
Un enfoque bayesiano para la extracción de características y agrupamiento en visión artificial

Miguel Ángel Cazorla Quevedo

Tesis de Doctorado

Facultad: Escuela Politécnica Superior

Director: Dr. Francisco Escolano Ruiz

2000

UNIVERSIDAD DE ALICANTE

Tesis doctoral

Un enfoque bayesiano para la extracción de características y agrupamiento en visión artificial

Presentada por
Miguel Ángel Cazorla Quevedo
Dirigida por
Dr. Francisco Escolano Ruiz

Departamento de Ciencia de la Computación e Inteligencia Artificial

Mayo del 2000

Para ti, Amparo
Para vosotros, padres

También para ti, Fran
(en tu memoria)

No hay sistemas visuales perfectos. Para construir un sistema óptimo, es necesario especificar claramente sus tareas y el coste de no realizarlas. Además, para construir una máquina visual óptima, uno debe usar tanto conocimiento como sea posible acerca del mundo y de los recursos computacionales disponibles.

[Grzywacz y Escolano, 2000]

Agradecimientos

Una tesis doctoral no surge como un producto aislado. Suele ser la consecución de un camino que a veces se bifurca en bastantes ocasiones, muchas más de las que quisiéramos, para llegar a una meta final. Este camino se empezó a andar con los cursos de doctorado realizados en el Departamento de Tecnología Informática y Computación de la Universidad de Alicante. Durante los cursos empecé a trabajar con dos personas que desde entonces han sido una fuente de constante ánimo y crítica. Me refiero a mi director de tesis Francisco Escolano y a Domingo Gallardo. Junto a ellos, y más tarde con Otto Colomina, fuimos dando forma al contenido final de este trabajo. Fueron muchos días de *pizarrón* y *latiguillos*.

Las personas integrantes del grupo de investigación al que uno pertenece da soporte y ánimos para seguir adelante. En concreto, Ramón Rizo como director, Isabel Alfonso, Pilar Arques, Patricia Compañ, Faraón Llorens y Rosana Satorre. Los recursos, tanto humanos como materiales, de este grupo han servido para que esta tesis se desarrollara de la mejor manera posible.

En el verano de 1999 realicé una estancia en el Smith-Kettlewell Eye Research Institute (SKREI) de San Francisco. Quisiera dedicar unas líneas a agradecer la tremenda amabilidad con la que me acogieron. En especial, agradecer a Norberto Grzywacz todos los esfuerzos realizados para que mi estancia fuera lo más agradable posible. El final del camino de esta tesis empezó a verse en dicha estancia. Alan Yuille y James Coughlan compartieron conmigo sus trabajos actuales. El laboratorio de Alan fue el lugar de múltiples reuniones, en las cuales se discutieron y sacaron adelante muchas de las propuestas aquí presentadas. También a ellos deseo agradecer el tiempo dedicado a ampliar mis conocimientos (en principio bastante escasos) sobre teoría bayesiana. El trabajo de Scott Konishi sirvió de base para la formulación bayesiana de la evidencia de aristas. Scott me asesoró sobre la mejor forma de obtener las mejores instantáneas que luego utilizamos en los experimentos.

Prefacio

Dotar a un robot móvil de las capacidades visuales necesarias para sobrevivir en un entorno dado no es una tarea fácil. Dicha supervivencia implica desarrollar con éxito ciertas tareas como auto-localización, evitación de obstáculos, aproximación al objetivo, etc., y el éxito en su realización depende de la capacidad de inferir el estado del mundo y la ubicación del robot en el mismo. Para ello es imprescindible disponer de representaciones y estrategias capaces de procesar y asimilar de manera eficiente y robusta el elevado flujo de información captado por el robot durante su movimiento. Reconstruir exhaustivamente el mundo, tal como enfatiza el paradigma clásico [Marr, 1982], no es una opción aceptable debido a su elevado coste computacional. Durante la última década, el paradigma de la visión activa [Aloimonos *et al.*, 1988], [Bajcsy, 1988], [Ballard, 1991], [Krotkov, 1987], [Yuille y Blake, 1992] viene abordando ésta y otras cuestiones relacionadas con el diseño de sistemas visuales artificiales óptimos en el sentido en que se enfatiza el carácter selectivo del procesamiento visual y su íntimo acoplamiento con la tarea a desarrollar, lo cual conlleva un conocimiento básico acerca del entorno y de los recursos computacionales disponibles.

La inferencia bayesiana proporciona, en combinación con la teoría de la información, un marco conceptual atractivo para formalizar el problema de la inferencia visual [Knill y Richards, 1996]. Asimismo, este planteamiento permite expresar de forma cuantitativa y cualitativa la optimalidad del sistema visual en el sentido en que el paradigma de la visión activa lo requiere. Por ejemplo, dada una tarea concreta (auto-localización, evitación de obstáculos, aproximación al objetivo) el teorema de Bayes proporciona una regla para cuantificar la bondad de la inferencia realizada por los distintos módulos del sistema que intervienen: filtrado [Geman y Geman, 1984], extracción de características [Zerubia y Chellapa, 1993], segmentación [Zhu y Yuille, 1996b], estimación de profundidad [Belhumeur, 1996], matching y reconocimiento [Kittler, 1997]. Dicha regla integra una métrica para identificar la compatibilidad de una solución propuesta con la entrada visual (verosimilitud) con el conocimiento previo del entorno disponible (información a priori). Por otro lado la teoría de la información [Cover y Thomas, 1991] aporta métricas (como la entropía, la información mutua o la distancia de Kullback-Leibler) y principios (como el principio MDL o de longitud de codificación mínima) que permiten, entre otras cosas: reducir la redundancia en la entrada visual [Bartlett *et al.*, 1998], reducir la complejidad de los modelos de contorno [Figueroa *et al.*, 1997], búsqueda de caminos óptimos [Coughlan y Yuille, 1998], determinar las características necesarias para construir modelos de imagen [Zhu *et al.*, 1997], cuantificar la efectividad de los filtros en la

extracción de características [Konishi *et al.*, 1999], así como seleccionar los módulos estrictamente necesarios para resolver la tarea, en la medida en que estos aportan información de utilidad [Rimey y Brown, 1992].

Esta tesis incorpora elementos de inferencia bayesiana y teoría de la información, en los distintos módulos (filtrado y detección de aristas, detección de puntos esquina y uniones, agrupamiento de uniones mediante búsqueda de caminos) que intervienen en la obtención eficiente de una representación geométrica robusta y adecuada para inferir parámetros de posicionamiento, en particular la orientación relativa del robot con respecto al entorno.

Preface

Endowing a robot with the visual capabilities to survive in a partial known environment is not a trivial task. Surviving implies to develop successfully several tasks such as self-localization, obstacle avoidance, target approaching, and so on, and the success depends on the capability of inferring the state of the world and the robot position in it. To do this inference it is essential to have good representations and strategies for processing and assimilating in an efficient and robust way the huge flow of information collected by the robot while it is moving around. Exhaustive world reconstruction, as the classic paradigm emphasizes [Marr, 1982], is not a acceptable option due to its high computational cost. During the last decade, the active vision paradigm [Aloimonos *et al.*, 1988], [Bajcsy, 1988], [Ballard, 1991], [Krotkov, 1987], [Yuille y Blake, 1992] has been addressing this and another questions related to the design of optimal artificial visual systems, but optimal in the sense that emphasizes selective visual processing, close coupling between vision and task, which implies a good knowledge about the environment and the available resources.

Bayesian inference provides, in combination with information theory, an attractive conceptual frame to formalize the visual inference problem [Knill y Richards, 1996]. This approach allows us to express in a quantitative and qualitative way the optimality of the visual system in the sense that the active vision paradigm requires. For instance, given a specific task (self-localization, obstacle avoidance, target approaching) the Bayes theorem provides a rule to quantify the goodness of a given inference performed by, for instance: filtering [Geman y Geman, 1984], feature extraction [Zerubia y Chellapa, 1993], segmentation [Zhu y Yuille, 1996b], depth estimation [Belhumeur, 1996], matching and recognition [Kittler, 1997]. Such a rule combines a metric to identify the compatibility of a proposed solution with the visual input (likelihood) and the prior knowledge. On the other hand, the information theory [Cover y Thomas, 1991] provides metrics (like entropy, mutual information and the Kullback-Leibler distance) and principles (like the MDL or minimum description length principle) which yields: reducing the redundancy of the visual input [Bartlett *et al.*, 1998], reducing the complexity of the contour models [Figuereido *et al.*, 1997], optimal path searching [Coughlan y Yuille, 1998], determining the necessary features to build image models [Zhu *et al.*, 1997], quantifying the effectiveness of the filters used in feature extraction [Konishi *et al.*, 1999], and selecting only those modules which are needed to solve a task by evaluating the utility of the information provide by these modules [Rimey y Brown, 1992].

This thesis includes elements of Bayesian inference and information theory

points in several visual modules (filtering and edge detection, corner and junctions detection, junction grouping by means of path searching) which are involved in obtaining an efficient and robust geometric representation which is adequate to infer positional parameters, like the robot relative orientation with respect to the environment.

Resumen

El marco de esta tesis se centra en la extracción de características y agrupamiento perceptual en visión. Nuestro propósito principal ha sido formular y experimentar con nuevos métodos para la extracción de características (principalmente aquellos relacionados con la identificación de puntos característicos y la clasificación de uniones) y el agrupamiento (realizando conexión entre uniones). El contexto de aplicación de estos métodos, la visión en robots autónomos, nos impone restricciones especiales sobre su aplicabilidad y así se observan los siguientes requerimientos: eficiencia, robustez y flexibilidad. Estos requerimientos están presentes en todas las técnicas aportadas en esta tesis: la clasificación de uniones se realiza mediante un método voraz que se basa en estadística robusta; el agrupamiento de uniones se realiza mediante un método de búsqueda de caminos con complejidad lineal media, el cual utiliza condiciones estadísticas de poda para restringir la búsqueda de caminos estables; por último, el agrupamiento de uniones encuentra la orientación relativa entre el robot y su entorno. Sin embargo, el rango de aplicación de los métodos propuestos no está restringido a inferir la orientación relativa del robot, y puede ser extendido a otras tareas como la segmentación de imágenes, estimación de profundidad o reconocimiento de objetos.

Esta tesis está estructurada en tres partes: clasificación de uniones, agrupamiento y cálculo de la orientación relativa. Los métodos presentados en cada parte son usados por las siguientes de acuerdo a una organización de complejidad incremental:

- En primer lugar, abordaremos el problema de la *clasificación de uniones*. Para ello nos apoyaremos en la detección de puntos característicos, es decir, de puntos donde la curvatura de los contornos es alta. También son puntos característicos aquellos puntos donde convergen varias aristas. Hemos comparado dos de los métodos de detección de dichos puntos, el operador de SUSAN y el de Nitzberg. Hemos determinado en qué situaciones utilizaremos uno y otro.

Siendo las uniones estructuras geométricas que emanan de puntos característicos, definimos dichas uniones de forma paramétrica. Formulamos el problema de su clasificación en términos de la búsqueda de los parámetros que mejor se ajustan a la evidencia de la imagen. Dicha búsqueda se basa en la minimización de funciones de energía. Definimos dos métodos basados en inferencia bayesiana.

Hemos realizado experimentos para comprobar la eficacia de estos méto-

dos, presentando resultados obtenidos al aplicarlos sobre imágenes tanto de interior como de exterior. Ambos métodos presentan buenos resultados en imágenes de interior. Sin embargo, cuando procesamos imágenes de exterior, y debido a la aparición de objetos no estructurados (árboles, carteles, etc.) el elevado número de puntos candidatos obtenidos genera un gran número de uniones ficticias. También hemos observado que los métodos propuestos son muy dependientes de la localización del centro de la unión. Malas localizaciones de dicho centro provocan errores en la clasificación de los límites angulares.

Hemos analizado la eficiencia y robustez de estos métodos propuestos frente a un método reciente basado en programación dinámica, Kona, obteniéndose un error medio y un tiempo de procesamiento menor cuando aplicamos nuestros operadores. Nuestro algoritmo de segmentación lineal detecta un menor número de falsos positivos, es decir, límites que en realidad no lo son. Sin embargo, nuestro algoritmo de modelización de aristas tiene un tiempo de computación menor y deja un menor número de límites sin detectar.

En algunas situaciones se detectan falsos positivos o límites inexistentes, como por ejemplo cuando se invaden dominios de otras regiones. Este caso se puede salvar recurriendo a la prolongación de la unión para buscar otra unión en la dirección que marca el límite de la sección angular. Esta propuesta es la que desarrollaremos en el siguiente punto de la tesis.

- El objetivo de la segunda parte de la tesis es definir *estrategias de conexión de uniones*, lo cual se plantea como la búsqueda de un camino, marcado por el contraste de la imagen. Dicha búsqueda se basa en un algoritmo A^* bayesiano, que nos permite centrarnos en la extracción del camino verdadero en una población de falsos caminos, en lugar de buscar el mejor camino en una población de caminos candidatos. Esta simplificación no asegura la optimalidad del resultado, pero permite obtener un coste lineal. El coste de un camino vendrá dado por la suma de dos componentes, la componente de intensidad y la geométrica. La componente de intensidad nos da una medida de la evidencia de arista. La geométrica nos permite modelar la rigidez de los caminos a explorar. Dentro de la componente de intensidad presentamos dos modelos a seguir. En uno de ellos utilizamos un filtro no lineal basado en el número de votos obtenidos para un segmento dado. Para ello se han tomado estadísticas previas, determinando la respuesta del filtro para un determinado operador de gradiente (en nuestro caso hemos utilizado SUSAN) cuando el segmento se encuentra encima de una arista y cuando se encuentra fuera de ésta. En el otro modelo, el de modelización de aristas, las estadísticas se han obtenido de un trabajo previo y modela-

mos tanto la componente de magnitud del gradiente como su orientación. Al igual que tenemos dos términos, consideramos dos umbrales, el de intensidad y el geométrico, que sirven para modelar la evidencia de gradiente y la rigidez del camino, respectivamente. El umbral de intensidad, nos modela la cantidad de evidencia que debe tener debajo un bloque de segmentos para considerarlo válido. Si aumentamos este umbral, sólo nos quedaremos con aquellos caminos que estén encima de un gradiente con mucha magnitud. Al disminuir el umbral permitimos que caminos con poca evidencia (magnitud de gradiente) sobrevivan a la poda. En cuanto al umbral de geométrico, cuando lo disminuimos, siempre dentro de los límites teóricos establecidos, permitimos caminos poco rígidos, mientras que si aumentamos el umbral sólo permitimos caminos con una variación de ángulo entre segmentos muy pequeña.

Una vez definidos los dos modelos, presentamos resultados obtenidos al aplicarlos sobre imágenes tanto de interior como de exterior. Al igual que en los métodos de clasificación de uniones, el método propuesto presenta buenos resultados en imágenes de interior, teniendo una respuesta no tan buena en entornos no estructurados.

El método propuesto es robusto frente a posibles errores en la localización de los límites angulares. Esto es debido a que inicializamos la búsqueda con un margen lo suficientemente grande para que se subsanen estos posibles errores. También es robusto a la no localización de un límite angular, ya que, dadas dos uniones existen dos formas de conectarlas: trazar caminos en ambas direcciones.

Hemos comprobado la ganancia en calidad de las uniones detectadas al aplicar el método de agrupamiento. Comprobamos que se produce una mejora notable en cuanto al número de límites de secciones angulares falsos eliminados. También mejora el error angular cometido en el cálculo de dichos límites. De esta forma estamos introduciendo una retroalimentación de un proceso de alto nivel en otro de un nivel inferior.

- Finalmente, abordamos la obtención del *ángulo de orientación de la cámara*. Partimos de un novedoso método de estimación que consiste en un proceso de votación basado en un criterio bayesiano, utiliza información bayesiana de pertenencia de un determinado píxel a una arista. Hemos realizado experimentos para comprobar su funcionamiento, proponiendo mejoras para solventar algunos problemas. Se ha comprobado que el tiempo de ejecución es excesivo para el dominio de trabajo en el que estamos interesados, la robótica móvil. Por ello y haciendo uso de nuestro método de agrupamiento de uniones, hemos desarrollado un nuevo método que hace uso de la mis-

ma formulación bayesiana. Este método presenta una mejora sustancial en cuanto a tiempo de computación y el error medio cometido es prácticamente despreciable.

Finalizaremos esta tesis con las conclusiones y planteamientos futuros, fundamentalmente en las siguientes líneas: mejora de la localización, aplicación en tareas de segmentación, reconocimiento y estimación de profundidad, y mejora de los métodos de cálculo de orientación, fundamentalmente en su aplicación dinámica.

Abstract

This thesis focuses on feature extraction and perceptual grouping in computer vision. Our main purpose has been to formulate and test new methods for feature extraction (mainly those related to corner identification and junction classification) and grouping (through junction connection). The context of application of these methods, robot vision, imposes special constraints over their applicability and thus, the following requirements are observed: efficiency, robustness, and flexibility. These requirements are present in almost all the techniques presented in this thesis: Junction classification is performed by greedy method that relies on sound statistics; Junction grouping is performed by a path-searching method of linear complexity on average, which is based on pruning conditions to constrain the search of stable paths, and such conditions rely on statistical information; and junction grouping yields speeding up the voting scheme used to find the relative orientation between the robot and the environment. However, the range of application of the proposed method is not restricted to infer relative orientation, and can be extended to other tasks like image segmentation, depth estimation or object recognition.

This thesis is structured in three parts which cover junction classification, grouping and relative orientation, yielding a bottom-up exposition:

- *Junction Classification*: It relies on combining corner detectors and template matching. Corner detection is performed through two well known operators: SUSAN and Nitzberg. These operators provide an initial localization of the junction center. Given this localization, junction classification is posed in terms of finding the set of angular sections, or wedges, that explain as better as possible the underlying evidence in the image. We propose two greedy methods which relies on elements of Bayesian inference. These methods are tested with indoor and outdoor images in order to identify their robustness to noise and also to bad localizations of junctions centers. Experimental results show that these methods are saver than other methods recently proposed, like Kona, and their computational efficiency is even better. However, we have detected some problems due to the local extent of junction detection. These problems and those derived from bad center localization can be aliviated through local-to-global interactions, and these interactions are inferred by grouping processes.
- *Grouping*: Using classified junctions as starting elements, this stage is performed by finding connecting paths between wedge limits belonging to pairs of junctions, and these paths exist when there is sufficient contrast or edge

support below them. Given that corners are usually associated to points of high curvature in the image, it can be assumed that connecting paths must be smooth if they exist. Contrast support and smoothness are quantified by a cost function. As it can be assumed that there will be no more than one path between two junctions through a pair of wedge limits, such a path can be found by the Bayesian version of the well known A^* algorithm. This method recently proposed searches the true path, according to the cost function, in a population of false paths, instead of the best paths among a population of possible paths. Such a simplification yields a computational average cost which is linear with the length of the path. We propose two versions of the cost functions and explore the possibilities of the method for the task at hand by selecting adequate thresholds which control the penalization introduced by the lack of support or smoothness. Penalized partial paths can be discarded by a pruning rule. We also modify this rule to provide stronger pruning although the admissibility of the algorithm is not guaranteed in this case. Grouping is completed by end-path conditions, and the robustness of the process is ensured by bidirectional search. Experimental results show that the grouping stage improves individual junction identification, and in this sense grouping gives feedback to a low level task. But grouping also provides a schematic representation which is useful to higher-level tasks like computing relative orientation from a single image.

- *Relative orientation:* Grouping results followed by straight line detection are very useful to compute the relative orientation between the observer and the environment. A recent proposal, which relies on bayesian inference, has proved that this task can be performed by using a single image, at least when horizontal viewing direction is assumed. This method uses perspective and assumes that lines follow a particular configuration known as Manhattan world. We have replaced the original pixel-based strategy by the edges resulting from grouping in the preceding stage, yielding robust and fast results.

This thesis is completed by final discussion and conclusions. We have identified several issues which can be improved in the future. Some of these issues are referred to the quality of the representations, and others refer to applying our grouping strategy in other tasks like segmentation or even object recognition.

ÍNDICE GENERAL

1	Introducción	27
1.1	Visión orientada a tareas frente a visión reconstructiva	29
1.2	Descripción del contenido de la tesis	30
2	Detección y clasificación de uniones	35
2.1	Métodos de detección de puntos característicos	35
2.1.1	Puntos característicos a partir de aristas	38
2.1.2	Puntos característicos a partir de la propia imagen	39
2.2	Comparativa de detectores de puntos característicos	41
2.2.1	El operador de Nitzberg	41
2.2.2	Descripción de SUSAN	42
2.2.3	Comparación de SUSAN y Nitzberg	46
2.3	Definición de uniones mediante Kona	50
2.4	Clasificación de uniones mediante segmentación	58
2.5	Clasificación de uniones mediante modelado de aristas	67
2.6	Discusión	76
3	Agrupamiento de uniones	79
3.1	Agrupamiento de características	81
3.2	Modelado de caminos	82
3.2.1	Definición de la componente de intensidad	83
3.2.2	Definición de la componente geométrica	86
3.3	Seguimiento de caminos	89
3.3.1	Algoritmo y restricciones	89
3.3.2	Finalización del camino	91
3.4	Agrupamiento de uniones	94
3.4.1	Conexión punto a punto	101
3.5	Variación de las componentes geométrica y de intensidad	103
3.6	Clasificación de uniones	107
3.7	Discusión	107

4	Obtención del ángulo de orientación de la cámara	111
4.1	Técnicas de obtención de puntos de fuga	112
4.2	Geometría tridimensional	114
4.3	Modelo bayesiano inicial	114
4.3.1	Estadísticas de las aristas	114
4.3.2	Evidencia de un píxel	115
4.3.3	Evidencia de todos los píxeles	115
4.3.4	Resultados	116
4.3.5	Etiquetado de los píxeles	123
4.3.6	Conclusiones del método	123
4.4	Utilización de la información de arista	125
4.4.1	Obtención de caminos rectos	125
4.4.2	Estadística de las aristas	126
4.4.3	Cálculo del ángulo de orientación	127
4.4.4	Resultados	127
4.4.5	Etiquetado de los segmentos	135
4.5	Discusión	135
5	Conclusiones. Líneas de trabajo futuro	139
I	Apéndices	143
A	Inferencia bayesiana	145
A.1	Teorema de Bayes	145
A.2	Principio de longitud de descripción mínima	147
A.3	Hipótesis de máxima verosimilitud y error cuadrático medio	148
B	Búsquedas dentro de un rango	149
B.1	Búsqueda unidimensional	150
B.2	Árboles de rango	151
C	Extracción de segmentos rectos	155

ÍNDICE DE FIGURAS

1.1	Robot utilizado en algunos de los experimentos.	28
1.2	Proceso completo de aplicación de los distintos métodos propuestos.	30
1.3	Ejemplo de obtención de uniones en una imagen de gris.	31
1.4	Ejemplo de agrupamiento de uniones.	32
1.5	Cálculo del ángulo de orientación de la cámara.	33
2.1	Aplicación del filtro de Canny modificando la varianza. De izquierda a derecha y de arriba a abajo se han utilizado varianzas de 1, 2 y 3.	37
2.2	Resultado de calcular los autovalores de una imagen. Arriba derecha: autovalor mayor (aristas), debajo izquierda: autovalor menor (esquinas), debajo derecha: puntos esquina detectados aplicando una umbralización y una búsqueda de máximo locales.	42
2.3	Área SUSAN.	43
2.4	Centro de gravedad.	45
2.5	$t=10$	45
2.6	$t=20$	45
2.7	$t=30$	45
2.8	Puntos esquinas detectados por SUSAN. Arriba izquierda: $t=10$, arriba derecha: $t=20$, abajo: $t=30$	46
2.9	Resultados al ir aumentando el parámetro de geometría (izquierda-derecha-abajo).	47
2.10	Ejemplo de ejecución del operador de SUSAN variando el ruido en la imagen. Izquierda: sin ruido, central: ruido gaussiano de varianza 20, derecha: ídem varianza 50.	48
2.11	Ejemplo de ejecución del operador de Nitzberg variando el ruido en la imagen. Izquierda: sin ruido, central: ruido gaussiano de varianza 20, derecha: ídem varianza 50.	48
2.12	Comparación del error medio cometido por los operadores de SUSAN y Nitzberg.	49
2.13	Comparación del número de falsos positivos detectados por los operadores de SUSAN y Nitzberg.	49
2.14	Comparación de tiempos de ejecución frente al tamaño de la imagen de los operadores de SUSAN y Nitzberg.	50

2.15	Parámetros de la unión.	51
2.16	Valor acumulado de los radios en imágenes reales.	52
2.17	Ejemplo de cálculo de la discretización del vector angular.	54
2.18	Valor acumulado de los radios en imágenes sintéticas.	55
2.19	Efecto no deseado al utilizar un radio incorrecto.	56
2.20	Resultado de aplicar el método Kona.	56
2.21	Resultado de aplicar el método Kona.	57
2.22	Dinámica del límite de una sección angular. En este caso, $P(\mathcal{I}_{\theta_s} \mu = 150, \sigma = 0) > P(\mathcal{I}_{\theta_s} \mu = 120, \sigma = 5)$, por lo que la posición del límite se desplaza hacia la izquierda.	59
2.23	Dinámica del límite de una sección angular.	62
2.24	Algoritmo de segmentación lineal mediante competición de regiones.	63
2.25	Error producido en la discretización.	64
2.26	Resultado de aplicar el clasificador de uniones.	65
2.27	Resultado de aplicar el clasificador de uniones.	66
2.28	Arriba: Imagen original y el resultado de calcular el $\log(P_{on}/P_{off})$ en cada píxel. Abajo: Cálculo de los valores de orientación (izquierda) y magnitud (derecha).	68
2.29	Estadísticas de las aristas: P_{on} arriba, P_{off} abajo.	69
2.30	Gráfica del log de P_{on}/P_{off} (arriba) y de la función utilizada para P_{ang} (abajo).	70
2.31	Transformación de la imagen con el método de P_{on}/P_{off}	72
2.32	Algoritmo de detección de uniones mediante modelado de aristas.	73
2.33	Resultado de aplicar el método de modelado de aristas.	74
2.34	Resultado de aplicar el método de modelado de aristas.	75
3.1	Imagen original (izquierda) junto con el resultado de calcular el gradiente (derecha).	80
3.2	Estadísticas de las aristas: P_{on} y P_{off} , junto con el logaritmo del ratio entre las dos.	84
3.3	Las posibles configuraciones del filtro (izquierda), nomenclatura de los píxeles (derecha arriba) y los diez test realizados a lo largo de una determinada configuración (derecha abajo).	85
3.4	Funciones de distribución de probabilidad: P_{on} y P_{off} , junto con el logaritmo del ratio entre las dos.	87
3.5	Funciones de distribución de probabilidad: $P_{\Delta G}$ y U , junto con el logaritmo del ratio entre las dos	88
3.6	Ejemplo de árbol de búsqueda. El factor de ramificación Q es tres.	89
3.7	Situaciones a detectar cuando nos quedamos sin caminos.	92
3.8	Ejemplo de inicialización, búsqueda y finalización de camino.	93

Índice de Figuras

23

3.9 Algoritmo de búsqueda de camino.	95
3.10 Generación de caminos.	96
3.11 Algoritmo de agrupamiento.	96
3.12 Ejemplos de aplicación del algoritmo de agrupamiento (componente de intensidad: filtro no lineal).	97
3.13 Ejemplos de aplicación del algoritmo de agrupamiento (componente de intensidad: filtro de modelado de aristas).	98
3.14 Ejemplos de aplicación del algoritmo de agrupamiento (componente de intensidad: filtro de modelado de aristas).	99
3.15 Ejemplos de aplicación del algoritmo de agrupamiento (componente de intensidad: filtro de modelado de aristas).	100
3.16 Aplicación del método de agrupamiento exhaustivo.	102
3.17 Imagen utilizada en los experimentos de variación de los umbrales de corte.	103
3.18 Variación del umbral de corte de la componente de intensidad.	104
3.19 Variación del umbral de corte de la componente de intensidad y geométrica.	105
3.20 Variación del umbral de corte de la componente geométrica.	106
3.21 Mejora producida al introducir el agrupamiento de uniones. Las líneas azules son las uniones detectadas y las verdes son los caminos del agrupamiento. En la imagen de la izquierda podemos observar como algunas de las uniones tienen límites que no se corresponden con regiones reales. El proceso de agrupamiento elimina estos límites. En la imagen de la derecha (centro) una unión con forma de T termina clasificándose como una L. En la imagen de la derecha (arriba) la unión original es eliminada al no tener ningún límite soporte de arista.	109
4.1 Sistema de coordenadas en el mundo de Manhattan.	112
4.2 Esfera de acumulación para el cálculo del punto de fuga.	113
4.3 Algoritmo de obtención del ángulo de orientación de la cámara.	116
4.4 Gráficas del MAP frente al ángulo de orientación. Estas gráficas se corresponden con las Figuras 4.5 y 4.9.	118
4.5 Resultado obtenido en imagen de exterior (1). El ángulo de orientación de la cámara obtenido fue de -18°	119
4.6 Resultado obtenido en imagen de exterior (2). El ángulo de orientación de la cámara obtenido fue de 32°	119
4.7 Resultado obtenido en imagen de exterior (3). El ángulo de orientación de la cámara obtenido fue de 20°	120

4.8	Resultado obtenido en imagen de exterior (4). El ángulo de orientación de la cámara obtenido fue de 8° .	120
4.9	Resultado obtenido en imagen de interior (1). El ángulo de orientación de la cámara obtenido fue de 37° .	121
4.10	Resultado obtenido en imagen de interior (2). El ángulo de orientación de la cámara obtenido fue de -15° .	121
4.11	Resultado obtenido en imagen de interior (3). El ángulo de orientación de la cámara obtenido fue de -13° .	122
4.12	Resultado obtenido en imagen de interior (4). El ángulo de orientación de la cámara obtenido fue de 10° .	122
4.13	Puntos etiquetados con modelo 4 (derecha) y magnitud del gradiente (izquierda).	123
4.14	Puntos etiquetados con modelo 4 (derecha) y magnitud del gradiente (izquierda).	124
4.15	Resultado erróneo debido a la mala calidad de la imagen obtenida.	125
4.16	Resultado erróneo ($\Psi = -15^\circ$) utilizando la suma de la evidencia para cada modelo.	127
4.17	Resultado correcto ($\Psi = 40^\circ$) utilizando el máximo valor de evidencia.	127
4.18	Función log posteriori. Izquierda: utilizando el máximo de la función. Derecha: utilizando la suma.	128
4.19	Algoritmo de obtención del ángulo de orientación de la cámara utilizando información de arista.	128
4.20	Resultado obtenido en imagen de exterior. El ángulo de orientación de la cámara obtenido fue de -15° .	129
4.21	Resultados obtenidos en imágenes de exterior. El ángulo de orientación de la cámara obtenido fue de 22° .	130
4.22	Resultados obtenidos en imágenes de exterior. El ángulo de orientación de la cámara obtenido fue de 9° .	131
4.23	Resultados obtenidos en imágenes de interior. El ángulo de orientación de la cámara obtenido fue de 34° .	132
4.24	Resultados obtenidos en imágenes de interior. El ángulo de orientación de la cámara obtenido fue de -13° .	133
4.25	Resultado obtenido en imagen de interior. El ángulo de orientación de la cámara obtenido fue de 13° .	134
4.26	Separación de aristas dependiendo de si están alineadas con el mundo de Manhattan o no.	136
4.27	Separación de aristas dependiendo de si están alineadas con el mundo de Manhattan o no.	137
B.1	Algoritmo de creación de un árbol binario.	150

Índice de Figuras

25

B.2	Árbol binario creado indicando el nodo de partición.	151
B.3	Obtención de los puntos dentro de un rango.	152
B.4	Árbol de rango con el árbol binario asociado.	153
B.5	Algoritmo de consulta de un árbol binario.	154
C.1	Ejemplo de obtención de los segmentos rectos a partir del resultado del agrupamiento.	156

