en la imagen hacia la izquierda, por tanto, los valores horizontales se encuentran descompensados (la mayoría de los puntos se encuentran en el mismo lado de la imagen), mientras que en el caso de los valores verticales existe una proporción parecida entre los que se encuentran en la parte superior e inferior de la imagen.

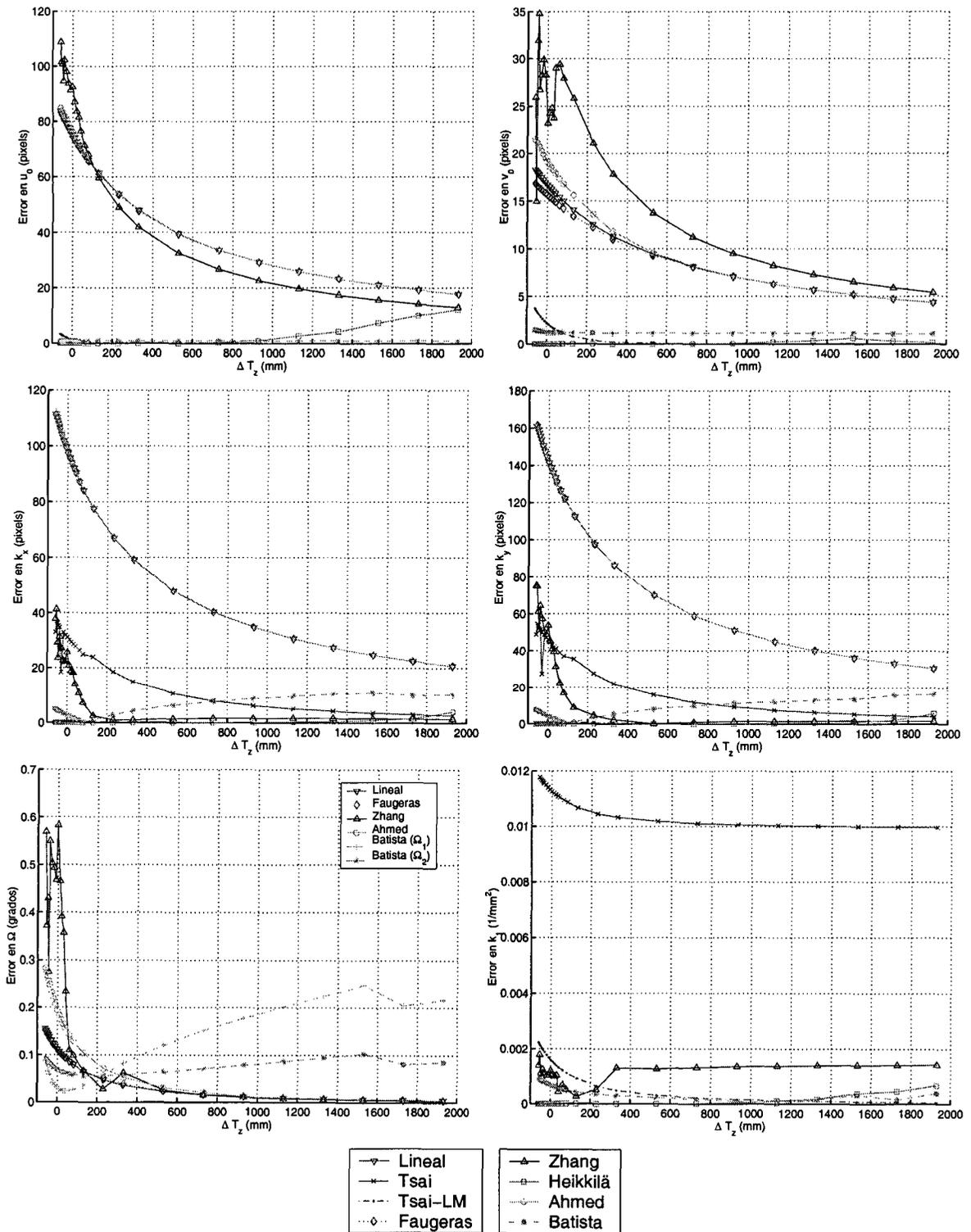


Figura 2.34: Error en el cálculo de los parámetros intrínsecos ante la variación de la distancia entre el patrón y la cámara usando imágenes con distorsión ($k_1 = -0,005$).

Las curvas de error de los factores de escalado de casi todos los métodos son similares a las del centro del eje óptico, salvo en el caso del método de Zhang que, aunque presenta el mismo comportamiento, el error es mucho más atenuado en relación al resto. En cambio, el método de Batista muestra un error mucho más acentuado y va aumentando a medida que aumenta la distancia, tal y como se apreciaba en las gráficas del experimento anterior. El método de Tsai va disminuyendo el error a medida que el patrón se va alejando de la cámara.

La mayoría de los métodos que obtienen el ángulo que forman los ejes del plano imagen (Ω) tienen un comportamiento parecido al observado en las gráficas anteriores de este experimento. Se puede observar que todos presentan un error relativamente alto (décimas de grado) cuando el patrón está muy cerca. Mientras que el error disminuye a medida que el patrón se aleja de la cámara, llegando a tener un error inferior a la centésima de grado a partir de los dos metros. La excepción es el método de Batista que si bien tiene un alto error en distancias cortas, presenta un mínimo (tanto en Ω_1 como en Ω_2) sobre los 60 cm para aumentar a partir de ese punto a medida que aumenta la distancia. Como en anteriores experimentos, el método de Tsai es el que peor estima el coeficiente de distorsión ya que obtiene el coeficiente con el signo cambiado. Sin embargo, este parámetro apenas parece verse afectado por la variación de la distancia en imágenes distorsionadas.

En los métodos de Ahmed, Faugeras y Lineal, a diferencia de las gráficas anteriores, el error en la estimación de la posición del patrón aumenta a medida que la distancia aumenta. Esto probablemente se debe al hecho de que al ser el valor de T_z cada vez mayor, las magnitudes del error son también mayores, pero el porcentaje de error disminuye. En el caso de T_x y T_y se justifica debido a que sus valores dependen de T_z (ec. 2.30). En estos dos parámetros, el resto de los métodos tiene un comportamiento parecido al obtenido para el centro del eje óptico, salvo que el error en el método de Zhang es, respecto a los demás, mucho más atenuado en las distancias más cortas. En el caso de T_z , el comportamiento se asemeja al mostrado en la gráfica de los factores de escalado.

En las gráficas donde se muestra el error en la estimación de la orientación del patrón se sigue conservando el comportamiento visto en el resto de las gráficas de este experimento. La excepción es el método de Zhang, en el que los errores en la rotación alrededor del eje X y alrededor del eje Z son muy pequeños.

En las dos gráficas que indican el promedio del error en la reconstrucción de los puntos (figura 2.36), el comportamiento de todos los métodos es semejante, a medida que aumenta la distancia disminuye el error. Destaca levemente la magnitud del error de los métodos de Zhang y Tsai-optimizado en distancias muy cortas: con una separación de 36 cm el error en las coordenadas 2D ronda los 2 pixels y en las coordenadas 3D es

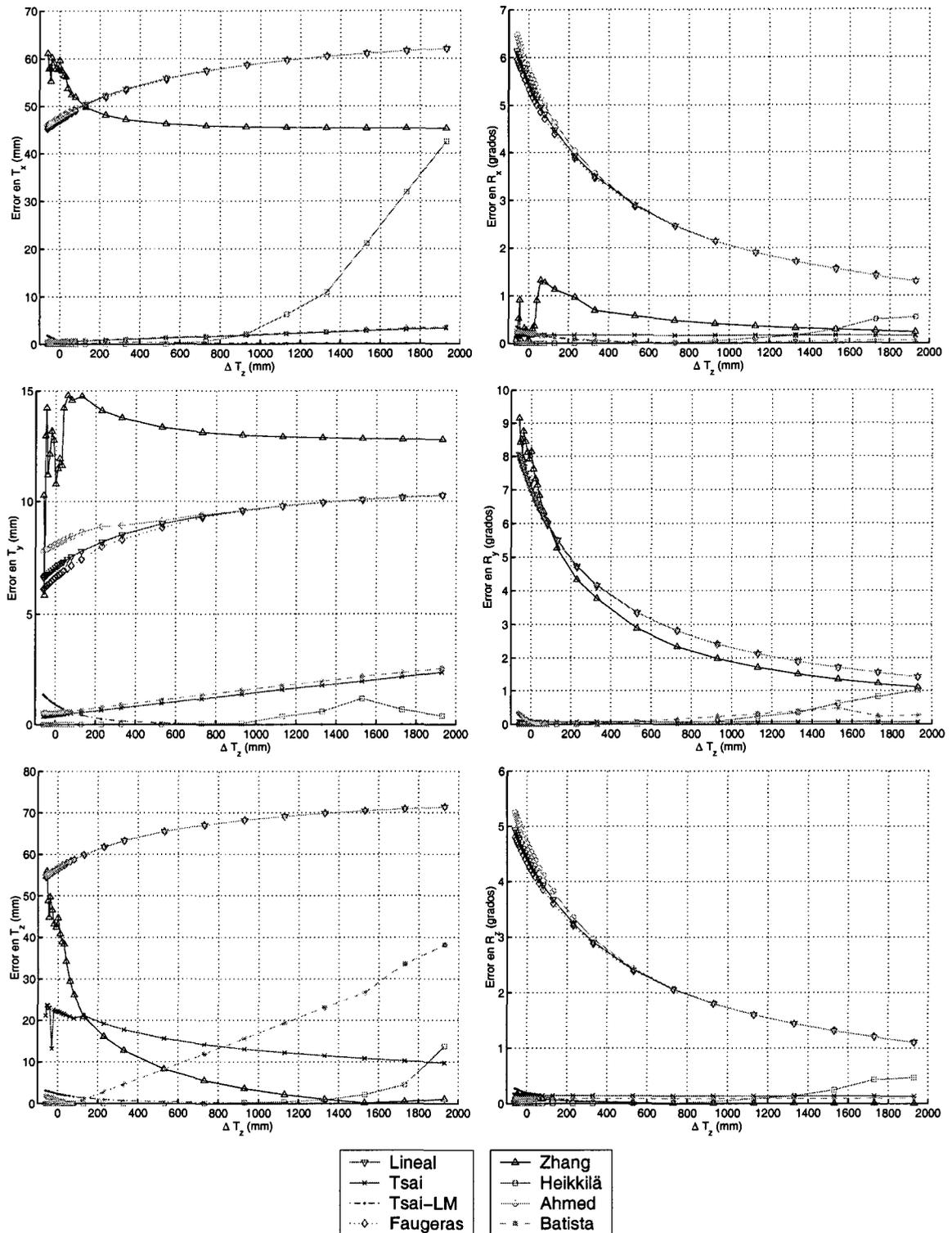


Figura 2.35: Error en el cálculo de los parámetros extrínsecos ante la variación de la distancia entre el patrón y la cámara usando imágenes con distorsión ($k_1 = -0,005$).

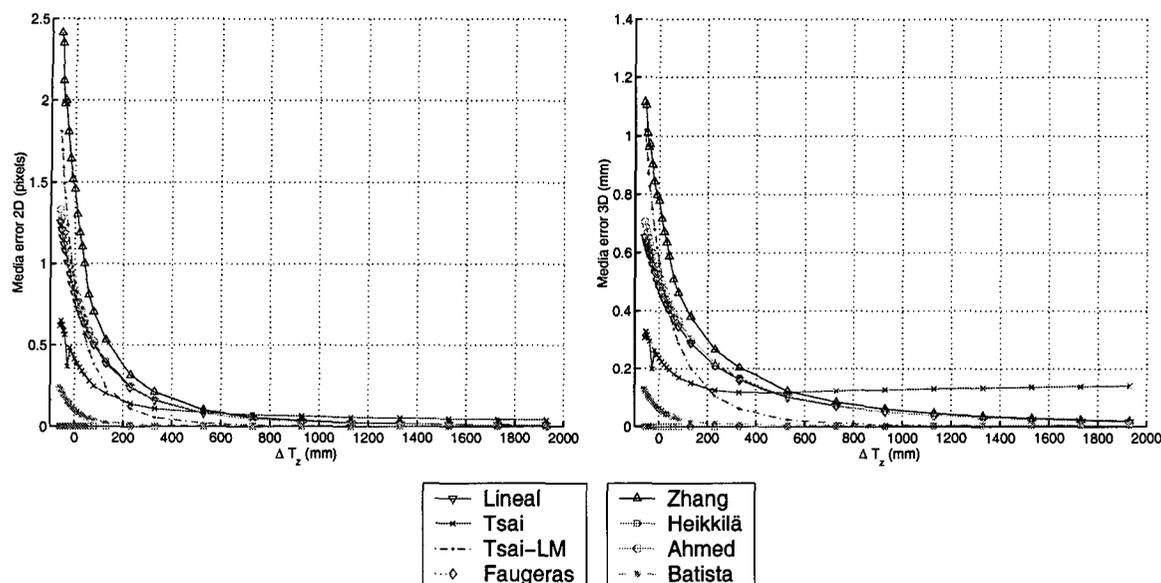


Figura 2.36: Error promedio ante la variación de la distancia entre el patrón y la cámara usando imágenes con distorsión ($k_1 = -0,005$).

de aproximadamente 1 mm.

Como era de esperar, aquellos métodos que no incluyen la distorsión en su modelo de cámara, disminuyen el error a medida que el patrón es más pequeño en la imagen. Por tanto, como los puntos del patrón se encuentran más lejos de los bordes de la imagen, el efecto de la distorsión radial es menor. A estos métodos se une el método de Zhang que tiene un comportamiento semejante. El resto de los métodos varían la forma de la curva de forma distinta en cada parámetro, pero existe una tendencia a disminuir el error a medida que aumenta la distancia. Esto es coherente con el experimento anterior y con el hecho de que estos métodos son los que mejor responden ante valores altos de distorsión.

Variación de la distancia del patrón (T_z) usando imágenes con ruido gaussiano en las coordenadas 2D

Se ha repetido el experimento de la variación de la distancia del patrón, pero esta vez se ha inyectado un ruido gaussiano de media nula y desviación típica 0.5 en las coordenadas 2D de la imagen. Al igual que en los casos anteriores el plano utilizado para la calibración del método de Batista ha variado entre 39.3 y 238.3 cm (Figura 2.33). El plano utilizado en la calibración del método de Zhang ha variado entre 35 y 234 cm en la primera vista, entre 36 y 235 cm en la segunda, y entre 37.5 y 236.5 cm en la tercera. El diedro utilizado en el resto de los métodos ha variado su distancia respecto a la cámara entre 36 cm y 235 cm. El experimento se ha repetido 50 veces en cada una

de las distancias de la muestra. En este caso, la distorsión es nula.

Este experimento se ha realizado para mostrar cómo influye en algunos métodos el ruido en las coordenadas de imagen, según el tamaño que tenga el patrón en la misma. Por ejemplo, el método de Batista se basa en la relación entre las coordenadas de puntos que forman un rectángulo para obtener los ángulos de rotación; si estas coordenadas están afectadas por ruido, la influencia del mismo será mucho mayor si los puntos están muy próximos entre sí.

En las figuras 2.37 y 2.38 se puede observar que todos los métodos se ven influenciados por la distancias del patrón. Cuando existe ruido en las coordenadas de la imagen, el error aumenta a medida que aumenta la distancia.

Al contrario de lo que se planteaba a priori, el método de Batista parece el más preciso en la obtención del centro del óptico, cuando el patrón se encuentra bastante alejado. Aunque como se mostró en el primer experimento, este método era el más estable en presencia de ruidos en las coordenadas de los puntos (tanto 2D como 3D). Esto se debe a que el valor que se obtiene para este parámetro es siempre muy cercano a la aproximación inicial que se le proporciona (centro de la imagen), y como en la configuración de la cámara usada en los experimentos sólo hay un pixel de diferencia entre ambos centros, es lógico que el error sea pequeño. También se muestra poco sensible el método de Tsai-LM. Por contra, llama la atención lo sensible que es el método de Ahmed en la estimación de este parámetro cuando el patrón está muy alejado de la cámara y las coordenadas de imagen tienen ruido.

En las gráficas que muestran el error en la estimación de los factores de escalado se observa que los métodos siguen un comportamiento semejante al caso del centro del eje óptico. Las excepciones vienen dadas por el método de Tsai, que no calcula los parámetros anteriores y por tanto no aparece en las gráficas, que se muestra como un método muy sensible, en comparación al resto, a que el patrón se encuentre muy separado de la cámara. El método optimizado de Tsai también presenta un alto error en relación a los demás métodos. En estos métodos la causa de este comportamiento se debe a que la distancia focal se estima en un segundo paso junto con la distancia del patrón y el coeficiente de distorsión, arrastrándose en esta fase los pequeños errores cometidos en la primera fase motivados por la fijación de algunos parámetros. Por otro lado, también se observa que la incidencia de la distancia en el método de Batista es ligeramente mayor que en el caso del centro óptico y, en cambio, en el caso del método de Ahmed es a la inversa.

Al igual que en otros experimentos, el método de Batista es el que peor estima el ángulo de separación entre los dos ejes del plano imagen, siendo la magnitud del error superior a 5° en los dos ángulos que estima si el patrón está muy alejado de la cámara. En

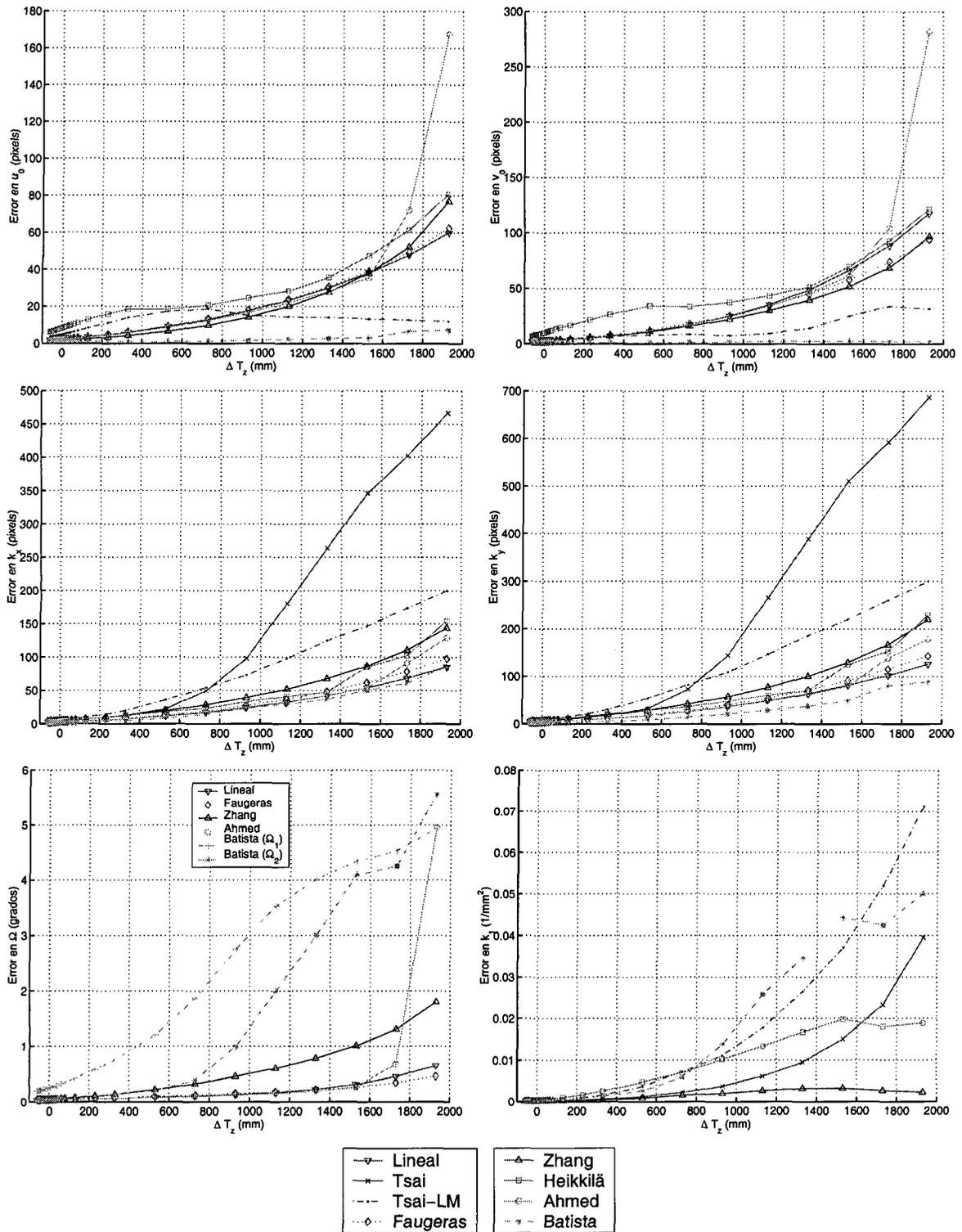


Figura 2.37: Error en el cálculo de los parámetros intrínsecos ante la variación de la distancia entre el patrón y la cámara usando imágenes con ruido de 0.5 pixels en las coordenadas 2D.

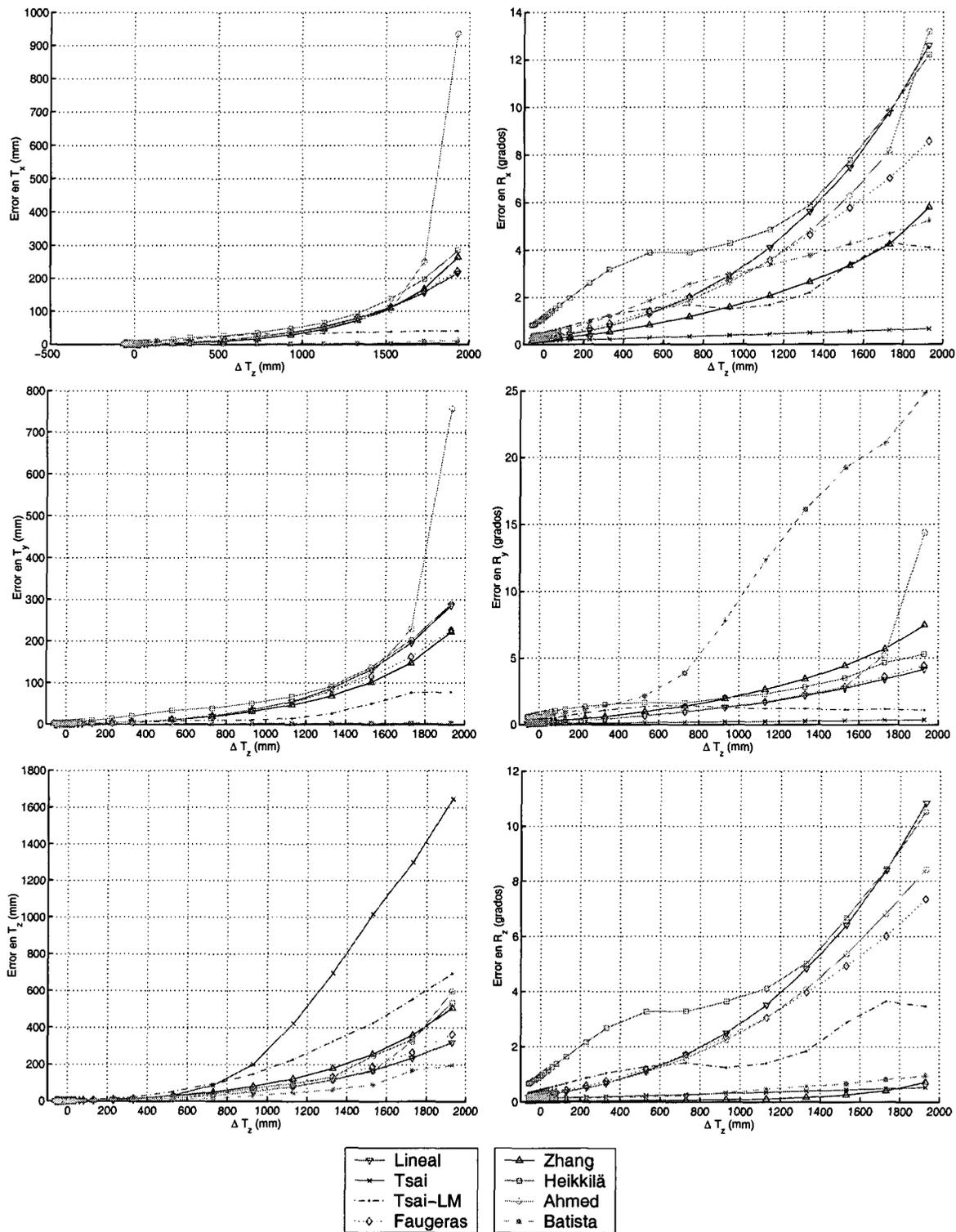


Figura 2.38: Error en el cálculo de los parámetros extrínsecos ante la variación de la distancia entre el patrón y la cámara usando imágenes con ruido de 0.5 pixels en las coordenadas 2D.

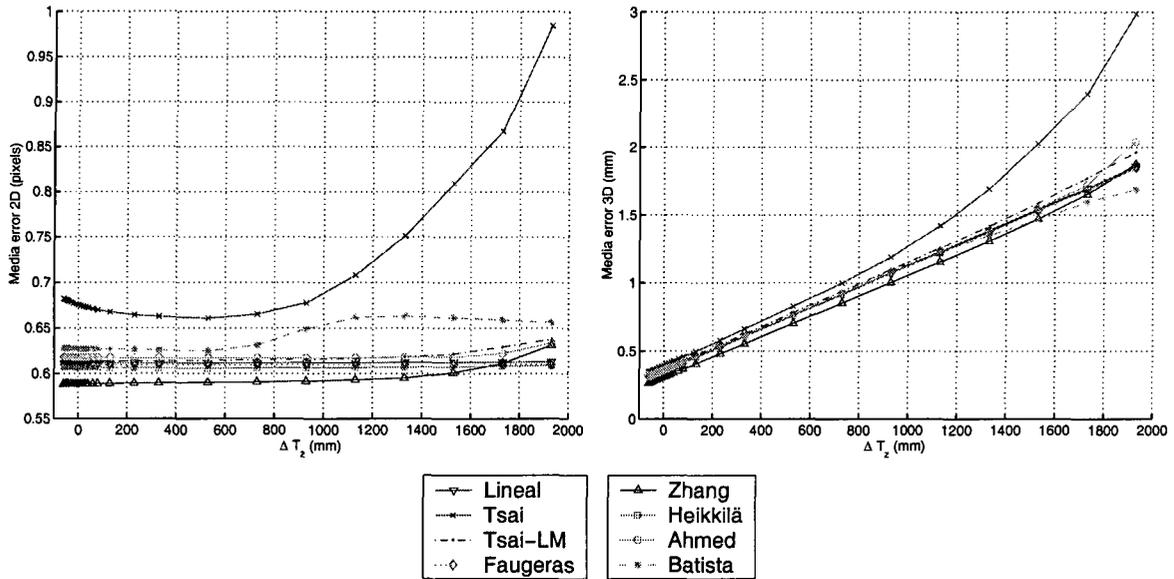


Figura 2.39: Error promedio ante la variación de la distancia entre el patrón y la cámara usando imágenes con ruido de 0.5 pixels en las coordenadas 2D.

todos los métodos se observa que el error aumenta a medida que se aumenta la separación entre la cámara y el patrón. En el caso de Ahmed tiene un buen comportamiento hasta los dos metros, pero a partir de ahí el error aumenta de forma desmedida, lo que coincide con la gráfica del centro óptico.

El resultado obtenido en el cálculo de la posición del patrón coincide con el mostrado en otros experimentos. Existe una relación directa entre las curvas de las gráficas del centro del eje óptico con las de T_x y T_y , y por otro lado, la gráfica de T_z es similar a la de k_y .

En la estimación de la orientación del patrón, el método que parece tener peor comportamiento es el de Batista, en especial en R_y , aunque en R_x también presenta un elevado error. Como se indicó en la introducción de este experimento, el error en estos parámetros depende mucho de la distancia. El motivo se debe a que cuanto más pequeña es la proyección del patrón en la imagen, tanto más pequeña es la relación entre los puntos del mismo, por lo que el ruido gaussiano inyectado afectará en mayor medida. El tercer ángulo (R_z) no se ve afectado debido a que es el único que no depende en su estimación del valor de la distancia focal. El más estable en todos los casos es el método de Tsai debido a que es al que menos le afecta el ruido al tener fijos algunos parámetros (centro del eje óptico), y estar los valores predefinidos de éstos cerca de los reales. En general, todos los métodos tienen un elevado nivel de error en distancias grandes: para una separación superior a los dos metros el error supera los cinco grados.

En las gráficas de la figura 2.39 se muestra, como en otros experimentos, que el error

promedio de todos los métodos es semejante entre sí. La excepción la marca el método de Tsai que presenta un error ligeramente superior al resto, aunque este aumenta de forma paulatina a medida que aumenta la distancia entre el patrón y la cámara. También se observa que el error en 2D es prácticamente constante, mientras que el error en 3D aumenta a medida que el patrón se separa de la cámara.

Como conclusión se puede indicar que en todos los métodos influye la distancia del patrón respecto a la cámara. Según esta distancia aumenta, el efecto del ruido que se ha inyectado sobre las coordenadas 2D de la imagen se incrementa. En general, los métodos son muy inestables (por ejemplo, es posible llegar a un error superior a los 50 pixels en el cálculo del centro del eje óptico, o superior a los 10 cm en la estimación de la posición del patrón, etc.) cuando la distancia es superior a los dos metros. Por tanto, de los resultados obtenidos se puede extraer como conclusión que, en ausencia de distorsión en la formación de las imágenes, la proyección del patrón debe ocupar la mayor porción de imagen posible, o sea, debe estar lo más cerca posible de la cámara.

Variación de la distancia del patrón (T_z) en imágenes con distorsión y ruido gaussiano en las coordenadas 2D

Una vez realizados los dos experimentos anteriores se ha observado que existen discrepancias en las conclusiones. Es decir, en el experimento en el que se variaba la distancia con imágenes distorsionadas, daba en muchos casos como resultado que el error era menor en las distancias más largas. En cambio, en el experimento anterior, en el que se variaba la distancia con imágenes en las que se había inyectado ruido gaussiano, el resultado mostraba que el error era menor cuando el patrón estaba muy cerca. En ambos casos, los resultados son lógicos, ya que la distorsión afecta en mayor medida en los bordes de la imagen, por tanto, cuando el patrón ocupa la mayor parte de la misma, y por otro lado el ruido en las coordenadas 2D afecta en mayor medida cuando los puntos del patrón están muy próximos entre sí.

Debido a estos resultados se ha planteado el siguiente experimento, en el que las imágenes, sobre las que se ha realizado la calibración, incluyen tanto la distorsión ($k_1 = -0,005$) como el ruido gaussiano de media nula y desviación típica 0.5 en las coordenadas 2D. Los patrones se han desplazado de la misma forma que en los experimentos anteriores de esta sección, y la prueba se ha repetido 50 veces en cada una de las distancias de la muestra. Como conclusión de este experimento se espera obtener cuál es la distancia adecuada para colocar el patrón para cada uno de los métodos.

En las figuras 2.40 a 2.42 se muestran los resultados de este experimento. Se puede observar que la parte izquierda de las gráficas (patrón próximo a la cámara) coincide

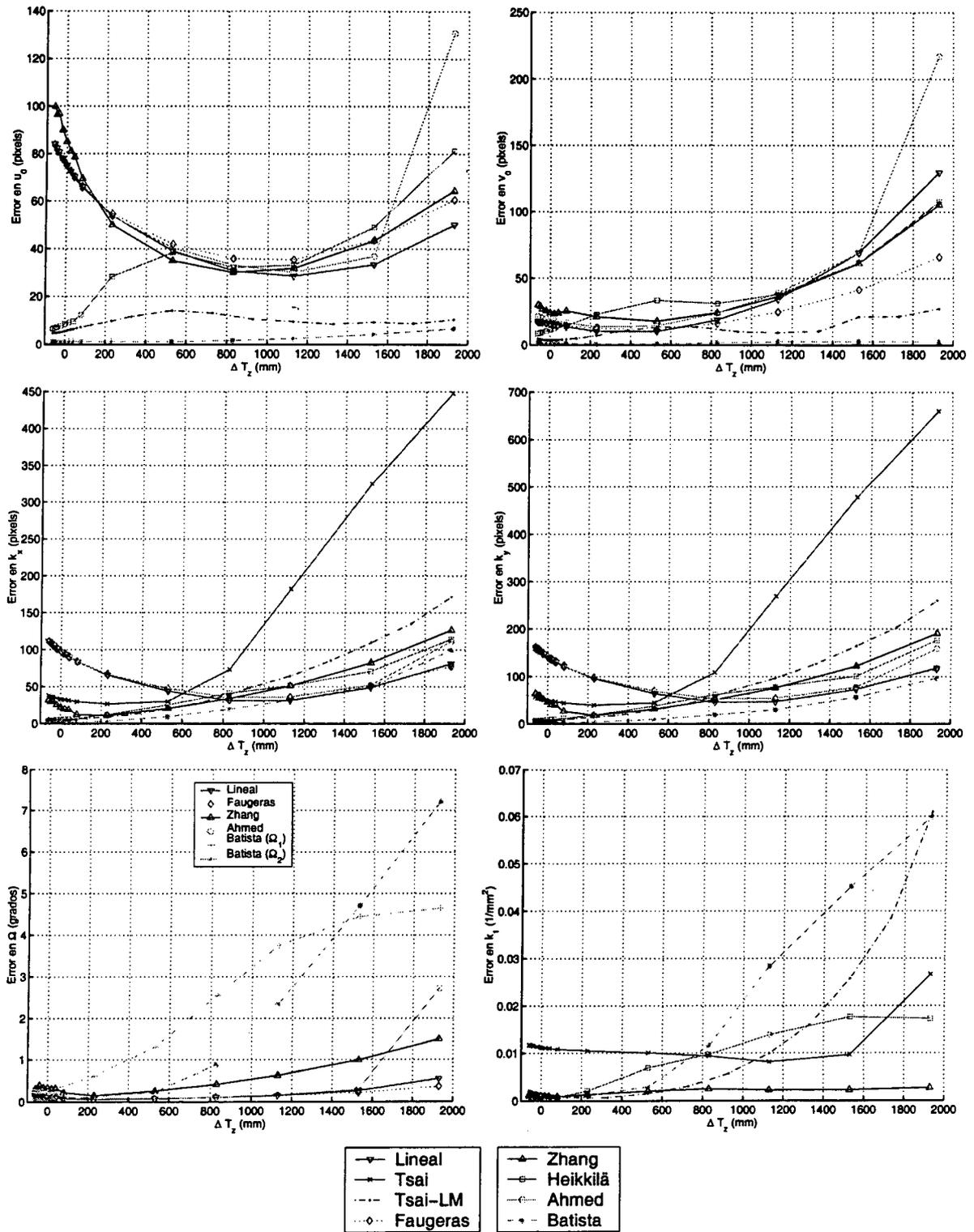


Figura 2.40: Error en el cálculo de los parámetros intrínsecos ante la variación de la distancia entre el patrón y la cámara usando imágenes con distorsión y ruido gaussiano en las coordenadas 2D.

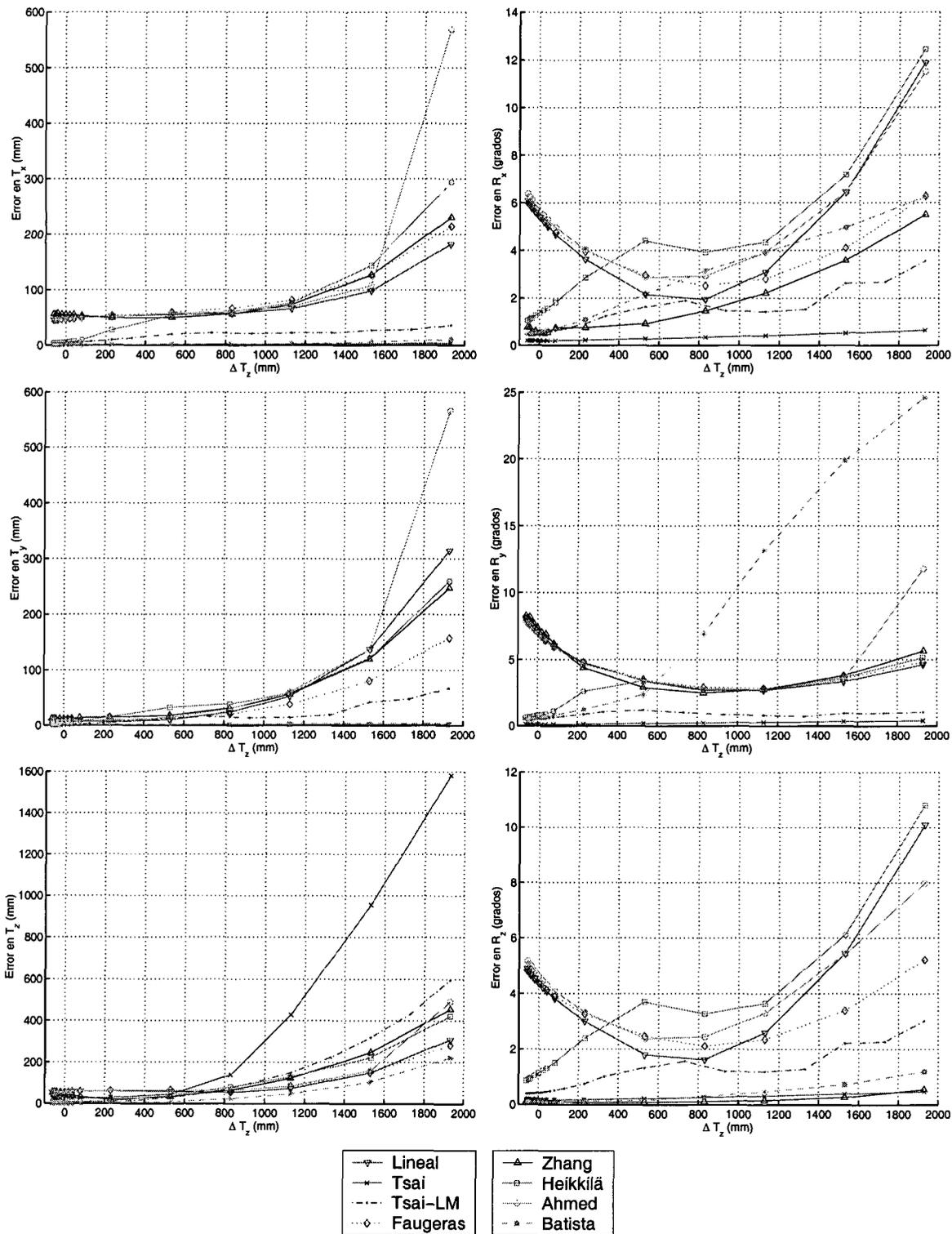


Figura 2.41: Error en el cálculo de los parámetros extrínsecos ante la variación de la distancia entre el patrón y la cámara usando imágenes con ruido de 0.5 píxeles en las coordenadas 2D.

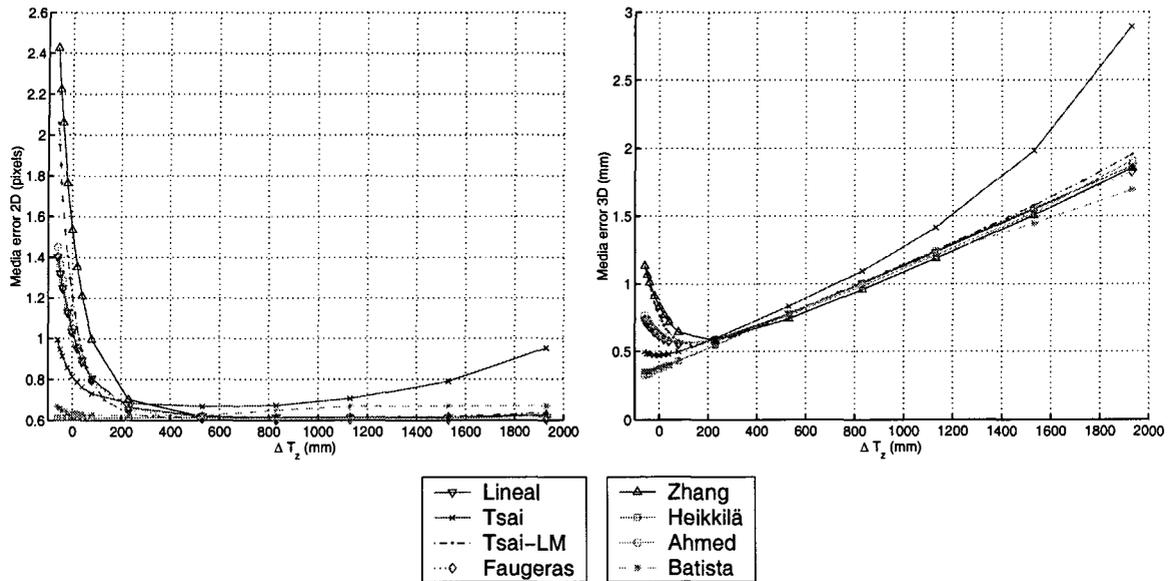


Figura 2.42: Error promedio ante la variación de la distancia entre el patrón y la cámara usando imágenes con ruido de 0.5 pixels en las coordenadas 2D.

con las gráficas del experimento en el que variaba la distancia del patrón haciendo uso de imágenes con distorsión. En cambio, la parte derecha de las gráficas (patrón lejos de la cámara) coincide con las gráficas del experimento anterior (uso de imágenes con ruido gaussiano en las coordenadas 2D). Se demuestra, por tanto, lo que se indicaba previamente, que cuando la proyección del patrón ocupa la mayor parte de la imagen, la distorsión domina y cuando el patrón ocupa una pequeña porción de la misma el factor más importante es el ruido en las coordenadas 2D.

También se puede comprobar que el error producido por la presencia de distorsión en la imagen, en general, es mucho menor que el producido por efecto del ruido en las coordenadas 2D, al menos con el valor del coeficiente de distorsión y el nivel de ruido utilizados en este experimento. Existe alguna excepción a este comportamiento, como en el caso de los métodos de Ahmed, Zhang, Faugeras y Lineal en parámetros como el centro del eje óptico, los factores de escala y la orientación del patrón, donde las curvas tienen forma de 'U' aplanada. Es decir, muestran los mayores índices de error tanto en las distancias más pequeñas como en las mayores y, en cambio, tienen un error mínimo en valores intermedios (entre 1 y 1.5 metros de separación). Esto se debe a que estos métodos son los que peor modelan la distorsión, y, por tanto, producen un error muy alto cuando el patrón ocupa la mayor parte de la imagen. Por tanto, se puede concluir que en aquellos métodos que modelan bien la distorsión el patrón debe abarcar el máximo campo posible en la imagen. En otro caso, es preferible colocar el patrón en distancias intermedias.

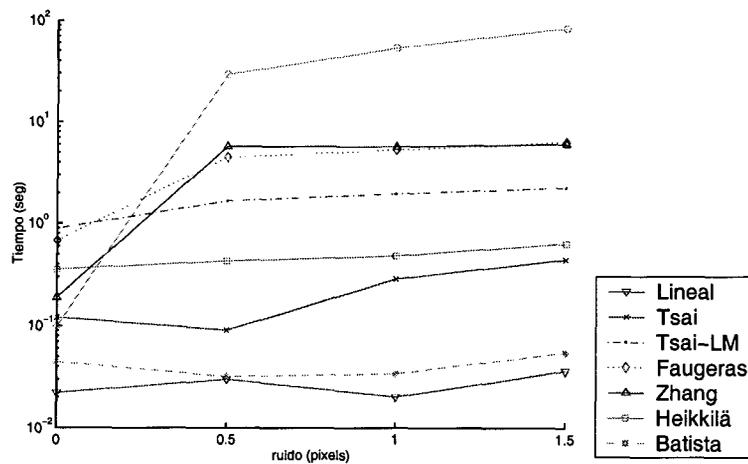


Figura 2.43: Tiempos de ejecución.

2.3.4. Comparación de tiempos de ejecución

En el rendimiento de los métodos también es importante medir su coste computacional. Para ello, se ha medido el tiempo que tarda cada uno de los métodos en realizar el proceso de calibración para distintos niveles de ruido en las coordenadas 2D de la imagen. En estas condiciones se ha medido el promedio que tarda cada método en realizar la computación después de repetir 50 veces el experimento en cada nivel de ruido. El simulador sobre el que se ha realizado el experimento ha sido desarrollado utilizando Matlab (versión 5.3.0.10183 (R11)) y ha sido llevado a cabo sobre un Pentium IV a 2Ghz y con 256Mb de RAM.

Los tiempos resultantes se muestran en la Figura 2.43. Como se puede ver el método más lento es el de Ahmed, cuyo resultado es proporcional al nivel de ruido inyectado; llegando a tardar, cuando este nivel es de 1.5 píxeles, cerca de minuto y medio en realizar la calibración. También son lentos los métodos de Faugeras y Zhang, que rondan los 5 segundos si las imágenes no están libres de ruido, y el método de Tsai-LM que ronda los 2 segundos. Los métodos de Tsai y Heikkilä apenas tardan unas décimas de segundo, y finalmente, los más rápidos son los métodos de Batista y Lineal que rondan las 5 y 3 centésimas de segundo, respectivamente.

Por último, también se puede desprender de la gráfica el hecho de que los tiempos de ejecución dependen de los datos suministrados (la cantidad de ruido presente en los datos) o del conocimiento de la configuración de la cámara (la cercanía de las aproximaciones iniciales introducidas en el método respecto a los valores reales). En general, los métodos aumentan el tiempo de ejecución en función del nivel ruido que tengan los datos, en especial los métodos iterativos que tardarán más en converger a la solución.

2.3.5. Otros resultados

Sobre cada uno de los métodos se han realizado otros experimentos en los que se ha podido observar la influencia de distintos parámetros: distancia focal (f), factor de proporción (s_x), orientación del patrón, traslación del mismo. En todos los casos se observa que los métodos que se ven más afectados por la variación de estos parámetros son los de Tsai y Batista. Como ya se vio en experimentos anteriores el motivo de la sensibilidad de estos métodos a la variación de cualquier parámetro estriba en fijar (o presuponer) de antemano los valores de alguno de los parámetros presentes en el modelo. No se han incluido las gráficas con los resultados por resultar redundantes con los experimentos ya descritos en este trabajo.

2.4. Conclusiones de los experimentos sobre el simulador

A partir de los resultados obtenidos se pueden distinguir perfectamente dos clases de métodos. Por un lado, los que basan sus cálculos en presuponer los valores de algunos parámetros o incluso fijarlos (en muchos casos suelen utilizar datos ofrecidos por el fabricante como valor prefijado o aproximado), y los que resuelven simultáneamente el cálculo de todos los parámetros, pudiendo emplear algún método lineal para obtener una aproximación inicial. Se observa que los incluidos en el primer grupo, son muy estables ante los ruidos inyectados tanto en las coordenadas de la imagen como en las coordenadas de los puntos del patrón. Pero para ello los valores de los parámetros deben ser excelentes aproximaciones a los valores reales, ya que, en caso contrario, los métodos son muy inestables. Los que se incluyen en el segundo grupo tienen en común que son bastante sensibles a los ruidos, pero prácticamente no les influye la variación en la configuración de la cámara, ya que no dependen de aproximaciones iniciales.

A partir del análisis de los resultados de cada uno de los métodos se pueden extraer una serie de conclusiones. Llama la atención que los métodos de Faugeras y de Ahmed, los cuales han usado como aproximación inicial el resultado obtenido por el método Lineal, tienen unos resultados que coinciden bastante con éste último, lo que se explicaría por la existencia de una fuerte dependencia en este tipo de métodos de la aproximación inicial realizada. Por otro lado, en el método de Tsai, el cuál suele utilizarse como punto de comparación cuando se propone un nuevo método, se observa un alto nivel de error para los parámetros estimados en el segundo paso del método (la distancia focal, la distancia del patrón y el coeficiente de distorsión), debido a que, en condiciones no-ideales (afectado por ruido), la distancia focal tiende a valores próximos a cero para

poder minimizar la función objetivo planteada (disparidad de imagen). Cabe destacar también el buen comportamiento que tienen los métodos de Tsai y Batista a la hora de incluir lentes con alta distorsión. En esta situación estos métodos son sólo mejorados por el método de Heikkilä, que, sin embargo, es el que peor responde si existe ruido en las coordenadas tanto de la imagen como del patrón. Esto es debido a que es el método que utiliza un mayor número de parámetros para describir el modelo de cámara.

También es reseñable el hecho de que, en condiciones libres de ruido, los únicos métodos que se ven alterados por la variación del centro del eje óptico son los métodos de Tsai y de Batista, aunque este último en menor medida. A ellos que se une los métodos de Tsai-optimizado y Heikkilä cuando se varía el grado de ortogonalidad del plano imagen. Estos resultados son lógicos si se tiene en cuenta que estos métodos tienen prefijados los valores de estos parámetros en la totalidad o en parte de su algoritmo.

Otra conclusión importante que se puede extraer a la vista de los resultados, es que los métodos son bastante sensibles a la variación de un parámetro que no esté en el modelo que utilizan. Por tanto, habría que decantarse en principio por los métodos que se basen en un modelo de cámara lo más completo posible. Sin embargo, ello implicaría resolver un mayor número de parámetros y que la calibración fuera mucho más sensible a los ruidos en los datos de entrada. Esto se ve reflejado al comparar las dos variantes del método de Tsai que se han utilizado en los experimentos. La versión clásica de este método, la cuál no incluye el centro del eje óptico, es más estable ante la presencia de ruidos que la versión optimizada, que sí lo incluye, como se puede observar en las gráficas del primer experimento.

Por otro lado, también se puede observar en los resultados lo importante que es la colocación del patrón respecto a la cámara, ya que de ello va a depender el que los métodos sean más precisos. Por ejemplo, en la mayoría de los métodos es mejor colocar el patrón lo más cerca de la cámara que ocupe la mayor porción de imagen, y lo más centrado posible. Aunque si la distorsión es muy apreciable, los resultados de todos ellos mejoran si se separa un poco el patrón de la cámara, debido a que el efecto de ésta es mayor en los bordes de la imagen. Obviamente, en caso de que la distorsión fuera importante sería aconsejable decantarse por un método diseñado específicamente para cámaras de focal muy corta como puede ser el de Heikkilä. El cual, por otra parte, no es un buen método si la cámara no sufre de distorsión o ésta es muy pequeña, debido a su sensibilidad al ruido en las coordenadas tanto 2D como 3D. En el caso de que la distorsión no fuese apreciable, el patrón debería estar colocado de tal forma que ocupara la mayor porción de imagen que fuera posible.

También se puede resaltar el hecho de que, en general, hay una coincidencia en la disparidad de imagen (distancia entre las coordenadas reales de imagen y las obtenidas

aplicando el modelo con los resultados de la calibración) en los métodos utilizados. Esto se debe al hecho de que todos ellos buscan minimizar la misma función (el promedio de la distancia euclídea entre las coordenadas reales y las reconstruidas utilizando el modelo), y lo que los diferencia es el modelo de cámara empleado y la forma de resolver la calibración.

En las gráficas resultantes de los distintos experimentos llama la atención la correlación existente en el nivel de error de algunos parámetros. Esto indica que el error que se comete en la estimación de uno de ellos se ve compensado en la misma proporción por otro parámetro. En concreto, se pueden observar correspondencias entre el centro del eje óptico y de la traslación en X e Y del patrón, y también entre los parámetros que incluyen la distancia focal (k_x, k_y) y la distancia a la que se encuentra el patrón (T_z).

A la vista de los resultados obtenidos en este trabajo se puede extraer como conclusión final que no existe un método netamente superior al resto. La preferencia por un método dependerá fundamentalmente de las circunstancias en las que se realice la calibración. Por ejemplo, si las imágenes tienen mucha distorsión, no es aconsejable usar los métodos de Faugeras, Ahmed, Lineal o Zhang; si los datos de entrada (ya sean las coordenadas 3D del patrón como las coordenadas 2D de su imagen) están contaminados con altos niveles de ruido, no es recomendable usar métodos como el de Heikkilä o el de Tsai-optimizado; Por otro lado, en el caso de no conocer con exactitud el valor de ciertos parámetros (centro del eje óptico, distancia focal, etc.), que sirven como aproximación inicial para métodos como el de Tsai o el de Batista, no es aconsejable el uso de ninguno de estos métodos.

Si bien no se puede afirmar que un método posee mayor estabilidad que el resto, sí existen otros criterios que pueden ser analizados y pueden dar indicios de que un método sea mejor que otro. Por un lado, desde el punto de vista del tiempo de ejecución, es claro que tanto el método Lineal como el de Batista superan con creces al resto. Por otro lado, si se analiza el tipo de patrón utilizado, puede que los métodos de Batista y de Zhang sean los más indicados ya que, al utilizar planos, reducen el número de fuentes de ruido en las coordenadas 3D; puesto que, el resto de los métodos necesitan que los puntos estén distribuidos en, al menos, dos planos, generalmente en forma de diedro, por lo que es necesario conocer con exactitud el ángulo que forman los dos planos.

Capítulo 3

Métodos clásicos: experimentos con datos reales

En este capítulo se estudia el comportamiento de los distintos métodos de calibración analizados en el capítulo anterior cuando se aplican sobre una cámara real. Los siguientes experimentos se han planteado de cara a observar, principalmente, cuál de los métodos es más estable. Teóricamente, el método más estable será aquel que, aunque cambien las condiciones externas, posición y orientación del patrón, o el número y tamaño de los puntos del patrón, ofrezca unos resultados similares en todos los casos en los parámetros intrínsecos. Esto se debe cumplir igualmente en el sentido contrario, es decir, que variando exclusivamente la configuración de la cámara, los parámetros extrínsecos deben permanecer constantes. También se analiza la robustez de los métodos evaluando los errores en la reconstrucción de las coordenadas 2D de los puntos en la imagen y las coordenadas 3D en el patrón.

Por otro lado, se estudian los efectos que producen elementos como el zoom, o la colocación del patrón respecto a la cámara, en los resultados de la calibración. En trabajos como [Willson, 1994] se establece que los parámetros que varían al desplazar la lente de la cámara cuando se realiza zoom son:

- La distancia focal. La variación en este parámetro es lógica ya que se produce un distanciamiento o acercamiento entre el plano imagen y el foco.
- El centro del eje óptico. Se producirá si no coincide la orientación del eje mecánico sobre el que se realiza el desplazamiento con el eje óptico.
- El coeficiente de distorsión radial. Es posible que al modificar la distancia focal este coeficiente de distorsión cambie.

En [Willson, 1994] se extraen estas conclusiones habiendo realizado la calibración con el método de Tsai. En los experimentos mostrados en el presente trabajo se trata de extender este análisis a otros métodos, intentando generalizar dichas conclusiones. Una cuestión que se plantea a la vista de los experimentos realizados por Willson es si los parámetros se ven afectados realmente por la variación del zoom, o por contra, la influencia es producto de que la distribución de los puntos del patrón sobre la imagen varía.

En el capítulo anterior, los resultados obtenidos en el simulador mostraban que existe una íntima relación entre algunos de los parámetros. Esta correspondencia aparece entre la componente horizontal del centro del eje óptico (u_0) con la traslación en X del patrón (T_x), o entre v_0 y T_y . También ocurre esto entre los parámetros que incluyen la distancia focal (k_x, k_y) y la distancia a la que se encuentra el patrón (T_z). Para corroborar estos resultados, se han planteado experimentos en los que se pueda estudiar como pequeñas variaciones en los parámetros extrínsecos pueden afectar al resultado final de la calibración.

Para realizar el análisis del funcionamiento de los distintos métodos sobre datos reales se han adquirido varios conjuntos de imágenes en diferentes condiciones. Como ya se ha comentado, los métodos de Batista y Zhang necesitan que los puntos del patrón se distribuyan sobre un plano en cada vista o imagen de calibración mientras que el resto de los métodos (Heikkilä, Faugeras, Lineal, Tsai y Tsai-LM) requieren que los puntos del patrón no se distribuyan sobre un plano, empleándose normalmente un diedro. Por este motivo, en los resultados de los parámetros extrínsecos sólo se comparan los métodos que usan el mismo patrón, es decir, los que usan un diedro.

En el estudio sobre datos reales no se ha incluido el método de Ahmed debido a que los resultados en el simulador coinciden con los obtenidos por los métodos de Faugeras y Lineal. Además, se muestra como un método excesivamente lento.

Este capítulo se organiza en varias secciones. En primer lugar, se describen las muestras de datos utilizadas para el estudio y se presentan los resultados de su calibración por cada uno de los métodos. A continuación se muestra el análisis realizado sobre los resultados obtenidos de la calibración. Este análisis se realiza desde varios puntos de vista. Primeramente, se estudia la precisión de cada método a partir del error en la reconstrucción de los puntos. A continuación se realiza un análisis de la estabilidad de cada método, tanto en los parámetros intrínsecos como en los extrínsecos. También se estudia la incidencia de elementos propios de la configuración del patrón y de la cámara, los cuales se estudian en sucesivas secciones. Se estudia también cómo influye el tamaño del área que ocupan los puntos del patrón en la imagen. Finalmente, se establecen una serie de conclusiones sobre los experimentos de calibración a la luz de los resultados

obtenidos en este capítulo.

El capítulo incluye una comparación detallada entre los resultados obtenidos sobre la cámara real y los obtenidos sobre el simulador. En este estudio se han repetido algunos de los experimentos desarrollados en este capítulo sobre el simulador, intentando reproducir los resultados imponiendo en el simulador unas condiciones lo más parecidas posibles a las vigentes sobre la cámara real. Por último, los resultados de este estudio son comparados con los obtenidos por otros autores en trabajos similares, y se presentan las conclusiones sobre métodos de calibración que usan un patrón como referente.

3.1. Descripción de la toma de datos y resultados de la calibración

Para la realización de las pruebas con datos reales se han utilizado imágenes obtenidas con una cámara Sony Evi-G21. Esta cámara usa un sensor de imagen del tipo 1/4" Hiper HAD Color CCD, siendo el tamaño del CCD para este tipo de sensores de 3.2 mm x 2.4 mm. Las imágenes formadas son de 768 x 576 pixels. Estas cámaras disponen de zoom motorizado pudiendo variar la distancia focal entre 4.5 y 13.5 mm, según el rango definido por el fabricante, y también permiten movimientos de azimut y elevación (hasta $\pm 30^\circ$ en azimut y $\pm 15^\circ$ en elevación).

Los patrones usados como referente en los experimentos están formados por círculos, ya que, como se indica en [Min et al., 2001], este tipo de patrón es preferible frente al formado por cuadrados. Para conseguir precisión subpixel se obtiene el centroide del círculo en la imagen. Hay que indicar que todos los puntos del patrón utilizado en cada caso caen dentro del campo de visión de la cámara en todas las imágenes que se han capturado para la realización de las distintas pruebas. De tal forma que, cuando la distancia focal se desplaza a su máximo valor, el patrón ocupa prácticamente todo el plano imagen.

En las gráficas resultantes de cada una de las pruebas se muestran los valores de los parámetros intrínsecos y extrínsecos obtenidos al calibrar cada conjunto de datos con los distintos métodos analizados. También se introducen diversas medidas de error basadas en la distancia del punto real respecto al reconstruido tanto en 2D como en 3D. En cada caso se muestra la distancia promedio y su desviación típica, así como la medida de error RMS.

Por último, es importante comentar que el método de Zhang modela la distorsión de forma diferente al resto de los métodos que la calibran. Por tanto, se ha optado, al igual que en los experimentos del simulador, por obtener un parámetro equivalente al

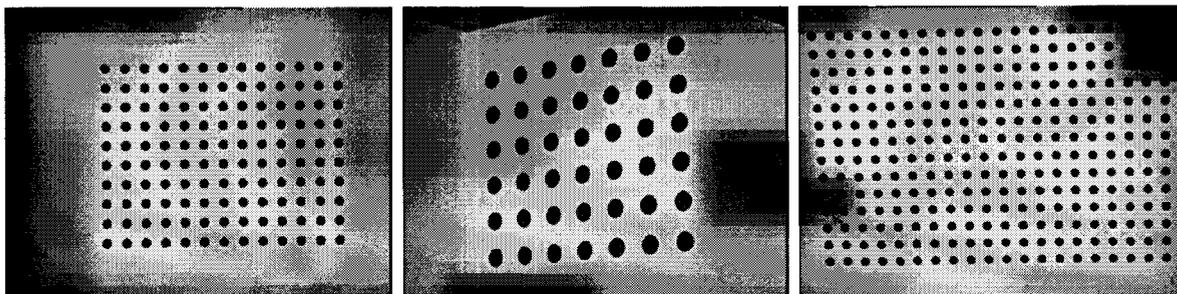


Figura 3.1: a) Patrón monopiano de 13x10 puntos. b) Patrón monopiano de 7x6 puntos. c) Patrón monopiano de 20x14 puntos.

del resto de los métodos para hacer sus resultados comparables a los de éstos. Para ello, se ha dividido el valor del coeficiente de distorsión por el cuadrado de la distancia focal para su representación en las gráficas.

Los mayoría de los estudios presentados en este capítulo han sido realizados a partir de la calibración llevada a cabo sobre dos conjuntos de imágenes. En ambos conjuntos se ha realizado la calibración de la cámara con diferentes grados de zoom y manteniendo el enfoque fijo. A continuación, se describen las condiciones en las que se han generado estos conjuntos de imágenes y se etiquetan los conjuntos para poderlos referenciar de aquí en adelante:

Conjunto 8x8. En este conjunto de imágenes, se ha empleado la misma configuración de patrón que la utilizada en el simulador. En esta configuración se trata de que todos los métodos procesen aproximadamente el mismo número de puntos. En concreto, para el método de Batista se ha utilizado un plano de 13x10 puntos, estando sus centros separados entre sí 1.75 cm (figura 3.1a). El patrón se colocó a menos de un metro de la cámara y aproximadamente paralelo al plano imagen. Para el método de Zhang se ha utilizado un plano de 7x6 puntos separados entre si 3 cm (figura 3.1b). Sobre este patrón se han capturado tres vistas con tres orientaciones diferentes en cada nivel de zoom aplicado, tal y como especifica el método. Para realizar el experimento con el resto de los métodos se ha utilizado un patrón de 8x8 puntos en cada plano de un diedro, siendo los puntos de un centímetro de diámetro y estando éstos separados entre sí dos centímetros (figura 3.2a). Los resultados de calibrar las imágenes tomadas de cada patrón con sus respectivos métodos se muestran en las figuras 3.3 - 3.6.

Conjunto 14x20. En este conjunto de imágenes, se han empleado unos patrones con mayor cantidad de puntos, del orden de cuatro veces más. En esta configuración se ha buscado que todos los patrones utilicen puntos con las mismas dimensiones

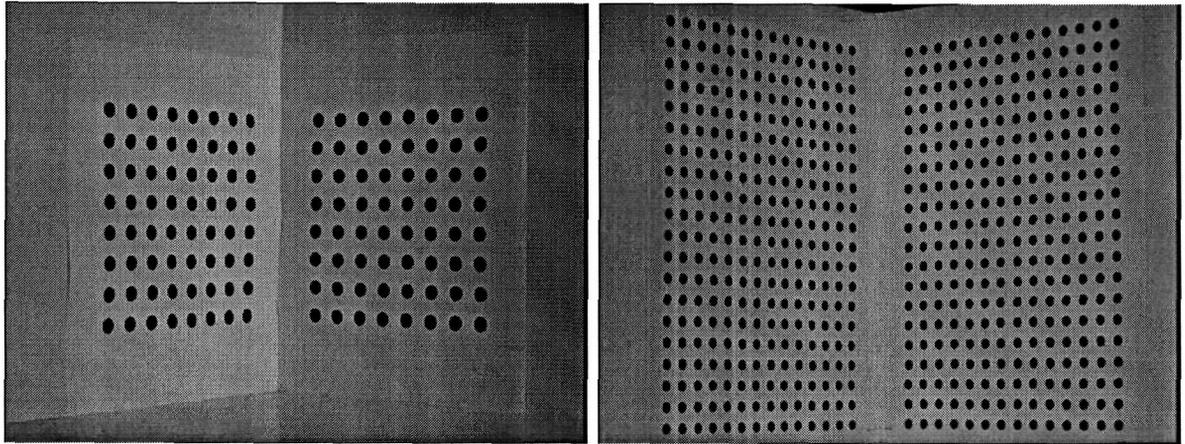


Figura 3.2: a) Patrón biplano de 8x8 puntos en cada plano. b) Patrón biplano de 14x20 puntos en cada plano.

y que la separación entre ellos sea la misma. En este caso los patrones usados con los métodos de Batista y Zhang sólo tienen la mitad de puntos; de hecho, es uno de los planos del diedro que se utiliza para el resto de los métodos. Por tanto, para este segundo conjunto, el método de Batista procesa un 50% menos de puntos por calibración que en el primer conjunto, mientras que el método de Zhang debe procesar un 50% más de puntos por calibración debido a que necesita tres vistas. En concreto, para el método de Batista y Zhang se ha utilizado un plano de 20x14 puntos (figura 3.1c). En la primera vista, que es coincidente para los dos métodos, el patrón se encuentra paralelo al plano imagen. En las otras dos vistas, que sólo utiliza el método de Zhang, el patrón ha rotado respecto al eje Y, y al eje X de la cámara, respectivamente. Para realizar el experimento con el resto de los métodos se ha utilizado un patrón de 14x20 puntos en cada plano de un diedro (figura 3.2b). En todos los casos los puntos tienen un tamaño de 6.5 mm y la distancia que separa sus centros es de 13 mm. Los resultados de calibrar las imágenes tomadas de cada patrón con sus respectivos métodos se muestran en las figuras 3.7 - 3.10.

Tanto en la figura 3.6 como la 3.10 muestran los resultados de la calibración empleando los Conjuntos 8x8 y 14x20, pero donde no se han incluido los resultados de los métodos de Heikkilä y Tsai-LM. Éste último sólo ha sido eliminado en la segunda figura. Esta eliminación se ha realizado debido a que estos métodos son los que producen mayor dispersión en los resultados y por tanto, no permiten observar con mayor detalle la evolución del resto de los métodos.

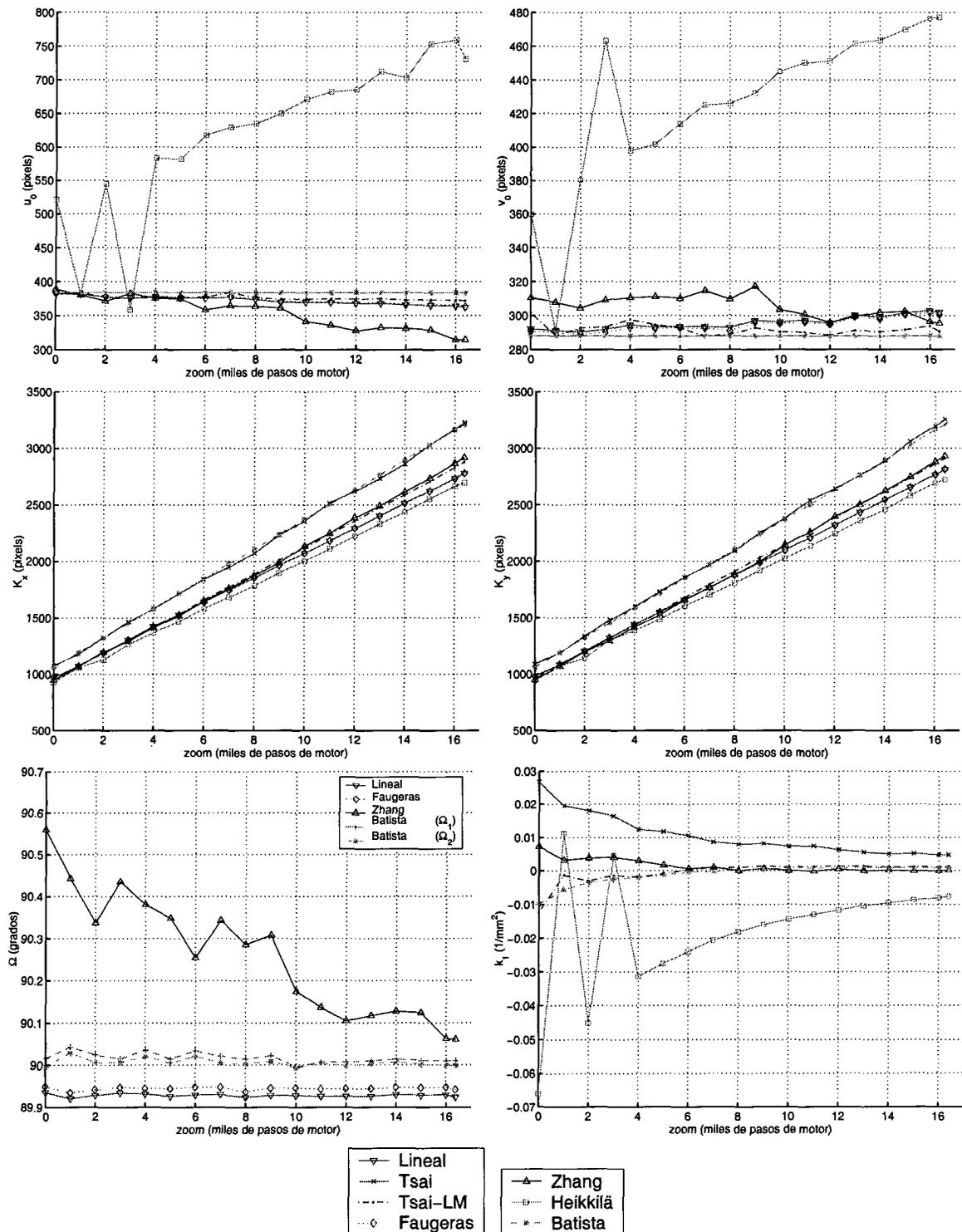


Figura 3.3: Resultados de la calibración del Conjunto 8x8. Valor de los parámetros intrínsecos para cada nivel de zoom.

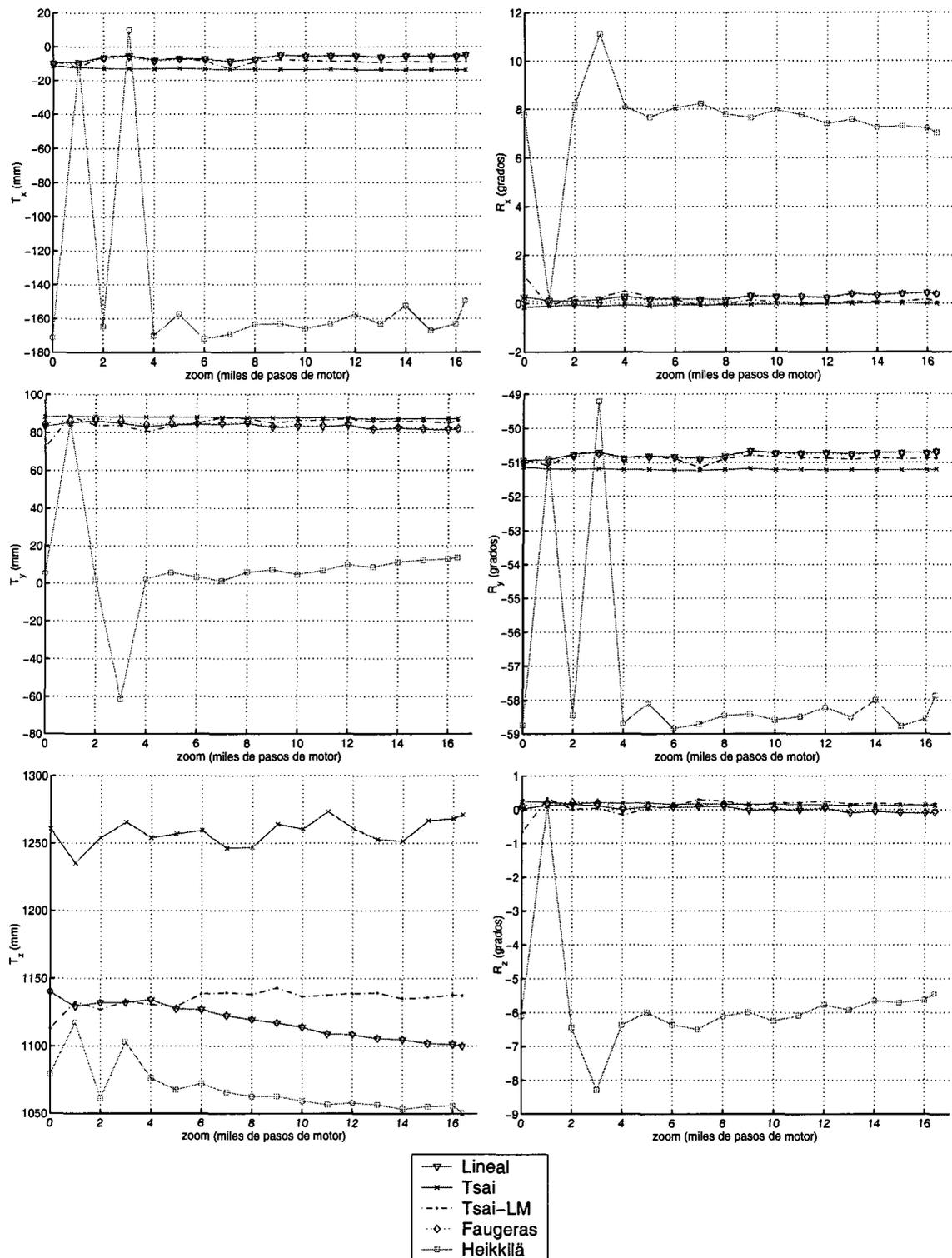


Figura 3.4: Resultados de la calibración del Conjunto 8x8. Valor de los parámetros extrínsecos para cada nivel de zoom.

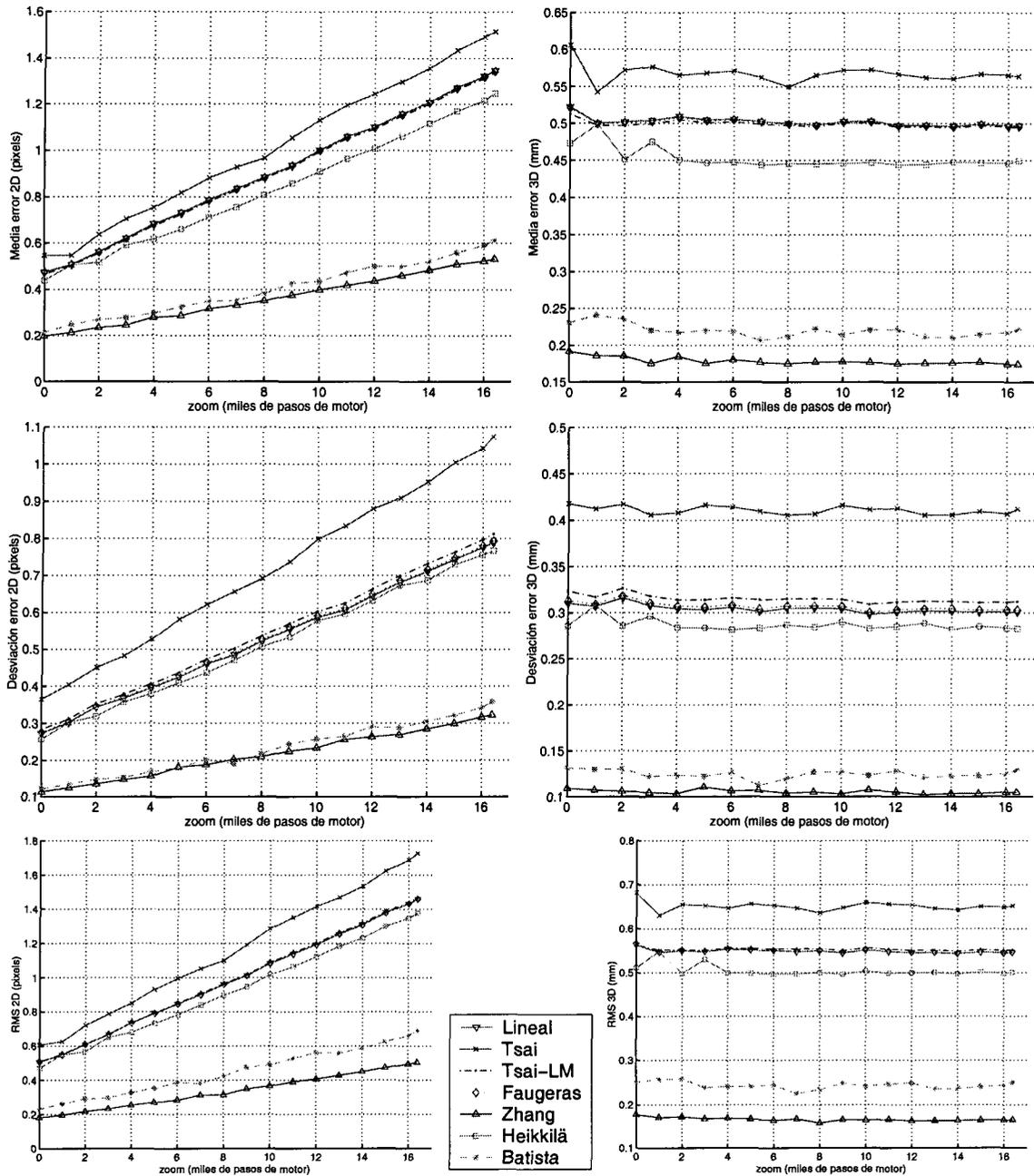


Figura 3.5: Resultados de la calibración del Conjunto 8x8. Error 2D (izquierda) y Error 3D (derecha) para cada nivel de zoom.

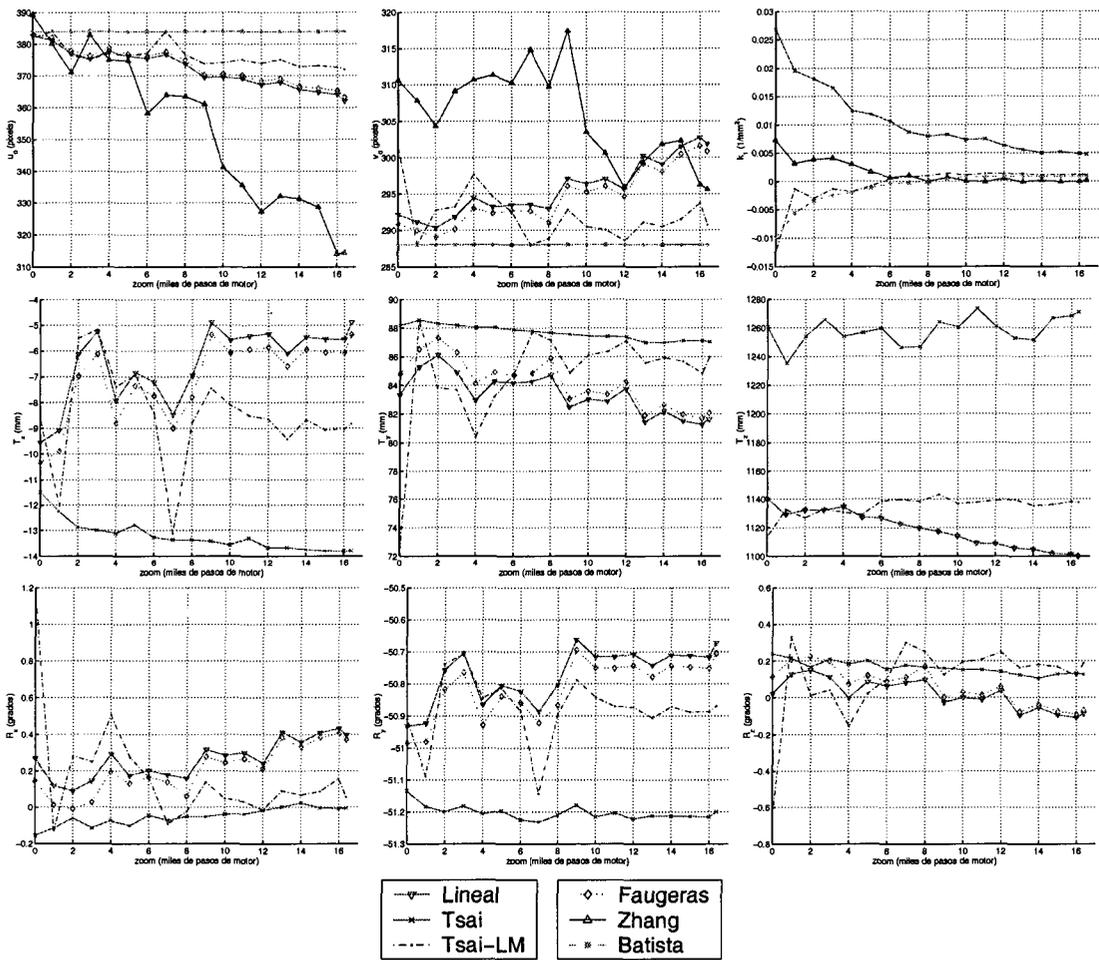


Figura 3.6: Algunos resultados de la calibración del Conjunto 8x8. Repetición de algunas gráficas de las figuras 3.3 y 3.4, pero excluyendo el método de Heikkilä.

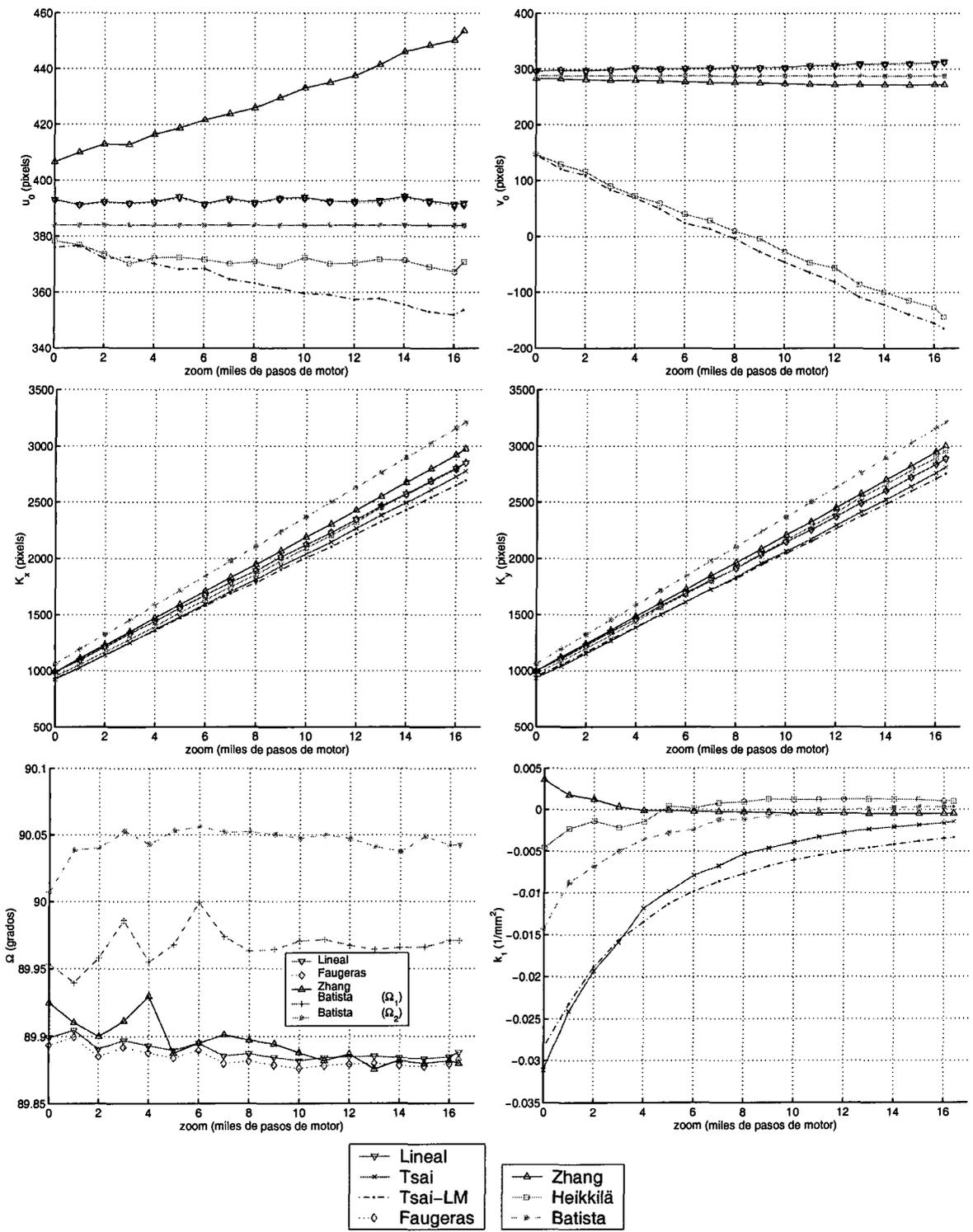


Figura 3.7: Resultados de la calibración del Conjunto 14x20. Valor de los parámetros intrínsecos para cada nivel de zoom.

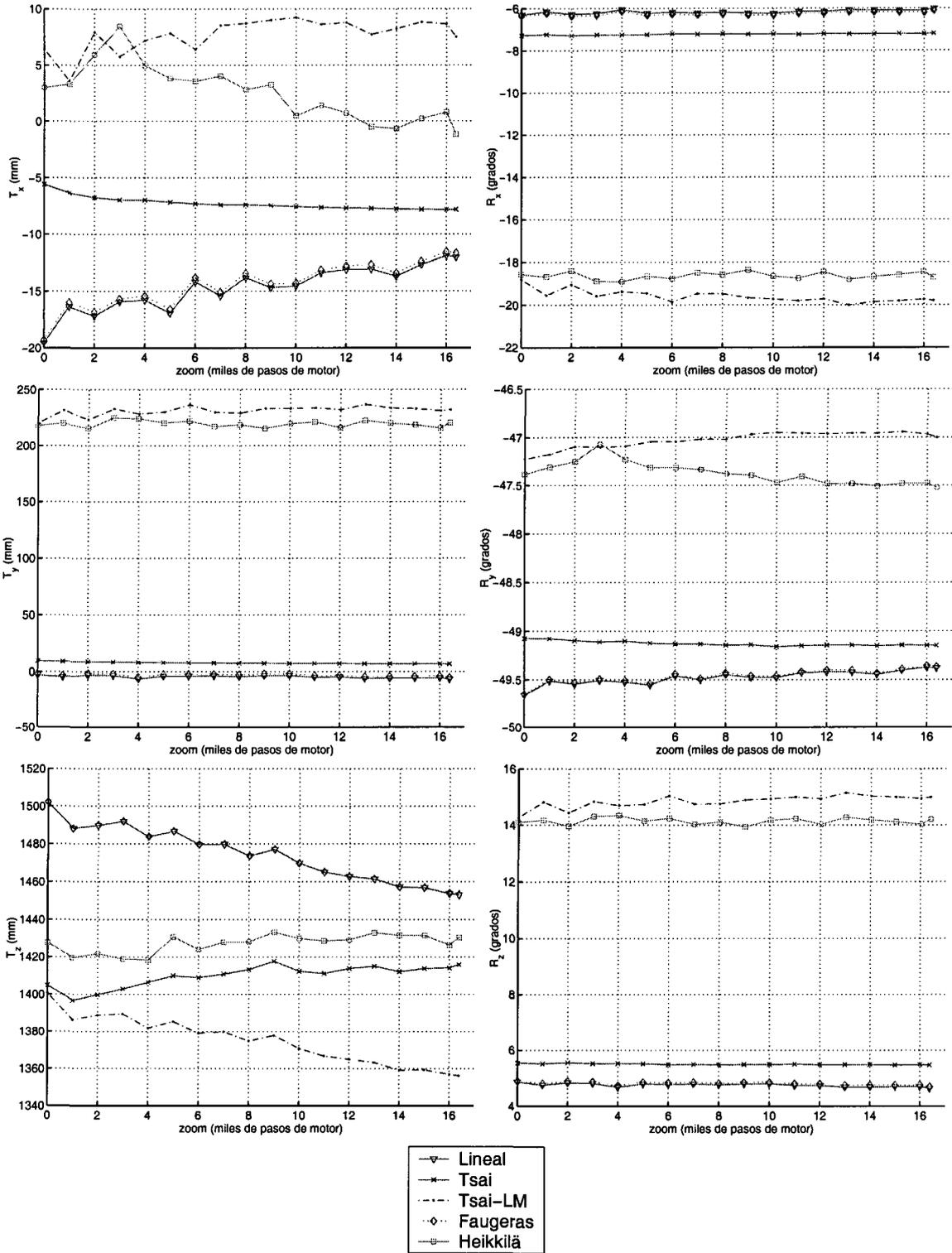


Figura 3.8: Resultados de la calibración del Conjunto 14x20. Valor de los parámetros extrínsecos para cada nivel de zoom.

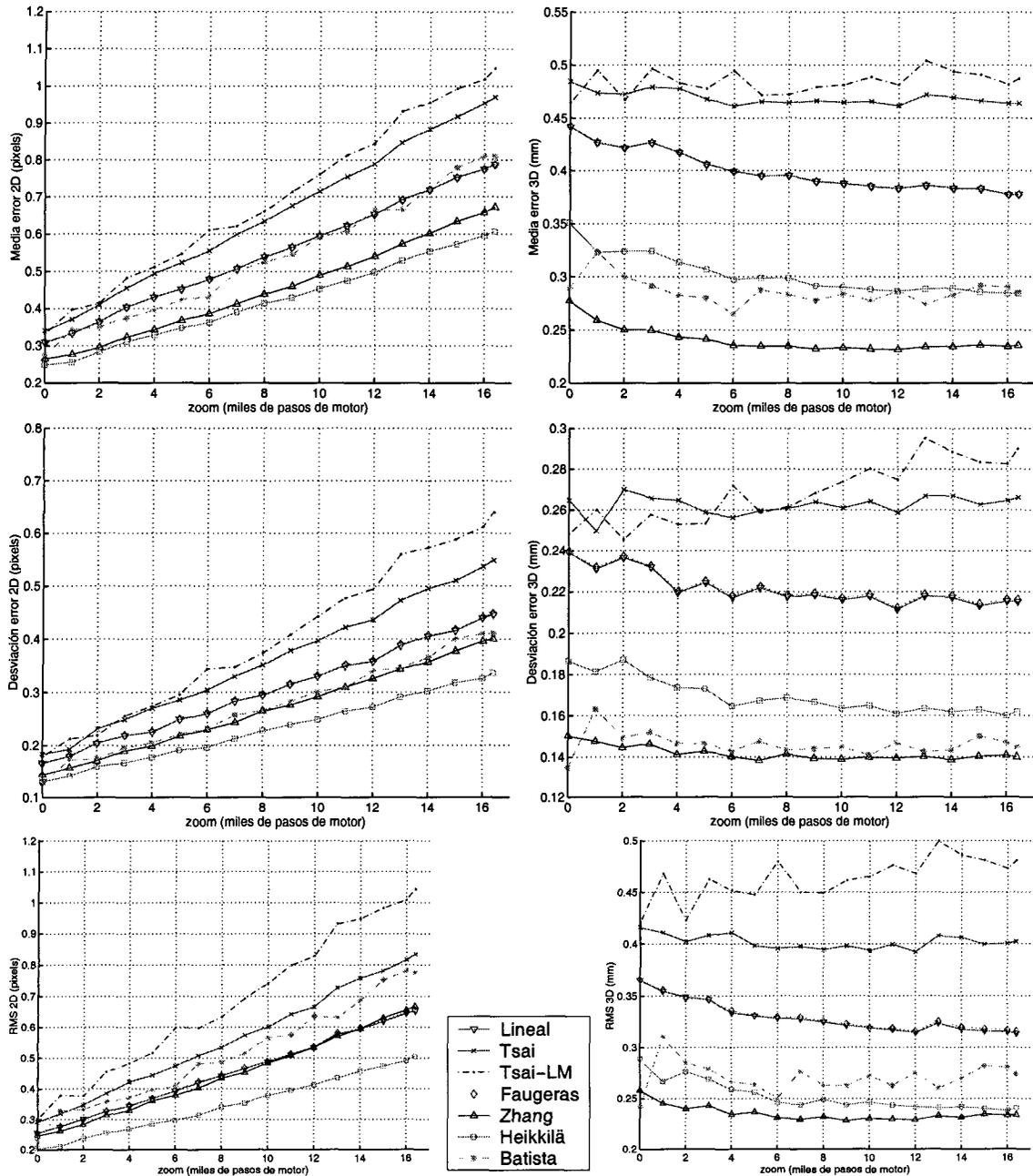


Figura 3.9: Resultados de la calibración del Conjunto 14x20. Error 2D (izquierda) y Error 3D (derecha) para cada nivel de zoom.

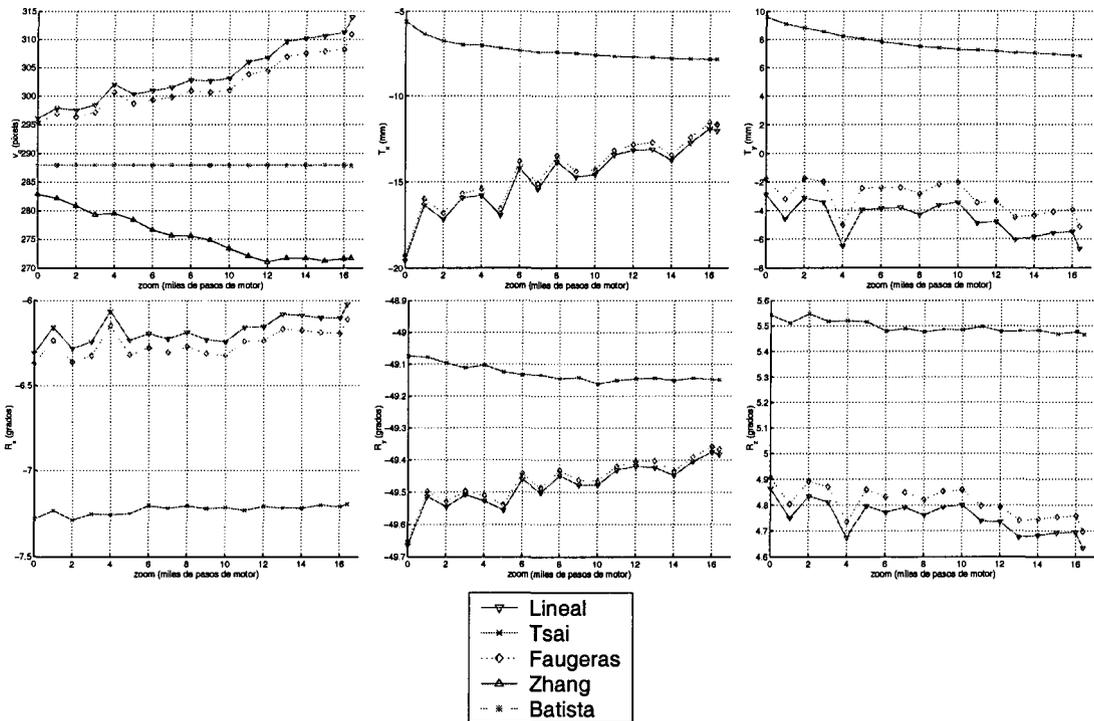


Figura 3.10: Algunos resultados de la calibración del Conjunto 14x20. Repetición de algunas gráficas de las figuras 3.7 y 3.8, pero excluyendo los métodos de Heikkilä y Tsai-LM.

3.2. Precisión de los métodos

En primer lugar, se analiza la precisión de cada uno de los métodos y se comparan sus resultados. Debido a que no se dispone de los datos reales de cada uno de los parámetros, a diferencia de lo que ocurría en los experimentos del simulador, no es posible conocer la precisión con que se recupera cada uno de los parámetros del modelo, sino sólo el ajuste global. Por tanto, la única forma de realizar este análisis es estudiando el error que se produce en la reconstrucción de las coordenadas de los puntos tanto en 2D (en la imagen) como en 3D (en el patrón).

Como se observó en el simulador, el problema que se presenta es que, en muchos casos, los errores de un parámetro se ven compensados con los errores de otro parámetro. Esto indica que se puede haber reconstruido perfectamente los puntos, pero el modelo obtenido no se ajusta a la realidad. Esto se debe a que los métodos de calibración intentan ajustar los valores de los distintos parámetros del modelo a un determinado conjunto de datos de entrada. Por tanto, la medida de error utilizada no reflejará al completo la bondad del método, sólo lo bien que se reconstruye ese conjunto de datos utilizado. Por eso, es probable que al repetirse el proceso de calibración en circunstancias en las que no cambie la configuración de la cámara, pero empleando otro conjunto de

puntos como patrón o colocando el patrón en otra posición, los parámetros intrínsecos obtenidos varían de resultado.

En primer lugar se analizan y comparan los errores cometidos por cada uno de los métodos en la reconstrucción de las coordenadas de los puntos tanto en 2D como en 3D. En la figura 3.5 se muestran las curvas de estos errores para el primer conjunto (Conjunto 8x8), y en la figura 3.9 para el segundo (Conjunto 14x20).

En general, se observa en gráficas que, tanto en el error de reconstrucción de los puntos en 2D (disparidad en la imagen) como en 3D (distancia sobre la posición real de los puntos sobre el patrón), que todos los métodos que usan el mismo patrón de referencia tienen resultados semejantes.

En el primer conjunto de datos (Conjunto 8x8) los métodos que utilizan patrones planos (Batista y Zhang) aparecen como los que menor error producen en la reconstrucción de los puntos, siendo éste aproximadamente la mitad de los anteriores. Esto puede ser debido a que el patrón utilizado elimina algunas de las fuentes de ruido que presentan los patrones en forma de diedro, esto es, la determinación precisa del ángulo que forman los dos planos del mismo o del desplazamiento respecto al origen de coordenadas común. El resto de los métodos tienen resultados similares. Dentro de esta similitud de resultados el error del método de Tsai es ligeramente superior al alcanzado por los otros, mientras que el método de Heikkilä presenta un error ligeramente inferior.

En la segunda muestra (Conjunto 14x20), los métodos basados en planos aumentan ligeramente la magnitud del error, en especial, en el error en la reconstrucción en las coordenadas 2D, mientras que los otros métodos disminuyen significativamente el error.

Al comparar estos resultados con los obtenidos en el simulador, se puede asegurar que la cámara parece no producir una distorsión elevada. Se llega a esta conclusión porque el método de Zhang, que, como ya se demostró, es muy sensible a valores altos del coeficiente de distorsión, se muestra, en cambio, como el método con menor error.

Sin embargo, el hecho de que el método de Batista presente muy buenos resultados no es sintomático, ya que es un método que ajusta muy bien los parámetros para que el error sea mínimo en todas las circunstancias. Esto no significa que calcule bien los distintos parámetros, sino que compensa adecuadamente los valores de estos para que el error de reconstrucción sea mínimo. En los resultados del simulador se observa que el punto débil de este método aparece cuando la configuración de la cámara dista de la aproximación inicial utilizada. Habitualmente se utiliza que el centro óptico coincida con el centro de la imagen, que el factor de proporción sea uno y que la ortogonalidad de los ejes del plano imagen sea perfecta. Pero aún cuando la configuración de la cámara esté alejada de estas condiciones iniciales, y los parámetros resultantes de la calibración disten de los reales, el error en la reconstrucción de los puntos apenas se ve afectado.

3.3. Estabilidad de los métodos

En teoría, los parámetros intrínsecos de la cámara deben permanecer constantes ante cualquier variación en la disposición del patrón en relación a la cámara. De igual forma, los parámetros extrínsecos de la cámara deben permanecer constantes ante cualquier variación en la configuración de la cámara. A partir de estas consideraciones es posible realizar un análisis de la estabilidad de cada método. Para llevar a cabo este análisis, se ha comprobado el grado de dispersión que ofrecen los resultados al variar sólo uno de los tipos de factores que inciden en los resultados. Es decir, al variar la disposición del patrón en relación a la cámara se ha medido la dispersión en los parámetros intrínsecos y al variar la configuración de la cámara se medido la dispersión en los parámetros extrínsecos.

3.3.1. Influencia de la configuración de la cámara en los parámetros extrínsecos

En este apartado se analiza la estabilidad de los parámetros extrínsecos ante variaciones en la configuración de la cámara. Para realizar este análisis se han utilizado como referencia los resultados de los parámetros extrínsecos (figuras 3.4 y 3.8) obtenidos de la calibración de los dos conjuntos de imágenes descritos previamente (Conjunto 8x8 y Conjunto 14x20). Se han utilizado estos conjuntos porque cada uno de ellos muestra un desplazamiento del zoom de la cámara, mientras permanecía fija la colocación del patrón respecto a la cámara.

En las tablas 3.1 y 3.2 se muestra los rangos de variación en la estimación de cada parámetro para cada uno de los métodos. De los resultados parece desprenderse que los parámetros extrínsecos son, en general, bastante inestables, destacando principalmente la distancia del patrón (T_z). Esto parece lógico ya que se está modificando la distancia focal.

	T_x (mm)	T_y (mm)	T_z (mm)	R_x (°)	R_y (°)	R_z (°)
Lineal	1.5/4.7	1.4/4.9	13/40	0.11/0.34	0.08/0.27	0.08/0.26
Tsai	0.6/2.3	0.5/1.6	9.6/38.1	0.05/0.18	0.02/0.10	0.04/0.13
Tsai-LM	1.9/7.9	3.6/16.3	6.7/29.9	0.29/1.29	0.10/0.44	0.22/1.00
Faugeras	1.5/5.0	1.7/5.6	13/40	0.13/0.42	0.09/0.29	0.10/0.32
Zhang	8.3/27.6	5.6/18.6	11.6/37.6	0.22/0.70	0.22/0.73	0.14/0.47
Batista	0.5/2.0	0.5/1.5	7.4/28.0	0.08/0.31	0.09/0.33	0.01/0.03
Heikkilä	53.3/182.0	25.4/146.3	17.6/66.7	2.03/10.98	2.75/9.63	1.60/8.42

Tabla 3.1: Desviación típica y rango máximo de variación (Desviación/Máximo) de los parámetros extrínsecos obtenidos de la calibración del Conjunto 8x8 por cada uno de los métodos.

	T_x (mm)	T_y (mm)	T_z (mm)	R_x (°)	R_y (°)	R_z (°)
Lineal	2.0/7.7	1.2/3.8	14.7/48.9	0.08/0.29	0.07/0.29	0.06/0.23
Tsai	0.6/2.2	0.8/2.8	5.7/20.8	0.03/0.09	0.03/0.09	0.02/0.08
Tsai-LM	1.4/5.6	4.1/16.6	13.04/44.5	0.30/1.20	0.08/0.28	0.21/0.88
Faugeras	2.0/7.7	1.1/3.4	14.7/48.8	0.08/0.26	0.07/0.29	0.06/0.21
Zhang	0.8/2.9	0.6/2.0	2.8/12.7	0.06/0.21	0.06/0.21	0.01/0.04
Batista	0.5/1.8	0.5/1.7	6.4/24.8	0.10/0.38	0.19/0.90	0.01/0.04
Heikkilä	2.5/9.6	2.9/9.5	4.8/14.9	0.16/0.54	0.12/0.46	0.12/0.40

Tabla 3.2: Desviación típica y rango máximo de variación (Desviación/Máximo) de los parámetros extrínsecos obtenidos de la calibración del Conjunto 14x20 por cada uno de los métodos.

Llama la atención el resultado del método de Heikkilä al calibrar el Conjunto 8x8, debido a que presenta una alta inestabilidad cuando la distancia focal es corta (menos de 4000 pasos del motor del zoom). También resalta el hecho de que tanto el método de Heikkilä como el de Zhang registren una ostensible mejora de los resultados al calibrar el Conjunto 14x20, mejorando en aproximadamente un 90 % su estabilidad respecto al otro conjunto de datos, en el cual presentan una alta dispersión en los resultados.

Parece que los métodos que presentan una mayor estabilidad son los métodos de Batista y Tsai. Esto es lógico, ya que en el método de Batista, las estimaciones que se obtienen de los parámetros intrínsecos son muy próximas a las aproximaciones iniciales aportadas; y en el caso del método de Tsai, el hecho de mantener el centro óptico en el centro de la imagen, hace que el resultado de casi todos los parámetros sea muy estable.

3.3.2. Desplazamiento horizontal del patrón

En este apartado se analiza la estabilidad de los parámetros intrínsecos ante variaciones en la disposición del patrón en relación a la cámara. Para realizar este análisis se ha realizado un experimento en el que el patrón se ha desplazado horizontalmente frente a la cámara, mientras que la configuración de la cámara ha permanecido fija. En este experimento se trata de comprobar cómo la posición del patrón puede influir sobre los resultados de los parámetros intrínsecos, en concreto sobre el centro del eje óptico.

Para realizar esta prueba se ha utilizado el patrón de 14x20 puntos en dos planos que forman 90° entre sí (figura 3.2b). Se han tomado 5 imágenes en las que se ha desplazando el patrón en horizontal, tanto hacia la derecha como a la izquierda siendo, los desplazamientos de 5 y 10 cm. hacia cada lado.

En general, se puede observar que en todos los parámetros intrínsecos (figura 3.11) existe una amplia dispersión de los resultados al variar la disposición del patrón. En algunos parámetros esta dispersión llega a ser desmesurada, como ocurre en el método de Zhang, a la hora de calcular u_0 y los factores de escalado. Esto mismo ocurre en la estimación de k_x y Ω_1 por parte del método de Batista debido a que el método no

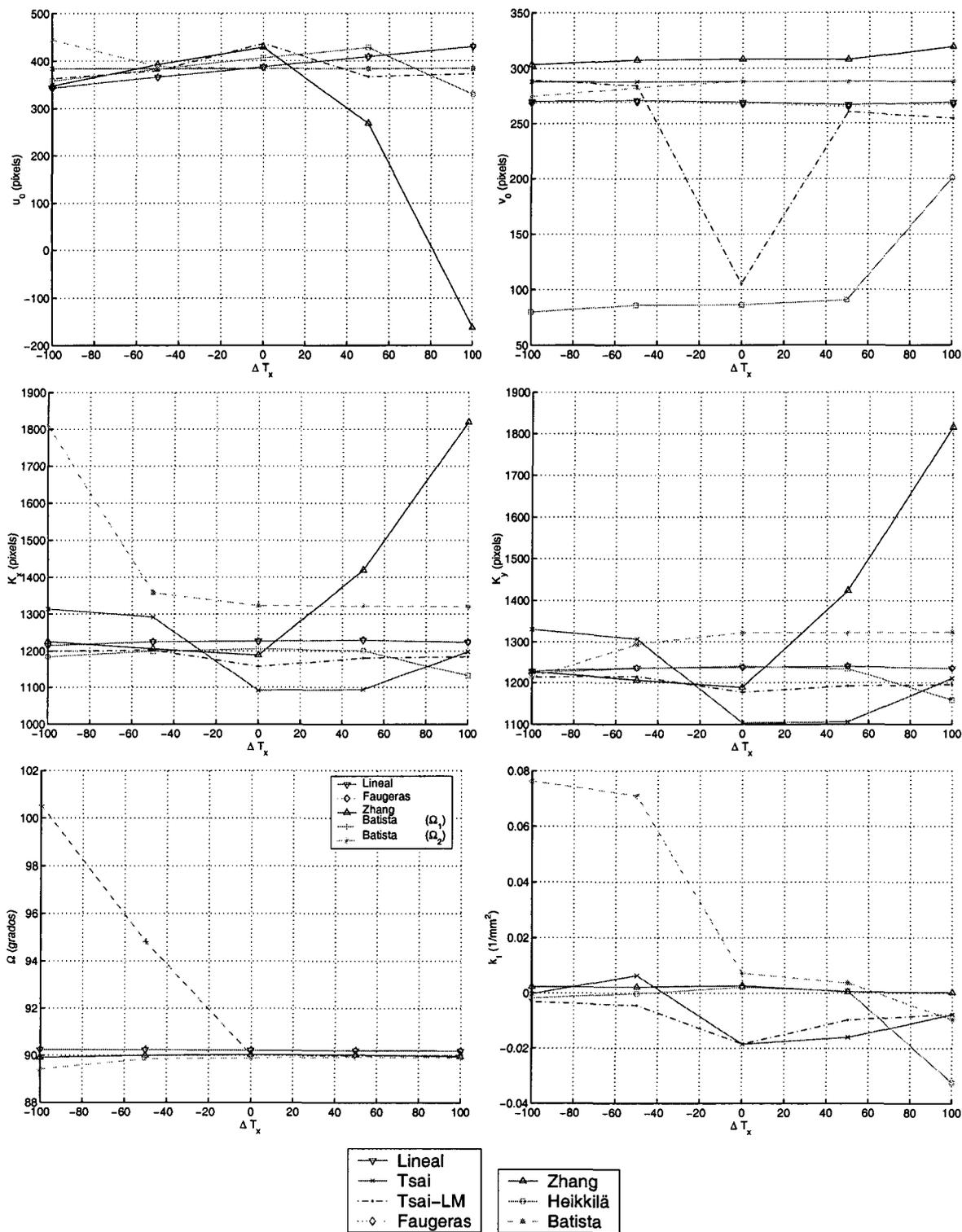


Figura 3.11: Valores de los parámetros intrínsecos obtenidos de la variación de T_x .

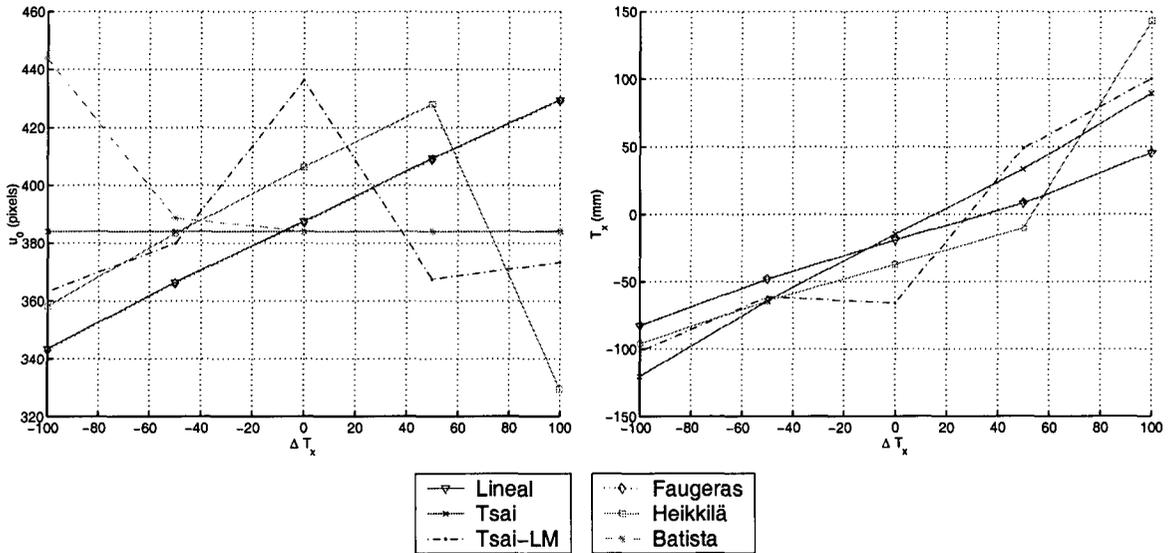


Figura 3.12: Resultados de la calibración obtenidos al variar T_x . Se muestran las estimaciones de u_0 sin el método de Zhang, para una mejor lectura de los resultados del resto de los métodos

converge, y en la estimación de v_0 por parte del método de Tsai-LM.

Sólo los métodos de Faugeras y Lineal parecen tener un comportamiento constante en todos los parámetros, salvo en la estimación de u_0 , en donde el valor de este parámetro varía en función del desplazamiento horizontal realizado en el patrón (figura 3.12).

Si se analiza la posición estimada del objeto (T_x) parece que el método de Tsai es el que estima con mayor precisión el desplazamiento del patrón realizado en la realidad, ya que el incremento que experimenta entre muestras (unos 5 cm) se corresponde con el realizado en la realidad.

Salvo en los casos de los métodos de Batista y Tsai, el comportamiento irregular de los métodos a la hora de obtener los parámetros intrínsecos no se ve reflejado en el error en la reconstrucción de los puntos del patrón.

En resumen, se puede afirmar que la colocación del patrón es determinante a la hora de estimar con precisión los parámetros intrínsecos. Se debe indicar, que el patrón se encontraba situado aproximadamente a 1.10-1.15 m de la cámara, lo que permitió que en todo momento estuviera dentro del campo de visión de ésta. Como se observó en los experimentos realizados en el simulador, el factor de la distancia es decisivo a la hora de obtener resultados precisos cuando los datos de entrada estaban contaminados con ruido, pero este rango de distancia no es lo suficientemente alto como para tener una incidencia determinante en los resultados.

Finalmente, si se comparan los dos experimentos realizados en esta sección se observa que la variación de los parámetros extrínsecos tiene mucha más incidencia en los resul-

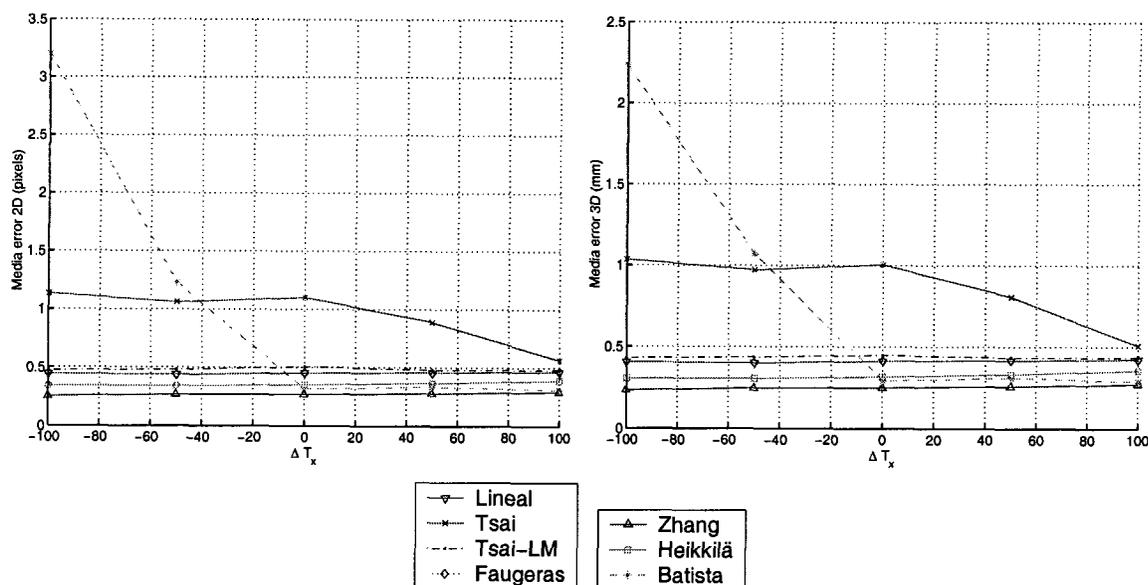


Figura 3.13: Error en la reconstrucción de los puntos obtenido al variar T_x .

tados de los parámetros intrínsecos de la que se produce en el proceso inverso, es decir, la dispersión de los resultados de los parámetros extrínsecos al variar los intrínsecos.

3.4. Variación en la configuración del patrón

Esta sección se analiza la influencia del número de puntos y el tamaño de los puntos sobre la fiabilidad de los resultados. Teóricamente, al variar estos elementos, los resultados en los parámetros intrínsecos deben ser similares. Por otro lado, como se ha mostrado en el experimento anterior, los niveles de error entre la calibración de los dos conjuntos de imágenes (Conjunto 8x8 y Conjunto 14x20) utilizados varían bastante.

Al comparar las figuras correspondientes a los parámetros intrínsecos de los dos conjuntos de imágenes (3.3 y 3.7) se puede observar que todos los parámetros intrínsecos varían bastante de una imagen a otra. Este efecto se repite en todos los métodos, salvo en el método de Batista que sólo se observa una pequeña variación en la estimación de la asimetría. El resultado de este método es lógico debido a que siempre obtiene estimaciones de los parámetros intrínsecos muy próximas a la aproximación inicial introducida en el método.

También se observa que todos los métodos tienen un comportamiento más irregular en las gráficas del primer conjunto de imágenes que en las del segundo motivado probablemente por la cantidad de puntos procesados.

En lo que se refiere a los parámetros extrínsecos (figuras 3.4 y 3.8), es llamativa la gran diferencia que existe en la estimación de la distancia a la que se encuentra el

patrón (T_z) entre los distintos métodos. Esta diferencia ronda los 20 cm de diferencia sobre un promedio de 120 cm en la primera muestra, y sobre los 10 cm de diferencia sobre un promedio de 145 cm en la segunda muestra. Hay que indicar que al haber mayor cantidad de puntos la estimación parece mejor, ya que los distintos métodos convergen a resultados más próximos.

Por otro lado, y al contrario de lo que podría parecer, no todos los métodos mejoran el resultado en términos de error 3D (nivel de error en la reconstrucción de los puntos) al aumentar el número de puntos. De hecho, los métodos de Batista y Zhang, los cuales trabajan con planos, obtienen peores resultados con el Conjunto 14x20. Hay que recordar que en este conjunto se emplea el doble de puntos en el caso del método de Batista y el sextuple en el caso de Zhang. El problema puede tener su origen en que los puntos utilizados son de menor dimensión (6.5 mm. de diámetro frente a los 8.75 mm. y 15 mm, respectivamente, del patrón utilizado en el otro conjunto de datos) y están situados más próximos unos a otros. Además, la distancia a la que se encuentra situado el patrón respecto a la cámara es aproximadamente 30 cm. mayor. Esto produce que el área en las imágenes adquiridas de los círculos del patrón sea inferior y, por tanto, la obtención del centroide de estos círculos sea más imprecisa.

Para comprobar la influencia que tiene el tamaño de los puntos del patrón sobre las coordenadas 2D de la imagen, se ha procesado una imagen variando el nivel de umbral aplicado. Como se puede observar en las figuras 3.1 y 3.2 los patrones están formados por puntos negros sobre fondo blanco. Para realizar el proceso de detección de la región que forma cada punto se utiliza un valor umbral que separe lo blanco de lo negro. Variando este umbral se consigue que la región que forma cada punto aumente o disminuya ligeramente de tamaño. Al estimar el centroide de la región que forma cada punto con diferentes niveles de umbral se observa que la dispersión media obtenida para la posición del centroide de cada punto del patrón es de 0.05 pixels, aunque llega a máximos de 0.19 pixels en u y a 0.10 pixels en v . Estos resultados indican que existe una pequeña influencia del tamaño sobre la precisión en la estimación de las coordenadas 2D la imagen.

También se ha analizado la influencia que puede tener la iluminación sobre la precisión en la obtención del centroide. Este factor es importante debido a que la cámara regula de forma automática la apertura del iris. La variación de la apertura influye sobre la cantidad de luz que incide sobre el sensor de la imagen, lo que hace que varíe el umbral aplicado en el procesamiento de la imagen. El segundo elemento que se ve afectado por este hecho es la profundidad de campo, que indica el rango de distancia en el cuál un objeto parece estar enfocado [Jiménez, 2000].

Para analizar la influencia de la iluminación en el cálculo del centroide de los puntos

en la imagen, se han capturado cuatro imágenes de un patrón sin variar ni su posición ni la configuración de la cámara. Las cuatro imágenes ha sido adquiridas en diferentes condiciones de iluminación. La prueba se ha realizado dos veces, una con el enfoque fijo y la otra con el enfoque automático.

Como resultado de las pruebas con el enfoque automático se obtiene que la distancia promedio entre los centroides correspondientes a un mismo punto del patrón en las cuatro muestras para cada uno de los puntos es de 0.11 pixels en el eje horizontal de la imagen y de 0.10 pixels en el vertical, alcanzando valores máximos de 0.26 y 0.20 respectivamente. En el caso de usar enfoque fijo, la incidencia es bastante inferior, siendo el promedio de 0.07 pixels en ambos ejes y los valores máximos de 0.18 y 0.17 respectivamente.

Como se puede observar es más conveniente usar imágenes con enfoque fijo para realizar la calibración ya que el cálculo del centroide de los puntos será más preciso. Además, el primer resultado concuerda con el obtenido por la variación del umbral.

Por tanto, se deduce que el nivel de ruido en las coordenadas 2D de la imagen no debe superar los 0.1 pixels, en el caso de que el enfoque sea fijo, y los 0.2 pixels, en el caso de que el enfoque pueda variar. Si se observan las gráficas de los experimentos realizados en el simulador (figuras 2.10 - 2.12) el error producido en cada uno de los parámetros y en la reconstrucción de los puntos con estos niveles de ruido no llega a ser significativo.

3.5. Variación del zoom

En este punto se analizará el comportamiento de los distintos parámetros ante el zoom de la cámara. El objetivo es conocer qué parámetros se ven afectados por este elemento y validar los resultados de Willson [Willson, 1994], generalizando sus conclusiones, ya que en sus experimentos sólo utiliza el método de Tsai.

En las pruebas realizadas sobre los dos conjuntos de datos mencionados hasta el momento se aumentó linealmente el grado de zoom aplicado. En las figuras 3.3 - 3.10 se muestran los resultados de la calibración con diferentes niveles de zoom. En el eje de abscisa de las gráficas se muestra el número de pasos de motor que se ha desplazado la lente.

En los resultados obtenidos en el simulador ya se observaba que el método de Heikkilä es muy sensible al ruido, triplicando en algunos parámetros el error cometido por el resto de los métodos. Debido a esta debilidad del método, parece razonable que sus resultados disten exageradamente del resto de los métodos en situaciones reales, donde, además, el nivel de la distorsión parece no ser importante. Se puede observar que este método difiere en sus resultados respecto al resto en todos los parámetros, excepto en k_x , k_y y

T_z , en los cuales se alcanzan valores similares. Esto concuerda también con el resultado del simulador que indicaba que en estos parámetros el nivel de error, aunque elevado, era más próximo al del resto de los métodos, siendo semejante al del método Zhang e inferior al del método de Tsai-LM.

También hay que destacar la poca robustez de los métodos de Heikkilä y Tsai-LM, porque en el segundo conjunto de imágenes obtienen valores negativos en la estimación de v_0 , resultado totalmente ilógico y que invalida los resultados globales. Esto provoca en la mayor parte de las gráficas exista una gran dispersión de los resultados de estos métodos respecto al resto. Debido a esto, en la figura 3.10 se han excluido los datos correspondientes a estos dos métodos.

Como ya se ha comentado, los resultados obtenidos al calibrar las imágenes del Conjunto 8x8 difieren bastante de los obtenidos al calibrar las imágenes del Conjunto 14x20. A continuación se analiza parámetro por parámetro los resultados obtenidos en ambos conjuntos.

En primer lugar, en el caso de centro del eje óptico, se observa que el método de Zhang es el que presenta una mayor variación de este parámetro (excluyendo los dos métodos ya comentados). El resultado muestra una variación entre 50 y 75 pixels en u_0 y entre 10 y 15 pixels en v_0 , siendo esta variación mucho más acentuada en los resultados del Conjunto 8x8. En los métodos de Faugeras y Lineal este efecto es más atenuado (20 pixel en u_0 y 10 pixels en v_0) y la variación más uniforme, incluso en los resultados del Conjunto 14x20 prácticamente no existe variación en u_0 . Se puede observar que estos dos métodos tienen resultados muy semejantes entre sí. Esto concuerda con los resultados del simulador y con el hecho de que algorítmicamente ambos métodos sean muy parecidos. Por último, se muestra que el método de Batista obtiene valores muy próximos al centro de la imagen. Lo cual también es lógico porque este método siempre obtiene resultados muy próximos a la aproximación inicial introducida.

Asimismo, la distancia focal varía en proporción al alejamiento que se realiza de la lente respecto del plano imagen. Pero, la estimación de esta distancia no es igual en todos los métodos, apreciándose dos grupos de curvas. Por un lado, las del método de Tsai y Batista, que muestran un valor muy próximo a la focal nominal dada por el fabricante, y por otro el resto, que obtienen un valor muy inferior a las anteriores (hasta cerca de 2 mm de diferencia cuando el valor es mayor). Parece que este resultado tiene relación con el hecho de que estos dos métodos mantienen fijo o prácticamente fijo el centro óptico sobre el centro de la imagen. Este resultado obtenido de la primera muestra de datos es equivalente al obtenido en la segunda muestra para el caso del método de Batista, pero no en el caso de Tsai que obtiene valores similares a los del resto de los métodos.

En la estimación de la asimetría destaca el hecho de que salvo en el caso del método

de Zhang, este es un parámetro que prácticamente no se ve afectado por el zoom de la cámara (sólo una pequeña variación de centésimas de grado).

Por otro lado, se observa que la magnitud del coeficiente de distorsión disminuye exponencialmente a medida que el número de pasos del motor del zoom aumenta. Solamente cabe destacar que el valor del coeficiente tiende a cambiar de signo al aumentar el zoom en muchos casos. Éste es de los pocos parámetros para los cuales los resultados obtenidos con uno de los conjuntos de imágenes coinciden con los del otro en casi todos los métodos.

Al igual que en las gráficas de los factores de escalado se distingue que el método de Tsai obtiene un resultado en la estimación de T_z muy distante al obtenido por el resto de los métodos. En cuanto al efecto de la variación del zoom los más afectados en este parámetro son los métodos de Faugeras y Lineal, en los cuáles aparece una variación que ronda los 4 cm. Hay que hacer notar que mientras en estos métodos decrece el valor de T_z al aumentar la distancia focal, en el de Tsai hay un ligero incremento de un par de centímetros.

Mientras el método Tsai presenta pequeños decrementos de 2 ó 3 mm. en la estimación de T_x y T_y cuando aumenta el número de pasos de motor del zoom, el resto de los métodos presentan una variación entre 5 y 7 mm. Dada la distancia a la que se encuentra el patrón no parece que estas alteraciones sean significativas. También se observa que el comportamiento del método de Tsai es menos irregular que el del resto de los métodos. Esto se debe a que hace coincidir el cálculo del centro óptico con el centro de la imagen.

Respecto a las medidas de error utilizadas se observa que el error en la reconstrucción 2D de los puntos va aumentando a medida que aumenta el zoom, aunque también coincide con la expansión del patrón en la imagen. En cambio, el error en la reconstrucción 3D permanece prácticamente constante o con un ligero decremento.

En resumen, todos los parámetros se ven afectados por el zoom. Algunos de forma clara y manteniendo siempre la misma tendencia como es el caso de k_1 , que disminuye exponencialmente su magnitud a medida que la lente se separa del plano imagen. En el caso del centro óptico también existe una tendencia, salvo alguna excepción, a separarse del centro de la imagen a medida que aumenta el grado de zoom. En estos dos casos ocurre que aunque la tendencia general es la misma, la dirección de cada uno de ellos es distinta. Es decir, hay métodos que obtienen valores de k_1 positivos y otros negativos, o incluso, el mismo método varía de un conjunto a otro. Lo mismo ocurre en el caso del centro óptico, ya que hay métodos que obtienen valores superiores al centro de la imagen y tienden a crecer al aumentar el zoom, y otros con valores inferiores, tendiendo a decrecer; incluso de un conjunto de imágenes a otro el mismo método cambia su tendencia. En cambio, en la estimación de la ortogonalidad de los ejes del plano imagen

no se aprecia un notorio desplazamiento de los valores obtenidos.

En la estimación de la posición del patrón también se producen ligeras variaciones de unos milímetros (centímetros en el caso de T_z), aunque la tendencia difiere de unos métodos a otros. En cuanto a la orientación también se producen en todos los métodos ligeras variaciones de décimas de grado, y al igual que en el caso anterior, las tendencias varían en función del método.

Quizás el efecto más destacable es el hecho de obtener un crecimiento constante del error en la reconstrucción de las coordenadas 2D, mientras que en las coordenadas 3D el error se mantiene constante. La causa de este efecto se estudiará en posteriores secciones.

3.6. Influencia del enfoque

Hasta el momento todas las calibraciones han sido llevadas a cabo sobre imágenes en las que se variaba el grado de zoom, pero en las que el enfoque permanecía constante. A priori, como ya se ha visto en apartados anteriores, el enfoque influye sobre el tamaño del punto, y por tanto, sobre la precisión en la obtención del centroide del mismo. Además, el enfoque produce una variación en los valores de los parámetros intrínsecos, ya que la distancia entre la lente y el plano imagen varía. En este apartado se trata de analizar si la influencia del enfoque es determinante en los resultados de la calibración.

En esta última prueba sobre movimiento del zoom, se ha repetido la calibración sobre un conjunto de imágenes capturadas en las mismas condiciones que el Conjunto 14x20, pero a la hora de adquirir las imágenes el enfoque se ajustaba automáticamente. La posición y orientación del patrón respecto a la cámara se ha mantenido igual a la adquisición de dicho conjunto de imágenes. Esta prueba sólo se ha realizado con los métodos que trabajan con el diedro.

Como se puede observar en las figuras 3.14-3.16, las diferencias con las gráficas obtenidas de la calibración de las imágenes del Conjunto 14x20 no son significativas. Respecto a este conjunto de datos, los datos obtenidos de la calibración muestran que la estimación del centro óptico varía entre 1 y 2 pixels, los factores de escalado unos 6 pixels (equivalente a una variación en la distancia focal de 0.025 mm), el valor de Ω permanece prácticamente igual. El parámetro más afectado es k_1 ya que la diferencia promedio ronda el 10 % del valor, siendo más acentuado en el caso del método de Heikkilä.

Hay que hacer notar que la calibración realizada por el método de Tsai-LM en la distancia focal máxima produce un valor discordante con el resto. Este resultado se explica porque el proceso iterativo de optimización queda atrapado en un mínimo local y no converge a los resultados esperados.

En cuanto a los parámetros extrínsecos también se obtiene que la incidencia es mí-

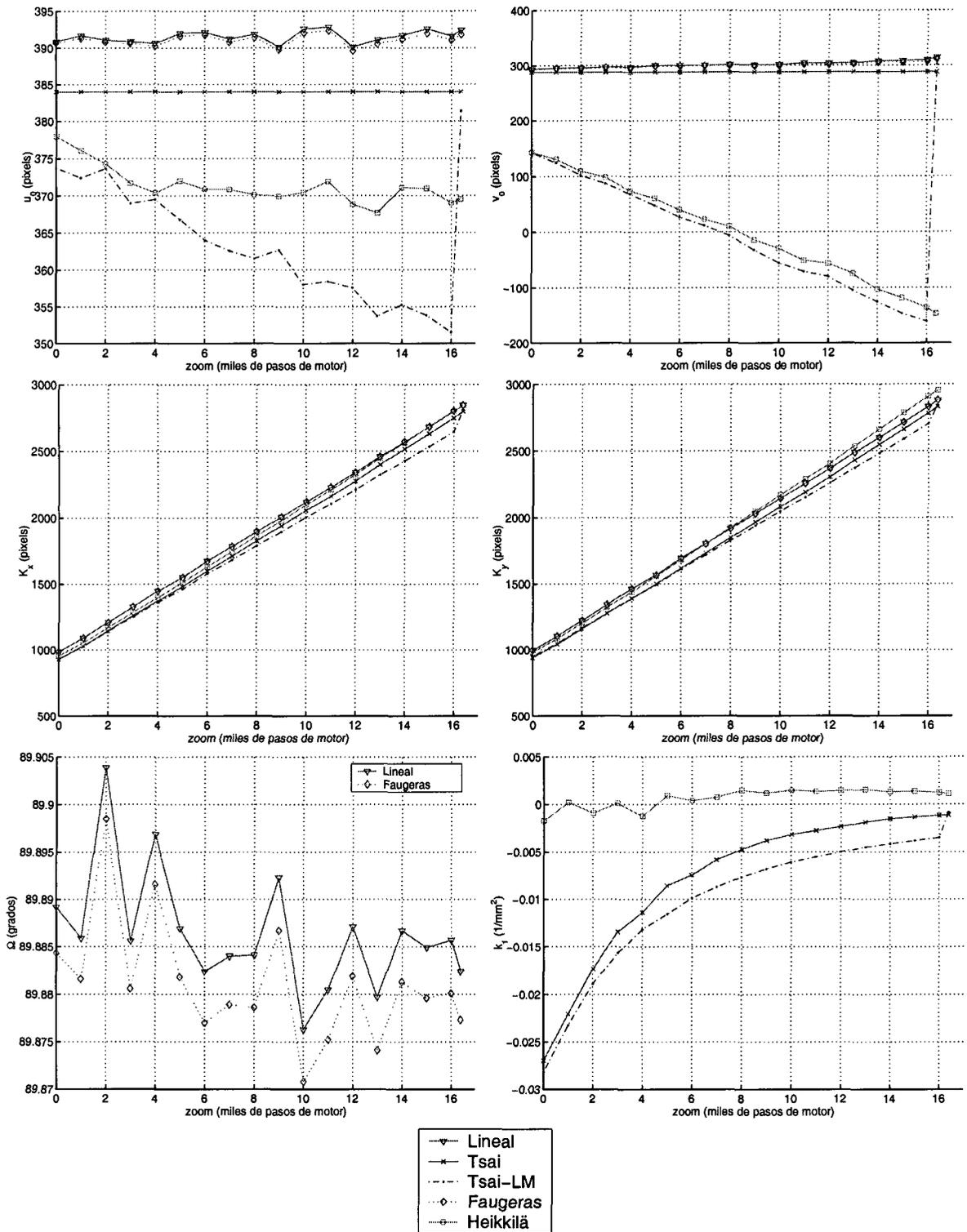


Figura 3.14: Parámetros intrínsecos obtenidos de la variación del zoom con enfoque automático.

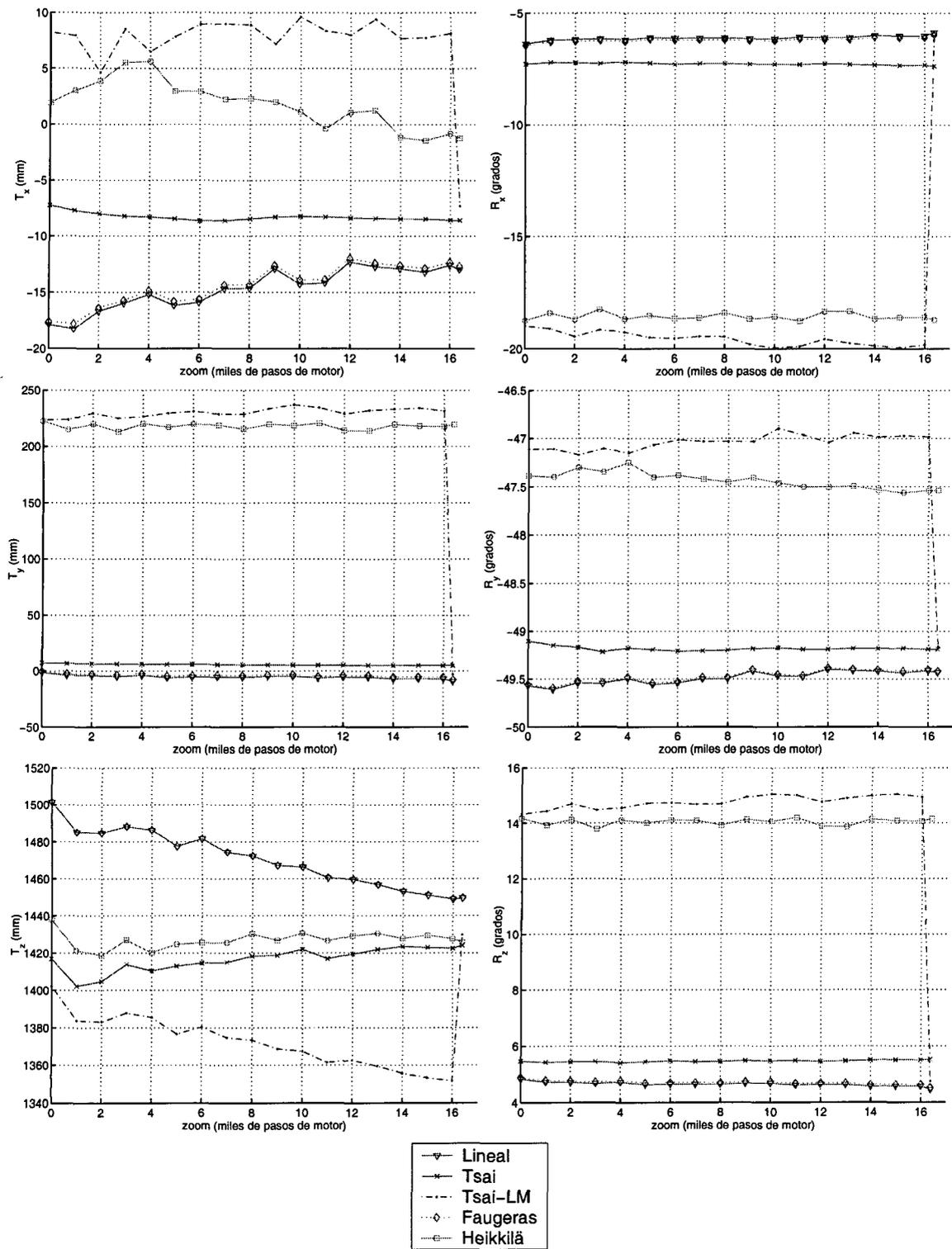


Figura 3.15: Parámetros extrínsecos obtenidos de la variación del zoom con enfoque automático.

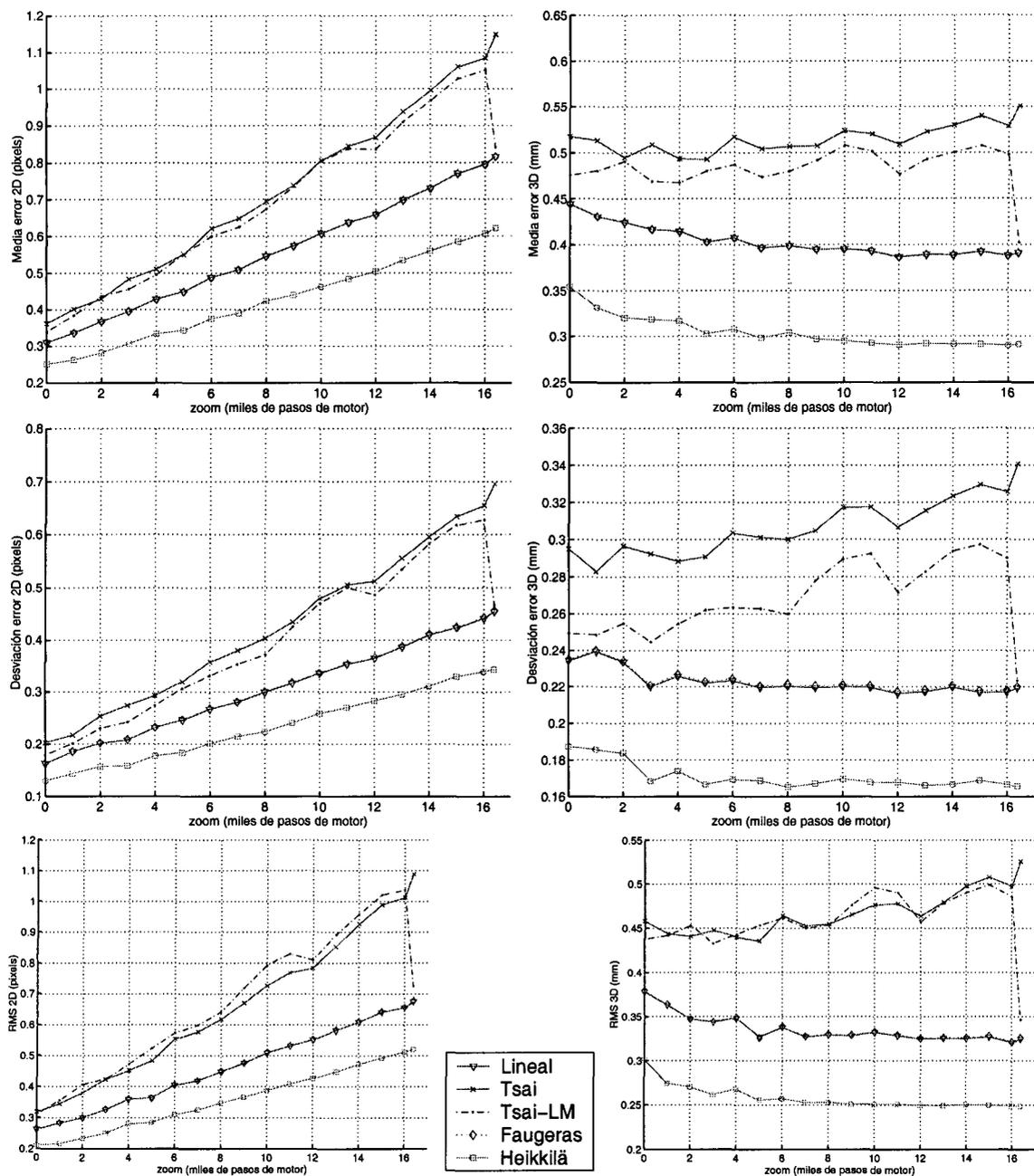


Figura 3.16: Error en la reconstrucción de los puntos producido al variar el zoom con enfoque automático.

nima. En la posición la diferencia se encuentra entre 1 y 2 mm para T_x y T_y , mientras que para T_z es de unos 4.5 mm (supone el 0,3 % de su valor). En la estimación de los ángulos de la orientación la diferencia promedio ronda los 0.1 grados.

A partir de estos datos se puede deducir que la influencia del enfoque no es significativa en el proceso de calibración, al menos en la cámara utilizada y bajo las condiciones en las que se han realizado estos experimentos.

3.7. Desplazamiento de la distancia del patrón respecto a la cámara y zoom

En los experimentos anteriores se ha mostrado cómo al modificar la distancia focal, prácticamente todos los parámetros varían su valor. Pero esta variación provoca que el objeto se vea más grande o más pequeño en la imagen. Por lo tanto, la cuestión que se puede plantear es si la variación de estos parámetros se debe realmente a la influencia del zoom o está motivada por el tamaño del patrón en la imagen.

Para resolver esta cuestión en este experimento se trata de observar cómo influye el movimiento del zoom en la calibración de un patrón que siempre ocupa la misma porción de imagen. Para ello, se han tomado 6 imágenes en las que el zoom se ha desplazado 1500 pasos de motor entre cada una de ellas y el patrón se ha ido alejando de la cámara de 10 en 10 centímetros. Se ha utilizado como patrón el diedro de 14x20 puntos en cada plano (figura 3.2b).

Los resultados se muestran en la figuras 3.17-3.19 e indican que el centro óptico depende de la variación del zoom, ya que, aún ocupando el patrón siempre la misma porción de imagen, sigue habiendo una variación en la estimación de este parámetro.

Otro de los parámetros que planteaba dudas en las calibraciones anteriores es el coeficiente de distorsión. En este caso se puede observar que al aumentar el valor del zoom sigue existiendo un decremento exponencial de la magnitud de k_1 , aunque también se puede observar que esta magnitud es unas diez veces inferior a las obtenidas en los casos anteriores.

En los datos obtenidos de los parámetros extrínsecos se puede comprobar que el método, de los estudiados que trabajan con patrón, que parece estimar mejor la variación de T_z es el método de Tsai-LM. Esto se debe a que es el único que presenta un continuo desplazamiento de aproximadamente 10 cm. (con un error inferior a medio centímetro) entre imágenes, correspondiéndose esta distancia al desplazamiento realizado en la realidad. En el extremo contrario se encuentran los métodos de Tsai y Batista, que difieren en más de un 10 % de la estimación realizada de cada desplazamiento.

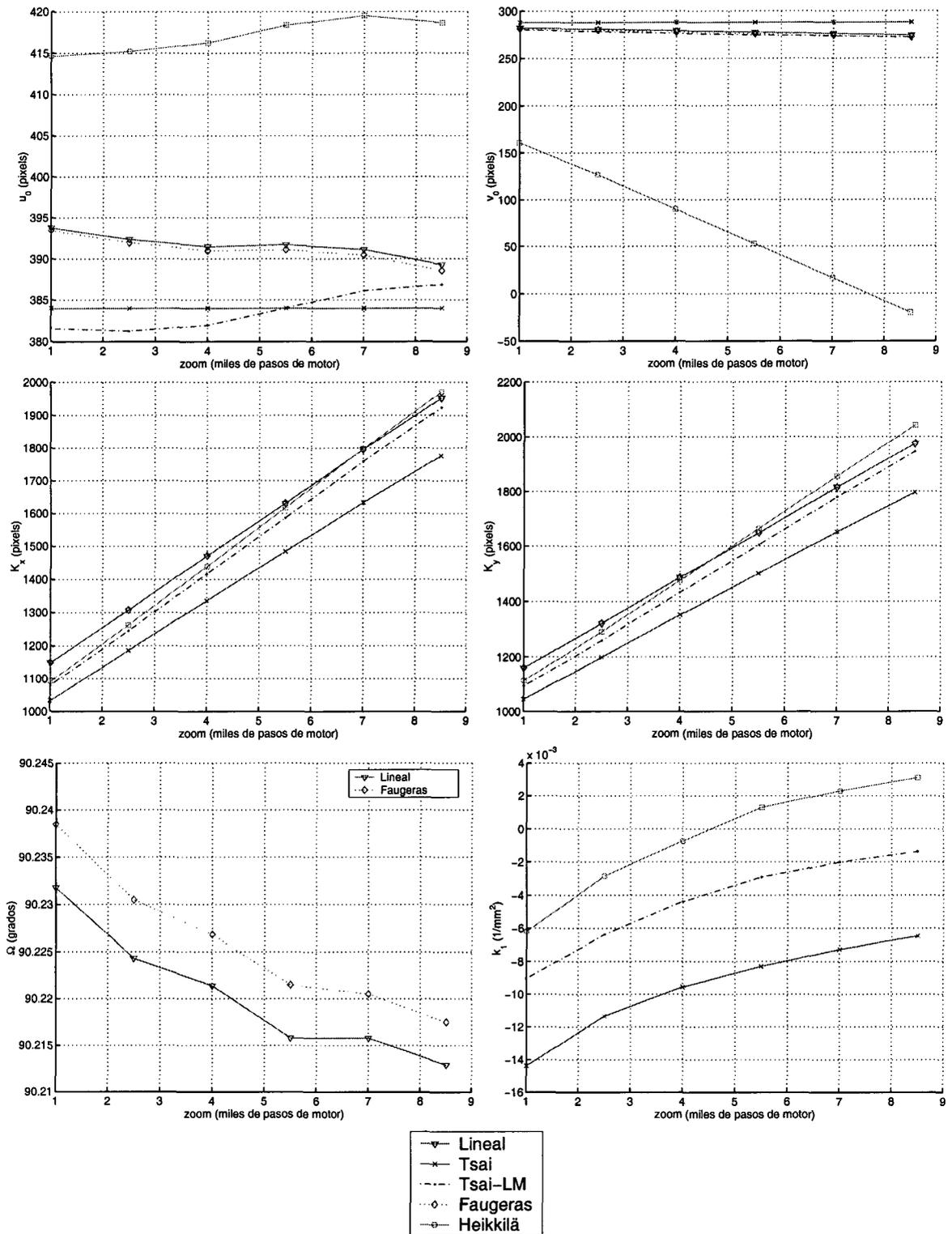


Figura 3.17: Parámetros intrínsecos obtenidos de la variación del zoom y T_z .

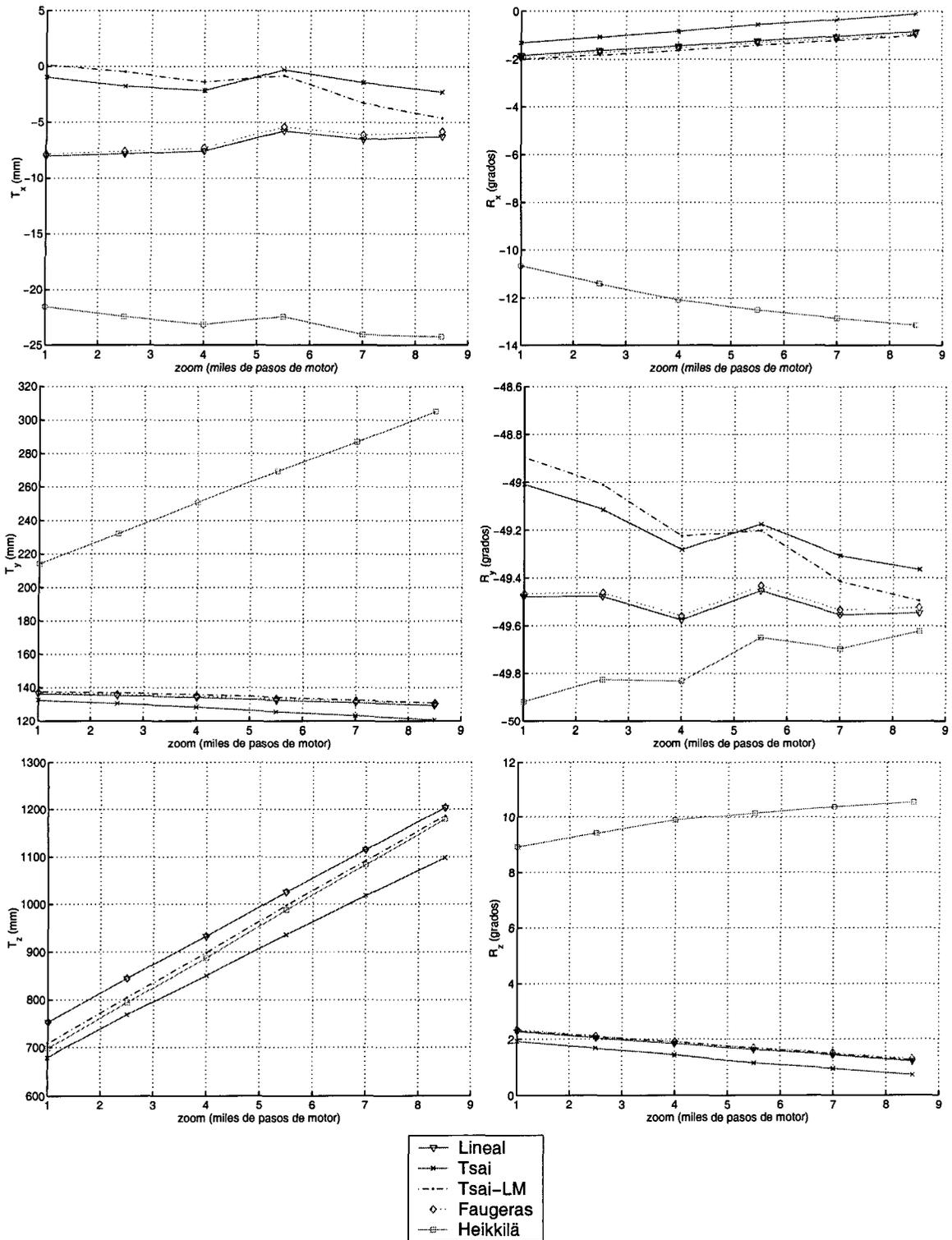


Figura 3.18: Parámetros extrínsecos obtenidos de la variación del zoom y T_z .

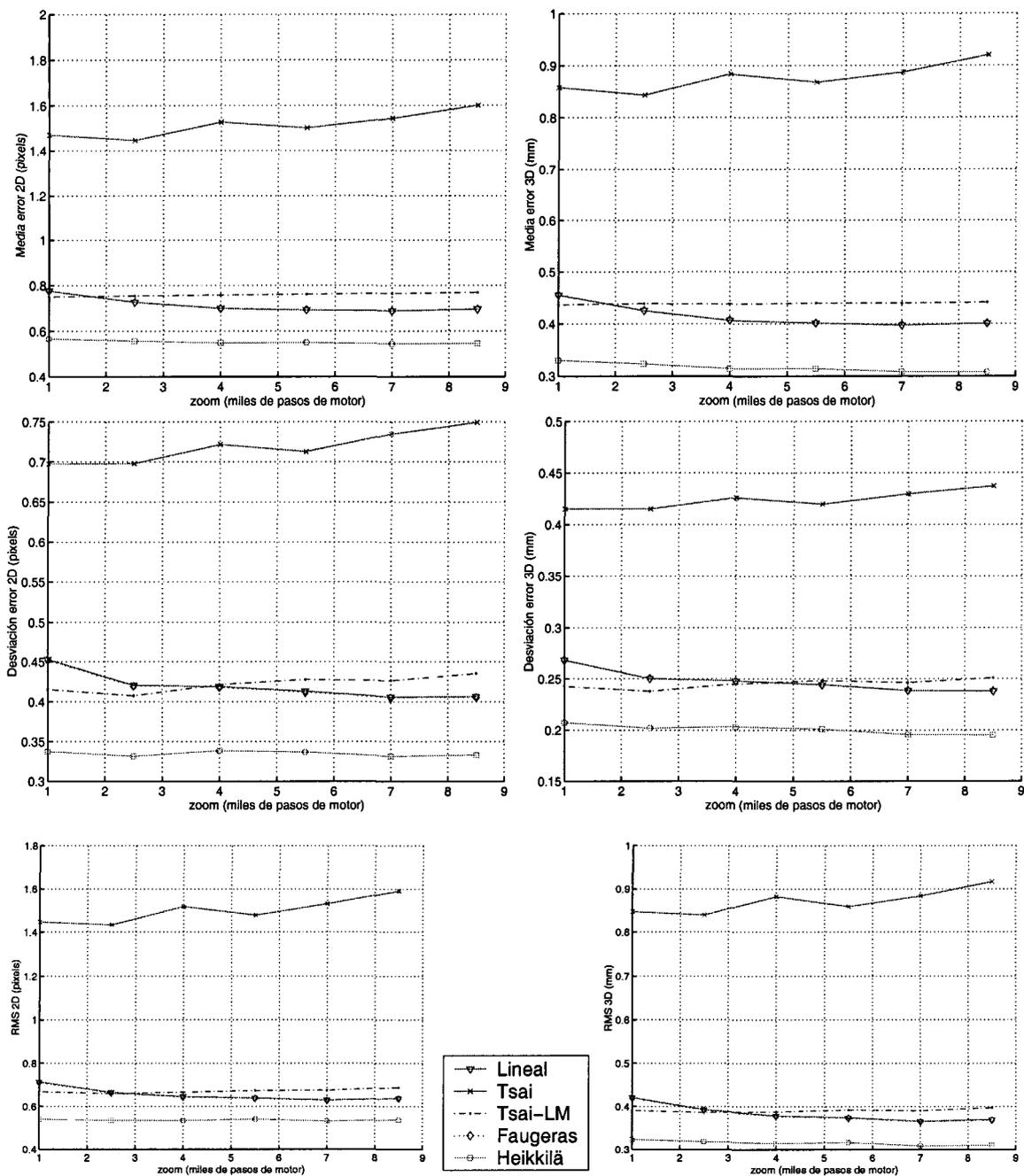


Figura 3.19: Error en la reconstrucción de los puntos producido al variar el zoom y T_z .

Por último, indicar que el error en la reconstrucción de las coordenadas 2D permanece constante ante el movimiento del zoom, esto se contrapone a los resultados obtenidos en las pruebas anteriores. Este resultado parece indicar que el incremento del error que se producía en los anteriores experimentos depende del espacio que ocupa el patrón en la imagen y no del zoom.

En resumen, se ha despejado la duda de si la variación en los parámetros producida cuando se desplaza el zoom está motivada por este elemento o porque se produce un cambio en el tamaño del área que ocupa el patrón en la imagen. Siendo el primer efecto dominante frente al inducido por el cambio del área que los puntos del patrón ocupan en la imagen.

3.8. Conclusiones del estudio sobre datos reales

De la calibración realizada sobre los diferentes conjuntos de imágenes se pueden extraer las siguientes conclusiones:

- La estimación del centro del eje óptico por cada uno de los métodos ha variado a lo largo de los experimentos. Debido a que la configuración de la cámara es la misma, entonces se debe concluir que sobre este parámetro influye la colocación del patrón respecto a la cámara (parámetros extrínsecos). Además este hecho se ve reforzado con el resultado del experimento en el que se desplazaba horizontalmente el patrón respecto a la cámara y se observaba cómo variaba la estimación del centro óptico. Como muestra ese experimento, esta influencia de los parámetros extrínsecos se extiende a todos los parámetros intrínsecos.
- También se ha demostrado que la variación de parámetros intrínsecos al variar el zoom incide en la estabilidad de los parámetros extrínsecos, aunque en menor medida que en el caso anterior.
- La estimación de la distancia a la que se encuentra el objeto (T_z) parece ser un aspecto crítico e inestable, ya que, en todos los casos existen grandes diferencias entre la estimaciones realizadas por los distintos métodos.
- Al variar la distancia focal (zoom) parece que se ven afectados el centro del eje óptico, que varía linealmente, y el primer coeficiente de distorsión radial, que decrementa de forma exponencial su magnitud a medida que aumenta el valor del zoom. Los parámetros extrínsecos se ven todos afectados por esta variación.
- El error en la reconstrucción de las coordenadas 2D de la imagen aumenta a medida que aumenta el número de pasos del motor del zoom. Esto puede deberse a que

existe una descompensación en la relación f/Z_c (ecuación 2.6), ya que cuando la distancia del patrón varía en la misma proporción que la distancia focal, se observa que este error se mantiene constante.

- Hay métodos como los de Heikkilä y Tsai-LM que se presentan como poco robustos, ya que, en determinadas configuraciones llegan a obtener valores negativos en la estimación del centro óptico. Esto también ocurre en el método de Zhang si el patrón se encuentra muy desplazado del centro de la imagen.
- El aumento del número de puntos no parece determinante en los resultados obtenidos, aunque sí parece que la precisión con que se pueda obtener la ubicación real de las coordenadas 2D de estos puntos en la imagen lo sea. Esto se ha podido comprobar al comparar los resultados del primer conjunto de imágenes con el segundo. En este último caso se puede observar que en algunos métodos el error de reconstrucción, tanto en 2D como 3D, es mayor que el obtenido con el diedro de 8x8 puntos en cada plano. Hay que tener en cuenta que en el segundo patrón utilizado hay más del cuádruple de puntos del primero. Pero por contra, el tamaño de los círculos es mucho menor (el diámetro pasa de 10 mm a 6.5 mm) y, además, el patrón está situado unos 30 cm más alejado de la cámara, esto hace que el tamaño del punto en la imagen sea mucho menor y por tanto que el cálculo del centroide sea más deficiente.
- Se observa que factores como la iluminación o el tamaño de los puntos tienen una pequeña incidencia en la obtención de las coordenadas 2D de los puntos en la imagen. También se observa que esta incidencia es menor cuando el enfoque permanece constante. A la luz de los resultados del experimentos realizados se puede deducir que la influencia del enfoque es poco significativa sobre los resultados.
- Si se comparan los resultados de la variación del zoom con los obtenidos por [Willson, 1994] se observa que hay una coincidencia en el hecho de que el centro óptico y el coeficiente de distorsión se ven afectados por esta variación. Los resultados muestran que estos parámetros se pueden obtener en función del nivel de zoom aplicado; es decir, el centro óptico se puede ajustar mediante un polinomio de primer grado y la distancia focal mediante un polinomio de segundo grado. A diferencia de los resultados obtenidos por este autor, se observa que también T_x y T_y varían su valor en función a la distancia focal, aunque sí que es cierto que esta variación es pequeña (inferior a un centímetro en la mayoría de los casos) y poco significativa (inferior al 1 % de la distancia).

Por último, a la cuestión que se planteaba al comienzo de cuál de los métodos era más preciso, se ha mostrado que el método de Zhang es el que mejor reconstruye la posición de los puntos del patrón. Pero, experimentos posteriores, añadidos a los realizados en el simulador, han demostrado que este factor no puede tomarse como una medida absoluta de la bondad de un método, ya que este error puede ser muy bajo y la estimación de los distintos parámetros puede ser deficiente. Esto es debido a que los errores en unos parámetros se ven compensados por errores en otros.

Del análisis de la estabilidad de los métodos se desprende que, en general, los métodos de calibración estudiados obtienen parametrizaciones que globalmente minimizan de forma satisfactoria el error de reconstrucción de las coordenadas tridimensionales de los puntos del patrón o el error 2D. Sin embargo, estas parametrizaciones, o grupos de valores de parámetros intrínsecos y extrínsecos, sólo son válidos en conjunto, es decir, estos grupos no pueden ser empleados de forma independiente. Este resultado es relevante por cuanto en numerosos problemas se calibra una cámara para obtener exclusivamente el conjunto de parámetros intrínsecos que la describe.

3.9. Experimentos en el simulador en las mismas circunstancias que los datos reales

Los resultados obtenidos en los apartados anteriores no son concluyentes a la hora de determinar cuál de los métodos es más robusto o preciso. Para tratar de responder a esta cuestión se han trasladado al simulador los experimentos realizados con datos reales.

En el simulador se ha tratado de reproducir los experimentos en circunstancias semejantes a los realizados con datos reales: la forma y el tamaño del patrón, la colocación aproximada de éste respecto a la cámara, y la configuración de la cámara usando los datos aportados por el fabricante.

En todos los experimentos realizados, y para todos los métodos utilizados, se ha usado la misma configuración de cámara para proyectar el patrón de calibración. El plano imagen es de 768x576 pixels, situando el centro del eje óptico en el centro de la imagen (384,288). Los ejes del plano imagen son ortogonales. El tamaño del CCD es de 3.2 mm x 2.4 mm, siendo, por tanto, los pixels cuadrados. No se ha introducido inicialmente ninguna distorsión.

Para realizar la calibración con aquellos métodos que necesitan que los puntos no se encuentren en el mismo plano (Lineal, Tsai, Tsai-LM, Faugeras y Heikkilä) se ha utilizado una distribución de puntos en forma de diedro con 8x8 puntos en cada cara

y separados entre sí 2 cm. Las coordenadas del centro del patrón respecto a la cámara son: $x = 0$ mm, $y = 5$ mm, $z = 1200$ mm; con una rotación de -45° en el eje Y_c .

Hay que indicar que el patrón ha tenido que ser desplazado ligeramente en la vertical. Esto se debe a que las dos versiones del método de Tsai fallan estrepitosamente cuando el centro del patrón se encuentra exactamente a la misma altura que el eje óptico de la cámara. Esto se produce porque el rango de la matriz usada en la resolución de la ecuación 2.18 dentro del primer paso del proceso de calibración tiene que tener rango siete (ya que hay siete incógnitas), y en estas circunstancias el rango de esta matriz tiene un valor inferior.

Para el método de Batista se ha utilizado un plano de 13×10 puntos separados entre sí 17.5 mm. El patrón se ha colocado casi paralelo al plano imagen, apenas inclinado un grado en cada eje, y a una distancia de 103 cm. Para el método de Zhang se ha empleado un plano de 7×6 puntos separados entre sí 30 mm. El patrón se ha colocado en tres orientaciones distintas.

3.9.1. Variación de la distancia focal

En este apartado se identifica cuál es la fuente de ruido que afecta en mayor medida a los resultados. Esta identificación puede ser importante porque hay métodos que son más sensibles al ruido en las coordenadas 2D que al producido en las coordenadas 3D, como es el caso del método de Batista, y en otros casos es a la inversa como ocurre en el resto de los métodos estudiados en este trabajo.

En primer lugar se realiza un análisis teórico de la incidencia de cada tipo de error, y en segundo lugar se corrobora experimentalmente en el simulador. Para realizar esto último, se ha variado en el simulador la distancia focal y se ha aplicado, por separado, ruido gaussiano a las coordenadas 2D y a las coordenadas 3D.

Para realizar el análisis teórico de los errores se han simplificado las ecuaciones suponiendo que el modelo está libre de distorsiones, que los pixels son cuadrados y que los ejes del plano imagen son perfectamente ortogonales.

Así, dadas las coordenadas $M_c = [X_c, Y_c, Z_c]^t$ de un punto respecto a la cámara, su proyección en la imagen $m = [u, v]$ en condiciones ideales se obtiene a partir de las siguiente relación (sólo se muestra el desarrollo en uno de los ejes de coordenadas ya que en el otro es equivalente):

$$u = d_x \cdot f \cdot \frac{X_c}{Z_c} + u_0 \quad (3.1)$$

y, por tanto el proceso inverso es:

$$X_c = \frac{u - u_0}{d_x} \cdot \frac{Z_c}{f} \quad (3.2)$$

Si se inyecta un ruido δ en las coordenadas 2D, entonces se tiene el siguiente desarrollo:

$$\hat{u} = d_x \cdot f \cdot \frac{X_c}{Z_c} + u_0 + \delta \quad (3.3)$$

mientras que a la inversa es:

$$\hat{X}_c = \frac{(u + \delta) - u_0}{d_x} \cdot \frac{Z_c}{f} \quad (3.4)$$

Si se estima el error en las coordenadas 2D como $|u - \hat{u}|$ y el de las coordenadas 3D como $|X_c - \hat{X}_c|$, entonces al sustituir en las ecuaciones anteriores se tiene que:

$$Error_{2D} = \delta \quad (3.5)$$

$$Error_{3D} = \frac{\delta \cdot Z_c}{d_x \cdot f} \quad (3.6)$$

Por tanto, se observa que cuando existe ruido en las coordenadas 2D de la imagen, el error en las coordenadas 3D es inversamente proporcional a la distancia focal y directamente proporcional a la distancia al patrón, mientras que el error en las coordenadas 2D es constante.

Por otro lado, Si un ruido Δ es inyectado en las coordenadas 3D, entonces se tiene el siguiente desarrollo:

$$\hat{u} = d_x \cdot f \cdot \frac{X_c + \Delta}{Z_c} + u_0 \quad (3.7)$$

mientras que a la inversa es:

$$\hat{X}_c = \frac{u - u_0}{d_x} \cdot \frac{Z_c}{f} + \Delta \quad (3.8)$$

Si se estima el error de reconstrucción de los puntos se tiene que:

$$Error_{2D} = \frac{d_x \cdot f \cdot \Delta}{Z_c} \quad (3.9)$$

$$Error_{3D} = \Delta \quad (3.10)$$

Por tanto, se observa que cuando existe ruido en las coordenadas 3D de la imagen, el error en el error en las coordenadas 3D es directamente proporcional a la distancia

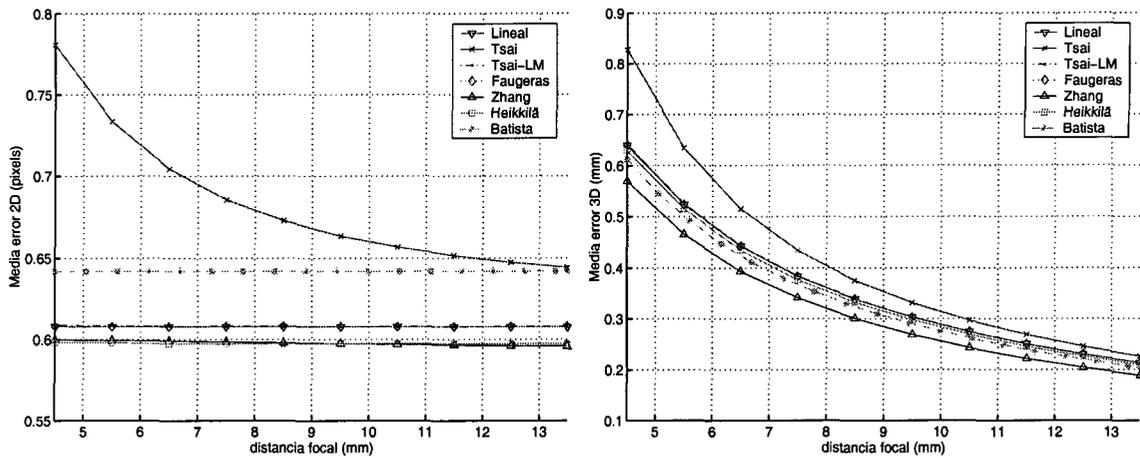


Figura 3.20: Error promedio en la reconstrucción en 2D y en 3D cuando varía la distancia focal en imágenes con ruido gaussiano de media nula y desviación típica de 0.5 pixels en las coordenadas 2D.

focal e inversamente proporcional a la distancia del patrón, mientras que el error en las coordenadas 3D es constante.

Estos análisis han sido corroborados en el simulador. En las gráficas de la figura 3.20 se muestran los resultados obtenidos al variar la distancia focal usando imágenes en las que se ha añadido un ruido gaussiano a las coordenadas 2D de la imagen de media 0 y desviación típica 0.5 pixels. A su vez en la figura 3.21 se muestran los resultados del mismo experimento, pero el ruido gaussiano inyectado de media 0 y desviación típica 0.5 mm es en las coordenadas 3D de los puntos del patrón. Cada experimento ha sido repetido 25 veces.

Como se puede observar, cuando las coordenadas 2D están contaminadas con ruido, el error en las coordenadas 3D se va decrementado a medida que el valor de la distancia focal va aumentando, mientras que el error en las coordenadas 2D permanece prácticamente constante. La excepción la constituye el método de Tsai que desciende su valor en ambos casos. En cambio, cuando el ruido se aplica a las coordenadas 3D el resultado es el inverso. Es decir, el error en la reconstrucción de las coordenadas 2D aumenta en proporción a la distancia focal, mientras que en las coordenadas 3D se mantiene constante.

Si se traslada este análisis a los resultados obtenidos en la calibración a partir de datos reales (figura 3.5), se puede ver que el error que varía es el producido en las coordenadas 2D. Por tanto, se puede deducir que las coordenadas utilizadas de los puntos 3D son mucho menos precisas que las coordenadas 2D. También se puede observar que el nivel de error es mayor que en los datos reales, sobre todo en los métodos de Batista y Zhang. Esto parece indicar que los datos reales deben tener un ruido inferior a los 0.5 mm en

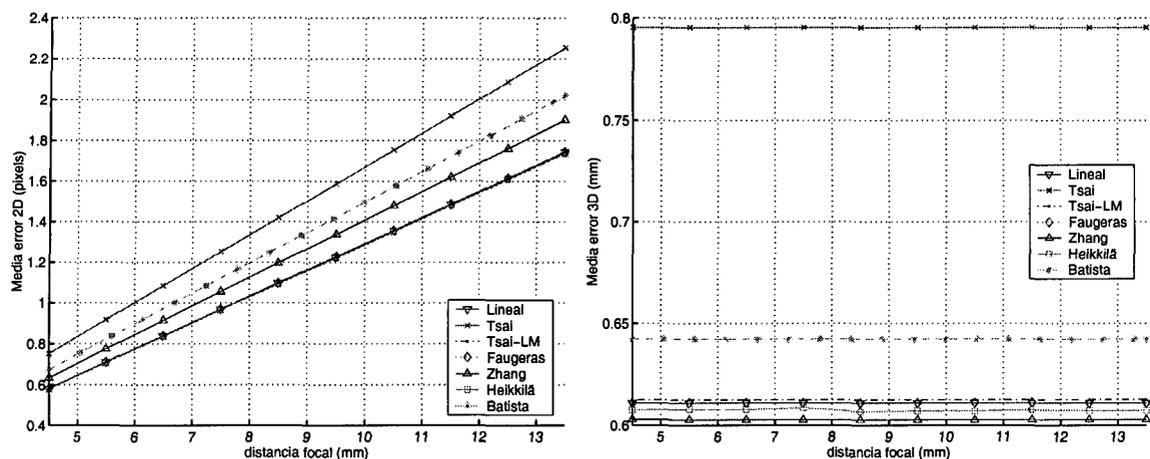


Figura 3.21: Error promedio en la reconstrucción en 2D y en 3D cuando varía la distancia focal en imágenes con ruido gaussiano de media nula y desviación típica de 0.5 mm en las coordenadas 3D.

las coordenadas 3D.

Los datos obtenidos también parecen indicar que el ruido en las coordenadas 3D que afecta a los resultados de los métodos de Batista y Zhang es inferior al del resto de los métodos. Esto es lógico si se observa el patrón utilizado. Estos dos métodos usan puntos coplanares como referente, en los cuales obtener la relación entre ellos produce menos incertidumbre que la producida cuando los puntos tienen una distribución no-coplanar (diedro). El problema de este último tipo de patrón es que aparecen posibles fuentes de error adicionales como son la precisión en la medida del ángulo entre los dos planos o la colocación del origen de coordenadas común.

3.9.2. Influencia de la distorsión

El efecto de la distorsión ya ha sido analizado en el apartado 2.3.2. Pero en el experimento presentado en aquella sección el patrón aparece bastante descentrado en la imagen. Este hecho puede tener gran influencia sobre los resultados ya que la influencia de la distorsión radial depende de la distancia de los puntos al centro. Debido a esta causa se ha repetido el experimento en un intento de observar, en condiciones semejantes a las que se han hecho los experimentos sobre datos reales, la influencia de este parámetro.

Para ello, se ha realizado el proceso de calibración en el simulador utilizando la configuración de cámara y patrones descritas al comienzo de la sección (3.9). Se ha utilizado una distancia focal de 13.5 mm, que es la máxima que describe el fabricante de la cámara utilizada, con el objetivo de que el patrón ocupe el mayor espacio en la imagen y, por tanto, el efecto de la distorsión esté más acentuado, ya que el patrón está

próximo a los bordes de la imagen. Además se ha inyectado un ruido gaussiano de media nula y desviación típica de 0.1 pixels a las coordenadas 2D de la imagen y un ruido de 0.4 mm a las coordenadas 3D de los puntos del patrón. El experimento se ha repetido 25 veces. El valor de k_1 ha variado entre -10^{-5} y -10^{-2} . Estos valores límite se han seleccionado en base a los resultados obtenidos en la sección 2.3.2, donde se observa que el error se mantiene prácticamente constante a partir de magnitudes inferiores a 10^{-4} , y en el caso del límite superior debido a que con esa magnitud la distorsión es muy apreciable en las imágenes, cosa que no ocurre en las imágenes obtenidas por la cámara utilizada.

La figura 3.22 se muestran los resultados más significativos. Las gráficas del error difieren de los resultados obtenidos en el otro experimento (figura 2.22). Se observa que el nivel de error es similar en todos los métodos y no existe un incremento significativo del error a medida que aumenta la magnitud de la distorsión. Sólo en los valores más altos existe un cambio en esta tendencia, en lo que parece el inicio de un crecimiento exponencial de algunas curvas (métodos de Faugeras, Lineal y Zhang). Por un lado, la igualdad en los resultados se debe al hecho de que en aquellos métodos que en el primer experimento se veían más afectados, al centrar el patrón, disminuyen drásticamente el error producido por la distorsión. Por otro lado, el hecho de que el error se mantenga prácticamente constante se debe a que la influencia de los ruidos inyectados a los datos de entrada es muy superior al efecto de la distorsión, salvo cuando la magnitud de ésta es muy elevada. También se observa que cuando el nivel de distorsión es alto el método que mejor reconstruye la posición de los puntos es el método de Heikkilä, lo cual es lógico ya que es un método orientado a este fin.

Lo que sí ofrece información es el análisis de las gráficas de los factores de escalado, en las se observa cómo los valores altos de la distorsión afectan significativamente a los métodos de Faugeras y Lineal. De hecho, a partir de magnitudes de k_1 superiores a 10^{-3} el valor obtenido dista bastante del calculado por el resto de los métodos. Este comportamiento se justifica al ser estos dos métodos los únicos que no modelan la distorsión. Por otro lado, si se observan las gráficas de los experimentos con datos reales (figura 3.3) los valores de estos métodos para los factores de escalado son similares a métodos como los de Zhang, Heikkilä o Tsai-LM. De todo ello se deduce que el valor de la distorsión radial de la cámara empleada debe tener una magnitud inferior a 10^{-3} . Esta conclusión se ve reforzada por el hecho de que el método de Zhang aumenta significativamente el error en la reconstrucción de los puntos, tanto 2D como 3D, cuando la distorsión es elevada, cuestión que no se produce en los datos reales donde este método era el que menor valor presenta en estas medidas de error.

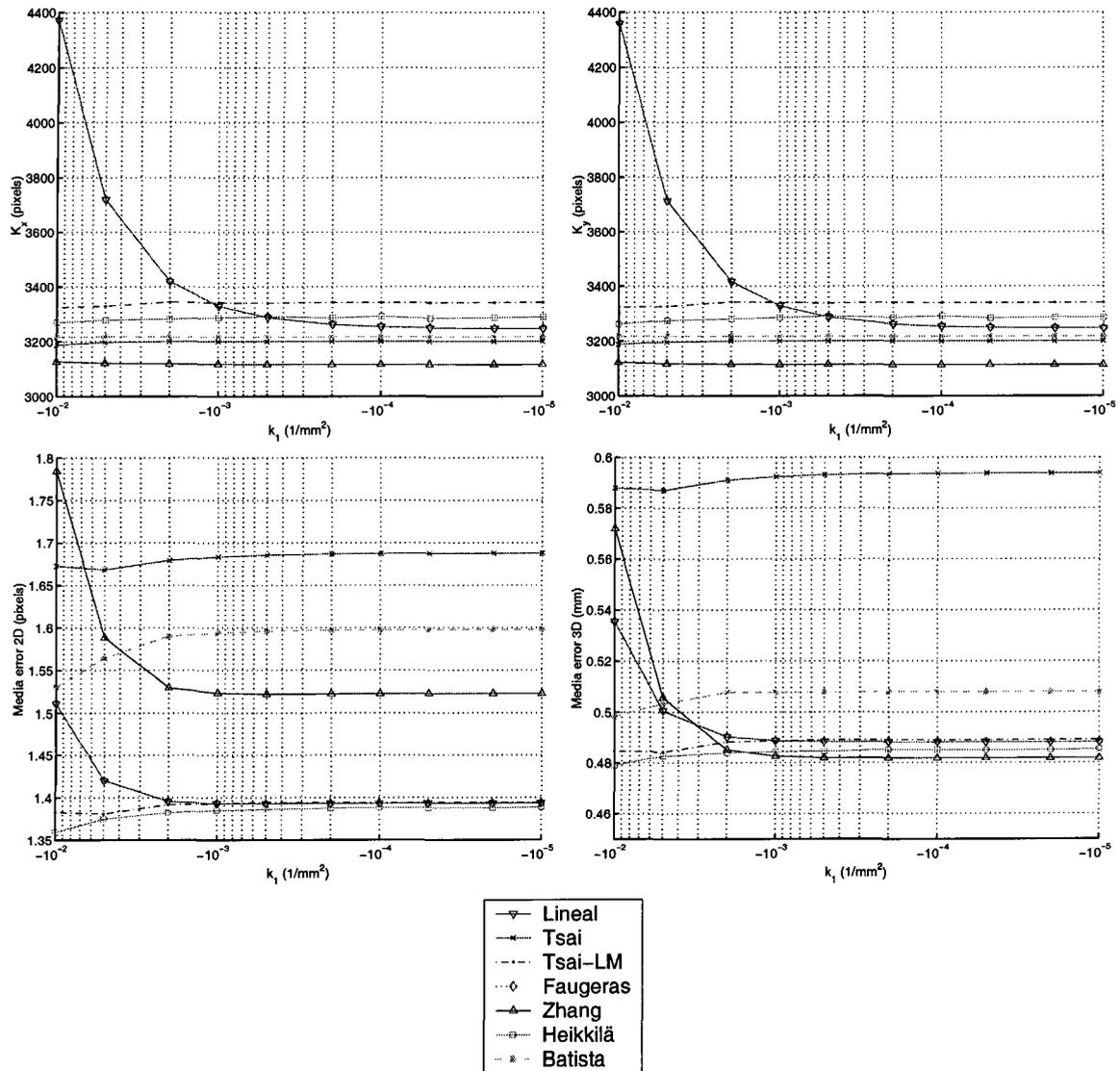


Figura 3.22: Factores de escalado y error de reconstrucción obtenidos de la variación del primer coeficiente de distorsión radial (k_1).

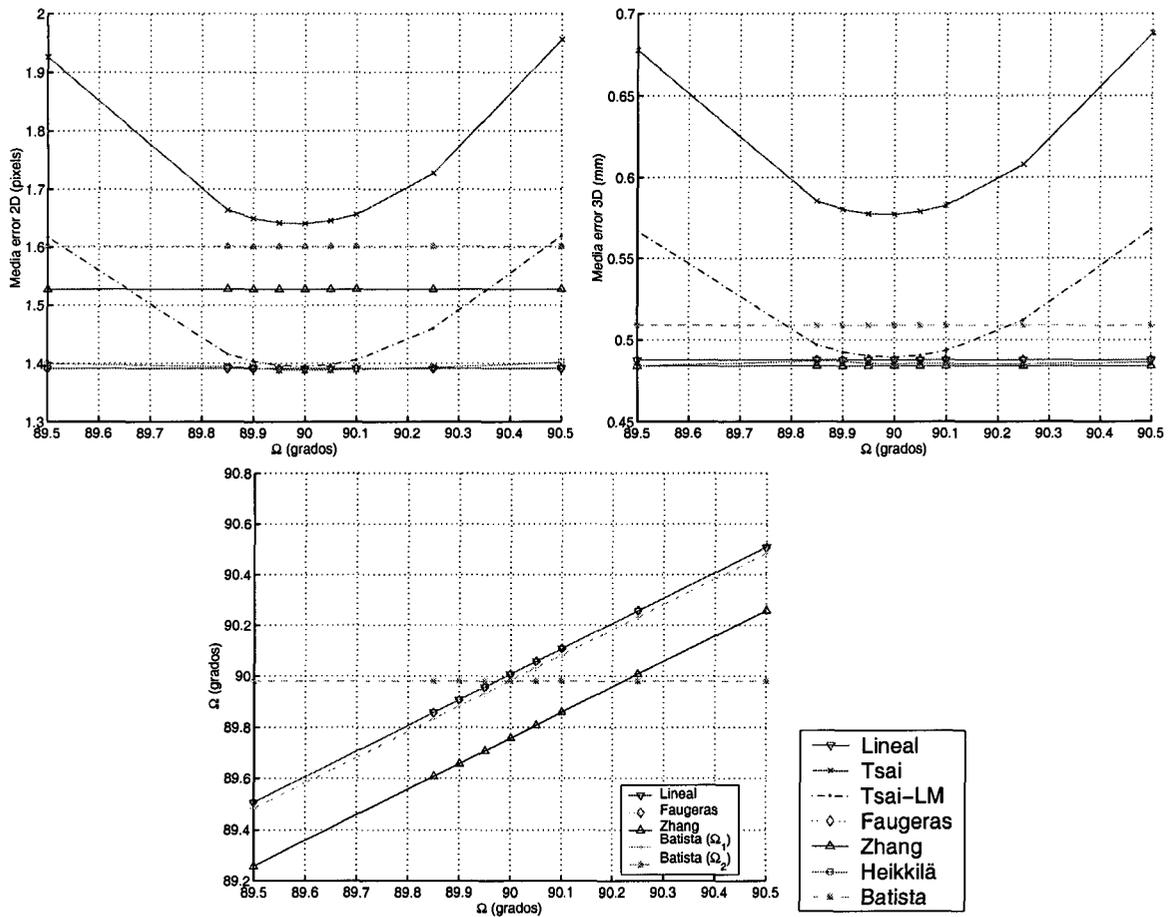


Figura 3.23: Error promedio en la reconstrucción en 2D y en 3D cuando varía el valor de Ω_1 .

3.9.3. Influencia de la ortogonalidad

Al igual que en el caso anterior, se ha analizado la influencia de la ortogonalidad de los ejes de la imagen en condiciones semejantes a las que se han hecho los experimentos sobre datos reales. Se ha encontrado que el error en la reconstrucción de los puntos de los métodos de Tsai y Tsai-LM se ve afectado por este parámetro, mientras que en el resto la incidencia es nula (figura 3.23).

Por otro lado, se observa que la estimación que realizan los diferentes métodos de este parámetro es bastante buena. Sólo el método de Zhang diverge de estos resultados en presencia de ruido, ya que, si bien la media obtenida es buena, tiene una desviación típica elevada, próxima a un grado, mientras que el resto no sobrepasa la décima de grado.

3.9.4. Reproducción de los experimentos reales con condiciones aproximadas

Hasta ahora, a lo largo de esta sección se ha ido analizando los resultados obtenidos sobre una cámara real mediante su reproducción parcial en el simulador, es decir, en cada experimento sólo se han estudiado aspectos concretos, intentando encontrar explicación a ciertos fenómenos.

La siguiente prueba lo que se plantea es reproducir en conjunto las condiciones de los datos reales para observar si los efectos producidos son los mismos. En concreto, se trata de reproducir los resultados obtenidos al calibrar el Conjunto 8x8 (figuras 3.3 - 3.6). Los valores promedio obtenidos en cada parámetro después de repetir en cada nivel de zoom 25 veces el experimento, así como los errores cometidos en la estimación de éstos, se muestran en las figuras 3.24 - 3.28.

Para realizar este experimento se han utilizado valores próximos a los devueltos por los métodos. En concreto, se ha utilizado la siguiente configuración de cámara: el centro óptico ligeramente desplazado del centro de la imagen (383,292), $k_1 = -0,001$, el factor de proporción próximo a la unidad (0.991) y la ortogonalidad de los ejes del plano imagen pequeña ($\Omega = 89,95$). A las coordenadas 2D de la imagen se les ha inyectado un ruido gaussiano de media cero y desviación típica 0.1 pixels, a las coordenadas 3D del patrón en forma de diedro un ruido de desviación típica 0.4 mm y a las coordenadas 3D de los patrones en forma de plano un ruido de 0.2 mm. Debido a que se ha utilizado una distancia focal efectiva inferior a la nominal en un 8%, los patrones utilizados se han acercado a la cámara unos centímetros en relación a los experimentos anteriores, de cara a que ocupen la misma porción de imagen que en los experimentos con datos reales.

Es importante resaltar el hecho de que se muestra el promedio de los valores obtenidos en cada una de las repeticiones del experimento. Esto produce que los resultados sean mucho más moderados que los que se puedan producir en una sola muestra como ocurre en los datos reales, es decir, que los resultados sean más próximo al valor utilizado en la configuración de la cámara. Por tanto, lo que es más interesante es analizar, en circunstancias próximas a las que se ha realizado la calibración real, que método calcula de forma más fiable (con menor error) cada parámetro y no los valores promedios obtenidos.

Los resultados obtenidos en la estimación del centro óptico concuerdan bastante con los obtenidos en los datos reales, ya que parece mantenerse la tendencia a separarse del centro de la imagen a medida que aumenta la distancia focal en todos los métodos excepto los de Batista y Tsai. Como excepción aparece que la curva del método de Heikkilä tiene un rango más próximo al resto de los métodos, y que las curvas de los

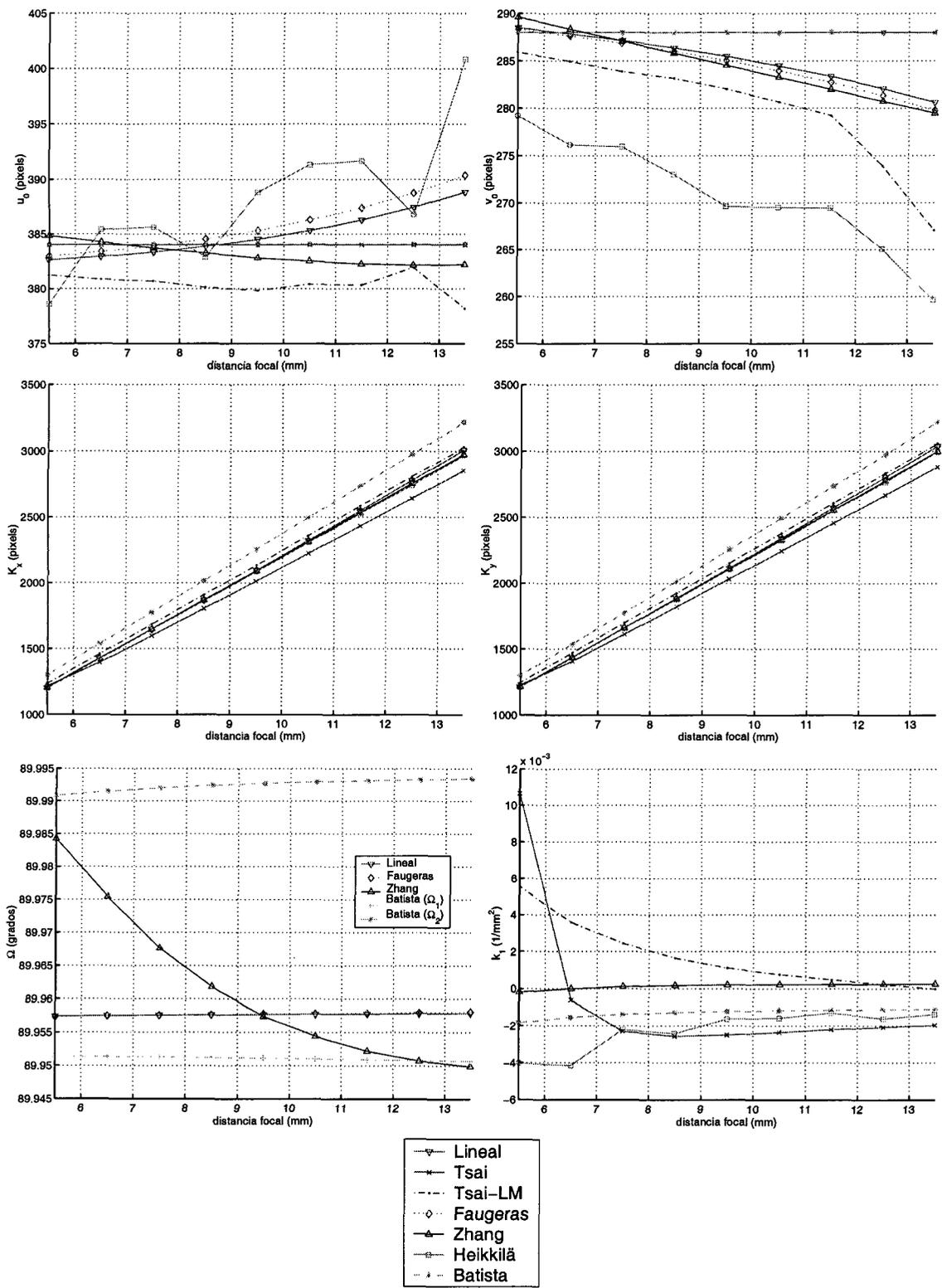


Figura 3.24: Valor promedio en los parámetros intrínsecos obtenidos de la variación del zoom usando una configuración de cámara semejante a la utilizada para obtener los datos reales.

métodos de Zhang y Tsai-LM tienen un comportamiento más homogéneo. Esto último puede deberse a que la gráfica muestra el valor promedio y no el valor de una sola muestra.

En la estimación de los factores de escalado se observa que el método de Batista dista del resto de los métodos, de manera semejante a lo que ocurre con los datos reales. En la estimación del coeficiente de distorsión aparece el mismo efecto que con los datos reales, es decir, un decremento exponencial de la magnitud a medida que aumenta la distancia focal.

En las gráficas del error de los parámetros intrínsecos (figura 3.25 se observa que los métodos que mejor estiman el centro del eje óptico en estas circunstancias son los de Tsai y Batista, ya que en la configuración de cámara utilizada este parámetro tiene un valor muy próximo al centro de la imagen. Por contra, los métodos de Zhang y de Heikkilä se muestran como los más imprecisos.

El factor de escalado muestra un error promedio elevado lo que indica que ninguno de los métodos es muy fiable. Los que presentan un mejor resultado son los métodos de Faugeras y Lineal con un error que ronda el 2%. En cambio, el método de Batista es el menos exacto debido a que la distancia focal efectiva dista mucho de la distancia focal nominal.

En el caso de la ortogonalidad el método que peor estima este parámetro es el método de Zhang. Por contra, este método, junto con el método de Batista, son los que mejor calculan el primer coeficiente de distorsión radial.

En la estimación de la posición del patrón sólo cabe destacar el comportamiento del método de Tsai en T_z que tiene un decrecimiento exponencial a medida que aumenta la distancia focal, resultado que no coincide con el de los datos reales. Este efecto, aunque en menor medida parece reproducirse en la estimación de R_x y R_z .

Por último, aunque las gráficas de los errores de reconstrucción no coinciden en magnitud, sí en la forma, observándose un continuo crecimiento del error en las coordenadas 2D y una estabilidad en el error 3D. Una excepción la presenta el método de Tsai que reproduce los efectos en la estimación de los parámetros extrínsecos.

3.10. Comparación de los resultados a partir de datos reales con los de datos simulados

En los resultados obtenidos a partir de una cámara real no se puede calcular el nivel de error producido en cada uno de los parámetros debido a que se desconoce el valor real de éstos. Sólo es posible comparar los resultados obtenidos en el simulador con

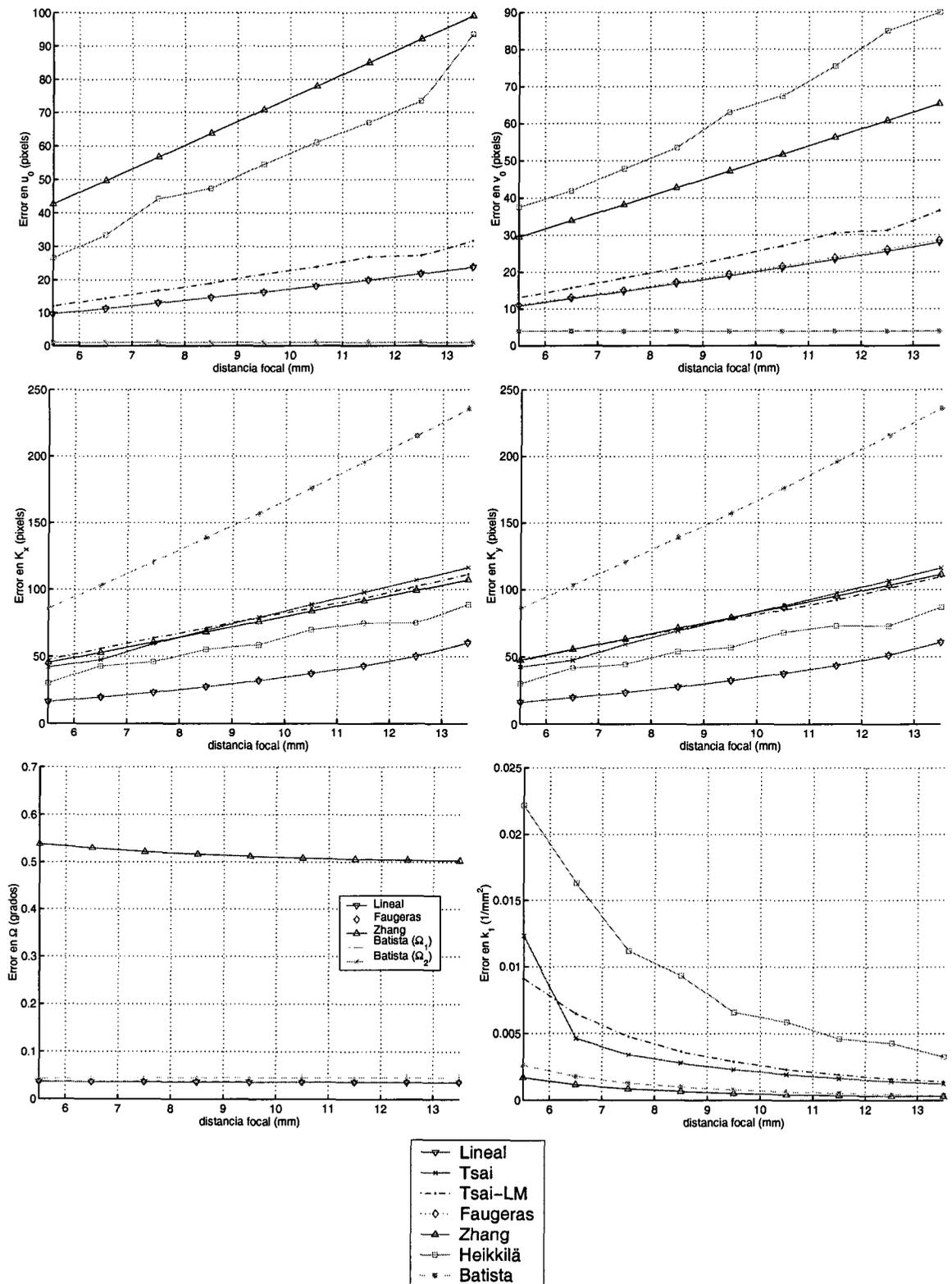


Figura 3.25: Error en la estimación de los parámetros intrínsecos obtenidos de la variación del zoom usando una configuración de cámara semejante a la utilizada para obtener los datos reales.

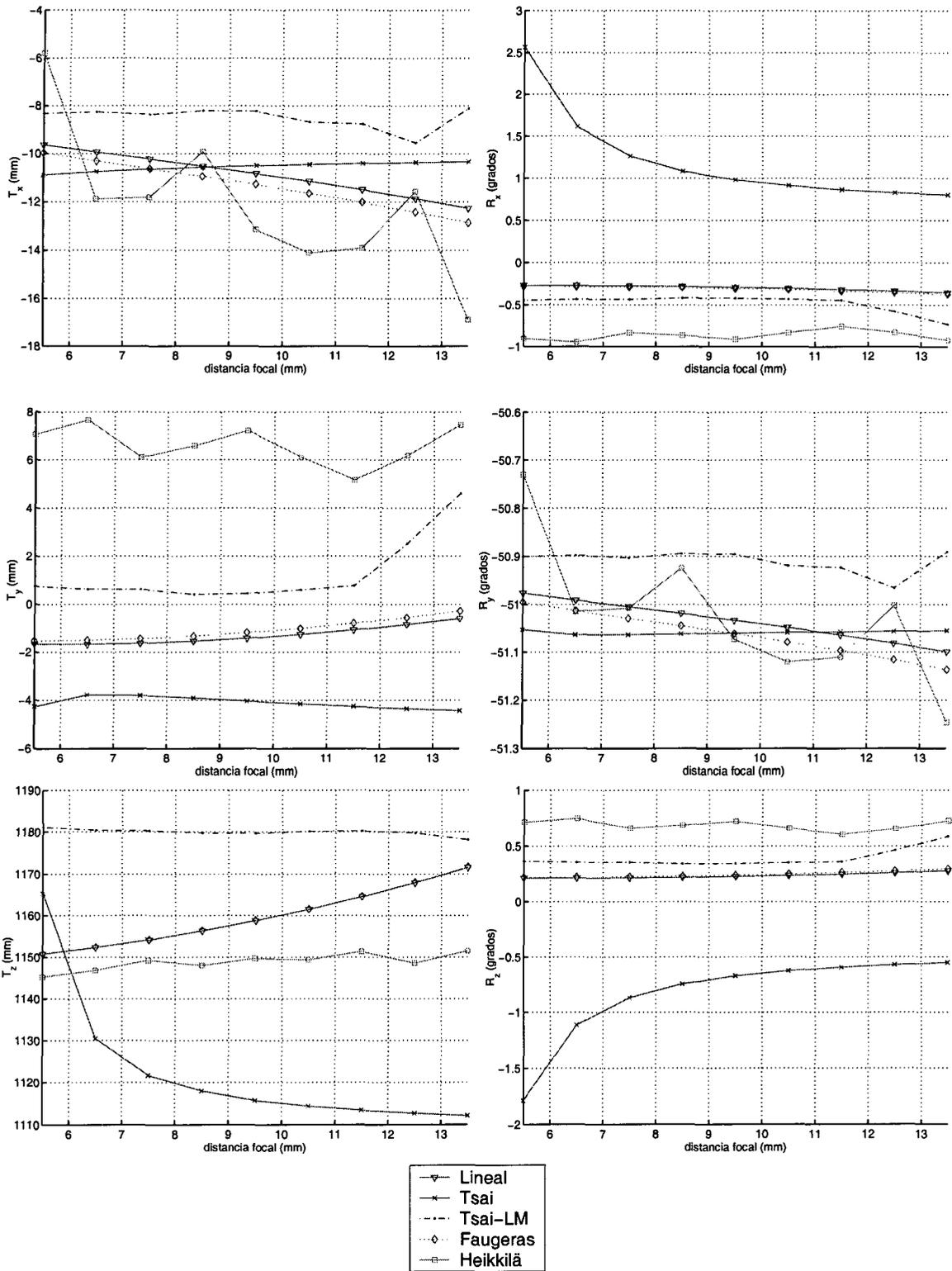


Figura 3.26: Valor promedio en los parámetros extrínsecos obtenidos de la variación del zoom usando una configuración de cámara semejante a la utilizada para obtener los datos reales.

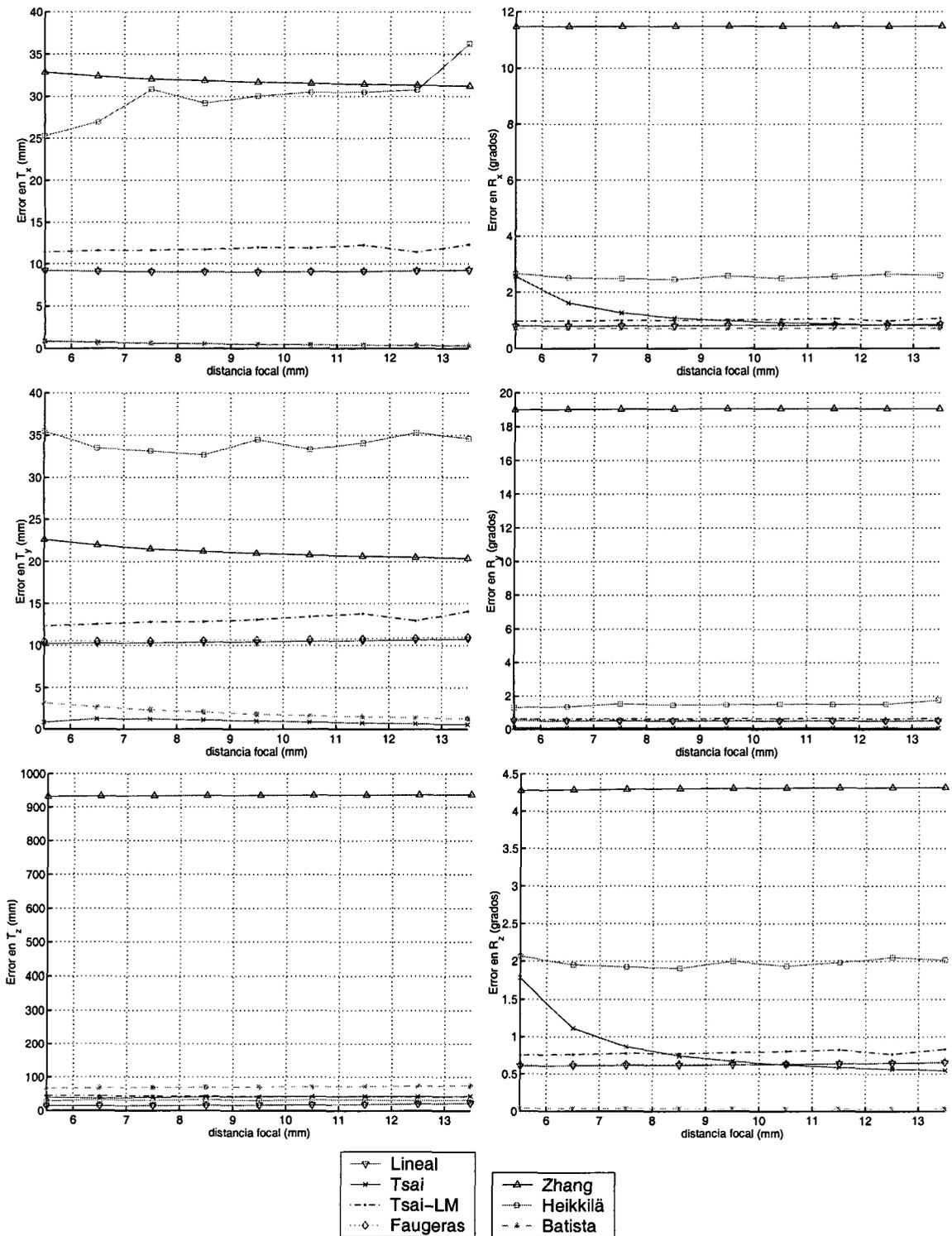


Figura 3.27: Error en la estimación de los parámetros extrínsecos obtenidos de la variación del zoom usando una configuración de cámara semejante a la utilizada para obtener los datos reales.

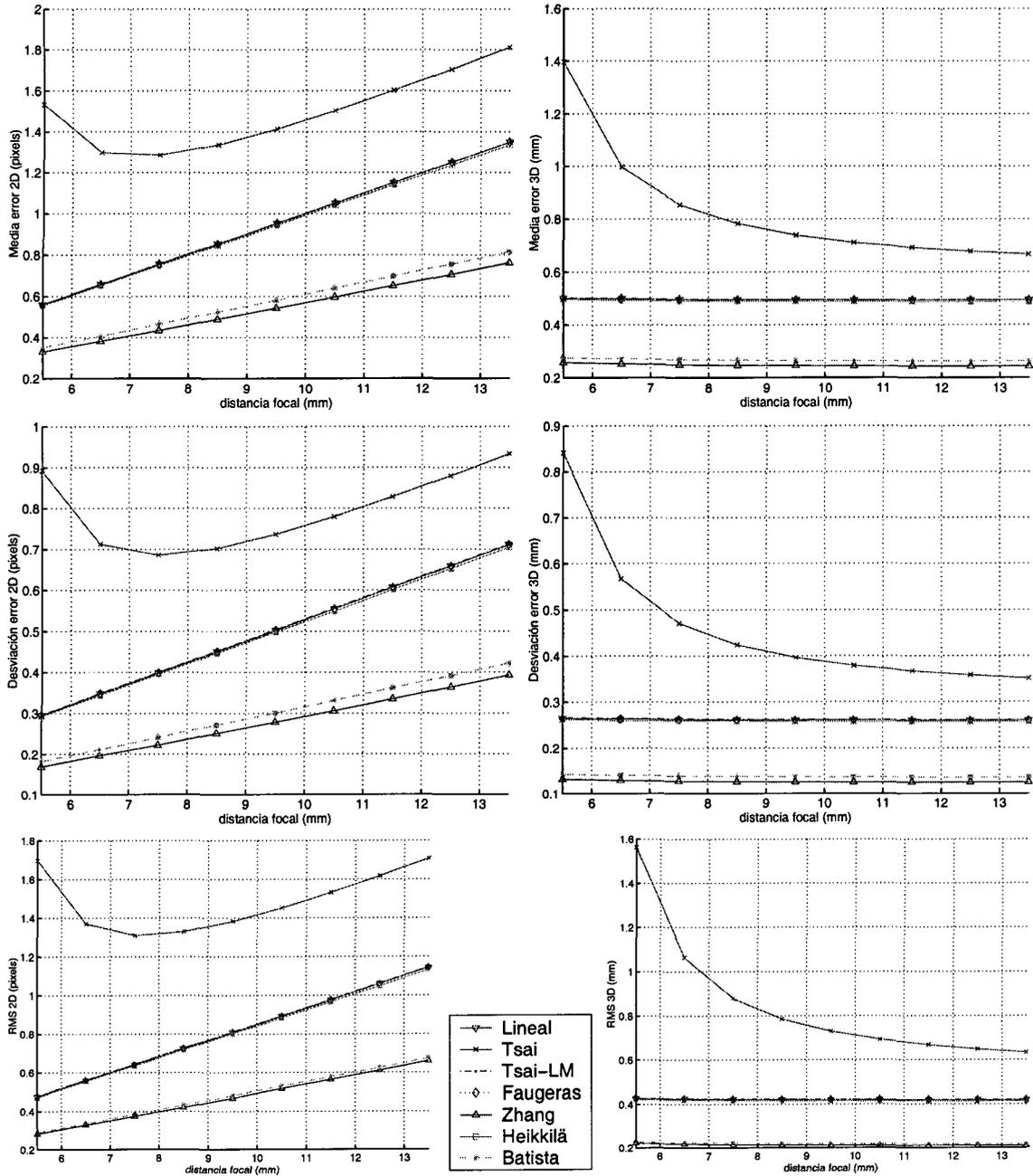


Figura 3.28: Error en la reconstrucción de las coordenadas 2D (izquierda) y 3D (derecha) de los puntos del patrón obtenidos de la variación del zoom usando una configuración de cámara semejante a la utilizada para obtener los datos reales.

aquellos que se refieren al error promedio en la reconstrucción tanto de los puntos 2D como 3D. En vista de lo cual se ha realizado un análisis de estos errores en uno y otro caso. En los siguientes párrafos se comentan algunos aspectos destacables obtenidos de la comparación de estos errores.

En la reconstrucción usando los datos obtenidos del primer patrón, un diedro con 8×8 puntos en cada plano que es el mismo que se ha utilizado en el simulador, el nivel de error en coordenadas 2D está en torno a 1-1.2 pixels y en coordenadas 3D sobre medio milímetro. En los resultados obtenidos con el simulador (figuras 2.12 y 2.15) se consiguen estos niveles de error cuando el ruido en las coordenadas 2D está en torno a los 0.75 pixels, o cuando el ruido en las coordenadas 3D es inferior al medio milímetro. Aunque es obvio que con la cámara real este nivel de error se produce por la confluencia de varios factores y no sólo por uno de ellos, como se analiza en cada uno de los experimentos realizados en el simulador. Por tanto, estos valores sirven para establecer una cota superior para el nivel de ruido. Además se observa que los métodos que utilizan planos (Zhang y Batista) presentan un error menor que el resto.

Posteriormente, en el simulador se han reproducido las condiciones con las que se ha realizado la calibración con datos reales. Estos experimentos han demostrado que, en la calibración realizada, la incidencia del ruido en las coordenadas 3D del patrón es mucho mayor que la producida por el ruido en las coordenadas 2D. Esta conclusión se alcanza después de analizar la forma de las curvas del error en la reconstrucción de los puntos, tanto en 2D como 3D.

El hecho de que el método de Zhang sea el que mejor resultado obtenga al realizar la reconstrucción de los puntos es indicativo de que el grado de distorsión de las imágenes no es elevado. Esto se debe a que este método degrada la calidad de su calibración cuando la magnitud de k_1 es superior a $5 \cdot 10^{-3}$. Esta hipótesis se apoya además en que los métodos de Faugeras y Lineal tienen valores próximos al resto en la estimación de los factores de escalado. Además, los valores del error 2D y del error 3D de estos métodos son próximos a otros que utilizan diedro como patrón.

En general, se observa que en el simulador se reproducen muchos de los efectos producidos en la calibración de la cámara real, permitiendo conocer qué factores son los que han incidido en mayor medida en los resultados.

3.11. Estudios realizados por otros autores

En este apartado, se comparan, en primer lugar, los resultados que se han presentado en este trabajo con los métodos estudiados por X. Armangué and J. Salvi en [Salvi et al., 2002] y [Armangué et al., 2000]. A continuación se revisan las conclusiones

de Lai [Lai, 1993], sobre el efecto de algunos parámetros internos sobre el resultado de los parámetros externos, a la luz de los resultados aquí obtenidos.

Para el primer caso, se han repetido en el simulador algunos de los experimentos usando el código generado por X. Armangué and J. Salvi (el cuál se puede obtener en: "http://eia.udg.es/%7Earmangué/research/"). Este código incluye una serie de métodos como son los de Hall [Hall et al., 1982], Weng [Weng et al., 1992] y varias versiones de los métodos de Faugeras y de Tsai. La descripción de estos métodos se puede ver en [Salvi et al., 2002] o [Armangué et al., 2000]. Se han seleccionado dos experimentos de los realizados para el presente trabajo, los cuales pueden permitir observar el comportamiento de estos otros métodos. Estos experimentos han sido realizados en las mismas condiciones que los descritos en la introducción a los experimentos. En primer lugar se ha elegido el primer experimento porque muestra la incidencia del ruido en una de las entradas de datos (coordenadas 2D) sobre los resultados. Y en segundo lugar, se ha seleccionado el experimento en el que se varía la distancia del patrón usando imágenes distorsionadas porque muestra la influencia de otros dos tipos de fuentes de error, por un lado la inclusión de un parámetro intrínseco que no tienen en cuenta todos los métodos (la distorsión de la lente) y por otro lado la variación de uno de los parámetros extrínsecos (la distancia del patrón).

Sobre estos métodos se ha variado el nivel de ruido en las coordenadas 2D de la imagen (repetición del primer experimento). Como resultado llamativo se comprobó que en los niveles de ruido más altos (por encima de 1.2 pixels de ruido). En los métodos de Faugeras con optimización de Newton Raphson y de Weng se disparaba con frecuencia el error debido a que los métodos no convergían. La versión de Faugeras-Toscani es ligeramente peor que la utilizada en este artículo, en la que se utiliza como aproximación inicial el resultado del método Lineal. En el caso del método de Tsai, usando el método de optimización de Newton-Raphson el resultado es semejante al ofrecido en este trabajo por el método de Tsai usando el método de optimización de Levenberg-Marquardt.

Como se puede observar en la tabla 3.3 los métodos empleados en [Salvi et al., 2002] tienen, en general, peor resultado cuando existe ruido en las imágenes que los métodos utilizados en el presente trabajo, excepto el de Heikkilä. En concreto, el que peor se comporta es el método de Weng. También tiene un comportamiento peor que el resto la versión de Faugeras con optimización de Newton-Raphson.

También se ha repetido sobre este conjunto de métodos el experimento en el que se realiza una variación de la distancia del patrón con imágenes que tienen un valor de $-0,005$ en el primer coeficiente de distorsión. Los resultados se muestran en las figuras 3.29 y 3.30. Se comprueba que los métodos que se ven más influenciados por la presencia de la distorsión son los métodos de Faugeras-Toscani, que obtiene valores para

	u_0	v_0	k_x	k_y
Lineal	3.5	4.1	5	7.5
Tsai	-	-	6.5	9.7
Faugeras	3.9	4.7	4.7	7.3
Zhang	3.1	5.1	9.1	13.3
Batista	2.2	1.6	7.9	3.8
Heikkilä	12.6	18.5	9.4	14.5
Ahmed	4.1	4.8	4.6	7.2
Tsai-LM	9.8	5	11.1	17.4
Tsai-NR	10.6	4.1	10.8	16.8
Faugeras-NR	18.9	11.8	34.9	53.2
Faugeras-Toscani	4.5	5.7	5.6	8.7
Weng	43.7	21.1	74.2	112.3

Tabla 3.3: Error promedio en los resultados de algunos parámetros intrínsecos (u_0, v_0, k_x, k_y) de cada método cuando el nivel de ruido en las coordenadas 2D de la imagen es de 0.9 pixels.

los parámetros semejantes a los obtenidos por los métodos de Ahmed, Lineal y Faugeras (figura 2.34), y de Weng, aunque en menor medida que el anterior experimento. En los demás, el efecto es prácticamente nulo. Como ocurría en los métodos utilizados en el presente trabajo, el error es mayor cuanto más cerca está el patrón de la cámara.

Por otro lado, Lai [Lai, 1993] analiza la incidencia de los errores en algunos parámetros intrínsecos (centro óptico, distorsión o factor de proporción) en los parámetros extrínsecos y concluye que la distorsión de la lente y el centro óptico tienen un efecto pequeño sobre la estimación de la posición y orientación del patrón, siendo mayor la influencia del centro óptico que la de la distorsión. También obtiene que el factor de proporción tiene una gran influencia sobre el cálculo de la posición del patrón.

Se puede observar que algunos de los resultados del presente trabajo contradicen las conclusiones obtenidas por él. Como se ha podido observar en los distintos experimentos existe una estrecha relación entre el centro del eje óptico y la posición del patrón, en concreto con T_x y T_y . Se puede comprobar cómo las gráficas obtenidas para u_0 se corresponden en forma a las de T_x y las de v_0 con las T_y . Las gráficas con los niveles de error obtenidos para estos parámetros tienen la misma forma, y por tanto, la magnitud del error de T_x y T_y estará directamente relacionada con la distancia a la que se encuentra el patrón y del error cometido en el centro óptico. Este resultado contradice a Lai, ya que en su trabajo se concluye que el centro del eje óptico apenas incide en la obtención de la posición del patrón.

En cambio, los experimentos realizados en este trabajo sí parecen confirmar, aunque con matices, otra de las conclusiones de Lai. En su artículo se indica que el patrón debe

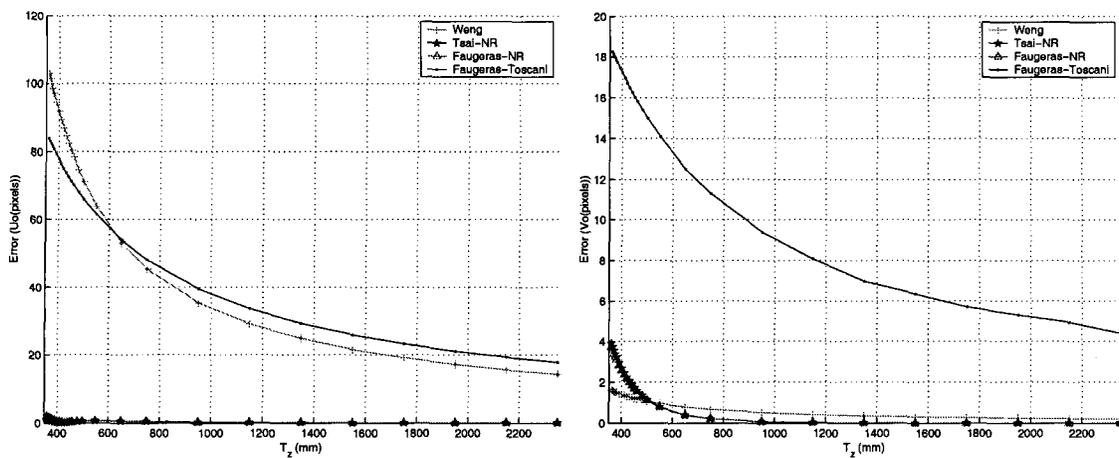


Figura 3.29: Error en el cálculo del centro del eje óptico ante la variación de la distancia del patrón en imágenes con distorsión.

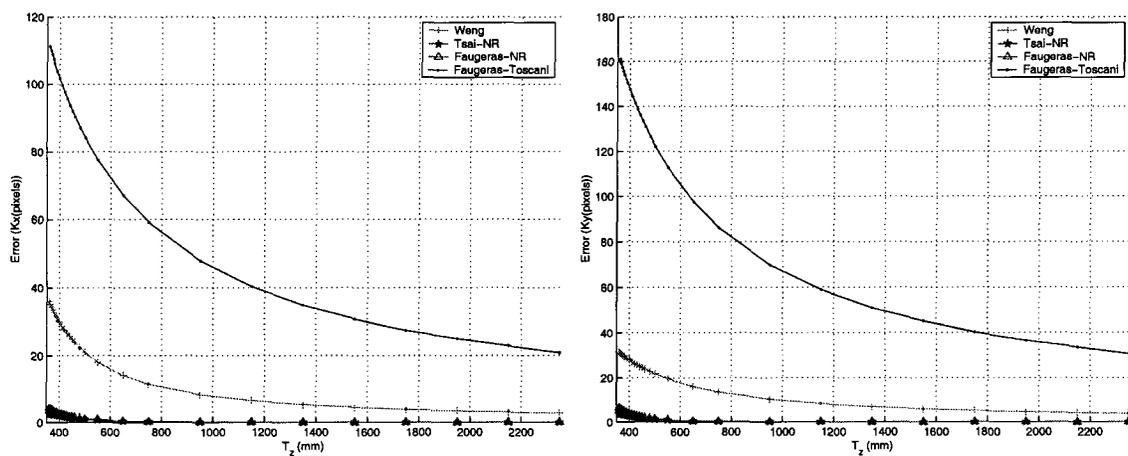


Figura 3.30: Error en el cálculo de los factores de escalado ante la variación de la distancia del patrón en imágenes con distorsión.

estar colocado de forma que ocupe la mayor proporción de imagen posible. Los resultados descritos en el presente trabajo de los experimentos realizados sobre la variación de la distancia del patrón respecto a la cámara indican que en presencia de distorsión es mejor colocar el patrón alejado de la imagen, pero que si la imagen de referencia tiene ruido gaussiano, es mejor colocar el patrón lo más cerca que sea posible. También se observa en el último de estos experimentos que la influencia del ruido es, en general, mucho mayor que la debida a la distorsión, al menos con la configuración de cámara utilizada. Por tanto, se concluye que la posición del patrón debe observar el siguiente compromiso: debe estar cerca de la cámara, pero si la distorsión es muy notoria, no debe ocupar toda la imagen, evitando en lo posible los bordes de la misma donde más acusado es el efecto de la distorsión.

3.12. Conclusiones

En este capítulo se muestran los resultados obtenidos en la calibración con los métodos estudiados a partir de varios conjuntos de muestras obtenidas con una cámara real. Después de analizar los errores cometidos por cada método en la reconstrucción de los puntos se observa que los métodos que realizan de forma más fiable esta reconstrucción son los métodos de Zhang y Batista. Este resultado es relevante, ya que estos dos métodos son los métodos que trabajan con planos en vez de usar un diedro. Este resultado es lógico ya que un diedro presenta un mayor número de posibles fuentes de ruido, como son la precisión con que se obtiene el ángulo que forman los dos planos o la ubicación del origen de coordenadas común.

Debido a que cuando se trabaja con una cámara real se desconocen los valores reales de los parámetros, en principio, no hay forma de apreciar la precisión con que se estima cada parámetro. En cambio, se ha podido analizar la estabilidad de estos parámetros en diferentes situaciones. Se ha observado cómo el desplazamiento del patrón (variación de los parámetros extrínsecos) influye en los parámetros intrínsecos y se ha observado que los resultados (sección 3.3.2) presentan una gran dispersión, cuando teóricamente debería permanecer constantes. Por tanto, los resultados de la calibración, es decir, la obtención de los parámetros intrínsecos de una cámara, dependen en gran medida de las condiciones en que se haya realizado el proceso. La principal conclusión que se puede extraer de esto es que los procesos de calibración consiguen buenas estimaciones para el conjunto de imágenes que se ha tomado como referencia, pero estos resultados no son válidos en general si se traslada a otras situaciones en las que se cambie la disposición del patrón respecto a la cámara.

Por otro lado, se ha mostrado que es posible reproducir en el simulador muchos de

los efectos observados al aplicar los métodos de calibración estudiados sobre una cámara real. De esta forma se ha podido establecer una guía de pasos o criterios a seguir para establecer en qué condiciones se ha realizado la calibración y por qué ocurren los efectos que aparecen en ésta.

Como se ha visto, en primer lugar se puede analizar la forma de las curvas de los errores, lo cual da pie a averiguar qué fuente de datos, las coordenadas 2D de la imagen o las coordenadas 3D de los puntos del patrón, tiene mayor incidencia en los resultados, y obtener una cuantificación aproximada de su valor. Esto puede ser útil, si la incidencia es muy alta, para saber qué conjunto de datos se debe corregir. Por ejemplo, para obtener un patrón donde se conozca con más precisión los puntos si el mayor ruido proviene de las coordenadas 3D.

Otra cuestión que se puede averiguar es el grado de distorsión de las imágenes analizando los valores de los factores de escalado de métodos como los de Faugeras y Lineal. Esto es debido a que como no modelan la distorsión, compensan su ausencia variando significativamente, respecto al resto de los métodos, el valor de los factores de escalado. Es decir, si en los resultados obtenidos de la calibración de una cámara se observa que, la estimación de los factores de escalado obtenida por estos métodos dista significativamente de los valores obtenidos por todos los demás métodos, es indicativo de que la distorsión de las imágenes es muy acentuada.

Por otro lado, la relación en la estimación de los factores de escalado obtenidos por el método Batista respecto al resto de los métodos, puede ser indicativo de la relación entre la distancia focal efectiva y la nominal aportada por el fabricante. Esto se debe a que al analizar las gráficas correspondientes a los factores de escalado se observa que el resultado de este método difiere bastante de todos los demás. El único que en determinados casos da un resultado semejante es el método de Tsai debido a que mantiene fijo el valor del centro óptico.

También se observa que, excepto el método de Zhang, el resto obtiene con bastante precisión el ángulo que forman los ejes del plano imagen (asimetría). En resumen, la unión de todas estas pistas es un indicativo de las condiciones en las que se ha realizado la calibración con datos reales.

Los experimentos realizados también han mostrado que el simulador construido se ha mostrado como una herramienta eficaz ya que ha permitido la reproducción de muchos de los efectos producidos sobre la calibración de una cámara real.

Por último, se han comparado los resultados del presente trabajo con los obtenidos en otros estudios de similares características, pero en los que sólo se analizan algunos de los aspectos recogidos en este trabajo. Respecto a los trabajos [Salvi et al., 2002] y [Armangué et al., 2000] se ha podido comparar los métodos utilizados en su comparativa

con los presentados en este trabajo debido a la disponibilidad de sus códigos en Internet. Se ha observado que los métodos utilizados en el presente trabajo demuestran ser más precisos. Por otro lado, hay que indicar que una de las conclusiones del presente trabajo, la correlación entre los resultados de parámetros intrínsecos y extrínsecos, contradice los resultados obtenidos por [Lai, 1993]. En cambio, sí existe coincidencia respecto a la disposición que debe tener el patrón en la imagen para que la calibración tenga los mejores resultados, es decir, el patrón debe colocarse de forma que abarque el mayor campo de imagen posible, salvo en condiciones de distorsión severa donde debe alejarse un poco para que su tamaño en la imagen disminuya.

Capítulo 4

Métodos de autocalibración

4.1. Introducción

Hasta ahora se ha realizado un estudio del funcionamiento de los métodos clásicos de calibración, los cuales se han basado en el uso de patrones de puntos conocidos. Un problema que presentan estos métodos es que no permiten una calibración on-line cuando se está utilizando la cámara en tareas visuales. Este hecho, añadido a que la configuración interna de una cámara puede variar (movimiento del zoom, enfoque, desajustes mecánicos, variaciones térmicas, etc.) durante la ejecución de dichas tareas, hace que en los últimos años se haya orientado una importante línea de trabajo al desarrollo de métodos que permitan una calibración automática, es decir, métodos de autocalibración.

Estos métodos no cuentan con puntos de referencia en los que se conoce la situación exacta de cada uno de ellos. En su lugar, se utiliza la correspondencia de puntos detectados en la escena a lo largo de una secuencia de imágenes. Esto permite que un sistema que lleve incorporada una cámara pueda realizar de forma automática la calibración de ésta en cualquier momento, ya que simplemente necesita adquirir una secuencia de imágenes de su entorno.

Los primeros trabajos en este sentido [Maybank and Faugeras, 1992] [Luong et al., 1992] [Hartley, 1994b] se basaron en la suposición de que la configuración de las cámaras con ópticas no motorizadas permanece constante durante largo tiempo y que sólo desajustes de tipo mecánico o térmico afectan a las cámaras. Por tanto, sólo se hacía necesaria la recalibración cada cierto tiempo para compensar estas pequeñas variaciones. Estos trabajos han aportado la ventaja de que el proceso de calibración se puede realizar de forma automática, permitiendo su realización en paralelo a otras tareas de visión.

Posteriormente se observó que estos métodos no cubrían algunas de las necesidades

que se empezaban a plantear. En concreto, el uso de cámaras con zoom variable o con el enfoque automático, hacían necesario el desarrollo de nuevos métodos, ya que, los parámetros intrínsecos de estas cámaras pueden variar constantemente. Las adaptaciones de los métodos anteriores a esta nueva circunstancia ha propiciado la aparición de nuevos métodos que puedan calibrar una cámara aunque entre una imagen y otra haya cambiado parte de la configuración de la cámara o incluso toda ella.

Sobre los métodos de autocalibración se puede realizar distintos tipos de clasificaciones: según el tipo de movimiento que es necesario realizar y según si los parámetros intrínsecos pueden variar o no.

- Tipos de movimientos

General: Permiten cualquier tipo de desplazamiento, orientado por tanto a cámaras situadas sobre plataformas móviles. En algunos casos, pueden tener problemas con movimientos específicos, por ejemplo, con traslaciones o rotaciones puras.

Rotaciones Puras: Orientados a cámaras fijas como pueden ser, por ejemplo, cámaras de vigilancia o para retransmisiones deportivas. Sólo se permite realizar giros, habitualmente azimut y elevación.

Traslaciones Puras: Sólo se efectúan traslaciones, por tanto, no existen cambios en la orientación, lo que permite simplificar las ecuaciones. En general, se intenta buscar que los desplazamientos sean ortogonales entre sí.

En un plano: Útil en aquellos casos en los que los desplazamientos se realizan en un mismo plano, aunque se permiten tanto rotaciones como traslaciones. Están orientados a cámaras situadas sobre objetos que se desplacen sobre una superficie plana.

- Parámetros intrínsecos:

Fijos: Son los métodos más antiguos, en los cuales se supone que la configuración de la cámara permanece estable entre las diferentes tomas de imagen. En el caso de que haya alguna variación se hace necesario repetir todo el proceso de calibración.

Variables: En general, se estudia principalmente el efecto de la variación del zoom. En algunos métodos se presupone que el zoom sólo afecta a la distancia focal, por lo que sólo este parámetro junto con los extrínsecos se consideran distintos de unas imágenes a otras. En otros, se supone que el zoom también

puede afectar al valor del centro del eje óptico, pudiéndose variar este parámetro también. Existen algunos métodos que toleran que otros parámetros pueden variar de una imagen a otra (ratio del pixel, coeficientes de distorsión, etc.).

El objetivo de este capítulo va a ser el estudio de los métodos de autocalibración para cámaras estáticas, es decir, cámaras que no se desplazan, aunque pueden tener movimientos de rotación. En este estudio se presentan los resultados de diversos experimentos tanto con datos simulados como con datos reales. Se han comparado cuatro métodos de autocalibración catalogados dentro del grupo que solamente realiza rotaciones: el método de McLauchlan [McLauchlan and Murray, 1996], el método de Seo [Seo and Hong, 1998] y el método de Agapito, del que se han implementado dos versiones, una lineal [de Agapito et al., 1999] y otra iterativa [de Agapito et al., 1998]. Se han utilizado tanto métodos que suponen constantes los parámetros intrínsecos durante la calibración como los que admiten la variación de éstos en la secuencia de imágenes utilizada para calibrar.

El estudio se centrará en analizar la influencia de diversos factores sobre los resultados de la calibración. Por un lado, aspectos propios de la configuración de la cámara, por otro lado, aspectos de la escena utilizada para el proceso de calibración y también aspectos propios del tipo de métodos utilizado en la comparativa. El estudio compara la precisión, la robustez y el coste computacional de cada uno de los métodos.

Respecto a la configuración de la cámara se examina como influyen factores que no son modelados en la mayoría de los métodos de autocalibración como es el caso de la distorsión, la ortogonalidad de los ejes del plano imagen o el factor de proporción. Hay que tener en cuenta que en la mayoría de los métodos se necesita que los ejes del plano imagen sean perfectamente ortogonales y que el factor de proporción sea conocido o prefijado, ya que esto permite aumentar el número de restricciones y compensar el aumento del número de parámetros que presenta este tipo de métodos. Además, se incluye un análisis de la incidencia de un elemento derivado del movimiento de la cámara, como es el desalineamiento entre los ejes de giro de la cámara y el punto focal.

El coste computacional de cada uno de los métodos se va a ver directamente influenciado por el número de imágenes que necesita cada método para que converjan sus resultados, y por la cantidad de puntos en correspondencia que sean procesados por imagen. En este análisis se estudiará también la influencia del ruido de cuantización de las imágenes o ruido 2D, y por último, se comparan los resultados de los métodos clásicos de calibración con los obtenidos en los experimentos sobre los métodos de autocalibración.

Este capítulo se organiza en varios apartados. En primer lugar, se realiza una revisión

bibliográfica de los métodos de autocalibración más destacados y de otros trabajos relacionados con el tema. Los métodos son catalogados siguiendo los tipos de clasificación descritos previamente. En segundo lugar, se realiza una descripción de los algoritmos correspondientes a los métodos de autocalibración implementados para realizar el presente estudio y se presentan los resultados de los experimentos realizados en un simulador. Posteriormente, estos experimentos se repiten sobre un sistema real y se presentará un análisis de sus resultados. Una vez realizados todos los experimentos se comparan los resultados obtenidos con los obtenidos por los métodos clásicos. Finalmente, se detallan las conclusiones más importantes sobre esta tipología de métodos de autocalibración de cámaras.

4.2. Revisión bibliográfica

A continuación se hace un recorrido histórico por aquellos trabajos más conocidos dentro del campo de la autocalibración de cámaras, y se lleva a cabo una catalogación en función de las características descritas previamente.

4.2.1. Movimiento en general

Parámetros intrínsecos fijos

En primer lugar se van a analizar los métodos que permiten cualquier tipo de desplazamiento, y dentro de éstos los que mantienen fijos los parámetros intrínsecos. Los métodos encasillados en esta subcategoría pueden dividirse a su vez en dos subgrupos según la metodología utilizada para resolver el problema.

Por un lado, existe un conjunto de métodos en los que la calibración se realiza en dos pasos. En primer lugar se obtienen las transformaciones epipolares que permiten describir las relaciones entre varias cámaras (o vistas). A continuación, se utilizan las ecuaciones de Kruppa para enlazar estas transformaciones con la imagen de la cónica absoluta, de las que se pueden extraer los parámetros intrínsecos de la cámara. En estos trabajos se obtiene la calibración de la cámara directamente a partir de la correspondencia de puntos de al menos tres vistas de la escena obtenidas desde distintas posiciones.

Entre los métodos que se basan en este algoritmo destacan [Maybank and Faugeras, 1992] y [Luong et al., 1992], que fueron los trabajos que sirvieron de punto de partida para el desarrollo de los métodos de autocalibración. En [Luong and Faugeras, 1997] se sigue la misma filosofía que en los anteriores métodos. En este trabajo se indica que el punto principal y el ángulo de los ejes del plano imagen (Ω) deben ser prefijados ya que inducen mucho error, además en los experimentos

realizados se observa que los parámetros intrínsecos se obtienen de forma menos precisa que los extrínsecos. El trabajo de Zeller [Zeller and Faugeras, 1996] es una generalización de [Luong et al., 1992] para una secuencia entera de vídeo y mejora su robustez debido a que usa herramientas estadísticas (matrices de covarianzas, etc). El método propuesto por Lourakis [Lourakis and Deriche, 1999] se diferencia de los anteriores en que utiliza la descomposición en valores singulares (SVD) de la Matriz Fundamental para simplificar las ecuaciones de Kruppa. De esta forma evita el uso de los epipolos que son los elementos que inducen una mayor cantidad de ruido. Los parámetros pueden ser recuperados usando un esquema de minimización no-lineal.

Por otro lado, existe otro conjunto de métodos en los que para obtener la reconstrucción métrica de la escena descomponen el proceso de calibración en varios pasos. En primer lugar, se obtiene la reconstrucción proyectiva de la escena, para lo que es necesario reconstruir la geometría epipolar (Matriz Fundamental y epipolos). A partir de las matrices de proyección de cada cámara se puede obtener la reconstrucción afín. Por último en base a ésta se puede calcular la reconstrucción métrica de la escena.

Dentro de este grupo se incluiría el método de Hartley [Hartley, 1994a]. Este método trabaja sobre una secuencia de varias imágenes y obtiene una reconstrucción métrica de la escena en los pasos mencionados. Propone un paso intermedio para pasar de la reconstrucción proyectiva a la afín que denomina reconstrucción cuasi-afín. Para obtener la reconstrucción afín se selecciona aquella reconstrucción cuasi-afín que permita obtener la matriz de parámetros intrínsecos. Para obtener la reconstrucción métrica utiliza un proceso iterativo (Levenberg-Marquardt [More, 1977]).

Pollefeys y Van Gool recurren a lo que denominan restricciones de los módulos para obtener la reconstrucción afín [Pollefeys et al., 1996b], para lo que necesitan al menos cuatro imágenes. Posteriormente, en [Pollefeys and Gool, 1997] utilizan una propiedad de la homografía infinita que aporta una restricción para obtener el plano en el infinito por cada par de imágenes, y no sólo respecto a una imagen de referencia como era el caso anterior, lo cual robustece el método y reduce a tres el número de imágenes necesarias. En [Heyden and Åström, 1996], Heyden y Åström simplifican el proceso pasando directamente de reconstrucción proyectiva a métrica usando una variante (respecto a [Luong et al., 1992]) de las ecuaciones de Kruppa, necesitando sólo 3 imágenes para conseguirlo. El último trabajo es el propuesto por Triggs [Triggs, 1997], que al igual que el anterior obtiene directamente la reconstrucción euclídea, pero utiliza el concepto de cuádrica absoluta para hacerlo. Es un método no-lineal y presenta problemas de convergencia y estabilidad.

Parámetros intrínsecos variables

En esta categoría se encuentran los métodos que, realizando movimientos de cualquier tipo, modelan la variación de parámetros intrínsecos durante la secuencia de imágenes utilizada para la calibración. Generalmente, cuando se realiza la consideración de que los parámetros intrínsecos varían durante el proceso de calibración es porque las cámaras utilizadas poseen zoom, y este mecanismo puede ser utilizado durante dicho proceso y también cuando la cámara dispone de autoenfoco. La cuestión que se plantea es qué parámetros varían cuando estos elementos cambian. En un intento de resolver esta duda Willson [Willson, 1994] realiza un análisis de la influencia del zoom y el enfoque sobre distintos parámetros. De las pruebas que realiza concluye que la variación de estos elementos no sólo afecta a la distancia focal sino que normalmente también afecta al centro del eje óptico. Propone como solución la creación de un mapa de los valores de los parámetros intrínsecos según los valores de los motores de zoom y enfoque.

Algunos métodos que se pueden incluir en este apartado son adaptaciones de otros presentados por los mismos autores y que se pueden ver en el punto anterior. Heyden en [Heyden and Åström, 1997] y [Heyden and Åström, 1998] presenta un método iterativo que denominan "Algoritmo de ajuste de paquetes" (bundle adjustment). En el primer caso [Heyden and Åström, 1997] se permite que varíe el centro óptico y en el segundo no [Heyden and Åström, 1998]. En este método es necesario evitar las traslaciones puras y los movimiento circulares para su correcto funcionamiento. Lourakis y Deriche en [Lourakis and Deriche, 2000] realizan una adaptación del procedimiento introducido en [Lourakis and Deriche, 1999], donde se utiliza la descomposición en valores singulares (SVD) de la Matriz Fundamental para simplificar las ecuaciones de Kruppa, al caso de que los parámetros intrínsecos varíen. Los parámetros se pueden recuperar usando un esquema de minimización no-lineal. En los experimentos suponen perfecta ortogonalidad y pixels cuadrados.

Pollefeys [Pollefeys et al., 1996a] [Pollefeys and Gool, 1999] utiliza un enfoque estratificado; es decir, recurre a la obtención de una proyección afín como paso intermedio para recuperar la reconstrucción métrica. Para esto usan lo que llaman las restricciones de los módulos. Para conseguir una buena inicialización requiere que el primer par de imágenes se obtenga con una traslación pura. Posteriormente, en [Pollefeys et al., 1999], muestra un método iterativo que explota las propiedades de la cuádrlica absoluta. Permite calibrar cámaras según el conocimiento que se tenga de los parámetros. En el caso extremo, sólo es necesario que los ejes del plano imagen sean ortogonales, pero se necesitan ocho imágenes para realizar la calibración. También propone un método lineal que puede ser usado en las primeras imágenes de la secuencia, pero se debe aumentar

el número de restricciones (se suponen conocidos algunos parámetros: centro del eje óptico, factores de proporción y focal). Los resultados de este método lineal sirven como aproximación inicial a los siguientes resultados. El segundo paso es un proceso iterativo y se basa en el uso de la cónica y la cuádrlica absolutas. El trabajo incluye mecanismos para la detección de movimientos críticos.

Hartley [Hartley, 1992] basándose en las propiedades de la matriz Esencial desarrolla un método en el que puede variar tanto el centro óptico como la distancia focal. Posteriormente en [Hartley et al., 1999] adapta el procedimiento para autocalibrar una secuencia de imágenes. Se simplifican las ecuaciones, y por tanto, el número de imágenes necesarias, al considerar los pixels cuadrados y el punto principal en el centro de la imagen. Sigue la línea marcada en [Hartley, 1994a] y es una extensión de [de Agapito et al., 1999] para el caso de cualquier tipo de movimiento. Se puede dividir la calibración en cuatro fases: reconstrucción proyectiva, obtención del plano en el infinito, obtención de la homografía infinita y reconstrucción métrica. Se utiliza un proceso iterativo en el cual se realiza una búsqueda exhaustiva (búsqueda en densidad) del plano en el infinito. Mendonça y Cipolla [Mendonça and Cipolla, 1999] extienden el método [Hartley, 1992] a varios casos: focal variable de un par estéreo, múltiples intrínsecos variables, y largas secuencias de imágenes.

Sturm [Sturm, 1997b] también supone que el centro se ve afectado por la variación del zoom. Muestra que los parámetros intrínsecos están interrelacionados y desarrolla un método en el que se realiza una pre-calibración donde se aprovecha esta circunstancia. La pre-calibración es un paso previo donde se intentan obtener las funciones que hacen depender el factor de escalado horizontal y el centro del eje óptico del factor de escalado vertical. Utiliza las ecuaciones de Kruppa para obtener varias estimaciones de la componente vertical del factor de escalado. Se selecciona la mejor de ellas y se obtiene el resto de los parámetros mediante las funciones obtenidas en el paso previo.

4.2.2. Desplazamientos en el mismo plano

Debido a lo costosos y sensibles que son los métodos que permiten cualquier tipo de desplazamiento se ha optado, en alguno casos, por aplicar algunas restricciones que simplifiquen los algoritmos y que disminuyan su sensibilidad al ruido. En esta línea aparecen los métodos que restringen el movimiento de la cámara, al menos durante el proceso de calibración, a un sólo plano. Es decir, sólo existen desplazamientos de la cámara sobre dos de los ejes de coordenadas contenidos en el plano y las rotaciones son exclusivamente alrededor del tercer eje (perpendicular al plano).

Luong y Faugeras [Luong and Faugeras, 1994] presentan un método de autocalibra-

ción que necesita tomar al menos tres vistas de la escena desde diferentes posiciones. La calibración se obtiene factorizando la Matriz Fundamental, en lugar de usar las ecuaciones de Kruppa. Un inconveniente de este método es que necesita una buena aproximación inicial, la cual puede ser obtenida de uno de los métodos que usan las ecuaciones de Kruppa.

Más recientemente, en [Faugeras et al., 2000] se desarrolla un método en el que primero se trata de obtener la proyección unidimensional de una cámara, usando para ello tensores trifocales y sólo se necesitan tres vistas de la escena. A partir de la proyección unidimensional se intenta transformar a 2D, con lo que se obtiene la calibración que se busca. El uso de tensores trifocales ya había sido utilizada con anterioridad por Armstrong en [Armstrong et al., 1996].

4.2.3. Movimientos específicos

Son métodos que necesitan determinados movimientos o conjunto de movimientos entre las tomas de imágenes. En la mayor parte de los casos los desplazamientos son traslaciones puras, es decir, no existen rotaciones. Esto permite simplificar la ecuaciones y hacer más sencillos los métodos. De hecho, se consigue así que estos métodos sean lineales. En este tipo de métodos resulta problemático, en determinados casos, el hecho de no poderse conocer con exactitud el desplazamiento realizado.

Dentro de este grupo se pueden encuadrar métodos como el de Ma [Ma, 1996] que necesita dos conjuntos de tres imágenes obtenidas realizando traslaciones perpendiculares entre sí; o el método de Yang [Yang et al., 1998] que necesita cuatro conjuntos de imágenes obtenidas de dos movimientos perpendiculares entre sí. Ambos métodos necesitan que los ejes del plano imagen sean ortogonales. Esta restricción es evitada por Li [Li et al., 2002], pero debe usar para ello la información planar existente en la escena y obtener al menos cinco conjuntos de muestras a partir de dos traslaciones ortogonales en un mismo plano.

En [Guohui and Baozong, 1998] se realiza la calibración en base a dos movimientos, primero una traslación pura y luego una rotación pura. El punto principal es obtenido previamente. Calculando el centro de la hipérbola que describe un punto cuando se rota la cámara sobre el eje X se obtiene v_0 (o sobre el eje Y para u_0). Para la autocalibración utilizan la estructura epipolar de pares de vistas.

En determinados casos hay tareas que se pueden realizar sin necesidad de conocer el valor de los parámetros que definen una cámara, o sea, no es necesario realizar la reconstrucción euclídea. Por eso en [Gaspard and Viéville, 1996] se plantean distintos niveles de reconstrucción de la escena (proyectiva, afín, euclídea) en función de la información

disponible. El algoritmo utilizado se basa en un análisis jerarquizado de la escena, en el que comprueban si pueden realizar la reconstrucción en alguno de los niveles mencionados. Para alcanzar una reconstrucción más completa que la proyectiva son necesarios determinados tipos de desplazamiento (rotaciones o traslaciones puras). En todo caso, es necesario conocer el factor de escalado y que los ejes del plano imagen sea ortogonales.

4.2.4. Sólo rotaciones

En esta sección se encuadran aquellos métodos desarrollados principalmente para la autocalibración de cámaras que pueden rotar, pero no desplazarse. Este es un caso bastante estudiado ya que muchas cámaras cumplen con estas especificaciones, como puede ser el caso de las cámaras de vigilancia o las que cubren eventos deportivos. El hecho de conocer que la traslación es nula permite eliminar tres parámetros del problema y, por tanto, simplificar los métodos. El problema que puede presentar este tipo de métodos es que la traslación no se puede asegurar que sea nula, ya que, debido a la propia mecánica de las cámaras su eje de giro no contiene al centro óptico. Por tanto, cuando se realiza una rotación sobre uno de los ejes aparece una pequeña traslación entre ambos puntos. Este efecto también se produce cuando se realiza un zoom con la cámara ya que existe un desplazamiento del centro óptico. La influencia de este desalineamiento será uno de los elementos analizados en este trabajo.

Parámetros intrínsecos constantes

El problema que presentan los métodos que suponen constantes los parámetros intrínsecos es que si varía alguno de estos parámetros obliga a reiniciar el proceso de calibración. En función de la información que se debe aportar al método se pueden distinguir dos subgrupos. Principalmente, la diferencia estriba en si se supone conocido, o no, el ángulo del desplazamiento realizado por la cámara, el cual se puede obtener de la lectura de los codificadores de posición de los motores.

Los primeros trabajos en esta línea suponían conocido este ángulo, como es el caso de [Du and Brady, 1993] [Stein, 1993] [Stein, 1995] [Viéville, 1994]. Por ejemplo Stein [Stein, 1993] trabaja con varios pares de vistas entre las que existe sólo una rotación alrededor de un eje. Tras obtener una aproximación inicial, y en base al error cometido, obtenido de la disparidad en la segunda imagen de cada par, se refinan los resultados de cada parámetro aplicando una búsqueda no-lineal. Se puede incluir la distorsión como parámetro en el modelo utilizado. Existen problemas si el eje de giro no es coincidente, o muy próximo, al centro de proyección. En el método de Vieville [Viéville, 1994] también se utiliza la rotación sobre un eje. Pero en su caso, la recuperación de los parámetros

se hace en dos pasos: primero se obtiene una estimación inicial mediante ajuste a un modelo polinomial (desarrollo de Taylor), y luego se minimiza una función objetivo (distancia pseudo-Mahalanobis entre puntos reales y estimados) mediante un método no lineal (Gauss-Newton).

Al igual que Stein, McLauchlan y Murray [McLauchlan and Murray, 1996] usan la idea de predecir la posición de los puntos en la siguiente imagen para después corregir el error rectificando los valores de los parámetros a calibrar. Para ello usa un filtro recursivo y, aunque necesita una aproximación inicial para los factores de escalado, los valores iniciales pueden distar hasta en un 50 % del valor real. Sólo se limita a obtener los valores de la rotación y de los factores de escalado, suponiendo conocido el resto. El proceso necesita un gran número de imágenes para converger a la solución y, como se mostrará es sumamente costoso en términos de memoria.

En el otro subgrupo cabe mencionar el trabajo de Hartley [Hartley, 1994b], que fue el primero en suponer desconocidos el valor de los ángulos de giro de la cámara entre vistas. El método necesita al menos tres vistas de la escena, aunque pueden ser menos si se realizan algunas suposiciones sobre los parámetros intrínsecos. Es un algoritmo no-iterativo y utiliza la factorización de Cholesky para extraer los parámetros intrínsecos, e indica que los resultados no distan mucho de la aplicación del método de optimización de Levenberg-Marquardt, aunque se puede usar éste para refinar los resultados. Este trabajo se ve ampliado en [Hartley, 1997b] donde describe el cálculo de la homografía entre imágenes, primero aplicando un método lineal y luego refinando el resultado mediante Levenberg-Marquardt.

Parámetros intrínsecos variables

En este apartado se encuentran los métodos que incluyen en el modelo la variación de parámetros intrínsecos durante la secuencia de imágenes utilizada para la calibración.

Tanto Agapito et al. [de Agapito et al., 1998] como Seo y Hong [Seo and Hong, 1998] [Seo and Hong, 1999] han desarrollado sendos métodos iterativos aplicando las restricciones que impone la homografía en el infinito. El primero usa Levenberg-Marquardt como método de optimización mientras que el segundo usa el método de Kanatani [Kanatani, 1998]. Todos estos métodos suponen fijos el punto principal y el factor de proporción, y los ejes del plano imagen son ortogonales.

Posteriormente, estos mismo autores han desarrollado variantes lineales de sus métodos [Seo and Hong, 2001] [de Agapito et al., 1999]. El primero de ellos supone conocido el centro del eje óptico, con lo que las ecuaciones se ven simplificadas y sólo con dos imágenes se puede realizar el proceso de calibración. En el caso de Agapito, desarrolla

primero un método lineal que se basa en el uso de la cónica absoluta y, al igual que en el caso anterior, sólo necesita dos imágenes si se conoce el centro del eje óptico. En segundo lugar, desarrolla un método completo [Agapito et al., 2000] que se descompone en varias fases: obtención de una aproximación inicial mediante el método lineal [de Agapito et al., 1999]; aplicación del método iterativo [de Agapito et al., 1998]; y refinamiento de los parámetros limitando los errores geométricos, para lo que se minimiza la Estimación de la Máxima Probabilidad y la Estimación del Máximo a Posteriori.

Xiong y Turkowski [Xiong and Turkowski, 1997] desarrollan un procedimiento orientado a reconstrucción del entorno para Realidad Virtual utilizando cámaras con lentes de ojo de pez. Se calibran todos los parámetros (incluida la distorsión que utiliza 3 coeficientes sobre un sistema de coordenadas polares). Se basa en un modelo de proyección equi-distante, donde el centro del eje óptico se puede calcular independientemente. Usan un procedimiento iterativo de optimización (Levenberg-Marquardt o Newton-Raphson) para resolver un sistema con 34 incógnitas, para lo que necesitan al menos 4 imágenes.

4.2.5. Otros trabajos relacionados

En este apartado se describen otros trabajos donde se tratan aspectos relacionados con la autocalibración. En primer lugar, son varios los estudios realizados sobre casos críticos, o qué tipo de desplazamientos provocan un alto nivel de error en los resultados de la calibración o que ésta no se pueda resolver.

Tanto Sturm [Sturm, 1997a] como Zisserman [Zisserman et al., 1998] realizan un análisis de estos casos críticos para algoritmos que permitan movimientos generales, pero en los que los parámetros intrínsecos permanezcan constantes. Se concluye que las siguientes posiciones y/u orientaciones de las cámaras constituyen secuencias de movimientos críticos: 1) el eje óptico se mantiene paralelo en distintas adquisiciones de imágenes, es decir, traslación pura y/o rotación exclusivamente sobre dicho eje (ciclomotor); 2) centro óptico colineal, es decir, traslación hacia adelante en la misma dirección del eje óptico pudiendo, incluso, haber rotación sobre el centro óptico; 3) el centro óptico se desplaza siguiendo una elipse o una hipérbola, siendo el eje óptico tangente a dicha elipse o hipérbola. Cuando se realiza un desplazamiento considerado crítico se hace necesario el uso de métodos específicamente desarrollados para ese tipo de movimiento, ya que los encuadrados en este grupo no los pueden resolver.

Un par de trabajos realizan este mismo análisis sobre algoritmos para la autocalibración en el caso de parámetros intrínsecos variables. Kahl et al. [Kahl et al., 2000] estudia de qué manera los distintos tipos de restricciones aplicadas sobre los parámetros intrínsecos influyen en la aparición de estos movimientos críticos. Concluyen que los

valores de los parámetros intrínsecos no influyen en que los movimientos sean críticos. Es decir, no existen configuraciones de cámara que lleven a situaciones críticas para la calibración, sino que es el tipo de desplazamiento realizado por ésta el que determina la aparición de dichas situaciones. En [Sturm, 2002], Sturm presenta una extensión de un trabajo anterior [Sturm, 1997a] para el caso de secuencias en las que varía la distancia focal y para sistemas estéreos.

En [Agapito et al., 2000] se hace un análisis de movimientos críticos para métodos de autocalibración basados en movimientos de rotación. Se observa que se producen resultados inestables cuando se realizan rotaciones puras sobre uno de los ejes que definen el plano imagen, o en situaciones que pueden dar lugar a varias soluciones. Por ejemplo, ocurre que pequeñas rotaciones con cámaras de focal larga pueden dar parecidos resultados a grandes rotaciones con cámaras de focal pequeña; o también, que los resultados de desplazamientos del centro óptico en un eje coincidan con los obtenidos con rotaciones de la cámara en el otro eje.

Por otra parte, existen otros trabajos que analizan la incidencia de algunos factores sobre los resultados globales. En concreto, en [Seo and Hong, 1999] se hace un estudio sobre un simulador y sobre datos obtenidos de cámaras reales acerca de cómo afectan a estos parámetros los errores en el cálculo del punto principal. Y [Agapito et al., 2000] incluye un análisis de los efectos de la distorsión en la autocalibración.

Dentro de este tipo de trabajos cabe destacar el análisis pormenorizado de la influencia del desalineamiento del centro óptico respecto al eje de giro en los métodos de autocalibración basados en sólo rotaciones.

En [Wang et al., 2001] se analiza teórica y experimentalmente esta circunstancia en un contexto donde los parámetros intrínsecos permanecen constantes entre las diferentes adquisiciones de imágenes. Como conclusiones obtienen que los efectos de este desalineamiento son elevados sobre los parámetros intrínsecos. Estos efectos pueden ser mitigados si los puntos se encuentran alejados de la cámara; el ángulo de rotación entre imágenes es alto, pero teniendo cuidado de que la región que solapa dos imágenes sea suficiente como para obtener un mínimo número de puntos correspondientes; o el número de imágenes se aumenta. El trabajo de Hayman y Murray [Hayman and Murray, 2002] avanza en la misma línea que el anterior y analiza el caso en el que la distancia focal pueda variar entre imágenes, aunque supone conocidos los valores del resto de los parámetros intrínsecos. Como conclusión obtiene que el error producido por el desalineamiento es inversamente proporcional a la distancia a la que se encuentran los objetos de la escena, y, por tanto, si esta distancia es lo suficientemente alta el error será mínimo.

Pollefeys y Van Gool [Pollefeys and Gool, 2000] realizan una recopilación bibliográfica de distintos métodos de autocalibración. Además, analizan algunas cuestiones rela-

cionadas con movimientos críticos.

4.3. Métodos de autocalibración

En esta sección se describen los algoritmos de cada uno de los métodos de autocalibración que se han utilizado en la comparativa. Todos los métodos seleccionados corresponden a métodos basados en movimientos de sólo rotación.

4.3.1. Método de McLauchlan [McLauchlan and Murray, 1996]

Una vez seleccionado una serie de puntos de atención en una vista, en este método se intenta predecir la posición de los puntos en la siguiente imagen para después corregir el error rectificando los valores de los parámetros a calibrar. Para ello se usa un filtro recursivo, aunque se necesita una aproximación inicial para los valores de los factores de escalado (horizontal y vertical). Se limita a obtener los valores de los factores de escalado y la orientación relativa entre el plano imagen y el eje de giro de la cámara. Todos estos parámetros se suponen constantes durante el proceso de autocalibración. El resto de los parámetros se suponen fijos y conocidos. Éstos son: el centro del eje óptico, que se supone situado en el centro de la imagen, los ángulos de giro de la cámara obtenidos a partir de la lectura de los codificadores de posición de los motores, los ejes del plano imagen se suponen completamente ortogonales, y la distorsión no se modela. El proceso necesita de un gran número de imágenes para converger a la solución. Si la distancia focal cambia durante el proceso hay que reinicializarlo.

El filtro utilizado se denomina Filtro de Dimensión de Estado Variable (VSDF). El vector de estado utilizado tiene dos componentes: una global (x) que incluye los parámetros de la calibración que se quieren obtener, y que permanecen constantes a lo largo de toda la secuencia (focales y orientación relativa); y la otra local (y_i) que incluye la posición 3D de cada uno de los puntos detectados en la escena. El tamaño de esta componente es variable ya que depende del número de puntos observados en cada instante. La actualización de este vector dependerá de una matriz (Q) que contiene las covarianzas del vector.

Para inicializar el método es necesaria la aportación de unos valores iniciales, en concreto los factores de escalado, que pueden obtenerse de los datos aportados por el fabricante de la cámara, permitiéndose, según los autores, una desviación de hasta el 50 % respecto a los valores reales. La orientación relativa se inicializa a cero y los valores de la matriz de covarianzas se estiman en función de la precisión que se prevea van a tener los datos. Según los autores estos valores pueden rondar los 30 pixels para los factores

de escalado y 0.1 radianes para la orientación. Los valores utilizados para la creación de esta matriz afectarán al tiempo de convergencia del método y su estabilidad.

La matriz de covarianzas se divide en $(n + 1) \times (n + 1)$ bloques (4.1). Estos bloques se clasifican en tres tipos: uno contiene la covarianza del vector de estado global (A); otro incluye las covarianzas del vector de estado local (C_i); y el último tipo son los que relacionan los dos vectores de estado (B_i). De estos dos últimos elementos existirá un bloque por cada uno de los puntos observados en la escena.

$$Q^{-1} = \begin{bmatrix} A & B_1 & \dots & \dots & B_n \\ B_1^t & C_1 & 0 & \dots & 0 \\ \vdots & 0 & \dots & \dots & \vdots \\ \vdots & \vdots & \dots & \dots & 0 \\ B_n^t & 0 & \dots & 0 & C_n \end{bmatrix} \quad (4.1)$$

Inicialmente la submatriz A contendrá en su diagonal la varianza prevista para cada uno de los elementos del vector de estado global.

Cada vez que en la secuencia de imágenes aparece un nuevo punto a procesar hay que añadir un nuevo estado local al vector de estados (y_{n+1}). Para ello, se estima su posición 3D respecto a la cámara utilizando las coordenadas de imagen del punto, los valores del vector de estado global estimado y los valores de azimut y elevación de la cámara para el instante de tiempo en el que es observado por primera vez. Se crea un nuevo bloque B_{n+1} y otro C_{n+1} , con lo que la matriz de covarianzas aumenta en un bloque las filas y las columnas. Estas submatrices se inicializan a cero cuando se crean.

Cada vez que un punto no es detectado en una serie de imágenes consecutivas hay que eliminar el elemento correspondiente del vector de estado local. El mismo procedimiento se aplica a las submatrices correspondientes. Por tanto, la matriz de covarianzas se decrementa en un bloque tanto en filas como en columnas. Antes de realizar estas operaciones hay que actualizar la submatriz de covarianzas del vector de estado global: $A = A - B_l C_l^{-1} B_l^t$.

Para actualizar el vector de estados hay que calcular la innovación o medida del error cometido en la predicción: $v_i = z_i - h_i(x, y_i)$, donde z_i son las coordenadas del punto en la imagen actual y $h_i(x, y_i)$ es la estimación de la posición actual del punto. Esta estimación se obtiene en función de la posición 3D estimada (y_i), los valores actuales del vector de estado global (x), y el azimut y la elevación de la cámara del instante en que se realizó la adquisición de la imagen. También hay que obtener las matrices con los jacobianos de la función de proyección con los vectores de estado: $D = \partial h / \partial x$ y $E = \partial h / \partial y$.

Las ecuaciones que actualizan el vector de estados y la matriz de covarianza son:

$$\begin{aligned}
A &= A + \sum_{i=1}^n D_i^t S_i^{-1} D_i \\
B_i &= B_i + D_i^t S_i^{-1} E_i \\
C_i &= B_i + E_i^t S_i^{-1} E_i \\
M &= A - \sum_{i=1}^n B_i C_i^{-1} B_i^t \\
xd &= M^{-1} \sum_{i=1}^n (D_i^t - B_i C_i^{-1} E_i^t) S_i^{-1} v_i \\
x &= x + xd \\
yd_i &= C_i^{-1} (E_i^t S_i^{-1} v_i - B_i^t xd) \\
y_i &= y_i + yd_i
\end{aligned} \tag{4.2}$$

donde S_i es una matriz que contiene en la diagonal las varianzas del ruido previsto para las coordenadas de imagen. Los autores suponen que este valor es de 0.5 pixels.

Este proceso se repite a lo largo de una secuencia de imágenes hasta que se minimice la siguiente función:

$$J = [xd; yd_i]^t Q^{-1} [xd; yd_i] + \sum_{i=1}^n v_i^t S_i^t v_i \tag{4.3}$$

donde el primer sumando representa la variación del vector de estado completo mientras que el segundo término está relacionado con el error cometido en la estimación de la posición de los puntos en la última adquisición.

4.3.2. Método Lineal de Agapito [de Agapito et al., 1999]

Es un método que permite la autocalibración de cámaras cuyo movimiento se limite a rotaciones, aunque pueden variar los parámetros intrínsecos durante la adquisición de las imágenes necesarias para realizar este proceso. El método es aplicable bajo una serie de restricciones: se supone que los ejes del plano imagen son perfectamente ortogonales, y que las imágenes no están afectadas por distorsión. Además, en el caso de que sólo exista un movimiento de azimut-elevación, o sea, que no exista ciclotorsión (roll), debe añadirse, al menos, una restricción más. En estos casos se parte de la suposición de que los pixels sean cuadrados, o que al menos se conozca la relación entre sus lados.

El método se basa en la obtención de la cónica absoluta a partir de la homografía entre imágenes mediante un algoritmo lineal. Por tanto, al ser un método lineal, su tiempo de ejecución es muy pequeño. Luego, se aplica un método como el de factorización de Cholesky para extraer los parámetros de la cámara. Teóricamente, aplicando las restricciones anteriormente mencionadas es necesario un mínimo de tres vistas desde distintas orientaciones para poder obtener los resultados.

Para calibrar la cámara hay que obtener en primer lugar, la homografía de cada ima-

gen respecto a una que se toma como referencia, normalmente la primera se la secuencia. Para esto se puede utilizar el método lineal propuesto por Hartley [Hartley, 1997b].

Una vez obtenidas las matrices (H_i) que relacionan cada una de las imágenes de la secuencia con la de referencia, se puede aplicar la siguiente ecuación para obtener la cónica absoluta (w):

$$w_i = H_i^{-t} w_0 H_i^{-1} \quad (4.4)$$

donde w_0 es la cónica de la imagen de referencia. Al ser esta una matriz simétrica, se puede representar su triangular superior como un vector de seis elementos ($a = [a_1, a_2, \dots, a_6]^t$). La matriz w_i es función de los parámetros intrínsecos de la cámara, los cuales se recogen en la matriz A (ecuación 2.15).

Por tanto, si los ejes del plano imagen son ortogonales ($\Omega = 90^\circ$) y los pixels son cuadrados ($k_x = k_y$), entonces w es equivalente a:

$$w = A^{-t} A^{-1} = \begin{bmatrix} 1/\alpha^2 & 0 & -u_0/\alpha^2 \\ 0 & 1/\alpha^2 & -v_0/\alpha^2 \\ -u_0/\alpha^2 & -v_0/\alpha^2 & 1 + (u_0^2 + v_0^2)/\alpha^2 \end{bmatrix} \quad (4.5)$$

donde α es el factor de escalado ($\alpha = k_x = k_y$). A la vista de esta equivalencia se pueden deducir dos ecuaciones para cada imagen: una en la que $w_i(1, 2) = 0$ ¹ y otra en la que $w_i(1, 1) = w_i(2, 2)$. El conjunto de todas las ecuaciones, incluyendo las ecuaciones sobre la imagen de referencia se almacena en la matriz E , donde cada fila es una ecuación. Éstas se forman aplicando estas restricciones a la ecuación 4.4, lo que resulta un sistema de ecuaciones de la forma $E \cdot a = 0$. Para resolverlo, basta con encontrar el vector a que minimice $\|Ea\|$ tal que $\|a\| = 1$. Esto se consigue descomponiendo en valores singulares la matriz E , y tomando como solución el autovector correspondiente al autovalor más pequeño. Al obtener a se reconstruye la matriz w_0 . Luego se aplica la ecuación 4.4 para obtener las matrices w_i de cada una de las imágenes.

Para extraer los parámetros intrínsecos de cada imagen (A_i) a partir de w_i se puede aplicar la factorización de Cholesky sobre esta matriz e invertir el resultado o bien se puede deducir directamente de la ecuación 4.5. Hay que tener en cuenta que para poder realizar esta operación, la matriz que contiene la cónica absoluta de cada imagen debe ser definida positiva. Por último, los parámetros extrínsecos, o sea, la matriz de rotación que relaciona cada imagen respecto a la utilizada como referencia, se pueden obtener de la siguiente ecuación:

$$R_i = A_0 H_i A_i^{-1} \quad (4.6)$$

¹ $w(j, k)$ es el elemento de la j -ésima fila y k -ésima columna de la matriz w

donde A_0 es la matriz de parámetros intrínsecos de la imagen de referencia y A_i es la matriz de la i -ésima vista.

En este proceso se puede variar el número de restricciones que se aplican al algoritmo en función del grado de conocimiento que se tenga sobre las cámaras y la cantidad de datos que haya disponible. Una de las restricciones que se suele aplicar es la de prefijar el valor del centro del eje óptico.

4.3.3. Método iterativo de Agapito [de Agapito et al., 1998]

Es una prolongación del método anterior. Se ha seleccionado para contrastar la rentabilidad de usar un método iterativo frente a uno lineal. Se aplica el método de optimización de Levenberg-Marquardt a las restricciones que impone la homografía en el infinito. Al igual que en el método lineal anterior, se contempla la variación de los parámetros intrínsecos, principalmente focal y centro óptico, aunque si este último permanece constante los resultados son sensiblemente mejores.

El método necesita una solución aproximada para iniciar el proceso, la cual se puede conseguir aplicando un método lineal. Los parámetros que se necesitan conocer a priori son el centro óptico y los factores de escalado, a partir de los cuales se forma la matriz A . En este trabajo se ha utilizado el método descrito en el punto anterior, propuestos por los mismos autores. A continuación se minimiza, usando Levenberg-Marquardt, la siguiente medida:

$$\sum_{i=1}^n \|A_i A_i^t - H_i A_0 A_0^t H_i\|_F^2 \quad (4.7)$$

donde $A_i A_i^t$ y $H_i A_0 A_0^t H_i$ son normalizadas mediante la norma de Frobenius para eliminar el factor de escala.

4.3.4. Método de Seo-Hong [Seo and Hong, 1998]

Es un método iterativo de autocalibración en el cual los parámetros intrínsecos (factores de escalado y centro óptico) pueden variar entre cada adquisición de imagen, pero necesita que los ejes del plano imagen sean ortogonales. Se basa en relacionar la homografía entre imágenes y obtener la cónica absoluta. Son necesarias al menos cuatro homografías y, por tanto, cuatro vistas de la escena, aunque este número se puede reducir aumentando las restricciones.

Para iniciar el proceso de calibración es necesario conocer las homografías que relacionan cada par de imágenes. Para ello se comenta en [Seo and Hong, 1999] que se puede usar el método de Kanatani [Kanatani, 1998], pero en el presente trabajo se ha utilizado

el mismo algoritmo que en el método de Agapito (método de Hartley [Hartley, 1997b]), ya que se trata de comparar los métodos aquí comentados y no comparar la forma de obtener las homografías.

Como todo método iterativo necesita una aproximación inicial a los valores de los parámetros intrínsecos que permitan que el método converja. Se propone un método lineal para obtener valores aproximados a los factores de escalado horizontal y vertical (k_x, k_y) de cada imagen. El centro óptico se supone inicialmente en el centro de la imagen.

Al igual que en otros métodos, éste se desarrolla a partir de la ecuación de la cónica absoluta (ver 4.4), de la cual se pueden extraer directamente los factores de escalado. Relacionando la imagen k -ésima con la imagen de referencia (primera imagen) mediante la homografía (H_k) se generan las siguientes ecuaciones:

$$\begin{bmatrix} h_{11}h_{21} & h_{12}h_{22} \\ h_{11}h_{31} & h_{12}h_{32} \\ h_{21}h_{31} & h_{22}h_{32} \end{bmatrix} \begin{bmatrix} k_x(0)^2 \\ k_y(0)^2 \end{bmatrix} = \begin{bmatrix} -h_{13}h_{23} \\ -h_{13}h_{33} \\ -h_{23}h_{33} \end{bmatrix} \quad (4.8)$$

donde h_{ij} son los elementos de la matriz H_k , y $k_x(0)$ y $k_y(0)$ son los factores de escalado horizontal y vertical de la imagen cero o de referencia. Uniendo todas las ecuaciones formadas para cada una de las imágenes adquiridas, se puede formar un sistema de ecuaciones a partir del cual se pueden deducir los factores de escalado de la imagen de referencia. Una vez obtenidos estos se pueden obtener los factores de escalado del resto de las imágenes:

$$k_x(k)^2 = \frac{k_x(0)^2 h_{11}^2 + k_y(0)^2 h_{12}^2 + h_{13}^2}{k_x(0)^2 h_{31}^2 + k_y(0)^2 h_{32}^2 + h_{33}^2} \quad (4.9)$$

$$k_y(k)^2 = \frac{k_x(0)^2 h_{21}^2 + k_y(0)^2 h_{22}^2 + h_{23}^2}{k_x(0)^2 h_{31}^2 + k_y(0)^2 h_{32}^2 + h_{33}^2} \quad (4.10)$$

El procedimiento incluye una fase de optimización no-lineal basada en la minimización de la siguiente función de error en la que se utiliza norma de Frobenius ($\|\cdot\|_F$):

$$E = \sum_{k=1}^N (\|R_k R_k^t - I\|_F^2 + \|R_k^t R_k - I\|_F^2) \quad (4.11)$$

donde $R_k = \frac{A_k^{-1} H_k A_0}{\det(A_k^{-1} H_k A_0)^{1/3}}$. Esta función trata de ajustar la matriz de rotación de cada una de las poses en las que han sido tomadas las imágenes, en relación a la imagen de referencia, a una matriz ortonormal.

4.4. Experimentos en el simulador

Se han realizado diversos experimentos con el objetivo de ver la precisión y robustez de cada uno de los métodos descritos. Estos elementos, junto con otros como el coste computacional de cada método, se utilizan para realizar una valoración comparativa de los métodos implementados.

Al igual que en el caso de los métodos clásicos se plantea el objetivo de averiguar en qué medida afectan distintos factores a los resultados proporcionados por estos métodos. Esto permitirá identificar las mejores condiciones de aplicación de cada uno de ellos. Además, se pretende estudiar las ventajas y desventajas de utilizar este tipo de métodos frente a los clásicos.

Con el fin de responder a estas cuestiones se han planteado en primer lugar una serie de experimentos sobre un simulador, donde el conocimiento de las condiciones experimentales permite un análisis profundo de la incidencia de cada factor sobre cada uno de los métodos. En segundo lugar, algunos de los experimentos han sido trasladados a un banco de pruebas real, donde se ha podido analizar en conjunto el funcionamiento de cada método.

En el simulador se analiza, por un lado, la incidencia de factores relativos a la formación de las imágenes, como el ruido de cuantización marcado por la resolución espacial del sensor CCD. Este factor afecta a la precisión con la que es posible recuperar las coordenadas de imagen, coordenadas 2D, de los puntos utilizados en la calibración. Esta imprecisión se modela como un ruido gaussiano de media nula. En este caso, no existe ruido en las coordenadas 3D (como ocurre en los métodos clásicos) debido a que la posición de los puntos en el espacio no se utiliza como dato de entrada en este tipo de métodos.

También se analizan aspectos relativos a las limitaciones que presentan los modelos de cámara utilizados. En este caso se analiza el efecto de la distorsión, ya que es un parámetro que se modela en muy pocos métodos de autocalibración; de hecho, ninguno de los métodos utilizados en el presente estudio lo contempla. Otro aspecto que no se modela es el desalineamiento entre los ejes de giro de la cámara y el centro óptico. Este factor es importante ya que al realizar cambios en la orientación en la cámara existe un desplazamiento del plano imagen que no se recoge en el modelo de movimiento contemplado por el método empleado. En trabajos como [Wang et al., 2001] y [Hayman and Murray, 2002] se indica que la influencia de este efecto es alta, pero que disminuye ostensiblemente si los puntos procesados se encuentran alejados de la cámara. También se analizan la influencia de parámetros como el grado de ortogonalidad de los ejes del plano imagen y el factor de proporción de los pixels debido a que en la mayoría

de los casos no se modela, porque se suponen muy próximos a su estado idóneo (pixels totalmente cuadrados).

El uso de cámaras que permiten realizar zoom o tienen enfoque automático hace que se susciten situaciones en las que los parámetros intrínsecos puedan variar. Por tanto, es potencialmente interesante analizar la influencia de estos factores, en este caso, modelando estos factores en función de la distancia focal.

El análisis de la influencia de cada uno de estos elementos sobre cada método permite examinar la robustez de éstos. También se analiza como inciden dichos elementos en la precisión de cada parámetro definido en el modelo utilizado por cada método, pudiendo por tanto, estudiar la precisión de éstos.

Finalmente, se analizan aspectos que influyen de forma determinante en el coste computacional de cada método. Se estudia, por un lado, la influencia del número de vistas o imágenes que son necesarias para que cada método converja. Por otro lado, se analiza cómo el número de puntos detectados afecta tanto a la precisión de los resultados como al tiempo de ejecución.

Para realizar los experimentos se ha utilizado el mismo simulador que en el caso de los métodos clásicos. En todos los experimentos realizados y para todos los métodos utilizados se ha usado la misma configuración de cámara para proyectar el patrón de calibración. Se ha utilizado una configuración próxima a la de la cámara que se utiliza en los experimentos reales para poder comparar los resultados. El plano imagen es de 768x576 pixels, situando el centro del eje óptico en el centro de la imagen (384,288). Inicialmente la distancia focal utilizada es de 4.5 mm. Esto permite un mayor campo de visión y por tanto que el número de puntos procesados sea mayor. Los ejes del plano imagen son ortogonales. El tamaño del CCD es de 3.2 mm x 2.4 mm, siendo, por tanto, los pixels cuadrados. De esta configuración se deduce que los factores de escalado, tanto horizontal como vertical, tendrán un valor de 1080 pixels. No se ha introducido inicialmente ninguna distorsión.

Para simular los puntos detectados se ha generado aleatoriamente una nube de 200 puntos distribuidos uniformemente sobre un cubo de 2 metros de lado y cuyo centro se sitúa a una distancia de dos metros y medio del plano imagen de la cámara. Por tanto, los puntos se distribuyen en rango que va desde 1.5 m hasta 3.5 m. De esta nube de puntos, aproximadamente unos cien puntos entran en el campo visual de la cámara modelada en cada una de las vistas.

Los métodos estudiados necesitan que la cámara permanezca en el mismo punto (sin desplazamientos), por lo que sólo permiten rotaciones. Por tanto, se ha generado una trayectoria en la que el movimiento del eje óptico de la cámara describe un cono de radio de 10° (figura 4.2). Entre vistas consecutivas de la secuencia se avanza un arco

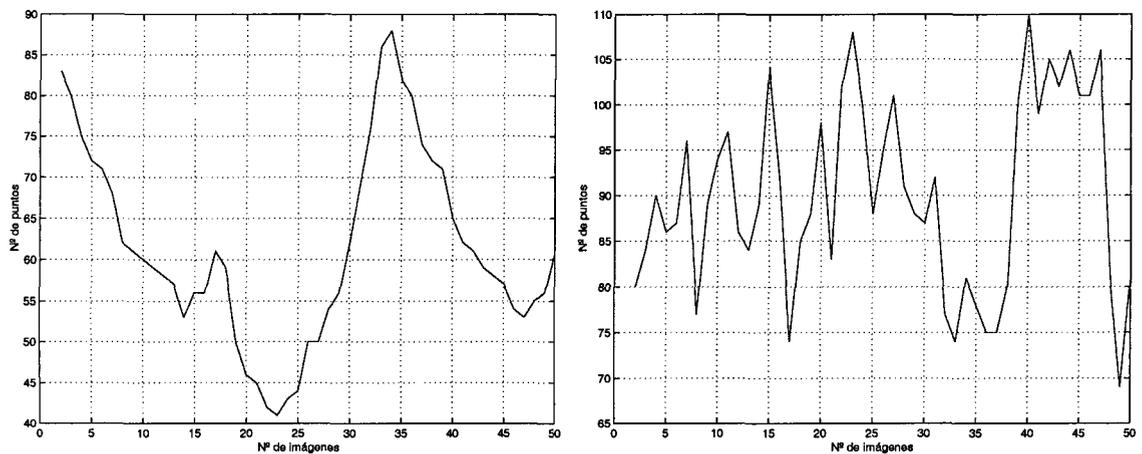


Figura 4.1: a) Número de puntos coincidentes entre cada vista y la primera (imagen de referencia). b) Número de puntos coincidentes entre cada vista y su precedente.

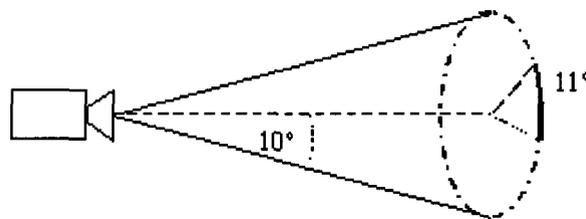


Figura 4.2: Movimiento realizado por la cámara en cada adquisición de imagen.

de 11° , respecto al último desplazamiento, dentro del movimiento circular realizado. En la figura 4.1 se puede observar el número de puntos en correspondencia entre pares de imágenes al efectuar este movimiento. Hay que hacer notar que, a excepción del método de McLauchlan-Murray, ninguno de los métodos utiliza el conocimiento del desplazamiento realizado por las cámaras, de hecho, es uno de los parámetros que obtiene el proceso de calibración (parámetros extrínsecos).

Los métodos implementados necesitan que a priori se fije el número de imágenes que van a ser procesadas, ya que salvo algunas cuestiones de bajo nivel, como puede ser el cálculo de la homografía, el resto del procedimiento se ejecuta cuando se ha adquirido la secuencia completa. La única excepción es el método McLauchlan que realiza un proceso imagen por imagen, y el proceso finaliza cuando se alcanza un mínimo, el error se estabiliza, o alcanza un número máximo predefinido de vistas procesadas.

Para cada uno de los métodos se ha evaluado en cada experimento el comportamiento de cada parámetro, tanto intrínsecos como extrínsecos, incluido en el modelo que utiliza y se han comparado entre sí los resultados de cada método. Como medida de error se ha utilizado el promedio de distancia entre el valor real empleado en la simulación y la estimación obtenida. Hay que hacer notar que, a diferencia de los métodos clásicos o de

calibración manual, los parámetros extrínsecos expresan la relación entre la posición de la cámara en cada una de las vistas respecto a la de referencia (primera imagen). Por último, se ha usado el error RMS (A.7) sobre la distancia entre las coordenadas 2D de la imagen generadas con el simulador y la reconstrucción obtenida mediante la calibración.

Por último, hay que indicar que el método de McLauchlan-Murray sólo aparece en algunas gráficas debido a que solamente coincide con el resto de los métodos en la estimación de los factores de escalado (k_x, k_y) , ya que el otro conjunto de parámetros que obtiene es el desalineamiento en la orientación de los ejes de coordenadas de la cámara (plano imagen) respecto al eje de giro de ésta, el cual no es obtenido por ningún otro método. También hay que recordar que el resto de los métodos utilizados obtiene un valor para los factores de escalado y para el centro óptico por cada una de las imágenes de la secuencia. Por tanto, para estos parámetros se muestra en las gráficas el promedio del error del conjunto de estimaciones de la secuencia entera. Es decir, se obtiene en cada vista la diferencia entre el valor real y el estimado de cada parámetro, y en las gráficas se muestra el promedio de estas diferencias.

4.4.1. Análisis del coste computacional

En esta sección se analizan aspectos que influyen de forma determinante en el coste computacional de cada método. Por un lado, se estudia la influencia del número de vistas o imágenes que son necesarias para que cada método alcance unos resultados aceptables. Por otro lado, se analiza cómo el número de puntos detectados afecta tanto a la precisión de los resultados como al tiempo de ejecución.

Antes de observar la medición de tiempos realizada para cada uno de los métodos, se analiza cuántas vistas y cuántos puntos por imagen son necesarios para que los métodos converjan y en función de esto estimar el coste computacional de cada método.

Variación del número de imágenes

En este apartado se realiza el análisis de la influencia del número de imágenes sobre el precisión de los resultados de la calibración y sobre el tiempo de ejecución. Para comprobar esta influencia se ha realizado un experimento donde el número de imágenes de la secuencia empleada en la calibración se ha ido incrementando. Durante toda la secuencia de imágenes utilizada para realizar la prueba, los parámetros intrínsecos se han mantenido fijos, aunque en los métodos preparados para ello se ha estimado el valor de cada parámetro de forma independiente para cada una de las vistas.

El rango de variación del número de imágenes en todos los casos ha sido entre 3 y 50. Un ruido gaussiano de media nula y desviación típica de 0.5 pixels ha sido añadido

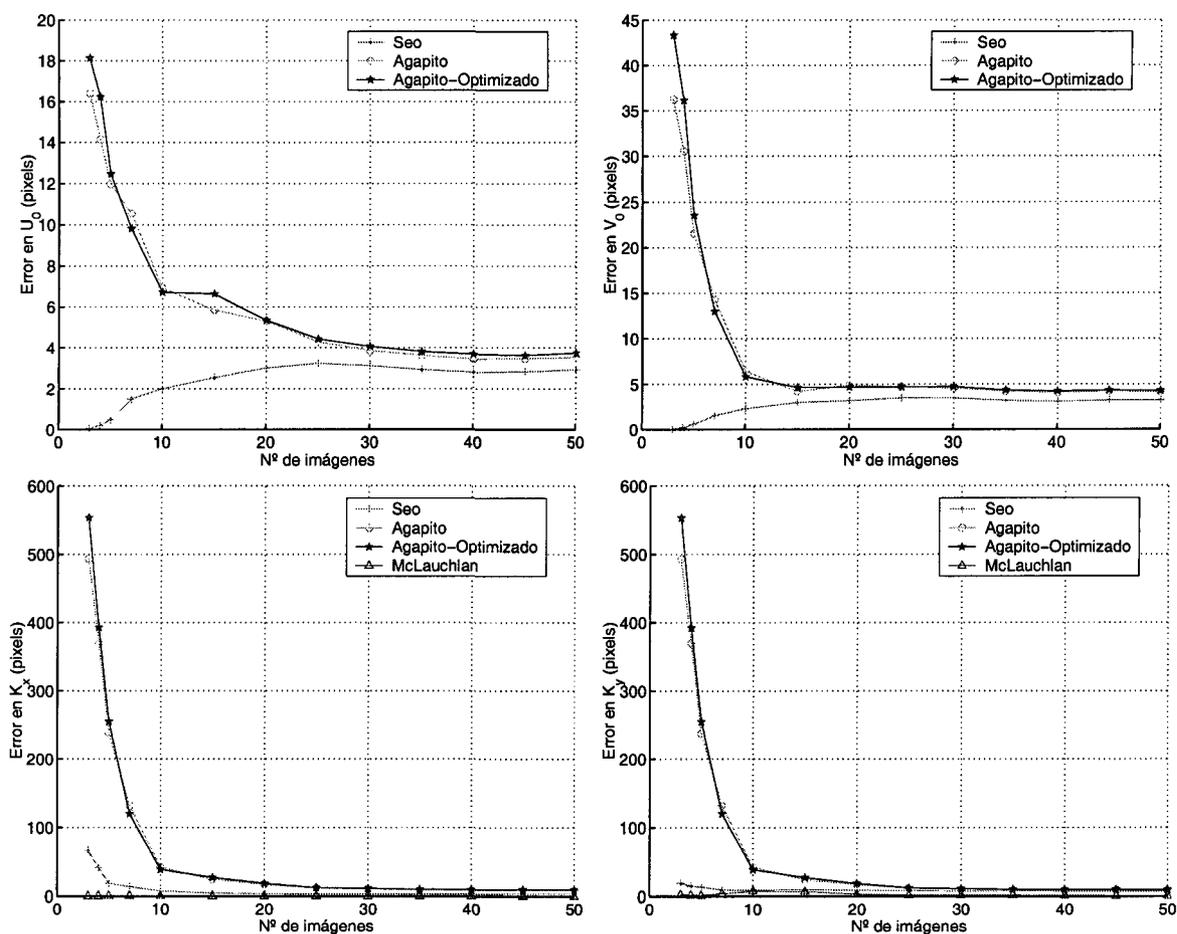


Figura 4.3: Error de cada parámetro intrínseco cuando varía el número de imágenes. Las imágenes tienen un ruido gaussiano de 0.5 pixels en sus coordenadas 2D. Los parámetros intrínsecos permanecen constantes en la secuencia.

a las coordenadas de imagen para simular el efecto de cuantización en el proceso de formación de la misma. Cada experimento se ha repetido 50 veces para cada método.

Los resultados obtenidos en este experimento se muestran en la figuras 4.3 y 4.4. En el primer conjunto de gráficas se pueden observar los errores cometidos por los métodos en la estimación de los parámetros intrínsecos, mientras que en el segundo grupo se muestran los errores en la estimación de los parámetros extrínsecos y el error RMS en las coordenadas 2D de la imagen.

Como se puede observar el resultado del método de Agapito, tanto la versión lineal como la iterativa, es peor que el obtenido por el método de Seo en todos los parámetros. También se observa que la versión optimizada de este método no mejora significativamente los resultados de la versión lineal, e incluso, obtiene un mayor error cuando el número de imágenes es pequeño (menos de cinco). En general, se observa que este método presenta errores muy grandes en la estimación de los parámetros si el número de

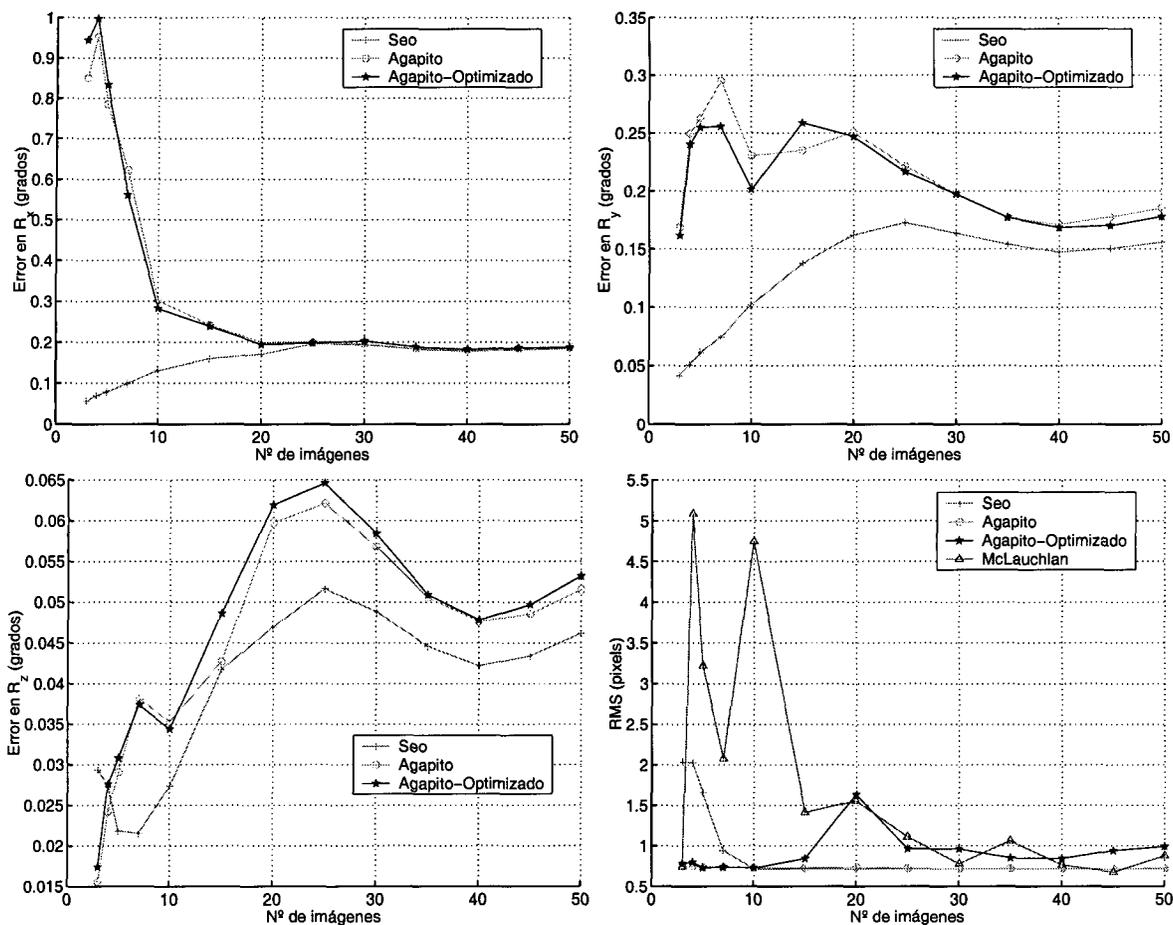


Figura 4.4: Error en la orientación y RMS sobre las coordenadas 2D cuando varía el número de imágenes. Las imágenes tienen un ruido gaussiano de 0.5 pixels en sus coordenadas 2D. Los parámetros intrínsecos permanecen constantes en la secuencia.

imágenes procesadas es inferior a 10.

Por otro lado, se puede ver que el error en la estimación del centro óptico en el método de Seo aumenta con el número de imágenes, y llega a converger con niveles de error semejantes a los del método de Agapito. En cambio, en los factores de escalado el resultado es el inverso, es decir, el error disminuye cuando aumenta el número de vistas procesadas. Este resultado parece deberse a que el método de Seo ajusta en un primer momento, y de forma lineal los factores de escalado, con lo que éstos absorben la mayor parte del error producido por el ruido en las coordenadas de imagen. Ésto, unido al hecho de haber usado Seo una configuración de cámara en la que el centro óptico se encuentra en el centro de la imagen, provoca que cuando la cantidad de muestras no es muy grande la estimación del centro óptico permanece muy cerca del centro de la imagen.

Por último, se observa que el método de McLauchlan obtiene excelentes resultados en

la estimación de los factores de escalado, lo que no es sorprendente por las condiciones del experimento, que son las idóneas para este método. En el siguiente experimento se puede observar la sensibilidad de dicho método a la variación de dicho parámetro. Llama la atención que, si bien este método modela muy bien los factores de escalado, en la gráfica del error RMS aparece como el peor método y, además, con un comportamiento muy irregular cuando el número de imágenes es inferior a 15. Esto parece deberse a que en el resto de los métodos se estiman tanto los parámetros extrínsecos como los intrínsecos, lo que posiblemente provoque el que unos parámetros compensen los errores de otros, tal y como se demostró en el análisis de los métodos clásicos. Sin embargo, en el método de McLauchlan, sólo se obtienen ciertos parámetros intrínsecos, por lo que los errores en la estimación de éstos no se ven compensados. Además, si se observa las pruebas documentadas en el trabajo de McLauchlan y Murray [McLauchlan and Murray, 1996] este método necesita de secuencias muy largas para que los resultados converjan, del orden de varios cientos, cuando las aproximaciones iniciales no son muy buenas.

Si se analizan los parámetros destaca el hecho de la relativa buena precisión que ofrecen los métodos de Seo y Agapito a la hora de estimar la orientación relativa de cada una de las vistas (parámetros extrínsecos), si se compara con la precisión de los parámetros intrínsecos.

Variación del número de puntos por imagen

Este experimento tenía como objetivo observar la respuesta de los distintos métodos cuando varían el número de elementos (puntos) detectados en una vista. Este factor es relevante debido a que, dependiendo del entorno donde se realice la calibración, el número de puntos detectados variará. Por tanto, si el número de puntos por imagen es pequeño es probable que deba compensar la situación aumentando el tamaño de la secuencia utilizada para realizar el proceso de calibración. Además, la cantidad de puntos procesados incide de forma directa en el tiempo de ejecución de cada método.

Para realizar la prueba se han generado secuencias de imágenes en las que se detectan desde 20 hasta 400 puntos por vista. A las imágenes se les ha inducido un ruido gaussiano de media cero y desviación típica de 0.5 pixels. Para cada valor del rango se ha repetido el experimento 50 veces. La distancia focal y el centro óptico permanecen constantes en las secuencias de imágenes. Para todos los métodos se han utilizado secuencias de 15 imágenes. Esta cifra se ha seleccionado en función de los resultados obtenidos en el experimento anterior. En el caso del método de McLauchlan el número de imágenes depende de cuando converge el método. En este experimento se ha observado que cuando el número de puntos es muy pequeño (20 puntos) necesita procesar varios cientos, pero

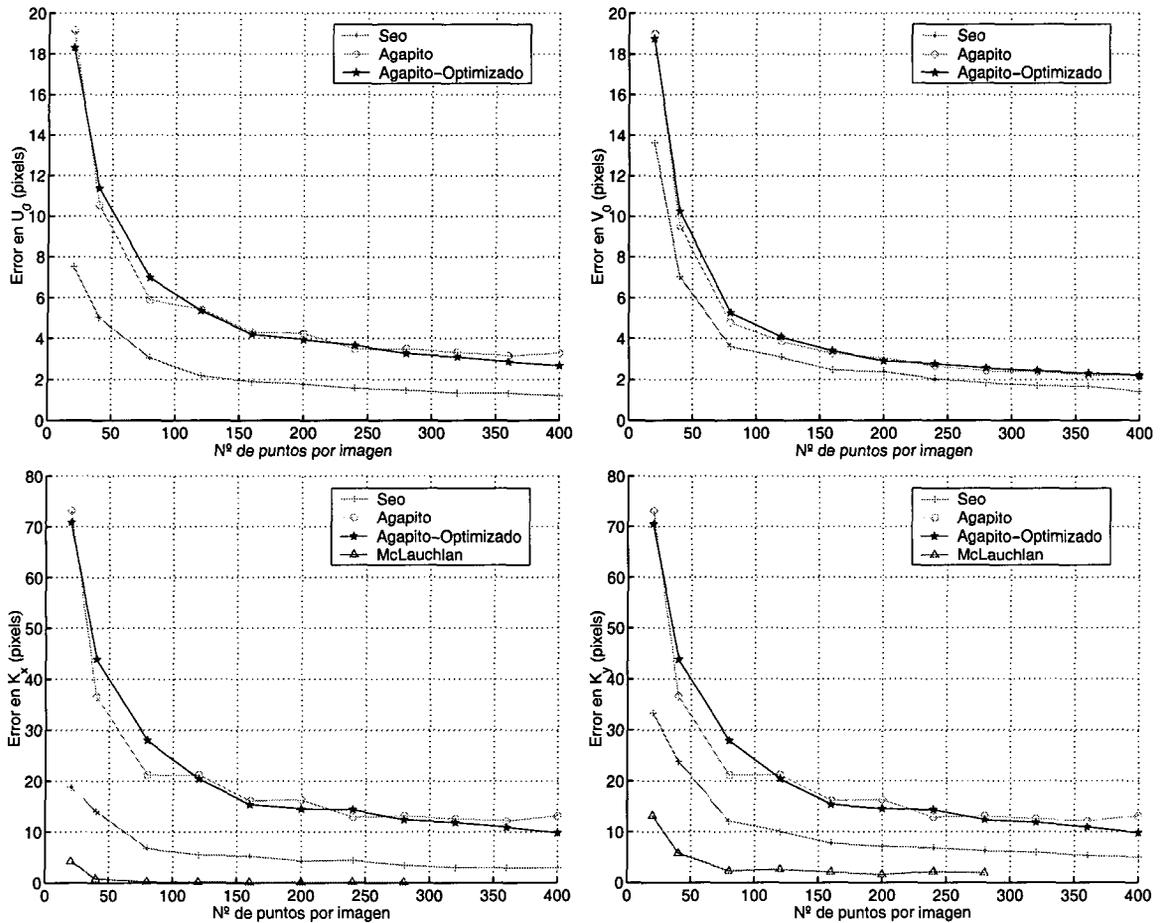


Figura 4.5: Error de cada parámetro intrínseco cuando varía el número de puntos detectados por imagen.

en el resto de los casos no supera las 30 imágenes, necesitando en algunos casos incluso menos 10.

Como era de esperar, todos los métodos tienen un comportamiento exponencial en el error (figuras 4.5 y 4.6), que va decreciendo a medida que el número de puntos detectados aumenta. Parece que a partir de los 150 puntos el error se estabiliza.

Es importante destacar que en el método de McLauchlan no se pudo superar los 300 puntos por imagen ya que agotaba la memoria disponible. Esto es un signo indicativo de la cantidad de recursos que necesita este método. Por otro lado, se observa que este método necesita muchos menos puntos, entorno a la mitad respecto a los otros, para que el error se estabilice. También hay que indicar que si el número de puntos es muy pequeño los métodos de Agapito tienen, a veces, problemas para resolver la calibración, al menos, con el nivel de ruido utilizado. El problema de este método es que si el sistema de ecuaciones que resuelve devuelve una matriz de la cónica absoluta que no es definida positiva no se pueden extraer los parámetros intrínsecos y, por tanto, no se puede resolver

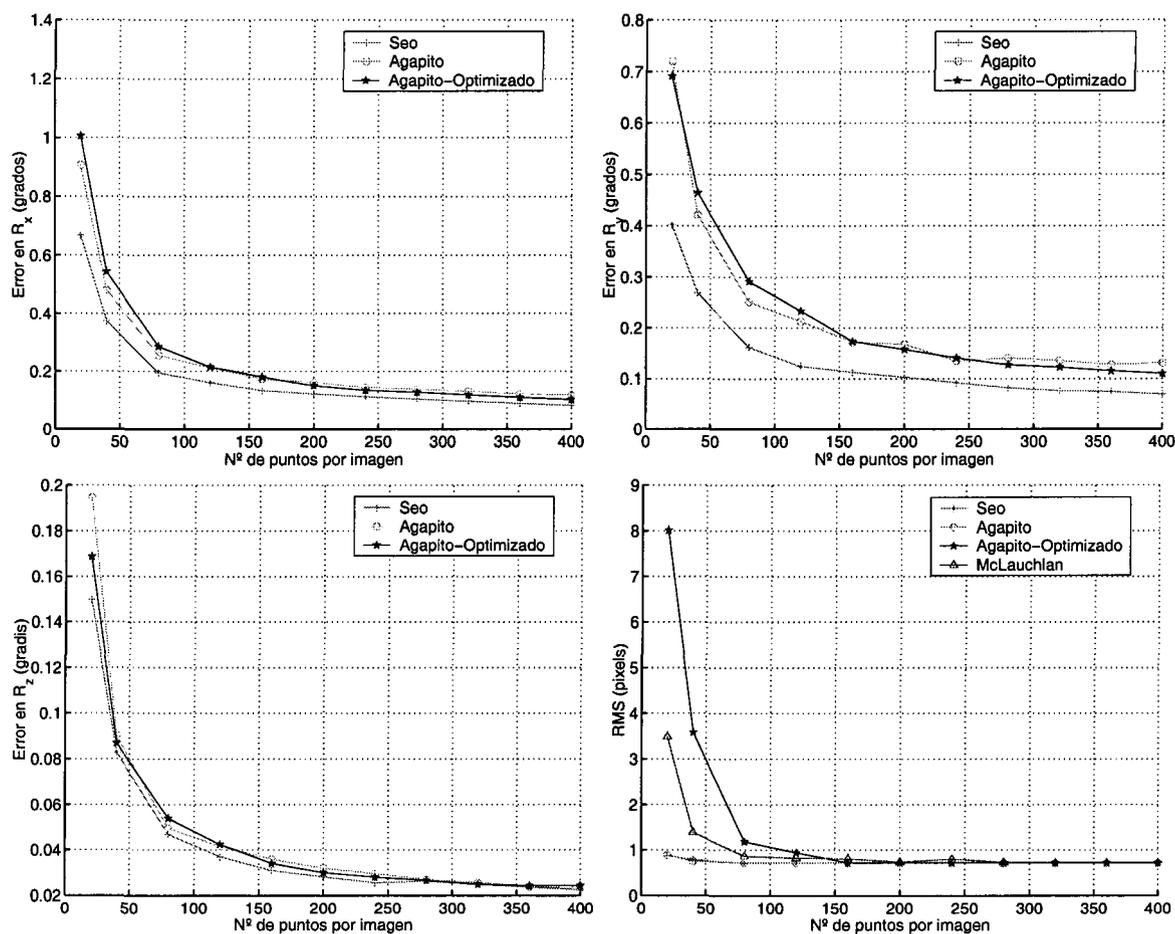


Figura 4.6: Error en la orientación y RMS sobre las coordenadas 2D cuando varía la distancia a lo largo del eje óptico (Z) entre el foco y el eje de giro de la cámara. Las imágenes están libres de ruidos.

la calibración.

Tiempos de ejecución

Al realizar los experimentos descritos en los dos apartados anteriores se ha medido el tiempo que ha tardado cada método en realizar la calibración. Hay que indicar, que estas medidas de tiempo no incluyen el tiempo necesario para la adquisición de cada imagen. Como los experimentos se han repetido varias veces, se ha medido el promedio que tarda cada método en realizar el proceso. Los experimentos se han realizado en Matlab (versión 5.3.0.10183 (R11)) y sobre un Pentium IV a 2Ghz. y con 256Mb de RAM.

En la figura 4.7 se muestran los tiempos de ejecución de cada uno de los métodos en función del número de imágenes. En la primera gráfica se muestra el tiempo total de ejecución, mientras que en la segunda gráfica se muestra el tiempo que transcurre

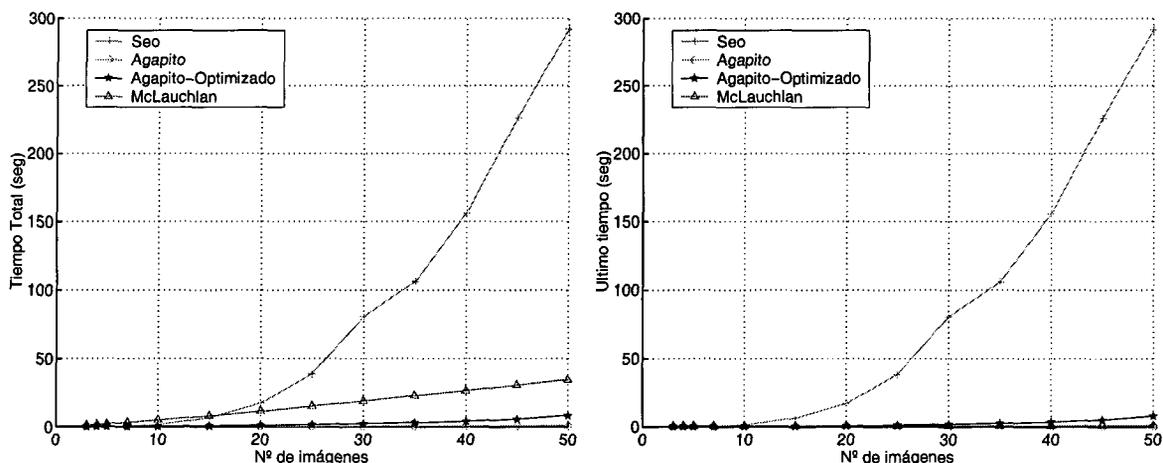


Figura 4.7: Tiempo de ejecución total y posterior a la última adquisición de imagen según el número de imágenes que componen la secuencia procesada.

después de la adquisición de la última imagen. El mostrar las dos gráficas se justifica por el hecho de que hay métodos que no realizan el proceso calibración hasta haber adquirido la última imagen de las necesarias para que el método funcione correctamente. Aunque entre adquisición y adquisición estos métodos pueden adelantar parte del proceso realizando tareas de bajo nivel como, por ejemplo, la estimación de la homografía. Obviamente, a estos tiempos habría que sumar el tiempo necesario para adquirir cada una de las imágenes.

En las gráficas se puede observar que, con diferencia, el método más lento es el de Seo, el cuál tiene un comportamiento exponencial. Con un pequeño número de imágenes es muy rápido, pero si la secuencia tiene más de 15 imágenes se torna excesivamente lento, necesitando para realizar la calibración de una secuencia de 50 imágenes cerca de 5 minutos.

El método de McLauchlan es relativamente lento ya que para secuencias de 50 imágenes necesita unos 40 segundos, además, hay que tener en cuenta que este método finaliza el proceso cuando convergen los resultados y que, por tanto, es habitual que utilice un mayor número de imágenes para realizar la calibración. La ventaja que tiene es que el coste computacional se distribuye de forma homogénea a lo largo del todo el proceso, cosa que no ocurre en los otros métodos donde la mayor parte de este proceso se realiza después adquirir la última imagen.

Como era de esperar, el método de Agapito lineal es el más rápido. No llega a consumir un segundo cuando el número de imágenes procesadas es de 50, siendo unas 10 veces más rápido que la versión optimizada; aunque en este último caso no sobrepasa el segundo en el caso de que el número de imágenes procesadas sea inferior a 20.

En la figura 4.8 se muestran los tiempos de ejecución de cada uno de los métodos en

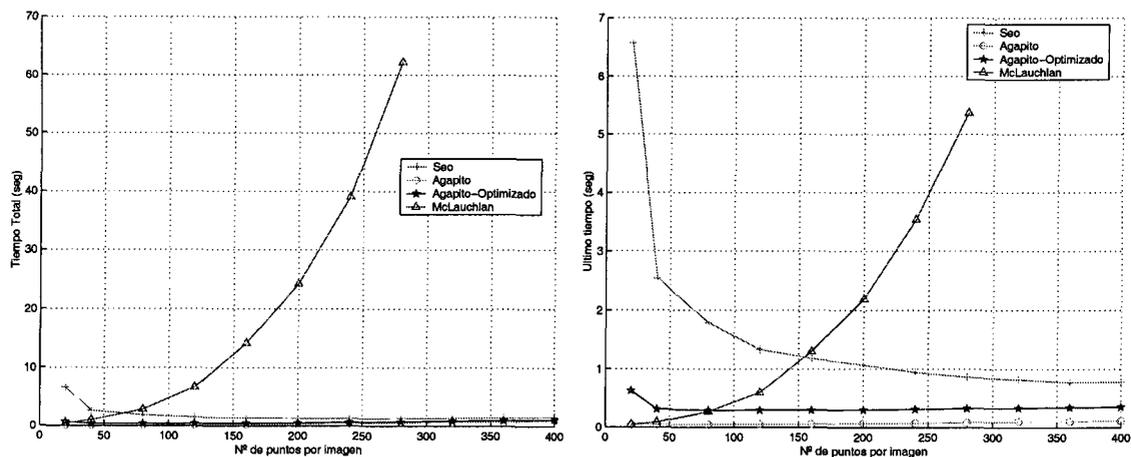


Figura 4.8: Tiempo de ejecución total y posterior a la última adquisición de imagen según el número de puntos procesados por imagen.

función del número de puntos detectados en cada vista de la secuencia.

Se observa que el método con mayor dependencia del número de puntos es el de McLauchlan, pero los resultados muestran que la carga computacional la distribuye de forma uniforme durante todo el proceso, mientras que el resto realizan la mayor parte del computo después de haber adquirido todas las imágenes.

Lo más llamativo de los resultados es el comportamiento de los métodos de Seo y Agapito optimizado, que requieren un mayor tiempo de ejecución cuando la cantidad de puntos procesados es pequeña. A priori, este resultado puede parecer ilógico, pero se justifica por el hecho de que al haber más puntos a procesar la parte lineal de estos métodos ajusta mejor los resultados, y por tanto, se necesitan menos pasos en la fase iterativa para alcanzar la convergencia de los resultados.

4.4.2. Estabilidad de los métodos

En este apartado se analiza la estabilidad de estos métodos, es decir, si a la hora de recuperar los parámetros intrínsecos de la cámara, éstos se mantienen constantes a lo largo de la secuencia, en la cual la orientación de la cámara ha ido variando.

Para ello se ha realizado un experimento sobre una secuencia de 40 imágenes, durante la cual los parámetros intrínsecos se han mantenido fijos. En los métodos preparados para ello se ha estimado el valor de cada parámetro de forma independiente para cada una de las vistas. En cambio, en el caso del método de McLauchlan el resultado es constante ya que devuelve un sólo valor para toda la secuencia. Un ruido gaussiano de media nula y desviación típica de 0.5 pixels ha sido añadido a las coordenadas de imagen para simular el efecto de cuantización en el proceso de formación de la misma. Cada

experimento se ha repetido 50 veces para cada método.

Los resultados obtenidos en este experimento se muestran en las figuras 4.9 - 4.11. En estas figuras se muestra tanto el valor promedio obtenido como el error en la estimación de cada parámetro. En la estimación del centro óptico se observa que todos los métodos presentan una pequeña dispersión en los resultados, cuando debería permanecer constante el valor obtenido de estos parámetros. El método que presenta la mayor dispersión en este parámetro es el método optimizado de Agapito, el cual muestra una variación máxima de 5 pixels. En el caso de los factores de escalado se observa que el resultado de todos los métodos es muy estable y apenas se obtienen diferencias entre ellos. Esta estabilidad en los parámetros intrínsecos está motivada porque los métodos de autocalibración realizan un ajuste global sobre la secuencia completa de imágenes.

En la estimación de la orientación se observa que el error es mínimo, apenas unas décimas de grado. En los valores obtenidos se puede apreciar el movimiento realizado por la cámara al desplazarse en azimut y elevación. Como puede observarse, se reproduce con precisión el movimiento descrito por la cámara.

4.4.3. Parámetros intrínsecos variables

Hay métodos como el de McLauchlan que modelan los parámetros intrínsecos como constantes. En cambio otros, como los métodos de Agapito y Seo, permiten estimar un valor distinto para algunos de estos parámetros en cada una de las vistas de la secuencia utilizada para la calibración. En este punto se trata de observar como incide esta circunstancia, principalmente en aquellos métodos que suponen que estos parámetros son fijos a lo largo de la secuencia. También se analiza si la variación, tanto de la distancia focal como del centro óptico, inciden en el resto de los métodos.

El estudio se centra en cómo pueden afectar dos elementos que se incluyen en muchas cámaras como son la presencia de enfoque automático o de zoom. Obviamente, estos elementos inciden de forma directa sobre la distancia focal, ya que producen desplazamientos bien de la lente o bien del plano imagen. Como ya se ha comentado, otro elemento que se ve afectado normalmente por estos elementos es el centro óptico [Willson, 1994].

Los experimentos presentados permiten analizar por separado y de forma aislada la incidencia de estos dos factores. Se ha realizado un primer experimento en el que se ha variado sólo la distancia focal y a continuación se ha realizado un segundo experimento variando sólo el centro óptico.

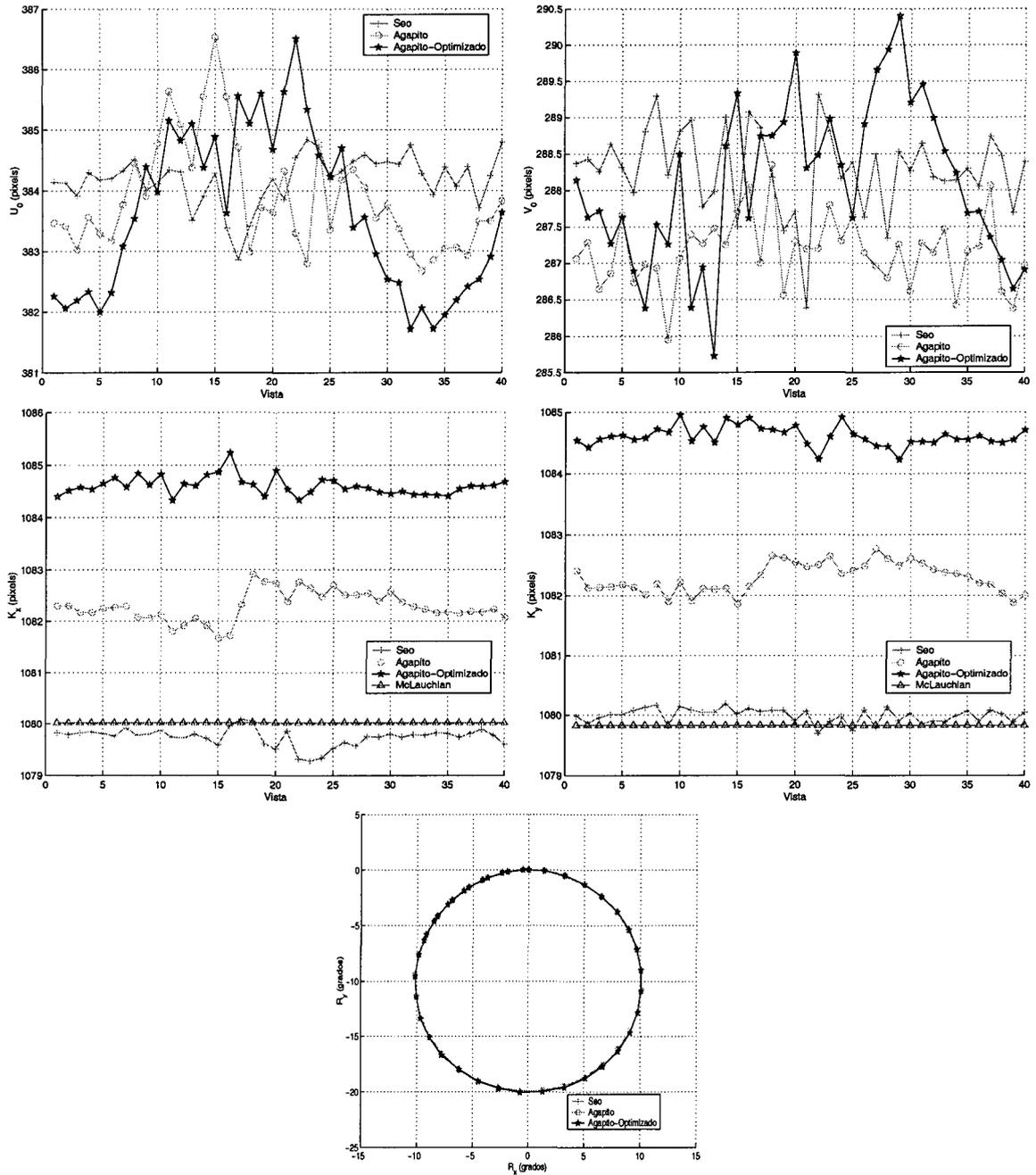


Figura 4.9: Valor promedio de cada parámetro para cada una de las vistas. Las imágenes tienen un ruido gaussiano de 0.5 píxeles en sus coordenadas 2D. Los parámetros intrínsecos permanecen constantes en la secuencia.

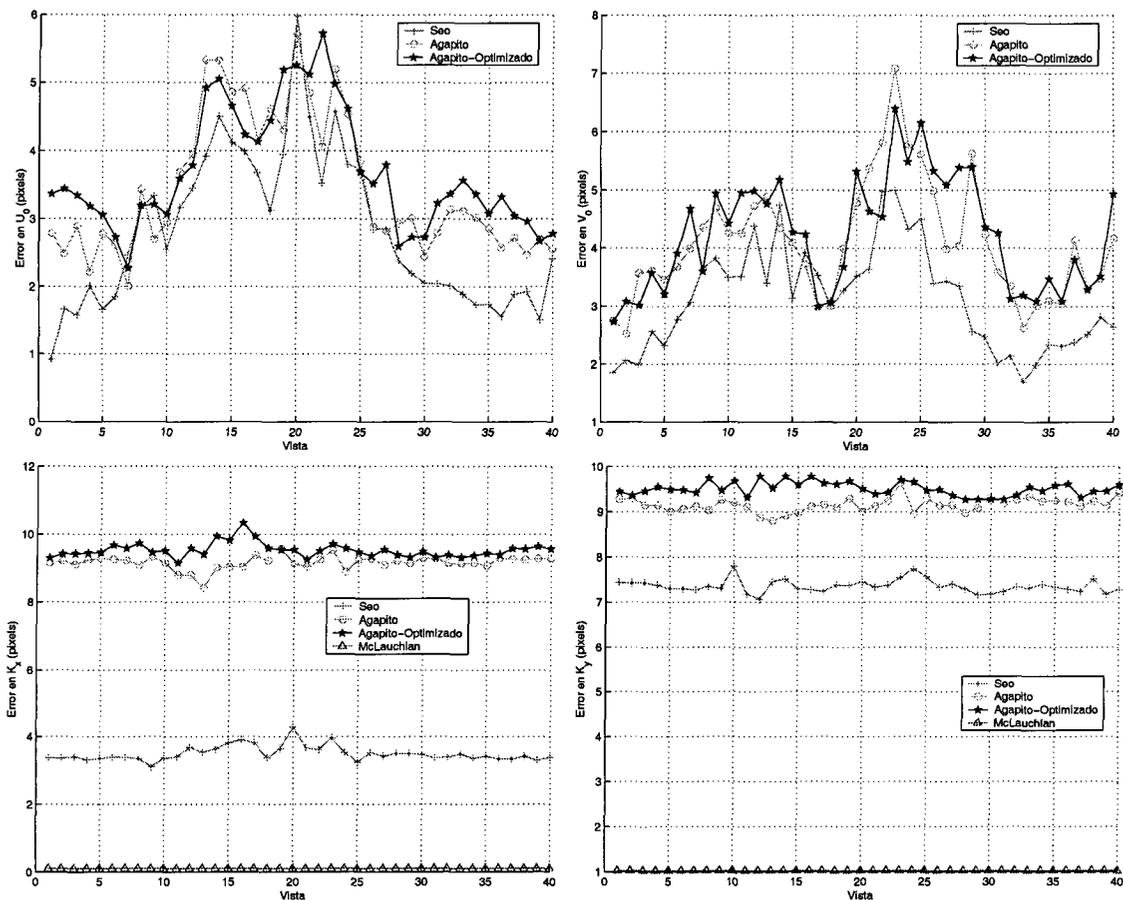


Figura 4.10: Error en la estimación de cada parámetro intrínseco para cada una de las vistas. Las imágenes tienen un ruido gaussiano de 0.5 pixels en sus coordenadas 2D. Los parámetros intrínsecos permanecen constantes en la secuencia.

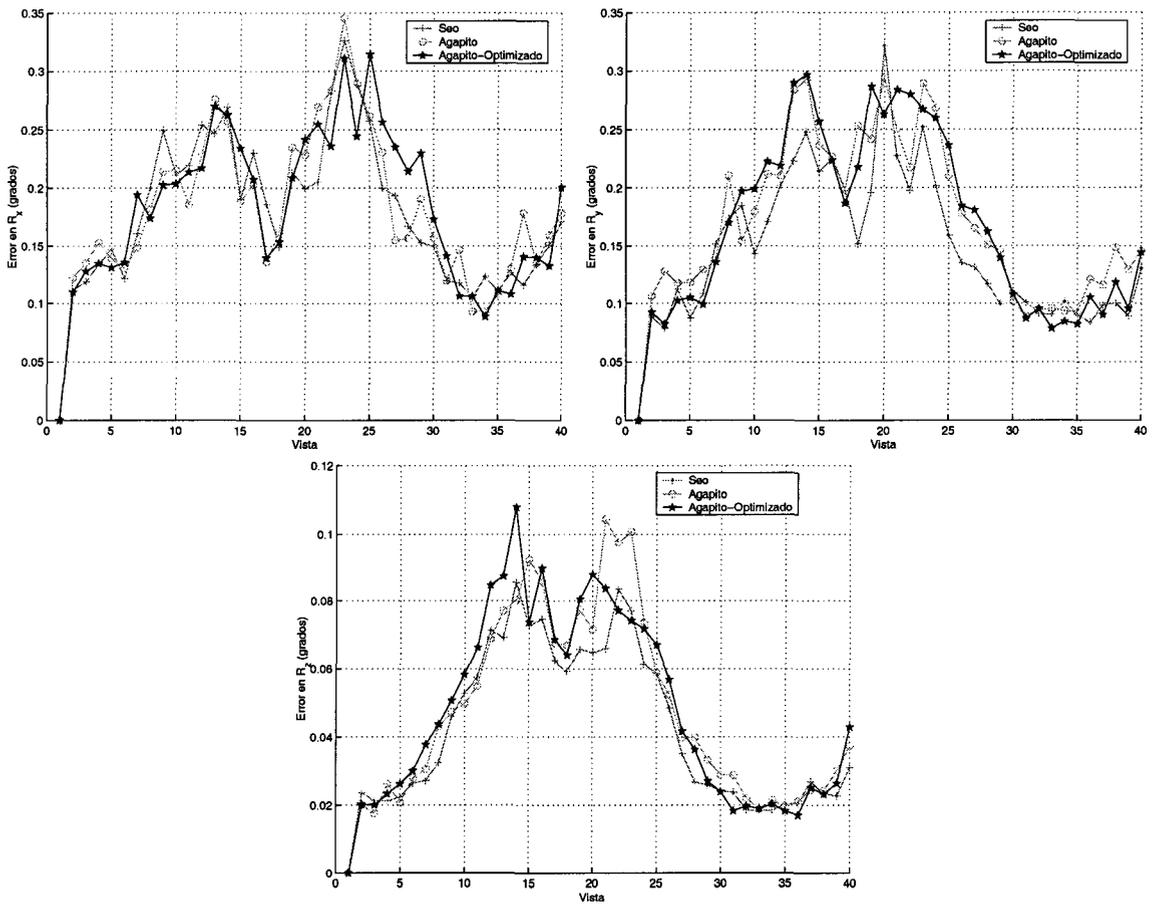


Figura 4.11: Error en la estimación de la orientación para cada una de las vistas. Las imágenes tienen un ruido gaussiano de 0.5 pixels en sus coordenadas 2D. Los parámetros intrínsecos permanecen constantes en la secuencia.

Distancia focal variable

En este apartado se analiza la influencia de la distancia focal sobre la estabilidad y precisión de los resultados. Para realizar la prueba se ha incrementado el valor de la distancia focal linealmente, siendo de 4.5 mm en la primera vista y de 13.5 en la última. Esto implica que cuando el número de imágenes es alto, en las últimas vistas el número de puntos procesados es pequeño debido que el campo de visión se reduce al aumentar la distancia focal. Por tanto, el error producido debe ser directamente proporcional al número de imágenes. En este caso, hay que indicar que los resultados del método de McLauchlan deben ser peores que los del resto debido a que modela la distancia focal como constante a lo largo de toda la secuencia.

Para realizar el experimento se ha utilizado una secuencia de 40 imágenes. Un ruido gaussiano de media 0 y desviación típica de 0.5 pixels ha sido añadido a las coordenadas de imagen para simular el efecto de cuantización en el proceso de formación de la misma. Cada experimento se ha repetido 20 veces para cada método.

Como se puede observar en las figuras 4.12 - 4.14 se observa que la dispersión de los resultados es mucho más amplia que la obtenida en el experimento anterior. Esta dispersión es más acentuada en la estimación del centro óptico, donde en algunos casos se supera los 100 pixels de amplitud en las curvas. En este parámetro se observa que el método que presenta mayor dispersión es el método lineal de Agapito. En general, se observa que el error aumenta a medida que la distancia focal es mayor debido a que el número de puntos procesados por imagen es menor, lo que genera una mayor incertidumbre en los resultados.

Como es lógico el método McLauchlan se muestra constante en su estimación ya que devuelve un sólo valor para los factores de escalado, mientras el resto de los métodos incrementan su valor en función de la variación de la distancia focal realizada. En cualquier caso, el resultado que devuelve este método es irrelevante y se incluye en la comparativa para servir de contraste a los resultados obtenidos por los otros métodos.

Centro óptico variable

En este apartado se analiza la influencia del centro óptico sobre la estabilidad y precisión de los resultados. Para realizar la prueba se ha variado de forma lineal durante la secuencia el valor del centro óptico en un rango de ± 20 pixels respecto al centro de la imagen, es decir, en la primera vista las coordenadas son (364,268) y en la última son (403,307). Para el experimento se ha utilizado una distancia focal de 4.5 mm. Por tanto, el hecho de que varíe uno de los parámetros intrínsecos durante la secuencia de imágenes utilizada para la calibración implica que los resultados del método de McLauchlan deben

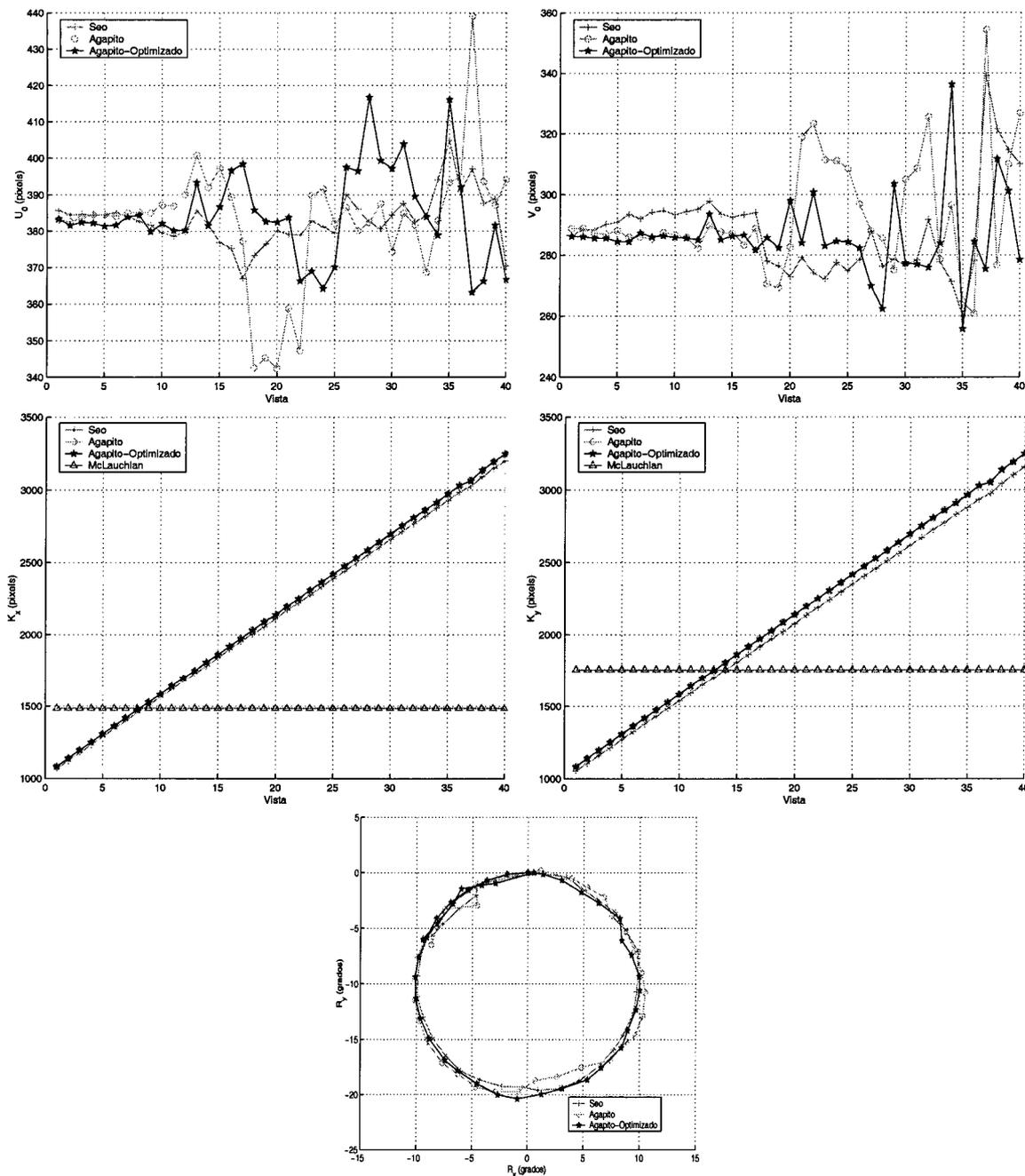


Figura 4.12: Valor promedio de cada parámetro para cada una de las vistas. Las imágenes tienen un ruido gaussiano de 0.5 pixels en sus coordenadas 2D. La distancia focal se incrementa linealmente a lo largo de la secuencia.

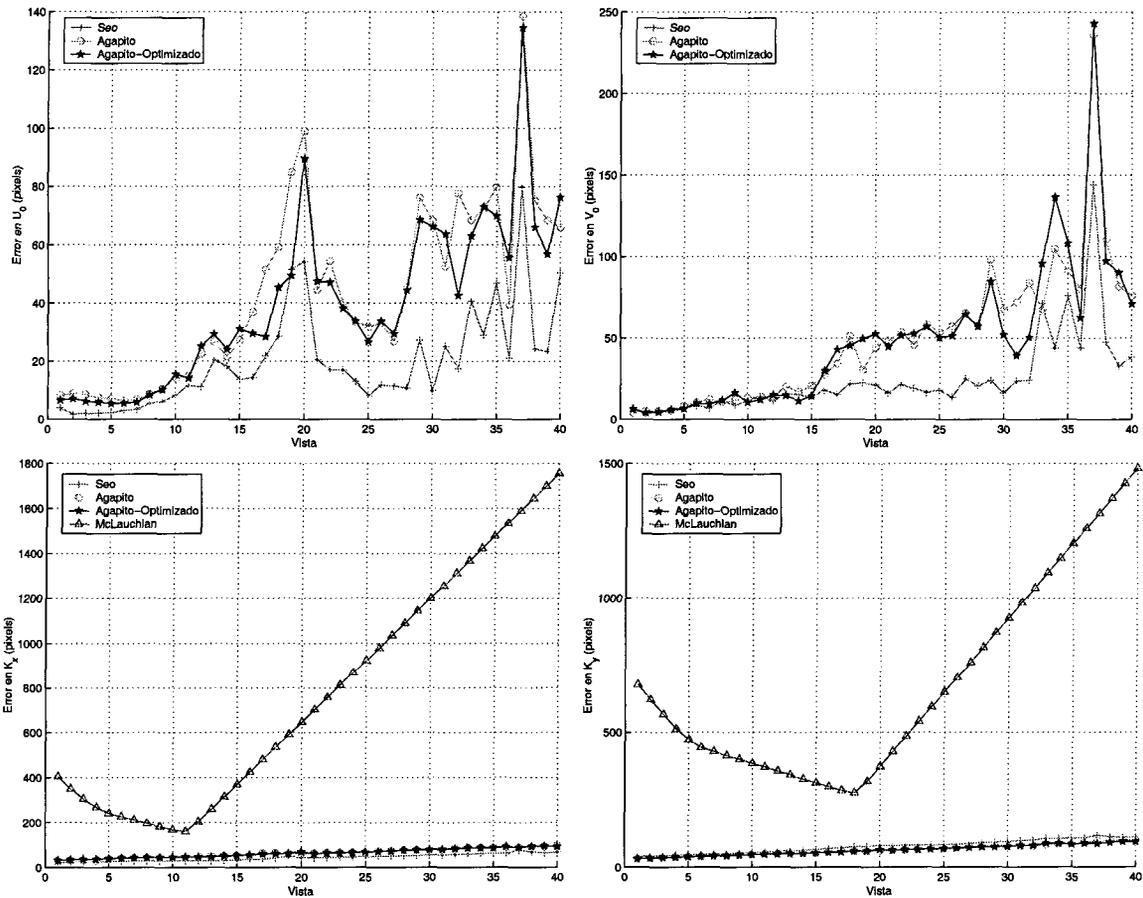


Figura 4.13: Error en la estimación de cada parámetro intrínseco para cada una de las vistas. Las imágenes tienen un ruido gaussiano de 0.5 píxeles en sus coordenadas 2D. La distancia focal se incrementa linealmente a lo largo de la secuencia.

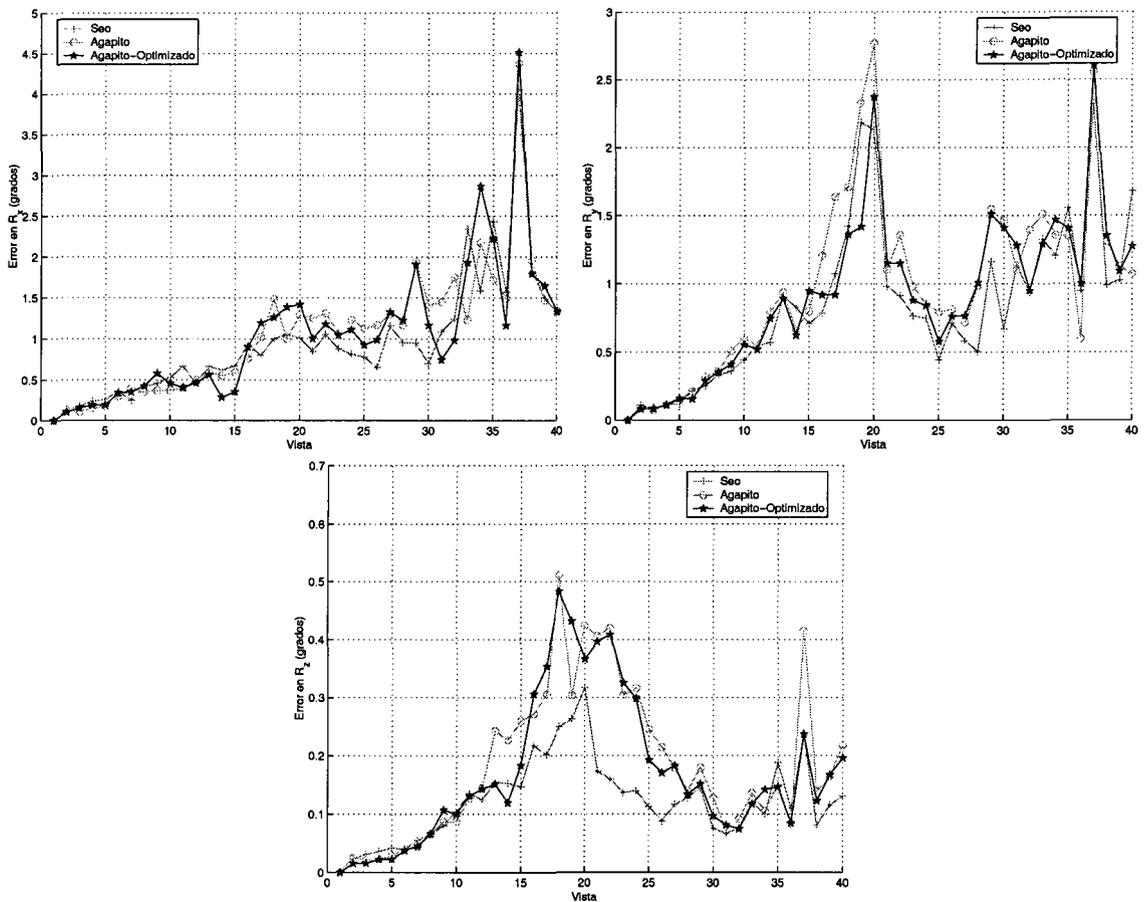


Figura 4.14: Error en la estimación de la orientación para cada una de las vistas. Las imágenes tienen un ruido gaussiano de 0.5 píxeles en sus coordenadas 2D. La distancia focal se incrementa linealmente a lo largo de la secuencia.

ser peores que los del resto tal y como ya se ha comentado.

Para realizar el experimento se ha utilizado una secuencia de 40 imágenes. Un ruido gaussiano de media 0 y desviación típica de 0.5 pixels ha sido añadido a las coordenadas de imagen para simular el efecto de cuantización en el proceso de formación de la misma. Cada experimento se ha repetido 30 veces para cada método. Los resultados obtenidos se muestran en las figuras 4.15 a 4.17.

En los resultados se observa que en el caso de los métodos de Agapito la variación de este parámetro no tiene influencia en los resultados. En cambio, el método de Seo se ve afectado principalmente en tres parámetros, todos ellos en la componente horizontal: u_0 , k_x y R_y . Aún estos parámetros el error es muy pequeño en comparación con el producido por la variación de la distancia focal. En cierta medida es lógico que el error en este método sea mayor que el método de Agapito. La razón de esto es que el método de Seo hace coincidir el centro óptico con el centro de la imagen en la primera fase del método, por tanto, al no coincidir éstos en la realidad, esto hace que los valores de los factores de escalado obtenidos en esta fase, los cuáles son utilizados posteriormente como aproximación inicial en el proceso de optimización, no sean precisos. Por último, y como era de esperar, el resultado del método de McLauchlan también se ve afectado por la variación del centro óptico durante la secuencia de imágenes, pero la incidencia es mucho menor que la producida por la distancia focal.

4.4.4. Influencia del nivel de ruido en las coordenadas 2D de la imagen

En este experimento se muestra el comportamiento de los distintos métodos ante el aumento del nivel de ruido inducido sobre las coordenadas 2D de la imagen. Este ruido se modela como un ruido gaussiano de media cero y la desviación típica varía entre cero y 1.5 pixels. En cada método se ha usado el número de imágenes que utilizan los autores en los experimentos que muestran en sus respectivos trabajos, de tal forma que para los métodos de Agapito la secuencia es de 30 imágenes, para el de Seo es de 4 imágenes y en el método de McLauchlan varía, ya que depende de cuando converge o se estabiliza el error (se ha puesto un máximo de 300 imágenes). Tanto el centro óptico como la distancia focal son constantes a lo largo de cada secuencia.

Los resultados obtenidos se muestran en la figuras 4.18 y 4.19. En el primer conjunto de gráficas se muestran los errores cometidos por los métodos en la estimación de los parámetros intrínsecos, mientras que en el segundo grupo se muestran el error en la estimación de los parámetros extrínsecos y el error RMS en las coordenadas 2D de la imagen.

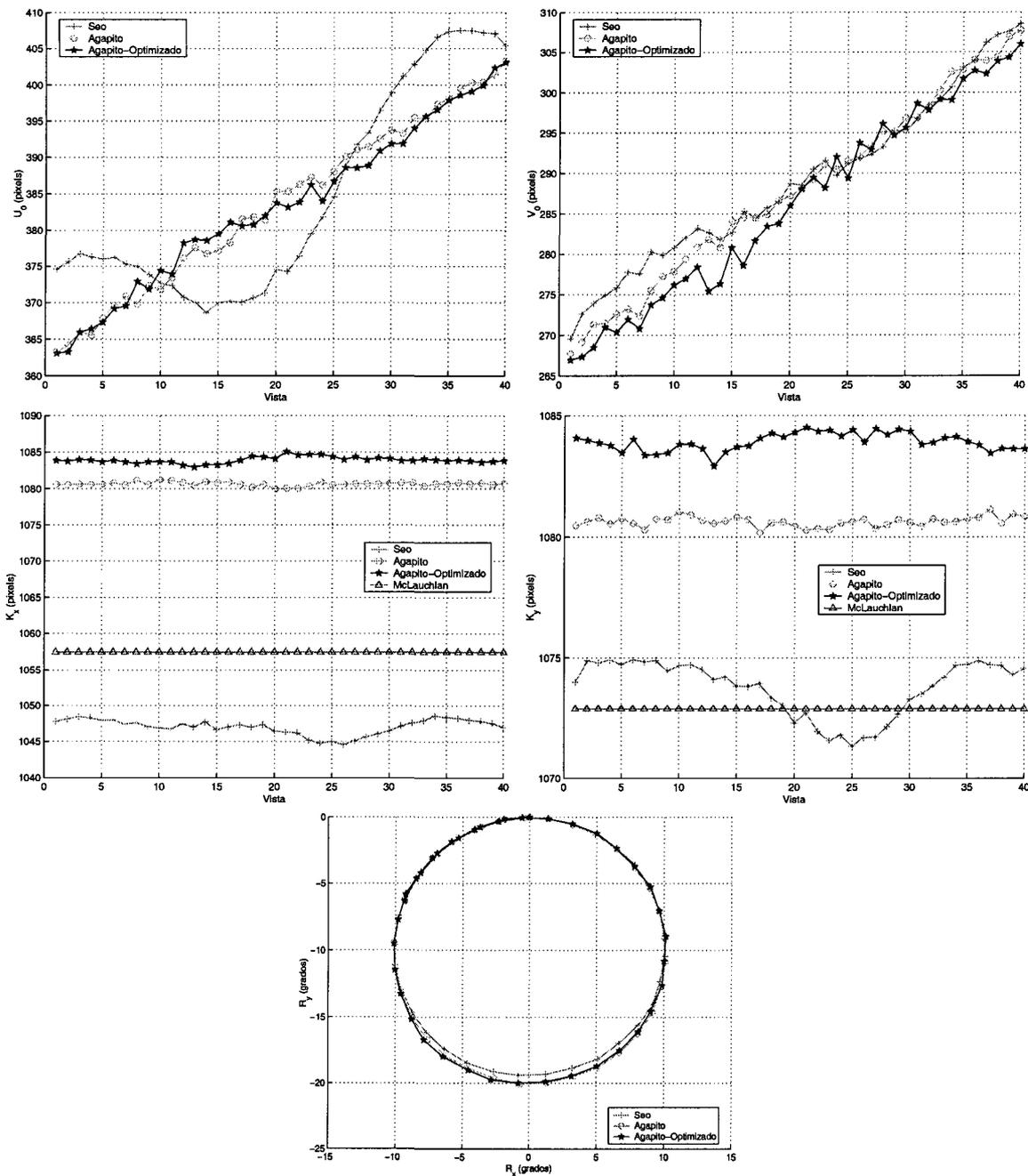


Figura 4.15: Valor promedio de cada parámetro para cada una de las vistas. Las imágenes tienen un ruido gaussiano de 0.5 pixels en sus coordenadas 2D. El centro óptico se incrementa linealmente a lo largo de la secuencia.

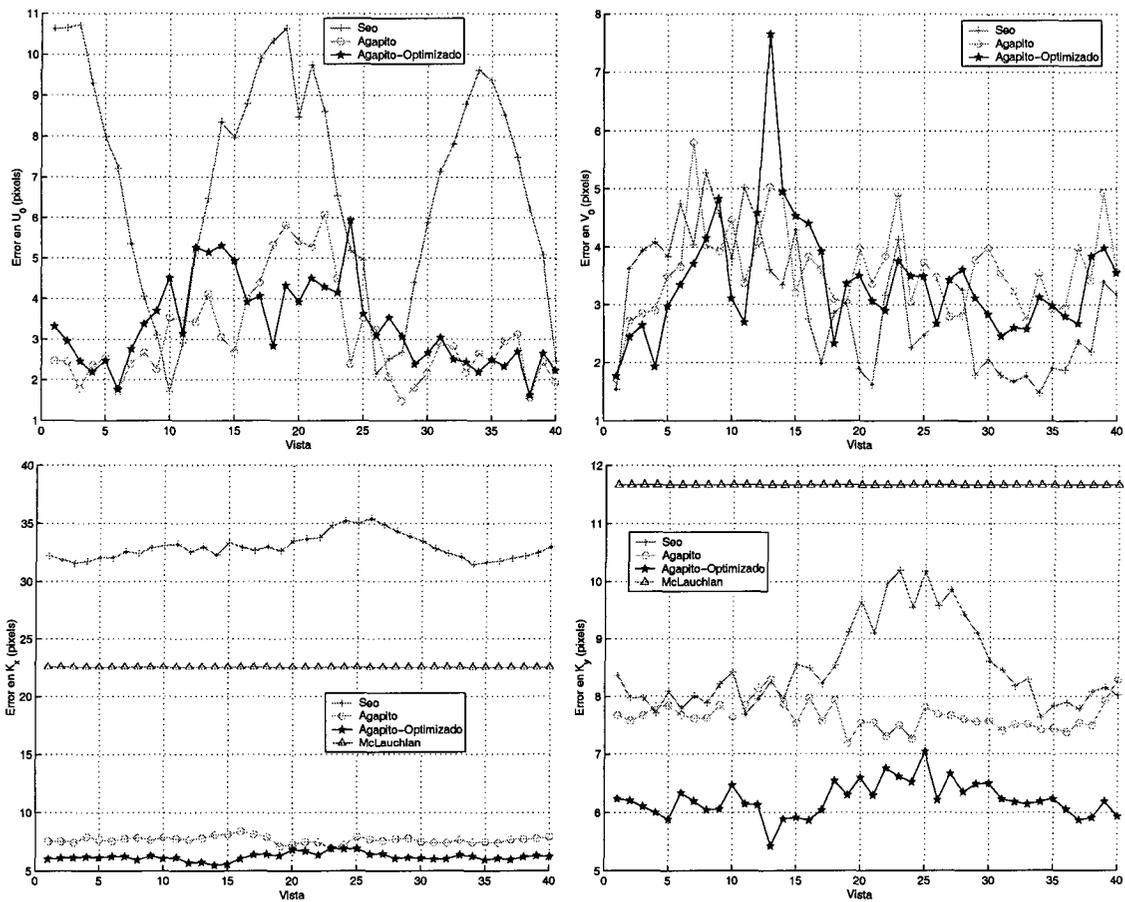


Figura 4.16: Error en la estimación de cada parámetro intrínseco para cada una de las vistas. Las imágenes tienen un ruido gaussiano de 0.5 pixels en sus coordenadas 2D. El centro óptico se incrementa linealmente a lo largo de la secuencia.

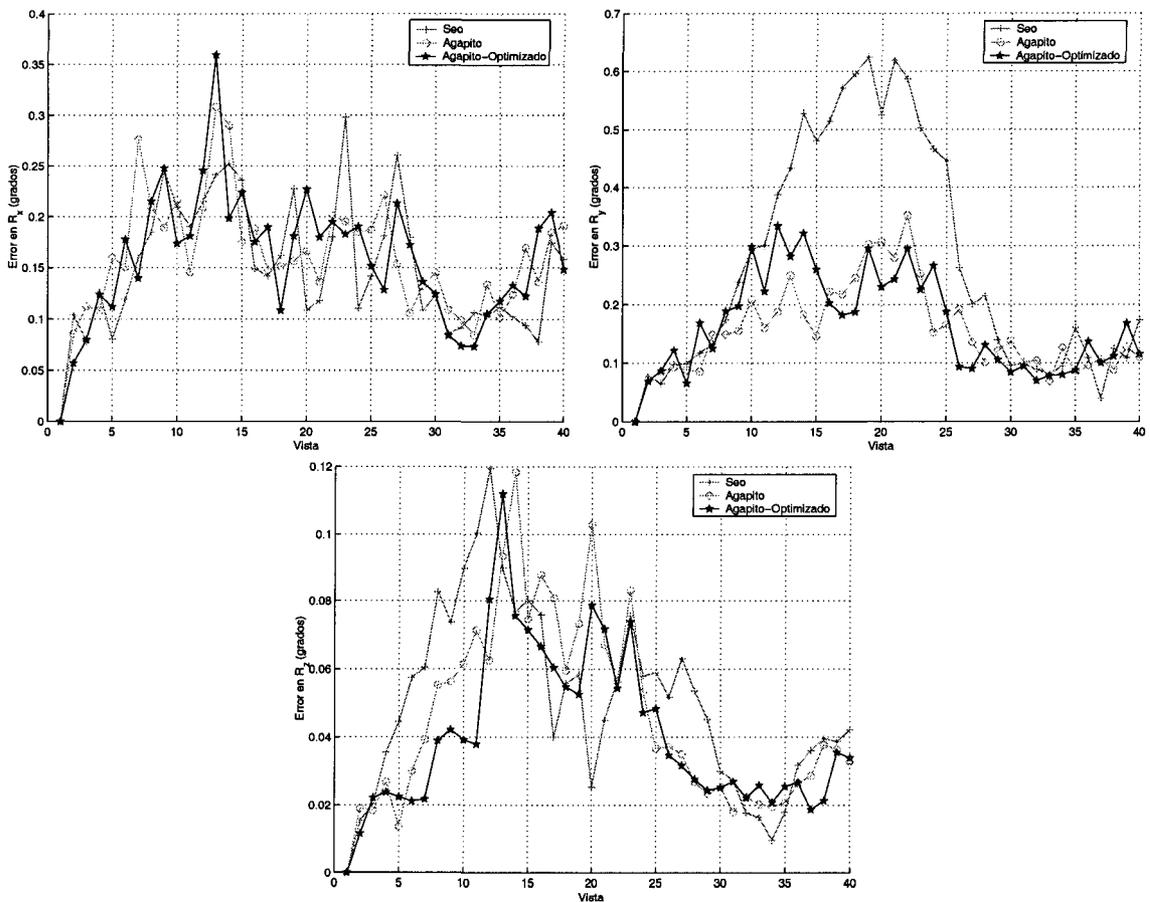


Figura 4.17: Error en la estimación de la orientación para cada una de las vistas. Las imágenes tienen un ruido gaussiano de 0.5 pixels en sus coordenadas 2D. El centro óptico se incrementa linealmente a lo largo de la secuencia.

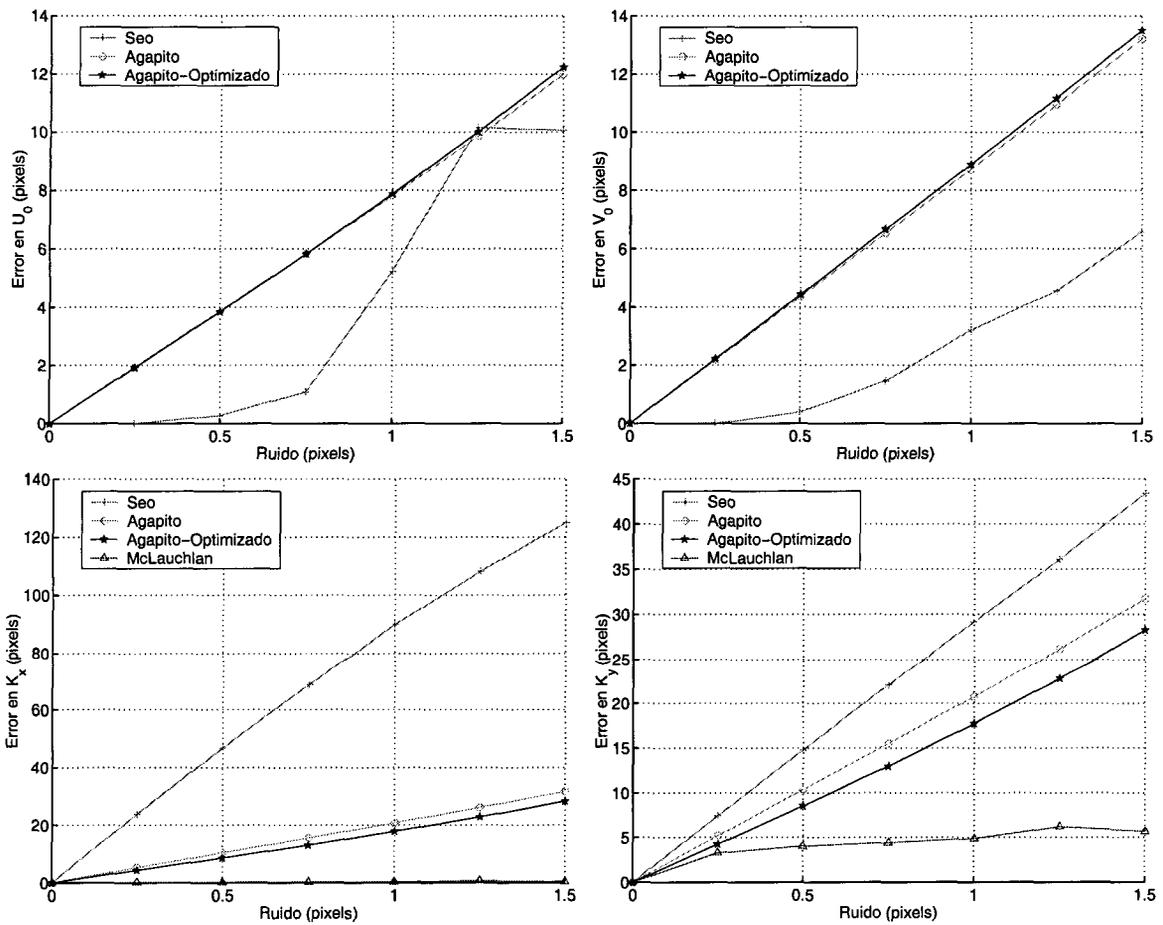


Figura 4.18: Error de cada parámetro intrínsecos cuando varía el nivel de ruido gaussiano en las coordenadas 2D.

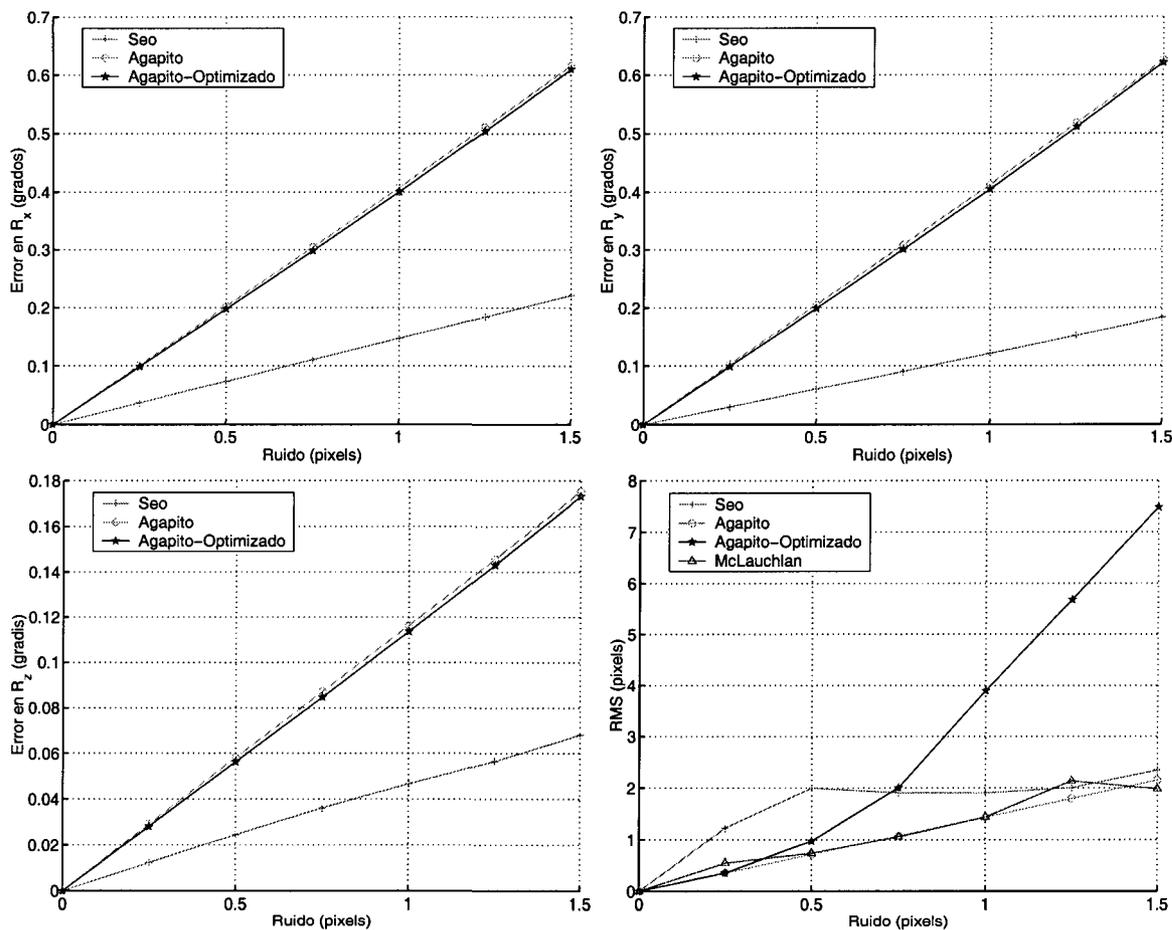


Figura 4.19: Error en la orientación y RMS sobre las coordenadas 2D cuando varía el nivel de ruido gaussiano en estas coordenadas.

Por lo general, se observa un crecimiento lineal y uniforme del error a medida que aumenta el nivel de ruido. Entre los resultados cabe destacar que el método de Seo tiene mejor comportamiento que los métodos de Agapito en todas las gráficas excepto en las relativas a los factores de escalado, aunque como se observó en los experimentos sobre la variación del número de imágenes, este factor es el que más influye en la sensibilidad del método. También es destacable el bajo nivel de error en la estimación de los factores de escalado en el método de McLauchlan, aunque hay que tener en cuenta que tanto la distancia focal como el centro óptico permanecen constantes en el experimento. Tras la realización de este experimento se ha comprobado que este método no ha necesitado más de 30 imágenes en ninguno de los casos.

Por último, llama la atención el alto nivel de error en la reconstrucción de los puntos (error RMS), en el método de Agapito optimizado cuando el nivel de ruido es alto.

4.4.5. Parámetros intrínsecos no modelados por los métodos

En esta sección se analiza cómo repercute en el resultado de los métodos los efectos producidos por parámetros que no son contemplados en su modelo. Hay que tener en cuenta que en los métodos de autocalibración se modelan muchos menos parámetros que en los métodos de calibración manual. En este caso se analiza la influencia de parámetros como la distorsión, la ortogonalidad de los ejes del plano imagen, el factor de escalado y el desalineamiento entre el eje de giro de la cámara y el punto focal.

Imágenes con distorsión

Aunque en algunos de los métodos analizados se indica que la distorsión se puede evitar realizando un preprocesamiento de la imagen donde se corrige dicho efecto, aquí se analiza la influencia de la distorsión (modelada sólo con la componente radial) sobre cada uno de los parámetros sin realizar dicho preprocesamiento.

Para realizar el estudio, se ha variado el valor del primer coeficiente de distorsión radial (k_1) entre $-0,02$ y -10^{-6} . Las imágenes han sido modeladas libres de ruidos y manteniendo la distancia focal y el centro óptico fijos. Para todos los métodos se han utilizado secuencias de 15 imágenes, excepto en el de McLauchlan en el que la longitud de la secuencia esta determinada por la convergencia del método.

Los resultados del experimento se pueden ver en las figuras 4.20 y 4.21. Las gráficas muestran los resultados en escala logarítmica debido al amplio rango que cubren los datos. Hay que indicar que en las gráficas los métodos de Agapito y Agapito-optimizado aparecen superpuestos.

En estas gráficas se observa que para una distorsión de magnitud mayor a $2 \cdot 10^{-3}$, el error en la obtención del centro óptico en los métodos de Agapito es desmesurado. Como en casos anteriores el error en la estimación de estos parámetros en el método de Seo es mucho menor. Sólo con magnitudes elevadas de k_1 (a partir de 10^{-2}) ofrece este método errores muy elevados.

Los factores de escalado son los parámetros más afectados por el efecto de la distorsión ya que con valores pequeños distorsión ($k_1 = -10^{-4}$) el error producido alcanza los 10 pixels/mm. La excepción es el método de McLauchlan que no presenta niveles altos de error hasta que la magnitud de k_1 es superior a 10^{-3} .

También cabe mencionar que la incidencia de la distorsión sobre los parámetros extrínsecos es inferior a la que tiene sobre los parámetros intrínsecos.

Es importante indicar que los resultados de los métodos de Agapito, tanto el lineal como el iterativo, no convergen en determinadas configuraciones de cámara si la magnitud de k_1 es elevada. Esto es debido a que la matriz resultante de la cónica absoluta

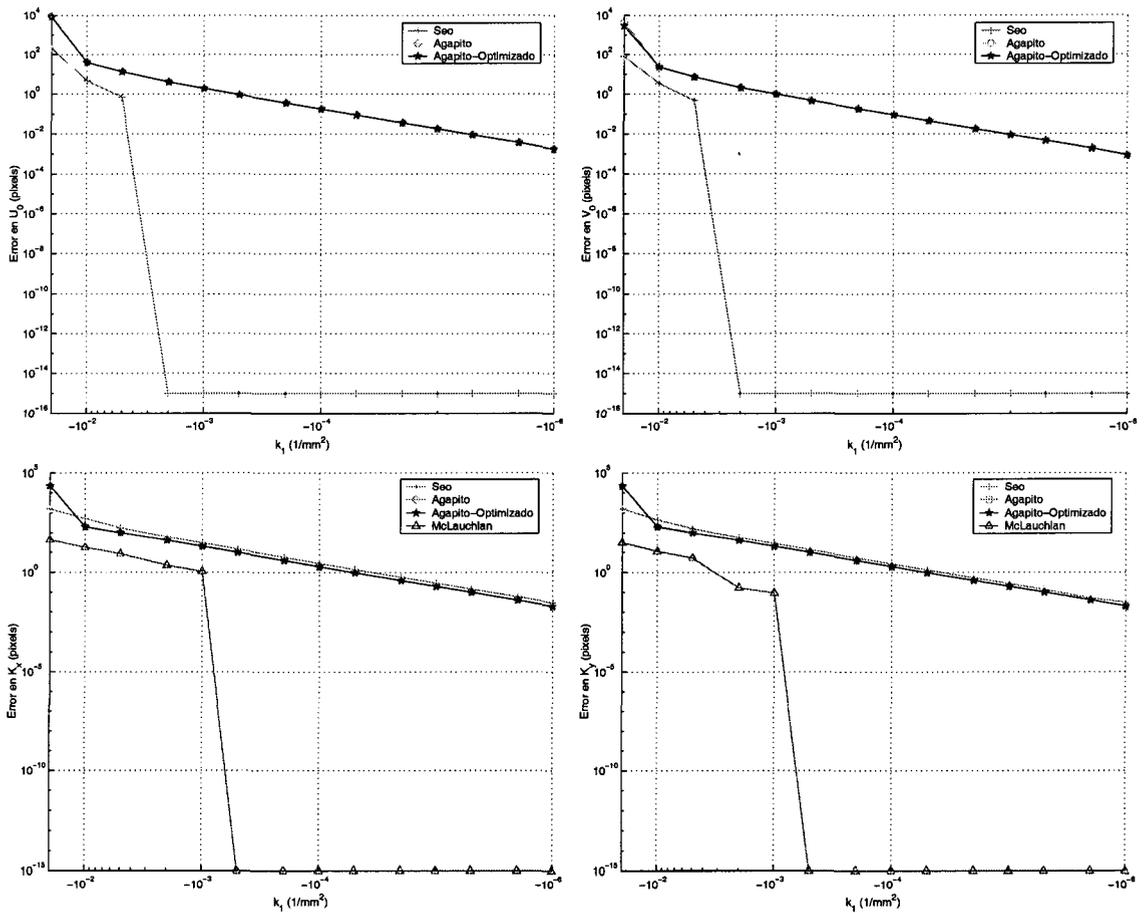


Figura 4.20: Error de cada parámetro intrínseco cuando varía el primer coeficiente de distorsión radial (k_1).

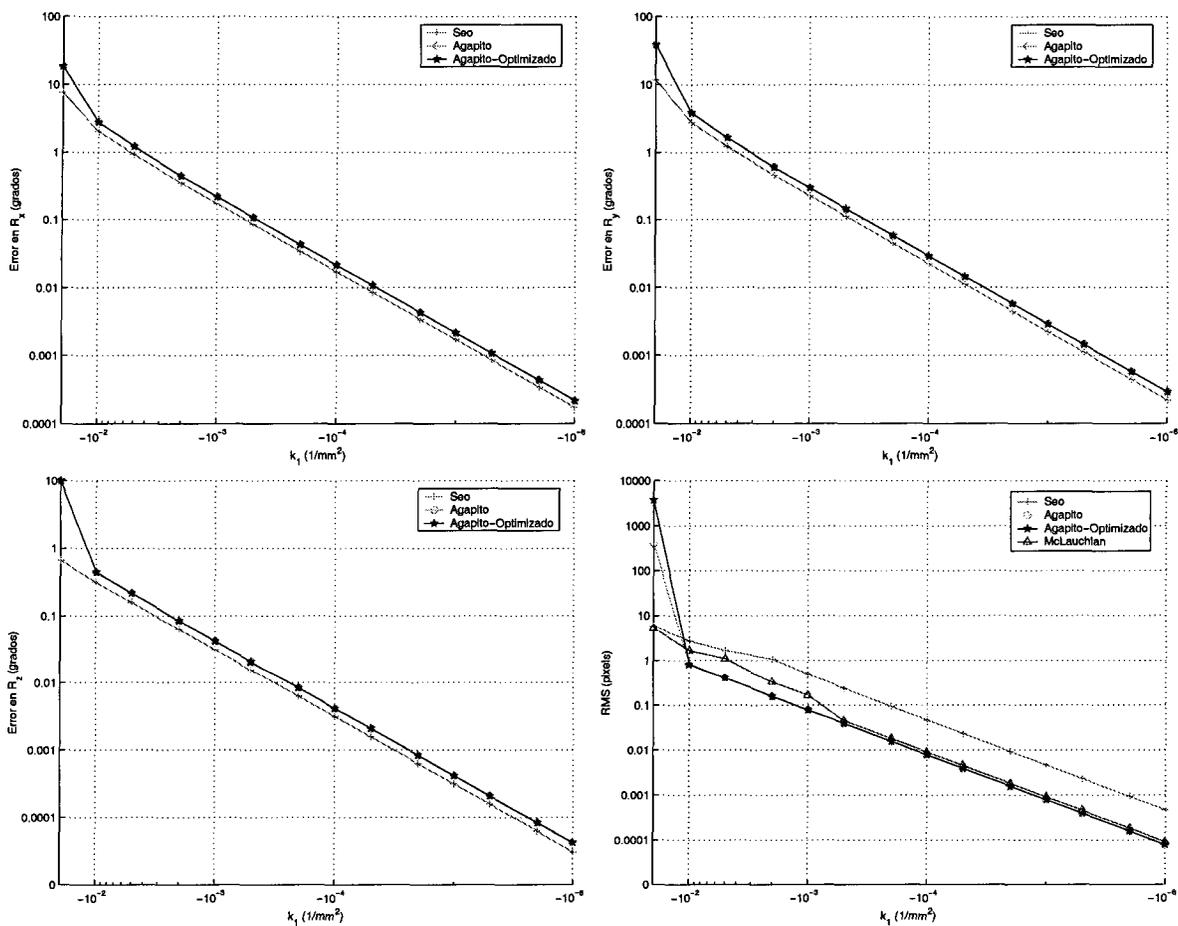


Figura 4.21: Error en la orientación y RMS sobre las coordenadas 2D cuando varía el primer coeficiente de distorsión radial (k_1).

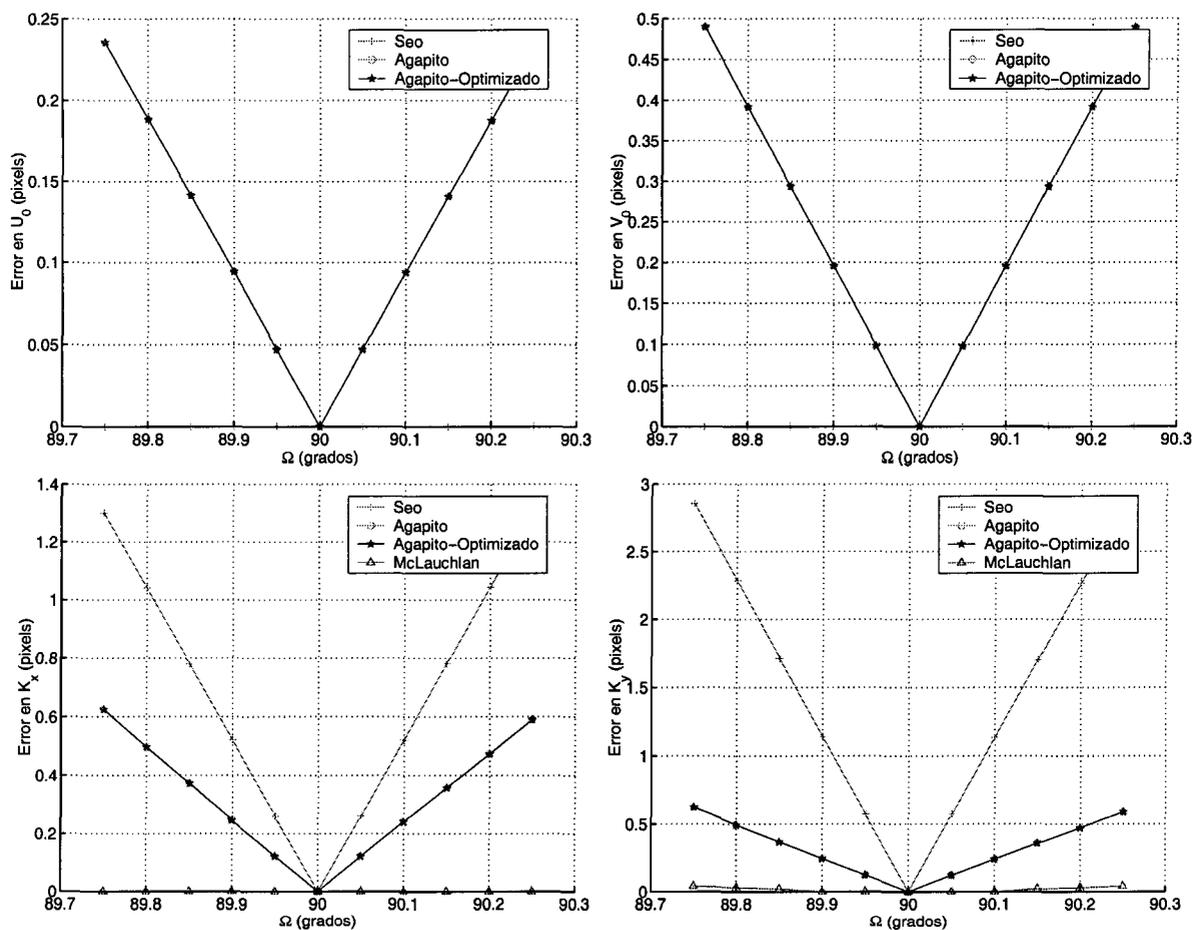


Figura 4.22: Error de cada parámetro intrínseco cuando varía el grado de ortogonalidad de los ejes del plano imagen.

no es definida positiva y, por tanto, no se puede extraer de ella la matriz de parámetros intrínsecos.

En resumen, es obvio que se hace necesario realizar un preprocesamiento de las imágenes donde se corrija la distorsión para que estos métodos puedan dar un resultado aceptable, debido a que se observa que este factor tiene una clara incidencia, provocando que el error en los resultados aumente exponencialmente con el nivel de distorsión.

Ortogonalidad de los ejes del plano imagen

En la mayoría de los métodos de autocalibración este parámetro se supone constante, asumiéndose que estos ejes son perfectamente ortogonales. Esto es debido a que proporciona una restricción adicional que permite que el sistema sea resoluble o que se disminuya el número de imágenes necesarias para que lo sea.

Para realizar el estudio, se ha variado el valor de Ω entre 89.75° y 90.25° . Las imágenes han sido modeladas libres de ruidos y manteniendo la distancia focal y el centro óptico

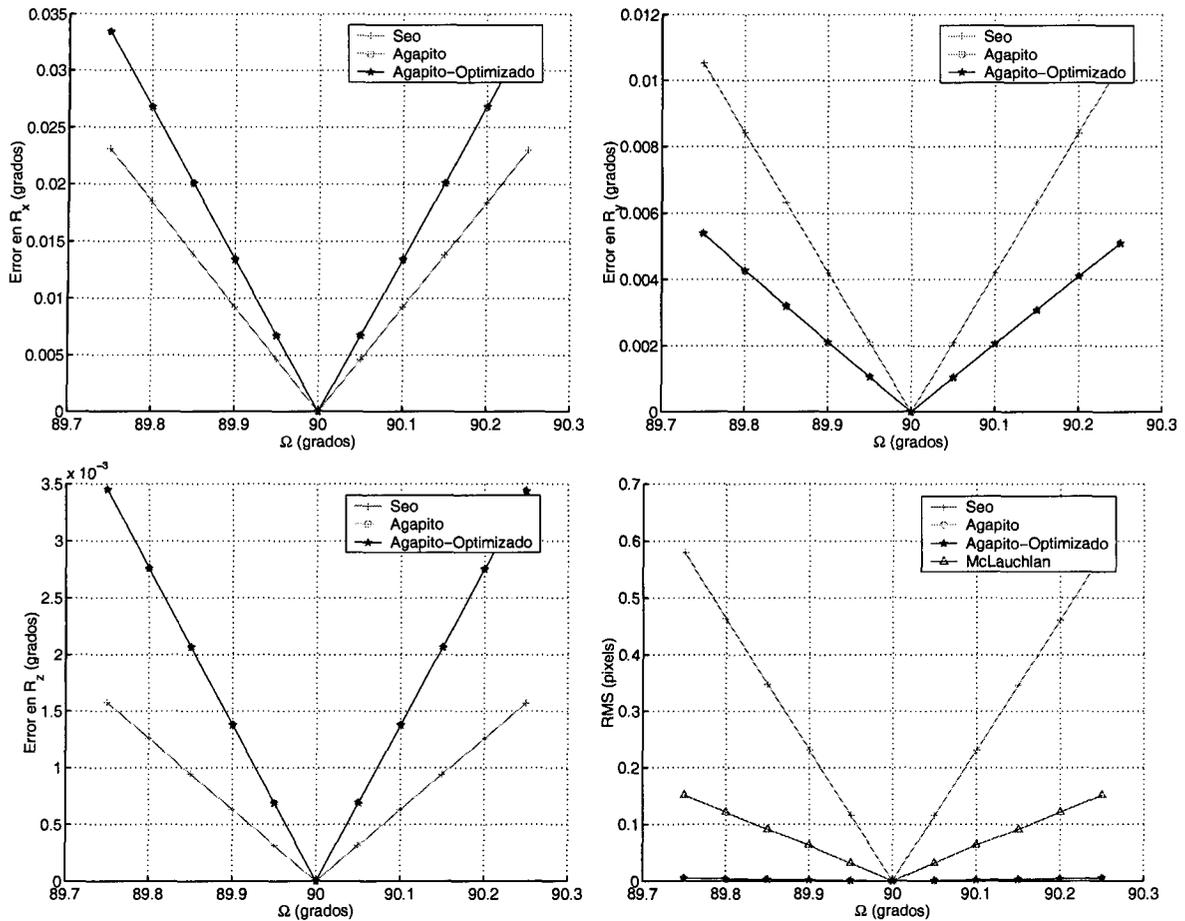


Figura 4.23: Error en la orientación y RMS sobre las coordenadas 2D cuando varía el grado de ortogonalidad de los ejes del plano imagen.

fijos. Para todos los métodos se han utilizado secuencias de 15 imágenes.

Como se puede ver en los resultados del experimento (figuras 4.22 y 4.23) la incidencia del error en cada uno de los parámetros es mínima y proporcional a la desviación respecto a la ortogonalidad de estos ejes. Aunque en cada uno de los parámetros la influencia es mínima, el error en reconstrucción de los puntos de la imagen es equivalente a haber inducido un ruido de 0.3 pixels en las coordenadas 2D.

En el caso del centro del eje óptico la curva del método de Seo se confunde con el eje de abscisas debido a que las imágenes están libres de ruido y el centro óptico coincide con el centro de la imagen. En cambio, en los factores de escalado el error del método de Seo es el de mayor magnitud. Como se ha mostrado en otros experimentos este parámetro es el que absorbe la mayor parte del error en este método. El error en la estimación de la orientación es de apenas unas centésimas de grado y en el caso de R_z mucho más atenuado que en los otros dos ángulos.

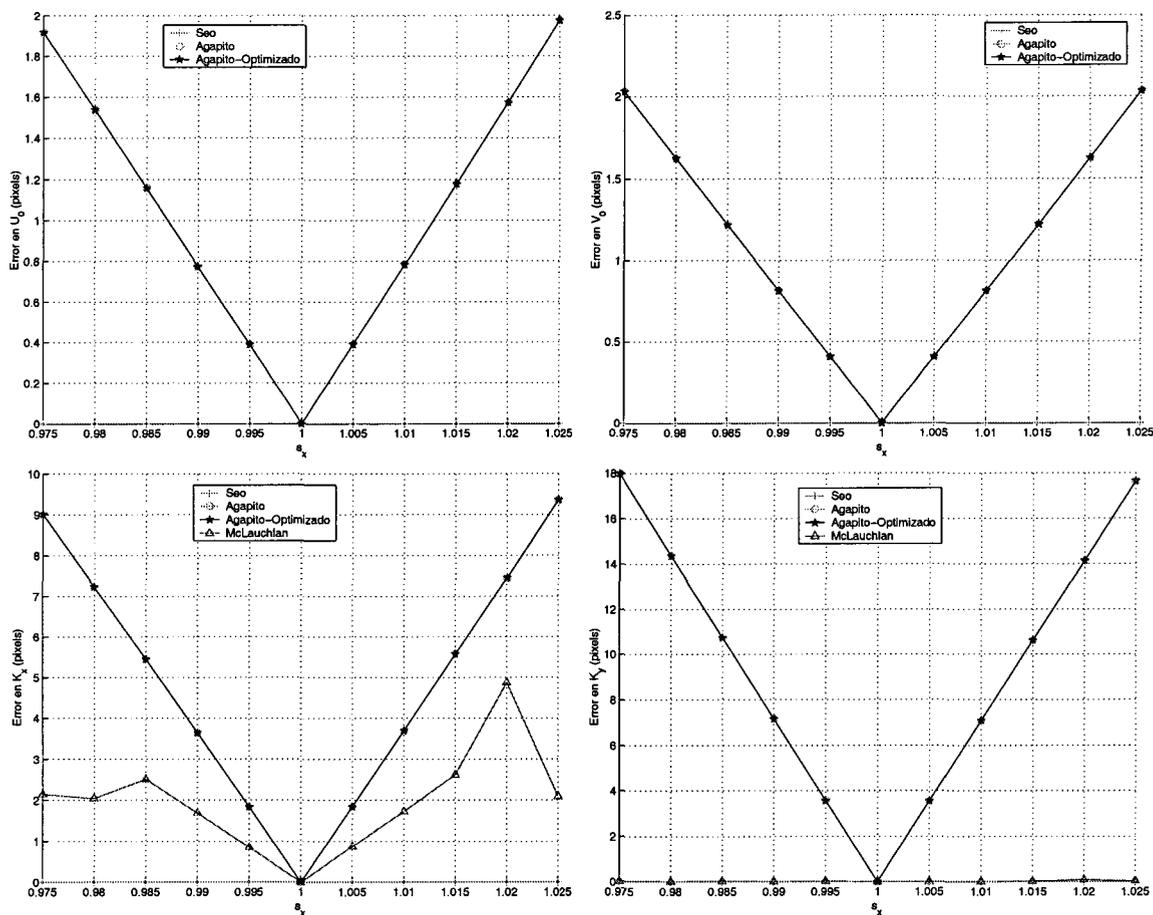


Figura 4.24: Error de cada parámetro intrínseco cuando varía el factor de proporción.

Factor de proporción

Es interesante realizar el análisis de este parámetro ya que muchos métodos de autocalibración que los pixels son perfectamente cuadrados o que el factor de proporción es conocido. Para realizar el estudio, se ha variado el valor de s_x entre 0.975 y 1.025. Las imágenes han sido modeladas libres de ruidos y manteniendo la distancia focal (4.5 mm) y el centro óptico fijos. Para todos los métodos se han utilizado secuencias de 15 imágenes, salvo, como en anteriores casos, con el método de McLauchlan.

En las figuras 4.24 y 4.25 se observa que la incidencia de este factor es significativa, aunque no determinante. En el caso del método de Seo el error es nulo en todos los parámetros, por lo que su curva esta superpuesta al eje de abscisas en todas las gráficas. Esto se debe al hecho de que las imágenes están libres de ruidos y distorsión, los ejes del plano imagen son ortogonales y centro óptico coincide con en el centro de la imagen.

En las gráficas las curvas de las dos variantes del método de Agapito aparecen superpuestas. En la estimación de cada uno de los parámetros el error de estos dos métodos es proporcional a la desviación de s_x respecto a la unidad. El método de McLauchlan se

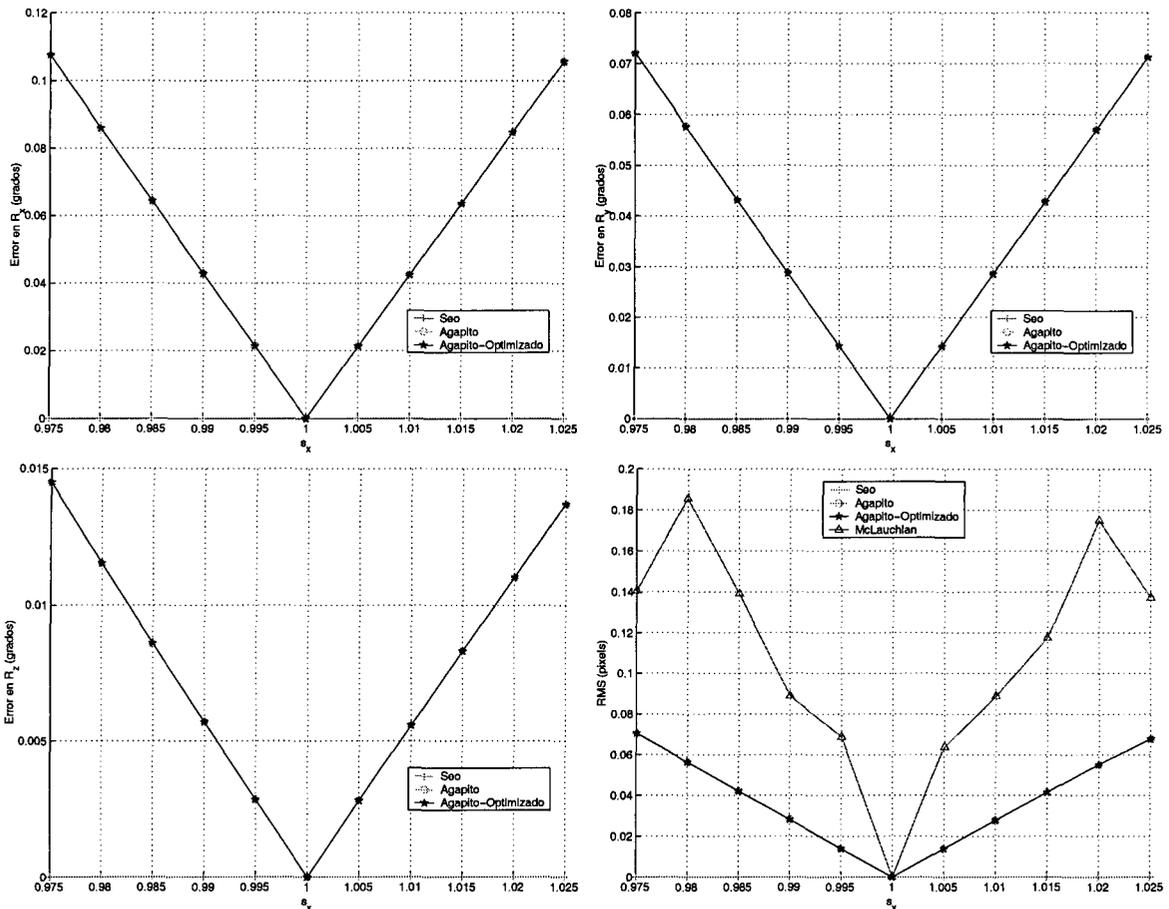


Figura 4.25: Error en la orientación y RMS sobre las coordenadas 2D cuando varía factor de proporción.

ve afectado principalmente en k_x , aunque en menor medida que los métodos de Agapito. Por último, en la gráfica del error RMS se observa que aunque el método de Agapito es el que peor estima cada uno de los parámetros obtiene menor error global que el método de McLauchlan debido a que los errores cometidos en los parámetros se compensan entre sí.

Desalineamiento entre los sistema de coordenadas de la cámara y del plano imagen

En este experimento se muestra cómo influye en cada parámetro el hecho de que la posición del foco no coincida con el centro de giro de la cámaras. En general, los métodos de autocalibración basados en rotaciones suponen que dichos puntos coinciden en el espacio. Esta coincidencia es difícil que se produzca en la realidad debido a razones mecánicas, que incluso se ven complicadas por el uso de lentes con zoom que, aún sin realizar rotaciones, desplazan el foco a lo largo del eje óptico. Tanto en

[Hayman and Murray, 2002] como en [Wang et al., 2001] se describe con más detalle este problema, además de presentar algunos resultados teóricos y prácticos que se intentarán corroborar en este experimento. Éste pretende extender el estudio realizado en esos trabajos al caso de imágenes con ruido, además de realizar un análisis completo de los efectos en cada uno de los parámetros que se obtienen.

Para realizar estos experimentos se han utilizado secuencias de 15 imágenes para todos los métodos, excepto en el de McLauchlan ya que este método finaliza cuando el error converge o se estabiliza. En cada secuencia los parámetros intrínsecos permanecen constantes.

Los experimentos se han realizado tanto para analizar los desplazamientos en la vertical (Y_c), como en la dirección del eje óptico (Z_c). En cada caso se ha estudiado el efecto en ausencia de ruido y con presencia de ruido gaussiano de media nula y desviación típica de 0.5 pixels en las coordenadas 2D de la imagen. En el caso de imágenes sin ruido el rango de variación de la posición del foco respecto al eje de giro de la cámara alcanza un centímetro en cada sentido, mientras en las imágenes con ruido se ha repetido cada prueba 50 veces y el rango alcanza el medio centímetro en cada sentido.

En primer lugar, se analizan los efectos producidos por el desalineamiento vertical entre el foco y el eje de giro en elevación. Los resultados obtenidos en el caso de imágenes sin ruido se muestran en la figuras 4.26 y 4.27. En el primer conjunto de gráficas se muestran los errores cometidos por los métodos en la estimación de los parámetros intrínsecos, mientras que en el segundo grupo se muestran el error en la estimación de los parámetros extrínsecos y el error RMS en las coordenadas 2D de la imagen.

Como se puede observar en todas las gráficas, la incidencia de este parámetro es mínima. El comportamiento es el mismo para los métodos de Seo y Agapito, ya que casi todas las curvas tienen forma de "V". La excepción la presenta el método de Seo que estima perfectamente el centro del eje óptico y por tanto su curva se confunde con el eje de abscisas, esto es debido a que este parámetro coincide con el centro de la imagen en la configuración utilizada para realizar la simulación. Por contra, en los factores de escalado el método de Seo es el que peor se comporta y en la estimación de la orientación los métodos de Agapito y Seo tienen resultados semejantes. Por otro lado, el método de McLauchlan tiene un error prácticamente nulo hasta los 3 mm de desalineamiento, pero cuando la separación es mayor aumenta aunque de forma irregular. Por último, se observa que las dos versiones del método de Agapito, las cuales se superponen en las gráficas, son las más precisas a la hora de reconstruir la posición de los puntos en la imagen (menor error RMS).

Las gráficas de las figuras 4.28 y 4.29, muestran los resultados de este experimento en imágenes con ruido gaussiano de media nula y desviación de 0.5 pixels, confirman la poca

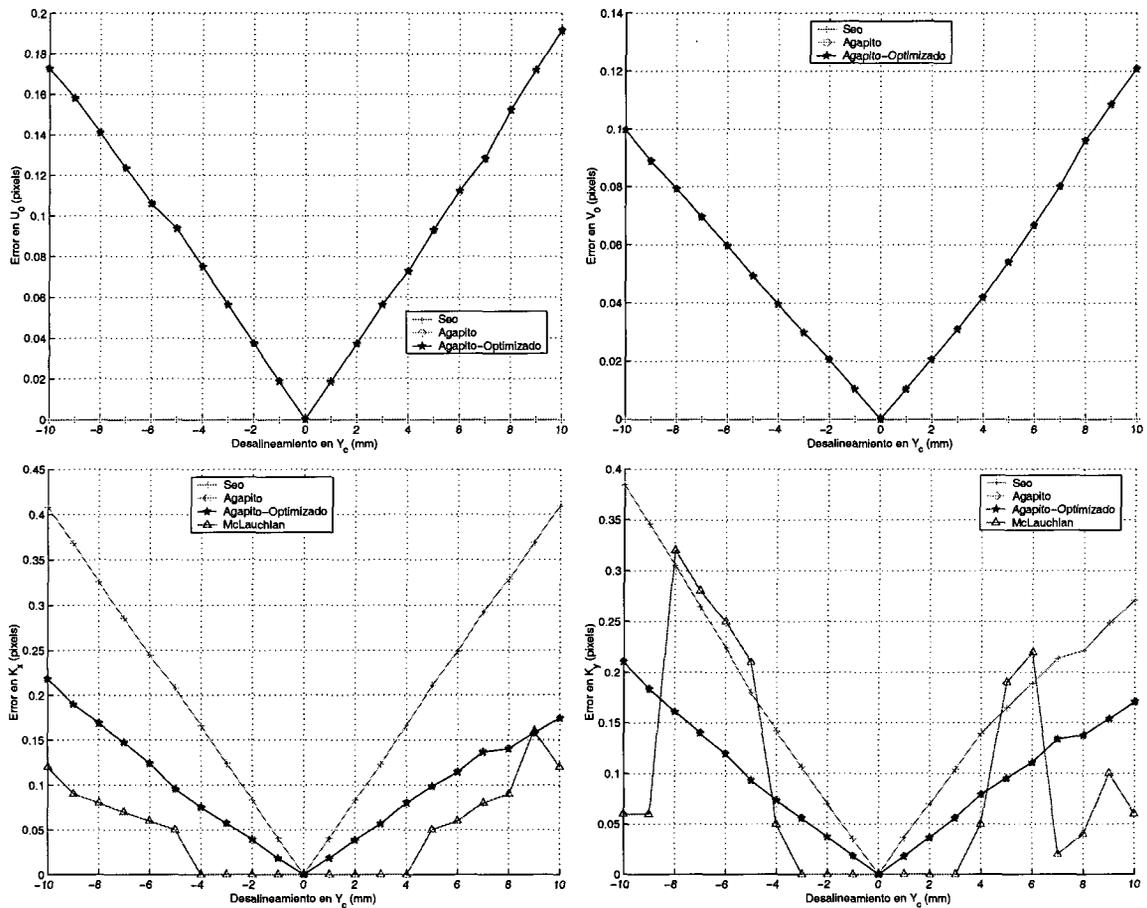


Figura 4.26: Error de cada parámetro intrínseco cuando varía la distancia vertical (Y) entre el foco y el eje de giro de la cámara. Las imágenes están libres de ruidos.

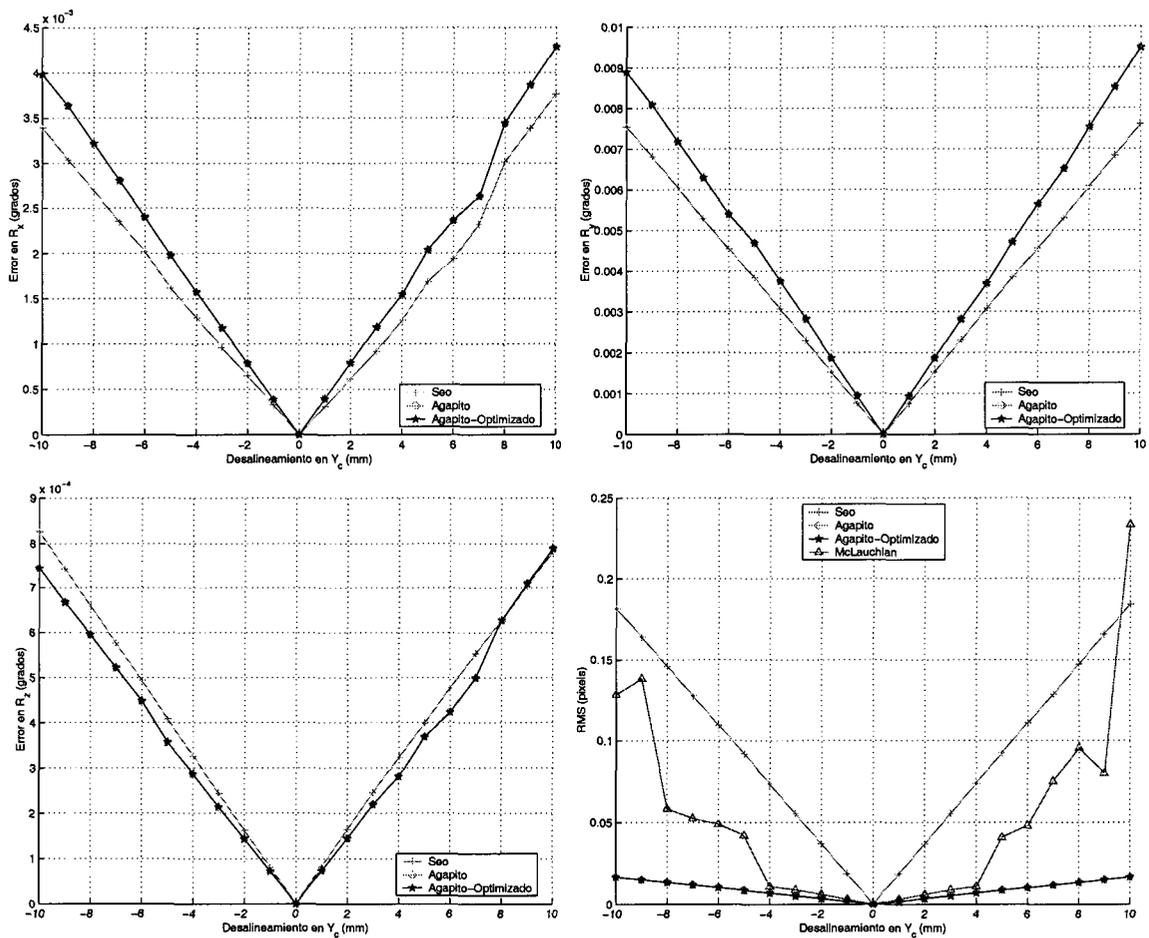


Figura 4.27: Error en la orientación y RMS sobre las coordenadas 2D cuando varía la distancia vertical (Y) entre el foco y el eje de giro de la cámara. Las imágenes están libres de ruidos.

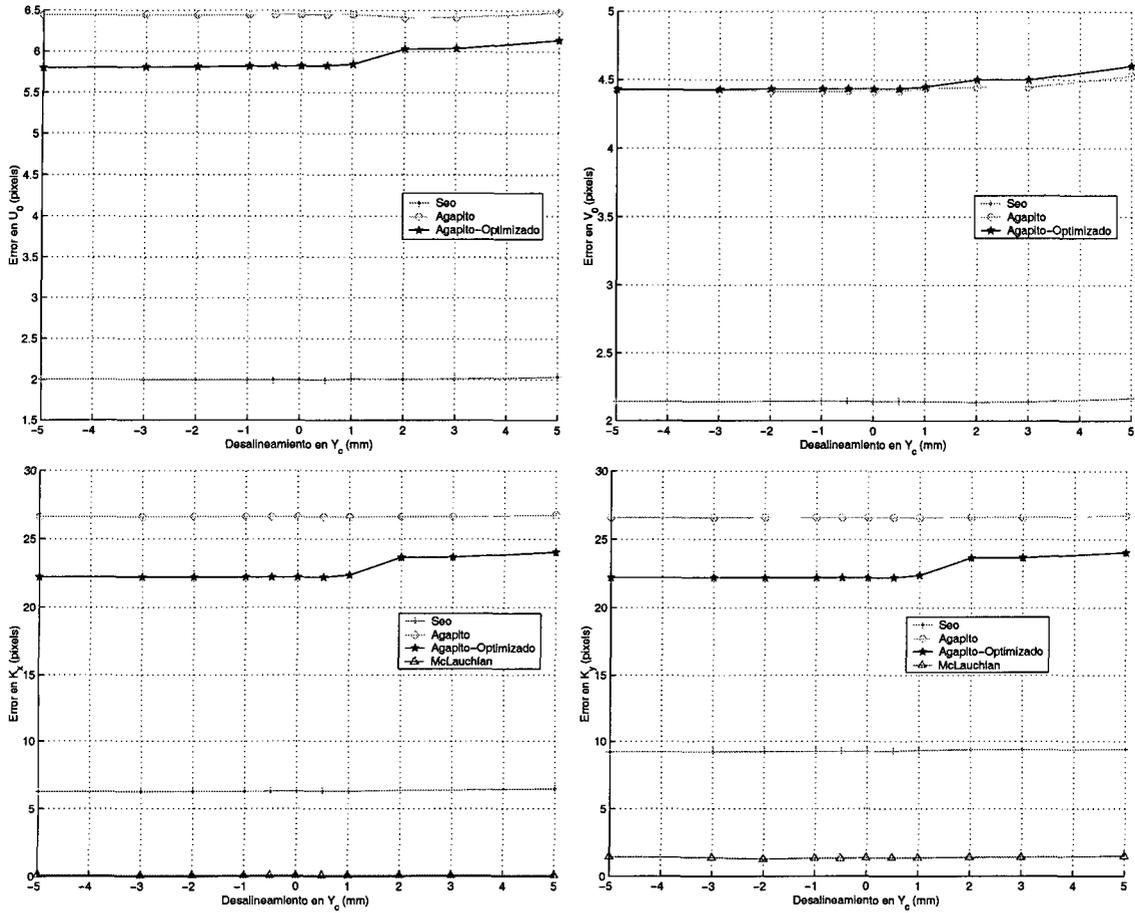


Figura 4.28: Error de cada parámetro intrínseco cuando varía la distancia vertical entre el foco y el eje de giro de la cámara. Las imágenes tienen un ruido gaussiano de media cero y desviación de 0.5 pixels.

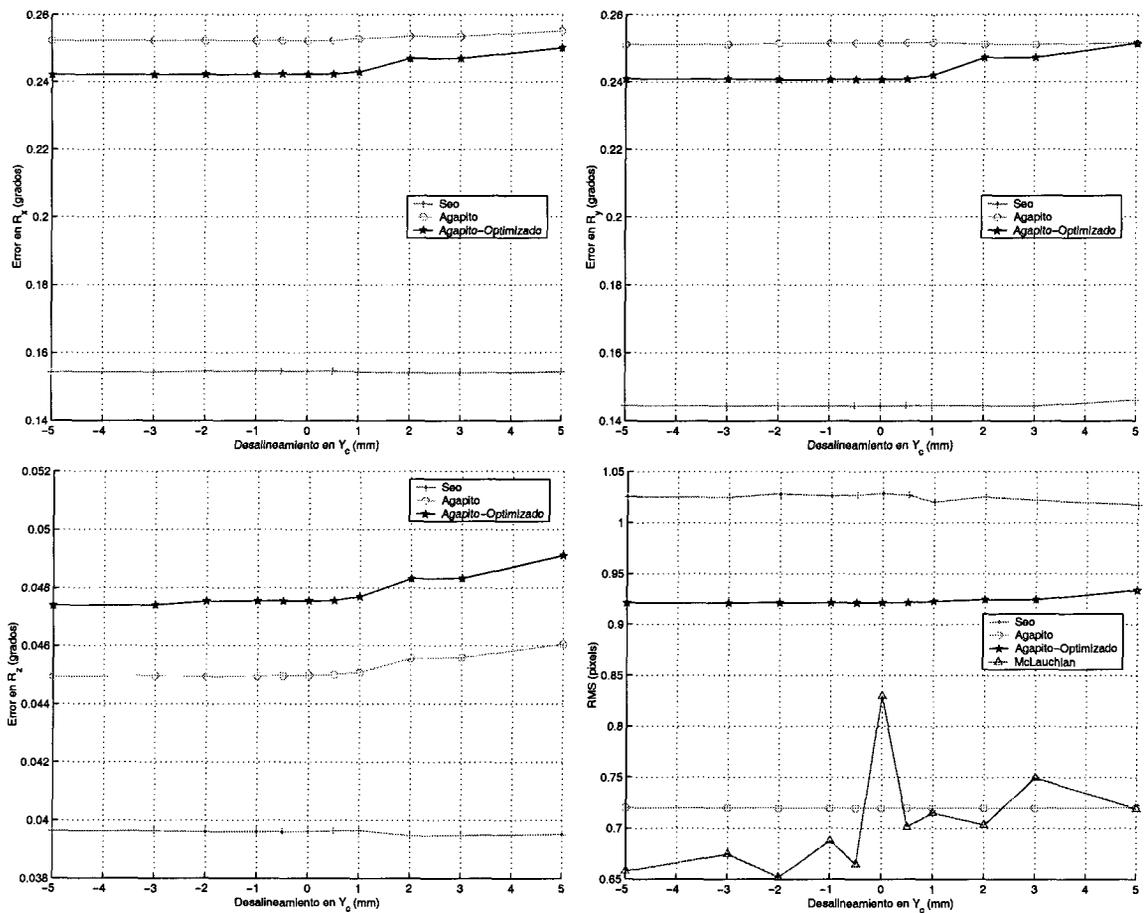


Figura 4.29: Error en la orientación y RMS sobre las coordenadas 2D cuando varía la distancia vertical entre el foco y el eje de giro de la cámara. Las imágenes tienen un ruido gaussiano de media cero y desviación de 0.5 pixels.

influencia de este factor ya que el error en la estimación de cada parámetro permanece prácticamente constante. Solamente el método optimizado de Agapito presenta pequeñas variaciones en el lado derecho de la curva en todas las gráficas.

En segundo lugar, se analiza el efecto producido por el desalineamiento en el eje Z_c (eje óptico) entre el foco y el centro de giro de la cámara. Este desalineamiento se produce, principalmente, por el desplazamiento del zoom. Los resultados obtenidos en el caso de imágenes sin ruido se muestran en la figuras 4.30 y 4.31. Como se puede observar en las gráficas la respuesta obtenida es semejante, en forma, al desalineamiento vertical, pero no así en magnitud ya que el error en la estimación de cada parámetro es mucho mayor que en el otro caso, sobre todo en los factores de escalado. Únicamente se puede destacar el hecho de que el método de Seo tiene menor error que los métodos de Agapito en casi todos los parámetros, la diferencia es especialmente llamativa en el caso de los factores de escalado.

Las gráficas de las figuras 4.32 y 4.33 muestran los resultados de este experimento en

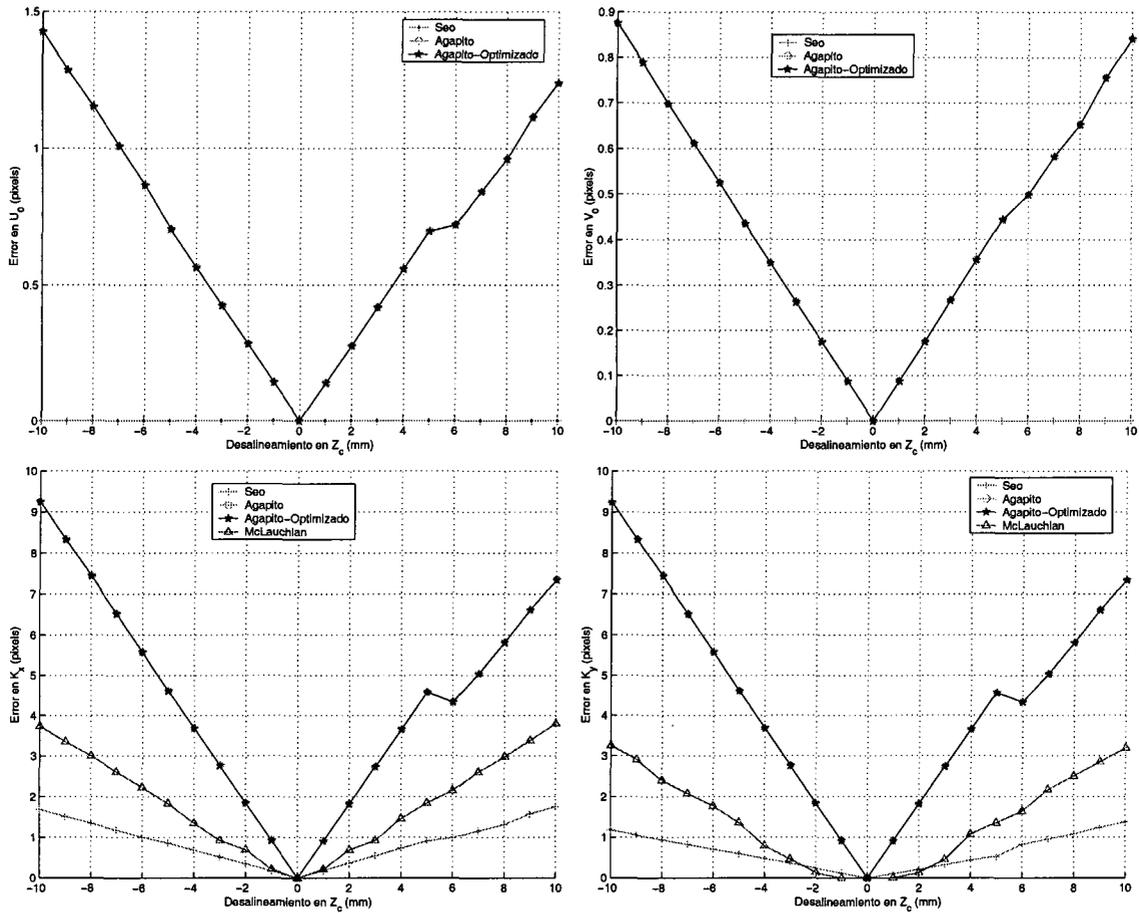


Figura 4.30: Error de cada parámetro intrínseco cuando varía la distancia a lo largo del eje óptico (Z) entre el foco y el eje de giro de la cámara. Las imágenes están libres de ruidos.

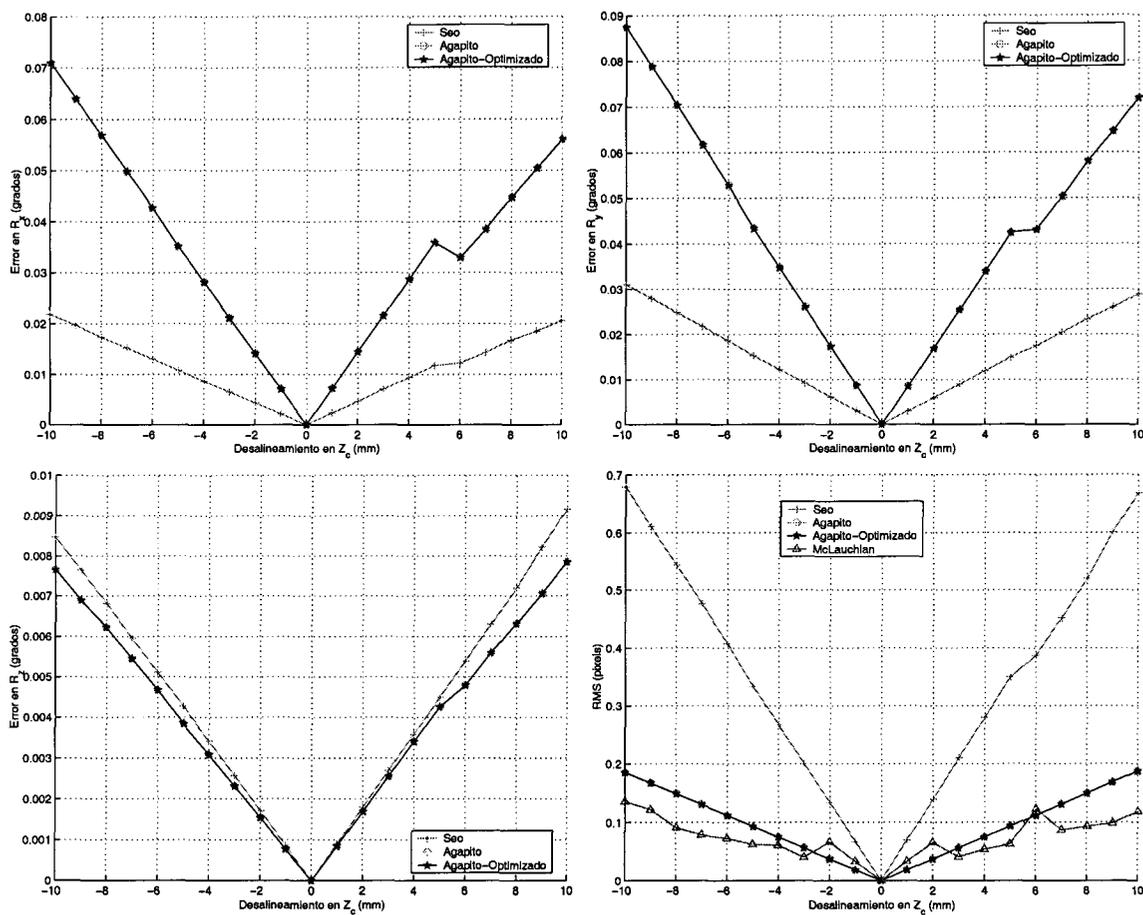


Figura 4.31: Error en la orientación y RMS sobre las coordenadas 2D cuando varía la distancia a lo largo del eje óptico (Z) entre el foco y el eje de giro de la cámara. Las imágenes están libres de ruidos.

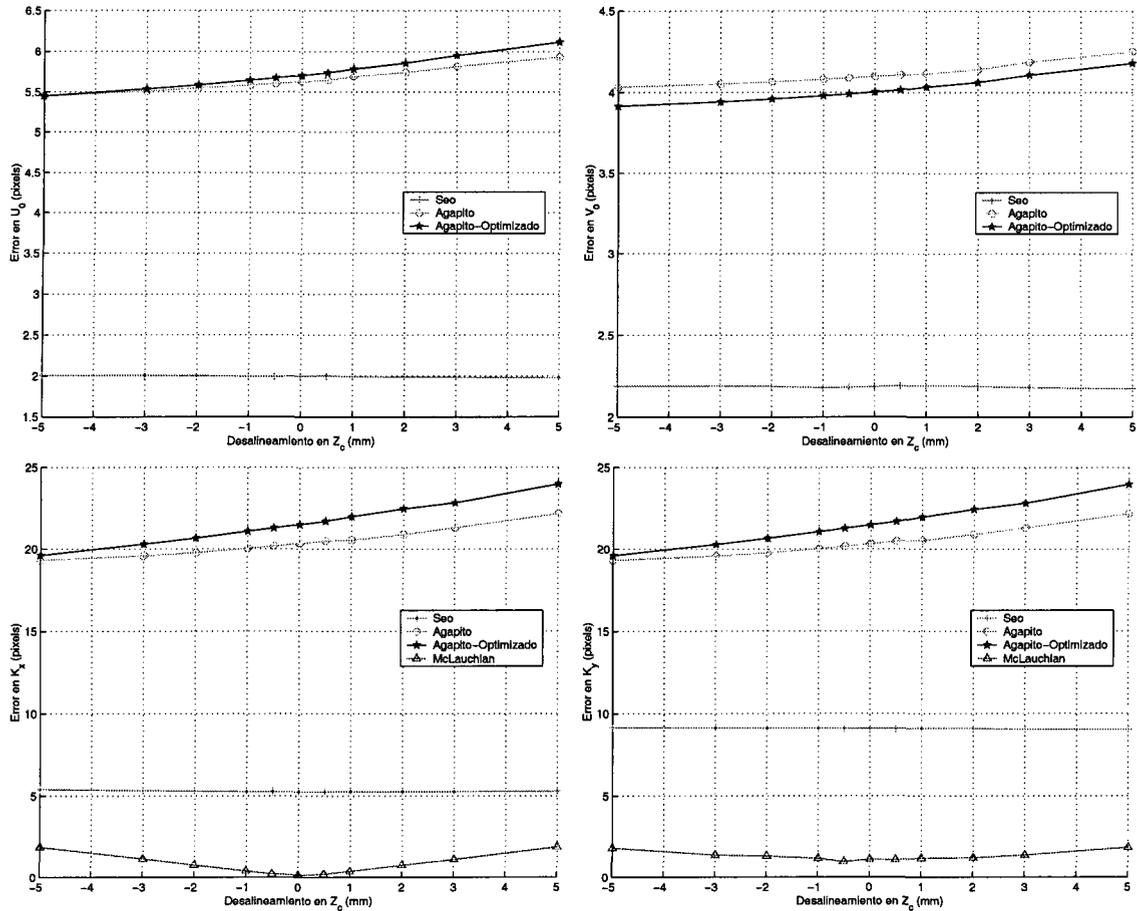


Figura 4.32: Error de cada parámetro intrínseco cuando varía la distancia en la dirección del eje óptico (Z) entre el foco y el eje de giro de la cámara. Las imágenes tienen un ruido gaussiano de media cero y desviación de 0.5 pixels.

imágenes con ruido gaussiano de media nula y desviación de 0.5 pixels. En estas gráficas se observa, al igual que en el caso anterior, que la incidencia del ruido es mucho mayor que la del desalineamiento ya que en la mayor parte de los parámetros el error apenas sufre variaciones. En el caso de los métodos de Agapito aparece un pequeño incremento del valor de cada parámetro y en el caso del método de McLauchlan existe un mínimo en el error en la estimación de los factores de escalado cuando el desalineamiento es nulo, sufriendo un pequeño incremento a medida que la distancia entre el eje de giro de la cámara y la posición del foco aumenta.

Según [Hayman and Murray, 2002] y [Wang et al., 2001] el error producido por este factor es inversamente proporcional a la distancia de los puntos de la escena. Esta conclusión, unida al hecho de que en este caso los puntos se encuentran bastante alejados, alrededor de dos metros y medio, explicaría el que la influencia del desalineamiento sea mínima.

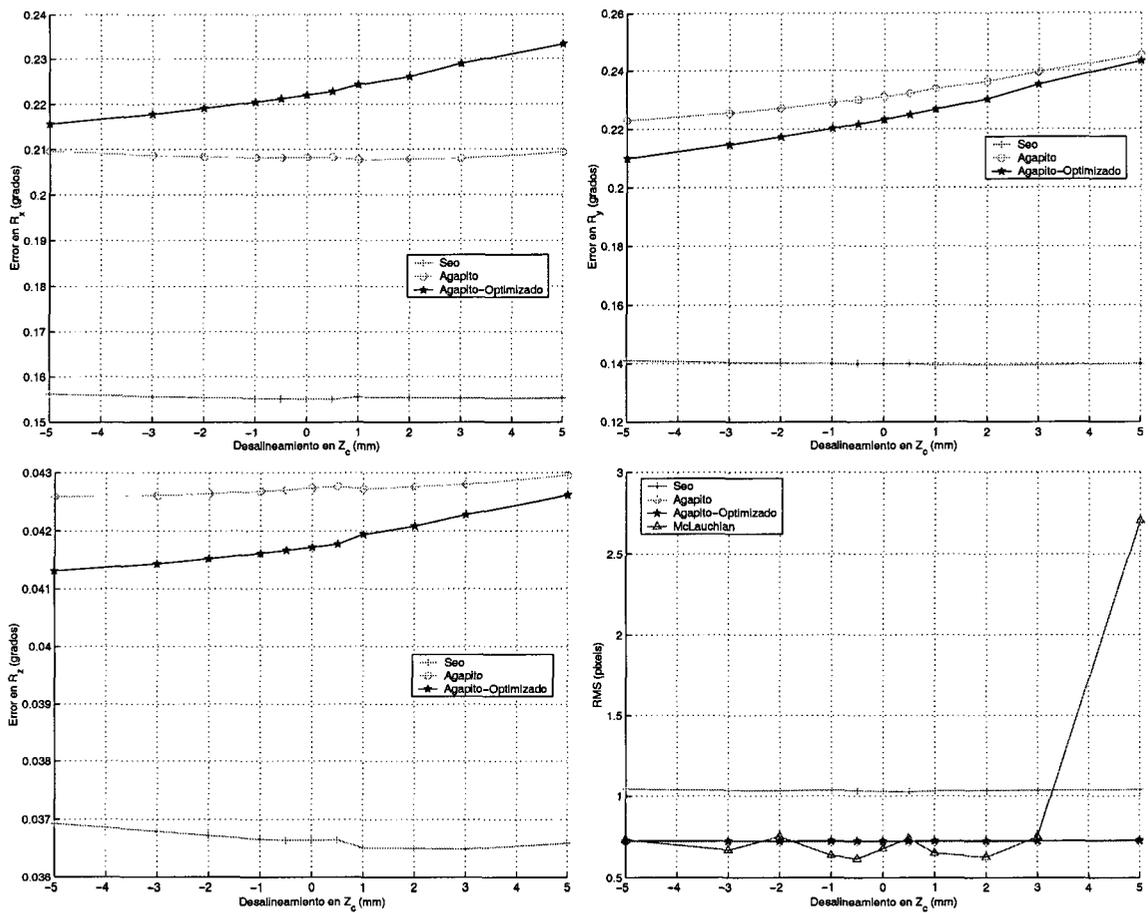


Figura 4.33: Error en la orientación y RMS sobre las coordenadas 2D cuando varía la distancia en la dirección del eje óptico (Z) entre el foco y el eje de giro de la cámara. Las imágenes tienen un ruido gaussiano de media cero y desviación de 0.5 pixels.

4.4.6. Conclusiones

Se han realizado diversos experimentos con el objetivo de ver la precisión, robustez y coste computacional de cada uno de los métodos descritos.

Los resultados obtenidos muestran que, como era de esperar, el método McLauchlan funciona muy bien siempre y cuando los parámetros intrínsecos se mantengan constantes. También falla, aunque en menor medida que el resto, cuando las imágenes presentan un alto grado de distorsión, o cuando elementos como la asimetría o el factor de proporción distan de sus valores idóneos, es decir, $\Omega = 90^\circ$, y $s_x = 1$. Por contra, es un método que presenta un alto coste computacional, ya que necesita muchas imágenes para que el método converja y no puede procesar elevadas cantidades puntos por vista. Por otra parte, un número elevado de puntos en un entorno "natural" no parece una condición realista.

El método de Seo presenta mejores resultados que el método de Agapito en la estimación del centro óptico y la orientación relativa de la cámara en cada una de las vistas. En cambio, en la estimación de los factores de escalado, suele presentar peores resultados. Esto se debe al algoritmo empleado, el cual estima mediante un método lineal los factores de escalado y luego realiza un ajuste global de todos los parámetros mediante un método iterativo. Puede tener problemas si el centro óptico dista del centro de la imagen. En cuanto al coste computacional, hay que indicar que este método no presenta tiempos muy elevados mientras se utilicen pocas imágenes, en cambio, si el número de vistas procesadas es elevado se torna excesivamente lento. Es un método adecuado si el número de puntos detectados por imagen es elevado ya que permite utilizar pocas vistas en el proceso de calibración y, por tanto, reducir el tiempo de ejecución.

Los métodos de Agapito distribuyen el error entre todos los parámetros de forma más homogénea que el método de Seo. También se observa que las dos versiones del método de Agapito presentan resultados similares entre sí en la mayoría de los casos. El método optimizado sólo parece mejorar ligeramente los resultados del método lineal al estimar k_x y k_y en situaciones donde el ruido en las coordenadas 2D es superior a 0.5 pixels, por lo que se deduce que es más rentable el uso del método lineal ya que presenta un menor coste computacional. Estos métodos, y en especial la versión lineal, presentan los tiempos de ejecución más bajos.

Como se puede observar, no existe un método que funcione mejor que el resto. Todo depende de los factores que intervengan en el proceso de calibración. Además, también se observa que existen parámetros que son estimados de mejor forma por determinados métodos.

En el simulador se ha analizado, por un lado, la incidencia de factores relativos a la

formación de las imágenes, como puede ser el producido por efecto de la cuantificación de las imágenes. También se han analizado aspectos relativos a las limitaciones que presentan los modelos de cámara utilizados relacionados con la distorsión de las imágenes, el grado de ortogonalidad de los ejes del plano imagen y el factor de proporción de los pixels. Se ha observado que de entre estos factores los que tienen mayor incidencia en los resultados son la distorsión y el ruido en las coordenadas 2D. El resto de los factores no presentan, en las condiciones utilizadas, una gran incidencia sobre los resultados.

Por otro lado, se observa que, relativamente, existen mayores niveles de error en los parámetros intrínsecos que en los extrínsecos. También se puede observar que el error RMS obtenido por cada método no está en relación con el error cometido en cada uno de los parámetros. Por lo que, al igual que en el caso de los métodos clásicos, parece que se produce una compensación de errores entre parámetros.

Capítulo 5

Métodos de autocalibración: experimentos con datos reales

En este capítulo se analiza el funcionamiento de los métodos autocalibración estudiados en el capítulo 4 cuando se aplican sobre una cámara real. Los siguientes experimentos se han planteado con el objetivo de caracterizar la robustez de estos métodos de autocalibración en un entorno real. Para ello, se analiza la estabilidad de los parámetros intrínsecos frente a variaciones de los parámetros extrínsecos. Además, también se comparan los resultados obtenidos por los métodos de autocalibración con los obtenidos por los métodos de calibración clásicos.

De los experimentos realizados en el simulador sólo se han podido trasladar dos de ellos a la calibración con datos reales. En concreto, se ha variado el número de imágenes en la secuencia utilizada para la calibración. En primer lugar manteniendo los parámetros intrínsecos fijos, lo cual permite analizar la estabilidad de los métodos, ya que el valor devuelto de la calibración para cada una de las vistas en estos parámetros debe ser siempre el mismo. Se supone que con la misma configuración de cámara los parámetros intrínsecos deben ser iguales en todos los casos, aunque cambien los parámetros extrínsecos. En segundo lugar, se repite el experimento variando la posición del zoom en cada una de las vistas adquiridas. Esto permite observar en qué parámetros incide este elemento y corroborar los resultados obtenidos con el simulador. En otro experimento se varía el número de puntos del patrón para estudiar cuáles son las condiciones del entorno idóneas en función del método e igualmente evaluar su coste computacional.

Para la realización de estos experimentos se han utilizado imágenes obtenidas con una cámara Sony Evi-G21. El tamaño del CCD según las indicaciones del fabricante es de 3.2 mm x 2.4 mm y el tamaño de las imágenes es de 768 x 576 pixels. Estas cámaras permiten realizar zoom, pudiendo variar la distancia focal entre 4.5 y 13.5 mm. Dispone

también de enfoque y apertura automáticos. Además cuenta con motores que permiten movimientos de azimut y elevación (hasta $\pm 30^\circ$ en azimut y $\pm 15^\circ$ en elevación).

En las pruebas se ha realizado el mismo movimiento descrito en el capítulo de experimentos de autocalibración en el simulador (sección 4.2), es decir, un movimiento circular realizado con el azimut y la elevación de la cámara.

Se ha utilizado un patrón para obtener los puntos de referencia que permiten realizar el proceso de calibración. Debe notarse que aunque se emplee un patrón, en estos experimentos no se utiliza ninguna información relativa a las coordenadas 3D de los puntos o sobre su posición relativa. Los motivos de utilizar un patrón de puntos son dos. Por un lado, se evita la aparición de falsos positivos, es decir, que en los resultados incidan errores motivados por no realizar la correspondencia de puntos entre cada una de las imágenes de forma adecuada. Por otro lado, permite medir el error en la reconstrucción 3D de dichos puntos, lo cuál será un indicativo de la precisión de cada método. Como patrón se ha utilizado un diedro de 7x6 puntos en cada plano, lo que significa un total de 84 puntos, entrando todos ellos dentro del campo de visión de la cámara en todas las imágenes. Los puntos tienen un diámetro de 1 cm y se encuentran separados entre sí 2 cm.

5.1. Influencia de la longitud de la secuencia

En este primer experimento se trata de determinar el número de imágenes necesario para que el resultado de los métodos converja. Al igual que en el simulador, se ha capturado una secuencia de 50 imágenes. En el experimento se ha variado el número de imágenes procesadas de esta secuencia.

Se han adquirido las secuencias de imágenes manteniendo el zoom y el enfoque fijos. Esto permite analizar el número de imágenes necesarias para que el método converja a un resultado estable. En el simulador se observó que los métodos convergen cuando la longitud de la secuencia se sitúa entre 15 y 20 imágenes.

Los resultados de realizar el proceso de calibración colocando el zoom en 1000 pasos de motor, siendo el máximo de 16384 pasos, se muestran en la figura 5.1. Esta posición del zoom equivale a una distancia focal bastante corta (5 mm) y donde el efecto de la posible distorsión es mayor, tal y como se ha mostrado en los experimentos realizados con los métodos clásicos. El patrón se ha situado aproximadamente a un metro de la cámara. El semiángulo del cono descrito por el eje óptico de la cámara en su movimiento es de 10° .

Estos resultados muestran que los métodos de Agapito, los cuales se superponen en las figuras, no logran converger en el caso del centro óptico con el número de imágenes

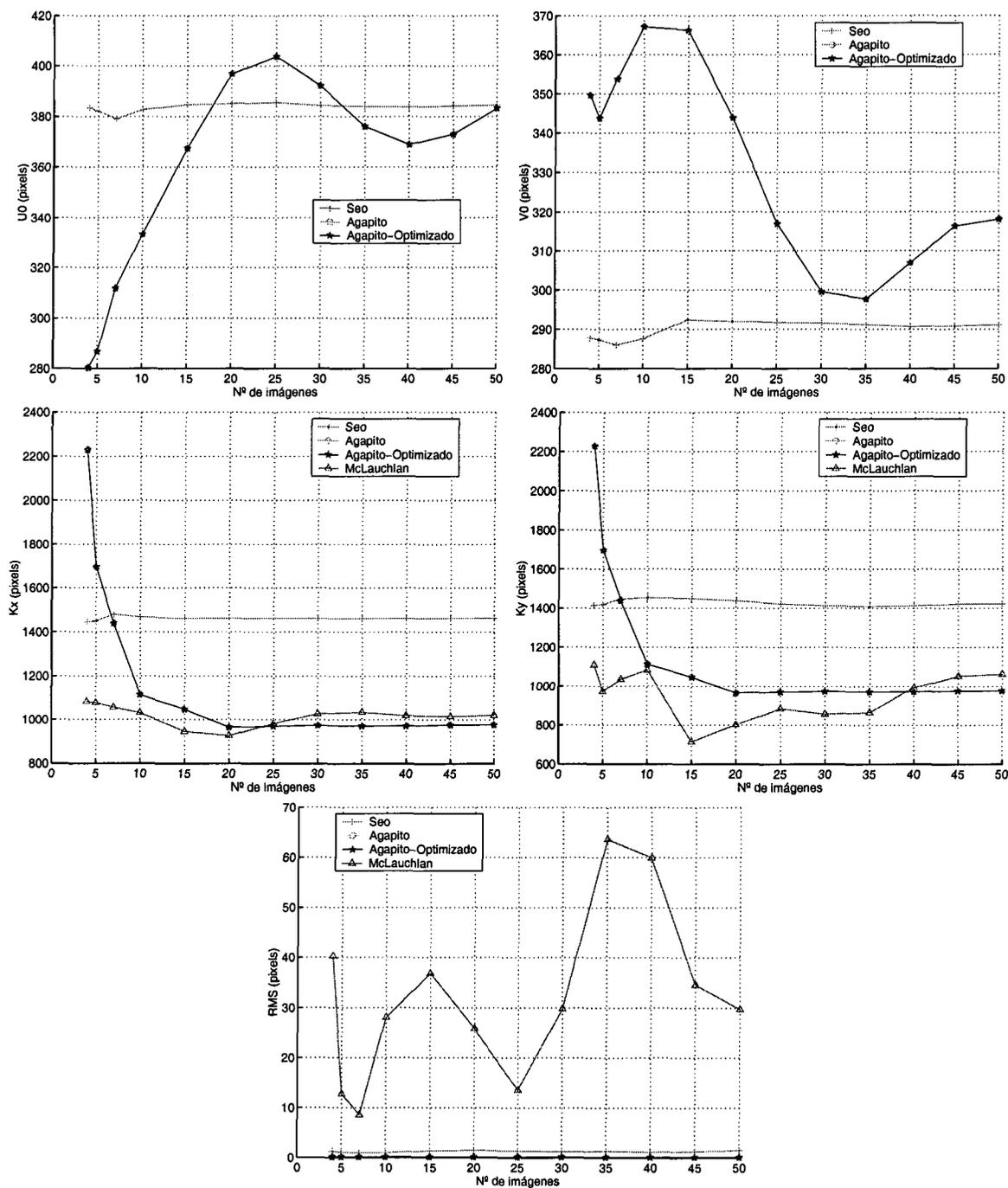


Figura 5.1: Valores obtenidos de los parámetros al variar el número de vistas procesadas. El motor del zoom tiene 1000 pasos (max. 16384), equivalentes aproximadamente a una focal de 5 mm.

procesadas. En cambio los factores de escalado convergen con secuencias de más de 20 imágenes. El método de McLauchlan también converge a la hora de estimar los factores de escalado, pero para hacerlo necesita en torno a 30 vistas en k_x y 45 en k_y . Por contra, el método de Seo presenta un resultado homogéneo a la hora de estimar los parámetros intrínsecos. Llama la atención la disparidad en las estimaciones realizados por los distintos métodos, en especial, en los factores de escalado, donde el método de Seo obtiene un resultado que dista bastante del resto. Además, el valor obtenido para los factores de escalado por los métodos clásicos (figuras (3.3 y 3.7) para la distancia focal utilizada (zoom igual a 1000 pasos de motor) se sitúa entre 1000 y 1200 pixels. Se puede observar que estos valores son más próximos a los obtenidos por los métodos de Agapito y de McLauchlan. La causa de el resultado tan diferente que obtiene el método de Seo está en el proceso de obtención de los parámetros que tiene este método; es decir, al obtener en una primera fase los factores de escalado estos absorben la mayor parte del error, mientras que en el resto de los métodos el error se distribuye de forma más homogénea entre todos los parámetros. Este hecho ya se ha mostró en el capítulo anterior en los experimentos realizados con el simulador.

El experimento se ha repetido pero utilizando una distancia focal mayor (aproximadamente 9.4 mm), es decir, se ha colocado el zoom en 9000 pasos motor, siendo el máximo de 16384 pasos. Al utilizar una distancia focal intermedia los resultados no se ven tan afectados por la distorsión. Como se ha aumentado la distancia focal ha sido necesario alejar el patrón, situándolo cerca de los dos metros y medio, y el semiángulo del cono descrito por el eje óptico de la cámara en su movimiento se ha reducido a 6° para que el total de puntos del patrón permanezca visible en todas las imágenes adquiridas. En la figura 5.2 se muestran los resultados al realizar el proceso de calibración.

En las gráficas se observa que en los métodos de Agapito la estimación obtenida cuando el número de imágenes es muy pequeño dista mucho del valor de convergencia. El método de McLauchlan necesita mayor número de imágenes que el resto para que el valor de los factores de escalado se estabilice, obteniéndose un valor bastante inferior al estimado por los otros métodos. En este método se observa que el error RMS es excesivamente alto cuando el número de imágenes procesadas es menor que 5. El método de Seo sigue mostrando un comportamiento semejante al caso anterior. En este caso, los métodos de Agapito y Seo muestran resultados mucho más próximos entre sí que en el caso anterior. La diferencia en los resultados de estos dos métodos ronda los 10 pixels en la estimación del centro óptico, en k_x es de alrededor de 150 pixels mientras que en k_y es de 20 pixels.

Por último, se observa que inestabilidad de los resultados cuando el número de imágenes es pequeño es mayor en la segunda muestra (figura 5.2) que en la primera muestra

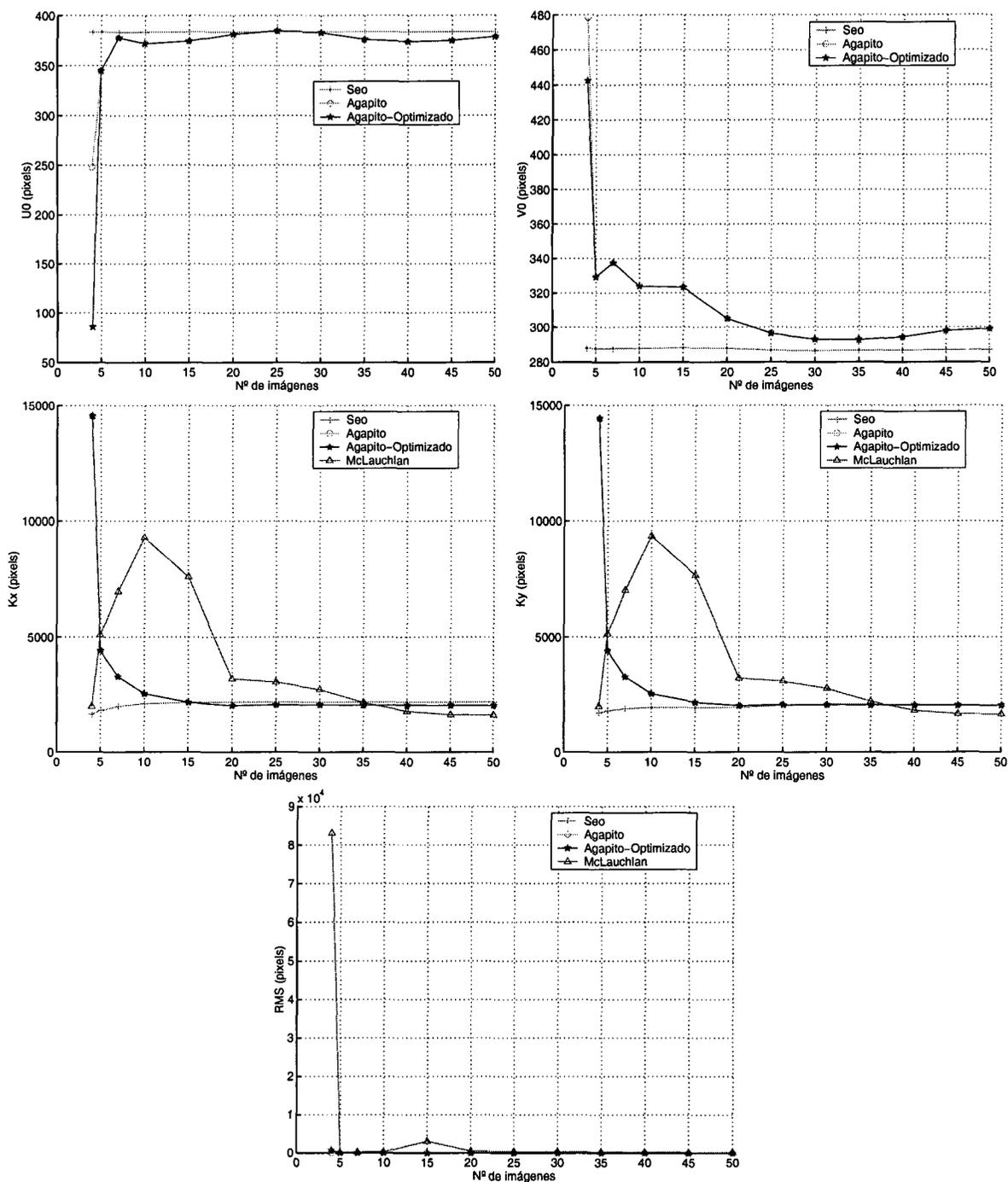


Figura 5.2: Valores obtenidos de los parámetros al variar el número de vistas procesadas. El motor del zoom tiene 9000 pasos (max. 16386), equivalentes aproximadamente a una focal de 9.4 mm.

(figura 5.1). Esto se debe a que al estar los puntos más alejados la precisión con que se obtienen los centroides de los mismos es menor, siendo, por tanto, el ruido 2D mayor.

5.2. Estabilidad de los resultados

En este apartado se analiza la estabilidad de estos métodos, es decir, si a la hora de recuperar los parámetros intrínsecos de la cámara, éstos se mantienen constantes a lo largo de la secuencia, en la cual la orientación de la cámara ha ido variando.

En la figuras 5.3 y 5.4 se representan los resultados obtenidos por cada método al calibrar una secuencia de 40 imágenes con un zoom de 9000, siendo por tanto, la distancia focal de aproximadamente 9.4 mm. El conjunto de puntos se encuentra a una distancia próxima a los tres metros. En las gráficas se muestra la estimación obtenida de cada parámetro en cada una de las 40 imágenes. El semiángulo del cono descrito por el eje óptico de la cámara en su movimiento es de 6° . Hay que decir, que las curvas del método de Agapito lineal no se observan en las gráficas debido a que coinciden plenamente con las de la versión optimizada.

Los resultados obtenidos muestran que, aunque los parámetros intrínsecos han permanecido constantes a lo largo de la secuencia, la estimación del centro del eje óptico presenta una alta inestabilidad para todos los métodos. Los métodos de Agapito se muestran mucho más inestables que el método de Seo, en especial en u_0 . El rango de variación observado llega hasta los 80 pixels en el caso de u_0 y de 110 pixels en el caso de v_0 . En cambio, los resultados del factor de escalado son mucho más estables, tanto en un método como en otro. En ninguno de los casos se alcanza los 20 pixels de rango, lo cual supone menos de 1 % del valor del parámetro.

En el caso del método de McLauchlan obtiene un valor único de los factores de escalado para toda la secuencia, por lo que su resultado es una recta perfecta. El resto de los métodos presentan una dispersión muy pequeña de los resultados. Llama la atención la gran diferencia de resultados obtenidos entre el método de McLauchlan y el resto, esto se ve reflejado, además, en la gráfica que muestra el error RMS donde aparece claramente como el peor método. Esto se debe a que la calibración a partir de la secuencia utilizada necesite mucho más de 40 imágenes para converger a un resultado aceptable. Entre las razones que impiden esta convergencia están: la precisión de las aproximaciones iniciales empleadas, la distribución de los puntos en la escena y la precisión de las varianzas utilizadas para cada parámetro (30 pixels para los factores de escalado y 0.1 radian para el desalineamiento en la orientación).

En cuanto a la estimación de la orientación relativa de la cámara se observa como las curvas de R_x y R_y describen el movimiento circular realizado. La dispersión en los

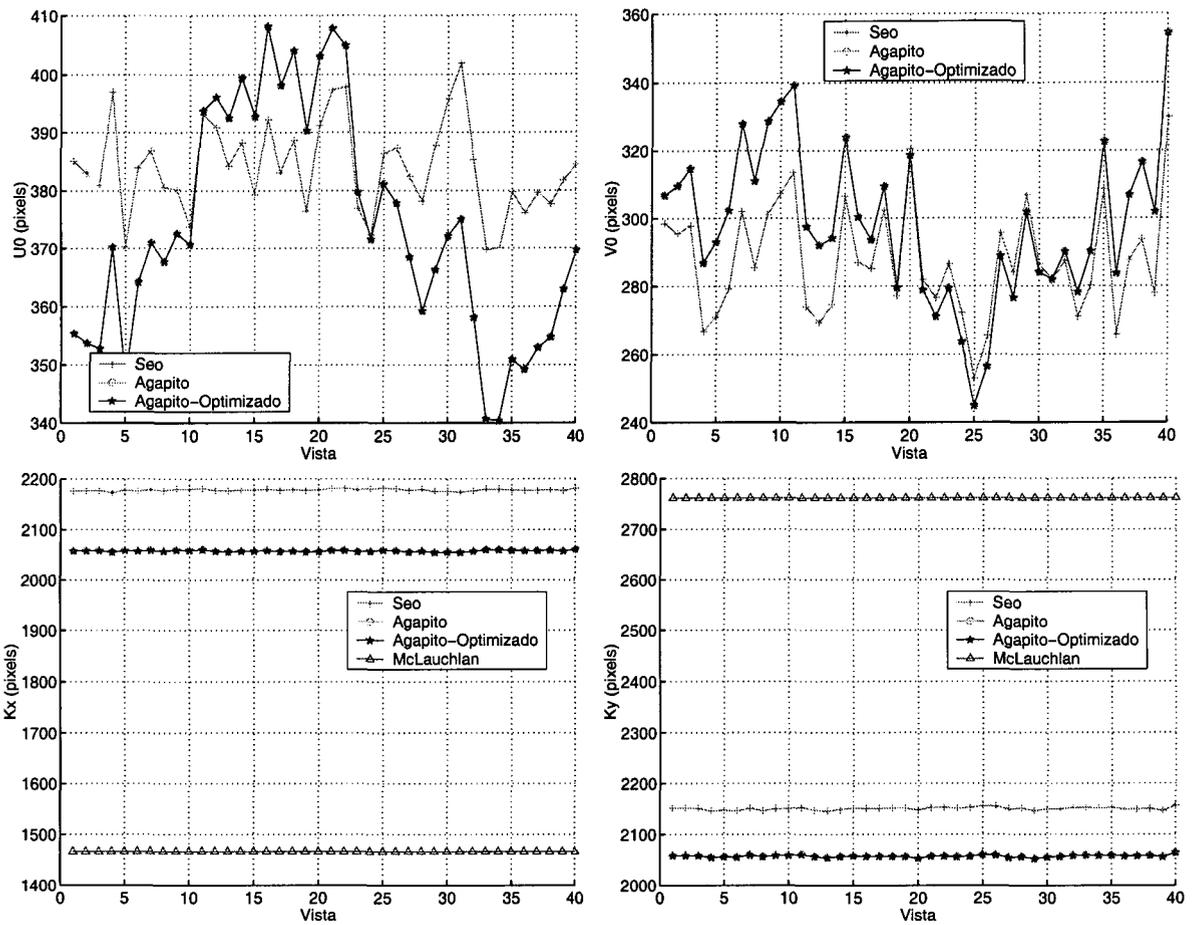


Figura 5.3: Valores obtenidos de los parámetros intrínsecos al calibrar una secuencia de 40 imágenes.

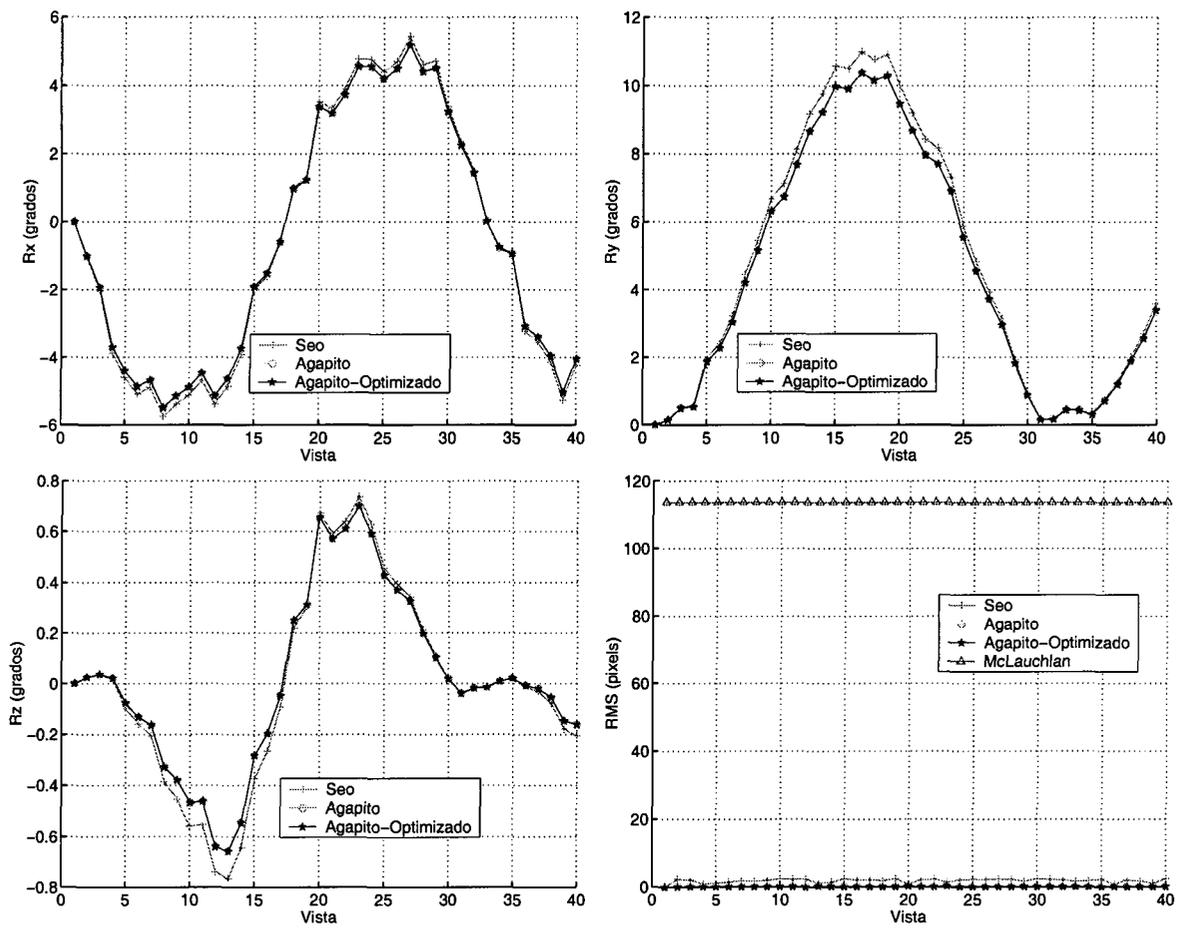


Figura 5.4: Valores obtenidos de la orientación y el error RMS al calibrar una secuencia de 40 imágenes.

resultados no parece muy alta. Llama la atención el hecho de que la estimación de R_z varíe más de un grado ya que se supone que no existe ciclotorsión. Esto sólo se puede explicar si los ejes de giro de la cámara no tienen la misma orientación que los ejes del plano imagen; es decir, que el movimiento de azimut no coincida con R_y y el de elevación con R_x .

5.3. Influencia del zoom

En este apartado se muestra la incidencia del zoom en los resultados de los métodos. En las figuras 5.5 y 5.6 se muestran los resultados de haber calibrado la cámara con una secuencia de 40 imágenes. En este caso, se ha incrementado linealmente el número de pasos del motor del zoom; es decir, en cada una de las vistas la distancia focal ha aumentado respecto a la anterior de forma lineal. El semiángulo del cono descrito por el eje óptico de la cámara en su movimiento es de 4° y como en los casos anteriores, todos los puntos del patrón son visibles en todas las imágenes adquiridas. Para observar la influencia del zoom, los resultados son comparados con los del apartado anterior.

En la determinación del centro óptico se observa que la inestabilidad es aún mayor que en el caso anterior. En los factores de escalado el método de McLauchlan obtiene, como es lógico, un valor constante, ya que este método sólo obtiene un valor de estos parámetros para toda la secuencia. Llama la atención que este método obtenga un valor próximo a cero, esto puede deberse a que el valor de la distancia focal no se ha mantenido constante durante la secuencia y es la única forma de hacer converger al filtro que utiliza este método. En el resto de los métodos se produce un incremento lineal del valor sin que se aprecie distorsión en los mismos, aunque es destacable la gran diferencia de resultados entre unos y otros.

En los ángulos de orientación relativa de la cámara se observa una mayor inestabilidad en los resultados respecto al experimento anterior. En cuanto a R_z también se observa que la amplitud de la curva obtenida es menor de medio grado, la mitad que en el caso anterior, siendo esto lógico ya que el radio del círculo descrito por la cámara en el movimiento realizado es más pequeño. En general, se observa que el método de Seo ha estimado un poco mejor la orientación ya que la amplitud de sus curvas está más próxima al radio utilizado en el movimiento de la cámara.

Por último, en la gráfica del error RMS se observa que el error del método de McLauchlan es muy alto, al igual que ocurre en el experimento anterior. Es llamativo el hecho de que en el método de Seo aumente ligeramente este error a medida que distancia focal va aumentando. Esto se produce porque el efecto de la distorsión se va haciendo más notorio a medida que el patrón se va acercando a los bordes de la imagen

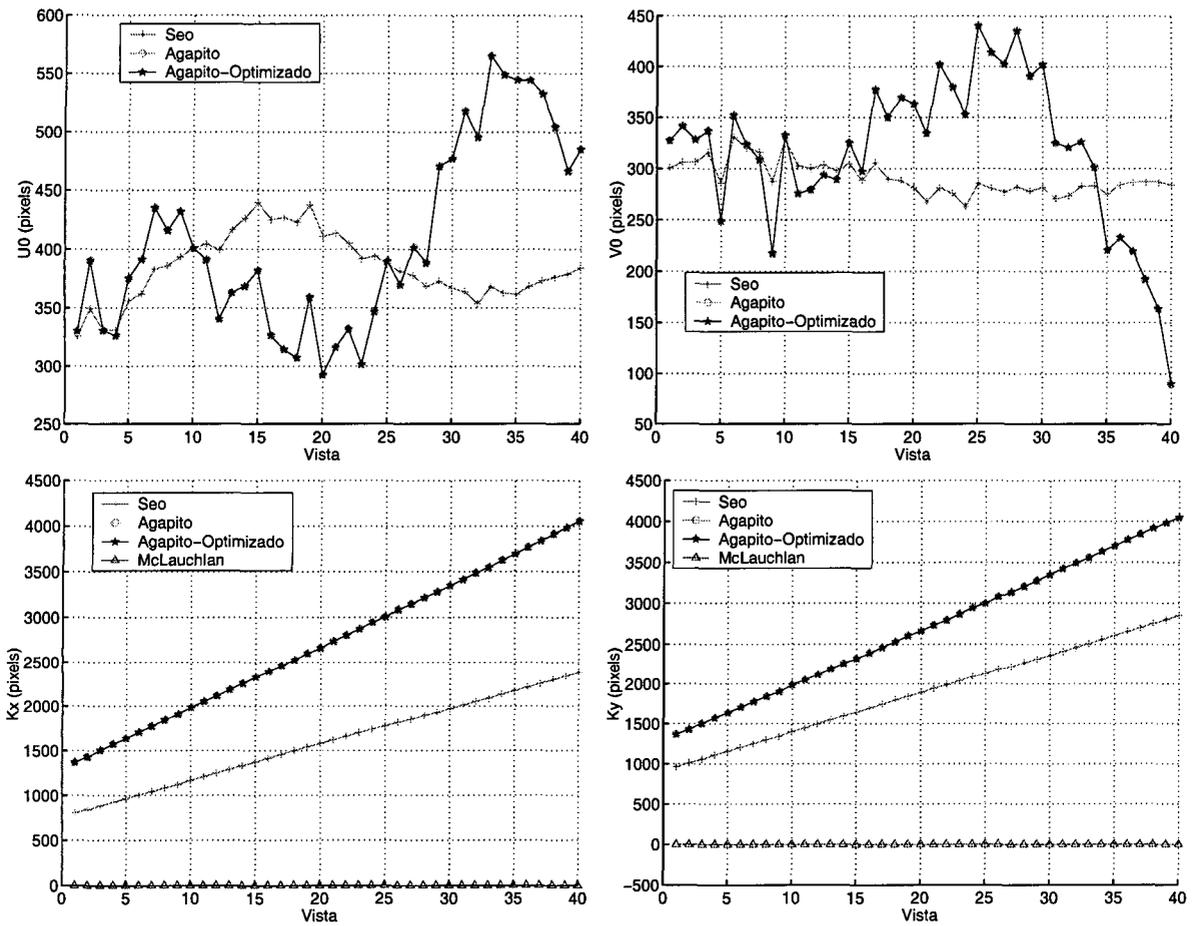


Figura 5.5: Valores obtenidos de los parámetros intrínsecos al calibrar una secuencia de 40 imágenes. El zoom aumenta linealmente.

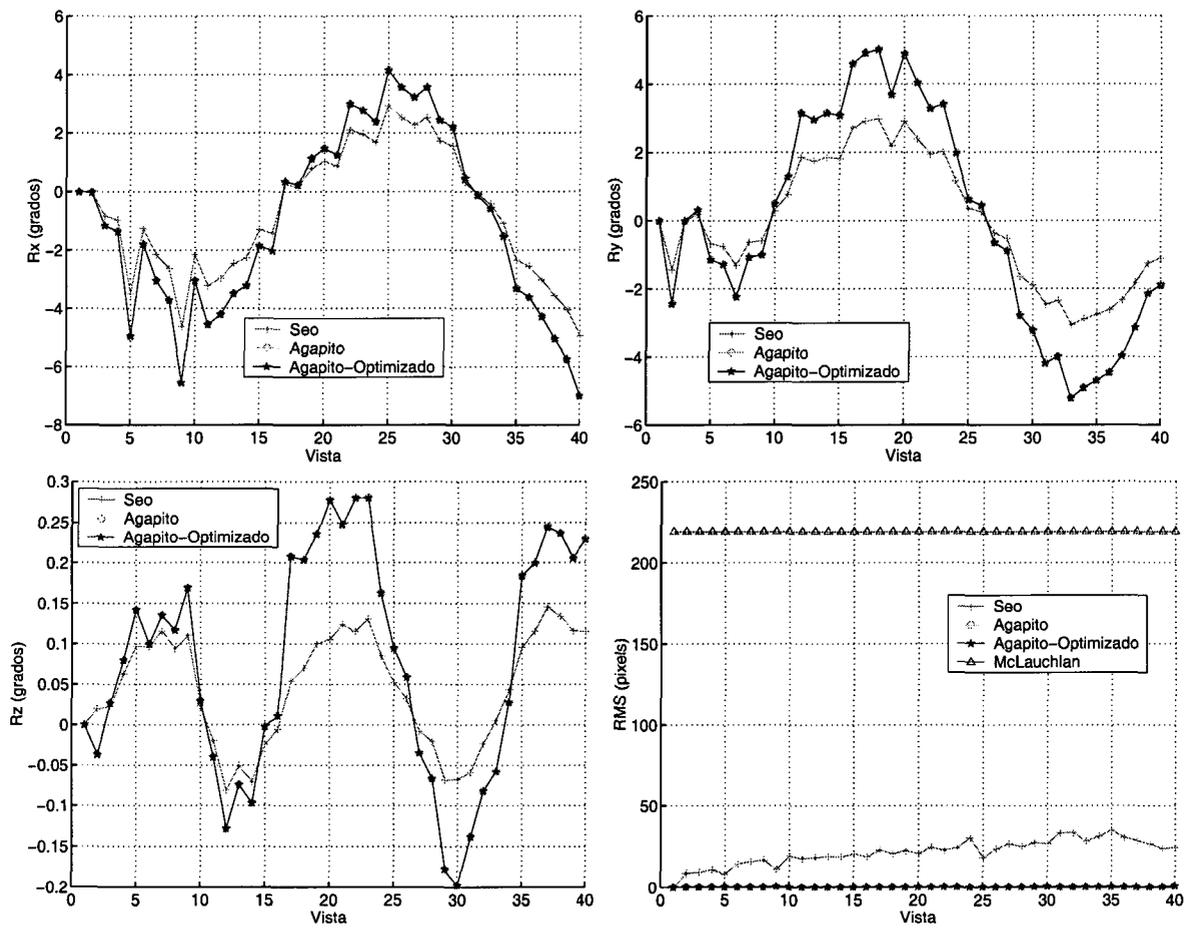


Figura 5.6: Valores obtenidos de la orientación y el error RMS al calibrar una secuencia de 40 imágenes. El zoom aumenta linealmente.

por causa del zoom. En los métodos de Agapito también se produce esta situación pero de forma más atenuado.

5.4. Reproducción de las condiciones de los experimentos reales en el simulador

El objetivo de este experimento era reproducir las condiciones de los datos reales sobre el simulador para observar si los resultados obtenidos por los métodos de autocalibración coinciden en ambos casos. Se han reproducido, aproximadamente, las condiciones con que se adquirió la segunda secuencia de imágenes empleada en la sección 5.1, donde los parámetros intrínsecos se mantienen fijos. En el experimento mostrado en este apartado se utilizó un patrón de 7x6 puntos en cada plano de un diedro. Además, se utilizó una distancia focal intermedia, siendo el número de pasos del motor del zoom de 9000.

La configuración de cámara que se ha utilizado en este experimento es la misma que se ha utilizado en los experimentos de calibración manual (sección 3.9.4). Por tanto, se han utilizado los siguientes valores: $k_1 = -0,001$, el factor de proporción próximo a la unidad (0.991) y la asimetría próxima a la perfecta ortogonalidad ($\Omega = 89,95$). A las coordenadas 2D de la imagen se les ha inyectado un ruido gaussiano de media cero y desviación típica 0.1 pixels. El centro óptico ligeramente desplazado del centro de la imagen (383,292). La distancia focal efectiva (8.69 mm) utilizada es inferior a la nominal (9.4 mm), que aproximadamente se corresponde con 9000 pasos de motor del zoom. Al igual que en el experimento con datos reales se ha utilizado un patrón con la misma configuración. El movimiento realizado tiene forma de cono al igual que en los anteriores casos, siendo el ángulo de la semiapertura de dicho cono de 6° .

El promedio de los valores obtenidos de cada parámetro intrínseco, después de repetir en cada nivel de zoom 25 veces este experimento, se muestra en la figura 5.7. En las gráficas se muestra el promedio de los valores debido a que todos los métodos, excepto el método de McLauchlan, estiman un valor por parámetro para cada una de las imágenes, aunque, como en este caso, los parámetros intrínsecos hayan permanecido constantes. Los errores cometidos en la estimación de estos parámetros se muestran en las figuras 5.8 - 5.9.

En las gráficas se observa que los métodos de Agapito distan mucho de su valor de convergencia cuando el número de imágenes es pequeño. Además, en el caso del centro óptico el resultado no llega a estabilizarse con el número de imágenes utilizado. En cambio, necesitan que la longitud de la secuencia tenga entre 20 y 25 imágenes para que los resultados de los factores de escalado converjan. La convergencia se produce a unos valo-

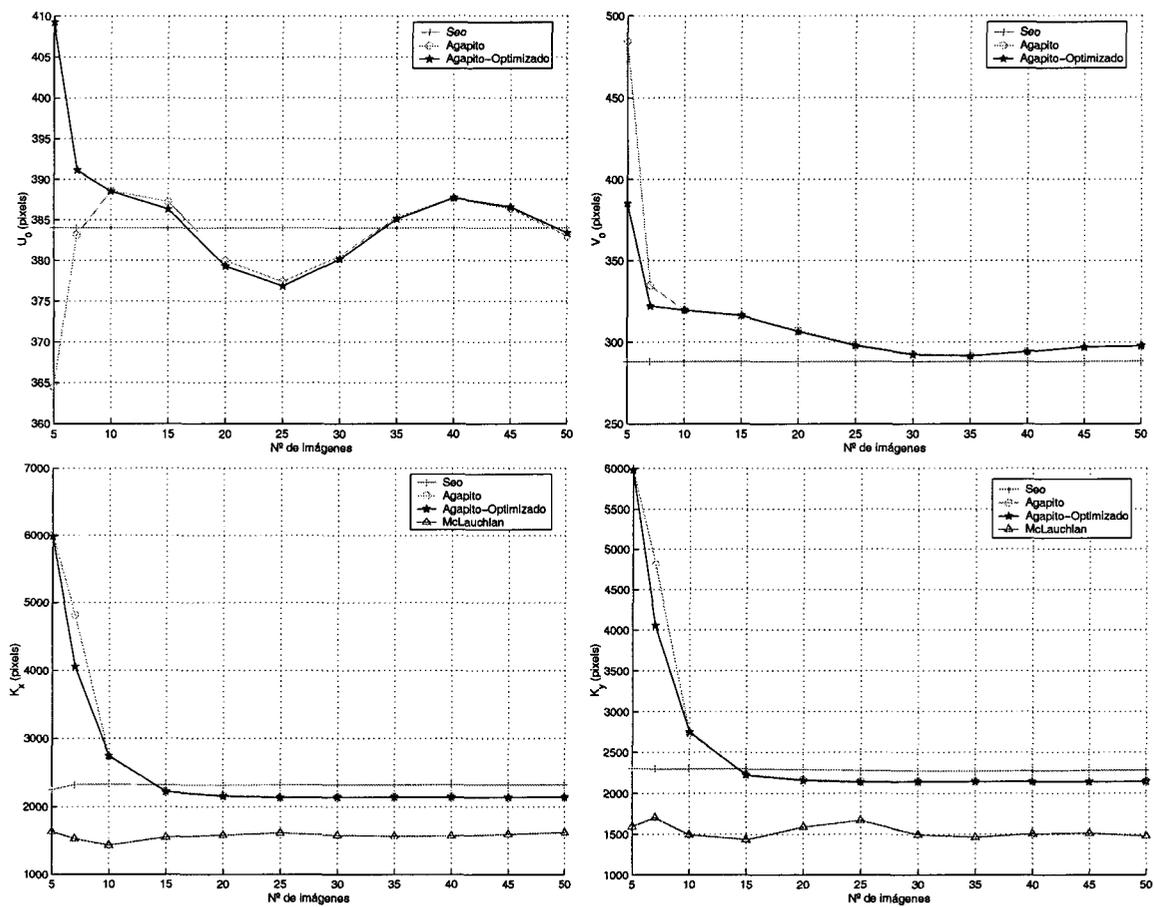


Figura 5.7: Valores promedio obtenidos de los parámetros intrínsecos al variar el número de vistas procesadas en condiciones próximas a las reales.

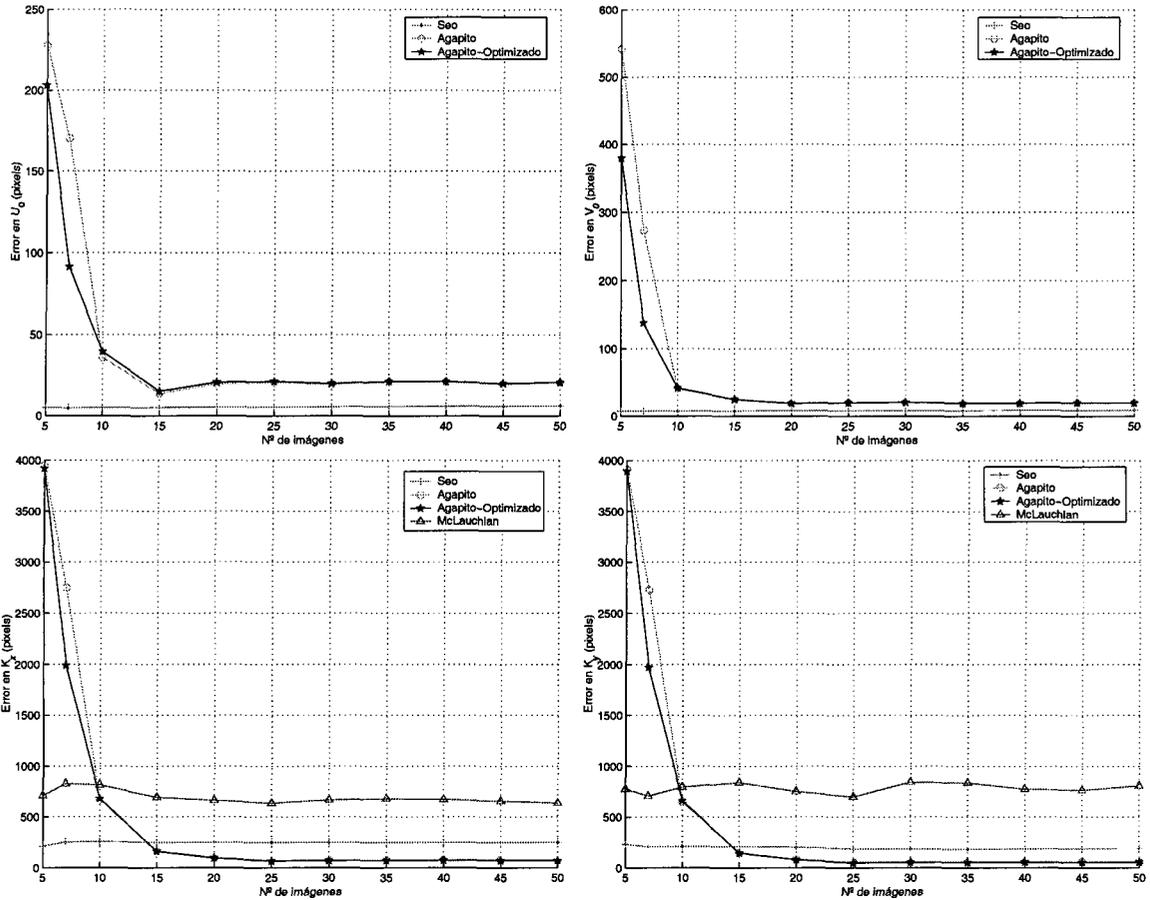


Figura 5.8: Errores obtenidos de los parámetros intrínsecos al variar el número de vistas procesadas en condiciones próximas a las reales

res no muy cercanos a los del método de Seo (unos 150 pixels de diferencia, equivalente a unos 0.6 mm de distancia focal), el cuál obtiene prácticamente el mismo valor a partir de 5 imágenes procesadas. El comportamiento de estos métodos es semejante, en este caso, al obtenido sobre la cámara real (figura 5.2). En cuanto al método de McLauchlan también se observa que el valor obtenido difiere bastante del resto de los métodos, en el caso de k_x unos 500 pixels, y en el caso de k_y la diferencia alcanza los 700 pixels.

En las gráficas que muestran el error cometido en cada uno de los parámetros se observa que en la estimación de los factores de escalado es algo deficiente, sobre todo en el caso del método de McLauchlan ofrece donde el resultado en la estimación de k_x y k_y sobrepasa los 800 pixels de error. El método de Seo ronda los 200 pixels. En el caso del método de Agapito tiene un error en ambos casos entre 60 y 70 pixels. Esto se debe a que hay bastante diferencia entre la distancia focal nominal y la efectiva que se han utilizado en la configuración de la cámara. Por otro lado, en el centro óptico y en la orientación el método de Seo es el que obtiene mejores resultados. El error en R_x está alrededor de medio grado y de 0.65° en R_y en el método de Seo cuando el método converge. Los

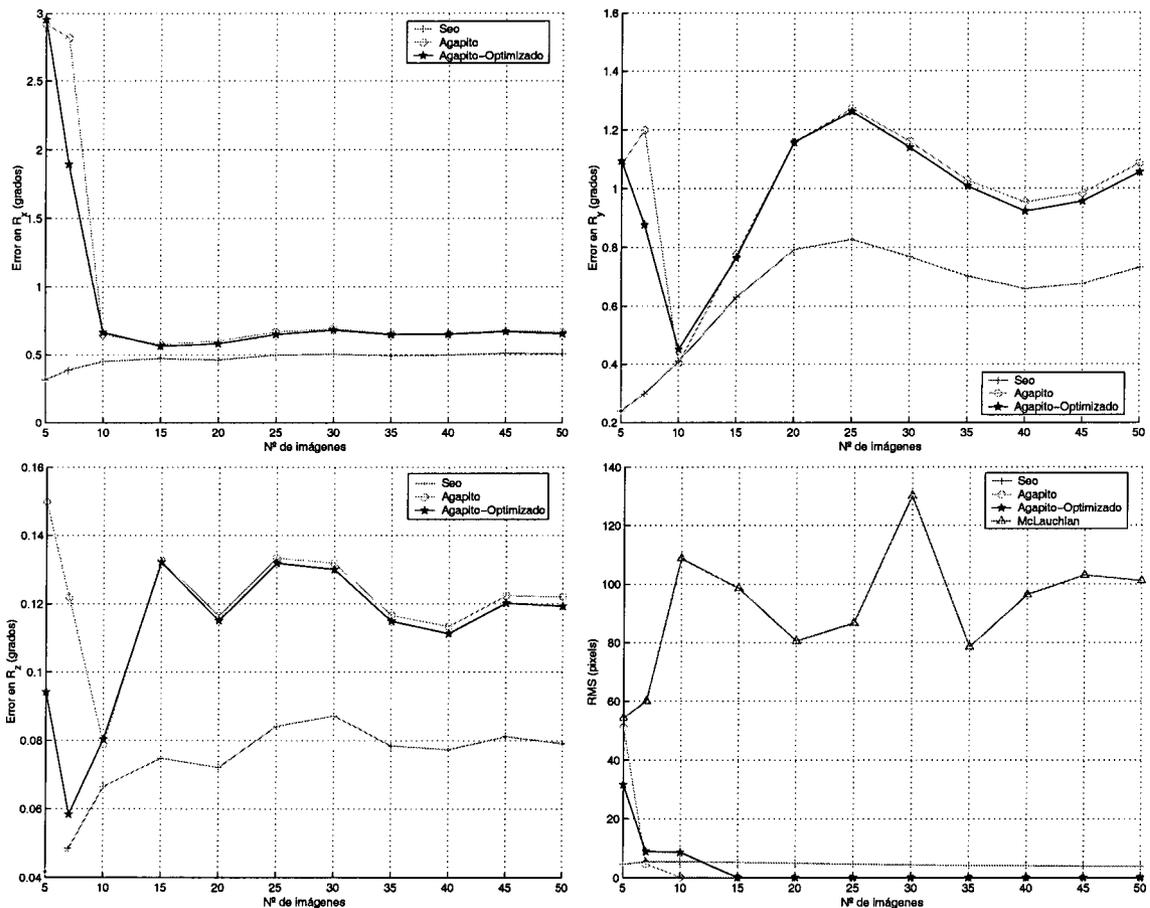


Figura 5.9: Errores obtenidos de la orientación y el error RMS al variar el número de vistas procesadas en condiciones próximas a las reales.

métodos de Agapito presentan un mayor error en estos parámetros, superando el 40% de la magnitud obtenido por el método de Seo. Por último, el error en R_z es mínimo cuando se logra la convergencia (alrededor de una décima de grado).

5.5. Comparativa de resultados con los métodos clásicos de calibración

Esta sección ofrece un análisis comparativo de los resultados obtenidos en los métodos de calibración manual con los métodos de autocalibración presentados en este capítulo.

La comparación se realiza en base a los resultados obtenidos en la estimación de algunos de los parámetros intrínsecos, ya que los extrínsecos obtenidos se refieren a cosas distintas. En los métodos de calibración manual los parámetros extrínsecos indican la colocación del patrón respecto a la cámara, mientras que en los de autocalibración se refieren a la posición de la cámara en las distintas vistas respecto a la utilizada como

referencia (la primera de ellas).

5.5.1. Experimentos en el simulador

En este apartado se comparan los resultados obtenidos por los dos tipos de métodos en el simulador. No obstante, hay que tener en cuenta que pocos experimentos son los que pueden ser comparados de un caso al otro, ya que son diferentes los factores que se analizan en cada tipo de calibración. Por ejemplo: en los métodos de autocalibración no influye el error en las coordenadas 3D debido a que no se usa este conocimiento para realizar la calibración; en los métodos de calibración manual no se analiza el número de imágenes ya que en la mayoría de los casos sólo se utiliza una imagen para realizar la calibración; o en los métodos de calibración manual no influye el desalineamiento entre el origen de coordenadas de la cámara (foco) con el del eje de giro de la misma.

Entre los resultados que sí es posible comparar se encuentran los obtenidos al variar el nivel de ruido en las coordenadas 2D de la cámara (figuras 2.10 - 2.12 vs 4.18 - 4.19). En general, se puede observar que los métodos de calibración manual obtienen mejores resultados, aproximadamente la mitad del error en los parámetros intrínsecos, que los de autocalibración. Sin embargo, los métodos de Tsai optimizado y Heikkilä son una excepción a este resultado general. Por otro lado, el método de McLauchlan es el que obtiene el mejor resultado, por lo menos en el caso de que durante el proceso de calibración el centro óptico y la distancia focal permanezcan constantes. También hay que decir que si bien el método de Seo tiene unos niveles de error en la estimación del centro óptico, producto de la no variación de este factor en el experimento realizado, semejantes a los métodos de calibración manual, en el caso de los factores de escalado supera ampliamente a todos los métodos, tanto de un tipo como de otro.

Si se compara la influencia de la distorsión sobre ambos tipos de métodos se observa que las dos variantes del método de Agapito son los más sensibles. El método de Seo tiene también peor comportamiento que los métodos clásicos en la estimación de los factores de escalado, aunque en la estimación del centro óptico obtiene resultados similares a los métodos clásicos, teniendo un comportamiento semejante al método de Zhang en este parámetro. Esto se justifica con lo comentado anteriormente de que es un método que realiza una primera fase donde obtiene una aproximación inicial a los factores de escalado, siendo este parámetro el que absorbe la mayor parte del ruido, aunque se realice una fase posterior de refinamiento que incluye todos los parámetros.

Por último, si se comparan los tiempos de ejecución se observa que si el número de imágenes es muy alto, el método de Seo es mucho más lento que cualquiera de los métodos clásicos (incluido el de Ahmed), aunque como lo habitual es que se use con

secuencias pequeñas, ya que el error suele estabilizarse en 15 o 20 imágenes, entonces el tiempo de ejecución sería inferior a 10 segundos, ligeramente superior a métodos como los de Faugeras o Zhang. En cambio, el método de McLauchan tiene tiempos semejantes a los de Ahmed, o incluso superiores ya que es un método que necesita secuencias bastante largas. El método de Agapito-optimizado tiene tiempos semejantes a los métodos de Heikkilä y Tsai si el número de imágenes no es muy alto: menos de un segundo si la secuencia es inferior a 25 imágenes. Por último el método lineal de Agapito, aunque para secuencias entre 20 y 25 imágenes tarda del orden de una tercera parte de lo que tarda su variante optimizada (ronda las tres décimas de segundo), su tiempo de ejecución es superior a los métodos de Batista y Lineal. Obviamente, estos resultados miden exclusivamente el esfuerzo computacional requerido por cada método. En el caso de los métodos de autocalibración habría que añadir a los tiempos de proceso el tiempo necesario para la adquisición de cada imagen. De esta forma, estos métodos necesitarían mucho más tiempo para realizar el proceso completo de calibración que el que emplean los métodos clásicos.

Como conclusión final hay que indicar que los métodos manuales proporcionan mejores resultados y en menos tiempo, pero en cambio, tienen el inconveniente de que necesitan la utilización de un patrón como referente. En cambio, los métodos de autocalibración tienen la ventaja de poder realizar el proceso de forma autónoma. Además, evitan una de los posibles fuentes de ruido que mayor incidencia puede tener sobre el resultado, que es la imprecisión en la posición de los puntos en el patrón. También hay que decir que los métodos de autocalibración utilizan modelos más simplificados de cámara que los métodos clásicos. Esto produce que estos métodos sean menos robustos en un mayor número de circunstancias, aquellas en las que exista una clara influencia de los parámetros que no modelan, en especial, la distorsión.

5.5.2. Experimentos con una cámara real

En este apartado se comparan los resultados obtenidos por los dos tipos de métodos a partir de datos reales. Para realizar esta comparativa se ha calibrado con los métodos clásicos una de las secuencias analizadas en las pruebas sobre datos reales de los métodos de autocalibración. En concreto, se han analizado dos situaciones, una donde los parámetros intrínsecos permanecen estáticos y otra donde varían.

En los resultados presentados en las siguientes apartados no aparece el método de Batista. Esto se debe a que este método presenta una alta dispersión de los resultados que obtiene, debido a que la calibración no converge a valores estables. La causa radica probablemente en que el patrón, además de encontrarse muy alejado (cerca de 3 metros),

aparece muy descentrado en las imágenes, cuestión que hace que este método se torne muy inestable.

Parámetros intrínsecos fijos

En este caso, se ha utilizado la secuencia de imágenes en la que se mantienen fijos los parámetros intrínsecos utilizando una distancia focal de aproximadamente 9.4 mm (posición del zoom en 9000 pasos de motor). El movimiento realizado con la cámara es el mismo que el descrito para los experimentos de autocalibración, siendo el semiángulo del cono descrito por el eje óptico de la cámara en su movimiento de 6° . En las figuras 5.10 - 5.12 se muestran los resultados obtenidos en cada una de las vistas. El resultado más llamativo es la gran dependencia que existe en todos los métodos de los parámetros intrínsecos sobre el valor de los parámetros extrínsecos. Esto ya se ha comentando en el capítulo dedicado a este tipo de métodos. En todo caso, hay que decir que los efectos de dispersión en todas las gráficas se pueden haber visto ampliados debido a que el patrón ocupa una parte relativamente pequeña de la imagen, y además, se encuentra descentrado debido al movimiento realizado por la cámara.

Los resultados en la estimación de los parámetros intrínsecos se comparan con los obtenidos en los métodos de autocalibración al procesar una secuencia de 40 imágenes (figura 5.3). Se observa que en ambos casos el centro óptico varía en función del movimiento realizado, el cuál tiene forma sinusoidal en cada eje y se identifica claramente en movimiento realizado por la cámara. Este efecto es común a los dos tipos de métodos, aunque es mucho más acusado (una mayor dispersión) sobre los métodos clásicos. La excepción la presenta el método de Zhang que tiene un comportamiento muy irregular, pero no sigue la forma sinusoidal del resto de las curvas.

En el caso de los factores de escalado se observa que los métodos de autocalibración son tan estables como los métodos de Faugeras y Lineal, no en vano, se basan en modelos de cámara similares. En general, se observa que la dispersión de los resultados es muy superior en los métodos clásicos, lo que se explica porque éstos realizan una calibración individualizada de cada imagen, mientras que los métodos de autocalibración realizan un ajuste global. Llama la atención la gran diferencia en el valor obtenido de estos parámetros.

Si se analiza el resultado de los parámetros extrínsecos se observa que se produce una alta dispersión de los resultados, en especial, en el método de Tsai cuando, teóricamente, el valor de T_z debería permanecer constante. Los métodos más estables siguen siendo los de Faugeras y Lineal. Estos resultados son coincidentes con los obtenidos en la estimación de los factores de escalado.

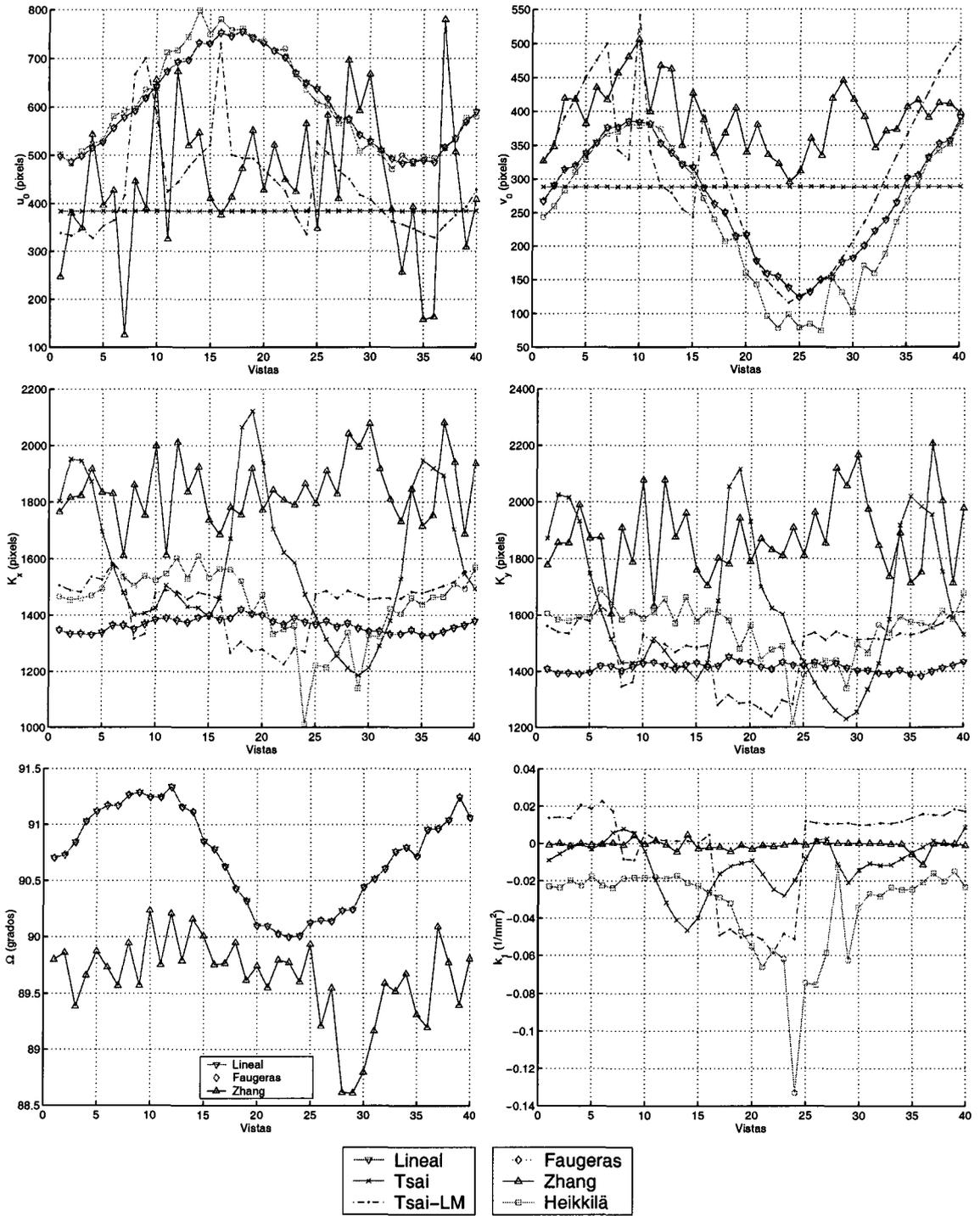


Figura 5.10: Estimación de los parámetros intrínsecos obtenidos por los métodos clásicos al variar los ángulos de azimut y elevación de la cámara.

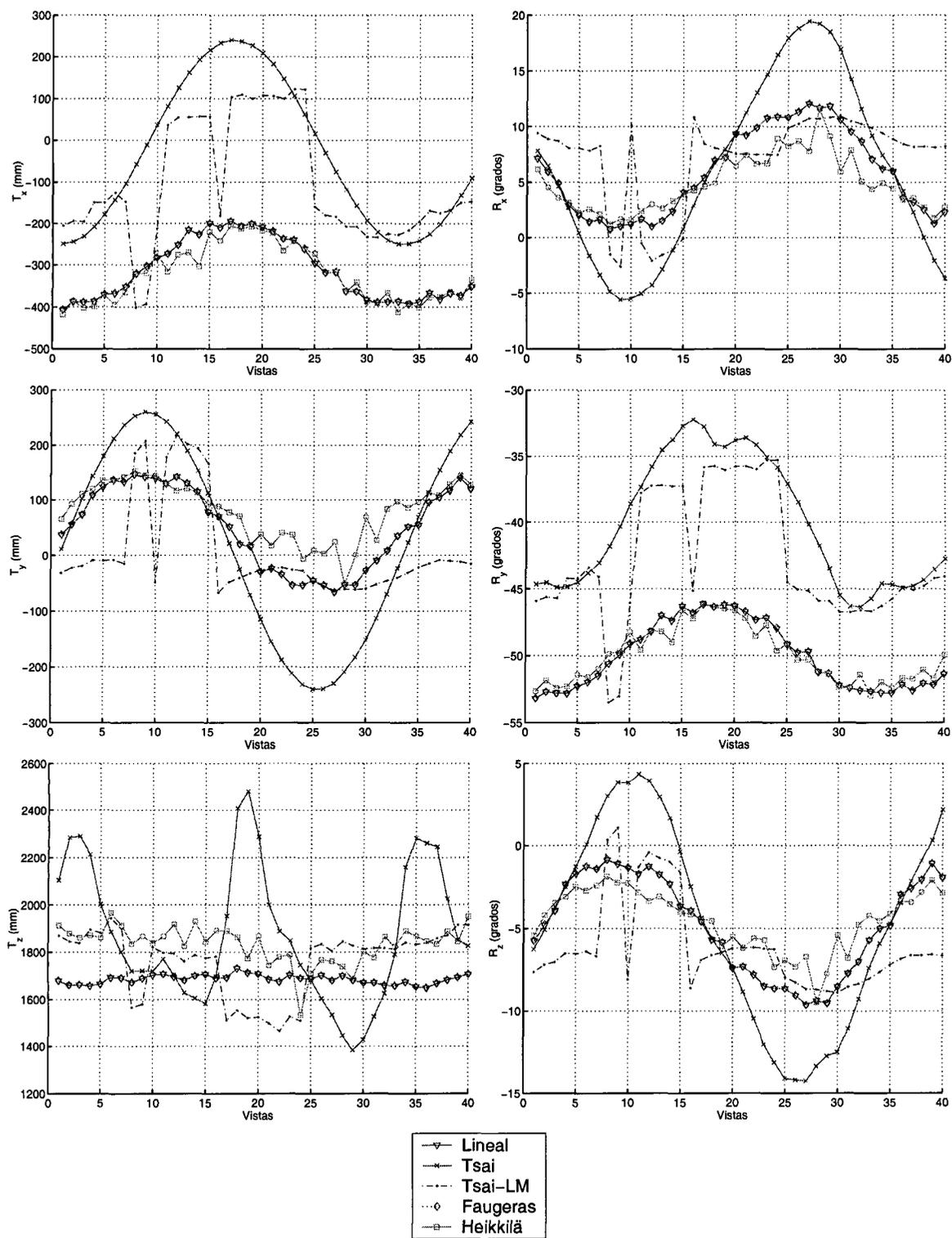


Figura 5.11: Estimación de los parámetros extrínsecos obtenidos por los métodos clásicos al variar los ángulos de azimuth y elevación de la cámara.

En teoría, la amplitud de las curvas de T_x y T_y debería ser como máximo de 30 cm. Este valor se ha obtenido en función de la distancia aproximada a la que se encuentra el patrón (cerca de 3 metros) y la semi-apertura del cono del movimiento realizado (6°). Ninguno de los métodos obtiene una amplitud próxima a este valor. Se observa que los métodos que presentan la amplitud más cercana a esta estimación son los métodos de Faugeras y Lineal (20 cm). Este resultado coincide con el hecho de que estos métodos son los más estables en la estimación de la distancia focal. El método de Heikkilä obtiene valores muy parecidos a estos dos métodos. Es destacable el resultado de Tsai ya que es el método que presenta un mejor ajuste a una curva sinusoidal, pero, en cambio, es el que presenta una mayor amplitud (cerca de 50 cm). Esto se debe a que este método mantiene fijos los valores del centro óptico.

Según el movimiento realizado por la cámara la amplitud máxima de R_x y R_y debería ser de 12° y R_z debería permanecer constante ya que no existe movimiento de ciclotorsión. Al igual que en la estimación de la posición, los métodos de Faugeras, Lineal y Heikkilä son los que obtienen los resultados más próximos a estos valores y el método de Tsai el que presenta las mayores amplitudes en la curvas. Destaca la mala estimación de todos los métodos en el caso de R_z .

Estos resultados en la estimación de la orientación no pueden ser comparados en magnitud con los obtenidos por los métodos de autocalibración (figura 5.4). Esto se debe a que la orientación obtenida por los métodos clásicos es respecto al patrón y en los de autocalibración lo es respecto a la primera imagen de la secuencia. Lo que sí se puede comparar es el movimiento realizado, analizando el rango de variación de las curvas. En este caso, se observa que en R_x y R_y los métodos de autocalibración también superan la amplitud de la curva mencionada antes. Donde estos métodos obtienen mejores resultados es en R_z ; aún así la dispersión obtenida es de 1.5 grados.

Al analizar el error en la reconstrucción de los puntos se observa que el método de Tsai presenta un comportamiento muy irregular en los resultados, aparte de obtener niveles altos de error. Aún así, el peor resultado lo obtiene el método de McLauchlan que alcanza los nueve pixels de error. El resto de los métodos de calibración obtienen un valor estable en torno a medio pixel, mientras que el resto de los de autocalibración presentan un error prácticamente nulo. Esta diferencia en los errores puede deberse a la influencia del ruido en las coordenadas 3D del patrón, ya que sólo influye en la calibración realizada por los métodos clásicos.

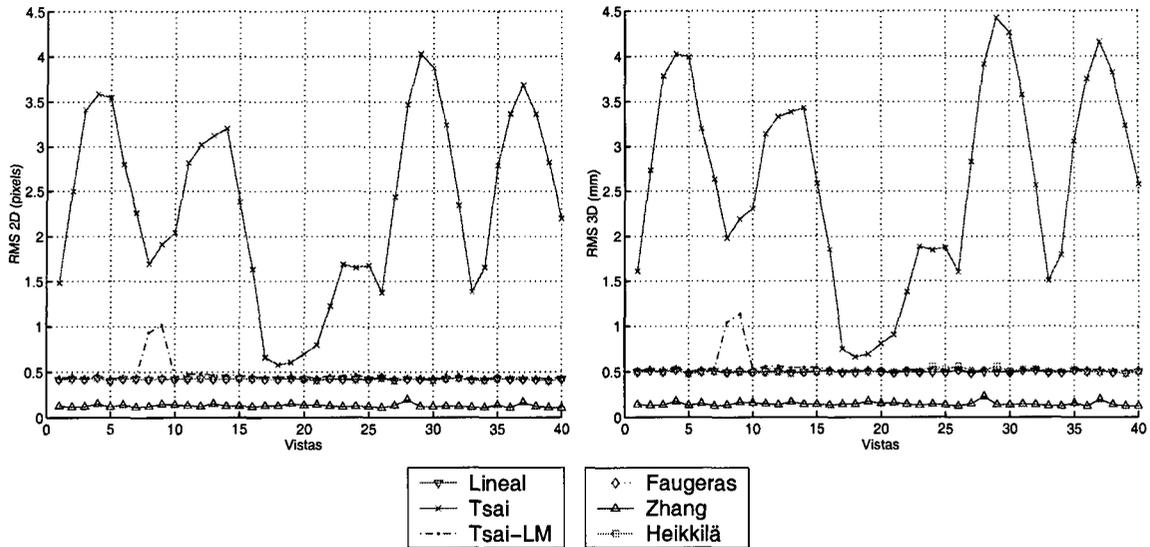


Figura 5.12: Error en la reconstrucción de las coordenadas de los puntos obtenidos por los métodos clásicos al variar los ángulos de azimut y elevación de la cámara.

Parámetros intrínsecos variables

En este caso, se ha utilizado la secuencia de imágenes en la que ha variado el zoom, aumentando la distancia focal de forma lineal. La distancia focal ha variado de 4.5 mm en la primera imagen a 13.5 mm en la última. El semiángulo del cono descrito por el eje óptico de la cámara en su movimiento es de 4° .

En las figuras 5.13 - 5.15 se muestran los resultados obtenidos en cada una de las vistas. Aquí también se puede observar que el valor estimado de los parámetros intrínsecos depende en gran medida de la orientación y posición del patrón respecto a la cámara, siendo éste un efecto común a todos los métodos.

En este caso las gráficas se comparan con las de la figura 5.5, observándose efectos parecidos a los obtenidos en el experimento anterior. En especial, se puede destacar que los factores de escalado tienen un comportamiento lineal en los métodos de auto-calibración no apreciándose inestabilidad en las curvas, mientras que en los métodos de calibración clásica sí aparece este efecto y en algunos casos bastante acentuado. En este caso se observa que la dispersión de los resultados es menor que en el experimento anterior debido que el desplazamiento de la cámara es menor.

El error en la reconstrucción de la posición de los puntos muestra una forma semejante a las gráficas mostradas en la figura 3.5, salvo el caso del método de Tsai y de Batista. Como se demostró en el estudio realizado sobre los métodos clásicos, estos dos métodos son los más sensibles a la desviación de la configuración de la cámara y del patrón respecto a la aproximaciones iniciales utilizadas.

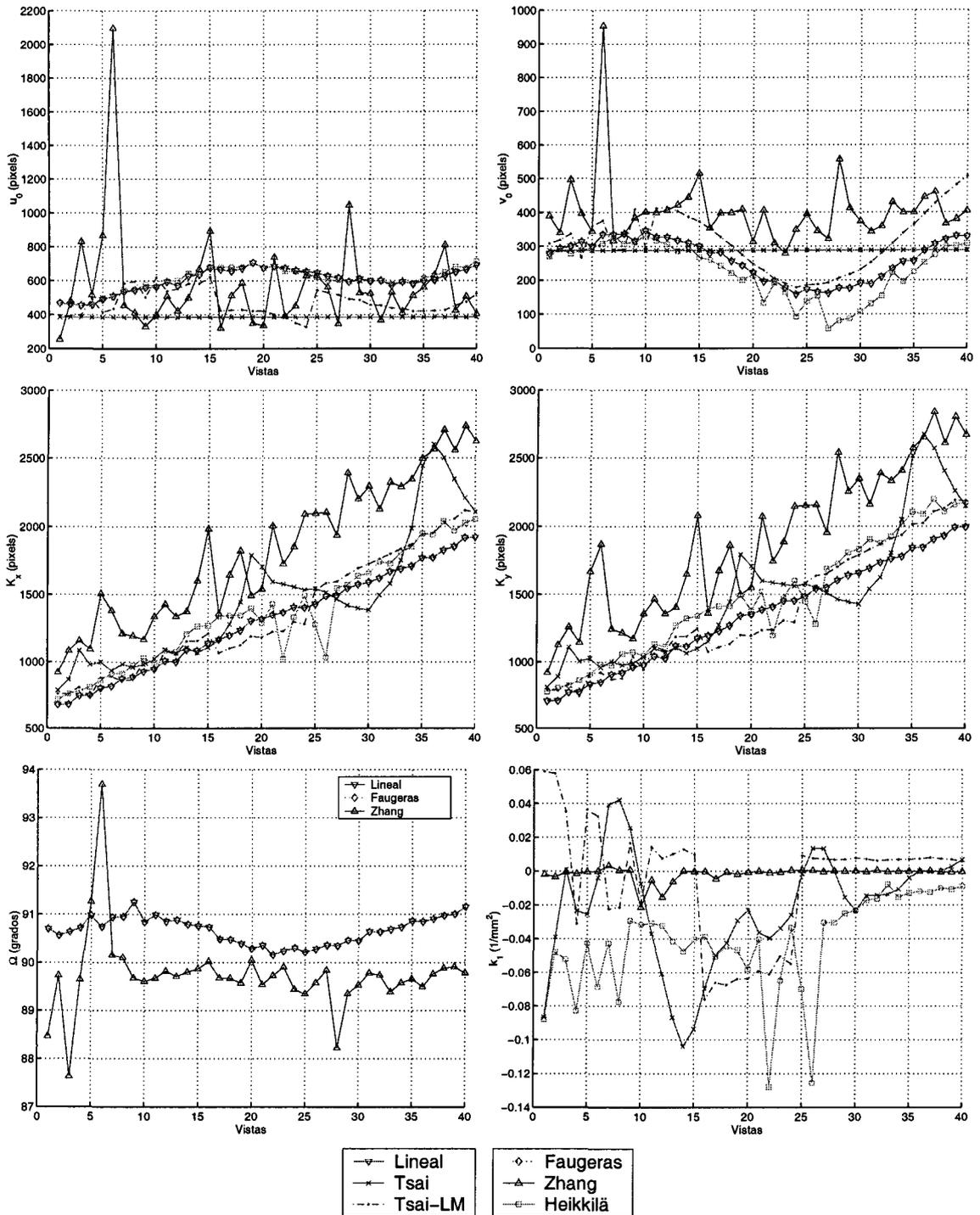


Figura 5.13: Estimación de los parámetros intrínsecos obtenidos al variar los ángulos de azimut y elevación de la cámara. El zoom varía linealmente.

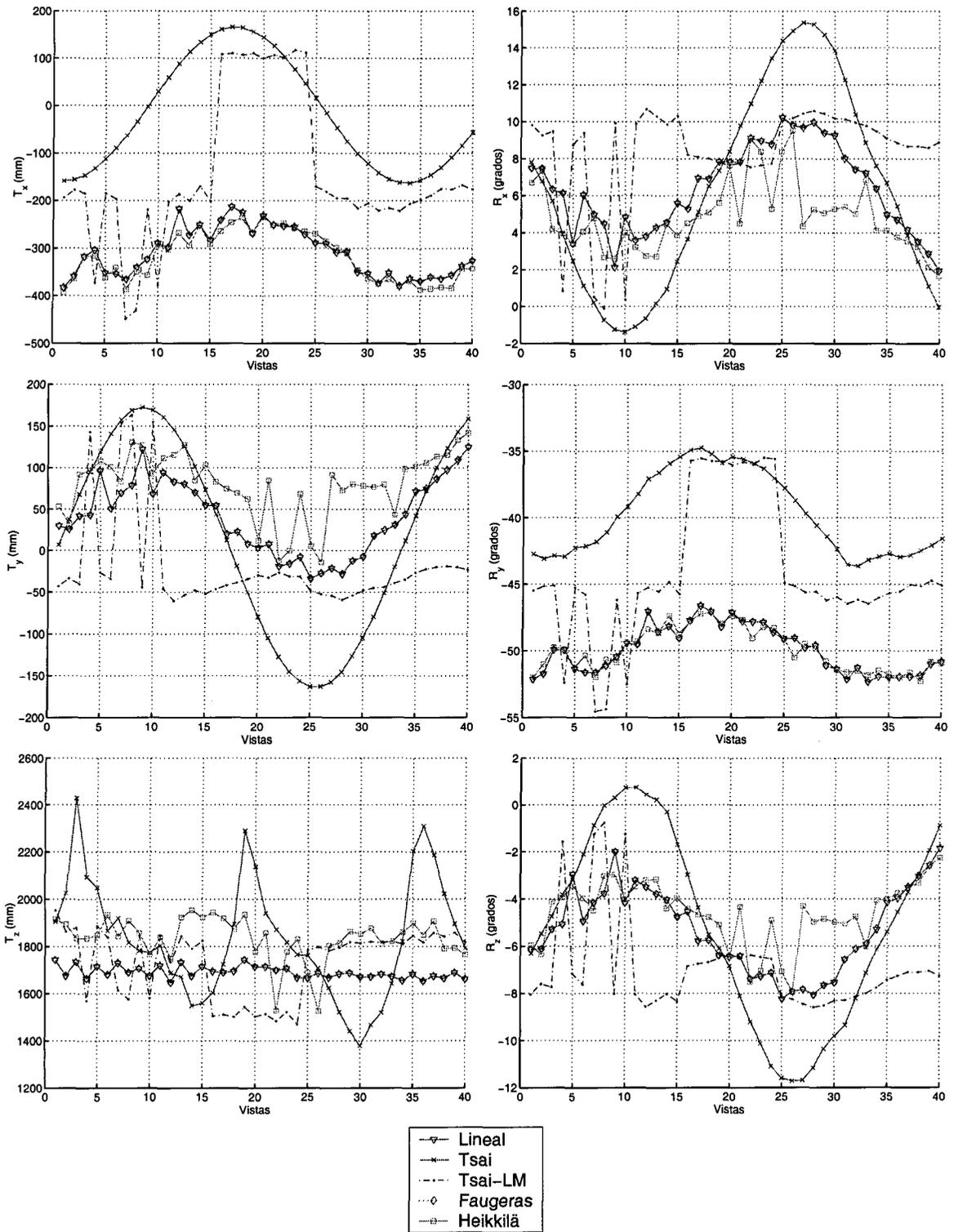


Figura 5.14: Estimación de los parámetros extrínsecos obtenidos al variar los ángulos de azimut y elevación de la cámara. El zoom varía linealmente.

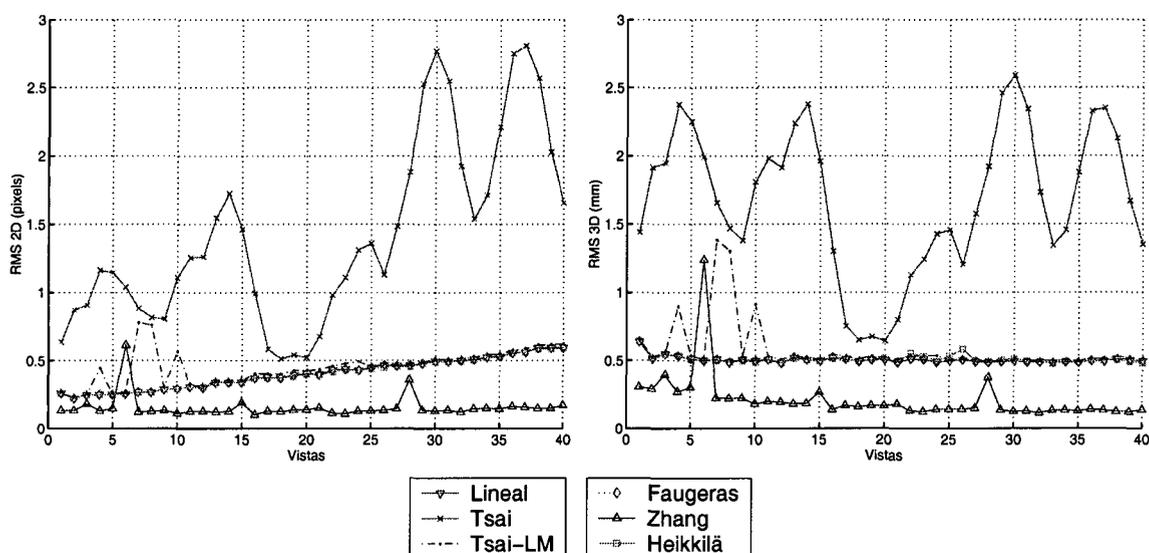


Figura 5.15: Error en la reconstrucción de las coordenadas de los puntos obtenidos al variar los ángulos de azimuth y elevación de la cámara. El zoom varía linealmente.

Por otro lado, se observa en la figura 5.15 que el error 2D aumenta progresivamente, mientras que el error 3D es estable. Esto es indicativo de que la mayor fuente de ruido se encuentra en las coordenadas 3D.

5.6. Conclusiones

En este capítulo se han analizado diversos experimentos con datos obtenidos de una cámara real con el objetivo de caracterizar la precisión y estabilidad de los métodos implementados, así como, medir el coste computacional de cada uno de ellos. Los resultados obtenidos por cada método han sido comparados entre sí. También se han examinado las diferencias en los resultados de los métodos de autocalibración respecto a los métodos clásicos.

En los experimentos sobre datos reales destaca la disparidad en los resultados obtenidos por los diferentes métodos, es decir, en la mayoría de los parámetros los valores obtenidos por cada método difieren bastante entre sí.

Los resultados obtenidos muestran que, como era de esperar, el método de McLauchlan funciona muy bien si los parámetros intrínsecos se mantienen constantes a lo largo de la secuencia utilizada para la calibración. Por contra, es un método que presenta un alto coste computacional, ya que necesita muchas imágenes para que el método converja y no puede procesar elevadas cantidades puntos. Además, también es un método muy dependiente de los valores iniciales que se le aporten, como es el caso del centro óptico o el movimiento de cámara realizado.

El método de Seo presenta mejores resultados que el método de Agapito en la estimación del centro óptico y la orientación relativa de la cámara en cada una de las vistas. En cambio, en la estimación de los factores de escalado, suele presentar peor resultado.

Los métodos de Agapito distribuyen el error entre todos los parámetros de forma más homogénea que el método de Seo. También se observa que las dos versiones del método de Agapito presentan resultados muy similares entre sí, por lo que se deduce que es más rentable el uso del método lineal ya que presenta un menor coste computacional.

Al comparar los resultados de los métodos de calibración clásicos con los de autocalibración se ha observado que los métodos clásicos proporcionan mejores resultados y en menos tiempo, pero en cambio, tienen el inconveniente de que necesitan la utilización de un patrón como referente y, además, que esté bien situado (centrado y a la distancia conveniente). En cambio, los métodos de autocalibración tienen la ventaja de poder realizar el proceso de forma autónoma. Además, evitan una de las posibles fuentes de ruido que mayor incidencia pueden tener sobre el resultado que es la imprecisión en la posición de los puntos en el patrón. Por último, indicar que en los resultados con datos reales se observa que los métodos de autocalibración son mucho más estables y precisos que los métodos clásicos. Esto se debe a que en los métodos clásicos se realiza un ajuste local de los parámetros, es decir, de cada imagen por separado y, en cambio, en los métodos de autocalibración se realiza un ajuste global de todas las imágenes.

Capítulo 6

Conclusiones

En este capítulo se resumen los principales aspectos recogidos en esta tesis. En primer lugar, se describe la metodología seguida en el estudio experimental y la aportación de las herramientas usadas para dicho estudio. A continuación se indican aquellos resultados más destacados obtenidos tanto por los métodos de calibración que utilizan un patrón como referente como los métodos de autocalibración. Hay que comentar que aquellos resultados de menor importancia sobre cada tipo de método se recogen en las conclusiones de cada uno de los capítulos precedentes. Posteriormente se han detallado aquellas consideraciones comunes a todos los métodos. Por último, se especifican las principales contribuciones del trabajo y las posibles líneas de actuación de cara al futuro.

Sobre la metodología

En primer lugar, en esta tesis se ha realizado una revisión bibliográfica extensa sobre los métodos de calibración y autocalibración de cámaras. Además, para los métodos que se han implementado para la realización de este estudio se ha presentado una amplia descripción de sus algoritmos, de forma que su notación ha sido unificada. Este hecho es resaltable debido a que en la bibliografía se utilizan diferentes notaciones para referirse a los mismos elementos, cuestión que a los lectores puede llevar a equívocos.

Se ha desarrollado un simulador que permite reproducir las condiciones de proyección de puntos del espacio tridimensional en el plano imagen. Esta herramienta se ha construido de forma que fuera versátil y flexible. Ha permitido la integración de código original de otros autores para su estudio en un entorno controlado, lo que ha posibilitado la comparación de resultados con los obtenidos por otros autores.

Los experimentos realizados se han realizado tanto sobre el simulador como sobre una cámara real, la cual está dotada de autoenfoco, zoom, y permite movimientos en azimut y elevación. Los experimentos sobre el simulador han permitido el análisis de la

precisión con que se obtienen cada uno de los parámetros del modelo de cámara utilizado. La realización de las pruebas en este entorno controlado ha permitido, además, analizar por separado diferentes fuentes de error sobre los resultados, cuestión que no es posible realizar sobre una cámara real debido a que no se conoce el valor real de los parámetros. El uso del simulador también ha permitido observar la incidencia de factores que no son controlables en una situación real, como puede ser el grado de distorsión de las imágenes, la ortogonalidad de los ejes del plano imagen, etc. Finalmente, ha permitido identificar qué parámetros del modelo se ven más afectados por cada uno de estos factores.

El simulador se ha mostrado como una herramienta efectiva ya que ha permitido la reproducción de muchos de los efectos obtenidos en la calibración de una cámara real, estableciéndose una serie de pasos a seguir en cualquier situación para identificar qué factores producen estos efectos. Al poder reproducir las condiciones reales con que se ha realizado la calibración se ha podido calificar la calidad con la que cada método estima cada uno de los parámetros del modelo.

Sobre los métodos de calibración y autocalibración

El trabajo se ha dividido en dos partes principales, en la primera se ha analizado el funcionamiento de varios métodos de calibración de cámaras que utilizan un patrón como referente, en la segunda, se ha estudiado los métodos de autocalibración. Entre los resultados obtenidos para los métodos de calibración clásicos cabe destacar las siguientes conclusiones:

- En primer lugar, en los métodos que utilizan un patrón como referente se ha analizado en el simulador qué situaciones son las menos favorables de cara a la aplicación de un cierto método de calibración. Como conclusión se ha obtenido que la preferencia por un método dependerá fundamentalmente de las circunstancias en las que se realice la calibración. Por ejemplo, si las imágenes tienen mucha distorsión, no es aconsejable usar los métodos de Faugeras, Ahmed, Lineal o Zhang; si los datos de entrada, ya sean las coordenadas 3D del patrón o las coordenadas 2D de su imagen, están contaminados con altos niveles de ruido, no es recomendable usar métodos como el de Heikkilä o el de Tsai-optimizado. Por otro lado, en el caso de no conocer con exactitud el valor de ciertos parámetros (centro del eje óptico, distancia focal, etc.), que sirven como aproximación inicial para métodos como el de Tsai o el de Batista, no es aconsejable el uso de ninguno de estos métodos pues son excesivamente dependientes de los mismos. En general, se observa que los métodos son bastante sensibles a aquellos parámetros que no incluyen en su modelo. En el método de Tsai se ha observado una fuerte dependencia a la aproximación

inicial introducida y que pequeñas desviaciones de ésta respecto a los valores reales provocan altos niveles de error en los resultados, en especial en los parámetros obtenidos en la segunda fase del método (distancia focal, distancia del patrón y coeficiente de distorsión). Este resultado es destacable por el hecho de que es el método más utilizado como punto de comparación cuando se propone un nuevo método.

- Si bien no se puede afirmar que un método posee mayor precisión que el resto, sí existen otros criterios que pueden ser utilizados para clasificar los métodos estudiados en base a sus características. Por un lado, desde el punto de vista del tiempo de ejecución, es claro que tanto el método Lineal como el de Batista son métodos más rápidos que el resto. Por otro lado, si se analiza el tipo de patrón utilizado, puede que los métodos de Batista y de Zhang sean los más indicados ya que, al utilizar planos, reducen el número de fuentes de ruido en las coordenadas 3D. En cambio, el resto de los métodos necesitan que los puntos estén distribuidos en, al menos, dos planos, generalmente en forma de diedro, por lo que es necesario conocer con exactitud el ángulo que forman los dos planos y la posición del origen de coordenadas común a los dos planos. Otro resultado que ha llamado la atención es que los métodos de Faugeras y Ahmed tienen resultados muy similares a los del método Lineal, pero su coste computacional es mucho más elevado, sin embargo, han demostrado ser los más estables (Faugeras y Lineal) cuando se les ha comparado con los métodos de autocalibración (mayor estabilidad de los parámetros intrínsecos frente a variaciones en los parámetros extrínsecos).
- En las gráficas obtenidas se ha observado una correspondencia entre los errores en la estimación del centro óptico y la traslación en horizontal y vertical del patrón. Esta correspondencia también se produce entre la distancia focal y la distancia del patrón respecto a la cámara.

Por otro lado, los resultados más destacables de los métodos de autocalibración se pueden resumir en:

- El método McLauchlan funciona muy bien siempre y cuando los parámetros intrínsecos se mantengan constantes. También falla, aunque en menor medida que el resto, cuando las imágenes presentan un alto grado de distorsión, o cuando elementos como la asimetría o el factor de proporción distan de sus valores idóneos. El método de Seo presenta mejores resultados que el método de Agapito en la estimación del centro óptico y la orientación relativa de la cámara en cada una de las vistas. En cambio, en la estimación de los factores de escalado, suele presentar

peores resultados. Esto se debe a que los métodos de Agapito distribuyen el error entre todos los parámetros de forma más homogénea que el método de Seo. Tampoco en este caso se puede concluir que un método funcione siempre mejor que el resto.

- En cuanto al coste computacional el método de McLauchlan necesita muchas imágenes para que el método converja y no puede procesar elevadas cantidades de puntos por vista. El método de Seo no presenta tiempos muy elevados mientras se utilicen pocas imágenes; en cambio, si el número de vistas procesadas es elevado se torna excesivamente lento. Los métodos de Agapito, en especial la versión lineal, son bastante rápidos.

En común a los dos tipos de métodos se ha encontrado que, tras la realización de los experimentos en el simulador, la medición del error en las coordenadas de los puntos no es un buen indicador de la bondad de un método. Se ha observado que, en la mayor parte de los casos, los parámetros del modelo no se ajustan a la realidad y que la desviación en la estimación de un parámetro se ve compensada por la desviación de otro, consiguiendo que ese error de reconstrucción sea mínimo. Por tanto, aunque este error sea pequeño, puede que no se obtengan con la misma precisión los parámetros intrínsecos, que es el principal objetivo de la calibración. Por este motivo se han realizado diversas pruebas de cara a averiguar cuál es la bondad de los resultados de los parámetros intrínsecos y, por extensión, la fiabilidad de la calibración realizada por cada método. Debido a que cuando se calibra una cámara real se desconocen los valores exactos de los parámetros intrínsecos, en principio, no hay forma de apreciar la precisión con que se estima cada parámetro.

En el estudio realizado sobre la calibración de una cámara a partir de imágenes reales se muestra que los resultados de cada método varían cuando se modifican las condiciones en las que se realiza la calibración. Es decir, después de analizar la estabilidad de los distintos parámetros en diferentes situaciones, se ha observado cómo el desplazamiento del patrón (variación de los parámetros extrínsecos) influye en los parámetros intrínsecos, observándose que los resultados presentan una gran dispersión, cuando teóricamente deberían permanecer constantes ya que la configuración de la cámara no varía. Por tanto, los resultados de la calibración dependen en gran medida de las condiciones que se haya realizado el proceso. La principal conclusión que se puede extraer de esto es que los procesos de calibración consiguen buenas estimaciones para el conjunto de imágenes que se han tomado como referencia. Esto es, el error de reconstrucción de los puntos es muy bajo, pero estos resultados dejan de ser válidos si la pose de la cámara cambia.

Como se ha podido comprobar, la calibración de una cámara es un problema muy complejo de resolver ya que muchos son los parámetros a determinar y muchos son los factores que influyen sobre los resultados. En este trabajo se han analizado la influencia de factores tales como perturbaciones en los datos de entrada, variaciones en la configuración de la cámara, en especial, aquellos aspectos que no parametriza el modelo utilizado, o la configuración y disposición del patrón utilizado como referente en aquellos métodos que lo emplean. Esto se ve reflejado en la gran disparidad de resultados obtenidos por los distintos métodos al calibrar el mismo conjunto de datos.

Consideraciones finales

En resumen, en esta tesis se ha demostrado no existe un método que sea superior al resto, por lo menos en todas las condiciones. A esta conclusión se ha llegado tras analizar el funcionamiento de doce métodos representativos de diversos tipos de resolución de este problema, de los cuales, ocho utilizan un patrón como referente y cuatro son de autocalibración. Este estudio comparativo ha examinado la incidencia de diversos factores en cada uno de los métodos. Por tanto, el objetivo de este trabajo ha sido identificar, en primer lugar, cuáles son las condiciones en las que cada tipo de método funciona mejor. Y, en segundo lugar, identificar los factores que inciden en el proceso de calibración de una cámara real para saber cuál de los métodos ofrece el resultado más fiable en esas circunstancias.

Tras la realización de los experimentos una cuestión que se debe responder es si los procesos de calibración empleados son fiables o no. La respuesta a esta cuestión tiene dos vertientes. Por un lado, sí se puede decir que son fiables en el sentido de que, salvo en determinadas circunstancias, el error global que se alcanza en el proceso de calibración es pequeño, pudiendo realizarse la reconstrucción 3D de objetos con elevada precisión (errores de milímetros). Pero, por otro lado, la respuesta es negativa, en el sentido de que los resultados de la calibración varían dependiendo de las condiciones en las que se realicen; es decir, los resultados obtenidos son buenas estimaciones para el conjunto de imágenes que se han tomado como referencia, pero pueden no serlo para otro conjunto diferente de muestras.

6.1. Futuras investigaciones

En el futuro se prevé ampliar el estudio en varios aspectos:

- Realizar un estudio similar donde se analicen métodos de autocalibración que permitan cualquier tipo de desplazamiento y comparar los resultados de éstos con

los obtenidos en el presente trabajo. Se trataría, en primer lugar, de observar y comparar la incidencia de diversos factores sobre diferentes métodos de este tipo, en la misma línea de lo realizado en esta tesis. En segundo lugar, se trataría de estudiar las ventajas y desventajas de diferentes tipos de métodos de autocalibración desde un punto de vista experimental. De igual forma, se puede extender este estudio a métodos muy específicos como son los orientados a cámara de ojo de pez, y su aplicación en tareas de robótica móvil.

- Estudiar nuevos esquemas de optimización multidimensional que traten problemas con múltiples mínimos locales y aplicarlos al problema de la calibración de cámaras. En esta línea parece muy atractiva la posibilidad de emplear técnicas de optimización basados en algoritmos genéticos.
- Experimentar con otras funciones objetivos. Volver a evaluar los métodos anteriores modificando la función objetivo aplicada en la fase de optimización numérica. En la mayoría de los métodos, la función objetivo empleada se basa en una suma promediada del error en la reconstrucción de los puntos en la imagen, aunque otras funciones objetivos son posibles [Hartley and Zisserman, 2000].

Apéndice A

Conceptos básicos

En este apéndice se pretende acercar al lector a una serie de conceptos que se manejan en el ámbito de la calibración de cámaras, así como presentar algunos elementos de la notación utilizada a lo largo del documento. Se describen una serie de operaciones o elementos básicos que son frecuentemente usados por métodos de calibración, así como el algoritmo empleado para su obtención. Hay que hacer notar que en la mayoría de los casos existen múltiples formas de estimar estos elementos, pero se ha intentado seleccionar las rutinas más simples, y por tanto más rápidas, ya que la velocidad puede ser un parámetro clave en especial en los métodos de autocalibración. La misma rutina de obtención de estos elementos se utilizó en todos los métodos implementados en la comparativa, ya que este trabajo se ha orientado a comparar los métodos de calibración y no la fiabilidad de estas rutinas básicas.

A.1. Descomposición en Valores Singulares (SVD - Singular Value Decomposition)

Se suele usar esta operación matricial para resolver sistemas de ecuaciones sobredeterminados del tipo $Ax = 0$, donde de otra forma se obtendría la solución trivial. Al no presentar este problema una solución exacta se recurre a minimizar la norma $\|Ax\|$ sujeta a $\|x\| = 1$. La solución viene dada por el autovector de $A^t A$ asociado al autovalor más pequeño, o equivalentemente, el vector singular derecho de la matriz A asociado con el menor valor singular de ésta.

A.2. Normalización Isotrópica

Debido a que en muchos casos existe una dependencia del valor de las coordenadas de los puntos para el cálculo de elementos como la Homografía o la Matriz Fundamental, se ha demostrado [Hartley, 1997a] que es necesaria una normalización de los datos de entrada para atenuar el efecto del ruido en estos datos sobre el cálculo del elemento de interés.

La normalización isotrópica trata que el promedio de la distancia del conjunto de puntos que se utiliza como datos de entrada sea igual a $\sqrt{2}$, esto significa que el punto promedio sea $[1, 1, 1]^t$. Debido a que habitualmente las coordenadas 2D de una imagen son homogéneas, sólo se realiza la normalización sobre los dos primeros elementos (u, v) de dichas coordenadas. Para realizar esta normalización, primero se traslada el origen de coordenadas al centroide de los puntos, para después escalar las coordenadas de los puntos de manera que su distancia promedio al centro sea $\sqrt{2}$. Esta transformación se aplica independientemente a cada conjunto de datos de entrada (imágenes) que haya.

A.3. Homografía

La homografía (H) es una matriz que relaciona dos proyecciones (coordenadas 2D de la imagen), bien obtenidas por dos cámaras distintas o bien obtenida por una cámara en dos posiciones distintas, de un mismo punto 3D estático en el espacio. Es decir, es una matriz 3x3 que relaciona cada par de puntos en correspondencia de dos imágenes tal que:

$$m_2 \doteq Hm_1 \quad (\text{A.1})$$

donde m_1 son las coordenadas 2D ($[u_1, v_1, w_1]^t$) en la primera imagen y m_2 las de la segunda imagen ($[u_2, v_2, w_2]^t$). Habitualmente estas coordenadas de imagen suelen estar normalizadas de forma que $w = 1$.

El algoritmo más frecuentemente usado para obtener la homografía a partir de los conjuntos de puntos en correspondencia de dos imágenes es el de *Transformación Lineal Directa* (DLT - Direct Linear Transformation) [Abdel-Aziz and Karara, 1971]. Este método lineal se suele usar como aproximación inicial a métodos de optimización donde se trata de minimizar funciones de costo, en unos casos basadas en distancias geométricas y otros casos basadas en errores algebraicos.

Para evitar el factor de proporcionalidad, la ecuación A.1 se puede reescribir como: $m_2 \otimes Hm_1 = 0$. Si se representa la matriz H por sus filas: $H = [h^1 h^2 h^3]^t$, entonces la

ecuación anterior se puede expresar como:

$$m_2 \otimes H m_1 = \begin{bmatrix} v_2 h^3 m_1 - w_2 h^2 m_1 \\ w_2 h^1 m_1 - u_2 h^3 m_1 \\ u_2 h^2 m_1 - v_2 h^1 m_1 \end{bmatrix} = 0 \quad (\text{A.2})$$

Hay que tener en cuenta que $h^i m_1$ es el producto escalar de estos dos vectores. Para obtener los elementos de H se puede transformar la ecuación anterior en el siguiente sistema de ecuaciones del tipo $Ah = 0$ y resolverlo mediante Descomposición en Valores Singulares:

$$\begin{bmatrix} 0^t & -w_2 m_1^t & v_2 m_1^t \\ w_2 m_1^t & 0^t & -u_2 m_1^t \\ -v_2 m_1^t & u_2 m_1^t & 0^t \end{bmatrix} \begin{bmatrix} h^{1t} \\ h^{2t} \\ h^{3t} \end{bmatrix} = 0 \quad (\text{A.3})$$

Hay que tener en cuenta que sólo dos de las tres ecuaciones son linealmente independientes, por tanto son necesarios como mínimo cuatro puntos en correspondencia para poder obtener la homografía.

Debido a que este algoritmo DLT no es invariante ante transformaciones de semejanza, se hace conveniente aplicar una transformación (S) de normalización a las coordenadas de cada imagen:

$$\hat{m}_1 = S_1 m_1 \quad , \quad \hat{m}_2 = S_2 m_2 \quad (\text{A.4})$$

Esta transformación se puede obtener utilizando una normalización isotrópica que se realiza antes de ejecutar el algoritmo DLT. Como los puntos están transformados, la homografía obtenida también estará transformada (\hat{H}). Por tanto, para obtener H hay que invertir la transformación de normalización: $H = S_2^{-1} \hat{H} S_1$. Otro método de obtener la Homografía aparece descrito en [Kanatani, 1998].

A.4. Medidas de error

Las siguientes medidas de error son utilizadas para cuantificar la bondad de los resultados obtenidos por los diferentes métodos:

Error en la reconstrucción de las coordenadas 2D (Error 2D). Mide la distancia euclídea promedio entre las coordenadas reales de los puntos en la imagen y su estimación.

$$Err_{2D} = \sum_{i=1}^N \sqrt{|u_i - \hat{u}_i|^2 + |v_i - \hat{v}_i|^2} \quad (\text{A.5})$$

donde $m_i = [u_i, v_i]^t$ son las coordenadas 2D reales de un punto en la imagen y $\hat{m}_i = [\hat{u}_i, \hat{v}_i]^t$ son las estimadas utilizando el modelo calibrado. N es el número de puntos procesados.

Error en la reconstrucción de las coordenadas 3D (Error 3D). Mide la distancia euclídea promedio entre las coordenadas tridimensionales reales de los puntos del patrón y su estimación.

$$Err_{3D} = \sum_{i=1}^N \sqrt{|X_i - \hat{X}_i|^2 + |Y_i - \hat{Y}_i|^2 + |Z_i - \hat{Z}_i|^2} \quad (A.6)$$

donde $M_i = [X_i, Y_i, Z_i]^t$ son las coordenadas tridimensionales reales de un punto del patrón y $\hat{M}_i = [\hat{X}_i, \hat{Y}_i, \hat{Z}_i]^t$ son las estimadas utilizando el modelo calibrado.

Error RMS. Esta una medida de error es función de la distancia euclídea (d) entre los puntos medidos y los estimados utilizando el modelo calibrado para reconstruirlo, y del número de puntos de la muestra (N). Los elementos del vector d se obtienen, por tanto, como: $d_i = \sqrt{|u_i - \hat{u}_i|^2 + |v_i - \hat{v}_i|^2}$. La formula del error es:

$$E_{RMS} = \frac{\|d\|}{\sqrt{N}} \quad (A.7)$$

Es parecida a las anteriores, lo único que el resultado es normalizado.

Bibliografía

- [Abdel-Aziz and Karara, 1971] Abdel-Aziz, Y. and Karara, H. (1971). Direct linear transformation into object space coordinates in close-range photometry. In *Proc. Symposium on Close-Range Photogrammetry*, pages 1–18.
- [Agapito et al., 2000] Agapito, L. D., Hayman, E., and Reid, I. (2000). Self-calibration of rotating and zooming cameras. Technical Report OUEL 0225/00, Department of Engineering Science, University of Oxford.
- [Ahmed et al., 1999] Ahmed, M., Hemayed, E., and Farag, A. (1999). Neurocalibration: a neural network that can tell camera calibration parameters. In *IEEE International Conference on Computer Vision*, pages 463–468.
- [Armangué et al., 2000] Armangué, X., Salvi, J., and Batlle, J. (2000). A comparative review of camera calibrating methods with accuracy evaluation. In *Proceedings of 5th Ibero-American Symposium on Pattern Recognition, SIARP 2000*, pages 183–194.
- [Armstrong et al., 1996] Armstrong, M., Zisserman, A., and Hartley, R. (1996). Self-calibration from image triplets. In *Proc. 4th European Conf. on Computer Vision, Lecture notes in Computer Science*, pages 3–16.
- [Basu and Licardie, 1993] Basu, A. and Licardie, S. (1993). Modeling fish-eye lenses. In *International Conference on Intelligent Robots and Systems*, pages 1822–1828.
- [Batista, 1996] Batista, J. (1996). Explicit monoplane camera calibration. Technical report, Institute of Systems and Robotics, University of Coimbra, Portugal.
- [Batista et al., 1998] Batista, J., Araújo, H., and de Almeida, A. (1998). Iterative multi-step explicit camera calibration. In *Proc. Sixth International Conference on Computer Vision ICCV98*, volume 15, pages 709–714.
- [de Agapito et al., 1999] de Agapito, L., Hartley, R., and Hayman, E. (1999). Linear calibration of a rotating and zooming camera. In *CVPR99*.
- [de Agapito et al., 1998] de Agapito, L., Hayman, E., and Reid, I. (1998). Self-calibration of a rotating camera with varying intrinsic parameters. In *In Proc. British Machine Vision Conference*, pages 105–114.
- [Du and Brady, 1993] Du, F. and Brady, M. (1993). Self-calibration of the intrinsic parameters of camera for active vision systems. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 477–482.

- [Faugeras, 1993] Faugeras, O. (1993). *Three-dimensional computer vision: a geometric viewpoint*. MIT Press.
- [Faugeras et al., 2000] Faugeras, O., Quan, L., and Strum, P. (2000). Self-calibration of a 1d projective camera and its application to the self-calibration of a 2d projective camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1179–1185.
- [Faugeras and Toscani, 1986] Faugeras, O. and Toscani, G. (1986). The calibration problem for stereo. In *Proc. of IEEE Computer Vision and Pattern Recognition*, pages 15–20.
- [Fischler and Bolles, 1981] Fischler, M. and Bolles, R. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. Assoc. Comp. Mach.*, 24(6):381–395.
- [Gaspard and Viéville, 1996] Gaspard, F. and Viéville, T. (1996). Hierarchical visual perception without calibration. Technical Report 3002, Institut National de Recherche en informatique et en automatique (INRIA).
- [Grattoni et al., 1991] Grattoni, P., Pettiti, G., and Pollastri, F. (1991). Geometric camera calibration: a comparison of methods. In *Proc. 5th Int. Conf. on Advanced Robotics*, pages 1775–1779.
- [Guohui and Baozong, 1998] Guohui, Z. and Baozong, Y. (1998). An algorithm for self-calibration in active vision. In *Proceedings of ICSP '98*, pages 932–935.
- [Hall et al., 1982] Hall, E., Tio, J., McPherson, C., and Sadjadi, F. (1982). Measuring curved surfaces for robot vision. *Computer Journal*, 15:42–54.
- [Hartley, 1992] Hartley, R. (1992). Estimation of relative camera positions for uncalibrated cameras. In *Proc. 2nd European Conference on Computer Vision. Lecture notes in Computer science 588*, pages 579–587.
- [Hartley, 1994a] Hartley, R. (1994a). Euclidean reconstruction from uncalibrated views. *Applications of Invariance in Computer Vision*, pages 237–256.
- [Hartley, 1994b] Hartley, R. (1994b). Self-calibration from multiple views with a rotating camera. In *Proc. of the Third European Conf. Computer Vision*, volume 1, pages 471–478.
- [Hartley et al., 1999] Hartley, R., Hayman, E., de Agapito, L., and Reid, I. (1999). Camera calibration and the search for infinity. In *The proceedings of the International Conference on Computer Vision*.
- [Hartley and Zisserman, 2000] Hartley, R. and Zisserman, A. (2000). *Multiple View Geometry in Computer Vision*. Cambridge University Press.
- [Hartley, 1997a] Hartley, R. I. (1997a). In defense of the eight-point algorithm. *Pattern analysis and machine intelligence*, 19(6):580–593.

- [Hartley, 1997b] Hartley, R. I. (1997b). Self-calibration of stationary cameras. *International Journal of Computer Vision*, 22(1):5–23.
- [Hayman and Murray, 2002] Hayman, E. and Murray, D. (2002). The effects of translational misalignment in the self-calibration of rotating and zooming cameras.
- [Hecht and Zajac, 1979] Hecht, E. and Zajac, A. (1979). *Optics*. Addison Wensley New York, 2nd edition.
- [Heikkilä and Silvén, 1996] Heikkilä, J. and Silvén, O. (1996). Calibration procedure for short focal length off-the-shelf ccd cameras. In *Proc. of The 13th International Conference on Pattern Recognition*, pages 166–170.
- [Heikkilä and Silvén, 1997] Heikkilä, J. and Silvén, O. (1997). A four-step camera calibration procedure with implicit image correction. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 1106–1112.
- [Heyden and Åström, 1996] Heyden, A. and Åström, K. (1996). Euclidean reconstruction from constant intrinsic parameters. In *Proc. International Conference on Pattern Recognition*, pages 31–42.
- [Heyden and Åström, 1997] Heyden, A. and Åström, K. (1997). Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In *Conf. Computer Vision and Pattern Recognition*, pages 438–443.
- [Heyden and Åström, 1998] Heyden, A. and Åström, K. (1998). Minimal conditions on intrinsic parameters for euclidean reconstruction. In *Proc. Asian Conference on Computer Vision*.
- [Jiménez, 2000] Jiménez, J. G. (2000). *Visión por Computador*. Editorial Paraninfo.
- [Kahl et al., 2000] Kahl, F., Triggs, B., and Åström, K. (2000). Critical motions for autocalibration when some intrinsic parameters can vary. *J. Math. Imaging and Vision*, 13(2).
- [Kanatani, 1998] Kanatani, K. (1998). Optimal homography computation with a reliability measure. In *Proceedings of the IAPR Workshop on Machine Vision Applications (MVA '98)*, pages 426–429.
- [Lai, 1993] Lai, J. Z. (1993). On the sensitivity of camera calibration. *Image and Vision Computing*, 11(10):656–664.
- [Li et al., 2002] Li, H., Zhang, H., Wu, F., and Z.Hu (2002). A new linear camera self-calibration technique. In *The 5th Asian Conference on Computer Vision*.
- [Lourakis and Deriche, 2000] Lourakis, M. and Deriche, R. (2000). Camera self-calibration using the kruppa equations and the svd of the fundamental matrix: the case of varying intrinsic parameters. Technical Report 3911, Institut national de recherche en informatique et en automatique.

- [Lourakis and Deriche, 1999] Lourakis, M. I. A. and Deriche, R. (1999). Camera self-calibration using the singular value decomposition of the fundamental matrix: From point correspondences to 3D measurements. Technical Report RR-3748, INRIA.
- [Luong and Faugeras, 1994] Luong, Q. and Faugeras, O. (1994). An optimization framework for efficient self-calibration and motion determination. In *Proc. International Conference on Pattern Recognition*, pages A:248–252.
- [Luong et al., 1992] Luong, Q., Faugeras, O., and Maybank, S. (1992). Camera self-calibration: Theory and experiments. In *Proc. European Conference on Computer Vision*, pages 321–334.
- [Luong and Faugeras, 1997] Luong, Q.-T. and Faugeras, O. (1997). Self-calibration of a moving camera from point correspondences and fundamental matrices. *International Journal of Computer Vision*, 22(3):5–40.
- [Ma, 1996] Ma, S. (1996). A self-calibration technique for active vision systems. *IEEE Trans on Robotics and Automation*, 12(1):114–120.
- [Martins et al., 1981] Martins, H., Birk, J., and Kelley, R. (1981). Camera models on data from two calibration planes. *Computer Vision Graphics and Image Proc.*, 17:173–180.
- [Maybank and Faugeras, 1992] Maybank, S. and Faugeras, O. (1992). A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151.
- [Mclauchlan and Murray, 1996] Mclauchlan, P. and Murray, D. (1996). Active camera calibration for a head-eye platform using the variable state-dimension filter. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):15–22.
- [Melem, 1994] Melem, T. (1994). *Geometrical modelling and calibration of video cameras for underwater navigation*. PhD thesis, Norges tekniske høgskole, Institutt for teknisk kybernetikk.
- [Mendonça and Cipolla, 1999] Mendonça, P. and Cipolla, R. (1999). A simple technique for self-calibration. In *Proc. Computer Vision and Pattern Recognition*, volume 1, pages 500–505.
- [Min et al., 2001] Min, S., Kittler, J., and Matas, J. (2001). Empirical evaluation of a calibration chart detector. *Machine Vision and Applications*, 12:305–325.
- [More, 1977] More, J. (1977). *The Levenberg-Marquardt Algorithm, Implementation, and Theory*. Numerical Analysis, G.A., ed., Springer-Verlag.
- [Pollefeys and Gool, 1997] Pollefeys, M. and Gool, L. V. (1997). A stratified approach to metric self-calibration. In *Conference on Computer Vision and Pattern Recognition*, pages 407–412.

- [Pollefeys and Gool, 1999] Pollefeys, M. and Gool, L. V. (1999). Stratified self-calibration with the modulus constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):707–724.
- [Pollefeys and Gool, 2000] Pollefeys, M. and Gool, L. V. (2000). Some issues on self-calibration and critical motion sequences. In *Proc. Asian Conference on Computer Vision (invited presentation)*, pages 893–898.
- [Pollefeys et al., 1996a] Pollefeys, M., Gool, L. V., and Oosterlinck, M. (1996a). Euclidean 3d reconstruction from image sequences with variable focal lengths. In *European Conf. Computer Vision*, pages 31–42.
- [Pollefeys et al., 1996b] Pollefeys, M., Gool, L. V., and Oosterlinck, M. (1996b). The modulus constraint: a new constraint for self-calibration. In *Proc. International Conference on Pattern Recognition*, pages 31–42.
- [Pollefeys et al., 1999] Pollefeys, M., Koch, R., and Gool, L. V. (1999). Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. *International Journal of Computer Vision*, 32(1):7–25.
- [Salvi et al., 2002] Salvi, J., Armangué, X., and Batlle, J. (2002). A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, pages 1617–1635.
- [Seo and Hong, 1998] Seo, Y. and Hong, K. (1998). Auto-calibration of a rotating and zooming camera. In *Proc. of IAPR workshop on Machine Vision Applications*, pages 17–19.
- [Seo and Hong, 1999] Seo, Y. and Hong, K. (1999). About the self-calibration of a rotating and zooming camera: Theory and practice. In *IEEE Proceedings - Vision, Image and Signal Processing*, pages 183–189.
- [Seo and Hong, 2001] Seo, Y. and Hong, K. (2001). Theory and practice on the self-calibration of a rotating and zooming camera from two views. *IEE Proc.-Vis. Image Signal Process*, 148(3):166–172.
- [Shah and Aggarwal, 1996] Shah, S. and Aggarwal, J. (1996). Intrinsic parameter calibration procedure for a (high-distortion) fish-eye lens camera with distortion model and accuracy estimation. *Pattern Recognition*, 29(11):1775–1788.
- [Stein, 1995] Stein, G. (1995). Accurate internal camera calibration using rotation, with analysis of sources of error. In *Proc. 5th Int'l Conf. on Computer Vision*, pages 230–236.
- [Stein, 1993] Stein, G. P. (1993). Internal camera calibration using rotation and geometric shapes. Technical Report AITR-1426, Massachusetts Institute of Technology.
- [Sturm, 1997a] Sturm, P. (1997a). Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction. In *Conf. Computer Vision and Pattern Recognition*, pages 1100–1105.

- [Sturm, 1997b] Sturm, P. (1997b). Self-calibration of a moving zoom-lens camera by pre-calibration. *Image and Vision Computing*, 15:583–589.
- [Sturm, 2002] Sturm, P. (2002). Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. *Image and Vision Computing*, 20:415–426.
- [Triggs, 1997] Triggs, W. (1997). Autocalibration and the absolute quadric. In *Conference on Computer Vision and Pattern Recognition*, pages 609–614.
- [Tsai, 1987] Tsai, R. Y. (1987). A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344.
- [Tsai, 1989] Tsai, R. Y. (1989). Synopsis of recent progress on camera calibration for 3d machine vision. *The Robotics Review*, pages 147–159.
- [Viéville, 1994] Viéville, T. (1994). Autocalibration of visual sensor parameters on a robotic head. *Image and Vision Computing*, 12(4):227–237.
- [Wang et al., 2001] Wang, L., Kang, S., Shum, H., and Xu, G. (2001). Error analysis of pure rotation-based self-calibration. In *Proc. 8th International Conference on Computer Vision*, volume I, pages 464–471.
- [Weng et al., 1992] Weng, J., Cohen, P., and Herniou, M. (1992). Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:965–980.
- [Willson, 1994] Willson, R. G. (1994). *Modeling and Calibration of Automated Zoom Lenses*. PhD thesis, The Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213.
- [Xiong and Turkowski, 1997] Xiong, Y. and Turkowski, K. (1997). Creating image-based vr using a self-calibrating fisheye lens. In *Conference on Computer Vision and Pattern Recognition*, pages 237–243.
- [Yang et al., 1998] Yang, C., Wang, W., and Hu, Z. (1998). An active vision based self-calibration technique. *Chinese Journal of Computers*, 5:428–435.
- [Yu et al., 1998] Yu, W., Lee, Y., and Chung, M. (1998). Self calibration of focal length and rotation parameters using cross ratio. In *Proceedings of the Intl. Conference on Intelligent Robots and Systems*, pages 1828–1833.
- [Zeller and Faugeras, 1996] Zeller, C. and Faugeras, O. (1996). Camera self-calibration from video sequences: the kruppa equations revisited. Technical Report 2793, Institut National de Recherche en informatique et en automatique (INRIA), INRIA Sophia Antipolis, BP 93, 06902 Sophia Antipolis Cedex, France.
- [Zhang, 1998] Zhang, Z. (1998). A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, Microsoft Research, Microsoft Corporation, Redmond, WA 98052.



- [Zhang, 2000] Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334.
- [Zisserman et al., 1998] Zisserman, A., Liebowitz, D., and Armstrong, M. (1998). Resolving ambiguities in auto-calibration. *Phil. Trans. Royal Soc. London A*, 356:1193–1211.