VALQUIRIA FENELON PEREIRA

Interpretação de Imagens com Raciocínio Espacial Qualitativo Probabilístico

São Paulo
2014

VALQUIRIA FENELON PEREIRA

Interpretação de Imagens com Raciocínio Espacial Qualitativo Probabilístico

Tese apresentada à Escola Politécnica da
Universidade de São Paulo para obtenção do
Título de Doutor em Engenharia.

São Paulo
2014

VALQUIRIA FENELON PEREIRA

Interpretação de Imagens com Raciocínio Espacial Qualitativo Probabilístico

Tese apresentada à Escola Politécnica da Universidade de São Paulo para obtenção do Título de Doutor em Engenharia.

Área de Concentração: Engenharia de Controle e Automação Mecânica

Orientador: Prof. Dr. Fábio Gagliardi Cozman

São Paulo
2014

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 23 de abril de 2014.

Assinatura do autor _____

Assinatura do orientador _____

Catálogo-na-publicação

Pereira, Valquiria Fenelon
Interpretação de imagens com raciocínio espacial
qualitativo
probabilístico / V.F. Pereira. -- versão corr. -- São Paulo, 2014.
174 p.

Tese (Doutorado) - Escola Politécnica da Universidade de
São Paulo. Departamento de Engenharia Mecatrônica e de
Sistemas Mecânicos.

1.Robótica 2.Interpretação de imagens 3.Raciocínio espacial
qualitativo probabilístico I.Universidade de São Paulo. Escola
Politécnica. Departamento de Engenharia Mecatrônica e de Sis-
temas Mecânicos II.t.

*A todos que me incentivaram a seguir
os meus sonhos. E ao universo que me
pôs nessa direção. E aos leitores que
foram escolhidos ou escolheram ler
essas páginas.*

AGRADECIMENTOS

Antes de começar a lista de agradecimentos, quero dizer que experimentei sensações extremas, acreditei e duvidei, sorri e chorei, me perdi e me encontrei, mas não estive só, sempre tive alguém que me desse apoio, ou diretamente com esse doutorado, ou indiretamente me ajudando a manter minha saúde mental. Espero que me perdoem aqueles que por falha da minha memória não forem citados aqui, por favor, saibam que em meu coração sou grata a todos que estiveram comigo durante essa jornada.

Ao meu orientador Prof Dr Fábio Cozman com sua paciência em tentar me ouvir e em me orientar em momentos cruciais e me indicar quais os caminhos prováveis a seguir onde a expectativa era maximar o sucesso e as descobertas. Ao Prof Dr Paulo Eduardo Santos que foi meu orientador na iniciação científica e no mestrado. Agradeço muito todos os conselhos, bate-papos, seu tempo e dedicação para me orientar nestes dez anos.

Aos meus amigos e colegas que me apoiaram, mas em especial ao meu amigo Murilo que me deu a oportunidade de passar alguns dias em sua casa e conseguiu uma mesa no Imperial College para eu trabalhar e preparar minha primeira apresentação de workshop, além de me ajudar na realização do último experimento e toda a ajuda entre esses eventos. E ao Angelo que me ajudou várias vezes com dicas e conversas que desataram alguns nós de programação. Aos colegas do laboratório de tomada de decisão ao Ochoa e ao Polastro que me ajudaram na primeira etapa desse trabalho. Ao pessoal do projeto Fapesp em especial a Anna e Renata que dispensaram algum tempo e ideias para soluções de alguns problemas encontrados no caminho.

À minha família, mãe e pai pelo apoio e incentivo para chegar até aqui e continuar seguindo em frente. Ao meu filho Édipo que algumas vezes segurou as pontas para que eu seguisse trabalhando sem culpa. Agradeço também ao Danisio que me ajudou a redescobrir a força para continuar, já que na hora que eu mais precisei ele reapareceu e mostrou a importância de realizar os sonhos. E tia Jô que fez perguntas sobre meu trabalho que nem sempre foram simples de responder.

À CAPES, pelo apoio financeiro. À Fapesp pelo financiamento do projeto temático ao qual meu trabalho esteve inserido. A instituição FEI e seus funcionários pelo apoio com material e infraestrutura necessários para realização dos testes, e a utilização do laboratório e do robô *PeopleBotTM*. À EPUSP por disponibilizar auxílio de custo para participar de conferências.

I want to give to Britta and Hannah my special thanks for all data and ideas that you have provided directly or indirectly for me in this work. And for all colleagues on Imperial College who received me in your laboratory, in special Pedro and Julia for help me on my first day on Edinburgh and for all folks who helped me with my first international presentation.

“ O senhor mire e veja. O mais importante e bonito, do mundo, é isto: que as pessoas não estão sempre iguais, ainda não foram terminadas – mas que elas vão sempre mudando. Afinam e desafinam. Verdade maior. O senhor saiba: eu toda a minha vida pensei por mim, forro, sou nascido diferente. Eu sou é eu mesmo. Divirjo de todo o mundo... Natureza da gente não cabe em nenhuma certeza. Quem sabe direito o que uma pessoa é? Antes sendo: julgamento é sempre defeituoso, porque o que a gente julga é o passado. Que isso foi o que sempre me invocou, o senhor sabe: eu careço de que o bom seja bom e o ruim, que dum lado esteja o preto e do outro o branco, que o feio fique bem apartado do bonito e a alegria longe da tristeza! Quero os todos pastos demarcados... Como é que posso com este mundo? A vida é ingrata no macio de si; mas transtraz a esperança mesmo do meio do fel do desespero. Ao que, este mundo é muito misturado...Se não, o senhor me diga: preto é preto? branco é branco? Ou: quando é que a velhice começa, surgindo de dentro da mocidade? ”

Grande Sertões Vereda, Guimarães Rosa

RESUMO

Um sistema artificial pode usar raciocínio espacial qualitativo para inferir informações sobre seu ambiente tridimensional a partir de imagens bidimensionais. Inferências realizadas com base em raciocínio espacial qualitativo devem ser capazes de lidar com incertezas. Neste trabalho investigamos a utilização de técnicas probabilísticas para tornar o raciocínio espacial qualitativo mais robusto a incertezas e aplicável a agentes móveis em ambientes reais. Investigamos uma formalização de raciocínio espacial com lógica de descrição probabilística em um subdomínio de tráfego. Desenvolvemos também um método que combina raciocínio espacial qualitativo com um filtro Bayesiano para desenvolver dois sistemas que foram aplicados na auto localização de um robô móvel. Executamos dois experimentos de auto localização; um utilizando a teoria de relações qualitativas percebíveis sobre sombra com filtro Bayesiano; e outro utilizando o cálculo de oclusão de regiões e o cálculo de direção com filtro Bayesiano. Ambos os sistemas obtiveram resultados positivos onde somente o raciocínio espacial qualitativo não foi capaz de inferir a localização do robô. Os experimentos com dados reais mostraram robustez aos ruídos e à informação parcial.

Palavras-chaves: Raciocínio Espacial Qualitativo, Filtro Bayesiano, Lógica Probabilística, Robótica Cognitiva, Auto Localização Qualitativa.

ABSTRACT

An artificial system can use qualitative spatial reasoning to obtain information about its tridimensional environment, from bi-dimensional images. Inferences produced by qualitative spatial reasoning must be able to deal with uncertainty. This work investigates the use of probabilistic techniques to make qualitative spatial reasoning more robust against uncertainty, and better applicable to mobile agents in real environments. The work investigates a formalization of spatial reasoning using probabilistic description logics in a traffic domain. Additionally, a method is presented that combines qualitative spatial reasoning with a Bayesian filter, to develop two systems that are applied to self-localization of mobile robots. Two experiments are described; one using the theory of perceptual qualitative relations about shadows; the other using occlusion calculus and direction calculus. Both systems are combined with a Bayesian filter producing positive results in situations where qualitative spatial reasoning alone cannot infer robot location. Experiments with real data show robustness to noise and partial information.

Key-words: Qualitative Spatial Reasoning, Bayesian Filter, Probabilistic Logic, Cognitive Robotics, Qualitative Self Localisation.

LISTA DE FIGURAS

Figura 1	– Diagrama conceitual de vizinhança (CND - <i>Conceptual Neighborhood Diagram</i>) das relações RCC-8. Note a continuidade das transições topológicas entre regiões (indicadas pelas setas), as mudanças de relações ocorrem sempre para uma relação vizinha.	39
Figura 2	– Representação gráfica das composições $R_1(a, b)$ e $R_2(b, c)$ antes e depois da transição representada pela seta, note que o conjunto resultante $R_3(a, c)$ são as relações possíveis (Tabela 2).	41
Figura 3	– Duas sequências de imagens (da esquerda para direita) onde um cone passa em frente (sequência superior) e por trás (sequência inferior) do cilindro, este é um exemplo da oclusão em ação. Note na ocorrência de oclusão os dois objetos formam uma única região na imagem.	42
Figura 4	– Mudança na percepção da oclusão causada pelo movimento do observador (paralaxe).	43
Figura 5	– Modelo gráfico das 20 relações básicas do cálculo sobre oclusão de região, o ROC-20.	44
Figura 6	– Estudos sobre sombra de Leonardo da Vinci. No esboço vê-se representados as sombras do objeto projetadas em outra superfície (sombra-projetada) e no próprio objeto (sombra-própria).	46
Figura 7	– Efeito na localização do objeto ocasionado pela posição da sombra.	47
Figura 8	– Bola na caixa. Mesmo na mesma localização da bola em (b) e (c) percebemos posicionamentos diferentes devido à sombra.	48
Figura 9	– Representação gráfica das relações básicas do PQRS.	49
Figura 10	– Imagens capturadas pelo <i>PeopleBotTM</i> . Representando as relações PQRS (a) NonOccludesDC(o,s), (b) NonOccludesEC(o,s), (c) PartiallyOccludesPO(o,s), (d) PartiallyOccludesTPP(o,s), (e) TotallyOccludesTPPI(o,s), (f) TotallyOccludesEQ(o,s) e (g) TotallyOccludesNTPPI(o,s).	50
Figura 11	– (a) Um objeto com sua sombra conectada e projetada numa superfície irregular, destaque para a parte da sombra relativa a parte superior do objeto; (b) Regiões distintas expressas pela observação do objeto e sua sombra, definidas pelas linhas de visão.	52
Figura 12	– Como a sombra aparece quando somente o observador altera sua posição para: (a) região 1, (b) região 2, (c) região 3, (d) região 4 e (e) região 5 representadas no mapa da Figura 11b	54
Figura 13	– Imagens capturadas no laboratório, onde podemos aplicar a fórmula de correspondência de sombra, note somente a sombra do objeto está externamente conectada à sua base e mais distante da fonte de luz.	57
Figura 14	– Por causa das sombras podemos inferir que há outro objeto totalmente ocluído pelo balde.	57
Figura 15	– Representação da projeção em quadrante, sendo os elementos representados como ponto (direito) e região (esquerda).	58
Figura 16	– Exemplo de representação CRALC em rede Bayesiana da inclusão e da definição.	67
Figura 17	– Exemplo de redes e tabelas de probabilidades condicionais geradas pelos construtores lógicos	67
Figura 18	– Exemplo de representação CRALC em rede Bayesiana da restrição de existencial.	68
Figura 19	– Exemplo de representação CRALC em rede Bayesiana da restrição universal.	68

Figura 20 – Exemplo gráfico de uma rede gerada a partir de uma terminologia τ , e sua característica para um domínio com dois indivíduos $[a, b]$	69
Figura 21 – Representação gráfica do cubo de cores RGB, o cone da esquerda representa como uma cor(80,200,130) está localizada neste sistema cartesiano.	73
Figura 22 – Representação gráfica do cone HSV.	73
Figura 23 – Gráfico representando o filtro de limiar e o limiar T	74
Figura 24 – Representação gráfica da dilatação e da erosão com uma matriz 3x3 com elemento estruturante E	75
Figura 25 – Exemplo da aplicação do componente conectado, a partir da imagem original (esquerda) um filtro limiar e morfológico aplicado (meio) o resultado na imagem com grupos de mesmo valor (representado pela cor).	76
Figura 26 – Representação gráfica do cenário de tráfego (mão francesa). (a) Representação da rua, faixa, e linhas divisórias. (b) Representação do veículo no cenário.	80
Figura 27 – (a) Cena Original. (b) Exemplo dos dados dos sensores para o cenário da cena original.	81
Figura 28 – Um sistema cardeal para vias: A está ao sul de B e oeste de C	82
Figura 29 – Domínio de tráfego e todas as situações possíveis para o veículo que esteja dentro das regras.	86
Figura 30 – Rede Bayesiana formada com conceitos e papéis embutidos no banco de conhecimento. Todos os nós em azul são evidências observadas pelos sensores.	89
Figura 31 – Exemplo de três cenários de tráfego, onde o carro está descendo uma via de mão única.	91
Figura 32 – Rede Bayesiana relacional representando um domínio de tráfego com papéis e indivíduos representando veículo, faixas, e linha divisória.	92
Figura 33 – Rede Bayesiana baseado na tabela de composição (Tabela 8)	93
Figura 34 – Representação espacial (a) das regiões considerando um ambiente real, e (b) representação do mapa qualitativo com a cinco regiões distintas, mostrando os lados direito e esquerdo e as regiões($Region_1$ até $Region_5$). As linhas de visão entre a fonte de luz L , o objeto O e sua sombra definem as fronteiras entre as regiões. Neste diagrama somente o topo da sombra é utilizado na definição das regiões.	101
Figura 35 – Uma representação da posição dos objetos no contexto deste experimento.	101
Figura 36 – Uma representação gráfica para a definição das direções <i>Left</i> , <i>Right</i> , <i>Front</i> e <i>Behind</i> (esquerda, direita, em frente, atrás) relativos à visão do robô através da imagem. Os eixos de coordenadas da imagem tem origem no canto acima-direito relativo à imagem.	103
Figura 37 – Diagrama representando nove regiões formadas pelas linhas de visão.	103
Figura 38 – As duas figuras representam as regiões e as relações espaciais percebidas em cada uma (a)As relações de oclusão e (b) as relações de direção para cada região, todas as relações percebidas por um observador ν . Somente a região $Region_9$ não é possível visualizar o par em nenhuma posição do observador.	104

Figura 39	– Exemplo de uma visualização para (a) $Region_1$, (b) $Region_2$, (c) $Region_3$, (d) $Region_4$, (e) $Region_5$, (f) $Region_6$, (g) $Region_7$, (h) $Region_8$, (i e j) $Region_9$ onde os objetos estão livres de oclusão, mas o campo de visão não permite visualizar os dois em uma mesma cena.	106
Figura 40	– Representações dos mapas qualitativos formados pelas linha de visão com duas configurações de posição que os três objetos podem assumir neste contexto.	107
Figura 41	– Imagens capturadas com o robô posicionado em distintos pontos na região 1. A imagem (c) mostra somente o topo da sombra em destaque. Note que o objeto não oclui a sombra.	108
Figura 42	– Imagens capturadas com o robô na região 2. Note que o objeto começa a ocluir a sombra. Todavia ainda se vê a sombra do topo do objetivo e as laterais da sombra do objeto indicados em (a).	109
Figura 43	– Imagens capturadas com o robô na região 3. A sombra nesta região aparece com um dos lados totalmente ocluido pelo objeto. Existe uma conexão entre a parte superior do objeto e sua sombra projetada.	109
Figura 44	– Imagens capturadas com o robô na região 4. A sombra nesta região quase desaparece ocluida pelo objeto. Contudo ainda existe uma conexão entre a parte superior do objeto e a sombra.	110
Figura 45	– Imagem capturada com o robô na região 5. A sombra do objeto não aparece na imagem por estar totalmente ocluida. O observador está na região equivalente ao ponto de vista da luz.	111
Figura 46	– Mapa qualitativo da disposição espacial dos objetos e a representação das regiões.	112
Figura 47	– Mapa qualitativo das áreas de atuação divididas em oito regiões.	117
Figura 48	– Diagrama conceitual de vizinhança (CND) das relações que determinam a região na qual o robô estava localizado. As transições estão representadas pelas setas e a mudança ocorre sempre para uma relação (região) vizinha.	118
Figura 49	– Representação das regiões como nós, e das transições entre regiões próximas como setas: setas contínuas regiões com fronteira, tracejadas regiões sem fronteira entre elas.	119
Figura 50	– Janelas para calibração manual dos parâmetros de visão utilizados no filtro de limiar para segmentar objeto e sombra, os valores escolhidos pela barra de rolagem ficam indicados à esquerda da barra. Da esquerda para direita da figura: define o limiar Th que separa objeto e sombra do ambiente; define o limiar que separa objeto da sombra; define o limiar que separa a sombra do objeto.	120
Figura 51	– Exemplo de sequência da segmentação de um quadro na detecção ROI: (a) canal V , (b) canal V após limiar acima α , (c) filtro morfológico, (d) retângulo verde indica regiões candidatas a região de interesse (ROI), (e) canal S após filtro passa-alta α_{sat} , (f) máscara ROI.	126
Figura 52	– Exemplo de sequência da segmentação de um quadro na detecção de objetos: (a) Imagem objeto e sombra depois de aplicado a máscara, (b) filtro de limiar abaixo e filtro morfológico, (c) imagem binária da base do objeto, (d) retângulo azul indica objetos encontrados na segmentação, (e) última máscara ROI criada pelo sistema de visão.	128
Figura 53	– Exemplo de sequência da segmentação de um quadro na detecção de sombras: (a) Imagem com objeto depois de aplicada a segunda máscara, (b) Imagem após filtros de limiar acima e morfológico de abertura, (c) retângulo limitador vermelho indicando a sombra localizada.	129

Figura 54 – Exemplo de segmentação do objeto e sua sombra. O topo da sombra também é detectado pelo sistema de visão conforme destacado nesta imagem.	130
Figura 55 – Exemplo de sequência da segmentação de um quadro na detecção ROI:(a) canal V , (b) canal V após limiar acima α , (c) filtro morfológico, (d) retângulo verde indica regiões candidatas a região de interesse (ROI), (e) canal S após filtro passa-alta α_{sat} , (f) máscara ROI.	132
Figura 56 – Imagem exemplo de busca de regiões externamente conectadas, para cada sombra todas as bases de objetos são verificadas onde as bordas se encontram intensidade aumenta e EC é verdade.	134
Figura 57 – Imagem exemplo para definição de relações topológica de oclusão.	138
Figura 58 – O ambiente utilizado para a coleta de imagens com sombra.	140
Figura 59 – Imagens em que o objeto e a sombra não estão completamente dentro do campo de visão do robô. O sistema somente com raciocínio espacial não poderia inferir uma localização a partir dessas imagens.	142
Figura 60 – Ambiente utilizado para captura e auto-localização com múltiplos objetos.	142
Figura 61 – Imagens capturadas com três pares, um par e nenhum par em cena	143
Figura 62 – Representação de um robô na região 3.	145

LISTA DE TABELAS

Tabela 1	– Definições e interpretações de algumas relações do RCC.	39
Tabela 2	– Tabela de composição para o RCC-8 com o conjunto das relações exclusivamente disjuntas possíveis $R_3(a, c)$ entre a e c a partir das relações $R_1(a, b)$ e $R_2(b, c)$	41
Tabela 3	– Matriz de confusão do sensor de captura de imagem em que colunas representam o tipo de linha divisória real enquanto que as linhas representam como a linha divisória foi classificada pelo sistema.	81
Tabela 4	– Matriz de confusão dos valores de probabilidade de resposta do sensor mapa digital em que colunas representam o tipo de rua enquanto que as linhas representam como a rua foi classificada pelo sistema.	82
Tabela 5	– Tabela com oito estados dos sensores indicando as dozes possibilidades de estado do veículo conforme representado na Figura 29. Com as coordenadas mudando relativas ao sentido de deslocamento do carro.	88
Tabela 6	– Resposta da rede Bayesiana representada na Figura 30 para a primeira questão: “qual faixa tem maior probabilidade do carro estar localizado”. As evidências referentes aos dados dos três sensores (A) estão expressas nas três primeiras colunas.	90
Tabela 7	– Resposta da pergunta dois a partir da rede Bayesiana da Figura 30: a probabilidade para a direção de a faixa ser <i>Down</i> dado às evidências A (expressas nas seis primeiras colunas).	91
Tabela 8	– Tabela de composição do cdc para o domínio de tráfego	93
Tabela 9	– Tabela que apresenta à quantidade de vezes que uma faixa pode ocorrer dada a combinação das características que podem ser percebidas pelos sensores.	94
Tabela 10	– Resposta da primeira questão: "qual faixa tem maior probabilidade do carro estar localizado?" As evidências (A) estão expressas nas seis primeiras colunas referentes aos dados dos três sensores.	94
Tabela 11	– Resposta da pergunta dois a partir da rede Bayesiana da Figura 32: a probabilidade para a direção de a faixa ser <i>Down</i> dado às evidências A (expressas nas seis primeiras colunas).	95
Tabela 12	– O modelo de movimento considerando que o robô tem uma navegação no sentido horário. Este modelo é baseado no tamanho qualitativo e na transição das regiões. O valores são alterados por uma função de mudança percebida w que representa a paralaxe causada pelo movimento do robô	119
Tabela 13	– Tabela com os valores que diferenciam as características em cada relação.	122
Tabela 14	– O resultado do experimento de auto localização utilizando um filtro Bayesiano e o Raciocínio Qualitativo Perceptual sobre Sombras.	146
Tabela 15	– Valores de acurácia do algoritmo baseado em conhecimento e o com o filtro Bayesiana qualitativo, onde '# imagens' é o número de imagens capturadas em cada região.	147

Tabela 16	– Matriz de confusão com os resultados do experimento de auto localização utilizando REQ probabilístico. As linhas são as regiões onde a imagem foi capturada e as colunas a resposta do sistema.	149
Tabela 17	– Tabela com os valores de acurácia do sistema REQ probabilístico.	150
Tabela 18	– Tabela com os valores de acurácia do sistema detalhado por região.	151
Tabela 19	– Tabela com os resultados de acurácia para o caso específico das regiões onde ocorre oclusão nas quais somente um objeto aparecem na imagem.	151
Tabela 20	– Resposta de total de imagens com resposta positiva para o REQ e o REQ probabilístico com o REQ probabilístico considerando as regiões vizinhas que estavam no percurso do robô como resposta positiva	152
Tabela 21	– Comparação que tem os valores de acurácia pelo total de objetos contidos na imagem, considerando a região vizinha que estava no percurso executado pelo robô como resposta positiva.	152

LISTA DE ABREVIATURAS E SIGLAS

<i>ARIA</i>	<i>ActivMedia Robotics Interface for Applications</i> ou Interface robótica <i>Activ-Media</i> para aplicações.
<i>CND</i>	<i>Conceptual Neighbourhood Diagram</i> ou diagrama conceitual de vizinhança.
<i>CDC</i>	Cálculo de Direção Cardeal.
<i>CRA$\mathcal{L}\mathcal{C}$</i>	Credal $\mathcal{A}\mathcal{L}\mathcal{C}$.
<i>DPC</i>	Distribuição de probabilidade condicional.
<i>HSV</i>	<i>Hue, Saturation, Value</i> ou espaço de cores com três canais, matiz, saturação e valor.
<i>OpenCV</i>	<i>Intel Open Source Computer Vision. Library</i> ou biblioteca de visão computacional com código aberto.
<i>PQRS</i>	<i>Perceptual Qualitative Relations about Shadows</i> ou relações perceptuais qualitativas sobre sombras.
<i>QSR</i>	<i>Qualitative Spatial Reasoning</i> ou REQ – Raciocínio Espacial Qualitativo.
<i>RCC</i>	<i>Region Connection Calculus</i> ou cálculo sobre conexão de região.
<i>REQ</i>	Raciocínio Espacial Qualitativo.
<i>RGB</i>	<i>Red, Green, Blue</i> ou espaço de cores com três canais, vermelho, verde e azul.
<i>ROC</i>	<i>Region Occlusion Calculus</i> ou cálculo sobre oclusão de região.
<i>ROI</i>	<i>Region Of Interest</i> ou região de interesse (regiões em uma imagem).
<i>SURF</i>	<i>Speeded Up Robust Features</i> .

LISTA DE SÍMBOLOS

Φ	Relações espaciais tridimensionais de oclusão.
Ξ	Relações espaciais tridimensionais de oclusão.
Ψ	Relações espaciais bidimensionais.
L	Constante que representa a fonte de luz.
Scr	Constante que representa a superfície onde uma sombra é projetada.
ν	Representa o observador.
s	Sombra.
s_k	Luminância dos pontos na imagem do instante k.
o	Objeto físicos.
t	Instante de tempo.
r	Regiões.
D_k	Diferença da luminância de uma imagem do instante k.

SUMÁRIO

1	INTRODUÇÃO	29
1.1	Trabalhos Relacionados	34
1.2	Organização	35
2	FUNDAMENTOS	37
2.1	Robótica Cognitiva	37
2.2	Raciocínio Espacial Qualitativo	38
2.2.1	Cálculo sobre conexão de regiões	38
2.2.2	Cálculo sobre oclusão de região	42
2.2.3	Relações qualitativas percebíveis sobre sombras	46
2.2.3.1	<i>Localização relativa</i>	51
2.2.3.2	<i>Correspondência entre objeto e sombra</i>	53
2.2.4	Cálculo de direções cardinais	56
2.3	Modelos Probabilísticos	58
2.3.1	Filtros Bayesianos	59
2.3.2	Redes Bayesianas relacionais	61
2.4	Lógica de Descrição Probabilística Credal $\mathcal{ALC}$	62
2.4.1	Lógica de descrição	62
2.4.2	Lógica de descrição probabilística	63
2.4.3	Credal $\mathcal{ALC}$ - $CR\mathcal{ALC}$	65
2.5	Localização Robótica	69
2.6	Interpretação de Imagens	70
2.7	Visão Computacional	72
2.7.1	Espaço de cores em imagens digitais	72
2.7.2	Filtro de limiar	74
2.7.3	Filtros morfológicos	74
2.7.4	Detectores de borda	75
2.7.5	Componentes conectados	76
I	Primeira Contribuição	77
3	RACIOCÍNIO ESPACIAL QUALITATIVO COM LÓGICA DE DESCRIÇÃO PROBABILÍSTICA	79
3.1	Descrição do Cenário de Tráfego	79
3.2	Definindo o Cenário de Tráfego em $CR\mathcal{ALC}$	82

3.3	Estratégias para Solução da Tarefa	87
3.3.1	Inferência com uma rede Bayesiana	87
3.3.2	Rede Bayesiana relacional com transitividade	92
3.4	Discussão dos Resultados	95

II Segunda Contribuição 97

4 RACIOCÍNIO ESPACIAL QUALITATIVO PROBABILÍSTICO PARA LOCALIZAÇÃO ROBÓTICA 99

4.1	Um Breve Histórico	99
4.2	Mapas Qualitativos	100
4.3	Definindo a Localização do Robô	107
4.3.1	Localização na região 1	107
4.3.2	Localização na região 2	108
4.3.3	Localização na região 3	110
4.3.4	Localização na região 4	110
4.3.5	Localização na região 5	111
4.3.6	Localização nas regiões R_1 até R_{28}	111
4.4	Filtro Bayesiano	115
4.4.1	Filtro Bayesiano sobre informações qualitativas do PQRS	115
4.4.2	Filtro Bayesiano para auto localização com múltiplos objetos	116
4.4.3	Modelo de imagem	117
4.4.4	Modelo de transição	117
4.5	Algoritmos Utilizados no Sistema de Auto Localização	120
4.5.1	Algoritmo de auto localização e auto calibração com PQRS probabilístico	120
4.5.2	Algoritmo de auto localização REQ probabilístico	123
4.5.3	Algoritmo detecção de sombra e objeto.	124
4.5.4	Algoritmo de detecção de objetos	130
4.5.5	Algoritmo de análise de correspondência	131
4.5.6	Algoritmo de reconhecimento de relações espaciais	135

5 EXPERIMENTOS EM RACIOCÍNIO ESPACIAL QUALITATIVO PROBABILÍSTICO PARA AUTO LOCALIZAÇÃO DO ROBÔ MÓVEL 139

5.1	Descrição dos Ambientes de Testes	139
5.1.1	Descrição do ambiente de teste com objeto-sombra	139
5.1.2	Descrição do ambiente de teste com múltiplos objetos	140
5.1.3	Procedimentos de coleta de imagens com sombra-objeto	140
5.1.4	Procedimento de coleta de imagens com múltiplos objetos	141
5.1.5	Modelo de imagens para sombra-objeto	143
5.1.6	Modelo de imagem com múltiplos objetos	144

5.1.7	Teste com o sistema de auto localização	144
5.2	Método de Avaliação	145
5.3	Resultados para Auto Calibração e Auto Localização com Sombras	145
5.4	Resultados para Auto Localização com Múltiplos Objetos	148
5.5	Discussão	153
6	CONCLUSÃO	157
	REFERÊNCIAS	161
	GLOSSÁRIO	171
A	TRABALHOS PUBLICADOS	173

1 INTRODUÇÃO

Não seria bom se nossos carros fossem capazes de funcionar de modo seguro no piloto automático, tornando acidentes de carro em algo do passado? Ou se robôs domésticos fossem assistentes inteligentes para tomar conta de reparos e cuidassem das tarefas de manutenção e limpeza da casa? Ou fossem capazes de auxiliar no cuidado de pessoas com necessidades especiais? Para fazer essas tarefas, um robô deve ser capaz de raciocinar sobre o mundo em que está inserido utilizando um conhecimento *a priori* e dados percebidos através de sensores (THRUN et al., 2005). A representação de conhecimento é uma área de Inteligência Artificial que trata do modo de representar e entender informações em computadores, para que os programas possam:

- ser utilizados na obtenção de informações que estão implícitas a partir das informações presentes;
- conversar com pessoas em linguagem natural;
- planejar atividades;
- resolver problemas em áreas em que o conhecimento humano é normalmente requerido (SHAPIRO, 2003).

Pesquisas em representação de conhecimento começaram por volta de 1958 com a investigação de como um computador poderia ser capaz de representar e utilizar um tipo de conhecimento de senso comum que temos quando decidimos como ir ao aeroporto por exemplo. Nos anos seguintes muitas pesquisas investigaram o uso das informações dos processos de linguagem natural. Nos anos 70' e 80's houve um interesse na pesquisa sobre o conhecimento em domínios específicos que era executado por humanos *experts*, tais como diagnóstico médico e identificação de componentes químicos, criando a área de pesquisa dos sistemas especialistas. Nos anos 80's e 90's começam pesquisar em detalhes alguns subdomínios específicos do conhecimento do dia a dia, tais como as teorias de tempo e espaço e em estruturas gerais do nosso conhecimento, levando a construção de ontologias com propósitos gerais de raciocínio espacial (SHAPIRO, 2003).

Raciocínio espacial qualitativo é um subcampo de pesquisa de representação de conhecimento que trata com o conhecimento sobre um domínio espacial infinito utilizando um conjunto finito de relações qualitativas. Um dos objetivos em particular é modelar o senso comum humano para entender o espaço (WOLTER; WALLGRÜN, 2012). Em nosso dia a dia a representação e raciocínio espacial tem um papel essencial. Mesmo que a abordagem quantitativa possa providenciar informações mais precisas, informações numéricas em geral são

desnecessárias ou indisponíveis para o ser humano (CHEN et al., 2013). O Raciocínio Espacial Qualitativo (REQ) procura formalizar o conhecimento espacial, baseado em relações primitivas definidas sobre entidades espaciais elementares. As teorias de raciocínio espacial utilizam muitos modelos espaciais que, com primitivas ontológicas e métodos de raciocínio, predizem, explicam e diagnosticam comportamentos de um sistema físico de modo qualitativo, sempre que uma avaliação quantitativa for intratável ou não avaliável (STOCK, 1997).

Em particular as pesquisas em REQ têm aplicações em áreas tais como: sistema de informação geográfica (FELICE; FOGLIARONI; WALLGRÜN, 2011), questões de linguagem espacial (SCHULZ et al., 2011), linguagem natural (HERMAN, 1999), entre outros campos. Em Cohn e Hazarika (2001) e Chen et al. (2013) encontra-se um apanhado mais detalhado desta área. Vários formalismos de relações espaciais qualitativas têm sido propostos para descrever diversos aspectos espaciais. Nesse trabalho nos concentraremos somente nos cálculos espaciais qualitativos topológicos representados pelo cálculo de conexão de regiões e sua extensão o cálculo de oclusão de regiões que representam as relações entre duas regiões de modo bidimensional e tridimensional. E os cálculos sobre relações de direções como o cálculo de direção cardinal que representa a relação de direção entre duas entidades espaciais.

A necessidade de cálculos que permitam uma representação e raciocínio espacial está presente em:

- planejamento e navegação robótica;
- interpretação de dados de entrada visuais;
- compreensão da linguagem natural.

Em todos esses casos a necessidade de representar e raciocinar sobre aspectos espaciais do mundo é uma característica chave. Um desafio em aplicar o raciocínio espacial qualitativo se encontra no fato de se tratar de teorias baseadas em lógica de primeira ordem, que sofrem com aplicações em mundo real devido aos ruídos dos sensores, às informações incompletas, ou seja, com a incerteza típica de ambientes reais. Ser capaz de raciocinar sobre a crença no que se percebe e na probabilidade de eventos é importante em muitas das aplicações onde o conhecimento do mundo é incerto (SHIRAZI; AMIR, 2007).

Em ambiente real, um agente trabalha com incerteza devido a uma observação parcial, ou não determinística, ou a combinação de ambas. A probabilidade fornece um modo, de resumir a incerteza de modo a ajudar a encontrar uma solução provável para o problema. O compromisso ontológico da lógica e da probabilidade são os mesmos: em ambos, o mundo é composto de fatos que se fazem ou não presentes em um caso particular, contudo são epistemologicamente diferentes: um agente lógico acredita que cada sentença é verdadeira ou falsa ou não tem opinião, enquanto que um agente probabilístico deve ter um grau de crença numérico entre 0 (falso) e

1 (verdadeiro). Deste modo uma lógica probabilística é representado por uma generalização semântica lógica cujo o valor verdadeiro da sentença é um valor de probabilidade.

Na última década houve um interesse crescente em combinar raciocínio probabilístico e teoria lógica, e em particular a lógica de descrição, a qual é fragmento da lógica de primeira ordem e teve sua origem nos anos 70's (MILCH; RUSSELL, 2007; NIKOLOV et al., 2007; CAMPOS; COZMAN; LUNA, 2009). Essa combinação não é motivada somente pelo interesse teórico puro, mas por sua relevância do ponto de vista de aplicações práticas, com a finalidade de equipar um sistema de raciocínio com inferências relacionais capazes de fazer também as avaliações probabilísticas (SANTOS et al., 2010; FENELON et al., 2010). Existem algumas lógicas probabilísticas distintas, como as propostas em Jaeger (1994b), Sebastiani (1994a), que foram definidas sobre um subconjunto de elementos do domínio. Estas lógicas têm dificuldade em tratar com asserções probabilísticas para indivíduos, como informação estatística sobre o domínio que não implica em informação sobre indivíduos; isto é conhecido como problema de inferência direta (BACCHUS et al., 1996). O problema da inferência direta está resolvido em Dürig e Studer (2005) pela adoção de probabilidade somente nas asserções.

Um modo alternativo de contornar o problema de inferência direta é atribuir probabilidades a subconjuntos de interpretações, como assumido em Laskey e Costa (2012), d'Ámato, Fanizzi e Lukasiewicz (2008) sendo também o cerne da lógica de descrição probabilística chamada credal $\mathcal{ALC}$ ($CR\mathcal{ALC}$) (COZMAN; POLASTRO, 2008) que foi utilizada neste trabalho (SANTOS et al., 2010; FENELON et al., 2010). O $CR\mathcal{ALC}$ é uma extensão probabilística da lógica de descrição $\mathcal{ALC}$ e como esta é baseada em conceitos, os quais representam conjuntos de indivíduos (tais como, Planta e Animal); e papéis, os quais representam relações entre dois indivíduos (tais como, paiDe ou amigoDe). Também tem um conjunto de operadores de intersecção, união e complemento além de alguns quantificadores (POLASTRO, 2012).

Para a localização robótica os métodos que têm maior sucesso são os que utilizam algoritmos probabilísticos, que são em geral criados sobre métodos de filtro Bayesiano (THRUN et al., 2005). Os filtros Bayesianos são algoritmos recursivos composto de duas etapas: a previsão e a atualização de medidas. A localização consiste em localizar as coordenadas de um robô relativo ao ambiente externo que em geral é representado por um mapa. O mapeamento robótico aborda o problema de aquisição de modelos espaciais dos ambientes físicos através de robôs móveis. Esse problema em geral é considerado importante quando se busca um robô móvel realmente autônomo. Apesar dos progressos significativos nesta área, ele ainda apresenta grandes desafios. Um problema que pode ocorrer utilizando esses algoritmos - a crença acabar convergindo para uma posição errada causado por informações ruidosas deixando a localização comprometida, em alguns casos o robô fica perdido ou toma decisões que são arriscadas à integridade física local e do robô. Neste trabalho os mapas são mapas fornecidos para o robô, as regiões do mapa qualitativo são definidas pelo tipo de relações percebidas entre os elementos do domínio. Pesquisas sobre abordagens e aplicações de localização robótica probabilística e mapeamento podem ser encontradas em Thrun et al. (2005) e Thrun (2002) respectivamente.

O objetivo deste trabalho é desenvolver um sistema de raciocínio espacial que seja capaz de trabalhar com incertezas e investigar sua aplicabilidade para a interpretação de cenas em ambientes envolvendo o movimento de veículos. Para que este objetivo seja atingido, as seguintes tarefas foram cumpridas:

1. Estudo das lógicas de descrição probabilísticas e formalismos de raciocínio espacial qualitativo (REQ).
2. Desenvolvimento de uma teoria lógica probabilística que permita representar cenas em termos de raciocínio espacial qualitativo (REQ).
3. Combinação de diferentes cálculos de raciocínio espacial com um filtro Bayesiano.
4. Desenvolvimento de filtro Bayesiano com informação qualitativa no problema de auto localização robótica qualitativa.
5. Validação empírica do sistema desenvolvido a partir de testes com um robô.

Uma lógica de raciocínio espacial qualitativo com probabilidade deve tornar um sistema de inferência mais robusto para tratar com ambientes reais. As teorias de raciocínio espacial qualitativo capacitam um sistema de inteligência artificial inferir informações sobre o ambiente tridimensional a partir de imagens, enquanto que técnicas probabilísticas auxiliam no tratamento de incertezas causadas pela falta de informação ou por informações parciais.

Para utilizar o raciocínio espacial qualitativo com lógica de descrição probabilística, utilizamos parte de um domínio de tráfego baseado no trabalho realizado por Hummel (2009), a qual já havia utilizado uma lógica de descrição como um formalismo de representação do conhecimento com raciocínio espacial para entender cena de um domínio de tráfego urbano. Investigou-se como formalizar esse domínio restrito de tráfego com a lógica de descrição probabilística *CRALLC* (COZMAN; POLASTRO, 2008). Neste caso tratamos do problema restrito às informações disponíveis a partir de sensores acoplados ao veículo e do conhecimento do domínio, inferindo uma nova informação tal como a faixa na qual o veículo estaria localizado. Realizamos as primeiras análises para descobrir como seria possível combinar uma interpretação com lógica de descrição probabilística e o raciocínio espacial aplicáveis em um robô móvel (carro autônomo móvel) (FENELON et al., 2010; SANTOS et al., 2010). O capítulo 3 contém mais detalhes sobre a utilização do *CRALLC* neste ambiente. Esse domínio serviu para uma primeira abordagem do uso de raciocínio espacial qualitativo com probabilidade.

Buscando uma aplicação com dados reais, estendeu-se com uma abordagem probabilística um sistema de auto localização previamente desenvolvido em trabalhos anteriores (SANTOS; DEE; FENELON, 2008; SANTOS; DEE; FENELON, 2009). O sistema trata do problema referente ao robô inferir sua localização qualitativa a partir da sua percepção visual das relações dos elementos alvos no ambiente. Contudo o mecanismo de inferência lógica sofria pela falta

de informação e pelo ruído dos dados do sistema de visão. Sem a capacidade para lidar com incerteza esse sistema estava a espera de um sistema de visão sem ruídos ou com aplicação somente teóricas em ambientes simulados e ou extremamente controlados. Este trabalho está dividido em duas partes em que como contribuição geral integraram sistemas de raciocínio espacial qualitativos com métodos de raciocínios probabilísticos.

Primeira contribuição: criar um método para combinar o cálculo de direções cardinais utilizado em um domínio de tráfego (HUMMEL; THIEMANN; LULCHEVA, 2008) com raciocínio probabilístico usando redes Bayesianas, em particular utilizando a lógica de descrição probabilística *CRA* (capítulo 3) (SANTOS et al., 2010; FENELON et al., 2010). Com os estudos realizados com *CRA* como contribuição constatamos alguns pontos que são essenciais para aplicação de *CRA* em um sistema robótico com raciocínio espacial, tal como a introdução de uma hierarquia de papéis que permita fazer composição de relações (seção 2.2.1) .

Segunda contribuição: é a mais importante, na qual foi proposto e utilizado um método que combina filtro Bayesiano com as lógicas de raciocínio espacial qualitativo que tratam de oclusão, tal como as relações espaciais qualitativas percebidas sobre sombras ou aquelas percebidas entre objetos. O método foi aplicado em sistemas de auto localização utilizando dados de imagens capturadas no ambiente real pelo robô móvel *PeopleBotTM* (capítulos 4 e 5) (FENELON et al., 2013).

As contribuições acima mencionadas foram temas em artigos publicados durante o doutorado. Os trabalhos relatados nestes artigos fazem parte dessa tese, o resumo destes artigos está transcrito no anexo A. Os títulos seguem relacionados abaixo em ordem cronológica:

- Probabilistic logic encoding of spatial domains (SANTOS et al., 2010).
- Encoding spatial domains with relational Bayesian networks (FENELON et al., 2010).
- Reasoning about shadows in a mobile robot environment (FENELON et al., 2012).
- Reconhecimento de Objetos Baseado em Contexto Utilizando a Lógica de Descrição Probabilística (FENELON-PEREIRA; SANTOS; COZMAN, 2012).
- A Qualitative-Probabilistic Approach to Autonomous Mobile Robot Self Localisation and Self Vision Calibration (FENELON et al., 2013)

Neste trabalho existem duas abordagens. Em comum existe um problema de localização sendo tratado utilizando algum tipo de raciocínio espacial qualitativo em uma abordagem probabilística. A primeira abordagem com lógica de descrição probabilística com cálculo de direção cardinal, a localização do veículo não é temporal e os dados para experimentos são simulados. Na segunda abordagem a localização de um robô foi realizada a partir de dados

extraídos de imagens capturadas pelo robô móvel enquanto navegava pelo ambiente, o sistema foi desenvolvido baseado na formalização sobre as relações de oclusão entre elementos do domínio combinado com um filtro Bayesiano. Apesar das semelhanças são dois problemas distintos e não existe uma comparação possível entre eles sobre seus resultados.

1.1 Trabalhos Relacionados

No estado da arte podemos encontrar alguns trabalhos onde as relações espaciais qualitativas são definidas com uma lógica de descrição (GUO; ZHANG, 2009; LUNA; CAMPOS; COZMAN, 2007; CRISTANI; GABRIELLI, 2009). Eles se concentraram no estudo teórico sem aplicar a nenhum problema em específico. Em nossa investigação utilizamos uma lógica de descrição probabilística para definir um domínio com base nas relações espaciais qualitativas (FENELON et al., 2010; SANTOS et al., 2010). Com o conhecimento adquirido nos estudos citados acima, direcionamos nossos esforços para viabilizar um sistema desenvolvido para auto localização de um robô móvel utilizando sombras (FENELON et al., 2012).

O domínio é formalizado utilizando a teoria das Relações Qualitativas Percebíveis sobre Sombras, a qual utiliza os dados de um sistema de visão que definem as relações entre a sombra e o objeto percebidas por um observador. Em Fenelon et al. (2012) um sistema de auto calibração foi proposto onde a partir do movimento do robô, ele mantinha uma crença de: ou continuar na mesma região, ou mudar para uma das regiões vizinhas. Com isso caso o parâmetro de visão não correspondesse à crença do robô, o sistema mudaria o parâmetro até encontrar um que confirmasse sua crença da localização. Todavia este sistema de calibração não funciona onde a imagem não tem informação ou estas estão incompletas (imagem sem sombra ou objeto). A abordagem probabilística que melhor se encaixou para tratar com as incertezas desse caso de auto localização e auto calibração foi um filtro Bayesiano (FENELON et al., 2013).

Para realizar a tarefa de localização as regiões são definidas por linhas de visão, que podem ser informalmente descritas como linhas tangenciais aos objetos chaves. Isto é similar ao encontrado em Fogliaroni et al. (2009), o qual propôs um modelo para navegação qualitativa de um agente autônomo. Em seu modelo Fogliaroni et al. (2009) considera objetos convexos como pontos de referências e o espaço navegável é decomposto em regiões. Mais especificamente, este modelo gera regiões em termos de porções de regiões que são visíveis, parcialmente visíveis e oclusas. A ordem em que os objetos são observados também é usada para qualificar as regiões. Em contraste com Fogliaroni et al. (2009) neste trabalho as teorias de raciocínio espacial qualitativo foram aplicadas em um domínio de robô móvel, na execução da tarefa de auto localização com uma abordagem probabilística para tratar com incertezas encontradas em ambientes reais.

Outro trabalho próximo ao descrito aqui é encontrado em Sjöö, Aydemir e Jensfelt (2012), no qual se utilizam duas relações "*on*" (sobre) e "*in*" (dentro), as quais permitem um sistema robótico móvel analisar a cena e servem como guia na localização de objetos por visão. Sjöö,

Aydemir e Jensfelt (2012) defende que as teorias de raciocínio espaciais qualitativos, por serem formalizados em geral em lógica de primeira ordem, são aplicáveis em sistemas computacionais, mas não são adaptáveis aos sistemas robóticos móveis que estejam em ambientes com incerteza. Entretanto este trabalho mostra que as teorias de raciocínio espacial são aplicáveis em robótica combinando com técnicas probabilísticas. Além disso, este trabalho não está restrito somente a relações topológicas: tal como o cálculo de conexão de regiões e suas extensões, também cálculos de direção e informações de distância foram utilizados. As incertezas são tratadas com um filtro Bayesiano que intuitivamente mantém uma crença na continuidade das ações do robô no mundo. Outra diferença entre nossos trabalhos é que estamos aplicando na tarefa de auto localização qualitativa.

1.2 Organização

O capítulo 2 apresenta um breve relato do atual estado da arte e as teorias envolvidas neste estudo (raciocínio espacial qualitativo, probabilidade Bayesiana, lógica de descrição probabilística, assim como algumas técnicas de visão computacional), procurando sempre relacionar com os objetivos dessa investigação. O capítulo 3 descreve nossas primeiras investigações implementando um domínio de tráfego formalizado em lógica de descrição probabilística com um cálculo de direções cardinais, onde analisamos como utilizar o *CRAFC* e como poderíamos associar probabilidades em sistemas lógicos espaciais, visando à aplicação em agentes móveis autônomos. No capítulo 4, a partir das conclusões derivadas dos experimentos preliminares, começamos o desenvolvimento de uma extensão probabilística de um sistema de auto localização para um robô móvel utilizando raciocínio espacial qualitativo. O capítulo 5 apresenta os resultados e discussões a respeito do sistema de auto localização qualitativa para um robô móvel. Enfim a conclusão e trabalhos futuros são encontrados no capítulo 6.

2 FUNDAMENTOS

Este capítulo descreve as teorias que foram utilizadas nesta pesquisa. Grande parte de nossas investigações envolve alguma teoria ou conceito de Raciocínio Espacial Qualitativo (REQ)¹ (seção 2.2) (STOCK, 1997), tais como: cálculo sobre conexão de regiões² (seção 2.2.1) (RANDELL; CUI; COHN, 1992), cálculo sobre oclusão de região³ (seção 2.2.2) (RANDELL; WITKOWSKI; SHANAHAN, 2001) que são base para uma teoria mais específica chamada relações qualitativas percebíveis sobre sombras⁴ (seção 2.2.3) (SANTOS; DEE; FENELON, 2008) e o cálculo de direção cardinal⁵ (seção 2.2.4) (FRANK, 1996). O raciocínio espacial qualitativo é utilizado por um agente munido de visão computacional para interpretação do mundo tridimensional a partir de informações de imagens bidimensionais (seção 2.2). A captura de informações ou mesmo a representação do conhecimento do mundo real tem a incerteza como característica intrínseca. Apesar de uma abordagem qualitativa ser uma técnica que trata com certo grau de incerteza, ainda assim em algumas situações não é possível representar as relações qualitativas apropriadamente (CHEN et al., 2013).

Para tratar com incerteza as abordagens probabilísticas (seção 2.3) utilizadas foram: as redes Bayesianas relacionais e um filtro Bayesiano. As redes Bayesianas relacionais fazem parte do mecanismo de inferência da lógica de descrição probabilística Credal *ALC* (CRALC) (seção 2.4) (COZMAN; POLASTRO, 2008). Também são abordados brevemente alguns tópicos relacionados às tarefas propostas, tais como: localização robótica (seção 2.5), interpretação de imagens (seção 2.6) e visão computacional (seção 2.7).

2.1 Robótica Cognitiva

Um desafio da robótica cognitiva é a integração de todos os níveis de uma arquitetura de controle que envolva, desde a execução e escolha de uma formalização integrada até os aspectos cognitivos, incluindo o de decidir qual tarefa executar e como executar (PIRRI; MENTUCCIA; STORRI, 2003). Um robô móvel autônomo trabalha continuamente com percepção e ação (NEHMZOW, 2003), assim um robô doméstico que tivesse que executar tarefas como cuidar de sua recarga de bateria, controlar a temperatura da casa, abrir a porta quando a campainha toca, teria que ter uma percepção do que está acontecendo na casa, e de si mesmo, e em situações onde é necessário executar tarefas concorrentes. Vários estudos são realizados buscando desenvolver técnicas melhores para capacitar um robô a executar tarefas diversas de modo autônomo. Kanda

¹ Do inglês: *Qualitative Spatial Reasoning* (QSR)

² Do inglês: *Regions Connection Calculus* (RCC)

³ Do inglês: *Region Occlusion Calculus* (ROC)

⁴ Do inglês: *Perceptual Qualitative Relations about Shadows* (PQRS)

⁵ Do inglês: *Cardinal Direction Calculus* (CDC)

et al. (2006) realizou pesquisas com robôs interagindo com pessoas, procurando melhorar a maneira das máquinas executarem tarefas de modo aceitável e confiável para os humanos. Pesquisas com foco em utilizar a percepção robótica são relatadas em Scherl e Levesque (1993) o qual utilizou as informações percebidas pelo robô para produzir conhecimento a partir de suas ações, enquanto que em Santos e Shanahan (2002a) os dados coletados pelos sensores do robô móvel serviram de base para a construção de um sistema de raciocínio espacial. Essas pesquisas são em geral realizadas em ambientes controlados ou executando apenas tarefas específicas. Ambientes controlados são utilizados em robótica devido ao elevado grau de incerteza que existe em ambientes humanos dinâmicos. Alguns pesquisadores acreditam que poderemos elevar a capacidade cognitiva robótica agregando os vários algoritmos desenvolvidos para efetuar diversas tarefas (MCCARTHY, 2007).

2.2 Raciocínio Espacial Qualitativo

O objetivo da área de raciocínio espacial qualitativo (REQ) é a formalização do conhecimento espacial a partir de entidades elementares tais como regiões espaciais (RANDELL; CUI; COHN, 1992), segmentos de linha (MORATZ; RENZ; WOLTER, 2000; SCHLIEDER, 1996), direções (FREKSA, 1992; LIGOZAT, 1998), entre outras (STOCK, 1997; COHN; HAZARIKA, 2001; CHEN et al., 2013). Uma síntese mais completa neste campo pode ser encontrada em (COHN; HAZARIKA, 2001; CHEN et al., 2013). Começaremos pelo cálculo sobre conexão de regiões que é uma das teorias mais conhecidas para representar e raciocinar com relações topológicas, além de ser a base para outras abordagens em raciocínio espacial qualitativo (CHEN et al., 2013).

2.2.1 Cálculo sobre conexão de regiões

Region Connection Calculus (RCC) (RANDELL; CUI; COHN, 1992) ou cálculo sobre conexão de regiões é um sistema axiomático sobre regiões espaciais em lógica de primeira ordem, que tem como base uma relação primitiva binária de conectividade simétrica e reflexiva. Para um melhor entendimento, dadas duas regiões x e y , a relação $C(x, y)$ representa que “ x está conectado com y ” sendo verdadeira se somente se as regiões x e y tiverem pelo menos um ponto em comum. A relação de conectividade obedece aos seguintes axiomas (RANDELL; CUI; COHN, 1992; GOTTS, 1994):

$$\begin{aligned} &\forall x \ C(x, x); \\ &\forall x \ y \ C(x, y) \rightarrow C(y, x). \end{aligned}$$

Assumindo a relação binária $C/2$ e as variáveis x, y e z para regiões espaciais define-se diferentes relações entre pares. A definição de algumas dessas relações e suas interpretações podem ser vistas na Tabela 1 (RANDELL; CUI; COHN, 1992; RENZ, 2002). As relações DC , EQ ,

Tabela 1: Definições e interpretações de algumas relações do RCC.

Definição da relação	Interpretação
$DC(x, y) \equiv_{def} \neg C(x, y)$	' x está desconectado de y '
$P(x, y) \equiv_{def} \forall z[C(z, x) \rightarrow C(z, y)]$	' x é parte de y '
$PP(x, y) \equiv_{def} P(x, y) \wedge \neg P(y, x)$	' x é parte própria de y '
$EQ(x, y) \equiv_{def} P(x, y) \wedge P(y, x)$	' $x = y$ ': ' x é idêntico a y '
$O(x, y) \equiv_{def} \exists z[P(z, x) \wedge P(z, y)]$	' x sobrepõe y '
$PO(x, y) \equiv_{def} O(x, y) \wedge \neg P(x, y) \wedge \neg P(y, x)$	' x sobrepõe parcialmente y '
$DR(x, y) \equiv_{def} \neg O(x, y)$	' x é distinta de y '
$EC(x, y) \equiv_{def} C(x, y) \wedge \neg O(x, y)$	' x está externamente conectado à y '
$TPP(x, y) \equiv_{def} PP(x, y) \wedge \exists z[EC(z, x) \wedge EC(z, y)]$	' x é uma parte própria tangencial de y '
$NTPP(x, y) \equiv_{def} PP(x, y) \wedge \neg \exists z[EC(z, x) \wedge EC(z, y)]$	' x é uma parte própria não tangencial de y '
$PI(x, y) \equiv_{def} P(y, x)$	' y é parte de x '
$PPI(x, y) \equiv_{def} PP(y, x)$	' y é parte própria de x '
$TPPI(x, y) \equiv_{def} TPP(y, x)$	' y é uma parte própria tangencial de x '
$NTPPI(x, y) \equiv_{def} NTPP(y, x)$	' y é uma parte própria não tangencial de x '

fonte:Randell, Cui e Cohn (1992), Renz (2002).

Figura 1: Diagrama conceitual de vizinhança (CND - *Conceptual Neighborhood Diagram*) das relações RCC-8. Note a continuidade das transições topológicas entre regiões (indicadas pelas setas), as mudanças de relações ocorrem sempre para uma relação vizinha.

Fonte: adaptada de Cui, Cohn e Randell (1992).

PO , EC , TPP , $NTPP$, $TPPI$ e $NTPPI$ formam um conjunto de relações mutuamente excludentes e disjuntas par a par. Este conjunto é referido como RCC-8 (RANDELL; CUI; COHN, 1992; STOCK, 1997; RENZ, 2002).

No contexto do raciocínio espacial qualitativo assume-se que as mudanças das relações entre regiões ocorrem de forma contínua. Todas as transições topológicas entre regiões são contínuas, ou seja, quando uma relação muda, essa mudança segue sempre para uma relação próxima ou vizinha. Na Figura 1 vemos uma representação gráfica de vizinhança das oito relações básicas do RCC (RCC-8). Nessa mesma figura as transições são representadas pelas setas; essa representação é conhecida como *diagrama conceitual de vizinhança* (CND) (CUI; COHN; RANDELL, 1992; RANDELL; CUI; COHN, 1992). A continuidade nas transições das relações é um conceito importante, fazendo com que o CND seja uma importante ferramenta para raciocinar em várias situações (STOCK, 1997).

Outra ferramenta para representar e raciocinar sobre as disjunções e conjunções de pares do conjunto de relações binárias, tanto espacial quanto temporal, é a tabela de composição (RANDELL; WITKOWSKI, 2002). Esta tabela contém as transitividades entre três regiões

distintas a, b e c , onde, dadas informações sobre as relações entre a e b e entre b e c , se extrai um conjunto de relações possíveis entre as regiões a e c . A Tabela 2 é uma tabela de composição para RCC-8, onde a primeira coluna representa as relações entre a e b ($R_1(a, b)$) e a primeira linha as relações entre b e c ($R_2(b, c)$). O conjunto das relações possíveis entre a e c ($R_3(a, c)$) estão representadas nas células definidas pela intersecção entre a coluna de $R_1(a, b)$ e a linha de $R_2(b, c)$ (RANDELL; WITKOWSKI, 2002).

Para exemplificar tanto a transitividade quanto a transição de relações utilizaremos um exemplo gráfico (Figura 2). Considerando as relações entre as três regiões (a, b e c) representadas à esquerda da Figura 2, onde $R_1(a, b)$ é uma relação $NTPP(a, b)$, caso ocorra uma transição da relação $R_1(a, b)$ pelo diagrama conceitual de vizinhança a próxima relação seria ou $EQ(a, b)$ ou $TPP(a, b)$, na Figura 2 a relação mudou para $TPP(a, b)$, enquanto que a relação $R_2(b, c)$ ($TPP(b, c)$) permaneceu igual. Nesta mesma figura também está representada a transitividade entre as três regiões (a, b e c). Pode se observar duas composições: a esquerda note que composição das relações $NTPP(a, b)$ e $TPP(b, c)$ resultam na relação $NTPP(a, c)$, enquanto a direita a composição resultante entre $TPP(a, b)$ e $TPP(b, c)$ são as duas relações $NTPP(a, c)$ e $TPP(a, c)$ referentes às duas possíveis transições. Esses resultados estão expressos na tabela de composição (Tabela 2).

A tabela de composição e o diagrama conceitual de vizinhança são ferramentas de inferência aplicáveis com outros cálculos espaciais. Neste trabalho o conceito de transição foi utilizado para inferir mudança de posição em relação ao contexto onde um observador está, ou seja, esse conceito de continuidade e vizinhança foi utilizado na criação do modelo de transição do sistema de auto localização (seção 4.4.4) (FENELON et al., 2013). A tabela de composição serve para inferir uma relação entre indivíduos em contextos nos quais essas relações não são percebidas diretamente pelos sensores. Neste trabalho, utilizamos uma tabela de composição para inferir relações entre um veículo e faixa da rua, das relações entre o veículo e linha divisória, e da faixa e a linha divisória a partir das relações percebidas, de informações do domínio e dos sensores (seção 3.2) (FENELON et al., 2010).

O cálculo de conexão de regiões (RCC) representa as relações mereotopológicas entre regiões espaciais que são independentes do ponto de vista do observador. Isto é diferente do cálculo de linhas visuais (*lines-of-sight calculus*) (GALTON, 1994) que a partir de um ponto de vista representa as posições relativas entre pares de corpos convexos não sobrepostos. Inspirado pelo cálculo de linhas visuais foi criada uma teoria conhecida como cálculo sobre oclusão de região (RANDELL; WITKOWSKI; SHANAHAN, 2001), a qual é uma extensão do RCC (RANDELL; CUI; COHN, 1992), proposta para representar as várias possibilidades de interposições entre quaisquer objetos dependendo do ponto de vista do observador. A seguir descreve-se brevemente esta teoria.

Tabela 2: Tabela de composição para o RCC-8 com o conjunto das relações exclusivamente disjuntas possíveis $R_3(a, c)$ entre a e c a partir das relações $R_1(a, b)$ e $R_2(b, c)$.

$R_1(a, b)$	$R_2(b, c)$							
	DC	EC	PO	TPP	TPPI	EQ	NTPP	NTPPI
DC	DC EC PO TPP TPPI EQ NTPP NTPPI	DC EC PO TPP NTPP	DC EC PO TPP NTPP	DC EC PO TPP NTPP	DC	DC	DC EC PO TPP NTPP	DC
EC	DC EC PO TPPI NTPPI	DC EC PO TPPI TPPI EQ	DC EC PO TPP NTPP	EC PO TPP NTPP	DC EC	EC	PO TPP NTPP	DC
PO	DC EC PO TPPI NTPPI	DC EC PO TPPI NTPPI	DC EC PO TPP TPPI EQ NTPP NTPPI	PO TPP NTPP	DC EC PO TPPI	PO	PO TPP NTPP	DC EC PO TPPI NTPPI
TPP	DC	DC EC	DC EC PO TPP NTPP	TPP NTPP	DC EC PO TPP TPPI EQ	TPP	NTPP	DC EC PO TPPI NTPPI
TPPI	DC EC PO TPPI NTPPI	EC PO TPPI NTPPI	PO TPPI NTPPI	PO TPP TPPI EQ	TPPI NTPPI	TPPI	PO TPP NTPP	NTPPI
EQ	DC	EC	PO	TPP	TPPI	EQ	NTPP	NTTPI
NTPP	DC	DC	DC EC PO TPP NTPP	NTPP	DC EC PO TPP NTPP	NTPP	NTPP	DC EC PO TPP TPPI EQ NTPP NTPPI
NTPPI	DC EC PO TPPI NTPPI	PO TPPI NTPPI	PO TPPI NTPPI	PO TPPI NTPPI	NTPPI	NTPPI	PO TPP TPPI EQ NTPP NTPPI	NTPPI

Fonte Randell e Witkowski (2002).

Figura 2: Representação gráfica das composições $R_1(a, b)$ e $R_2(b, c)$ antes e depois da transição representada pela seta, note que o conjunto resultante $R_3(a, c)$ são as relações possíveis (Tabela 2).

Fonte: autor.

Figura 3: Duas sequências de imagens (da esquerda para direita) onde um cone passa em frente (sequência superior) e por trás (sequência inferior) do cilindro, este é um exemplo da oclusão em ação. Note na ocorrência de oclusão os dois objetos formam uma única região na imagem.

Fonte: autor.

2.2.2 Cálculo sobre oclusão de região

Region occlusion calculus (ROC) (RANDELL; WITKOWSKI; SHANAHAN, 2001) ou cálculo sobre oclusão de região representa os eventos ou relações de oclusão espacial. A oclusão espacial ocorre quando um corpo se interpõe a outro a partir de um ponto de vista. Eventos de oclusão nos ajudam a determinar onde está o contorno de um objeto, ou inferir por que um objeto não pode ser visto e o que precisamos fazer para torná-lo visível. O ROC estende o RCC-8 para raciocinar sobre distâncias relativas entre dois corpos a partir dos eventos de oclusão e a transição entre relações de oclusão.

O ROC pode ser usado para modelar o efeito de paralaxe causado por ambos os movimentos do objeto ou do observador. Como exemplo, vamos considerar que as imagens dos objetos são regiões. Note que, na Figura 3, quando ocorre uma oclusão, nas imagens os objetos aparecem formando uma única região. O ROC capacita à inferência de que existem (neste caso) dois objetos sob alguma relação de oclusão.

Um evento de oclusão espacial pode ser visto na Figura 3 com movimento realizado pelo objeto. A partir de um ponto de vista fixo: o objeto cônico passa pela frente do objeto cilíndrico na sequência superior da imagem, enquanto que na sequência inferior passa por trás (Figura 3). As transições das relações ROC podem ajudar a determinar modelos de movimento, pois quando se muda o ponto de vista ocorre paralaxe (Figura 4), essa informação pode ser utilizada para derivar um modelo objetivo do mundo de um ponto de vista naturalmente restrito (RANDELL; WITKOWSKI; SHANAHAN, 2001). Para distinguir entre um corpo físico e a região do espaço por ele ocupada, e distinguir entre um corpo físico e sua projeção em relação a um ponto de vista, o ROC utiliza-se de duas funções: *region* e *image*. Pode-se entender a função *region* como o mapeamento de um corpo físico e sua região de ocupação. A função *image* representa o mapeamento relativo a uma projeção bidimensional do corpo físico a partir de um determinado

Figura 4: Mudança na percepção da oclusão causada pelo movimento do observador (paralaxe).

Fonte: adaptado de (RANDELL; WITKOWSKI; SHANAHAN, 2001).

ponto de vista. A distinção formal é apresentada pelas duas funções: ‘*region(x)*’ lida como “a região ocupada por *x*” e ‘*image(x, ν)*’ lida como “a imagem de *x* em relação ao ponto de vista *ν*” (RANDELL; WITKOWSKI, 2002).

Com a relação primitiva *TotallyOccludes*(*x, y, ν*) define-se um evento de oclusão total entre dois corpos, essa relação é lida como “*x* oclui totalmente *y* com relação ao ponto de vista *ν*”. *TotallyOccludes/3* é axiomatizada como transitiva e irreflexiva (RANDELL; WITKOWSKI; SHANAHAN, 2001). Para descrever o significado geométrico de uma oclusão total, considere que uma *linha*(*p*₁, *p*₂, *p*₃) é definida de modo que, os pontos *p*₁, *p*₂ e *p*₃ estejam em uma reta, com *p*₂ posicionado estritamente entre *p*₁ e *p*₃. Então, uma região *x* oclui totalmente uma região *y* do ponto de vista *ν* se e somente se para todo o ponto *p* em *y*, existir um ponto *q* em *x* tal que *linha*(*ν, q, p*) seja verdade e não exista nenhum outro ponto *p*’ em *y* e *q*’ em *x*, com os quais *linha*(*ν, p*’, *q*’) seja verdadeiro. Vários axiomas que constituem o ROC serão apresentados a seguir (RANDELL; WITKOWSKI; SHANAHAN, 2001):

$$\begin{aligned} \forall x y z \nu \ [& [TotallyOccludes(x, y, \nu) \wedge P(region(z), region(y))] \\ & \longrightarrow TotallyOccludes(x, z, \nu)], \end{aligned} \quad (2.1)$$

isto é, se *x* oclui totalmente *y*, *x* oclui totalmente qualquer parte de *y*.

$$\begin{aligned} \forall x y \nu \ [& TotallyOccludes(x, y, \nu) \longrightarrow \forall z \ [P(region(z), region(y))] \\ & \longrightarrow \neg TotallyOccludes(z, x, \nu)], \end{aligned} \quad (2.2)$$

Figura 5: Modelo gráfico das 20 relações básicas do cálculo sobre oclusão de região, o ROC-20.

Fonte: Randell e Witkowski (2002).

isto é, se x oclui totalmente y nenhuma parte de y oclui totalmente x .

$$\begin{aligned}
 & \forall x y \nu [TotallyOccludes(x, y, \nu) \\
 & \longrightarrow \forall z u [[P(region(z), region(x)) \wedge P(region(u), region(y))] \\
 & \longrightarrow \neg TotallyOccludes(u, z, \nu)],
 \end{aligned} \tag{2.3}$$

isto é, se x oclui totalmente y nenhuma parte de y oclui totalmente parte de x .

$$\begin{aligned}
 & \forall x \nu \exists y z [P(region(y), region(x)) \wedge P(region(z), region(x)) \wedge \\
 & TotallyOccludes(y, z, \nu)],
 \end{aligned} \tag{2.4}$$

isto é, todo x tem uma parte que oclui totalmente outra parte de x . Este axioma garante que o corpo tem profundidade.

$$\forall x y \nu [TotallyOccludes(x, y, \nu) \longrightarrow P(image(y, \nu), image(x, \nu))], \tag{2.5}$$

isto é, se x oclui totalmente y , a imagem de y é parte da imagem de x .

A seguir define-se uma relação de oclusão enfraquecida como $Occludes(x, y, \nu)$ a qual é lida como “ x oclui y com relação ao ponto de vista ν ”, utilizada para incluir outras relações de oclusão mais específicas tais como as oclusões parcial e mútua:

$$Occludes(x, y, \nu) \equiv \exists z u [P(region(z), region(x)) \wedge P(region(u), region(y)) \wedge TotallyOccludes(z, u, \nu)], \quad (2.6)$$

isto é, x oclui y do ponto de vista ν se pelo menos existe uma parte de x que oclui totalmente pelo menos uma parte de y .

$$PartiallyOccludes(x, y, \nu) \equiv Occludes(x, y, \nu) \wedge \neg TotallyOccludes(x, y, \nu) \wedge \neg Occludes(y, x, \nu), \quad (2.7)$$

isto é, x oclui parcialmente y do ponto de vista ν , mas y não oclui x .

$$MutuallyOccludes(x, y, \nu) \equiv Occludes(x, y, \nu) \wedge Occludes(y, x, \nu), \quad (2.8)$$

isto é, x e y ocluem um ao outro do ponto de vista ν .

No cálculo sobre oclusão de região (ROC) também ocorrem as relações inversas identificadas pelo expoente -1 . Essas relações são encontradas em Randell, Witkowski e Shanahan (2001), Randell e Witkowski (2002) e não serão listadas aqui. O ROC reduz a descrição de alinhamentos possíveis entre os objetos a vinte relações básicas. Por isso, esse cálculo é conhecido também por ROC-20 (WITKOWSKI et al., 2001; RANDELL; WITKOWSKI, 2002). A Figura 5 é uma representação gráfica dessas vinte relações. O ROC-20 (RANDELL; WITKOWSKI; SHANAHAN, 2001) utiliza um conjunto de relações binárias baseado no RCC-8 (RANDELL; CUI; COHN, 1992) e estende o cálculo de linhas visuais (GALTON, 1994), para modelar o relacionamento espacial entre corpos no espaço tridimensional e suas imagens bidimensionais, as quais são obtidas a partir de um ponto de vista. Essas relações são definidas pelo esquema de axiomas abaixo (RANDELL; WITKOWSKI; SHANAHAN, 2001).

$$\begin{aligned} \Phi\Psi(x, y, \nu) &\equiv \Phi(x, y, \nu) \wedge \Psi(image(x, \nu), image(y, \nu)); \\ \Xi\Psi^{-1}(x, y, \nu) &\equiv \Xi(y, x, \nu) \wedge \Psi(image(y, \nu), image(x, \nu)); \end{aligned}$$

onde:

$$\Phi = NonOccludes, \text{ então } \Psi \in \{DC, EC\};$$

$$\Phi = TotallyOccludes, \text{ então } \Psi \in \{EQ, TPPI, NTPPI\};$$

$$\Phi = PartiallyOccludes, \text{ então } \Psi \in \{PO, TPP, NTPP\};$$

$$\Phi = MutuallyOccludes, \text{ então } \Psi \in \{PO, EQ, TPP, NTPP\};$$

e onde:

$$\Xi = TotallyOccludes, \text{ então } \Psi \in \{EQ, TPPI, NTPPI\};$$

$$\Xi = PartiallyOccludes, \text{ então } \Psi \in \{PO, TPP, NTPP\};$$

$$\Xi = MutuallyOccludes, \text{ então } \Psi \in \{TPP, NTPP\}.$$

Figura 6: Estudos sobre sombra de Leonardo da Vinci. No esboço vê-se representados as sombras do objeto projetadas em outra superfície (sombra-projetada) e no próprio objeto (sombra-própria).

Fonte: Vinci (1970).

Com os axiomas acima as relações de oclusão em ambiente tridimensional podem ser interpretadas em um plano. Assim, analisando as relações espaciais que ocorrem entre regiões distintas em uma imagem podemos deduzir as relações espaciais no ambiente real, incorporando uma noção de profundidade e distância comparativa entre os objetos. Profundidade e distância ambas são essenciais em teorias lógicas quando utilizadas em aplicações práticas em robótica (WITKOWSKI et al., 2001).

Neste trabalho, o mecanismo de inferência de auto localização extraiu informações do ambiente a partir das relações percebidas nas imagens capturadas por um robô (capítulo 4). Deste modo, o ROC esteve presente na formalização tanto dos eventos de oclusão entre objetos quanto dos eventos de oclusão entre objeto e sombra. Considerando as relações ROC entre objeto e sombra a partir de um ponto de vista, foi proposta em Santos, Dee e Fenelon (2008), Santos, Dee e Fenelon (2009) uma nova teoria de raciocínio espacial, específica sobre informações de sombras, chamada de *Relações qualitativas percebíveis sobre sombras (Perceptual Qualitative Relations about Shadows - PQRS)*, a qual será descrita a seguir.

2.2.3 Relações qualitativas percebíveis sobre sombras

Para uma breve definição, uma sombra é formada quando a luz é obstruída, em sua trajetória, por um objeto opaco. Ela pode ser classificada em sombra-própria e sombra-projetada. A sombra-projetada é aquela que ocorre quando a sombra forma-se em uma superfície, enquanto que na sombra-própria a superfície de projeção é o próprio objeto que a provocou (Figura 6), uma análise da estrutura geométrica das sombras pode ser encontrada em Knill, Mamassian e Kersten (1997). Concentraremos na sombra-projetada que será referida neste trabalho somente como sombra. Para se formar uma sombra é necessária a combinação de três elementos básicos. Esses elementos são a fonte de luz, a superfície na qual a sombra é projetada e um objeto opaco. Um observador e cada um desses elementos são importantes, pois, qualquer alteração na localização de algum destes provocará um movimento da sombra.

Figura 7: Efeito na localização do objeto ocasionado pela posição da sombra.

Fonte: Kersten, Mamassian e Knill (1994).

Na Figura 7 aparecem três imagens que (quando mostradas em sequência) dão a impressão errônea de que o quadrado do meio move-se em profundidade, enquanto que de fato somente a sombra se movimenta. O curioso é que o movimento da sombra poderia ser ocasionado tanto pelo movimento do observador, ou pelo movimento da superfície na qual está projetada, ou devido à fonte de luz, contudo nessa demonstração são os quadrados que parecem se mover. Isso sugere que a percepção humana prefere a informação da sombra em detrimento a outros indicativos de movimento do objeto, ignorando inclusive a falta de alteração de tamanho do mesmo (KERSTEN; MAMASSIAN; KNILL, 1994).

Também há evidências de que as sombras são particularmente expressivas em indicar profundidade em cenas dinâmicas onde os objetos e suas sombras estão em movimento (MAMASSIAN; KNILL; KERSTEN, 1998). De fato, nos estudos realizados por Kersten, Mamassian e Knill (1994), relacionando objetos em movimento e a percepção relativa à profundidade, demonstrou-se que uma simples mudança no movimento da sombra é o suficiente para alterar drasticamente a nossa percepção da trajetória do objeto como mostra a Figura 8. Observando a Figura 8 vemos em (a) a bola posicionada próxima à superfície a frente da caixa, enquanto que em (b) e (c) apesar da bola estar na mesma posição, devido à sombra projetada, em (c) percebemos a bola longe da superfície na frente da caixa e em (b) vemos a bola no fundo da caixa próxima a superfície.

As sombras são utilizadas em aplicações de computação gráfica, por sua capacidade em melhorar o realismo das imagens e por nos ajudar a perceber ambientes 3D em cenas 2D (HASENFRATZ et al., 2003), entretanto, em visão computacional sombras são na maioria das vezes descartadas (PRATI et al., 2003). A teoria de relações qualitativas percebíveis sobre sombra (PQRS) foi criada com a proposta de utilizar as informações de sombra em sistemas artificiais, tal como robótica móvel (SANTOS; DEE; FENELON, 2008).

O PQRS é uma teoria fundamentada no cálculo sobre oclusão de região e assume um contexto onde os elementos são: o observador, o objeto, a fonte de luz, a sombra e a superfície onde a sombra é projetada. A fonte de luz (única e fixa acima do observador e do objeto) será representada pelo símbolo constante L . Em seguida Scr representa a superfície onde a sombra é projetada. O observador será representado com o símbolo ν . Os conjuntos de variáveis são representados como objetos físicos: $o_1, \dots, o_n$; instantes de tempo: $t_1, \dots, t_n$; regiões espaciais:

Figura 8: Bola na caixa. Mesmo na mesma localização da bola em (b) e (c) percebemos posicionamentos diferentes devido à sombra.

Fonte: Kersten, Mamassian e Knill (1994).

$r_1, \dots, r_n$; e sombras: $s_1, \dots, s_n$. Assume-se que as variáveis são universalmente quantificadas a menos que informado explicitamente sobre o contrário.

Formado por um subgrupo do ROC, a teoria das relações qualitativas percebíveis sobre sombras (PQRS) é um conjunto composto pelas relações de um objeto o e sua sombra s representadas graficamente na Figura 9. A seguir as interpretações e definições das relações básicas do PQRS, são exemplificadas com imagens capturadas em nosso laboratório (Figura 10) (a conexão com a base do objeto é desconsiderada):

$$\begin{aligned} NonOccludesDC(o, s, \nu) \equiv & NonOccludes(o, s, \nu) \wedge \\ & DC(image(o, \nu), image(s, \nu)), \end{aligned} \quad (2.9)$$

do ponto de vista ν não existe oclusão entre o objeto o e a sombra s conforme representado na Figura 10a, onde a sombra s não está ocluída e na imagem aparece desconectada do objeto o do ponto de vista ν ;

$$\begin{aligned} NonOccludesEC(o, s, \nu) \equiv & NonOccludes(o, s, \nu) \wedge \\ & EC(image(o, \nu), image(s, \nu)), \end{aligned} \quad (2.10)$$

do ponto de vista ν existe oclusão parcial da sombra s pelo objeto o conforme representado na Figura 10b, onde objeto o não oclui a sombra s porém suas bordas estão conectadas do ponto de vista ν ;

$$\begin{aligned} PartiallyOccludesPO(o, s, \nu) \equiv & PartiallyOccludes(o, s, \nu) \wedge \\ & PO(image(o, \nu), image(s, \nu)), \end{aligned} \quad (2.11)$$

do ponto de vista ν existe oclusão parcial entre o objeto o e a sombra s conforme representado na Figura 10c, onde o objeto o oclui a sombra s de modo que na imagem a sombra s aparece

Figura 9: Representação gráfica das relações básicas do PQRS.

Fonte: adaptada pelo autor de Randell e Witkowski (2002).

parcialmente ao lado do objeto o do ponto de vista ν ;

$$\begin{aligned} \text{PartiallyOccludesTPP}(o, s, \nu) \equiv & \text{PartiallyOccludes}(o, s, \nu) \wedge \\ & \text{TPP}(\text{image}(o, \nu), \text{image}(s, \nu)), \end{aligned} \quad (2.12)$$

do ponto de vista ν existe oclusão parcial da sombra s pelo objeto o conforme representado na Figura 10d, onde objeto o oclui parte da sombra s tangenciando uma das bordas da sombra s na imagem do ponto de vista ν ;

$$\begin{aligned} \text{TotallyOccludesTPPI}(o, s, \nu) \equiv & \text{TotallyOccludes}(o, s, \nu) \wedge \\ & \text{TPPI}(\text{image}(o, \nu), \text{image}(s, \nu)), \end{aligned} \quad (2.13)$$

do ponto de vista ν existe oclusão total da sombra s pelo objeto o conforme representado na Figura 10e, onde objeto o oclui totalmente a sombra s a qual é parte tangencial do objeto o na imagem do ponto de vista ν ;

$$\begin{aligned} \text{TotallyOccludesEQ}(o, s, \nu) \equiv & \text{TotallyOccludes}(o, s, \nu) \wedge \\ & \text{EQ}(\text{image}(o, \nu), \text{image}(s, \nu)), \end{aligned} \quad (2.14)$$

do ponto de vista ν existe oclusão total da sombra s pelo objeto o conforme representado na Figura 10f, onde objeto o oclui totalmente a sombra s a qual é igual ao objeto o na imagem do ponto de vista ν ;

$$\begin{aligned} \text{TotallyOccludesNTPPI}(o, s, \nu) \equiv & \text{TotallyOccludes}(o, s, \nu) \wedge \\ & \text{NTPPI}(\text{image}(o, \nu), \text{image}(s, \nu)), \end{aligned} \quad (2.15)$$

do ponto de vista ν existe oclusão total da sombra s pelo objeto o conforme representado na Figura 10g, onde objeto o oclui totalmente a sombra s a qual não tangencia o objeto o na imagem do ponto de vista ν .

Além das relações formalizadas acima, a teoria PQRS adota a primitiva $\text{Shadow}(s, o, \text{Scr}, L)$ que representa a sombra s do objeto o , provocada pela fonte de luz L sobre a superfície Scr .

Figura 10: Imagens capturadas pelo *PeopleBotTM*. Representando as relações PQRS (a) NonOccludesDC(o,s), (b) NonOccludesEC(o,s), (c) PartiallyOccludesPO(o,s), (d) PartiallyOccludesTPP(o,s), (e) TotallyOccludesTPPI(o,s), (f) TotallyOccludesEQ(o,s) e (g) TotallyOccludesNTPPI(o,s).

Fonte:(SANTOS; DEE; FENELON, 2008).

Pode-se declarar que a sombra s do objeto o é a mesma região da superfície Scr que está totalmente ocluída pelo objeto o do ponto de vista da fonte de luz L . O axioma que restringe a relação *Shadow/4* é apresentado pela fórmula (2.16) abaixo:

$$\begin{aligned} Shadow(s, o, Scr, L) \leftrightarrow & \neg \exists o' (o \neq o') \wedge Occludes(o', o, L) \wedge \\ & PO(region(s), region(Scr)) \wedge TotallyOccludes(o, s, L). \end{aligned} \quad (2.16)$$

O axioma (2.16) diz que a sombra s do objeto o é a região na superfície Scr que está totalmente ocluída por o do ponto de vista da fonte de luz L , caso não exista outro objeto o' ocluindo o do ponto de vista da fonte de luz. A partir deste formalismo podem-se provar alguns teoremas sobre fatos do senso comum, os quais podem ser encontrados em Santos, Dee e Fenelon (2008), Fenelon et al. (2012).

Até agora as descrições das relações de oclusão foram feitas do ponto de vista da fonte de luz. Porém, também é possível raciocinar sobre sombras a partir do ponto de vista de um observador. Com o PQRS podemos inferir informações tais como percepção de profundidade e a localização relativa do observador, ambas descritas em (SANTOS; DEE; FENELON, 2008; SANTOS; DEE; FENELON, 2009). Por ser relevante a esse trabalho a seguir descreveremos somente a localização relativa.

2.2.3.1 Localização relativa

Quando o observador tem dentro do seu campo visual um objeto com sua sombra correspondente, podemos por dedução determinar a posição do observador em relação ao objeto e a fonte de luz, ou seja, sua localização relativa. É possível definir relações entre sombra e objeto a partir do ponto de vista de um observador. Essas relações dependem da posição em que o observador se encontra, portanto esta informação pode ser utilizada em um mecanismo de auto localização do robô. A partir da descrição de oclusão da sombra, cinco regiões podem ser observadas e definidas pelas linhas de visão entre a fonte de luz L , o objeto o e sua sombra s , como representado na Figura 11b.

Para definir formalmente esta localização, assumimos a relação $located(r, \nu, o, s)$ que representa “um observador ν localizado na região r em relação ao objeto o e a sombra s ”. As relações são definidas a partir da sombra projetada pela parte superior do objeto (Figura 11a). O termo $r (r \in 1, 2, 3, 4, 5)$ representa as regiões mostradas na Figura 11b, então para definir a localização em cada região temos que (SANTOS; DEE; FENELON, 2009):

$$\begin{aligned} located(1, \nu, o, s) \leftarrow & Shadow(s, o, Scr, L) \wedge \\ & NonOccludesDC(o, s, \nu) \wedge \nu \neq o, \end{aligned} \quad (2.17)$$

Figura 11: (a) Um objeto com sua sombra conectada e projetada numa superfície irregular, destaque para a parte da sombra relativa a parte superior do objeto; (b) Regiões distintas expressas pela observação do objeto e sua sombra, definidas pelas linhas de visão.

Fonte: Santos, Dee e Fenelon (2008).

a fórmula (2.17) define que um observador ν na região $r = 1$ observa que: o não oclui e está desconectado de s , e s não é sombra de ν (Figura 12a);

$$\begin{aligned} located(2, \nu, o, s) &\leftarrow Shadow(s, o, Scr, L) \wedge \\ NonOccludesEC(o, s, \nu) &\wedge \nu \neq o, \end{aligned} \quad (2.18)$$

a fórmula (2.18) define que um observador ν na região 2 observa que: o não oclui e está externamente conectado de s , e s não é sombra de ν (Figura 12b);

$$\begin{aligned} located(3, \nu, o, s) &\leftarrow Shadow(s, o, Scr, L) \wedge \\ PartiallyOccludesPO(o, s, \nu) &\wedge \nu \neq o, \end{aligned} \quad (2.19)$$

a fórmula (2.19) define que um observador ν na região 3 observa que: o parcialmente oclui e sobrepõe s , e s não é sombra de ν (Figura 12c);

$$\begin{aligned} located(4, \nu, o, s) &\leftarrow Shadow(s, o, Scr, L) \wedge \\ TotallyOccludesTPPI(o, s, \nu) &\wedge \nu \neq o, \end{aligned} \quad (2.20)$$

a fórmula (2.20) define que um observador ν na região 4 observa que: o oclui totalmente s e este é parte tangencial de o , e s não é sombra de ν (Figura 12d);

$$\begin{aligned} located(5, \nu, o, s) &\leftarrow Shadow(s, o, Scr, L) \wedge \\ TotallyOccludesNTPPI(o, s, \nu) &\wedge \nu \neq o, \end{aligned} \quad (2.21)$$

a fórmula (2.21) define um observador ν na região 5 observa que: o oclui totalmente s e este é parte não tangencial de o , e s não é sombra de ν (Figura 12e).

Esta formalização das relações entre objetos e suas sombras do ponto de vista de um observador facilita caracterizar qualitativamente o espaço em volta do objeto. Iremos analisar cada região com fotos capturadas pelo *PeopleBotTM* no laboratório do centro universitário da FEI. As imagens mantêm o foco no objeto e na sua sombra projetada na parede enquanto o ponto de vista do observador será deslocado da região 1 até a região 5 (Figura 12).

Na Figura 12a é observada uma sombra completa não oclusa pelo objeto onde ambos (sombra e objeto) estão desconectados (*NonOccludesDC*) caracterizando a região 1. Quando o observador se desloca em direção à região 2, a sombra parece se deslocar ao encontro de seu objeto. No momento em que o deslocamento do observador provoca (na imagem) o encontro das bordas do objeto e sua sombra, o observador se encontra na região 2, quando nesta região o objeto não oclui a sombra mas ambos estão próximos, deste modo podemos considerá-los externamente conectados (*NonOccludesEC*) (Figura 12b).

Na Figura 12c o observador está posicionado na região 3, pois a sombra está parcialmente oclusa pelo objeto, e na imagem o objeto aparece sobrepondo a sombra parcialmente. A região 4 tem sua superfície definida pela linha tangencial à fonte de luz e ao objeto, nela a sombra está a ponto de desaparecer. Note pela Figura 12d que a sombra está oculta pelo objeto na região 4, a qual é a fronteira onde a sombra irá desaparecer por completo se o agente estiver se movimentando na direção da região 5 (Figura 12e), ou reaparecerá se estiver indo na direção da região 3 (*TotallyOccludesTPPI*) (Figura 12c). Na região 5 estamos no mesmo campo de visão da fonte de luz, nesta região a sombra está projetada numa superfície totalmente oclusa pelo objeto e, portanto, não é possível vê-la (*TotallyOccludesNTPPI*) (Figura 12e).

Toda a teoria do PQRS é baseada no fato de que as sombras são corretamente relacionadas aos objetos. Portanto relacionar uma sombra com o objeto que a provocou é necessário para a aplicação correta da teoria, pois uma inferência com a sombra de outro objeto conduz ao erro. Uma solução automática (parcial) para este problema de correspondência é proposta a seguir.

2.2.3.2 Correspondência entre objeto e sombra

Uma sombra pode estar ou não conectada ao objeto que a projetou. Neste trabalho foram desconsiderados os casos em que um objeto suspenso no ar terá a sua sombra projetada desconectada do mesmo. Para formalização de situações em um ambiente de laboratório usamos uma hipótese simplificadora de que em nossos experimentos o objeto está imóvel sobre o chão. A sombra sempre estará conectada ao objeto que a provocou. Desta maneira, nas imagens, as regiões que representam o objeto e sua sombra estão externamente conectadas. A correspondência entre o objeto e sua sombra foi formalizada com base no PQRS, porém, vamos ver antes algumas restrições para a formalização.

A fórmula 2.16 garante que somente existe sombra s se houver o objeto o , uma superfície Scr e uma fonte de luz L , além desses elementos outras restrições devem ser observadas, são elas:

Figura 12: Como a sombra aparece quando somente o observador altera sua posição para: (a) região 1, (b) região 2, (c) região 3, (d) região 4 e (e) região 5 representadas no mapa da Figura 11b

Fonte:autor.

- a fonte de luz é única, e está em posição superior ao objeto e ao observador;
- os objetos estão sobre uma superfície fixa, e não flutuam;
- o observador estará sempre em posição vertical (em pé).

O objeto sempre está externamente conectado (*EC*) com sua sombra. É, portanto, conveniente representar o objeto em base e topo: $Base(x, z, \nu)$ significa “ x é base de z do ponto de vista de ν ” e $Top(y, z, \nu)$ significa “ y é topo de z do ponto de vista de ν ”. *Base/3* obedece aos seguintes axiomas:

$$\forall x \nu \neg Base(x, x, \nu), \quad (2.22)$$

isto é, um objeto completo x não é base dele próprio do ponto de vista ν ;

$$\begin{aligned} \forall x y \nu Base(x, y, \nu) \rightarrow \exists z (x \neq z) \& TTP(region(x), region(y)) \wedge \\ TTP(region(z), region(y)) \wedge EC(region(x), region(z)) \rightarrow Top(z, y, \nu), \end{aligned} \quad (2.23)$$

isto é, x e z são diferentes e ambos são tangenciais parte própria de y , ambos estando desconectados sendo x a base de y então z será o topo de y , esse axioma separa o objeto em duas sub partes: base e topo do ponto de vista de ν ;

$$\forall x y \nu [Base(x, y, \nu) \rightarrow \neg Top(x, y, \nu)], \quad (2.24)$$

isto é, caso uma parte x seja a base de y está não poderá ser o topo de y ;

$$\forall x y z \nu [[Base(x, y, \nu) \wedge Base(y, z, \nu)] \rightarrow Base(x, z, \nu)], \quad (2.25)$$

isto é, sendo x base de y e sendo y base de z então por transitividade x é base de z ;

$$\begin{aligned} \forall x y \nu Base(x, y, \nu) \rightarrow \exists z [P(region(z), region(x)) \wedge Base(z, y, \nu)] \wedge \\ \neg \exists u [P(region(u), region(x)) \wedge Top(u, y, \nu)], \end{aligned} \quad (2.26)$$

isto é, qualquer parte z de x será uma base de y também e não existe nenhuma parte u de x que será o topo y ;

$$\begin{aligned} \forall x y \nu Base(x, y, \nu) \wedge Top(z, y, \nu) \rightarrow P(image(x, \nu), image(y, \nu)) \wedge \\ P(image(z, \nu), image(y, \nu)) \wedge EC(image(x, \nu), image(z, \nu)), \end{aligned} \quad (2.27)$$

isto é, que a base x e topo z do objeto y aparecem na imagem como parte de y e externamente conectados. Assim para definir que determinada sombra s corresponde ao objeto o utilizamos a fórmula 2.28:

$$\begin{aligned} Corresponde(s, o, Scr, L) \rightarrow N(L, o, s) \wedge \exists b \nu Base(b, o, \nu) \wedge \\ NTPP(s, Scr) \wedge NTPP(b, Scr) \wedge EC(s, b). \end{aligned} \quad (2.28)$$

A fórmula (2.28) diz que s é sombra do objeto o quando:

- a sombra s está conectada externamente à base b do objeto o ;
- e a fonte de luz L se encontra mais próxima de o do que de s ;
- e que tanto a base b quanto a sombra s estejam ambas sobre a mesma superfície.

A proximidade da fonte de luz é representada pela primitiva de proximidade $N(L, o, s)$ a qual é lida como “a fonte de luz L está mais próxima do objeto o do que da sombra s ” (SANTOS; DEE; FENELON, 2009). Essa formalização é suficiente para fazer a correspondência da sombra com seu objeto nas situações consideradas neste trabalho.

Na Figura 13 podem ser observadas as restrições da fórmula (2.28). Nesta figura pelo menos parte da sombra e a base do objeto estão localizadas na mesma superfície (o piso). Na Figura 13a, cada objeto tem somente uma sombra externamente conectada à sua base e esta se encontra mais distante da fonte de luz. Já na Figura 13b um dos objetos tem duas sombras conectadas à sua base, neste caso a proximidade $N(L, o, s)$ servirá para confirmar qual sombra corresponde ao objeto. Na Figura 13c pode-se ver outro tipo ambiguidade, pois nela um dos objetos tem três sombras conectadas a ele, destas somente a sombra que ele provocou é verdadeira para a fórmula (2.28). Além disso, nesta imagem um dos objetos está ocluído por outro do ponto de vista do observador, assim na segmentação aparecem como um objeto com duas sombras conectadas, uma na base e outra não, a partir de situações como esta se pode deduzir a oclusão, como a ocorre na Figura 14.

A formalização da correspondência entre objeto e sombra serviu como base para criar uma função que determina o par sombra-objeto, que utilizamos no sistema de auto localização baseado na localização qualitativa descrita na seção 2.2.3.1. A teoria do PQRS foi utilizada anteriormente em um algoritmo de auto localização aplicado em um robô móvel (SANTOS; DEE; FENELON, 2008; SANTOS; DEE; FENELON, 2009), o qual neste trabalho foi estendido com técnicas probabilísticas. As teorias de raciocínio espacial qualitativo, além dos cálculos topológicos, lidam com direções. As relações de direções descrevem onde uma entidade espacial está posicionada relativa à outra entidade. A seguir vamos descrever o cálculo de direções cardinais.

2.2.4 Cálculo de direções cardinais

Cardinal direction calculus ou cálculo de direções cardinais (CDC) (FRANK, 1996) é um formalismo para raciocínio sobre direções entre objetos espaciais. O cálculo tem nove relações básicas: norte (*north* (n)), leste (*east* (e)), oeste (*west* (w)), sul (*south* (s)), noroeste (*northwest* (nw)), nordeste (*northeast* (ne)), sudeste (*southeast* (se)), sudoeste (*southwest* (sw)) e igual (*equal* (eq)) que corresponde a região central neutra (Figura 15). Em Frank (1996), essas regiões são resultados da divisão do espaço em duas abordagens diferentes, que pode ser em forma de cone ou em quadrantes. Para os dois sistemas Frank (1996) fez uma tabela de composição, e ao final concluiu que a forma de quadrante é melhor que a forma de cone por ter um número maior

Figura 13: Imagens capturadas no laboratório, onde podemos aplicar a fórmula de correspondência de sombra, note somente a sombra do objeto está externamente conectada à sua base e mais distante da fonte de luz.

Fonte: autor.

Figura 14: Por causa das sombras podemos inferir que há outro objeto totalmente ocluído pelo balde.

Fonte: Santos, Dee e Fenelon (2008).

Figura 15: Representação da projeção em quadrante, sendo os elementos representados como ponto (direito) e região (esquerda).

Fonte: Ligozat (1998).

de deduções exatas. Na Figura 15 encontra-se uma representação da abordagem em quadrantes para os objetos sendo representados como pontos ou como regiões.

A principal tarefa do CDC refere-se a inferir a direção entre dois objetos A e C , a partir do conhecimento da direção entre A e (outro objeto) B e entre B e C . Essa inferência vem da tabela de composição entre as nove relações básicas, onde o resultado pode ser um conjunto de possibilidades exclusivamente disjuntas (LIGOZAT, 1998). O conceito de vizinhança também é aplicável devido à topologia das relações definidas pela proximidade entre as regiões formadas, como por exemplo, a transição que pode ocorrer a partir da relação sudeste se é o subconjunto composto pelas relações s (sul), e (leste), se (sudeste) e eq (igual) (LIGOZAT, 1998). As relações de direção foram utilizadas nas tarefas de domínio de tráfego, e de auto localização com múltiplos objetos. No domínio de tráfego (seção 3.2) foi utilizado principalmente o cálculo de direção cardinal. Para auto localização do robô com múltiplos objetos a transição de relação teve um papel relevante, e o cálculo de distância foi utilizado junto com cálculo topológico ROC para formalizar a localização (seção 4.2).

As teorias de raciocínio espacial qualitativo aqui descritas serão utilizadas para inferir informações a partir de imagens. Nosso objetivo é criar um sistema que interprete imagens e as informações sejam utilizadas por um robô. Estando esse robô em ambiente real, o sistema terá que lidar com incertezas, as quais são causadas ou pelo ruído dos sensores ou pela falta de informação. Para lidar com incerteza optamos por agregar uma abordagem probabilística.

2.3 Modelos Probabilísticos

Uma finalidade da probabilidade em robótica é lidar com incerteza, estimando valores quantitativos do estado a partir de dados do sensor que não são diretamente observáveis, mas que podem ser inferidos. Dados dos sensores, controle e o estado do robô e de seu ambiente são quantidades modeladas como variáveis aleatórias. Uma variável aleatória Y pode carregar

informação sobre outra variável aleatória X , através de probabilidades condicionais (eq. (2.29)):

$$P(x|y) = P(X = x|Y = y). \quad (2.29)$$

A regra de Bayes relaciona uma probabilidade condicional à sua inversa de modo que:

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)}. \quad (2.30)$$

A regra de Bayes tem um papel importante em inferência probabilística em geral. Pois considerando que x é uma quantidade que se quer inferir a partir de y , a probabilidade $P(x)$ é referente a distribuição de probabilidade a priori, e y é chamado de dado de informação (como dados dos sensores). A probabilidade $P(x|y)$ é a distribuição de probabilidade posterior sobre X . A regra de Bayes oferece um modo conveniente para calcular a probabilidade posterior $P(x|y)$ utilizando a verossimilhança $P(y|x)$ e a probabilidade a priori $P(x)$. Ou seja, para inferir x a partir do dado y do sensor, a regra de Bayes permite especificar qual a probabilidade dos dados y assumindo que x ocorra. Em robótica, a probabilidade $P(y|x)$ é frequentemente criada como um modelo que descreve com algum nível de abstração como um determinado estado da variável X causa o valor de sensor Y (THRUN et al., 2005).

Também é importante notar que o denominador na regra de Bayes não depende de x , assim o fator $P(y)^{-1}$ é uma constante para qualquer valor de x na probabilidade posterior $P(x|y)$. Por isso a constante pode ser escrita como um fator de normalização, que em geral é denominado como η :

$$P(x|y) = \eta P(y|x)P(x). \quad (2.31)$$

Neste trabalho usaremos o símbolo η para indicar que o resultado foi normalizado para 1.

Podemos condicionar a regra de Bayes com uma variável aleatória arbitrária $Z = z$, como na equação (2.32):

$$P(x|y, z) = \frac{P(y|x, z)P(x|z)}{P(y|z)}. \quad (2.32)$$

Também podemos definir independência condicional através da igualdade.

$$P(x, y|z) = P(x|z)P(y|z). \quad (2.33)$$

Neste trabalho, utilizamos dois modelos: filtros Bayesianos e redes Bayesianas relacionais. A seguir começaremos descrevendo brevemente o filtro Bayesiano.

2.3.1 Filtros Bayesianos

Filtros Bayesianos são algoritmos que calculam a distribuição de probabilidades sobre crenças (*bel*) a partir dos dados medidos e do controle. Utiliza X_t para representar o estado atual, o qual será assumido como não observado, ou seja, o qual não se pode observar diretamente por nenhum sensor. Enquanto Z_t representa o conjunto de evidências medidas que podem ser

observadas. A observação no instante t é $Z_t = z_t$ para um conjunto de valores de z_t (STUART; PETER, 2010). Com um conjunto de variáveis de estado e evidência para um dado problema, o próximo passo é especificar como o mundo evolui (modelo de transição) e como as variáveis de evidência adquirem seus valores (modelo do sensor).

O modelo de transição determina a distribuição de probabilidade sobre os valores atuais das variáveis de estado, dados os seus valores prévios, o qual é representado por $P(x_t|x_{0:t-1})$. O conjunto $x_{0:t-1}$ não tem limites de tamanho com o crescimento de t , mas considerando a suposição de Markov - o estado atual depende somente de um número fixo e finito de estados anteriores. Ou seja, um estado tem informação suficiente para fazer com que o futuro seja condicionalmente independente do passado, e então temos que o modelo de transição é uma distribuição condicional (eq. (2.34)):

$$P(x_t|x_{0:t-1}) = P(x_t|x_{t-1}). \quad (2.34)$$

Em um modelo de transição as mudanças são governadas por regras que não mudam com o tempo, e assim são caracterizadas como uma tabela de probabilidades condicionais (STUART; PETER, 2010). Em nosso trabalho, os valores da tabela de transição se baseiam no conceito de transição das relações espaciais (seção 2.2.1). Em robótica quando o controle u_t tem influência no modelo de transição do estado, a distribuição condicional é dada por:

$$P(x_t|x_{t-1}, u_t). \quad (2.35)$$

Para o modelo de sensor, as variáveis de medidas z_t podem depender tanto das variáveis de evidências anteriores $z_{0:t-1}$ quanto da variável do estado atual x_t , mas qualquer estado atual deve ser o suficiente para gerar o valor atual do sensor. Deste modo, aplicando a suposição de Markov temos a equação (2.36):

$$P(z_t|x_t, z_{0:t-1}) = P(z_t|x_t), \quad (2.36)$$

assim, $P(z_t|x_t)$ é o modelo de sensor. Em nosso trabalho as evidências são as relações espaciais qualitativas percebidas a partir das imagens, por isso nosso modelo de sensor é chamado de modelo de imagem.

Um filtro Bayesiano tem duas etapas: a previsão ($\overline{bel}(x_t)$) e a atualização das medidas ($bel(x_t)$) (THRUN et al., 2005). Na etapa de previsão a crença sobre o estado atual de um robô é condicionado ao controle $u_{0:t}$ e todas as evidências anteriores $z_{0:t-1}$ sem a evidência medida z_t (eq. (2.37)):

$$\overline{bel}(x_t) = P(x_t|z_{0:t-1}, u_{0:t}). \quad (2.37)$$

Note que na previsão de medidas a medida do estado atual não foi considerada, ao incluir a medida do estado atual temos a segunda etapa que é a atualização das medidas ($bel(x_t)$) onde

o estado atual previsto é ponderado pela probabilidade de uma atuação de controle u_t e uma medida z_t realizada neste estado x_t (eq. (2.38)):

$$bel(x_t) = P(x_t|z_{0:t}, u_{0:t}). \quad (2.38)$$

O filtro Bayesiano é recursivo, e calcula a crença do estado x_t baseado na crença do estado anterior x_{t-1} e do controle u_t (quando existente). Na etapa de previsão a distribuição de crença $\overline{bel}(x_t)$ (previsão) do robô para o estado x_t é obtido pela soma do produto de duas distribuições: o modelo de transição $P(x_t|x_{t-1}, u_t)$ e a distribuição de crença $bel(x_{t-1})$ do estado anterior, deste modo temos que:

$$\begin{aligned} \overline{bel}(x_t) &= \sum_{x_{t-1}} P(x_t|x_{t-1}, u_t) bel(x_{t-1}) \\ &= \sum_{x_{t-1}} P(x_t|x_{t-1}, u_t) P(x_{t-1}|z_{0:t-1}, u_{0:t-1}). \end{aligned} \quad (2.39)$$

Na etapa de atualização das medidas o filtro Bayesiano multiplica a distribuição de crença prevista $\overline{bel}(x_t)$ pelo modelo de sensor para a evidência z_t observada. Isto é realizado para cada hipotético estado posterior x_t . O valor encontrado ainda não é um valor de probabilidade, o qual deve ser normalizado por uma constante de normalização η . Isso resulta na distribuição de crença final do estado $bel(x_t)$ no instante t (form. (2.40)). Ao iniciar um algoritmo de filtro Bayesiano a distribuição de crença $bel(x_0)$ para o instante t_0 em geral recebe valores uniformes, caso não exista informação sobre valores da distribuição de crença inicial.

$$\begin{aligned} P(x_t|z_{0:t}, u_{0:t}) &= \eta P(z_t|x_t) \sum_{x_{t-1}} P(x_t|x_{t-1}, u_t) P(x_{t-1}|z_{0:t-1}, u_{0:t-1}) \\ bel(x_t) &= \eta P(z_t|x_t) \overline{bel}(x_t). \end{aligned} \quad (2.40)$$

Neste trabalho o filtro Bayesiano foi utilizado na tarefa de auto localização qualitativa do robô. As evidências medidas são as relações espaciais qualitativas percebidas através de imagens, e o estado representa a região onde o robô está localizado, e não se utilizou a variável de controle do robô. A próxima seção tratará das redes Bayesianas relacionais.

2.3.2 Redes Bayesianas relacionais

Redes Bayesianas são estruturas de grafos direcionados acíclicos no qual os nós representam variáveis aleatórias X_i e os arcos representam a dependência entre as variáveis conectadas (PEARL, 1988). Aciclicidade é uma característica da rede Bayesiana. Outra característica das redes Bayesianas é que cada nó X tem uma distribuição de probabilidade condicional (DPC) $dpc(X)$ associada a ele. Uma DPC é definida pelos pais do nó X ($Pa(X)$), deste modo $dpc(X) = P(X|Pa(X))$. Para construir uma rede Bayesiana, a condição de Markov é suposta.

A condição de Markov diz que um nó é condicionalmente independente de todos os outros nós na rede, dados seus pais, filhos e pais dos filhos (STUART; PETER, 2010). Com uma

rede Bayesiana, pode-se raciocinar sobre o domínio, inferindo informações quando observado o valor de alguma variável. O processo de inferência calcula a distribuição de probabilidade posterior, para um conjunto de nós de consulta, dado os valores para alguns nós de evidência (KORB; NICHOLSON, 2003).

Uma rede Bayesiana relacional para um vocabulário S é dada por uma rede Bayesiana que contém um nó para toda relação $r \in S$. Os nós para relações n -ária $r \in S$ estão associados com uma fórmula de probabilidade $F_r(x_1, \dots, x_n)$ sobre todos os símbolos dos pais de r ($Pa(r)$) (JAEGER, 1997). A fórmula de probabilidade indica como calcular a probabilidade de $r(v)$ para qualquer elemento v no domínio D . Quando uma inferência é requerida para alguma relação $r(v)$, ou seja, uma requisição para a probabilidade de $r(v)$ condicionada por algum evento observado, uma rede Bayesiana proposicional é gerada; cada nó desta rede representa a instanciação de uma relação (COZMAN et al., 2004). O método de redes Bayesianas relacionais foi desenvolvido por Jaeger (1994b) e está incluso no pacote de domínio público do sistema Primula ⁶ (JAEGER et al., 2006). Redes Bayesianas relacionais fazem parte do mecanismo de inferência da lógica de descrição probabilística Credal $\mathcal{ALC}$ descrita a seguir.

2.4 Lógica de Descrição Probabilística Credal $\mathcal{ALC}$

A lógica credal $\mathcal{ALC}$ (CR $\mathcal{ALC}$) (COZMAN; POLASTRO, 2008) é uma extensão probabilística da lógica de descrição $\mathcal{ALC}$ (SCHMIDT-SCHAUSS; SMOLKA, 1991). CR $\mathcal{ALC}$ integra a eficiência no tratamento de incertezas das redes Bayesianas, com a representação relacional da lógica de descrição (COZMAN; POLASTRO, 2008). Nessa seção iremos falar sobre o CR $\mathcal{ALC}$, mas antes vamos começar por um breve relato da lógica de descrição e algumas extensões probabilísticas.

2.4.1 Lógica de descrição

Lógicas de descrição compartilham algumas características, sendo baseadas em *conceitos*, como Planta e Animal; em *papéis*, como paiDe e amigoDe; em *indivíduos*, como Manuel e Maria. Conceitos e papéis podem ser entendidos como relações unárias e binárias respectivamente. Intersecção, união e complementação de conceitos são usualmente permitidas, bem como formas restritas de quantificação com uso de papéis. A inferência básica em lógicas de descrição é: será o conceito A um conjunto vazio de indivíduos, ou haverá um indivíduo em A ? Uma detalhada discussão de aplicações de lógicas de descrição, bem como de seus aspectos teóricos, pode ser encontrada no *Description Logic Handbook*, (BAADER et al., 2003).

Lógicas de descrição não representam incertezas, sejam essas incertezas decorrentes de ambiguidade, de falta de conhecimento, falta de precisão, ou de variações estatísticas. Por isso, um significativo número de propostas pode ser encontrado na literatura, onde lógicas de

⁶ <http://people.cs.aau.dk/~jaeger/Primula/>

descrição são combinadas com raciocínio sobre incerteza. É significativa a existência de um *Workshop on Uncertain Reasoning for the Semantic Web*, ligado à *Semantic Web Conference*.

Lógicas de descrição com representação de incertezas são por vezes baseadas em lógica difusa (STRACCIA, 2005), por vezes baseadas na teoria de Dempster-Shafer (NIKOLOV et al., 2007), e por vezes baseadas na teoria de probabilidades. Além da literatura em lógicas de descrição probabilísticas propriamente ditas, existem muitos trabalhos recentes sobre bases de dados dedutivas (LAKSHMANAN; SADRI, 1994; LAKSHMANAN; SHIRI, 2001; PARSONS, 1996) e ontologias com suporte para probabilidades (FUKUSHIGE, 2004; HOLI; HYVÖNEN, 2005; HUNG; GETOOR; SUBRAHMANIAN, 2007; MITRA; NOY; JAISWAL, 2005; POOL; AIKIN, 2004; UDREA; SUBRAHMANIAN; MAJKIC, 2006; YANG; CALMET, 2005).

Heinsohn (1994), Jaeger (1994a), Jaeger (1994b) e Sebastiani (1994b) propuseram inclusões probabilísticas em diferentes lógicas. As lógicas probabilísticas propostas por estes autores consideravam que probabilidades são expressas sobre o domínio, ou seja, a álgebra que recebe a medida de probabilidades é formada por subconjuntos do domínio.

Dürig e Studer (2005) adotaram a lógica $\mathcal{ALC}$ sem inclusões probabilísticas, porém considerando probabilidades independentes sobre indivíduos. Lukasiewicz (2008) propõe outra estratégia, adotar uma relação de consequência não-monotônica (GIUGNO; LUKASIEWICZ, 2002; LUKASIEWICZ, 2008) e permitindo com isso tanto inclusões probabilísticas quanto probabilidades sobre asserções. As propostas de Lukasiewicz (2008) são as lógicas de descrição probabilísticas mais expressivas na literatura, tendo como desvantagem sua semântica extremamente particular.

2.4.2 Lógica de descrição probabilística

As lógicas de descrição probabilísticas descritas acima compartilham o uso de probabilidade sobre o domínio. Nenhuma das lógicas garante que um único número real pode ser atribuído a uma probabilidade $P(C(a)|D(b))$, onde C e D são conceitos e a e b são indivíduos. Esse tipo de situação é muito comum em lógica probabilística (HALPERN, 2003; HANSEN; JAUMARD, 1996). A manipulação de *conjuntos* de probabilidades é motivada pelo interesse em realismo, flexibilidade e generalidade matemática, e também pela dificuldade em garantir unicidade quando probabilidades podem ser livremente associadas às fórmulas.

Outra característica das lógicas de descrição já mencionadas é que nenhuma pode expressar julgamentos de independência estocástica. Isso se deve, fundamentalmente, à dificuldade computacional de manipular simultaneamente conjuntos de probabilidades e restrições não lineares representadas por julgamentos de independência. Nas últimas décadas uma série de lógicas probabilísticas foram propostas com uma abordagem diferente: relações de independência são usadas para *reduzir* a complexidade computacional, ao serem usadas para garantir unicidade de probabilidades e para estruturar fórmulas (FRIEDMAN; KOLLER; PFEFFER, 1999; FUHR,

1995; KOLLER; PFEFFER, 1998; JAEGER, 1997; POOLE, 1993; NGO; HADDAWY, 1997; WELLMAN; BREESE; GOLDMAN, 1992).

Em particular, tem havido esforço significativo para combinar construções lógicas com redes Bayesianas ou campos de Markov (GETOOR; TASKAR, 2007; MILCH; RUSSELL, 2007). Essa abordagem pode também ser encontrada em várias lógicas de descrição probabilísticas. A lógica P-CLASSIC usa uma coleção de redes Bayesianas como semântica para qualquer conjunto de fórmulas (KOLLER; PFEFFER, 1997). Novamente, probabilidades são impostas sobre o domínio e asserções não podem ser associadas a probabilidades. O mesmo ocorre com a *Tiny Description Logic* de Yelland (1999), com a linguagem BayesOWL de Ding e Peng (2004), Ding, Peng e Pan (2006), e com a lógica proposta por Staker (2002).

Uma abordagem diferente é usada por Laskey e Costa (2012): sua proposta é usar fragmentos de redes Bayesianas para representar certos blocos fundamentais em lógica de primeira-ordem, e combinar esses blocos conforme a terminologia de interesse. Essa abordagem foi bastante geral, com inclusões e asserções probabilísticas — no entanto a dificuldade é que um algoritmo de inferência não foi derivado. Uma abordagem similar foi adotada por Nottelmann e Fuhr (2004), Nottelmann e Fuhr (2006), os quais partem de uma versão probabilística de Prolog e a especializam para uma lógica de descrição restrita.

As lógicas de Laskey e Costa (2012), Nottelmann e Fuhr (2004), Nottelmann e Fuhr (2006), são relevantes por usarem, diferentes das demais, uma semântica baseada em probabilidades sobre interpretações. Isso significa que, em lugar de impor probabilidades sobre o domínio, a lógica impõe probabilidades sobre as possíveis interpretações $\mathcal{I}$. Uma semântica baseada em probabilidades sobre interpretações pode facilmente lidar com probabilidades como $P(C(a)|D(b))$; para calcular $P(C(a)|D(b))$, devemos calcular a probabilidade de encontrar $C(a)$ e $D(b)$ entre todas as interpretações para as quais vale $D(b)$.

Uma semântica deve ser precisa para evitar o quanto possível problemas de ambiguidade na base de conhecimento (HUMMEL, 2009). Uma semântica baseada em interpretações pode interpretar uma inclusão probabilística (form. (2.41)), de pelo menos duas maneiras. Primeiro, como uma afirmação probabilística sobre cada indivíduo (form. (2.42)). Segundo, como uma afirmação sobre validade de sentenças (form.(2.43)):

$$P(C|D) \in [\alpha, \beta]; \quad (2.41)$$

$$\forall x : P(C(x)|D(x)) \in [\alpha, \beta]; \quad (2.42)$$

$$P(\forall x : C(x) | \forall y : D(y)) \in [\alpha, \beta]. \quad (2.43)$$

A segunda interpretação (form.(2.43)) é coerente com uma variedade grande de lógicas probabilísticas que empregam probabilidades sobre sentenças (na literatura em inteligência artificial, iniciadas por Nilsson (1986)).

A lógica de descrição $\mathcal{ALC}$ é uma das mais utilizadas (SCHMIDT-SCHAUSS; SMOLKA, 1991). Nela um conceito pode ser definido recursivamente através de *conjunção* ($C \sqcap D$), *disjunção* ($C \sqcup D$) e *negação* ($\neg C$) de outros conceitos, bem como *restrição existencial* ($\exists r.C$) e *restrição de valor* ($\forall r.C$) usando papéis e conceitos. Uma *inclusão*, $C \sqsubseteq D$, indica que o conceito D contém o conceito C . Uma *definição*, $C \equiv D$, indica que os conceitos C e D são equivalentes. Um conjunto de inclusões e definições é uma *terminologia*; comumente se assume que uma terminologia é acíclica (um conceito não pode se referir a ele mesmo através de uma sequência de inclusões e definições). A semântica é dada por um conjunto $\mathcal{D}$ chamado *domínio* e um *mapa* $\mathcal{I}$ que associa cada indivíduo a um elemento de $\mathcal{D}$, cada conceito a um subconjunto de $\mathcal{D}$, e cada papel a uma relação binária em $\mathcal{D} \times \mathcal{D}$ de forma que:

- $\mathcal{I}(C \sqcap D) = \mathcal{I}(C) \cap \mathcal{I}(D)$,
- $\mathcal{I}(C \sqcup D) = \mathcal{I}(C) \cup \mathcal{I}(D)$,
- $\mathcal{I}(\neg C) = \mathcal{D} \setminus \mathcal{I}(C)$,
- $\mathcal{I}(\exists r.C) = \{x \in \mathcal{D} \mid \exists y : (x, y) \in \mathcal{I}(r) \wedge y \in \mathcal{I}(C)\}$,
- $\mathcal{I}(\forall r.C) = \{x \in \mathcal{D} \mid \forall y : (x, y) \in \mathcal{I}(r) \rightarrow y \in \mathcal{I}(C)\}$.

Assim, uma inclusão $C \sqsubseteq D$ é satisfeita se e somente se $\mathcal{I}(C) \subseteq \mathcal{I}(D)$, e uma definição $C \equiv D$ é satisfeita se e somente se $\mathcal{I}(C) = \mathcal{I}(D)$. Por exemplo, $C \sqsubseteq (\exists \text{ temParente.Mulher}) \sqcap (\forall \text{ compraDe.}(\text{Peixe} \sqcup \text{Fruta}))$ indica que C contém indivíduos que tem um parente mulher e que compram frutas ou peixes (ou ambos).

Nos sistemas baseados em lógica de descrição, a base de conhecimento é formada por dois elementos, a $\mathcal{TBox}$ e a $\mathcal{ABox}$. A $\mathcal{TBox}$ introduz a terminologia, ou seja, o vocabulário aplicado a um domínio. Vocabulários são os conceitos, indivíduos e papéis. A $\mathcal{ABox}$ contém os dados que serão inseridos sobre indivíduos em termos de conceitos e papéis. As declarações na $\mathcal{TBox}$ e $\mathcal{ABox}$ podem ser identificadas por fórmulas em lógica de primeira ordem. Além de armazenar terminologias e inserções, a lógica de descrição também raciocina sobre eles (BAADER; NUTT, 2002).

2.4.3 Credal $\mathcal{ALC}$ - $\text{CR}\mathcal{ALC}$

O vocabulário básico do $\text{CR}\mathcal{ALC}$, assim como do $\mathcal{ALC}$, contém indivíduos, conceitos (conjunto de indivíduos) e papéis (relações binárias entre indivíduos), conforme descrito na Seção 2.4.1. Na versão probabilística do $\mathcal{ALC}$ ($\text{CR}\mathcal{ALC}$), do lado esquerdo das inclusões ou definições aparecem somente os conceitos. Dado um nome de conceito C , um conceito D e um

nome de papel r , as seguintes asserções probabilísticas são possíveis:

$$P(C) \in [\underline{\alpha}, \bar{\alpha}], \quad (2.44)$$

$$P(C|D) \in [\underline{\alpha}, \bar{\alpha}], \quad (2.45)$$

$$P(r) \in [\underline{\beta}, \bar{\beta}]. \quad (2.46)$$

Escrevemos $P(C|D) = \alpha$ quando $\underline{\alpha} = \bar{\alpha}$, $P(C|D) \geq \underline{\alpha}$ quando $\underline{\alpha} < \bar{\alpha} = 1$, e assim por diante. Para garantir a aciclicidade, não é permitido a nenhum conceito utilizar a si mesmo em inclusões ou definições determinísticas (ou probabilísticas). A fórmula (2.44) pode ser descrita assim: para qualquer x pertencente à $\mathcal{D}$, a probabilidade que x pertença à interpretação de C está no conjunto $[\underline{\alpha}, \bar{\alpha}]$ (form. (2.47)):

$$\forall x \in \mathcal{D} : P\left(\left\{\mathcal{I} : x \in \mathcal{I}(C)\right\}\right) \in [\underline{\alpha}, \bar{\alpha}]. \quad (2.47)$$

Informalmente, podemos representar a semântica como $\forall x \in \mathcal{D} : P(C(x)) \in [\underline{\alpha}, \bar{\alpha}]$. A semântica das fórmulas (2.45) e (2.46) é dado por (form. (2.48) e form. (2.49) respectivamente):

$$\forall x \in \mathcal{D} : P(C(x)|D(x)) \in [\underline{\alpha}, \bar{\alpha}], \quad (2.48)$$

$$\forall (x, y) \in \mathcal{D} \times \mathcal{D} : P(r(x, y)) \in [\underline{\beta}, \bar{\beta}]. \quad (2.49)$$

Dado um domínio finito, um conjunto de sentenças em CRALC especifica probabilidades sobre todos os conceitos e papéis instanciados. Em geral, um conjunto de probabilidades é determinado pelo conjunto de sentenças, por exemplo, se for especificado o intervalo de valor de probabilidade de C como $P(C) \in [0, 2; 0, 3]$, quer dizer que todos os valores dentro do intervalo são permitidos. Poucas suposições garantem que uma única distribuição de probabilidade seja especificada para um conjunto de sentenças: nome-único, probabilidade-pontual na avaliação, rigidez dos nomes (COZMAN; POLASTRO, 2008). Com essa suposição um domínio finito e um conjunto de sentenças especificam uma única rede Bayesiana sobre todos os conceitos e papéis instanciados. E para verificar a probabilidade de um papel ou conceito instanciado em particular, pode-se gerar essa rede Bayesiana e então usá-la para inferências probabilísticas (SANTOS et al., 2010; FENELON et al., 2010).

Com CRALC pode-se especificar uma rede Bayesiana relacional com elementos de lógica. O algoritmo do CRALC gera uma rede Bayesiana a partir da $\mathcal{T}\text{Box}$ onde os conceitos são definidos em lógica de descrição $\mathcal{ALC}$ probabilística. Para formar a rede os conceitos são traduzidos como nós de tal modo que os conceitos à esquerda são pais dos conceitos à direita. As asserções probabilísticas (form.(2.44), form.(2.45) e form.(2.46)) geram uma rede de forma direta.

Uma definição $A \equiv C$ é traduzida para $P(A|C) = 1$ e $P(A|\neg C) = 0$, ou seja, se A e C são idênticos, logo dado C então A (Figura 16(a)). A inclusão $A \sqsubseteq C$ equivale a $P(A|\neg C) = 0$, se A está em C logo dado $\neg C$ então $\neg A$, nada diz diretamente sobre a probabilidade $P(A|C)$,

Figura 16: Exemplo de representação CRALC em rede Bayesiana da inclusão e da definição.

Fonte: Imagem modificada de Polastro (2012).

Figura 17: Exemplo de redes e tabelas de probabilidades condicionais geradas pelos construtores lógicos

Fonte: Polastro (2012).

um valor α deve ser definido para está probabilidade $P(A|C) = \alpha$ (Figura 16(b)). Seguindo esse raciocínio na Figura 17 apresenta um exemplo das distribuições de probabilidades condicionais (DPC, definido anteriormente) para $P(E|A \sqcup \neg B) = 0.8$, $F \equiv A \sqcup \neg B$ e $G \sqsubseteq A \sqcup \neg B$ onde as interrogações são probabilidades desconhecidas cujo valor deve ser definido.

As restrições existenciais $\exists r.C$ são escritas na forma de $\bigcup_{i=1}^n r_i \cap C_i$. Ou seja, como $\exists r(x, y).C(y) \equiv \exists y r(x, y) \sqcap C(y)$ que pode ser escrita como uma disjunção de conjunções $((r(x, 0) \sqcap C(0)) \sqcup (r(x, 1) \sqcap C(1)) \sqcup \dots \sqcup (r(x, n) \sqcap C(n)))$. A representação Bayesiana equivalente à $A(x) \sqsubseteq \exists r(x, y).C(y)$ é a $P(A(x)|r(x, y) \sqcap C(y)) = \alpha$ e $P(A(x)|\neg(r(x, y) \sqcap C(y))) = 0$ (Figura 18).

Para as restrições universais $\forall r.C$ são escritas na forma $\bigcap_{i=1}^n r_i \rightarrow C_i$. Ou seja, como $\forall r(x, y).C(y) \equiv \forall y r(x, y) \rightarrow C(y) \equiv \forall y \neg r(x, y) \sqcup C(y)$ que pode ser escrito como uma conjunção de disjunções $((\neg r(x, 0) \sqcup C(0)) \sqcap (\neg r(x, 1) \sqcup C(1)) \sqcap \dots \sqcap (\neg r(x, n) \sqcup C(n)))$. A representação Bayesiana equivalente à $A(x) \sqsubseteq \forall r(x, y).C(y)$ é $P(A(x)|\neg r(x, y) \sqcup C(y)) = \alpha$ e $P(A(x)|\neg(\neg r(x, y) \sqcup C(y))) = 0$ (Figura 19). Em ambos α é o valor da probabilidade que deve ser definido. Esse é o modo como o CRALC traduz as asserções lógicas de conceitos e papéis, e

Figura 18: Exemplo de representação CRALC em rede Bayesiana da restrição de existencial.

Fonte: Imagem modificada de Polastro (2012).

Figura 19: Exemplo de representação CRALC em rede Bayesiana da restrição universal.

Fonte: Imagem modificada de Polastro (2012).

foi utilizado como base para a construção manual de uma rede Bayesiana relacional com papéis de relações de direções cardinais neste trabalho (seção 3.3.2).

Como exemplo considerando uma terminologia τ com conceitos A , B , C e D . Supondo que:

$$\begin{aligned}
 P(A) &= \alpha_1, \\
 B &\sqsubseteq A, \\
 P(D | \forall r.A) &= \alpha_2, \\
 C &\sqsubseteq B \sqcup \exists r.D.
 \end{aligned}$$

Na figura 20 os grafos são exemplos dessa terminologia τ em rede Bayesiana e com a sua configuração para um domínio com dois indivíduos.

Ao escrever o domínio em CRALC algumas restrições devem ser observadas. Aciclicidade é uma restrição para as lógicas probabilísticas que utilizem rede Bayesiana. Em CRALC os conceitos primitivos são definidos com uma probabilidade a priori. Caso a probabilidade a priori seja desconhecida, adota-se uma probabilidade neutra de 50%. Do mesmo modo são definidos

Figura 20: Exemplo gráfico de uma rede gerada a partir de uma terminologia τ , e sua característica para um domínio com dois indivíduos $[a, b]$.

Fonte: Cozman e Polastro (2008).

os papéis (relação entre dois indivíduos). A partir das definições dos conceitos primitivos e dos papéis, novos conceitos podem ser criados. Ou seja, com esses conceitos uma hierarquia de objetos pode ser formada. Um domínio pode ser definido a partir dos conceitos anteriores e das restrições. Em Polastro et al. (2010) uma ontologia definindo domínios de ambientes foi desenvolvida com *CRALC*. Uma formalização similar foi utilizada no reconhecimento de objetos baseado em contexto (FENELON-PEREIRA; SANTOS; COZMAN, 2012).

2.5 Localização Robótica

Estamos presenciando um movimento do emprego da tecnologia robótica das fábricas para aplicações em ambientes domésticos, escritórios e lugares públicos. Nesses ambientes em geral as tarefas a serem executadas exigem agentes robóticos que se desloquem com segurança (KRUSE et al., 2013). Assim capacidade de navegar pelo ambiente é importante em robótica móvel. Uma navegação autônoma necessita de habilidades tais como, localização, planejamento, construção ou interpretação de mapa e controle de movimento (SIEGWART; NOURBAKHS, 2004; NEHMZOW, 2003). A localização robótica consiste em estimar as coordenadas do robô relativas ao ambiente externo. O robô tem as informações do mapa e deve se localizar nele utilizando os dados dos sensores. Neste trabalho o termo localização qualitativa refere-se ao método de localização no qual o mapa é baseado em características qualitativas do domínio, diferente do mapa métrico.

Em uma das primeiras referências sobre abordagem para localização qualitativa, um mapa topológico foi construído onde as bordas das regiões eram definidas por um conjunto de linhas conectando pares de pontos de referências (LEVITT; LAWTON, 1990). Esta ideia serviu de inspiração para várias representações espaciais. Dentre elas a representação espacial chamada panorama, que restringe a localização de um ponto relativamente a ordenação visual de objetos observados em torno do robô (SCHLIEDER, 1996). A noção da navegação qualitativa é definida em Schlieder (1996), pelas alterações na ordenação das informações, que ocorre quando o observador atravessa cada uma das linhas virtuais definidas pelos objetos de referência. Essa

mesma ideia foi aplicada a um domínio da *RoboCup* e está relatada em (WAGNER; VISSER; HERZOG, 2004).

Para a construção de mapas qualitativos Santos, Dee e Fenelon (2008) apresentou a teoria PQRS em que o mapa qualitativo divide o espaço em cinco regiões baseado em uma formalização lógica de oclusão e observação de sombras, Fogliaroni et al. (2009) propôs um modelo de visibilidade e oclusão de objetos para gerar as regiões em espaço navegável. McClelland, Campbell e Estlin (2013) introduziu um método para a construção de mapa utilizando as posições relativas dos pontos de referência. Estes métodos são geralmente baseados em algum tipo de conceptualização humana do espaço, alguns deles fundamentados em descobertas recentes em psicologia cognitiva (WOLTER; FREKSA; LATECKI, 2008). No entanto, eles são incapazes de lidar com o ruído do sensor, o que limita a sua aplicabilidade a domínios bem controlados e deterministas. Um panorama mais completo de representações qualitativas para localização do robô é apresentado em (WOLTER; FREKSA; LATECKI, 2008).

Os métodos de localização de maior sucesso em robótica são baseados em algoritmos probabilísticos. Algoritmos de localização probabilísticos são geralmente definidos como filtros Bayesianos (THRUN et al., 2005). Como já vimos, na seção 2.3.1, o filtro Bayesiano é um algoritmo recursivo que consiste em duas partes: a previsão e atualização das medidas. Esta definição básica resume obras sobre localização de Markov (FOX; BURGARD; THRUN, 1999) e os procedimentos de localização que utilizam filtros de Kalman (CHEN, 2012). Este último é um caso especial da localização de Markov, onde os modelos de sensores de movimento são funções Gaussianas lineares. Métodos conhecidos como algoritmos de Monte Carlo (DELLAERT et al., 1999) estendem a localização básica de Markov por aplicação de um filtro de partículas para representar a distribuição dos estados possíveis. Dado o movimento do robô as partículas são atualizadas de acordo com a previsão do próximo estado e são redimensionadas de acordo com a percepção do agente. Essas ideias têm uma alta aceitação no campo da robótica (DELLAERT et al., 1999). Portanto, um levantamento completo das obras recentes relacionadas com algoritmos de localização probabilísticos é uma tarefa difícil. Pesquisas das abordagens iniciais para localização robótica probabilística podem ser encontrados em Thrun (2002), Thrun et al. (2005).

2.6 Interpretação de Imagens

A primeira abordagem lógica para interpretação de imagens foi proposta por Reiter e Mackworth (1989). Nesse arcabouço, três conjuntos de axiomas são propostos, restringindo assim o processo de interpretação de imagens. Dessa forma, interpretação de imagens torna-se um procedimento de satisfação de restrições. O sistema SIGMA (MATSUYAMA; HWANG, 1990) utiliza-se das ideias propostas por (REITER; MACKWORTH, 1989) e, adaptando-as a um procedimento de raciocínio baseado em abdução de hipóteses (POOLE; GOEBEL; ALELIUNAS, 1987) gera descrições de imagens aéreas. Algumas das propriedades do formalismo subjacente a

este sistema foram posteriormente exploradas por Schroeder e Neumann (1996), e revisitadas e incorporadas em lógica descritiva por Neumann e Möller (2008) na década seguinte.

Em Shanahan (1996), um formalismo lógico é proposto para definir rigorosamente a transição das informações provenientes dos sensores de um robô móvel para símbolos representando a existência, localização e a forma dos objetos observados. Essa proposta, porém, trata da interpretação de cenas observadas de um ponto de vista global, em que os objetos sempre permanecem visíveis. Por outro lado, os trabalhos reportados em Santos e Shanahan (2002b), Santos e Shanahan (2003), Santos (2007) consideram mudanças em um ambiente dinâmico, representado a partir de um ponto de vista egocêntrico.

Partindo de princípios análogos, Hazarika e Cohn (2002) descrevem um sistema para a assimilação de cenas utilizando o conceito de *histórias espaços-temporais* (isto é, regiões espaços-temporais representando o desenvolvimento temporal de relações topológicas entre objetos físicos). Esse modelo é baseado em uma proposta anterior para a construção automática de modelos de eventos a partir de informações visuais (FERNYHOUGH; COHN; HOGG, 2000). O trabalho proposto em Bennett, Cohn e Magee (2005) avalia múltiplas possibilidades de histórias para explicar uma dada sequência de imagens, e por um critério de votação elege ao final a história mais consistente com a interpretação da sequência. Em contraste à utilização de histórias, em Santos et al. (2008) aplica-se um formalismo de ações, baseado em uma lógica de processos (BONNER; KIFER, 1994) cuja semântica espelha o processo de transições de estados ocorrido nas observações dos objetos em cenas dinâmicas.

Pesquisas na área de avaliação de sequências de imagens (*image sequence evaluation* (NAGEL, 2000)) utilizam-se de algumas definições de raciocínio probabilístico, porém, por não mencionarem formalismos lógicos, sofrem com a falta de definições rigorosas de semântica e métodos de inferências. Neste trabalho um formalismo em raciocínio espacial qualitativo foi utilizado para extrair das imagens capturadas por um robô informações espaciais que possibilita inferir a localização do robô enquanto navega pelo ambiente. A sequência de imagens foi analisada com filtro Bayesiano utilizando informações das relações espaciais do domínio (capítulo 4) (FENELON et al., 2013).

Em um cenário de tráfego, por exemplo, o reconhecimento de faixa é uma tarefa realizada por imagens. As pesquisas em reconhecimento de faixa têm se concentrado tradicionalmente, em estimar as propriedades geométricas das faixas em frente do veículo utilizando equipamentos de imagens (MCCALL; TRIVEDI, 2006). Em Hummel (2009) encontramos uma tentativa de inferir propriedades funcionais das faixas, tal como a direção permitida da rua (por exemplo: ir para cima ou para baixo, relativo a um sistema de coordenadas fixo), a permissão de mudar de direção, ou interação com outros veículos. Avaliação destes dados funcionais para todas as faixas são realizadas utilizando lógica de descrição e raciocínio espacial a partir dos dados coletados de sensores *on-board* no veículo (HUMMEL, 2009). Neste trabalho, parte desse domínio foi utilizado para formalizar com *CRALLC* um algoritmo capaz de inferir em qual faixa o veículo

estava localizado e quais as mãos de direção de todas as faixas (SANTOS et al., 2010; FENELON et al., 2010). A seguir falaremos brevemente das técnicas de visão computacional utilizadas neste trabalho.

2.7 Visão Computacional

Uma imagem digital é composta de um número finito de elementos conhecidos como *pixels*, cada *pixel* tem uma localização na imagem e um valor particular. Portanto, podemos considerar a imagem digital como uma função $f(x, y)$ onde x e y são coordenadas espaciais do *pixel*. A amplitude da função f em *pixel* é o valor discretizado (finito) que chamamos de intensidade ou grau de cinza. Em geral as imagens em grau de cinza possuem somente um canal de informação com seus valores variando de zero a 255, e as imagens coloridas consistem de três canais de informação.

A visão computacional pode ser definida por níveis de processo. Os processos de visão do nível baixo são caracterizados por melhorar a qualidade da imagem e envolvem operações tais como redução de ruídos, aumento do contraste, ou destaque de algum atributo. O processamento de imagem de nível médio tem como característica o fato das entradas geralmente serem imagens e as saídas serem atributos extraídos dessas imagens, tais como bordas, contornos e identificações de características individuais. As tarefas executadas neste nível são segmentação, classificação dos objetos individuais, entre outros. Finalmente, o processamento de nível alto envolve o reconhecimento de objetos coerentemente, como em análise de imagem executada por funções cognitivas (GONZALEZ, 2001).

2.7.1 Espaço de cores em imagens digitais

As imagens digitais utilizam modelos de representação de cores conhecidos como espaços de cores. Pelas características de hardware e dos tipos de dados analisados serão utilizados, neste trabalho, os espaços de cores RGB (*red, green, blue*) e HSV (*hue, saturation, value*).

O modelo RGB tem três canais de informação de cores: vermelho (255,0,0), verde (0,255,0) e azul (0,0,255). É baseado em um sistema de coordenadas cartesianas, e seu subespaço de cores é representado por um cubo (conforme mostrado na Figura 21). Esse cubo é um sólido composto de $(2^8)^3 \approx 16.000.000$ de cores. Podemos simplificar dizendo que RGB é bom para geração de imagens coloridas, porém é limitado para descrição de cores (GONZALEZ, 2001).

O modelo HSV ou HSI é baseado em um sistema de coordenadas cilíndricas, composto pela cromaticidade representadas pelo matiz e a saturação, e pelos componentes de intensidade ou brilho. As três propriedades deste modelo são representadas por um cone conforme Figura 22. O matiz (*hue*) **H** é definido pelo o ângulo entre 0 e 2π relativo ao eixo do cone, matiz igual à zero representa o vermelho puro, $2\pi/3$ o verde puro e $4\pi/3$ o azul puro. A saturação **S** varia de 0 a 1 a partir do centro do cone e representa a tonalidade da cor, portanto alteração na saturação causa

Figura 21: Representação gráfica do cubo de cores RGB, o cone da esquerda representa como uma cor(80,200,130) está localizada neste sistema cartesiano.

Fonte: http://upload.wikimedia.org/wikipedia/commons/0/03/RGB_farbwuerfel.jpg último acesso em 20/02/2009.

Figura 22: Representação gráfica do cone HSV.

Fonte: http://4.bp.blogspot.com/_7YvslLFdkCI/TO44RD7Q24I/AAAAAAAAAFs/MiAsTPjSNTA/s1600/450px-HSV_Color_Cone.png último acesso em 29/01/2014.

alteração de tonalidade de cor, como exemplo, a cor rosa é um vermelho com uma saturação baixa. A intensidade ou valor **V** varia no eixo vertical do cone e representa o brilho, variações em **V** são percebidas como sombreamento, quando o valor de **V** é zero, a imagem aparece como preto, independente da intensidade dos canais de cromaticidade (GONZALEZ, 2001; SHAPIRO; STOCKMAN, 2001).

Uma consequência destes canais de informação em HSV é que eles são ideais para o desenvolvimento de algoritmos, pois a separação em canais facilita a aplicação de filtros para a separação de características (GONZALEZ, 2001; SHAPIRO; STOCKMAN, 2001). A transformação de um modelo de cor RGB para HSV e vice-versa é não linear, por brevidade não detalharemos sua matemática. Em nosso trabalho as imagens capturadas em RGB foram convertidas para HSV, por terem características úteis para segmentar imagem: com sombras utilizamos informações dos canais de intensidade **V** e saturação **S**; e com as caixas coloridas usamos as informações dos canais de matiz **H** e saturação **S**. A técnica principal utilizada para separar as informações foi o filtro de limiar.

Figura 23: Gráfico representando o filtro de limiar e o limiar T .

Fonte: adaptada de Gonzalez (2001).

2.7.2 Filtro de limiar

Um dos mais simples métodos para separar características é o filtro de limiar. O filtro de limiar, também conhecido como *threshold*, consiste em escolher alguns dos *pixels* como alvos (*foreground*), que é geralmente um objeto ou região de interesse, os outros *pixels* não escolhidos serão o fundo (*background*). Por exemplo, dada uma distribuição em escala de cinza em uma imagem, e assumindo que um tom de cinza T seja escolhido como valor que separa os *pixels* em grupos, então, pode-se definir que todos os *pixels* maiores ou iguais a T serão o alvo enquanto os menores serão o fundo. Pode-se classificar o filtro de limiar como: passa alta $f(x, y) \geq T$; passa baixa $f(x, y) \leq T$; passa faixa $T_{min} \leq f(x, y) \leq T_{max}$; e rejeita faixa $f(x, y) < T_{min}$ ou $f(x, y) > T_{max}$ (Figura 23) (GONZALEZ, 2001; SHAPIRO; STOCKMAN, 2001).

Com esse filtro podemos obter uma imagem binária (preto e branco), ou uma imagem truncada onde o limiar torna-se a intensidade máxima da imagem, ou o fundo igual à zero mantendo a intensidade do alvo. Esse é um filtro útil para segmentar imagem na qual os objetos têm características bem definidas em um dos canais de informação (GONZALEZ, 2001; SHAPIRO; STOCKMAN, 2001).

2.7.3 Filtros morfológicos

Os filtros morfológicos trabalham com a forma das imagens, seus operadores básicos são dilatação, erosão, fechamento e abertura. São utilizados para fechar lacunas, ou eliminar pequenas áreas. A dilatação funciona do seguinte modo: a partir de uma imagem binária $\mathbf{B}$ e um elemento estruturante $\mathbf{E}$, formam uma nova imagem dilatada, ao passar o elemento $\mathbf{E}$ sobre a imagem $\mathbf{B}$ a nova imagem é formada com os pontos da união de $\mathbf{B}$ e $\mathbf{E}$. Por outro lado, na erosão um ponto somente será mantido na nova imagem erodida se todos os pontos de $\mathbf{E}$ forem também pontos de $\mathbf{B}$. Uma explicação mais detalhada pode ser encontrada em Shapiro e Stockman (2001).

Procurando facilitar o entendimento, tem-se uma representação da dilatação e da erosão na Figura 24. Pode-se utilizar um filtro morfológico de erosão para eliminar pequenas áreas e logo

Figura 24: Representação gráfica da dilatação e da erosão com uma matriz 3x3 com elemento estruturante E .

Fonte: <http://www.cs.princeton.edu/~pshilane/class/mosaic/> último acesso em 31/01/2014.

após, um filtro morfológico de dilatação para manter as características do objeto. A combinação desses filtros nesta sequência é conhecida como operador morfológico de abertura. O operador morfológico de fechamento tem sequência inversa (dilata-erode) ao operador morfológico de abertura, é utilizado para completar pequenos espaços vazios (SHAPIRO; STOCKMAN, 2001). Neste trabalho após aplicar o filtro de limiar, o filtro morfológico de abertura foi empregado para eliminar pequenas regiões mantendo a forma das regiões maiores.

2.7.4 Detectores de borda

Os detectores de bordas funcionam identificando as variações acentuadas da intensidade de *pixels*. Matematicamente, eles identificam transições bruscas da função $f(x, y)$, que são medidas pela derivada primeira de $f'(x, y)$ ou mesmo pela derivada segunda $f''(x, y)$, o processamento de $f'(x, y)$ é realizado utilizando a magnitude do gradiente. Para uma imagem $f(x, y)$, o gradiente de f nas coordenadas (x, y) é definido por um vetor coluna bidimensional:

$$\nabla f = \begin{pmatrix} \frac{\delta f}{\delta x} \\ \frac{\delta f}{\delta y} \end{pmatrix} \quad (2.50)$$

e a magnitude deste vetor é dada por,

$$mag(\nabla f) = \left[\left(\frac{\delta f}{\delta x} \right)^2 + \left(\frac{\delta f}{\delta y} \right)^2 \right]^{\frac{1}{2}}. \quad (2.51)$$

Os operadores para detecção de bordas que utilizam gradientes são Sobel, Prewitt, Laplace, Canny. Mais detalhes sobre estes detectores de borda podem ser encontrados em (GONZALEZ, 2001; SHAPIRO; STOCKMAN, 2001).

Figura 25: Exemplo da aplicação do componente conectado, a partir da imagem original (esquerda) um filtro limiar e morfológico aplicado (meio) o resultado na imagem com grupos de mesmo valor (representado pela cor).

Fonte: autor.

2.7.5 Componentes conectados

A segmentação por componente conectado verifica todos os *pixels* de uma imagem, comparando se os *pixels* vizinhos são semelhantes. A partir desta avaliação cria-se uma nova imagem, onde cada grupo terá o mesmo valor. Um exemplo desta segmentação pode ser vista na Figura 25. Da esquerda para direita, vê-se a imagem original em RGB, o objeto extraído com filtro de limiar e a imagem segmentada por componentes conectados.

Note que o filtro de componente conectado é útil para separar e distinguir objetos, assim é possível extrair informações como área e quantidade de elementos. As informações extraídas a partir deste filtro podem ser utilizadas em etapas de nível alto, como classificação e decisão. Os componentes conectados foram importantes em nosso trabalho, pois com eles determinam-se as regiões de interesse na imagem. Com os detectores de borda definem as bordas dessas regiões, e essas informações serão utilizadas para extrair quais as relações entre as regiões na imagem. No próximo capítulo vamos falar sobre o trabalho desenvolvido no domínio de tráfego, referente à primeira contribuição.

Parte I

Primeira Contribuição

3 RACIOCÍNIO ESPACIAL QUALITATIVO COM LÓGICA DE DESCRIÇÃO PROBABILÍSTICA

Este capítulo apresenta uma formalização de um domínio espacial na forma de raciocínio espacial qualitativo codificado em uma lógica de descrição probabilística, para responder perguntas sobre um domínio de tráfego. A razão de utilizar lógica de descrição probabilística para este problema é desenvolver um subconjunto tratável do raciocínio espacial qualitativo capaz de manipular incertezas de forma rigorosa, bem como permitir a representação de conhecimento de alto nível. Para o raciocínio espacial qualitativo utilizamos um subconjunto do cálculo de direção cardinal (FRANK, 1996; LIGOZAT, 1998) (seção 2.2.4) e a lógica de descrição probabilística *CRALLC* (COZMAN; POLASTRO, 2008) (seção 2.4). Como domínio, consideramos um cenário onde um veículo experimental transita por uma via e está equipado com três sensores: um mapa digital, um GPS e uma câmera de vídeo.

A partir deste domínio, a tarefa do sistema consiste em responder duas questões: “qual faixa corresponde aquela em que o veículo está?”; e “qual mão de direção cada faixa permite?”. Essa tarefa foi proposta com o objetivo de investigar como utilizar o *CRALLC* com raciocínio espacial qualitativo e quais seriam suas limitações. Uma formalização do domínio foi desenvolvida para responder essas questões e foi testado com os estados de informações dos sensores, já que para este domínio não tivemos acesso aos dados reais dos sensores. A seguir apresentaremos o cenário de tráfego restrito que utilizamos, o qual foi inspirado no domínio de tráfego de Hummel (2006).

3.1 Descrição do Cenário de Tráfego

Nosso domínio consiste de um cenário de tráfego onde uma rua é dividida em faixas. Cada faixa da rua tem somente uma mão de direção: ou sobe, ou desce. Entre cada par de faixas existe uma linha divisória, a qual é ou uma linha tracejada, ou uma linha contínua (Figura 26a). O cenário também contém um veículo experimental equipado com três sensores: um mapa digital, um GPS e uma câmera de vídeo (Figura 26b). O formalismo deve estimar as propriedades funcionais da rua onde o veículo está, utilizando os sensores do veículo para responder a seguintes questões:

1. qual faixa corresponde aquela em que o veículo está? - A resposta é um elemento do conjunto: $\{1, \dots, n\}$, onde n é o número de faixas da rua. Esta tarefa vem do fato que os

Figura 26: Representação gráfica do cenário de tráfego (mão francesa). (a) Representação da rua, faixa, e linhas divisórias. (b) Representação do veículo no cenário.

Fonte: Fenelon et al. (2010).

receptores do GPS são capazes de determinar a rua em que o veículo está, mas não em qual faixa (HUMMEL, 2006).

2. qual mão de direção cada faixa permite? - A resposta é se a mão de direção da faixa é: ou sobe, ou desce relativo ao sistema de coordenada da rua.

Os dados do sensor de vídeo têm maior grau de incerteza. Este sensor consegue obter poucas informações das faixas adjacentes àquela em que o veículo está. Uma das razões para tal fato é porque o campo de visão que está à frente do carro tem um ponto cego sobre grande parte das faixas adjacentes. Além disso, as linhas divisórias de faixa, as marcações no chão (como setas de direção) e sinais de trânsito podem não aparecer na imagem devido à oclusão causada por outros veículos. Contudo oclusão não será um problema que iremos tratar neste domínio.

Algumas propriedades não são explicitamente dadas na forma de símbolos, mas precisam vir do contexto de um motorista humano (tal como a regra de trânsito da mão-direita). Estas observações são refletidas pelas entradas dos sensores disponíveis. Os sensores disponíveis para resolver a tarefa aqui proposta são:

Câmara de vídeo - um algoritmo de reconhecimento de marcas de divisão baseado em vídeo reconhece a linha de divisão e a classifica como linha tracejada ou linha contínua. Das informações obtidas pelo sensor de vídeo, uma matriz de confusão com os valores probabilísticos de reconhecimento e classificação do tipo de faixa está especificada na Tabela 3.

Figura 27: (a) Cena Original. (b) Exemplo dos dados dos sensores para o cenário da cena original.

Fonte: Santos et al. (2010).

Tabela 3: Matriz de confusão do sensor de captura de imagem em que colunas representam o tipo de linha divisória real enquanto que as linhas representam como a linha divisória foi classificada pelo sistema.

Vídeo: Linha divisória		
Classificação		
	Contínua	Tracejada
Contínua	0,80	0,07
Tracejada	0,20	0,93

Fonte: Santos et al. (2010), Fenelon et al. (2010).

GPS - este sensor informa a rua em que o veículo está, assim como define a direção em que ele se desloca. As direções são discretizadas como “going up” ou “going down” relativo ao sistema de coordenadas da rua (HUMMEL, 2009).

Mapa digital - este sensor informa quantas faixas a rua tem e se esta é uma rua de mão única ou de mão dupla. A Tabela 4 representa a matriz de confusão com a taxa de acerto das informações desse sensor.

Os valores da Tabela 4 são baseados nos resultados do algoritmo comparados com o evento real (*ground truth*) (HUMMEL, 2009), enquanto que os dados da Tabela 3 foram estimados. A Figura 27a mostra uma imagem da rua do ponto de vista do carro, enquanto que a Figura 27b tem uma representação dos dados dos sensores para o mesmo ponto de vista.

Neste domínio tratamos com as direções globais (enquanto que o domínio é inspirado em uma cena dinâmica natural de tráfego), assume-se que cada rua tem seu próprio sistema de direção cardinal, no qual a direção “Down-Up” (Desce-Sobe) em vez de “South-North” (Sul-Norte) vai da origem da via até seu fim, seguindo a linha central da rua. Em outras palavras, “Down-Up” representa a direção entre dois objetos *A* e *B* na via e são definidos como uma

Tabela 4: Matriz de confusão dos valores de probabilidade de resposta do sensor mapa digital em que colunas representam o tipo de rua enquanto que as linhas representam como a rua foi classificada pelo sistema.

Mapa digital: Classificação da rodovia		
	Mão única	Duas mãos
Mão única	0,99	0,01
Duas mãos	0,01	0,99

Fonte: Santos et al. (2010), Fenelon et al. (2010).

Figura 28: Um sistema cardinal para vias: A está ao sul de B e oeste de C .

Fonte: Santos et al. (2010), Fenelon et al. (2010).

projeção da linha de A para B , que não é necessariamente a linha central da rua. A direção “East-West” (Leste-Oeste), referida como “right-left” (direita-esquerda) é definida como todos os pontos da via que estão na linha ortogonal que tangencia a linha central naquele ponto. A Figura 28 mostra um exemplo de como são definidas estas direções.

Com o conhecimento das características da rua e das regras de trânsito é possível inferir novas informações, somando-se às informações dos sensores. Na próxima seção apresentaremos uma formalização deste domínio em *CRALC*, a fim de estimar propriedades funcionais da rua utilizando os sensores do veículo. A taxonomia dos conceitos e papéis relevantes à tarefa de trafegar é apresentada a seguir.

3.2 Definindo o Cenário de Tráfego em *CRALC*

Com o domínio definido em lógica de descrição probabilística, a base de conhecimento referente as características e regras do cenário será formalizada em uma *TBox*. O conceito Lane (faixa) é definido utilizando duas primitivas, Up (sobe) e Down (desce) (form. (3.1)). O conceito Divider (linha divisória) é definido como a união de DashedDivider (linha tracejada)

e **SolidDivider** (linha contínua) (form. (3.2)), e o conceito **Vehicle** (veículo) é definido como **GoingUp** (subindo a rua) ou **GoingDown** (descendo a rua) (form.(3.3)) :

$$\text{Lane} \equiv \text{Up} \sqcup \text{Down} \quad (3.1)$$

$$\text{Divider} \equiv \text{DashedDivider} \sqcup \text{SolidDivider} \quad (3.2)$$

$$\text{Vehicle} \equiv \text{GoingUp} \sqcup \text{GoingDown} \quad (3.3)$$

Nas fórmulas (3.4)–(3.8) e (3.10) uma abreviação $\text{disjoint}(t_1, t_2, \dots, t_n)$ foi utilizada para representar o conjunto de declarações sobre termos exclusivamente disjuntos, por exemplo, $\text{disjoint}(t_i, t_j)$ significa que t_i contém indivíduos que não são t_j , assim $t_i \sqsubseteq \neg t_j \forall i, j \in 1, \dots, n, i \neq j$, onde o conceito t_i inclui a negação do conceito t_j e ambos são conceitos diferentes.

$$\text{disjoint}(\text{Vehicle}, \text{Divider}, \text{Lane}) \quad (3.4)$$

$$\text{disjoint}(\text{Up}, \text{Down}) \quad (3.5)$$

$$\text{disjoint}(\text{DashedDivider}, \text{SolidDivider}) \quad (3.6)$$

$$\text{disjoint}(\text{GoingUp}, \text{GoingDown}) \quad (3.7)$$

$$\text{disjoint}(\text{OnOneWayRoad}, \text{OnTwoWayRoad}) \quad (3.8)$$

A taxonomia dos papéis consiste das relações do cálculo de direções cardinais (CDC) (seção 2.2.4). Definimos um CDC inspirado no formalismo proposto em Ligozat (1998), onde objetos espaciais são pontos em um espaço bidimensional e as direções cardinais entre dois objetos A e B são definidas como duas projeções de uma linha reta de A para B : uma no eixo Sul-Norte (Down-Up) e outra no eixo Oeste-Leste (left-right).

Considerando que o nosso domínio não tem cruzamento, somente utilizamos três das nove relações básicas de direção cardinal: leste (*east* (**e**)), oeste (*west* (**w**)) e igual (*equal* (**eq**)). Vale informar que a relação **eq** representa o fato de um veículo estar localizado em uma determinada faixa. Deste modo o conceito **cdc** é definido como a união dos papéis disjuntos **e**, **w**, **eq** (fórmulas (3.9) e (3.10)) abaixo:

$$\text{cdc} \equiv \text{e} \sqcup \text{w} \sqcup \text{eq} \quad (3.9)$$

$$\text{disjoint}(\text{e}, \text{w}, \text{eq}) \quad (3.10)$$

A seguir, utilizando os conceitos e papéis acima descritos será formulado um conjunto de conceitos mais restritos construídos sobre a regulamentação de trânsito. Começaremos formalizando a semântica de tráfego para a mão de direção da faixa, conforme definido pelas fórmulas (3.11) e (3.12). Considerando que este domínio tem o sentido de circulação pela direita (mão francesa), em uma via de mão dupla a faixa na extremidade direita da rua só é permitido aos veículos trafegarem subindo (**GoingUp**), enquanto que a faixa na extremidade esquerda da rua permite somente o oposto disso (**GoingDown**). A fórmula (3.11) define **GoingUp** quando a leste do veículo existe uma linha divisória contínua **SolidDivider** e não existe nenhuma faixa

Lane ou a leste existe uma linha divisória tracejada **DashedDivider** e uma faixa **Up**. De modo semelhante, a fórmula (3.12) define que **GoingDown** está dentro do conjunto onde a oeste do veículo existe uma linha divisória contínua **SolidDivider** e não existe nenhuma faixa **Lane**, ou a oeste existe uma linha divisória tracejada **DashedDivider** e uma faixa **Down**.

$$\text{GoingUp} \sqsubseteq \exists e.(\text{SolidDivider} \sqcap \neg \text{Lane}) \sqcup \exists e.(\text{DashedDivider} \sqcap \text{Up}) \quad (3.11)$$

$$\text{GoingDown} \sqsubseteq \forall w.(\text{SolidDivider} \sqcap \neg \text{Lane}) \sqcup \exists w.(\text{DashedDivider} \sqcap \text{Down}) \quad (3.12)$$

As fórmulas (3.13) e (3.14) referem-se às linhas divisórias, as quais se situam nas bordas das faixas. Neste cenário um **DashedDivider** é uma linha tracejada que separa faixas com mesma mão de direção, enquanto um **SolidDivider** é uma linha contínua que separa faixas com direções opostas ou indica a borda da pista.

$$\text{DashedDivider} \sqsubseteq (\exists e.\text{Up} \sqcap \exists w.\text{Up}) \sqcup (\exists e.\text{Down} \sqcap \exists w.\text{Down}) \quad (3.13)$$

$$\begin{aligned} \text{SolidDivider} \sqsubseteq & (\neg \exists e.\text{Lane} \sqcap \exists w.\text{Lane}) \sqcup (\neg \exists w.\text{Lane} \sqcap \exists e.\text{Lane}) \sqcup \\ & (\exists e.\text{Up} \sqcap \exists w.\text{Down}) \sqcup (\exists e.\text{Down} \sqcap \exists w.\text{Up}) \end{aligned} \quad (3.14)$$

Na fórmula (3.13) o conceito **DashedDivider** está no conjunto em que a leste e a oeste da linha ou ambas as faixas são **Up** ou ambas são **Down**. Na fórmula (3.14) **SolidDivider** pode ocorrer quando ou não existe uma faixa a leste da linha e existe a oeste, ou não existe uma faixa a oeste da linha e existe a leste, ou a leste a faixa é **Up** e a oeste é **Down**, ou a oeste a faixa é **Up** e a leste é **Down**. Finalmente, a fórmula (3.15) descreve uma via de mão única, no qual todas as faixas relacionadas com a rua ou são **Up** ou são **Down**, a fórmula (3.16) define uma via de mão dupla, onde existe pelo menos uma faixa **Up** e uma faixa **Down** relacionadas com a rua.

$$\text{OnOneWayRoad} \sqsubseteq \forall cdc.\text{Up} \sqcup \forall cdc.\text{Down} \quad (3.15)$$

$$\text{OnTwoWayRoad} \sqsubseteq \exists cdc.\text{Up} \sqcap \exists cdc.\text{Down} \quad (3.16)$$

Os sensores que equipam o veículo são conceitos que vamos adicionar a nossa $\mathcal{T}\text{Box}$ como subclasses que define **Sensor** (form. (3.17)). A fórmula (3.18) define o sensor de classificação de linha divisória **SensedDashedDivider** pela classificação da faixa detectada (**SensedDashedDivider** ou **SensedSolidDivider**).

$$\text{Sensor} \equiv \text{SensedOnOneWayRoad} \sqcup$$

$$\text{SensedOnTwoWayRoad} \sqcup \text{SensedDivider} \quad (3.17)$$

$$\text{SensedDivider} \equiv \text{SensedDashedDivider} \sqcup \text{SensedSolidDivider} \quad (3.18)$$

As probabilidades podem ser representadas como probabilidades condicionais, de modo que a probabilidade de um evento real ocorrer é vinculada à leitura de dados do sensor. As tabelas (Tabelas 3 e 4) são matriz de confusão com os valores das probabilidades da detecção de um evento. Por exemplo, para a definição dos conceitos de linhas divisórias pode ser formulado

como um conjunto de probabilidades condicionais, utilizando os valores da Tabela 3, como a seguir (form.(3.19) - form.(3.22)):

$$P(\text{DashedDivider}|\text{SensedDashedDivider}) = 0,93 \quad (3.19)$$

$$P(\text{SolidDivider}|\text{SensedDashedDivider}) = 0,07 \quad (3.20)$$

$$P(\text{DashedDivider}|\text{SensedSolidDivider}) = 0,20 \quad (3.21)$$

$$P(\text{SolidDivider}|\text{SensedSolidDivider}) = 0,80 \quad (3.22)$$

As fórmulas apresentadas até aqui formam a nossa base de conhecimento referente às regras e características do domínio, ou seja, a $\mathcal{T}\text{Box}$ especificada. Enquanto que os dados dos sensores GPS, mapa digital e câmera de vídeo formam a $\mathcal{A}\text{Box}$ conforme segue um exemplo:

- *GPS*: o veículo é apresentado como um indivíduo v e o GPS informa que v está subindo a rua (GoingUp), a qual deve ter pelo menos uma faixa com direção Up, assumindo que o motorista sempre dirija na direção correta.

v : Vehicle

v : GoingUp

- *Mapa Digital*: mostra que v está dirigindo em uma via de mão dupla ou de mão única; e informa o número de faixas (neste contexto sempre três faixas). As faixas são apresentadas como indivíduos l_1, l_2, l_3 , o menor número, por convenção, corresponde à faixa mais a leste.

v : SensedOnTwoWayRoad

Lane = $\{l_1, l_2, l_3\}$

(l_1, l_2) : e

(l_2, l_3) : e

- *Câmera de vídeo*: uma linha divisória tracejada foi reconhecida e classificada pelo sensor e está a leste de v :

d_1 : SensedDashedDivider

(d_1, v) : e

No *CRA $\mathcal{L}\mathcal{C}$* as informações da $\mathcal{A}\text{Box}$ serão representadas como evidências da rede Bayesiana formada com os dados da $\mathcal{T}\text{Box}$. Com o domínio formalizado, vamos descrever como a tarefa foi resolvida.

Figura 29: Domínio de tráfego e todas as situações possíveis para o veículo que esteja dentro das regras.

Fonte: adaptado de Santos et al. (2010) e Fenelon et al. (2010).

3.3 Estratégias para Solução da Tarefa

Representando os estados deste domínio a Figura 29 exhibe uma matriz de imagens onde cada linha tem um tipo de rua e as colunas representam a posição do carro em uma das faixas. Algumas considerações sobre esse domínio:

- existe um único carro;
- o sentido de circulação e sinais horizontais seguem as leis brasileiras de trânsito ;
- o motorista não desrespeita as leis de trânsito, anda sempre dentro da faixa;
- a rua tem um número fixo de três faixas.

Procuramos formalizar o domínio observando também algumas restrições impostas pelo software que a partir da terminologia em *CRALC* gera automaticamente uma rede Bayesiana para inferência. Contudo, não foi possível representar diretamente as fórmulas (3.9) e (3.10), pois apesar destas serem parte da definição básica do *CRALC*, a hierarquia de papéis não está no escopo do *ALC* e, conseqüentemente, não está no *CRALC*. Por isso fizemos duas abordagens, publicadas em Santos et al. (2010) e em Fenelon et al. (2010) para cumprir a tarefa de desenvolver um mecanismo de inferência, o qual dado a percepção dos sensores seja capaz de responder: em qual faixa o carro está e qual a mão de direção de cada faixa.

3.3.1 Inferência com uma rede Bayesiana

Para utilizar o mecanismo de inferência do *CRALC* as informações do cálculo de direção sobre o domínio ficaram implícitas nas definições de cada faixa e de suas possíveis direções. Ou seja, nenhuma asserção de conceito que utilizaram papéis faz parte desta *TBox*. Assim, a solução proposta para utilizar o *CRALC* foi deixar todos os papéis implícitos, proposicionalizando a ontologia e evitando as definições cíclicas (SANTOS et al., 2010).

As relações entre indivíduos $\exists eq.Lane$ ("existe um veículo na faixa") foram incluídas na base de conhecimento como novos conceitos, tais como *Onl1*, *Onl2* e *Onl3* (form.(3.23), form.(3.24) e form.(3.25) respectivamente) que indica em qual faixa o veículo se encontra (*On* faixa (*li*, $i = \{1, 2, 3\}$)). Enquanto que para a direção de cada faixa foram utilizados os conceitos *Upli* e *Downli* (para, $i = \{1, 2, 3\}$), indicando que uma faixa *i* tem mão de direção subindo ou descendo respectivamente (form. (3.26) até form.(3.31)). Como hipótese simplificadora somente uma linha divisória é percebida pela câmera.

$$\begin{aligned} Onl1 \equiv & (GoingDown \sqcap OnOneWayRoad \sqcap DashedDivider \sqcap \neg Onl2) \sqcup \\ & (GoingUp \sqcap SolidDivider). \end{aligned} \quad (3.23)$$

Tabela 5: Tabela com oito estados dos sensores indicando as dozes possibilidades de estado do veículo conforme representado na Figura 29. Com as coordenadas mudando relativas ao sentido de deslocamento do carro.

		OnOneWayRoad		OnTwoWayRoad	
		SolidDivider	DashedDivider	SolidDivider	DashedDivider
GoingDown	L1	0	1	0	0
	L2	0	1	0	1
	L3	1	0	2	0
GoingUp	L1	1	0	2	0
	L2	0	1	0	1
	L3	0	1	0	0

Fonte: autor.

$$\begin{aligned} \text{Onl2} \equiv & (\text{OnOneWayRoad} \sqcap \text{DashedDivider}) \sqcup \\ & (\text{OnTwoWayRoad} \sqcap \text{DashedDivider}). \end{aligned} \quad (3.24)$$

$$\begin{aligned} \text{Onl3} \equiv & (\text{GoingDown} \sqcap \text{SolidDivider}) \sqcup \\ & (\text{GoingUp} \sqcap \text{OnOneWayRoad} \sqcap \text{DashedDivider} \sqcap \neg \text{Onl2}). \end{aligned} \quad (3.25)$$

$$\text{Upl1} \equiv (\text{GoingUp} \sqcap \text{OnOneWayRoad}). \quad (3.26)$$

$$\begin{aligned} \text{Upl2} \equiv & (\text{GoingUp} \sqcap \text{OnOneWayRoad}) \sqcup \\ & (\text{GoingUp} \sqcap \text{OnTwoWayRoad} \sqcap \text{Onl2}). \end{aligned} \quad (3.27)$$

$$\text{Upl3} \equiv (\text{GoingUp} \sqcap \text{OnOneWayRoad}) \sqcup \text{OnTwoWayRoad}. \quad (3.28)$$

$$\text{Downl1} \equiv (\text{GoingDown} \sqcap \text{OnOneWayRoad}). \quad (3.29)$$

$$\begin{aligned} \text{Downl2} \equiv & (\text{GoingDown} \sqcap \text{OnOneWayRoad}) \sqcup \\ & (\text{GoingDown} \sqcap \text{OnTwoWayRoad} \sqcap \text{Onl2}). \end{aligned} \quad (3.30)$$

$$\text{Downl3} \equiv (\text{GoingDown} \sqcap \text{OnOneWayRoad} \sqcap \text{Onl3}) \sqcup \text{OnTwoWayRoad}. \quad (3.31)$$

Nesta abordagem, o ponto de vista da câmera é relativo ao carro, ou seja, sempre é identificada a faixa que está do lado do passageiro. Portanto, significa que pelas coordenadas da rua a linha divisória que será classificada está: a oeste do carro caso este se desloque para o sul, e a leste do carro caso o sentido de deslocamento seja norte. Considerando os doze estados representados na Figura 29, a Tabela 5 mostra para cada combinação de resultados dos sensores quantas vezes o carro está na faixa. Essa tabela auxiliou na definição dos conceitos *Onl1*, *Onl2* e *Onl3* (form.(3.23), form.(3.24) e form.(3.25) respectivamente), a partir da disjunção de conjunções das respostas dos sensores.

Figura 30: Rede Bayesiana formada com conceitos e papéis embutidos no banco de conhecimento. Todos os nós em azul são evidências observadas pelos sensores.

Fonte: Santos et al. (2010).

Quando um conceito é definido por disjunção exclusiva, como na fórmula (3.4), isso poderia gerar um ciclo, que não é permitido nas redes Bayesianas. Assim sendo, nesta abordagem as disjunções exclusivas tal como $disjoint(A, B, C)$ foram representadas utilizando declarações como $A \sqsubseteq \neg C$ e $B \sqsubseteq \neg A \sqcap \neg C$. Essa representação foi aplicada para as disjunções exclusivas referentes: a fórmula (3.7); a resposta do sensor de linhas divisórias, para garantir a disjunção exclusiva da fórmula (3.6); e para os conceitos Onli. Mais algumas adaptações realizadas na ontologia apresentada na seção 3.2 foram:

- A disjunção exclusiva da fórmula (3.8) é representada pela probabilidade do (SensedTwoWayRoad).
- A fórmula 3.1 que define Lane é definida como fórmula (3.32);
- e a fórmula 3.33 é a que define Vehicle.

$$\text{Lane} \equiv \text{GoingUp} \sqcup \text{GoingDown} \quad (3.32)$$

$$\text{Vehicle} \equiv \text{OnOneWayRoad} \sqcup \text{OnTwoWayRoad} \quad (3.33)$$

Dado o formalismo apresentado na seção 3.2 e as considerações acima, o sistema gerou automaticamente a rede Bayesiana para um único indivíduo (Figura 30). Os nós referentes à percepção dos três sensores (GoingUp, GoingDown, SensedTwoWayRoad, SensedDashedDivider

Tabela 6: Resposta da rede Bayesiana representada na Figura 30 para a primeira questão: “qual faixa tem maior probabilidade do carro estar localizado”. As evidências referentes aos dados dos três sensores (A) estão expressas nas três primeiras colunas.

GPS	Mapa	Câmera	rede Bayesiana (Figura 30)
GoinUp	SensedOnTwoWayRoad	SensedDashedDivider	$\text{argmax}_{li} P((v : \text{Onli}))$
0	0	0	l3
0	0	1	l1 $\sqcup$ l2
0	1	0	l3
0	1	1	l2
1	0	0	l1
1	0	1	l2 $\sqcup$ l3
1	1	0	l1
1	1	1	l2

Fonte: Santos et al. (2010).

e **SensedSolidDivider**) são nós que recebem sua informação de valor a partir da $\mathcal{A}\text{Box}$. Por exemplo, caso o sensor perceba que o carro está subindo a rua então o nó **GoingUp** será uma evidência verdadeira, representado na rede Bayesiana da Figura 30 com a cor azul. Nesta configuração existem oito combinações de evidências para esses sensores. Com a rede da Figura 30 é possível responder as questões:

1. Em qual faixa (li) o veículo (v) tem a maior probabilidade de estar? Essa inferência é dada por $\text{argmax}_{li} P((v : \text{Onli}))$.
2. Qual a direção de cada faixa? Por exemplo, qual a probabilidade para cada faixa li de ter o sentido de direção descendo **Downli**? Essa inferência é dada por $\forall i : P(li : \text{Downli})$.

Consultando a rede Bayesiana da Figura 30 para os oito estados possíveis dos três sensores, obtivemos as respostas para as questões 1 e 2. A Tabela 6 apresenta a resposta sobre qual é a provável faixa na qual o veículo v está, enquanto a Tabela 7 apresenta a resposta sobre o sentido de direção da faixa.

A última coluna da Tabela 6 mostra as faixas na qual o veículo s provavelmente está ($\text{argmax}_{li} P((v : \text{Onli}))$), dado as evidências A dos três sensores representadas nas três primeiras colunas. Cada sensor indica uma de duas respostas mutuamente exclusivas, assim a primeira coluna representa os dados do sensor **GPS (GoingUp)** com a informação de direção em que o carro se movimenta. Na coluna seguinte estão os dados do mapa digital (**SensedTwoWayRoad**) indicando se é uma rodovia de mão única ou uma rodovia de mão dupla. Enquanto que na terceira coluna estão os dados da câmera (**SensedDashedDivider**) que faz uma classificação de linha divisória como linha contínua ou linha tracejada.

A primeira linha da Tabela 6, por exemplo, representa o estado onde o **GPS** indica que o carro está descendo a rua, e o mapa digital indica que é rua de mão única, e o sistema de visão classifica a linha divisória como linha contínua. Dado estas evidências o nó **Onl3** tem a

Tabela 7: Resposta da pergunta dois a partir da rede Bayesiana da Figura 30: a probabilidade para a direção de a faixa ser *Down* dado às evidências *A* (expressas nas seis primeiras colunas).

GPS	Mapa	Câmera	l1		l2		l3	
GoingUp	SensedOn TwoWayRoad	Sensed SolidDivider	P(Downl1 A)	P(Upl1 A)	P(Downl2 A)	P(Upl2 A)	P(Downl3 A)	P(Upl3 A)
0	0	0	0,99	0,01	0,99	0,01	1,00	0,00
0	0	1	0,99	0,01	0,99	0,01	1,00	0,00
0	1	0	0,01	0,99	0,76	0,24	1,00	0,00
0	1	1	0,01	0,99	0,95	0,05	1,00	0,00
1	0	0	0,00	1,00	0,01	0,99	0,01	0,99
1	0	1	0,00	1,00	0,00	1,00	0,01	0,99
1	1	0	0,00	1,00	0,61	0,39	0,99	0,01
1	1	1	0,00	1,00	0,09	0,91	0,99	0,01

Fonte: Santos et al. (2010).

Figura 31: Exemplo de três cenários de tráfego, onde o carro está descendo uma via de mão única.

Fonte: Santos et al. (2010).

maior probabilidade entre todos nós *Onli* da rede Bayesiana (Figura 30). Na Figura 31c temos uma ilustração desde caso. Na segunda linha da Tabela 6, tanto o **GPS** quanto o mapa digital continuam como descritos anteriormente, contudo a visão mudou e agora percebe uma linha tracejada. Neste caso existem duas hipóteses com igual probabilidade ¹ representadas nas Figuras 31a e 31b, as quais mostram um veículo na faixa l1 e na faixa l2 respectivamente. Os outros casos na Tabela 6 são análogos. A Tabela 6 tem alguns casos em que o sistema responde duas em vez de um faixa, isso devido ao fato de termos nesta configuração somente oito estados de sensores para doze estados deste domínio. Os resultados são os esperados quando comparamos com a Tabela 5.

A Tabela 7 tem os valores de probabilidade da direção permitida em cada faixa li, dado as evidências representadas nas três primeiras colunas. Tomando como exemplo a última linha, a probabilidade é maior para *Upl1*, *Upl2* e *Downl3*, as quais são consistentes com as evidências de que o veículo está subindo uma rua de mão dupla e percebe uma linha tracejada. Para a outra configuração das evidências as direções de cada faixa são dadas de modo similar nesta tabela.

Figura 32: Rede Bayesiana relacional representando um domínio de tráfego com papéis e indivíduos representando veículo, faixas, e linha divisória.

Fonte: Fenelon et al. (2010).

3.3.2 Rede Bayesiana relacional com transitividade

Na seção anterior, devido às restrições do *CRALC* em representar transitividade, todo o raciocínio espacial ficou implícito, gerando uma rede específica para aquele domínio. Para resolver isso, uma segunda rede Bayesiana relacional foi gerada manualmente. Esta rede representada na Figura 32 teve como base as regras de tradução do *CRALC* (seção 2.4). Esta abordagem tem como objetivo verificar a possibilidade de utilizar os papéis para representar as relações transitivas entre indivíduos. Nessa abordagem uma rede foi gerada para cada indivíduo: três faixas, um veículo e uma linha divisória. A geração da rede relacional resultante não é uma tarefa manual trivial e esta rede (Figura 32) foi simplificada para utilizar a linguagem de descrição probabilística. Os indivíduos têm diferentes relações entre eles, essas relações na maioria dos casos são percebidas pelos sensores, outras são inferidas pelo raciocínio espacial formalizado para o domínio.

Como qualquer *TBox* em *CRALC* pode ser representada por uma rede Bayesiana para inferência, as hierarquias de papéis e os conceitos disjuntos na *TBox* foram manualmente transformados em uma rede Bayesiana relacional. Conceitos foram traduzidos em nós da rede de tal modo que os conceitos à esquerda foram traduzidos como pais daqueles à direita. Para a disjunção, foi criado um nó *Xor* (ou exclusivo) que fica sempre evidenciado como verdadeiro na rede. Podemos então determinar se uma faixa sobe ou desce utilizando as fórmulas (3.13), (3.14) e (3.16). Apesar destas fórmulas definirem os conceitos utilizando papéis, neste trabalho elas foram definidas a partir de probabilidades condicionais dadas as evidências dos sensores de serem *GoingUp*, *GoingDown*, *OnOneWayRoad*, *OnTwoWayRoad*. O mesmo procedimento foi usado para indicar o tipo da linha divisória que está relacionada a cada faixa; ou seja, a linha divisória que está a leste da faixa. Considerou-se uma visão global e o sistema de coordenadas fixo a rua, independente da direção do carro.

¹ Diferem na terceira casa decimal

Tabela 8: Tabela de composição do cdc para o domínio de tráfego

domínio de trafego	e(b,c)	w(b,c)	eq(b,c)
e(a,b)	e(a,c)	eq(a,c) e(a,c) w(a,c)	e(a,c)
w(a,b)	eq(a,c) e(a,c) w(a,c)	w(a,c)	w(a,c)
eq(a,b)	e(a,c)	w(a,c)	eq(a,c)

Fonte: adaptado de Skiadopoulos e Koubarakis (2001).

Figura 33: Rede Bayesiana baseado na tabela de composição (Tabela 8)

Fonte: Fenelon et al. (2010).

Nesta abordagem, focamos no cálculo de direção cardinal (cdc) para determinar em qual faixa o veículo está localizado. Utilizamos uma ferramenta de raciocínio espacial qualitativo a tabela de composição criando uma tabela de composição para o cdc deste domínio (Tabela 8). A Figura 33 refere-se tabela de composição (Tabela 8) representada como uma rede Bayesiana. Essa rede foi parte do estudo para verificar como fazer inferências em uma rede Bayesiana utilizando as informações de transitividade das relações observadas entre três indivíduos. A partir das considerações das características do domínio, com a Tabela 8 pode-se determinar a $P(eq(vehicle, lane))$ utilizando os dados da ABox; por exemplo, se o sensor indicar uma linha divisória a leste do veículo ($e(vehicle, divider)$), inferindo as linhas de divisão relacionadas a cada faixa ($e(lane, divider)$), então podemos calcular a probabilidade de cada faixa ser aquela na qual o veículo está localizado. Intuitivamente, caso o veículo e a faixa estejam ambos à leste da linha divisória percebida pelo sensor, logo o veículo está na faixa.

$$e(vehicle, divider) \sqcap e(lane, divider) \longrightarrow eq(vehicle, lane) \quad (3.34)$$

Dada à formalização apresentada na seção 3.2 (e as considerações acima), gerou-se a rede Bayesiana relacional representada na Figura 32, onde os nós em destaque (vermelho) são as variáveis observadas, por exemplo, estado dos sensores. A rede Bayesiana da Figura 32 é

Tabela 9: Tabela que apresenta à quantidade de vezes que uma faixa pode ocorrer dada a combinação das características que podem ser percebidas pelos sensores.

		OnOneWayRoad		OnTwoWayRoad	
		SolidDivider	DashedDivider	SolidDivider	DashedDivider
GoingDown	L1	1	0	0	0
	L2	0	1	1	0
	L3	0	1	1	1
GoingUp	L1	1	0	2	0
	L2	0	1	0	1
	L3	0	1	0	0

Fonte: autor.

Tabela 10: Resposta da primeira questão: "qual faixa tem maior probabilidade do carro estar localizado?" As evidências (A) estão expressas nas seis primeiras colunas referentes aos dados dos três sensores.

GPS	Mapa	Câmera	rede Bayesiana (Figura 32)
GoinUp	SensedOnTwoWayRoad	SensedDashedDivider	$\text{argmax}_{li} P((v, li : eq A))$
0	0	0	l1
0	0	1	l2 $\sqcup$ l3
0	1	0	l2 $\sqcup$ l3
0	1	1	l3
1	0	0	l1
1	0	1	l2 $\sqcup$ l3
1	1	0	l1
1	1	1	l2

Fonte: Fenelon et al. (2010).

formada por cinco redes conectadas. Além das informações da $\mathcal{A}\text{Box}$, existem as evidências do conceito de disjunção (nós Xor) e dos nós que indicam se a rede representa uma faixa, um veículo ou uma linha divisória.

Com as redes da Figura 32 é possível responder as duas questões:

1. $\text{argmax}_{li} P((v : li : eq))$ $i = 1, 2, 3$, ou seja, em qual faixa (li) o veículo (v) tem a maior probabilidade de estar localizado?
2. $\forall i : P(li : \text{Down})$, ou seja, qual a probabilidade para cada faixa li permitir a direção Down (descendo)?

Consultando a rede da Figura 32, para os oito estados dos três sensores obtivemos as respostas para as questões 1 e 2 apresentadas pelas Tabelas 10 e 11 respectivamente.

A Tabela 10 mostra qual é a faixa mais provável em que o veículo v está dirigindo ($\text{argmax}_{li} P((v, li : eq))$), considerando as evidências A dos três sensores representas nas três primeiras colunas os dados representando. A primeira linha da Tabela 10, por exemplo, representa o estado onde o GPS indica GoingDown, o mapa digital indica ser uma rua de mão única e o sistema de visão indica uma linha contínua. Dadas estas evidências o nó li com a maior

Tabela 11: Resposta da pergunta dois a partir da rede Bayesiana da Figura 32: a probabilidade para a direção de a faixa ser *Down* dado às evidências *A* (expressas nas seis primeiras colunas).

GPS	Mapa	Câmera	I1		I2		I3	
GoingUp	SensedOn TwoWayRoad	Sensed SolidDivider	P(DownI1 A)	P(UplI1 A)	P(DownI2 A)	P(UplI2 A)	P(DownI3 A)	P(UplI3 A)
0	0	0	0,98	0,02	0,99	0,01	1,00	0,00
0	0	1	0,98	0,02	0,99	0,01	1,00	0,00
0	1	0	0,50	0,50	0,50	0,50	1,00	0,00
0	1	1	0,05	0,95	0,50	0,50	1,00	0,00
1	0	0	0,00	1,00	0,00	1,00	0,00	1,00
1	0	1	0,00	1,00	0,00	1,00	0,00	1,00
1	1	0	0,00	1,00	0,50	0,50	0,99	0,01
1	1	1	0,00	1,00	0,50	0,50	0,99	0,01

Fonte: Fenelon et al. (2010).

probabilidade a partir da rede Bayesiana relacional da Figura 32 é o nó I1. Este caso está representado na Figura 31a. Na segunda linha da Tabela 10, o GPS e o mapa digital continuam como descrito anteriormente, contudo a visão percebe uma linha tracejada. Neste caso, tem duas hipóteses com igual probabilidade²: I2 e I3, representados nas Figuras 31b e 31c respectivamente. Todos os outros casos na Tabela 10 são análogos.

A Tabela 11 representa a resposta para a questão 2 mostrando os valores de probabilidades para cada faixa Ii (DownIi, UplIi), dadas as evidências das três primeiras colunas. Pegando a primeira linha como exemplo, a probabilidade de ser DownIi é maior para as faixas I1, I2 e I3, o que é consistente com as evidências percebidas pelos sensores de que o veículo está descendo uma rua (GoingDown) de mão única (SensedOnOneWayRoad). Todos os outros casos nesta tabela são semelhantes. Nosso objetivo foi utilizar os cálculos de direções cardinais como informações para novas inferências em domínios reais. Portanto, esta abordagem serviu para verificar que é possível trabalhar com transitividade das relações entre indivíduos.

3.4 Discussão dos Resultados

A rede Bayesiana da Figura 30 foi gerada automaticamente para somente um indivíduo. Esta rede foi bem sucedida, pois as respostas eram as esperadas, sendo consequência direta da formalização do domínio assumido. Dado este sucesso inicial, supomos que há uma extensão adequada de CRALC capaz de representar e raciocinar sobre domínios espaciais de qualquer sistema de raciocínio espacial qualitativo. Para investigar esta suposição, seguimos empregando a linguagem de descrição CRALC como base, mas com alguns recursos adicionais para lidar com conceitos disjuntos e hierarquias de papéis que aparecem nos domínios espaciais, resultando na rede Bayesiana relacional da Figura 32, a qual gerou resultados dentro do esperado.

Neste trabalho investigou-se a formalização de um domínio espacial como extensão de uma lógica de descrição probabilística básica, CRALC. Fomos capazes de usar a expressividade de um formalismo relacional da lógica de descrição ALCC, com o tratamento de incerteza fornecido pelas redes Bayesianas, com o qual o modelo de sensor foi codificado. Essa formalização

² Diferem na terceira casa decimal

foi o primeiro trabalho a utilizar cálculos de raciocínio espacial qualitativo estendido com uma abordagem probabilística.

Uma discussão sempre levantada é que o algoritmo atual tem a restrição da definição acíclica. Essa foi uma das primeiras dificuldades encontradas na formalização da *TBox*. Cada conceito e papel deveriam ser definidos sem gerar ciclos e em uma sequência tal, que um conceito que esteja à direita tenha sido definido anteriormente. Esse desafio foi superado renomeando alguns conceitos.

Em *CRALC* fórmulas de inclusão para papéis têm uma dificuldade adicional, uma vez que eles podem levar às definições circulares. Hierarquias de papéis também não fazem parte da *ALC* básico, mas foram necessários para representar a relação entre as pistas e as suas direções e entre o veículo e sua faixa de rodagem. Na formalização apresentada neste trabalho, evitamos os papéis na formalização do domínio deixando-os implícitos, representando os papéis assumidos pelos indivíduos como conceitos, o que permitiu uma representação da relação espacial no domínio utilizando o algoritmo *CRALC*. Numa segunda abordagem incluímos as regras de composição proposionalizando os papéis. Isto é devido à natureza do cálculo de direção cardinal, em que um conjunto de relações espaciais somente uma relação acontece com o mesmo par de indivíduos. A mesma característica do formalismo raciocínio espacial subjacente deve ser explorada em outros domínios.

Além de fornecer uma solução para o uso de inclusões de papéis baseada em composição e hierarquias de funções dentro do *CRALC*, este trabalho contribui com a formalização de precisão do sensor como fórmulas *TBox* probabilísticos em *CRALC*. Nesta formalização, uma precisão estimada em particular foi determinada pelo modelo do sensor a partir do qual se originou. A maior dificuldade desta tarefa, que ainda enfrentamos no nosso trabalho, é a representação de relações transitivas, as quais são peças importantes do conhecimento espacial.

Uma consideração a fazer: formalizamos o domínio, porém não tivemos acesso a dados reais dos sensores para estendê-lo a uma aplicação real. O *CRALC* ainda estava em desenvolvimento, e sem a hierarquia de papéis um sistema com raciocínio espacial para aplicar em robô móvel não seria uma tarefa a ser realizada em médio prazo. Mas a partir do conhecimento adquirido com as formalizações dos domínios com *CRALC* codificando um raciocínio espacial, decidimos aplicar as inferências com redes Bayesianas a outro domínio com o qual tínhamos mais contato e que precisava do reforço probabilístico para lidar com as incertezas dos sensores e do ambiente real. No próximo capítulo vamos mostrar este sistema de auto localização de um robô móvel que foi estendido combinando um filtro Bayesiano que tem como evidências as relações qualitativas entre elementos do ambiente.

Parte II

Segunda Contribuição

4 RACIOCÍNIO ESPACIAL QUALITATIVO PROBABILÍSTICO PARA LOCALIZAÇÃO ROBÓTICA

Neste capítulo vamos propor e explorar um método que combina teorias espaciais qualitativas com filtro Bayesiano, para executar a tarefa de auto localização qualitativa de um robô móvel. O método aqui apresentado foi aplicado em dois domínios com diferentes mapas qualitativos. Em ambos os domínios as informações foram extraídas das imagens capturadas pelo robô. O método aqui proposto segue os seguintes passos:

- definição do mapa qualitativo a partir de teorias espaciais topológicas de oclusão aplicáveis aos elementos marcos do domínio;
- formalização da localização em cada região do mapa aplicando as teorias raciocínio espacial qualitativo (REQ) que melhor se apliquem aos elementos marcos do domínio;
- adaptação do algoritmo de filtro Bayesiano às necessidades de inferência, ao tipo de elementos do domínio e na formalização da localização e do mapa;
- criação do sistema com os algoritmos para segmentação, reconhecimento de relações espaciais, localização com REQ, inferência com filtro Bayesiano.

4.1 Um Breve Histórico

Em Santos, Dee e Fenelon (2009) foi apresentada uma teoria lógica para auto localização qualitativa de um robô utilizando a teoria das relações qualitativas percebíveis sobre sombras, descrita na seção 2.2.3. As técnicas qualitativas lidam com certo grau de incerteza, seja porque as informações quantitativas são desnecessárias ou indisponíveis. Apesar disto, durante experimentos com um sistema robótico (SANTOS; DEE; FENELON, 2008; SANTOS; DEE; FENELON, 2009) percebemos que o alto grau de incerteza dos dados de visão impediam que as relações qualitativas fossem representadas de modo apropriado. Por exemplo, a alteração da intensidade luminosa do ambiente impedia que o sistema de visão segmentasse sombra e objeto de modo satisfatório.

As inferências com PQRS dependem da segmentação da imagem baseada em filtros de limiares descritos na seção 2.7.2. Para obter uma boa segmentação é necessário encontrar um limiar adequado para as mudanças de posição do robô. Em Fenelon et al. (2012) foi proposto inferir o limiar utilizando as informações de localização atual, as regiões vizinhas e o movimento

executado pelo robô. A partir dessas informações uma previsão sobre o robô continuar na mesma região ou de mudar de região era inferida, quando o PQRS retornava uma região diferente daquela prevista, o sistema fazia uma varredura procurando por um parâmetro que confirmasse a crença na localização, assim executando a auto calibração. Baseado nos trabalho apresentado na seção 3.3.2, em uma primeira tentativa para obter um mecanismo de previsão probabilístico, esboçou-se uma rede Bayesiana criada a partir da formalização em PQRS do domínio apresentado em Fenelon et al. (2012). Percebeu-se, no entanto que o que estávamos propondo era semelhante a um filtro Bayesiano. Nesta pesquisa o PQRS combinado com filtro Bayesiano foi a abordagem que melhor se encaixou a aplicação no robô móvel.

Assim, o que fizemos foi estender a teoria PQRS, combinando-a com um filtro Bayesiano. O filtro Bayesiano é utilizado para estimar a localização do robô e também para melhorar a segmentação, determinando o limiar dado o estado atual do robô. Contudo neste trabalho só foi possível aplicar o PQRS em ambiente com um único objeto. Em ambiente com múltiplos objetos, aplicamos o raciocínio espacial que utilizou a informação de oclusão entre pares de objetos no lugar das relações de oclusão entre objeto e sua sombra. O filtro Bayesiano com as informações espaciais qualitativas foi utilizado principalmente para estimar a auto localização relativa às regiões definidas nos mapas qualitativos.

4.2 Mapas Qualitativos

Seguindo o método, o primeiro passo é definir o mapa qualitativo referente a cada domínio. As regiões foram definidas utilizando as linhas de visão com base na oclusão entre os pares de elementos marcos. Foram utilizados para formalização da relação de oclusão entre elementos, o PQRS para os elementos sombra e o cálculo de oclusão de regiões (ROC) para os três pares de caixas.

Na seção 2.2.3.1 foi descrito como a teoria PQRS define as regiões pelas linhas de visão entre a fonte de luz L , um objeto o e sua sombra s , a partir da descrição de oclusão da sombra. No entanto, para uma aplicação prática, as regiões unidimensionais 2 e 4 foram expandidas conforme representado pelas áreas hachuradas na Figura 34a. Essa modificação resulta em uma representação qualitativa mais apropriada para que o robô execute a tarefa de auto localização. O mapa qualitativo (Figura 34b) também representa que a região da sombra causada pela parte superior do objeto chamada de topo da sombra ($top(s)$) será utilizada para definir as regiões. Aumentando a quantidade de objeto-sombra mais regiões são definidas no mapa, onde sub-regiões são formadas da sobreposição das regiões relativas aos diferentes objetos. Como não foi possível para este trabalho utilizar um número maior que um objeto-sombra, optou-se por utilizar múltiplos objetos e a percepção da relação de oclusão entre eles.

Com múltiplos objetos as regiões são definidas para cada par de objetos utilizando as relações topológicas como o cálculo de oclusão de região tal qual no PQRS. Para distinguir mais

Figura 34: Representação espacial (a) das regiões considerando um ambiente real, e (b) representação do mapa qualitativo com a cinco regiões distintas, mostrando os lados direito e esquerdo e as regiões ($Region_1$ até $Region_5$). As linhas de visão entre a fonte de luz L , o objeto O e sua sombra definem as fronteiras entre as regiões. Neste diagrama somente o topo da sombra é utilizado na definição das regiões.

Fonte: (a) Santos, Dee e Fenelon (2008) (b) Fenelon et al. (2012).

Figura 35: Uma representação da posição dos objetos no contexto deste experimento.

Fonte: autor.

as regiões utilizamos tanto um cálculo topológico quanto um cálculo de direção. Combinando as relações de direção relevantes para esse sistema com as relações de oclusão. As relações de direção que utilizamos considerando os dois objetos a e b do ponto de vista do observador ν são:

- $Left(a, b, \nu)$ significando que “ a está à esquerda de b do ponto de vista de ν ”;
- $Right(a, b, \nu)$ significando que “ a está à direita de b do ponto de vista de ν ”;
- $Front(a, b, \nu)$ significando que “ a está na frente de b do ponto de vista de ν ”;
- $Behind(a, b, \nu)$ significando que “ a está atrás de b do ponto de vista de ν ”.

As relações $Left$ e $Right$ são simétricas entre elas (form.(4.1)) e exclusivamente disjuntas considerando o par, assim como as relações $Front$ e $Behind$ (form.4.4)). Essas quatro relações são assimétricas isoladamente (form. (4.2), (4.3), (4.5) e (4.6)). Considerando um par de objetos convexos, os quais do ponto de vista ν a está à esquerda de b então do mesmo ponto de vista b não estará a esquerda de a . Essa considerações estão formalizadas nos axiomas abaixo:

$$Left(a, b, \nu) \Leftrightarrow Right(b, a, \nu); \quad (4.1)$$

$$Left(a, b, \nu) \rightarrow \neg Left(b, a, \nu); \quad (4.2)$$

$$Right(a, b, \nu) \rightarrow \neg Right(b, a, \nu); \quad (4.3)$$

$$Front(a, b, \nu) \Leftrightarrow Behind(b, a, \nu); \quad (4.4)$$

$$Front(a, b, \nu) \rightarrow \neg Front(b, a, \nu); \quad (4.5)$$

$$Behind(a, b, \nu) \rightarrow \neg Behind(b, a, \nu). \quad (4.6)$$

Uma imagem representa a cena do ponto de vista ν do observador ($PeopleBot^{TM}$), portanto neste trabalho as direções acima foram definidas considerando as seguintes restrições:

- A posição da imagem é relativa à posição vertical do robô $PeopleBot^{TM}$, correspondendo à posição ereta do observador.
- A câmera está fixa, a imagem corresponde à posição no espaço, ou seja, as posições "acima" e "abaixo" seguem as definições usuais.
- Os objetos são convexos, sem buraco e sem movimento.
- A imagem tem um sistema de coordenadas com a origem dos eixos na lateral direita superior (relativo à imagem).

Com a Figura 36 vamos exemplificar como são verificadas essas direções a partir do ponto de vista do observador. Considerando os eixos de coordenadas da imagem, a lateral esquerda e a base do objeto pode-se inferir a relação entre o par de objetos a e b . Assim, para identificar se um objeto está à direita ou à esquerda do outro objeto compara-se a distância de

Figura 36: Uma representação gráfica para a definição das direções *Left*, *Right*, *Front* e *Behind* (esquerda, direita, em frente, atrás) relativos à visão do robô através da imagem. Os eixos de coordenadas da imagem tem origem no canto acima-direito relativo à imagem.

Fonte: autor.

Figura 37: Diagrama representando nove regiões formadas pelas linhas de visão.

Fonte: autor.

ambos em relação à origem do eixo x . O objeto mais próximo está posicionado à esquerda do objeto mais distante da origem de x , enquanto este está à direita do objeto mais próximo da origem de x . Para identificar qual o objeto está à frente ou atrás comparamos a distância da base relativa à origem do eixo y de ambos os objetos. O objeto mais distante está à frente do objeto mais próximo à origem do eixo y , enquanto que este está atrás do objeto mais distante da origem do eixo y . Pela Figura 36 do ponto de vista do observador podemos dizer que a está à esquerda e em frente de b , enquanto que b está à direita e atrás de a .

Nove regiões são definidas a partir das linhas de visão entre dois objetos conforme diagrama da Figura 37. As linhas de visão são tangentes internas e externas relativas aos dois objetos (A, B). Do ponto de vista de um observador ν que esteja sobre as linhas de visão tan-

Figura 38: As duas figuras representam as regiões e as relações espaciais percebidas em cada uma (a) As relações de oclusão e (b) as relações de direção para cada região, todas as relações percebidas por um observador ν . Somente a região *Region₉* não é possível visualizar o par em nenhuma posição do observador.

Fonte: autor.

gentes internas, se um objeto acabou de encostar ao outro então a relação *NonOccludesEC* é percebida. Quando o observador está sobre uma linha tangente externa perceberá a relação *TotallyOccludesTPP* ou o seu inverso, significando que um dos objetos acabou de ser totalmente ocluído pelo outro objeto. No entanto, diferente do que ocorreu com as regiões similares definidas com sombras (PQRS), não vamos expandir essas regiões, ou seja, as linhas de fronteira não serão definidas como região. Podemos observar também que existe simetria entre as regiões assim como as regiões definidas pelo PQRS.

Neste contexto, as relações entre os objetos *A* e *B* para localização são definidas pelas relações de oclusão e pela posição entre as imagens dos objetos no frame. A localização do robô em quase todas as regiões é definida a partir do que o observador pode perceber na imagem sobre o tipo de relação entre pares de objetos (Figura 38). Deste modo, a localização do observador numa determinada região é definida pela conjunção de relações do par de objetos percebida. Abaixo as fórmulas de localização e com exemplos de imagens representadas na Figura 39 em que *A* é a caixa azul e *B* a caixa verde:

$$\begin{aligned} located(Region_1, \nu, A, B) \leftarrow & \text{PartiallyOccludesPO}(A, B, \nu) \wedge \\ & \text{Left}(A, B, \nu) \wedge \text{Front}(A, B, \nu). \end{aligned} \quad (4.7)$$

A fórmula (4.7) significa que: quando na *Region₁* do ponto de vista ν - *A* oclui parcialmente *B* e *A* está à esquerda e a frente de *B* (Figura 39a).

$$\begin{aligned} located(Region_2, \nu, A, B) \leftarrow & \text{TotallyOccludesNTPP}(A, B, \nu) \wedge \\ & \text{Front}(A, B, \nu). \end{aligned} \quad (4.8)$$

A fórmula (4.8) significa que: quando na *Region₂* do ponto de vista ν - *A* oclui totalmente *B* e *A* está na frente de *B*. Ou seja, somente *A* é visto e está mais próximo do observador (Figura 39b).

$$\begin{aligned} located(Region_3, \nu, A, B) \leftarrow & \text{PartiallyOccludesPO}(A, B, \nu) \wedge \\ & \text{Right}(A, B, \nu) \wedge \text{Front}(A, B, \nu). \end{aligned} \quad (4.9)$$

A fórmula (4.9) significa que: quando na $Region_3$ do ponto de vista ν - A oclui parcialmente B e A está à direita e na frente de B (Figura 39c).

$$located(Region_4, \nu, A, B) \leftarrow NonOccludesDC(A, B, \nu) \wedge Right(A, B, \nu). \quad (4.10)$$

A fórmula (4.10) significa que: quando na $Region_4$ do ponto de vista ν - A e B estão livres de oclusão e A está à direita de B (Figura 39d).

$$located(Region_5, \nu, A, B) \leftarrow PartiallyOccludesPO^{-1}(A, B, \nu) \wedge Right(A, B, \nu) \wedge Behind(A, B, \nu). \quad (4.11)$$

A fórmula (4.11) significa que: quando na $Region_5$ do ponto de vista ν - B oclui parcialmente A e A está à direita e atrás de B (Figura 39e).

$$located(Region_6, \nu, A, B) \leftarrow TotallyOccludesNTPPI^{-1}(A, B, \nu) \wedge Behind(A, B, \nu). \quad (4.12)$$

A fórmula (4.12) significa que: quando na $Region_6$ do ponto de vista ν - B oclui totalmente A e A está atrás de B (Figura 39f).

$$located(Region_7, \nu, A, B) \leftarrow PartiallyOccludesPO^{-1}(A, B, \nu) \wedge Left(A, B, \nu) \wedge Behind(A, B, \nu). \quad (4.13)$$

A fórmula (4.13) significa que: quando na $Region_7$ do ponto de vista ν - B oclui parcialmente A e A está à esquerda e atrás de B (Figura 39g).

$$located(Region_8, \nu, A, B) \leftarrow NonOccludesDC(A, B, \nu) \wedge Left(A, B, \nu). \quad (4.14)$$

A fórmula (4.14) significa que: quando na $Region_8$ do ponto de vista ν - A e B estão livres de oclusão e A está à esquerda de B (Figura 39h).

$$located(Region_9, \nu, A, B) \leftarrow NonOccludesDC(A, B). \quad (4.15)$$

A fórmula (4.15) significa que: quando na $Region_9$ do ponto de vista ν - A e B estão livres de oclusão. Contudo, nesta região somente um observador com campo de visão de 360° poderia observar A e B na mesma imagem (Figuras 39i e 39j respectivamente).

Mapas qualitativos podem ser formados a partir da disposição espacial de pares de objetos no ambiente. Para analisar o comportamento do algoritmo de raciocínio espacial qualitativo probabilístico colocamos em nosso ambiente mais um objeto formando três pares em cena (Figuras 40a e 40b). No mapa qualitativo o número de regiões depende da disposição relativa dos objetos. Por exemplo, com os três objetos alinhados 14 regiões são formadas (Figura 40a), esta configuração não será utilizadas neste trabalho. Com os objetos dispostos de forma triangular formando as 28 regiões representadas no mapa qualitativo da Figura 40b, as quais são resultados da sobreposição das nove regiões de cada um dos três pares de objetos.

Figura 39: Exemplo de uma visualização para (a) *Region*₁, (b) *Region*₂, (c) *Region*₃, (d) *Region*₄, (e) *Region*₅, (f) *Region*₆, (g) *Region*₇, (h) *Region*₈, (i e j) *Region*₉ onde os objetos estão livres de oclusão, mas o campo de visão não permite visualizar os dois em uma mesma cena.

Figura 40: Representações dos mapas qualitativos formados pelas linha de visão com duas configurações de posição que os três objetos podem assumir neste contexto.

Fonte: autor.

A seguir o próximo passo do método: a definição de localização para cada região conforme o mapa qualitativo da Figura 34b e exemplos das relações entre objeto e sombra com imagens capturadas pelo robô móvel *PeopleBotTM* em cada região. Seguido da definição de localização referente as 28 regiões do mapa da Figura 40b.

4.3 Definindo a Localização do Robô

Nesta seção começamos a definição da localização nas cinco regiões relativas ao mapa da Figura 34b, com alguns detalhes sobre cada região além de fotos para ajudar a visualizar as relações entre sombra e objetos. A localização do mapa da Figura 40b será somente formalizadas por seus axiomas, pois por ter exemplificado as regiões e localização para um par de objetos na seção anterior, não repetiremos os detalhes de cada região.

4.3.1 Localização na região 1

A localização de um robô na região 1 é definida pelo fórmula (4.16):

$$\begin{aligned} located(1, \nu, o, s) \leftarrow & Shadows(s, o, Scr, L) \wedge \\ & NonOccludesDC(o, top(s), \nu) \wedge \nu \neq o. \end{aligned} \quad (4.16)$$

Do ponto de vista do observador, nesta região, o topo da sombra nunca é ocluído pelo objeto que a projetou. Uma representação do topo da sombra pode ser vista na Figura 41c onde somente a região que representa o topo da sombra está destacada em vermelho. Nas outras imagens toda a região equivalente à sombra do objeto está destacada em vermelho para uma melhor

Figura 41: Imagens capturadas com o robô posicionado em distintos pontos na região 1. A imagem (c) mostra somente o topo da sombra em destaque. Note que o objeto não oclui a sombra.

Fonte: autor (c) Fenelon et al. (2012).

visualização. O topo da sombra sempre aparece desconectado do objeto em qualquer ponto da região 1 (Figura 41). Cada região tem algumas características típicas percebidas na imagem. Como exemplo de características desta região: a área da sombra é maior que 90% da área do objeto; no eixo y da imagem a sombra está abaixo do objeto. Esta região consiste da maior parte da área de atuação do robô (Figura 34b), nela o objeto está sempre entre a fonte de luz e o robô. As regiões 2 são suas únicas regiões vizinhas.

4.3.2 Localização na região 2

A região 2 possui a segunda maior região, considerando-se a soma de ambas regiões 2 direita e esquerda (Figura 34b). A localização na região 2 é definida pelo fórmula (4.17):

$$\begin{aligned} located(2, \nu, o, s) &\leftarrow Shadow(s, o, Scr, L) \wedge \\ &NonOccludesEC(o, top(s), \nu) \wedge \nu \neq o. \end{aligned} \quad (4.17)$$

Nesta região o objeto começa a ocluir a sombra, ainda assim, a maior parte da sombra pode ser percebida do ponto de vista de um observador. A imagem (Figura 42) aponta para as duas laterais

Figura 42: Imagens capturadas com o robô na região 2. Note que o objeto começa a ocluir a sombra. Todavia ainda se vê a sombra do topo do objetivo e as laterais da sombra do objeto indicados em (a).

Fonte: autor.

Figura 43: Imagens capturadas com o robô na região 3. A sombra nesta região aparece com um dos lados totalmente ocluído pelo objeto. Existe uma conexão entre a parte superior do objeto e sua sombra projetada.

Fonte: autor.

da sombra, as quais podem ser observadas na sombra projetada enquanto o observador estiver na região 2, entretanto uma delas aparece parcialmente. Na Figura 42 as fronteiras foram marcadas com barbante, e a fronteira com a região 1 é aquela na qual a oclusão da lateral da sombra se inicia. Outras características típicas desta região são: o tamanho da sua área está entre 50% e 90% do tamanho da área do objeto por causa da oclusão; em relação ao eixo y da imagem as bases dos retângulos que contêm e definem o objeto e sombra na imagem estão aproximadamente no mesmo nível; a região do topo da sombra fica totalmente desconectada do objeto. Esta região faz fronteira com região 1 e também é vizinha da região 3.

Figura 44: Imagens capturadas com o robô na região 4. A sombra nesta região quase desaparece oclusa pelo objeto. Contudo ainda existe uma conexão entre a parte superior do objeto e a sombra.

Fonte: autor.

4.3.3 Localização na região 3

A região 3 possui a terceira área em extensão. A localização na região 3 é definida pela fórmula (4.18):

$$\begin{aligned} located(3, \nu, o, s) \leftarrow & Shadow(s, o, Scr, L) \wedge \\ & PartiallyOccludesPO(o, top(s), \nu) \wedge \nu \neq o. \end{aligned} \quad (4.18)$$

Do ponto de vista do observador, o topo da sombra está ocluído parcialmente pelo objeto. Nas imagens somente uma das laterais da sombra pode ser observada (Figura 43). Outros exemplos de características típicas desta região são: a área da sombra é menor que 50% da área do objeto; devido à oclusão parcial, o topo da sombra está conectado a lateral do objeto. A região 3 fica localizada entre as regiões 2 e 4.

4.3.4 Localização na região 4

Esta região é uma das menores regiões da área de atuação do robô (Figura 34b). A localização na região 4 é definida pela fórmula (4.19):

$$\begin{aligned} located(4, \nu, o, s) \leftarrow & Shadow(s, o, Scr, L) \wedge \\ & TotallyOccludesTPPI(o, top(s), \nu) \wedge \nu \neq o. \end{aligned} \quad (4.19)$$

Do ponto de vista do observador, somente uma pequena parte da sombra do objeto é percebida. Quase todo o topo da sombra está ocluído pelo objeto. Na imagem (Figura 44) a relação entre as áreas da sombra e do objeto difere da relação na região 3 principalmente pelo tamanho da área da sombra. A sombra é bem menor que o objeto, aparecendo na imagem como uma região estreita de sombra na lateral do objeto. Esta região é vizinha das regiões 3 e 5.

Figura 45: Imagem capturada com o robô na região 5. A sombra do objeto não aparece na imagem por estar totalmente oclusa. O observador está na região equivalente ao ponto de vista da luz.

Fonte: autor.

4.3.5 Localização na região 5

A região 5 é a menor de todas as regiões da área de atuação do robô (Figura 34b). Esta região está na posição oposta a região 1 e como esta é única. Tendo como vizinhas somente as regiões 4 esquerda e direita. A localização na região 5 é definida pela fórmula (4.20):

$$\begin{aligned} located(5, \nu, o, s) \leftarrow & Shadow(s, o, Scr, L) \wedge \\ & TotallyOccludesNTPPI(o, top(s), \nu) \wedge \nu \neq o. \end{aligned} \quad (4.20)$$

Do ponto de vista do observador, a sombra do objeto existe, mas não pode ser percebida, pois está totalmente oclusa pelo objeto. A característica principal desta região é que somente o objeto pode ser observado (Figura 45) na cena. Com essa região a formalização para a localização no mapa da Figura 34b está completa.

4.3.6 Localização nas regiões R_1 até R_{28}

Para o mapa da Figura 46 os elementos marcos são caixas reconhecidas pelas suas cores diferenciadas azul, verde e vermelho, representadas pelas letras b , g e r que formam três pares: azul e verde (b, g); azul e vermelho (b, r); verde e vermelho (g, r), cada par com nove regiões distintas. A posição das caixas, a distância entre elas e sua posição exata no ambiente são desconhecidas. Um robô pode se localizar em uma das regiões a partir da observação das relações dos objetos. Para isso utilizamos $located(R_i, \nu, b, g, r)$ que significa que um observador ν está em determinada região R_i , sendo $i = \{1, 2, \dots, 28\}$, relativa às três caixas, b, g e r . As fórmulas 4.21 a 4.48 utilizam de disjunções de conjunções de relações espaciais para cada par de objetos para definir a localização nas 28 regiões:

$$\begin{aligned} located(R_1, \nu, b, g, r) \leftarrow & PartiallyOccludesPO^{-1}(b, g, \nu) \wedge Right(b, g, \nu) \wedge Behind(b, g, \nu) \vee \\ & NonOccludesDC(b, r, \nu) \wedge Left(b, r, \nu) \vee \\ & NonOccludesDC(g, r, \nu) \wedge Left(g, r, \nu); \end{aligned} \quad (4.21)$$

Figura 46: Mapa qualitativo da disposição espacial dos objetos e a representação das regiões.

Fonte: autor.

$$\begin{aligned} located(R2, \nu, b, g, r) \leftarrow & TotallyOccludesNTPPI^{-1}(b, g, \nu) \wedge Behind(b, g, \nu) \vee \\ & NonOccludesDC(g, r, \nu) \wedge Left(g, r, \nu); \end{aligned} \quad (4.22)$$

$$\begin{aligned} located(R3, \nu, b, g, r) \leftarrow & PartiallyOccludesPO^{-1}(b, g, \nu) \wedge Left(b, g, \nu) \wedge Behind(b, g, \nu) \vee \\ & NonOccludesDC(b, r, \nu) \wedge Left(b, r, \nu) \vee \\ & NonOccludesDC(g, r, \nu) \wedge Left(g, r, \nu); \wedge Front(g, r, \nu) \end{aligned} \quad (4.23)$$

$$\begin{aligned} located(R4, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Left(b, g, \nu) \wedge Behind(b, g, \nu) \vee \\ & NonOccludesDC(b, r, \nu) \wedge Left(b, r, \nu) \vee \\ & NonOccludesDC(g, r, \nu) \wedge Left(g, r, \nu) \wedge Front(g, r, \nu); \end{aligned} \quad (4.24)$$

$$\begin{aligned} located(R5, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Left(b, g, \nu) \wedge Behind(b, g, \nu) \vee \\ & NonOccludesDC(b, r, \nu) \wedge Left(b, r, \nu) \vee \\ & PartiallyOccludesPO(g, r, \nu) \wedge Left(g, r, \nu) \wedge Front(g, r, \nu); \end{aligned} \quad (4.25)$$

$$\begin{aligned} located(R6, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Left(b, g, \nu) \vee \\ & TotallyOccludesNTPP(g, r, \nu) \wedge Front(g, r, \nu); \end{aligned} \quad (4.26)$$

$$\begin{aligned} located(R7, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Left(b, g, \nu) \vee \\ & NonOccludesDC(b, r, \nu) \wedge Left(b, r, \nu) \vee \\ & PartiallyOccludesPO(g, r, \nu) \wedge Right(g, r, \nu) \wedge Front(g, r, \nu); \end{aligned} \quad (4.27)$$

$$\begin{aligned}
located(R8, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Left(b, g, \nu) \vee \\
& NonOccludesDC(b, r) \wedge Left(b, r, \nu) \wedge Front(b, r, \nu) \vee \\
& NonOccludesDC(g, r) \wedge Right(g, r, \nu) \wedge Front(g, r, \nu);
\end{aligned} \tag{4.28}$$

$$\begin{aligned}
located(R9, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Left(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Right(g, r, \nu);
\end{aligned} \tag{4.29}$$

$$\begin{aligned}
located(R10, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Left(b, g, \nu) \vee \\
& PartiallyOccludesPO(b, r, \nu) \wedge Left(b, r, \nu) \wedge Front(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Right(g, r, \nu);
\end{aligned} \tag{4.30}$$

$$\begin{aligned}
located(R11, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Left(b, g, \nu) \\
& \vee TotallyOccludesNTPP(b, r, \nu) \wedge Front(b, r, \nu)
\end{aligned} \tag{4.31}$$

$$\begin{aligned}
located(R12, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Left(b, g, \nu) \wedge Front(b, g, \nu) \vee \\
& PartiallyOccludesPO(b, r, \nu) \wedge Right(b, r, \nu) \wedge Front(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Right(g, r, \nu);
\end{aligned} \tag{4.32}$$

$$\begin{aligned}
located(R13, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Left(b, g, \nu) \wedge Front(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Right(b, r, \nu) \wedge Front(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Right(g, r, \nu);
\end{aligned} \tag{4.33}$$

$$\begin{aligned}
located(R14, \nu, b, g, r) \leftarrow & PartiallyOccludesPO(b, g, \nu) \wedge Left(b, g, \nu) \wedge Front(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Right(b, r, \nu) \wedge Front(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Right(g, r, \nu);
\end{aligned} \tag{4.34}$$

$$\begin{aligned}
located(R15, \nu, b, g, r) \leftarrow & TotallyOccludesNTPP(b, g, \nu) \wedge Front(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Right(b, r, \nu);
\end{aligned} \tag{4.35}$$

$$\begin{aligned}
located(R16, \nu, b, g, r) \leftarrow & PartiallyOccludesPO(b, g, \nu) \wedge Right(b, g, \nu) \wedge Front(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Right(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Right(g, r, \nu);
\end{aligned} \tag{4.36}$$

$$\begin{aligned}
located(R17, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \wedge Front(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Right(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Right(g, r, \nu) \wedge Behind(g, r, \nu);
\end{aligned} \tag{4.37}$$

$$\begin{aligned}
located(R18, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Right(g, r, \nu);
\end{aligned} \tag{4.38}$$

$$\begin{aligned}
located(R19, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Right(b, r, \nu) \vee \\
& PartiallyOccludesPO^{-1}(g, r, \nu) \wedge Right(g, r, \nu) \wedge Behind(g, r, \nu);
\end{aligned} \tag{4.39}$$

$$\begin{aligned}
located(R20, \nu, b, g, r) \leftarrow & NonOccludesDC(b, r, \nu) \wedge Right(b, r, \nu) \vee \\
& TotallyOccludesNTPPI^{-1}(g, r, \nu) \wedge Behind(g, r, \nu);
\end{aligned} \tag{4.40}$$

$$\begin{aligned}
located(R21, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Right(b, r, \nu) \wedge Behind(b, r, \nu) \vee \\
& PartiallyOccludesPO^{-1}(g, r, \nu) \wedge Left(g, r, \nu) \wedge Behind(g, r, \nu);
\end{aligned} \tag{4.41}$$

$$\begin{aligned}
located(R22, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Right(b, r, \nu) \wedge Behind(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Left(g, r, \nu) \wedge Behind(g, r, \nu);
\end{aligned} \tag{4.42}$$

$$\begin{aligned}
located(R23, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \vee \\
& PartiallyOccludesPO^{-1}(b, r, \nu) \wedge Right(b, r, \nu) \wedge Behind(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Left(g, r, \nu) \wedge Behind(g, r, \nu);
\end{aligned} \tag{4.43}$$

$$\begin{aligned}
located(R24, \nu, b, g, r) \leftarrow & TotallyOccludesNTPPI^{-1}(b, r, \nu) \wedge Behind(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Left(g, r, \nu) \wedge Behind(g, r, \nu);
\end{aligned} \tag{4.44}$$

$$\begin{aligned}
located(R25, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \vee \\
& PartiallyOccludesPO^{-1}(b, r, \nu) \wedge Left(b, r, \nu) \wedge Behind(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Left(g, r, \nu) \wedge Behind(g, r, \nu);
\end{aligned} \tag{4.45}$$

$$\begin{aligned}
located(R26, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Left(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Left(g, r, \nu);
\end{aligned} \tag{4.46}$$

$$\begin{aligned}
located(R27, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Left(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu);
\end{aligned} \tag{4.47}$$

$$\begin{aligned}
located(R28, \nu, b, g, r) \leftarrow & NonOccludesDC(b, g, \nu) \wedge Right(b, g, \nu) \vee \\
& NonOccludesDC(b, r, \nu) \wedge Left(b, r, \nu) \vee \\
& NonOccludesDC(g, r, \nu) \wedge Right(b, g, \nu).
\end{aligned} \tag{4.48}$$

Assim, com a formalização utilizando raciocínio espacial qualitativa para localização em cada uma das regiões relativas aos mapas qualitativos, o filtro Bayesiano vem para completar nosso algoritmo raciocínio espacial qualitativo probabilístico, e será descrito na próxima seção.

4.4 Filtro Bayesiano

Essa etapa consiste em adaptar o filtro Bayesiano às necessidades de resposta do sistema. O tipo de técnica de visão utilizada deve ser considerado caso o filtro Bayesiano tenha influência em algum parâmetro, como foi o caso no domínio com informação de sombras que além da localização precisou de uma interferência em seu parâmetro de visão. O modelo de imagem deve ser criado a partir das relações que definem as localizações, enquanto que o modelo de transição a partir do conceito de um diagrama de vizinhança sobre um mapa qualitativo.

O filtro Bayesiano é um algoritmo recursivo dividido em duas etapas, a etapa de previsão e a de atualização de medidas. Em ambos os filtros o estado s_t sempre indica a região onde o robô está localizado e as informações espaciais qualitativas obtidas a partir das imagens são evidências, as quais são representadas por (e_t) . A seguir os filtros Bayesianos adaptados para cada um dos domínios.

4.4.1 Filtro Bayesiano sobre informações qualitativas do PQRS

Para o mapa da Figura 34b as evidências são termos das relações do PQRS. O sistema deve inferir em qual a região s_t o robô está localizado, e também o limiar Th utilizado para segmentar objeto e sombra nas imagens, ambos devem maximizar a distribuição de crença do estado atual $bel(s_t)$ ou seja $P(s_t, Th|e_{0:t})$ que é a probabilidade da região s_t e o limiar Th condicionada a todas as evidências percebidas $(e_{0:t})$. Para calcular essa distribuição de crença utilizamos:

$$\begin{aligned}
bel(s_t) &= P(s_t, Th|e_{0:t}) \\
&= P(s_t, Th|e_{0:t-1}, e_t) \\
&= \eta P(e_t, Th|s_t, e_{0:t-1}) P(s_t|e_{0:t-1}),
\end{aligned} \tag{4.49}$$

onde η é a constante de normalização. Com a hipótese de Markov, o estado seguinte é independente das evidências anteriores $e_{0:t-1}$; assim:

$$bel(s_t) = \eta P(e_t, Th|s_t) P(s_t|e_{0:t-1}). \tag{4.50}$$

Esta fórmula (4.50) representa a atualização de medidas. O primeiro termo $P(e_t, Th|s_t)$ é o modelo da imagem (form.(4.50)) e incorpora a nova evidência e_t . O segundo termo é a previsão

para o estado atual, ou seja, a crença posterior $P(s_t|e_{0:t-1}) = \overline{bel}(s_t)$, esta deve ser calculada antes de incorporar a evidência e_t .

A crença posterior $P(s_t|e_{0:t-1})$ representa o passo de previsão do estado s_t , obtido previamente pelo condicionamento no estado anterior s_{t-1} , conseqüentemente,

$$\begin{aligned}\overline{bel}(s_t) &= P(s_t|e_{0:t-1}) \\ &= \sum_{s_{t-1}} P(s_t|s_{t-1}, e_{0:t-1})P(s_{t-1}|e_{0:t-1}) \\ &= \sum_{s_{t-1}} P(s_t|s_{t-1})P(s_{t-1}|e_{0:t-1}).\end{aligned}\tag{4.51}$$

Na fórmula (4.51) o termo $P(s_t|s_{t-1})$ é o valor de probabilidade da transição no estado s_t , e representa o modelo de transição. Para obter $P(s_{t-1}|e_{0:t-1})$ marginalizamos em Th a distribuição de crença em $bel(s_{t-1})$ (form.(4.49)):

$$P(s_{t-1}|e_{0:t-1}) = \sum_{Th} P(s_{t-1}, Th|e_{0:t-1}).\tag{4.52}$$

Portanto, a crença posterior ou predição é resultado da combinação do modelo de transição com a crença do estado anterior, assim substituindo a fórmula (4.52) na fórmula(4.51) temos que (form. 4.53):

$$\overline{bel}(s_t) = P(s_t|e_{0:t-1}) = \sum_{s_{t-1}} P(s_t|s_{t-1}) \sum_{Th} P(s_{t-1}, Th|e_{0:t-1}).\tag{4.53}$$

E finalmente, substituindo fórmula (4.54) na fórmula (4.50) temos a crença no estado atual dada por (form. 4.54):

$$\begin{aligned}bel(s_t) &= \eta P(e_t, Th|s_t) \sum_{s_{t-1}} P(s_t|s_{t-1}) \sum_{Th} P(s_{t-1}, Th|e_{0:t-1}) \\ &= \eta P(e_t, Th|s_t) \overline{bel}(s_t).\end{aligned}\tag{4.54}$$

Na seção seguinte está descrito o filtro Bayesiano para a auto localização com múltiplos objetos.

4.4.2 Filtro Bayesiano para auto localização com múltiplos objetos

Para o mapa qualitativo da Figura 46 o filtro Bayesiano foi utilizado para inferir a região onde o robô está localizado. Essa localização é representado pelo estado s_t que maximiza a distribuição de crença $bel(s_t)$. Com um valor do modelo de imagem para as dados de relações espaciais percebidas e_t , e a distribuição de crença posterior $\overline{bel}(s_t)$, temos a distribuição de crença do estado atual para todos os s_t dada pela fórmula (4.55):

$$bel(s_t) = \eta P(e_t|s_t) \overline{bel}(s_t).\tag{4.55}$$

A etapa de previsão calcula a distribuição de crença posterior $bel(s_t)$ baseado na distribuição de crença do estado anterior $bel(s_{t-1})$. A previsão $\overline{bel}(s_t)$ é obtida pela soma do produto

Figura 47: Mapa qualitativo das áreas de atuação divididas em oito regiões.

Fonte: autor.

do modelo de transição com a distribuição de crença do estado anterior (form. (4.56)):

$$\overline{bel}(s_t) = \sum_{s_{t-1}} P(s_t | s_{t-1}) bel(s_{t-1}). \quad (4.56)$$

A seguir a descrição sobre como os modelos de imagem e transição foram formados.

4.4.3 Modelo de imagem

O modelo de imagem é um dos fatores que fazem parte da etapa de atualização de informação de nosso filtro Bayesiano. O modelo de imagem é o valor de probabilidade de que uma relação seja percebida na imagem dado a região em que o robô está. A partir da localização em cada região formalizada nas seções anteriores, as evidências são as relações entre os elementos observadas pelo robô.

Para o sistema com sombras as evidências são as relações entre a sombra e objeto percebidas pelo robô. A localização foi formalizada pela teoria PQRS. O sistema de visão computacional depende do valor de limiar. Assim o modelo deve indicar (para cada limiar) a frequência em que as evidências ocorrem para cada região. O modelo de imagem é uma matriz com a porcentagem de relações por região para cada limiar.

Como no sistema de auto localização com sombras (seção 4.4.1). A formalização da localização feita com a disjunção de conjunções de relações espaciais entre os pares de elementos (caixas). O modelo de imagem indica os valores de probabilidade de o sensor perceber corretamente a evidências e_t dado que ele está localizado na região relacionada a essa evidência s_t representado pelo termo $P(e_t | s_t)$.

4.4.4 Modelo de transição

O modelo de transição representa a transição do robô ao passar de uma região para outra. Esse modelo é utilizado na etapa de previsão do filtro Bayesiano. Ao considerar o mapa qualitativo da Figura 47 a probabilidade de mudança entre duas regiões depende em parte do tamanho da região. Assim, uma região grande tem probabilidade maior que o robô continue

Figura 48: Diagrama conceitual de vizinhança (CND) das relações que determinam a região na qual o robô estava localizado. As transições estão representadas pelas setas e a mudança ocorre sempre para uma relação (região) vizinha.

Fonte: adaptada de Cui, Cohn e Randell (1992).

nessa região, isto é representado como uma constante (numeral) na Tabela 12. A outra parte depende do movimento realizado pelo robô.

O deslocamento do robô pela área de atuação segue o conceito espacial de transição contínua, do mesmo modo que o conceito de transição de relações do diagrama conceitual de vizinhança (CND) discutido na seção 2.2. A Figura 48 representa um CND das transições das relações que são percebidas em cada uma das regiões do mapa da Figura 47. Em cada região ocorre uma relação diferente entre sombra e objeto (Figura 48). Podemos afirmar que cada relação tem um conjunto de características que a distingue. Algumas dessas características podem ser compartilhadas, principalmente durante uma transição entre as relações vizinhas. Ao analisar as imagens com sombra e objetos as relações entre ambos mudam conforme o robô se desloca. Por esse motivo, utilizamos um fator w que é o valor proporcional ao conjunto de características observadas para cada relação vinculada a uma região. Na Tabela 12 este fator modula a probabilidade da mudança de estado. Por exemplo, considerando duas regiões vizinhas, quando o robô se desloca em direção a uma das regiões vizinhas as características desta região aumentam enquanto que as características daquela que o robô se afasta diminuem.

O modelo de transição para múltiplos objetos foi criado de modo análogo ao modelo de transição para o mapa da Figura 47 (sombra-objeto). Para o mapa (Figura 46) o modelo de transição é uma matriz 28×28 com os valores de probabilidade da transição de uma região para outra do mapa. Cada valor deste modelo compõe-se de um número real fixo associado a uma função. O valor real fixo é baseado na configuração do mapa (Figura 46) considerando a vizinhança e o tamanho da região. Intuitivamente em regiões maiores a probabilidade de continuar nelas é maior. A função está relacionada a percepção de movimento através da imagem, quando o robô se aproxima de uma fronteira, algumas características que definem as relações espaciais podem ser observadas, o que permite distinguir entre as regiões vizinhas, aquela que o robô está se direcionando, reforçando o valor de probabilidade do modelo de transição. Na Figura 49 uma ilustração topológica das 28 regiões como nós, e as transições entre regiões vizinhas representadas pela seta contínua, a seta tracejada são para regiões que não têm fronteiras, mas estão próximas.

A próxima etapa refere-se a determinar os algoritmos baseados nas formalizações de localização que fazem parte do sistema responsável pela inferência de localização do robô.

Tabela 12: O modelo de movimento considerando que o robô tem uma navegação no sentido horário. Este modelo é baseado no tamanho qualitativo e na transição das regiões. O valores são alterados por uma função de mudança percebida w que representa a paralaxe causada pelo movimento do robô

Estados para sentido horário		$s_t = Region_1$	$s_t = Region_2$	$s_t = Region_3$	$s_t = Region_4$	$s_t = Region_5$
$s_{t-1} = Region_1$ right		$0,90w$	$0,10w$	$0,00$	$0,00$	$0,00$
$s_{t-1} = Region_2$ right		$0,01w$	$0,80w$	$0,19w$	$0,00$	$0,00$
$s_{t-1} = Region_3$ right		$0,00$	$0,01w$	$0,75w$	$0,24w$	$0,00$
$s_{t-1} = Region_4$ right		$0,00$	$0,00$	$0,01w$	$0,70w$	$0,29w$
$s_{t-1} = Region_5$		$0,00$	$0,00$	$0,00$	$0,35w$	$0,65w$
$s_{t-1} = Region_4$ left		$0,00$	$0,00$	$0,29w$	$0,70w$	$0,01w$
$s_{t-1} = Region_3$ left		$0,00$	$0,24w$	$0,75w$	$0,01w$	$0,00$
$s_{t-1} = Region_2$ left		$0,19w$	$0,80 w$	$0,01 w$	$0,00$	$0,00$
$s_{t-1} = Region_1$ left		$0,99 w$	$0,01w$	$0,00$	$0,00$	$0,00$

Fonte: Fenelon et al. (2013).

Figura 49: Representação das regiões como nós, e das transições entre regiões próximas como setas: setas contínuas regiões com fronteira, tracejadas regiões sem fronteira entre elas.

Fonte: autor.

Figura 50: Janelas para calibração manual dos parâmetros de visão utilizados no filtro de limiar para segmentar objeto e sombra, os valores escolhidos pela barra de rolagem ficam indicados à esquerda da barra. Da esquerda para direita da figura: define o limiar Th que separa objeto e sombra do ambiente; define o limiar que separa objeto da sombra; define o limiar que separa a sombra do objeto.

Fonte: autor.

4.5 Algoritmos Utilizados no Sistema de Auto Localização

Neste sistema, existe um algoritmo principal que captura as imagens, outro responsável pelo reconhecimento dos objetos, outro pelo reconhecimento das relações espaciais entre os elementos, outro pela inferência de localização com raciocínio espacial que retorna ao algoritmo principal a evidência que será utilizada pelo filtro Bayesiano. Ambos os algoritmos principais (Algoritmo 1 e Algoritmo 3) inicializam-se com um valor de distribuição de crença igual para todas as regiões, isso porque o robô pode iniciar em qualquer região do mapa. O Algoritmo 1 executa uma etapa na qual um operador deve executar uma calibração manual dos parâmetros de visão. Para cada imagem capturada os algoritmos principais chamam repetidamente o algoritmo que retorna a evidência com máxima probabilidade $P(E)$ (Algoritmos 2 e 4).

4.5.1 Algoritmo de auto localização e auto calibração com PQRS probabilístico

O Algoritmo 1 inicializa o sistema com uma distribuição de crença posterior uniforme sobre a posição do robô (linha 1). A linha 2 chama a função de calibração dos parâmetros de visão, os limiares (inclusive Th) são escolhidos pelo operador através de três janelas como aquelas na Figura 50, assumindo assim que o parâmetro Th inicial garante uma segmentação confiável.

A calibração é realizada no canal de intensidade luminosa (V) sobre a imagem capturada pelo robô. Cada uma destas janelas tem uma barra, onde é possível variar os valores de 0 a 255. O parâmetro Th é o limiar que separa na imagem o plano de fundo dos alvos, ou seja, aquele que deve ser alterado durante o processo. Os outros dois parâmetros β_s e β_o são limiares para segmentar sombra e objeto respectivamente e não são alterados depois de calibrados (Algoritmo 1 - linha 2). Após efetuar a calibração dos parâmetros de visão, inicializa-se a captura

de imagens do ambiente para auto localização. Os frames capturados em sequência são indicados no algoritmo por t .

Algorithm 1 *THRESHOLD – AND – POSITION()*

```

1:  $\overline{bel}(S_1) = P(S_0) = [0.2, 0.2, 0.2, 0.2, 0.2]$ 
2: Calibration( $th, Scene, \nu$ )
3: while (1) do
4:    $e_t \leftarrow \max PERCEPTION - ACTION(th, image, \nu)$ 
5:    $bel(S_t) = P(S_t, Th|e_{0:t}) = P(e_t, Th|S_t) * \overline{bel}(S_t)$ 
6:    $(s_t, th) \leftarrow \arg \max_{s_t, th} P(S_t, Th|e_{0:t})$ 
7:    $P(S_t|e_t) \leftarrow \sum_{Th} P(S_t, Th|e_{0:t})$ 
8:    $\overline{bel}(S_{t+1}) = \sum_{s_t} P(S_{t+1}|s_t)P(s_t|e_t)$ 
9: end while

```

Algorithm 2 *PERCEPTION – ACTION*($th, Scene, \nu$)

```

1: Segment Scene using the threshold th to obtain a caster O and the top of its shadow top(S).
2: if caster and shadow were found then
3:    $P(e = DC) \leftarrow NonOccludesDC(O, top(S), \nu)$ 
4:    $P(e = EC) \leftarrow NonOccludesEC(O, top(S), \nu)$ 
5:    $P(e = PO) \leftarrow PartiallyOccludesPO(O, top(S), \nu)$ 
6:    $P(e = TPPI) \leftarrow TotallyOccludesTPPI(O, top(S), \nu)$ 
7: else
8:   if only caster was found then
9:      $P(e = NTPPI) \leftarrow TotallyOccludesNTPPI(O, top(S), \nu)$ 
10:  else
11:    return  $e_{t-1}$ ;
12:  end if
13: end if
14: return  $\arg \max_e P(E)$ ;

```

O Algoritmo 2 retorna a relação espacial que teve o maior número de características percebidas. As características extraídas da imagem $I(x, y)$ referentes ao reconhecimento do par objeto-sombra. As informações de características utilizadas para definir o tipo de relação que ocorre entre o par são:

Tabela 13: Tabela com os valores que diferenciam as características em cada relação.

Característica	DC	EC	PO	TPPI
1	maior que -1	entre(-25 , 23)	entre(-37,-4)	menor que -14
2	maior que -72	entre(-136,-48)	menor que -66	entre(-100,-78)
3	entre(113,201)	maior que 72	entre(33,157)	menor que 37
4	maior que 83	entre(68,163)	entre(76,205)	menor que 104
5	maior que 0,64	maior que 0,71	entre(0,33,0,79)	menor que 0,41
6	maior que 9600	entre(4525,17403)	entre(1903,8011)	menor que 2376
7	entre(11,82)	maior que 11	entre(4,54)	entre(2,38)
8	menor que 87	entre(4,88)	maior que 1	menor que 71
9	zero	menor que 16	menor que 25	menor que 40
10	maior que 11	entre(32,126)	entre(6,132)	entre menor 118

Fonte: autor.

1. A distância em y entre as bases dos retângulos limitadores da sombra e do objeto.
2. A distância em y entre o topo do retângulo limitador da sombra e a base do objeto.
3. A largura do retângulo limitador da sombra.
4. A altura aumentada em 40% do retângulo limitador da sombra.
5. A razão entre os dois valores acima (altura/ (1.4xlargura)).
6. A área do retângulo limitador da sombra.
7. Total de pontos conectados entre a sombra e a base do objeto.
8. Total de pontos conectados entre a sombra e a lateral do objeto.
9. Total de pontos conectados entre o topo da sombra e o topo do objeto.
10. Total de pontos conectados entre a sombra e o objeto.

Na Tabela 13 estão os valores que cada uma das características listadas acima pode assumir dependendo da relação espacial entre o objeto e sombra. A primeira coluna representa as características, enquanto que as colunas seguintes representam cada relação espacial observável entre sombra e objeto. Cada linha representa uma característica e os valores que ela pode assumir na ocorrência de cada relação espacial (DC , EC , PO , $TPPI$). Por exemplo, na primeira linha a característica 1 é referente a distância relativa em y entre as bases dos retângulos limitadores da sombra e do objeto. Em DC (maior que -1) de modo qualitativo significa que a base do retângulo limitador da sombra está na mesma altura ou abaixo da base do retângulo do objeto. De modo análogo para EC a base do retângulo limitador da sombra pode estar um pouco abaixo

ou um pouco acima, enquanto que estará sempre acima da base do retângulo do objeto para as relações espaciais PO e $TPPI$. Notem que alguns valores podem ser comuns para mais de uma relação espacial. As características de 1 a 6 referem-se à posição relativa entre os retângulos que envolvem a sombra e o objeto. As características de 7 a 10 são relativos ao número de pontos conectados entre o objeto e a sombra.

A evidência pode ser uma das relações DC , EC , PO , $TPPI$ ou $NTPPI$. Caso o objeto não tenha sombra na imagem, então somente a relação $NTPPI$ é analisada. Assim, no algoritmo 1 dada a evidência (linha 4), a linha 5 calcula a crença para todas as cinco regiões representadas por S_t , no intervalo de limiar $([Th - 5, Th + 5])$. Com a crença das regiões e os dois modelos, as próximas linhas do algoritmo 1 calculam a previsão e o limiar Th para serem utilizados em $t + 1$. A linha 6 dá o resultado sobre a região atual (s_t) e o limiar (Th) que será utilizado no próximo frame. O filtro Bayesiano manipula dois modelos de probabilidades, o modelo de imagem e o modelo de transição. As duas linhas finais no algoritmo 1 são referente a previsão, a crença atual para todos estados (dada a evidência) é calculada na linha 7. O modelo de transição que diz respeito à transição entre as regiões está representado na linha 8 ($P(S_{t+1}|s_t)$). Com os valores do modelo de transição (Tabela 12) e aqueles calculados na linha 7, na linha 8 é calculado a previsão da distribuição de crença de cada estado para o próximo frame.

4.5.2 Algoritmo de auto localização REQ probabilístico

O Algoritmo 3 inicializa o vetor de previsão com uma distribuição de crença uniforme para todas as regiões (linha 1) (seção 4.4.2). O Algoritmo 3 entra em repetição na linha 2 ficará executando a auto localização repetidamente, a partir de chamadas ao Algoritmo 4 na linha 3, que retorna um vetor de evidências.

O Algoritmo 4 retorna como evidências as regiões com máxima possibilidade a partir da inferência sobre a localização com os dados de imagem, e conhecimento do contexto formalizado utilizando o cálculo de oclusão combinado com um cálculo de direção. Na linha 6 do Algoritmo 4 a função Θ calcula todas as localizações possíveis em valores no intervalo $[0, 1]$. Esse valor foi definido pelo número de predicados das fórmulas 4.21 até 4.48 que são verificados. Por exemplo, uma fórmula que tem todas as suas relações observadas recebe valor máximo 1, enquanto que aquela que tem 3 de 4 predicados observados recebe o valor de 0,75. Este algoritmo diferente do utilizado no PQRS todos os máximos são considerados, deste modo o resultado pode indicar uma ou mais regiões como evidências. A quantidade de regiões que a inferência lógica retorna como resposta é inversamente proporcional à quantidade de dados na imagem observados (Algoritmo 4 - linha 6). Essas evidências serão utilizadas pelo filtro Bayesiano.

O filtro Bayesiano calcula a distribuição de crença da localização do robô ($bel(S_t)$) em todas as regiões s_t e todas as evidências e_t (linha-4). Para cada evidência será calculado uma distribuição de crença para todos os estados s_t com o produto valor do modelo de imagem e a crença anterior (linha 4). No Algoritmo 3 a localização do robô s_t é aquela que tiver maior valor

de crença (linha 5). A última etapa é calcular a previsão para $t + 1$ ($\overline{bel}(s_{t+1})$), marginalizando em s_t o produto da crença $P(s_t|e_t)$ com o modelo de transição $P(s_{t+1}|s_t)$.

Algorithm 3 *REGION – POSITION()*

```

1:  $\overline{bel}(S_1) = P(S_0) = [\frac{1}{28}, \dots, \frac{1}{28}]$ 
2: while (1) do
3:    $E_t \leftarrow \max \text{PERCEPTION} - \text{ACTION}(image, \nu)$ 
4:    $bel(S_t) = P(S_t|E_{0:t}) = \eta P(E_t|S_t) * \overline{bel}(s_t)$ 
5:    $(st, e_t) \leftarrow \arg \max P(S_t|E_t)$ 
6:    $\overline{bel}(S_{t+1}) = \sum_{s_t} P(s_{t+1}|s_t)P(s_t|e_t)$ 
7: end while
```

Algorithm 4 *PERCEPTION – ACTION(Scene, ν)*

```

1: Segmentação e detecção de caixas.
2: Reconhecimento das relações espaciais
3: for all Para todas regions  $R_i$  do
4:    $P(e = Ri) \leftarrow \Theta(\text{located}(Ri, \nu, b, g, r))$ 
5: end for
6: return  $\arg \max_e P(E)$ ;
```

Esse é o sistema de modo geral, que depende de um sistema de visão, capaz de reconhecer os elementos e as características que determinam as relações entre eles. A seguir vamos descrever esses algoritmos começando pelos algoritmos de reconhecimento dos elementos alvos.

4.5.3 Algoritmo detecção de sombra e objeto.

O Algoritmo 5 refere-se ao sistema de visão computacional responsável pela detecção e reconhecimento de sombras e objetos. A função de detecção foi desenvolvida para um ambiente predominantemente claro e objeto preto. As técnicas utilizadas são aquelas bem estabelecidas e disponíveis na biblioteca *OpenCV* (seção 2.7).

O primeiro filtro de limiar utilizado é o mais importante. Esse filtro elimina os elementos da cena que não sejam o objeto e a sombra. A imagem RGB é convertida para o espaço de cores HSV. A imagem HSV é dividida em seus canais resultando em três imagens em escala de cinza H, S e V . Um filtro de limiar passa baixa (seção 2.7.2) é aplicado na imagem V (Figura 51a). Todos os pixels com valor abaixo do valor de limiar th são preservados enquanto os outros recebem o valor zero (form.(4.57)). Como resultado objeto, sombra e alguns elementos do ambiente permanecem na imagem de objeto e sombra (Figura 51b). Os elementos do fundo que permanecem são derivados da penumbra que tem valor da intensidade luminosa na mesma faixa que a sombra.

$$\text{Imagem com objeto e sombra } (x, y) = \begin{cases} V(x, y), & \text{se } V(x, y) < \alpha ; \\ 0 & \text{em outro caso.} \end{cases} \quad (4.57)$$

Algorithm 5 Detecta sombra e objeto ($th, Scene$)

```

1:  $H, S$  e  $V \leftarrow$  Converte  $Scene$  em RGB para HSV e separa em canais ;
2: if ( $V(x, y) > th$ ) then
3:    $V(x, y) \leftarrow zero$ ;
4: end if
5:  $DEFINE - REGIAO(V, verde)$ ;
6: if ( $S(x, y) < \alpha_{sat}$ ) then
7:    $S(x, y) \leftarrow zero$ ;
8: end if
9: for  $i = 0$  to  $i = total - de - ROI$  do
10:  if (ROI equivalente em  $S$  tem mais de 1% de saturação) then
11:     $ROI[i] \leftarrow true$ ;
12:  else
13:     $ROI[i] \leftarrow false$ ;
14:  end if
15: end for
16: Criar máscara com  $ROI[True]$ 
17: Imagem de objeto e sombra  $\leftarrow$  máscara em  $V$ 
18: Imagem de objeto  $\leftarrow$  filtro de passa baixa na Imagem de objeto e sombra com limiar  $\beta_o$ .
19: Imagem de objeto  $\leftarrow$  filtro morfológico de fechamento na Imagem de objeto;
20:  $DEFINE - REGIAO(Imagem de objeto, azul)$ ;
21: for all  $ROI(x, y)$  do
22:  if Imagem de objeto( $x, y$ )  $> 0$  then
23:    ROI com objeto  $\rightarrow$  máscara;
24:  end if
25: end for
26: Imagem de sombra  $\leftarrow$  máscara em Imagem de objeto e sombra;
27: Imagem de sombra  $\leftarrow$  filtro de passa alta na Imagem de sombra com limiar  $\beta_s$ .
28: Imagem de sombra  $\leftarrow$  filtro morfológico de abertura na Imagem de sombra;
29:  $DEFINE - REGIAO(Imagem de sombra, vermelho)$ ;

```

Algorithm 6 $DEFINE - REGIAO(Imagem, Cor)$

```

1: Segmentação da Imagem por componentes conectados;
2: if ( $Cor = vermelho$ ) then
3:  Guarda informações() das regiões de sombra segmentadas;
4: else
5:  if ( $Cor = azul$ ) then
6:    Guarda informações() das regiões de objetos segmentadas;
7:  else
8:    if ( $Cor = verde$ ) then
9:      Guarda informações() das regiões de interesse  $ROI$  segmentadas;
10:      $total - de - ROI \leftarrow$  total de  $ROI$  encontrados.
11:    end if
12:  end if
13: end if
14: informações ( área, contorno da área, retângulo que contém área, quantidade encontrada);

```

Figura 51: Exemplo de sequência da segmentação de um quadro na detecção ROI: (a) canal V, (b) canal V após limiar acima α , (c) filtro morfológico, (d) retângulo verde indica regiões candidatas a região de interesse (ROI), (e) canal S após filtro passa-alta α_{sat} , (f) máscara ROI.

Fonte: autor.

O próximo passo é procurar na imagem resultante por regiões que são candidatas a serem as regiões de interesse (ROI) ¹. Para isso é aplicado um filtro morfológico de abertura. Depois são encontradas as regiões conectadas (seção 2.7.5). O filtro morfológico de abertura serve para eliminar pequenos grupos conectados que são considerados como ruídos na imagem. Por exemplo, a imagem da Figura 51b tem inicialmente 322 grupos de componentes conectados, este número é reduzido para dez depois de aplicado o filtro morfológico de abertura, a Figura 51c mostra o resultado desse procedimento. As áreas da imagem candidatas a ROI são escolhidas dentre aquelas que possam ter o objeto. No exemplo da Figura 51d, dentre os dez grupos conectados dois passaram pelo filtro e estão indicadas com retângulos na imagem. Uma máscara é criada com as regiões candidatas a ROI.

Neste trabalho, a região de interesse deve conter pelo menos parte do objeto alvo. Na imagem a região equivalente ao objeto alvo além da baixa intensidade luminosa, possui alto valor de saturação. Aplica-se um filtro de limiar passa alta na imagem S equivalente ao canal de saturação, onde somente os valores altos permanecem (Figura 51e). As regiões ROI na imagem do objeto e sombra que têm mais de 1% de saturação nas regiões equivalentes da imagem S são consideradas verdadeiras. Esse filtro serve para eliminar boa parte daquelas regiões candidatas onde somente a penumbra está presente. No caso da Figura 51e a região de penumbra permaneceu, pois existe uma sobreposição dos retângulos que indicam a região de ROI, o que fez com que parte do objeto estivesse dentro de ambas as regiões na imagem de saturação. Com as regiões ROI verdadeiras uma nova máscara é criada para ser utilizada na detecção de objetos (Figura 51f).

Para detectar o objeto, primeiro aplica-se a nova máscara na imagem com objeto e sombra, e como resultado permanece as áreas que podem ter o objeto alvo (Figura 52a). Como o objeto é preto, sua intensidade luminosa é mais baixa que a intensidade luminosa da sombra, então é aplicado o filtro de limiar passa baixa, com o limiar β_o definido na calibração inicial. Como mostra a fórmula (4.58), os valores abaixo do limiar β permanecem e os outros ficam com valor zero. Na imagem resultante é aplicado um filtro morfológico de fechamento (Figura 52b) para logo em seguida encontrar os componentes conectados, os quais definem a região do objeto.

$$Imagem\ com\ objeto\ e\ sombra\ (x, y) = \begin{cases} V(x, y), & se\ V(x, y) < \beta_o ; \\ 0 & em\ outro\ caso. \end{cases} \quad (4.58)$$

As informações da região de cada objeto são guardadas para serem utilizadas posteriormente. As informações são características da região do objeto, tais como: tamanho de área, dados do menor retângulo que contém a área do objeto, dados do contorno, se objeto está encostado nas bordas da imagem, dados da base do objeto (Figura 52c) e o total de objetos encontrados. A Figura 52d apresenta o exemplo de uma imagem com o retângulo desenhado indicando o objeto alvo encontrado. Novamente as regiões ROI são analisadas e aquelas que não têm todo o objeto

¹ do inglês: Region of Interest (ROI)

Figura 52: Exemplo de sequência da segmentação de um quadro na detecção de objetos: (a) Imagem objeto e sombra depois de aplicado a máscara, (b) filtro de limiar abaixo e filtro morfológico, (c) imagem binária da base do objeto, (d) retângulo azul indica objetos encontrados na segmentação, (e) última máscara ROI criada pelo sistema de visão.

Figura 53: Exemplo de sequência da segmentação de um quadro na detecção de sombras: (a) Imagem com objeto depois de aplicada a segunda máscara, (b) Imagem após filtros de limiar acima e morfológico de abertura, (c) retângulo limitador vermelho indicando a sombra localizada.

Fonte: autor.

dentro são apagadas da máscara. Somente os retângulos com objetos são mantidos criando uma nova máscara (Figura 52e).

Aplicando a nova máscara na imagem com sombra e objeto obtém-se uma imagem que mantém as regiões que possuem objeto alvo (Figura 53a). Este procedimento equivale a uma subtração de fundo deixando as regiões com os elementos alvos. Quando comparamos a Figura 52a (que mostra a imagem obtida com a primeira máscara) com a Figura 53a, nota-se que toda a região de penumbra não existe mais, permanecendo somente as regiões compreendidas pelo retângulo limitador da região ROI com os objetos alvos.

Para segmentar a sombra, é aplicado o filtro de limiar acima com o limiar β_s determinado na calibração inicial. Os valores de intensidade maiores que β_s permanecem e todos os outros vão para zero conforme fórmula (4.59).

$$Imagem\ sombra(x, y) = \begin{cases} V(x, y), & \text{se } V(x, y) > \beta_s; \\ 0 & \text{em outro caso.} \end{cases} \quad (4.59)$$

Figura 54: Exemplo de segmentação do objeto e sua sombra. O topo da sombra também é detectado pelo sistema de visão conforme destacado nesta imagem.

Fonte: Fenelon et al. (2013).

Como resultado da fórmula (4.59) obtém-se uma imagem onde somente a região equivalente a sombra permanece. Nesta imagem com sombras é aplicado um filtro morfológico de abertura para eliminar ruídos (Figura 53b), e achando os componentes conectados para que seja determinada a região equivalente a sombra e suas informações. Caso uma região de sombra se encontre totalmente dentro do retângulo que limita o objeto, esta não será considerada como sombra, pois em geral é um reflexo da luz no objeto. Para os demais casos, guardam-se as características das regiões de sombra, tal como área, os dados do menor retângulo que contém a área da sombra, se a sombra está encostada nas bordas da imagem e a quantidade de sombras existentes na imagem, a região equivalente ao topo da sombra. Uma sombra tem valores de intensidade diferentes, e a intensidade dos pixels do topo da sombra são mais altos. Assim, com um filtro de limiar é possível localizar o topo da sombra.

Como exemplo a Figura 53c mostra a sombra encontrada limitada pelo retângulo vermelho, enquanto que a Figura 54 mostra a localização do objeto e da sombra limitados pelos retângulos azul e vermelho respectivamente. A região equivalente ao topo da sombra está destacada em branco (Figura 54). Esse foi o algoritmo responsável pela detecção de sombra e objeto, a seguir o algoritmo para detecção de objetos.

4.5.4 Algoritmo de detecção de objetos

O sistema de visão para detectar as caixas coloridas foi desenvolvido de forma análoga ao utilizado no sistema de sombras. Na linha 1 do Algoritmo 7 a imagem em RGB (Figura 55a) é convertida para o espaço de cores HSV e dividida em três imagens com as informações de cada canal H , S e V . Como as caixas têm alto grau de saturação (Figura 55b), um filtro de limiar passa alta foi utilizado no canal de saturação S (linha 2, Figura 55c), as regiões com alta saturação são definidas como regiões de interesse ROI (linha 6). Uma máscara é criada com ROI

Algorithm 7 Detecta objeto (*Scene*)

```

1:  $H, S$  e  $V \leftarrow$  Converte  $Scene$  em RGB para HSV e separa em canais ;
2: if ( $S(x, y) < \alpha_{Sat}$ ) then
3:    $S(x, y) \leftarrow zero$ ;
4: end if
5:  $DEFINE - REGIAO(S, verde)$ ;
6: if (ROI equivalente em  $S > 1\%$  de saturação) then
7:   ROI  $\leftarrow$  true;
8: end if
9: máscara  $\leftarrow$  todos os ROI[True]
10: Imagem de objetos  $\leftarrow$  máscara em  $V$ 
11: Imagem de objeto vermelho  $\leftarrow$  filtro de passa faixa vermelho em Imagem de objetos.
12: Imagem de objeto verde  $\leftarrow$  filtro de passa faixa verde em Imagem de objetos.
13: Imagem de objeto azul  $\leftarrow$  filtro de passa faixa azul em Imagem de objetos.
14:  $DEFINE - REGIAO$ (Imagem de objeto vermelho, azul);
15:  $DEFINE - REGIAO$ (Imagem de objeto verde, azul);
16:  $DEFINE - REGIAO$ (Imagem de objeto azul, azul);

```

definindo regiões onde as caixas possam estar presentes (Figura 55e). Em todos os casos para definir as regiões (ROI, objeto, sombra) foi utilizado o Algoritmo 6: logo após aplicar um filtro de limiar (Figura 55c); utilizamos um filtro morfológico de abertura que elimina parte do ruído na imagem; para em seguida localizar as regiões conectadas (Figura 55d). A localização de cada região conectada na imagem é determinada pelo menor retângulo que a contém (Figura 55f).

Os objetos têm cores diferentes, vermelho, azul e verde, assim três imagens são criadas aplicando a máscara no canal de matiz H . Aplicando um filtro de limiar passa faixa no canal de matiz (H) somente as cores dentro da faixa permanecem na imagem. Aplicando o filtro morfológico e componente conectado, são encontradas as regiões equivalentes a cada objeto e suas informações são armazenadas. Com as informações de posição, de cor do objeto, verifica-se a profundidade média da área central da região do objeto na imagem de profundidade. Com as informações tanto de sombra e objeto quanto de objetos a próxima etapa é determinar os pares e a relação entre eles. Vamos começar pelo algoritmo que analisa a correspondência entre a sombra e o objeto.

4.5.5 Algoritmo de análise de correspondência

Depois de encontrar o objeto e a sombra na imagem, é importante determinar os pares corretamente. Por isso procuramos utilizar uma solução que apresentasse uma base apoiada em um formalismo teórico. Na seção 2.2.3.2 foi sugerida uma solução para o problema de correspondência baseada no PQRS, a qual diz que: é sombra do objeto aquela que está ligada à base deste objeto, com a fonte de luz mais próxima do objeto do que da sombra, e tanto objeto quanto pelo menos parte da sombra estejam sobre a mesma superfície. O algoritmo 8 representa o pseudocódigo utilizado para correspondência.

Figura 55: Exemplo de sequência da segmentação de um quadro na detecção ROI: (a) canal V, (b) canal V após limiar acima α , (c) filtro morfológico, (d) retângulo verde indica regiões candidatas a região de interesse (ROI), (e) canal S após filtro passa-alta α_{sat} , (f) máscara ROI.

Fonte: autor.

O Algoritmo 8 serve para mais de um objeto e sombra em cena, e funciona do seguinte modo: com os dados armazenados pelo algoritmo de reconhecimento de sombra e de objeto (Algoritmo 5), o sistema verifica se há conexão externa entre cada sombra e a base de todos os objetos (Algoritmo 9). Em caso de haver uma conexão, analisamos se fonte de luz está mais próxima do objeto do que da sombra. As informações de conexões serão armazenadas para serem utilizadas como uma das características para determinar o tipo de relação entre a sombra e o objeto.

Na Figura 56 encontramos um exemplo de como é verificada a conexão externa das sombras com a base dos objetos utilizando detectores de borda (seção 2.7.4). Na Figura 56a existem dois objetos e consequentemente duas sombras na imagem, note que o objeto em forma de cone tem duas sombras conectadas externamente a ele. As Figuras 56c e 56e mostram que as conexões entre as duas sombras e o cone foram detectadas, contudo somente uma das sombras está mais distante da fonte de luz do que o cone. Em alguns casos, a sombra na imagem não

Algorithm 8 Análise de Correspondência (*Dados*)

```

1: Declaração de Variáveis: Booleano Conectado
2: for all objeto do
3:   while sombras[j] do
4:     Conectado  $\leftarrow$  CONNECTIVIDADE(objeto[i], sombra[j]);
5:     if Conectado = True then
6:       sombra conectada ao objeto[i]  $\leftarrow$  sombra[j];
7:       objeto conectado à sombra[j]  $\leftarrow$  objeto[i];
8:       incrementa informação de total de conectado[i];
9:       incrementa j;
10:    end if
11:  end while
12:  incrementa i;
13: end for
14: while objeto[i] do
15:   if total de conectado=1 then
16:     if Conectado com base = True then
17:       while sombras[j] do
18:         sombra correspondente do objeto[i]  $\leftarrow$  sombra[j];
19:         correspondência  $\leftarrow$  true;
20:       end while
21:     end if
22:   else
23:     if total de conectado > 1 then
24:       lista de proximidade, dúvida  $\leftarrow$  objeto[i];
25:       while lista dúvida do
26:         if objeto mais próximo da luz do que sombra then
27:           sombra correspondente do objeto[i]  $\leftarrow$  sombra[j];
28:         end if
29:       end while
30:     end if
31:   end if
32: end while

```

se conecta a nenhuma base dos objetos em cena. Esses casos são casos especiais, tais como sombras provocadas por um objeto que esteja fora da cena ou oculto por outro objeto. Os casos de oclusão e de objeto com duas sombras não ocorrem nos experimentos relatados neste trabalho, pois utilizamos somente um único objeto alvo em cena.

Toda a sombra e objeto que são correspondentes entre eles formam o par objeto-sombra, e algumas informações sobre o par também são armazenadas, tal como a posição relativa entre os retângulos que definem a sombra e o objeto. As informações coletadas serão utilizadas para definir as relações entre o objeto e a sombra. Essas relações são as evidências utilizadas para auto localização do robô, como descrito na seção 4.5.1. A seguir o algoritmo que infere a relação entre pares de objetos.

Figura 56: Imagem exemplo de busca de regiões externamente conectadas, para cada sombra todas as bases de objetos são verificadas onde as bordas se encontram intensidade aumenta e EC é verdade.

Fonte: autor.

Algorithm 9 *CONNECTIVIDADE*(*Regiao1*, *Regiao2*)

```

1: Imagem 1  $\leftarrow$  contorno Região1  $\leftarrow$  100;
2: Imagem 2  $\leftarrow$  contorno Região2  $\leftarrow$  100;
3: Imagem Conectado  $\leftarrow$  Soma(Imagem 1, Imagem 2);
4: Imagem Conectado contorno - 100;
5: if Conta (pixel  $\neq$  zero) > zero then
6:   return True;
7: else
8:   return False;
9: end if

```

4.5.6 Algoritmo de reconhecimento de relações espaciais

A partir das informações dos objetos podemos acrescentar outras e descobrir o tipo de relação entre pares de objetos. Para cada par é verificado quais das relações de direção (posição e profundidade) e relações topológicas (oclusão) são percebidas. Para determinar se o par tem uma relação *Right* ou *Left* compara-se a posição dos objetos relativa ao eixo x da imagem (Figura 36). Para determinar se o par tem uma relação *Front* ou *Behind* compara-se a distância de ambos os objetos em relação ao robô. Utilizamos principalmente as informações da imagem de profundidade do sensor *Kinect*, quando não é possível obter a leitura deste sensor utilizamos as localizações dos objetos em cena relativas ao eixo y (Figura 36).

Para determinar a conexão entre os objetos, assim como no problema objeto e sombra (seção 4.5.5), as bordas são somadas em uma imagem e depois subtraídas permanecendo somente aqueles *pixels* onde as bordas se sobrepuseram (Figura 57). Todos os pixels que sobram são contados, caracterizando a conexão entre objetos, a Figura 57h mostra a conexão que ocorre na Figura 57f. Quando os objetos estão próximos, mas ainda não estão conectados, alguns pontos de conexão começam a aparecer. Essa é uma característica que pode indicar movimento, outra característica é a distância entre o par, ou seja, quando dois objetos se aproximam ou se afastam entre si na sequência de imagens.

As informações de conexão, distância e profundidade são parâmetros utilizados para determinar se os objetos estão parcialmente sobrepostos (*PartiallyOccludesPO* ou *PartiallyOccludesPO*⁻¹). Utilizamos uma variável para indicar a condição de oclusão (*ocluindo*) (algoritmo 7), nas cenas em que somente um objeto aparece, caso o par desse objeto esteja na condição de oclusão, infere-se que o objeto invisível está totalmente ocluído por aquele objeto em cena (*TotallyOccludesNTPP* ou *TotallyOccludesNTPPI*). Finalmente, quando entre os objetos não existe conexão e o par está livre de oclusão a relação percebida é a *NonOccludesDC*. Neste experimento as relações percebidas são informações para definir a localização. Quanto mais informações menor será o número de evidência retornada para o algoritmo 3 que será utilizada na etapa de atualização de informação (linha 3 algoritmo 3) do algoritmo de auto localização.

Algorithm 10 Relações espaciais (*Scene*)

```
1: for all todos os pares  $(A, B)$  do
2:   if  $A.x < B.x$  then
3:      $Left(A, B)$ 
4:   else
5:      $Right(A, B)$ 
6:   end if
7:   if  $A$  está mais próximo do robô que  $B$  then
8:      $Front(A, B)$ 
9:   else
10:    if  $B$  está mais próximo do robô que  $A$  then
11:       $Behind(A, B)$ 
12:    end if
13:  end if
14:  if  $CONNECTIVIDADE(A, B) = \text{True}$  then
15:     $ocluindo \leftarrow \text{True}$ 
16:    if  $Front(A, B)$  then
17:       $PartiallyOccludesPO(A, B)$ 
18:    else
19:      if  $Behind(A, B)$  then
20:         $PartiallyOccludesPO^{-1}(A, B)$ 
21:      end if
22:    end if
23:  else
24:    if  $ocluindo = \text{True}$  then
25:      if  $Front(A, B)$  then
26:         $TotallyOccludesNTPP(A, B)$ 
27:      else
28:        if  $Behind(A, B)$  then
29:           $TotallyOccludesNTPPI(A, B)$ 
30:        end if
31:      end if
32:       $NonOccludesDC(A, B)$ 
33:       $ocluindo \leftarrow \text{False}$ 
34:    end if
35:  end if
36: end for
```

Neste capítulo mostramos um método de raciocínio espacial qualitativo probabilístico para a interpretação espacial do ambiente tridimensional a partir de imagens bidimensionais para aplicar em um robô móvel em ambiente real na execução da tarefa de auto localização. No próximo capítulo faremos experimentos com esse método utilizando dados reais de um robô móvel.

Figura 57: Imagem exemplo para definição de relações topológicas de oclusão.

Fonte: adaptado de Santos et al. (2010) e Fenelon et al. (2010).

5 EXPERIMENTOS EM RACIOCÍNIO ESPACIAL QUALITATIVO PROBABILÍSTICO PARA AUTO LOCALIZAÇÃO DO ROBÔ MÓVEL

Neste capítulo descreveremos os procedimentos adotados para execução dos experimentos relativos à localização do robô. Os resultados obtidos com PQRS probabilístico e com REQ probabilístico estão relatados nas seções 5.2 e 5.4. Como os resultados foram avaliados está descrito na seção 5.3.

5.1 Descrição dos Ambientes de Testes

Em ambos os casos o ambiente foi preparado para que o robô tivesse uma área de atuação livre para navegação. No primeiro ambiente o objeto foi disposto de modo a garantir que o mesmo estivesse a vista do observador sem obstrução por parte de qualquer outro elemento. Enquanto que no segundo ambiente as caixas foram dispostas de forma que o robô pudesse navegar entre elas (Figura 60). Em ambos os casos o robô *PeopleBotTM* capturou imagens enquanto navegava pelo ambiente. Os sistemas de auto localização foram desenvolvidos em computadores com ambiente de desenvolvimento Visual Studio 2010 instalado. Para o sistema de visão computacional foi utilizada a biblioteca *Open Source Computer Vision (OpenCV)*, baseada na plataforma C++.

5.1.1 Descrição do ambiente de teste com objeto-sombra

As imagens foram capturadas em um laboratório (Figura 58). Neste ambiente interno existem alguns móveis de escritório, piso opaco e claro, uma fonte de luz predominante, e um objeto cilíndrico preto. O objeto que causa a sombra tem forma geométrica bem definida e contínua (sem buracos), além da cor escura que serve para distinguir o plano de fundo do objeto e sombra. Para ter uma sombra bem definida, utilizamos uma luminária em tripé com lâmpada halógena de 1000 Watts. Com essa configuração, a fonte de luz pode ser colocada em posições distintas. Entretanto, para garantir um ambiente cognitivamente relevante (MAMASSIAN; GOUTCHER, 2001) mantivemos a fonte de luz sempre acima do objeto e do observador. A distância entre a fonte de luz e o objeto define o raio do círculo e a forma da área de atuação.

Figura 58: O ambiente utilizado para a coleta de imagens com sombra.

Fonte: Santos, Dee e Fenelon (2008).

No papel de observador utilizamos um robô móvel *ActivMedia PeopleBotTM* com processador *Pentium III*, 512 MB de memória, velocidade de 533 Hz, com uma câmera acoplada, com a qual as imagens foram capturadas. O robô foi programado e teleoperado utilizando a biblioteca de controle *ActivMedia Robotics Interface for Applications* (ARIA). Barbantes esticados no chão serviram para a demarcação das regiões fisicamente.

5.1.2 Descrição do ambiente de teste com múltiplos objetos

O ambiente interno, utilizado nos experimentos reportados a seguir, é um espaço entre as salas existentes no quinto andar do prédio K do Centro Universitário da FEI (Figura 60). São utilizadas três caixas coloridas, cada caixa tem cor única (vermelho ou verde ou azul). Para demarcar as regiões (para fins de comparação com a resposta automática do sistema) utilizamos barbantes, os quais representam as linhas de visão fisicamente.

Para capturar imagens foi utilizado um *Kinect* disposto sobre um robô móvel *ActivMedia PeopleBotTM*. Ambos foram controlados por um algoritmo de controle e captura instalado em um notebook com o *Robot Operating System* (ROS). Todos os testes foram realizados *off-line* utilizando as imagens coletadas conforme procedimento descrito a seguir.

5.1.3 Procedimentos de coleta de imagens com sombra-objeto

Em ambiente de teste preparado conforme descrito na seção 5.1.1, o robô foi colocado em uma região dentro da área de atuação na qual o objeto e sombra estavam dentro do seu campo

visual. Isso porque é importante que a primeira imagem tenha objeto e sombra por completo, pois essa imagem será utilizada em uma calibração manual do limiar para segmentação. Depois que se inicia o algoritmo de captura e controle do robô, o operador tem autonomia para parar de salvar os frames capturados, assim como marcar em qual a região o robô está localizado durante a captura da imagem. Deste modo, para evitar erro na anotação, as imagens deixavam de serem salvas quando o robô passava de uma região para a outra, o que garantiu uma anotação correta da região em que a imagem foi capturada. Para cada imagem salva os dados relacionados a essa imagens tais como: o nome da imagem, a região real, a velocidade e posição relativa ao início de cada coleta são armazenados em arquivo.

Esse conjunto de imagens foi utilizado para o experimento de auto localização e auto calibração e por esse motivo foi mantida a sequência de captura. Por limitações do robô, a captura foi realizada em dias e horários diferentes, recomeçando a captura sempre do ponto que finalizou a sequência anterior. Deste modo a sequência de transição espacial foi mantida. O robô capturou 587 imagens passando por todas as regiões pelo menos uma vez. Contudo, em aproximadamente 33% das imagens não foi possível manter por completo o objeto e a sombra dentro do campo de visão do robô (Figura 59).

Todas as 587 imagens RGB com resolução de 640 x 480 e os dados em arquivo, foram utilizados para fazer os experimentos *off-line*. Além desse conjunto em sequência, havia em arquivos outras imagens de capturas anteriores, as quais estavam separadas por região onde foram capturadas, algumas dessas imagens foram utilizadas para criar o modelo de imagem. Na próxima seção vamos descrever o algoritmo de auto localização e auto calibração.

5.1.4 Procedimento de coleta de imagens com múltiplos objetos

Depois de preparar o ambiente (seção 5.1.2), o robô *PeopleBotTM* foi remotamente controlado para navegar por todas as 28 regiões. A posição inicial é aquela escolhida pelo operador que controla o robô navegando pelo ambiente entre as caixas salvando imagens RGB e imagens binárias de profundidade e os dados de hodometria. Para este experimento somente as imagens RGB e de profundidade capturadas com o sensor *Kinect* foram utilizadas. Esse procedimento é semelhante ao procedimento realizado no ambiente com sombra 5.1.1, porém não foi possível anotar as regiões onde o robô estava passando durante o experimento. Também não é necessária a calibração dos parâmetros de visão. Em duas coletas foram capturadas aproximadamente 18000 frames formando dois conjuntos de sequências de imagens com resolução de 640 x 480.

Em cada coleta foram executadas aproximadamente três voltas pelo ambiente alterando o circuito, porém procurando manter os objetos no campo visual do robô. O primeiro conjunto com 9033 imagens foi utilizado para criar o modelo de imagem (seção 5.1.6), enquanto que o segundo com 9483 imagens foi utilizado no experimento *off-line* de auto localização. Considerando o número de objetos que aparecem em cena, os dois conjuntos de imagens juntos têm 9% de

Figura 59: Imagens em que o objeto e a sombra não estão completamente dentro do campo de visão do robô. O sistema somente com raciocínio espacial não poderia inferir uma localização a partir dessas imagens.

Fonte: autor.

Figura 60: Ambiente utilizado para captura e auto-localização com múltiplos objetos.

Fonte: autor.

Figura 61: Imagens capturadas com três pares, um par e nenhum par em cena

Fonte: autor. .

imagens com três objetos, 42% com dois objetos, e com um ou nenhum objetos 45% e 5% respectivamente.

Neste procedimento não foram salvas as informações de localização real enquanto o robô capturava a imagem. Por isso a tabela (*gold standard*) que será utilizada para comparar a resposta do sistema foi elaborada manualmente. Cada imagem foi analisada pelo especialista que anotou os dados da região onde o robô estava localizado, quais objetos visíveis em cena, além das relações espaciais entre os pares. Para isso contou com a ajuda das informações do ambiente em cena e do barbante, o qual foi o principal parâmetro utilizado para reconhecer em qual região o robô estava localizado no momento da captura do frame. Distinguir uma região de outra foi possível utilizando a oclusão do barbante pela caixa. O conjunto que foi utilizado no experimento de auto localização é composto por 12% de imagens com três pares, 46% de imagens com um par, e 42% de imagens sem pares, destas imagens ou há somente uma caixa (39%) ou nenhuma (3%) presente em cena. A Figura 61 tem um exemplo para cada caso. Com as imagens coletadas e antes de colocar o sistema para inferir localização, fizemos o modelo de imagens para cada um dos domínios.

5.1.5 Modelo de imagens para sombra-objeto

Com imagens capturadas pelo *PeopleBotTM* de todas as regiões:

- Escolhemos somente imagens nas quais tanto o objeto quanto a sombra estavam completamente em cena. Este conjunto garante que existe sempre objeto-sombra para ser avaliado pelo sistema de visão.
- Todas as imagens separadas por região foram analisadas pelo algoritmo 2, para cada valor de limiar. Este algoritmo é o responsável por determinar qual o tipo de relação espacial qualitativa é percebida entre o objeto e a sombra.

O conjunto de imagens utilizadas para o modelo não está em sequência e pode apresentar diferenças de luminosidade por ter sido capturado em dias, horários diferentes, inclusive a posição da luz e do objeto dentro da sala não são sempre os mesmos. Esse conjunto foi montado

com imagens de um banco de imagens capturadas anteriormente, ou seja, não são as mesmas imagens utilizadas para o experimento de auto localização e calibração.

5.1.6 Modelo de imagem com múltiplos objetos

Uma forma de construir o modelo de imagem é separar um banco de imagens de cada região. Essas imagens devem conter pelo menos um par de objetos em cena, inclusive quando existe a oclusão. Pelas fórmulas (4.21) até (4.48) a localização em cada região está relacionada a uma combinação de relações entre pares percebíveis por um observador que esteja nesta região. Com o banco de imagens verificamos para cada imagem a frequência com que as combinações de relações aparecem para cada região. No banco de imagens utilizado as regiões, R_9 , R_{15} , R_{18} , R_{27} , R_{28} todas as imagens têm somente um ou nenhum objeto em cena, consequentemente nenhum par para extrair a relação pela imagem. Então uma opção foi definir os valores de probabilidade por aproximação.

Usamos os dados de uma sequência de imagens e a região na qual essas imagens foram capturadas (seção 5.1.4). Contamos o número de respostas de regiões do algoritmo de localização (Algoritmo 4) para cada região do mapa (Figura 46). Assim as evidências e_t são as regiões na qual as relações do sistema de visão podem ser percebidas. Um robô está em qualquer uma das regiões quando captura a imagem e isto é uma verdade. Deste modo, por exemplo, se em uma dada imagem quatro evidências apresentam a mesma possibilidade de ser a região em que o robô está localizado pelo algoritmo 4, por uma aproximação então cada uma recebe o valor de 0,25, enquanto que as outras evidências recebem zero. Todos os valores são somados por evidência e depois normalizados por região. Este é o modelo de imagem utilizado neste experimento.

5.1.7 Teste com o sistema de auto localização

O sistema composto pelos algoritmos relatados no capítulo 4.5 foi codificado em C++ e para as técnicas de visão computacional utilizamos a biblioteca *OpenCV*¹. Com o banco de imagens, os dados sobre essas imagens, o modelo de imagem e o modelo de transição rodamos os experimentos de auto localização de modo *off-line*. Em nossos experimentos um robô móvel *Pioneer PeopleBotTM* coletou fotos em um ambiente interno (Figura 58). O robô foi tele operado e poderia começar em qualquer uma das regiões indicadas na Figura 47. Com os algoritmos descritos na seção 5.1.7 o sistema utiliza a crença sobre a localização do robô e o que é percebido a partir da imagem do cenário, fornecendo o estado do robô e alterando o parâmetro de limiar do sistema de visão. O sistema de visão converte as imagens digitais capturadas pela câmera do robô móvel para o espaço de cores HSV, e a detecção do objeto e da sombra utiliza principalmente um filtro de limiar no canal de intensidade luminosa V .

¹ Open Source Computer Vision - <http://opencv.org/> - último acesso 06/02/2014

Figura 62: Representação de um robô na região 3.

Fonte: autor.

5.2 Método de Avaliação

Para avaliar o sistema foi analisada a sua acurácia. Para definir uma auto localização como correta comparou-se a resposta do sistema com uma base de conhecimento verdadeiro (*gold standard*). A base de conhecimento quando utilizando as informações de oclusão de sombras foram marcadas durante a captura das imagens, enquanto para o ambiente com múltiplas caixas essa tabela foi criada manualmente por um especialista. A acurácia é a proporção de resultados corretos, dentre todos os resultados e é dada pela fórmula (5.1):

$$acurácia = \frac{resultados\ corretos}{todos\ resultados} \quad (5.1)$$

Esse tipo de análise foi utilizada em ambos os experimentos de auto localização. Os resultados foram avaliados tanto para verificar a acurácia por região quanto acurácia do sistema todo. Por exemplo, todas as imagens que o robô capturou enquanto estava na região 3 (Figura 62), são consideradas como resultado correto quando o sistema retorna a região 3, e incorreto quando retorna qualquer outra região. Em ambos os casos comparamos os resultados do sistema com extensão probabilística com o sistema utilizando somente o raciocínio espacial qualitativo. Os resultados para ambos os testes estão descritos nas seções 5.3 e 5.4.

5.3 Resultados para Auto Calibração e Auto Localização com Sombras

O sistema PQRS probabilístico foi aplicado em uma sequência de 587 imagens. As imagens foram capturadas em um ambiente interno contendo um objeto com sombra (seção 5.1.1). O resultado do procedimento de localização (no mapa mostrado na Figura 47) está apresentado na matriz de confusão na Tabela 14. Cada coluna representa a resposta do sistema, enquanto que

Tabela 14: O resultado do experimento de auto localização utilizando um filtro Bayesiano e o Raciocínio Qualitativo Perceptual sobre Sombras.

Regiões	$Region_1$	$Region_2$	$Region_3$	$Region_4$
$Region_1$	97%	3%	0%	0%
$Region_2$	25%	74%	1%	0%
$Region_3$	9%	21%	66%	4%
$Region_4$	6%	0%	8%	86%
$Region_5$	0%	0%	35%	65%

Fonte: Fenelon et al. (2013).

as linhas representam a região na qual o robô estava (de acordo com a tabela de referência *gold standard*).

A Tabela 14 tem os valores de resposta do sistema PQRS probabilístico em porcentagem para cada região onde as imagens foram capturadas. Na região $Region_1$ o sistema apresentou uma acurácia de 97% e os 3% restantes o sistema deu como resposta a região vizinha $Region_2$. Em $Region_2$ o sistema obteve uma acurácia de 74%, os 26% restantes o sistema deu como resposta uma das regiões vizinhas $Region_1$ e $Region_3$ em um total de 25% e 1% respectivamente. Já para a região $Region_3$ a acurácia foi de 66%, sendo que achou estar localizado nas vizinhas $Region_2$ e $Region_4$ em 21% e 4% das imagens respectivamente, em 9% das imagens deu como resposta a $Region_1$. Para a $Region_4$ a acurácia do sistema foi de 86%, com sistema se localizando na vizinha $Region_3$ em 8% e na região $Region_1$ em 6%. A acurácia do sistema na região $Region_5$ não existe já que esta região não foi avaliada por este sistema (que utiliza a relação objeto-sombra visíveis na imagem), como localização apresentou 65% na vizinha $Region_4$ e 35% na região $Region_3$.

A Tabela 15 compara os resultados obtidos utilizando o PQRS com uma base de conhecimento com aqueles obtidos com o aqui proposto PQRS probabilístico. O sistema PQRS com uma base de conhecimento tem a informação da região na qual o robô está localizado e procura um limiar que confirme essa informação para esta localização, ou para as vizinhas (FENELON et al., 2012).

Na Tabela 15, a linha refere-se à localização real de cada imagem (regiões $Region_1$ até $Region_5$). A coluna #imagens representa a quantidade de imagens capturadas por região. Enquanto que a coluna "#inf completa" apresenta a porcentagem de imagens na qual a sombra e o objeto estão completamente dentro da imagem. Para compor os valores dessa coluna ficaram de fora: imagens sem sombra ou objetos; imagens só com sombras, imagens nas quais a sombra ou o objeto estão encostados em uma das laterais da imagem. Os resultados de acurácia obtidos com PQRS estão representados na coluna "baseado no conhecimento". Os resultados da coluna "baseado na crença" são referentes aos resultados de acurácia obtidos com PQRS probabilístico.

Tabela 15: Valores de acurácia do algoritmo baseado em conhecimento e o com o filtro Bayesiana qualitativo, onde '# imagens' é o número de imagens capturadas em cada região.

Regiões	# imagens	#inf completa (%)	baseado no conhecimento (%)	baseado na crença (%)
<i>Region₁</i>	225	54	66	97
<i>Region₂</i>	171	65	34	74
<i>Region₃</i>	138	77	80	66
<i>Region₄</i>	36	92	44	86
<i>Region₅</i>	17	94	59	0
Global	587	67	58	80

Fonte: Fenelon et al. (2013).

Como exemplo, a última linha mostra que para 587 onde 67% tinham informação completa na imagem o desempenho foi de 58% com PQRS para 80% de acurácia com o PQRS probabilístico.

O sistema PQRS consegue lidar com algum grau de incerteza, na Tabela 15 podemos ver o reflexo disso nas linhas referentes às regiões *Region₁* e *Region₃* onde o sistema obteve uma porcentagem de acerto maior que o número de imagens com informação completa. Ainda assim, com exceção da região *Region₃* o sistema PQRS probabilístico apresentou desempenho superior e em presença de incerteza em todas as regiões. A Tabela 14 mostra que o maior número de falsos negativos para a *Region₃* foi relatado com o robô localizado na região 2 (21% dos casos), parte deste erro é devido ao problema de fronteira.

Na fronteira entre duas regiões o número de características que indica a relação entre objeto-sombra é igual para ambos os locais. O problema de fronteira ocorre porque ao escolher qual a relação espacial (evidência) com máxima probabilidade somente aquela na qual o valor foi analisado primeiro passa como evidência para o filtro Bayesiano. Esta escolha teve reflexos positivos nas regiões *Region₁*, *Region₂* e *Region₄*, entretanto na Região *Region₃* a acurácia decaiu em comparação com experimentos anteriores.

A região *Region₅* é um caso a parte já que para estes experimentos optamos por codificar o sistema sem mecanismo para inferir esta região. Isso devido ao modelo de movimento, no qual a relação de oclusão total (*Totally Occludes*) não foi considerada na composição da função w , já que esta função é calculada com a existência de sombras. Como a área dessa região é pequena, o efeito negativo sobre o desempenho global foi de aproximadamente 2,8%.

O desempenho global deste experimento apresentou uma elevação da acurácia passando de 58% no PQRS para 80% no PQRS probabilístico. Os resultados com o sistema PQRS probabilísticos foram animadores, contudo os testes de auto localização foram realizados somente com um objeto (mapa Figura 47). Com mais de um objeto o mapa tem mais regiões formadas

pela sobreposição de regiões. Os testes com múltiplas caixas serviram também para mostrar a adaptabilidade do método considerando relações entre os elementos diferentes de sombra. Na próxima seção (seção 5.4) os resultados relativos aos experimentos de auto localização com REQ probabilístico.

5.4 Resultados para Auto Localização com Múltiplos Objetos

O sistema REQ probabilístico foi aplicado em uma sequência de 9483 imagens. As imagens foram capturadas em um ambiente interno contendo três objetos como marcos (seção 5.1.2). O resultado do procedimento de localização (no mapa mostrado na Figura 46) está apresentado na matriz de confusão na Tabela 16. Na primeira coluna estão as regiões de R1 até R28. Todas as demais colunas representam as respostas do sistema, enquanto que as linhas representam a região na qual o robô estava (de acordo com a tabela de referência *gold standard*).

A Tabela 16 tem os valores de resposta do sistema REQ probabilístico em porcentagem para cada região onde as imagens foram capturadas. Na diagonal desta tabela estão representados os valores de acurácia, com acurácia acima de 90% as regiões R2, R3, R4, R6, R9, R10, R11, R12, R13, R15, R16, R22, R21, R24, R25, com acurácia entre 70% e 90% as regiões R1, R5, R7, R14, R19, R20, R23, e com acurácia menor que 70% as regiões R8, R17, R18, R26, R27, R28. As regiões R1, R2, R3, R9, R10, R11, R12, R13, R14, R15, R16, R24, R22, R25 são casos nos quais ou o sistema respondeu como localização a própria região, ou deu como resposta uma região vizinha.

A Tabela 17 refere-se aos resultados de acurácia dos sistemas de auto localização relativos ao mapa qualitativo da Figura 46. Com essa tabela podemos comparar os resultados dos sistemas baseados em raciocínio espacial com e sem a extensão probabilística. A primeira coluna (#objetos) refere-se ao número de objetos alvos visíveis na imagem: nenhum (0), somente um (1), um par (2), três pares (3) e na última linha o total de imagens utilizadas neste experimento. Na quarta coluna (# imagens) os valores representam a quantidade de imagens coletas com o número de objetos alvos visíveis indicados na primeira coluna. Os valores apresentados na segunda e na terceira colunas são referentes às imagens com as quais o sistema retornou corretamente a localização do robô, na segunda coluna (REQ) estão as respostas do sistema utilizando somente o raciocínio espacial qualitativo para inferência, enquanto que na terceira coluna (REQ probabilístico) estão às respostas do sistema utilizando o raciocínio espacial qualitativo com um filtro Bayesiano. Nas duas últimas colunas (%REQ, %REQ probabilístico) estão os valores de acurácia representados em porcentagem para cada abordagem.

Pela Tabela 17, a primeira linha mostra que de 299 imagens sem nenhum objeto somente o sistema PQRS probabilístico acertou em uma única imagem a localização do robô. A última linha mostra que no total de 9483 imagens os totais com resposta de localização correta foram

Tabela 16: Matriz de confusão com os resultados do experimento de auto localização utilizando REQ probabilístico. As linhas são as regiões onde a imagem foi capturada e as colunas a resposta do sistema.

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22	R23	R24	R25	R26	R27	R28	
R1	0,839	0,054	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R2	-	0,991	0,009	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R3	-	0,007	0,971	0,022	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R4	-	0,024	0,003	0,969	0,003	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R5	-	0,208	-	-	0,703	0,089	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R6	-	-	-	-	-	0,984	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R7	-	0,207	-	-	-	0,034	0,736	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,016	-	
R8	-	0,089	-	-	-	-	0,002	0,391	0,354	0,053	-	-	-	-	-	-	-	0,015	-	-	-	-	-	-	-	0,011	0,011	-	
R9	-	-	-	-	-	-	-	1,000	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,075	0,021	-	
R10	-	-	-	-	-	-	-	-	0,922	-	0,078	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R11	-	-	-	-	-	-	-	-	-	0,992	0,008	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R12	-	-	-	-	-	-	-	-	-	0,044	0,956	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R13	-	-	-	-	-	-	-	-	-	-	0,021	0,925	0,054	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R14	-	-	-	-	-	-	-	-	-	-	-	0,833	0,167	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
R15	-	-	-	-	-	-	-	-	-	-	-	-	1,000	0,068	0,932	-	-	-	-	-	-	-	-	-	-	-	-	-	
R16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,002	0,530	0,099	0,001	-	-	0,001	0,004	0,221	-	-	-	-	-	
R17	-	-	-	-	-	-	-	-	0,129	-	-	-	-	-	-	-	-	0,226	-	-	-	-	-	0,098	-	-	-	-	
R18	-	-	-	-	-	-	-	0,676	-	-	-	-	-	-	-	-	0,087	-	0,739	0,109	-	0,043	-	0,022	-	-	-	-	
R19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,004	0,133	-	0,760	0,004	0,087	-	-	-	-	-	-	
R20	-	-	-	-	-	-	-	-	0,008	-	-	-	-	-	-	-	-	-	-	-	0,971	0,029	-	-	-	-	-	0,004	
R21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,019	0,923	0,058	-	-	-	-	-	
R22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,018	-	-	-	-	-	0,875	0,054	-	-	-	
R23	-	-	-	-	-	-	-	0,054	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,913	0,004	0,028	0,032	-	
R24	-	-	-	-	-	-	-	0,024	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1,000	-	-	-	
R25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0,002	0,634	0,026	-	
R26	0,003	-	-	-	-	-	-	-	0,101	-	-	-	-	-	-	-	-	0,236	-	-	-	-	-	-	-	-	-	-	
R27	-	-	-	-	-	-	-	0,016	0,690	-	-	-	-	-	-	-	-	0,284	-	-	-	-	-	-	-	-	0,002	0,008	
R28	-	-	-	-	-	-	-	-	0,860	-	-	-	-	-	-	-	-	0,103	-	-	-	-	-	-	-	-	0,009	0,028	-

Fonte: autor.

Tabela 17: Tabela com os valores de acurácia do sistema REQ probabilístico.

#objetos	REQ	REQ probabilístico	#imagens	%REQ	%REQ probabilístico
0	0	1	299	0,0%	0,3%
1	440	817	3867	11,4%	21,1%
2	3670	3727	4228	86,8%	88,2%
3	970	978	1089	89,7%	89,8%
total	5087	5523	9483	53,6%	58,2%

Fonte: autor.

53,6% (5087) para o sistema utilizando somente o raciocínio espacial para inferência e 58% (5523) para o REQ com o filtro Bayesiano. As outras linhas da Tabela 17 são análogas para imagens com um, dois ou três objetos.

A Tabela 18 tem os valores de acurácia detalhados por região, as linhas têm informações sobre as regiões especificadas na primeira coluna (Regiões). A segunda coluna (#imagens) é referente ao total de imagens capturas em cada região, enquanto que as quatro colunas seguintes são referentes ao total de imagens pelo número de objetos visíveis (#objetos na imagem). As sétima (REQ) e oitava (PREQ) colunas representam o total de vezes que os sistemas acertaram a localização, estes valores de acurácia têm sua representação em porcentagem nas nona (%REQ) e décima (%PREQ) colunas. Por exemplo na região R18 foram capturadas 509 imagens todas com um único objeto e somente o REQ probabilístico conseguiu inferir corretamente em 22,6% (115) das imagens.

A Tabela 19 refere-se somente as regiões na qual ocorre oclusão e as imagens em que somente um objeto aparece na imagem. A coluna *regiões* são as regiões onde ocorre a relação de oclusão total entre um dos pares ($TotallyOccludesNTPP$, $TotallyOccludesNTPPI^{-1}$). A segunda coluna (#imagens com 1 objeto) tem os totais de imagens capturadas nessas regiões em que somente uma caixa aparece na imagem contudo existem duas caixas em cena. A coluna REQ apresenta o total de vezes em que o sistema REQ acertou a localização. Por fim, a última coluna (PREQ) apresenta o total de vezes em que o sistema REQ probabilístico acertou a localização do robô. Como exemplo a última linha mostra que de 515 imagens em região de oclusão total com somente um objeto em 440 e 457 vezes a localização foi inferida corretamente pelos sistemas REQ sem e com filtro Bayesiano respectivamente.

Vamos fazer outra análise agora no que se refere à localização em regiões próximas, como uma resposta qualitativa caso uma região próxima seja dada essa deve ser considerada. Na matriz de confusão (Tabela 16) notem que em grande parte as respostas para a região e vizinhas ocorrem, por exemplo, considerando a vizinhança na região R1 o sistema teria acertado em todas as respostas. Mas vamos analisar considerando também o percurso, no caso em que o robô tem como resposta uma região que além de vizinha é aquela para a qual estava se deslocando. A Tabela 20 mostra uma comparação de acurácia por região, na última coluna (%PREQ Viz.) representa os valores para o REQ probabilístico considerando esse percurso e a vizinhança, a

Tabela 18: Tabela com os valores de acurácia do sistema detalhado por região.

Regiões	#images	#objetos na imagem				REQ	PREQ	%REQ	%PREQ
		0	1	2	3				
R1	112	0	3	78	31	96	94	85,7%	83,9%
R2	115	0	55	60	0	115	114	100,0%	99,1%
R3	415	0	3	346	66	399	403	96,1%	97,1%
R4	290	0	0	46	244	281	281	96,9%	96,9%
R5	101	0	9	87	5	71	71	70,3%	70,3%
R6	64	0	64	0	0	46	63	71,9%	98,4%
R7	87	0	7	80	0	61	64	70,1%	73,6%
R8	1452	0	661	698	93	557	568	38,4%	39,1%
R9	194	1	193	0	0	0	194	0,0%	100,0%
R10	102	0	4	71	27	94	94	92,2%	92,2%
R11	124	0	50	74	0	124	123	100,0%	99,2%
R12	91	0	3	30	58	84	87	92,3%	95,6%
R13	240	0	0	105	135	220	222	91,7%	92,5%
R14	54	0	7	31	16	45	45	83,3%	83,3%
R15	162	0	148	14	0	162	162	100,0%	100,0%
R16	44	0	7	37	0	37	41	84,1%	93,2%
R17	1651	0	577	986	88	859	875	52,0%	53,0%
R18	509	0	509	0	0	0	115	0,0%	22,6%
R19	46	0	1	15	30	34	34	73,9%	73,9%
R20	263	0	103	160	0	201	200	76,4%	76,0%
R21	34	0	0	4	30	29	33	85,3%	97,1%
R22	52	0	0	0	52	48	48	92,3%	92,3%
R23	56	0	3	5	48	49	49	87,5%	87,5%
R24	252	0	95	157	0	231	230	91,7%	91,3%
R25	64	0	2	32	30	61	64	95,3%	100,0%
R26	1970	0	725	1109	136	1183	1248	60,1%	63,4%
R27	503	3	497	3	0	0	1	0,0%	0,2%
R28	436	295	141	0	0	0	0	0,0%	0,0%
Global	9483	299	3867	4228	1089	5087	5523	53,6%	58,2%

Fonte: autor.

Tabela 19: Tabela com os resultados de acurácia para o caso específico das regiões onde ocorre oclusão nas quais somente um objeto aparecem na imagem.

regiões	#imagens com 1 objeto	REQ	PREQ
R2	55	55	55
R6	64	46	63
R11	50	50	50
R15	148	148	148
R20	103	66	66
R24	95	75	75
Total	515	440	457

Fonte: autor.

Tabela 20: Resposta de total de imagens com resposta positiva para o REQ e o REQ probabilístico com o REQ probabilístico considerando as regiões vizinhas que estavam no percurso do robô como resposta positiva .

Regiões	#images	REQ	PREQ	PREQ Viz.	%REQ	%PREQ	%PREQ Viz.
1	112	96	94	110	85,7%	83,9%	98,2%
2	115	115	114	115	100,0%	99,1%	100,0%
3	415	399	403	412	96,1%	97,1%	99,3%
4	290	281	281	283	96,9%	96,9%	97,6%
5	101	71	71	80	70,3%	70,3%	79,2%
6	64	46	63	63	71,9%	98,4%	98,4%
7	87	61	64	66	70,1%	73,6%	75,9%
8	1452	557	568	1159	38,4%	39,1%	79,8%
9	194	0	194	194	0,0%	100,0%	100,0%
10	102	94	94	102	92,2%	92,2%	100,0%
11	124	124	123	124	100,0%	99,2%	100,0%
12	91	84	87	91	92,3%	95,6%	100,0%
13	240	220	222	237	91,7%	92,5%	98,8%
14	54	45	45	54	83,3%	83,3%	100,0%
15	162	162	162	162	100,0%	100,0%	100,0%
16	44	37	41	44	84,1%	93,2%	100,0%
17	1651	859	875	1033	52,0%	53,0%	62,6%
18	509	0	115	115	0,0%	22,6%	22,6%
19	46	34	34	45	73,9%	73,9%	97,8%
20	263	201	200	224	76,4%	76,0%	85,2%
21	34	29	33	33	85,3%	97,1%	97,1%
22	52	48	48	52	92,3%	92,3%	100,0%
23	56	49	49	52	87,5%	87,5%	92,9%
24	252	231	230	231	91,7%	91,3%	91,7%
25	64	61	64	64	95,3%	100,0%	100,0%
26	1970	1183	1248	1294	60,1%	63,4%	65,7%
27	503	0	1	129	0,0%	0,2%	25,6%
28	436	0	0	245	0,0%	0,0%	56,2%
Global	9483	5087	5523	6813	53,6%	58,2%	71,8%

Fonte: autor.

Tabela 21: Comparação que tem os valores de acurácia pelo total de objetos contidos na imagem, considerando a região vizinha que estava no percurso executado pelo robô como resposta positiva.

	total de imagem por objeto	REQ	REQ probabilístico	REQ probabilístico vizinhança
0	299	0,0%	0,3%	70,9%
1	3867	11,4%	21,1%	44,1%
2	4228	86,8%	88,2%	90,2%
3	1089	89,7%	89,8%	99,3%
geral	9483	53,6%	58,2%	71,8%

Fonte: autor.

coluna “PREQ Viz.” tem o número de imagens com resultado positivo. Os valores para o REQ e o REQ probabilístico estão nas colunas adjacentes para comparação, essa tabela é similar a Tabela 18. Neste caso o resultado geral representado na última linha da Tabela 20 passou de 58,2% para 71,8%. Notem que desse modo o desempenho na região $R1$ de 98,2% nem sempre a região vizinha é aquela na qual o robô está se dirigindo.

Outra tabela para comparação que tem os valores de acurácia pelo total de objetos contidos na imagem, semelhante à Tabela 17 e a Tabela 21, onde na coluna “REQ probabilístico vizinhança” estão os valores de acurácia considerando o percurso e a vizinhança. Para o filtro Bayesiano, a falta de objeto na imagem faz com que todas as regiões sejam analisadas, assim o modelo de transição e a crença posterior (previsão) são o que decide qual região, o que acredito elevou para um acerto (ou região do percurso) de 70,9%, quando com uma caixa, algumas regiões não são analisadas (onde ocorre a oclusão total da caixa que está na imagem), assim ocorre casos em que a crença em uma região vizinha diferente daquela que o robô está realmente se deslocando ocorre, por isso a acurácia ficou em torno de 44,1%, para as imagens com mais informações desempenho foi superior a 90%.

5.5 Discussão

Neste capítulo foram apresentados dois sistemas de auto localização aplicado em sistema robótico. Esses sistemas executam de modo automático os algoritmos que executam tarefas semelhantes como: primeiro reconhecer e extrair as características dos elementos alvos em uma imagem; segundo utilizar o raciocínio espacial qualitativo para extrair as relações entre os elementos; para em seguida com a formalização do domínio e com os cálculos de raciocínio espacial inferir informações a partir das imagens, e utilizar essas informações como evidências em um filtro Bayesiano. A crença com maior valor é a resposta do sistema, podendo também atuar na definição de parâmetros como foi o caso do sistema de visão para segmentação de sombras e objetos. O filtro Bayesiano desempenha um papel relevante principalmente na acurácia das respostas quando as informações dos sensores foram incompletas ou inexistentes.

O sistema de PQRS probabilístico utilizou o filtro Bayesiano tanto para auto localização como para auto calibração. Quando tratando com sombras, o sistema de visão necessita de uma calibração diferente conforme muda de posição no ambiente, elevando o grau de incerteza das informações das imagens. Um sistema de calibração com PQRS precisa saber qual região onde está localizado, para encontrar um parâmetro de visão que confirme essa informação. Enquanto que o PQRS probabilístico trabalha com a crença na região, e tanto a região quanto o parâmetro que têm maior valor de crença são escolhidos como resposta do sistema. No ambiente com caixas a incerteza ocorre principalmente devido a informação parcial ou inexistente sobre os pares de objetos.

Em ambos os experimentos mapas qualitativos foram criados a partir da relação de oclusão percebida entre elementos. Quando os elementos estão fixos em um ambiente, um observador em movimento irá notar alterações na relação espacial entre os elementos do ambiente. Formalizamos a localização neste mapa utilizando como base o PQRS (seção 2.2.3) e o ROC (seção 2.2.2). A auto localização é qualitativa e relativa a elementos que estejam no ambiente. Existe um conhecimento da disposição qualitativa dos elementos sem informação sobre a posição exatas deles.

O controle manual não deve interferir na tarefa de auto localização. A auto localização é realizada com inferência a partir dos dados das imagens. Quando a posição é relativa ao ponto de luz e ao objeto, a informação vem da relação sombra-objeto. O robô vai sendo levado de uma região a outra e faz a sua localização pelo que observa. A partir da observação ele vai formando as suas crenças que podem ser reforçadas ou não. Quando o robô tem longa sequência de tempo sem informação, a distribuição de crença do filtro vai ficando uniforme por todas as regiões. Quando isso acontece, ele já não consegue identificar corretamente sua localização. Caso tivesse autonomia poderia parar e observar em volta do ambiente, isso aumentaria a crença em uma das regiões.

O código atual foi desenvolvido para testar as hipóteses sobre a abordagem de raciocínio espacial qualitativo probabilístico para interpretar imagens de forma mais robusta. O sistema para múltiplos objetos está em fase de desenvolvimento, os experimentos aqui relatados foram realizados para verificar se o método de raciocínio espacial qualitativo probabilístico poderia ser utilizado em ambientes com mais elementos.

A segmentação de sombras e de objetos é uma parte essencial quanto ao sucesso da auto localização. A visão tem uma grande influência na localização, se a segmentação não conseguir separar o ambiente do objeto e sombra, isso pode corromper a informação utilizada para determinar as relações entre objeto e sombra na imagem. Assim, partes das informações que sejam alteradas pelos filtros irão alterar os resultados do raciocínio. Nesse trabalho uma calibração é realizada através do software no início dos testes. A partir dessa calibração com um filtro Bayesiano o parâmetro de visão responsável por extrair sombra e objeto da imagem é auto calibrado pelo sistema em tempo de execução. Para múltiplos objetos a escolha de caixa coloridas teve como objetivo neutralizar os problemas de visão.

Diferente do desempenho do PQRS probabilístico que teve uma melhora de 20% relativa ao sistema PQRS, no experimento de auto localização no mapa da Figura 46 teve uma melhora de aproximadamente 5%. Esses são resultados de testes preliminares, que foram realizados com aproximações dos modelos de filtros Bayesianos para verificar a transferência do método utilizado especificamente para informações espaciais relativas às sombras.

O modelo de imagem aqui utilizado é uma aproximação. Um modelo de imagem deve conter as probabilidades das disjunções de conjunções das relações de cada par serem percebidas em cada região. Procuramos manter o mesmo princípio do modelo de transição do PQRS, no

qual o movimento é percebido através da alteração nas características percebidas nas imagens, porém neste experimento perceber o transição através da mudança de relação espacial entre pares foi prejudicada pelo alto número de imagens sem pares. A Tabela 18 mostra que em algumas regiões nenhuma imagem foi capturada com pares.

Outro motivo que causa essa diferença, é que ao realizar a captura da sequência de imagens, nossa atenção foi que o robô passasse por todas as regiões sem atentar para o fato que em algumas não foram mantidos os pares no campo de visão. O filtro Bayesiano fez o que foi proposto, capacitando o sistema de raciocínio espacial qualitativo a lidar com informações parciais. Na Tabela 16 a maior parte dos resultados foi ou a região ou uma região próxima à região que o robô estava localizado. Nas Tabelas 20 e 21 mostram como o filtro fez o trabalho e com alguns ajuste pode ficar mais eficaz.

As imagens com somente um objeto visível são as informações parciais típicas do experimento de auto localização no mapa da Figura 46. Isso porque com exceção dos casos em que ocorre uma oclusão total em cena, com apenas um objeto um sistema utilizando somente o REQ não é capaz de inferir uma localização. Com o filtro Bayesiano o sistema pode obter alguns resultados positivos na auto localização com informação parcial da imagem. Isso fica evidente quando comparamos a Tabela 17 na linha referente a um único objeto na imagem com a última linha da Tabela 19. Os resultados utilizando REQ são os mesmos 440 nas duas tabelas, como esperado já que o mecanismo de inferência REQ utiliza relações espaciais dos pares percebidas em cenas. O sistema REQ probabilístico acertou a localização em 817 imagens, das quais 457 são referentes às regiões onde uma relação espacial de oclusão total ocorre entre os objetos. Assim, o REQ probabilístico mesmo com informação parcial foi capaz de inferir corretamente a localização em 360 casos.

6 CONCLUSÃO

Os resultados relatados foram animadores pois, além de demonstrarem a possibilidade de utilizar técnicas probabilísticas com raciocínio espacial, combinar ambos faz com que os sistemas tenham uma melhora nas suas respostas.

No capítulo 3 começamos a pesquisa com lógica de descrição probabilística com cálculo de direção cardinal onde desenvolvemos um mecanismo de inferência espacial qualitativo com *CRALLC* para um subdomínio de tráfego. Na seção 3.3 propomos duas soluções para desenvolver este mecanismo. Em uma solução utilizamos o software para criar uma rede Bayesiana automaticamente para as inferências, formalizando com *CRALLC* os papéis referentes as relações espaciais implicitamente codificados como conceitos, e a outra solução a partir de uma tabela de composição das relações de direção cardinais permitiu que criássemos uma rede Bayesiana relacional colocando os papéis de relações espaciais entre objetos de modo explícito. Esses mecanismos de inferência não foram testados com dados reais por falta de acesso aos mesmos. Contudo essa pesquisa contribuiu tanto para verificar como poderíamos utilizar um raciocínio espacial com técnicas probabilísticas, quanto para perceber as restrições do *CRALLC* para aplicações em robótica onde uma hierarquia de papel fosse necessária.

A tabela de composição é útil quando a inferência não vem direto das informações dos sensores. Esse conceito foi utilizado nas inferências modeladas em *CRALLC* onde não existe um sensor que localiza a faixa exata onde o carro está. O diagrama conceitual de vizinhança é uma ferramenta que carrega o conceito de transição continua entre as relações. Esse conceito de transição é um conceito inerente às tarefas aplicadas em robótica móvel tal como navegação, rastreamento. Utilizamos esse conceito para criar os modelos de transição, que foram utilizados para inferir transição do robô pelas regiões e a previsão para os próximos estados.

Os sistemas utilizados no capítulo 4 foram codificados a partir dos algoritmos e funções desenvolvidos especificamente sobre os axiomas da teoria de raciocínio espacial qualitativo probabilístico. Estes sistemas estão mais robustos aos ruídos nos dados de leitura do sensor e as informações parciais típicas de ambientes reais. Com a formalização dos domínios criamos mecanismos de inferências para obter informações relevantes sobre o ambiente tridimensional a partir de informações bidimensionais.

Os resultados do capítulo 5 mostram que combinar um sistema baseado em raciocínio espacial qualitativo com filtro Bayesiano aumenta a acurácia do sistema. Ao adaptarmos o algoritmo para realizar a auto localização em dois contextos diferentes, pudemos perceber que independente do contexto, a combinação de um filtro de Bayesiano ao raciocínio espacial qualitativo é a abordagem que obteve melhor desempenho. Os resultados do PQRS probabilístico e do REQ probabilístico levam a um próximo passo, colocar o robô para navegar de forma

autônoma, e isso pode levar o sistema a melhorar a sua acurácia. Começamos a perceber como o diagrama conceitual de vizinhança e tabela de composição são úteis nas aplicações de raciocínio espacial qualitativo com técnicas probabilísticas.

Um sistema foi construído de modo a permitir que o robô se localize relativo ao espaço onde se encontra a partir dos objetos que sejam referência. Esse sistema possibilita o avanço para um sistema em que o robô atue em seu autocontrole, modificando a suas crenças, essa hipótese vem do fato de termos obtidos resultados bons na auto calibração da imagem. Imagem tem um grande número de possibilidades e pode ser utilizada principalmente como suporte para desambiguação ou aumentar o que é percebido por outros sensores, para que o tempo de ação de um robô seja melhorado. Deste modo os métodos espaciais qualitativos devem ser utilizados não como substitutos dos algoritmos numéricos tradicionais, mas como complemento de modo a adicionar um nível de raciocínio em situações em que as informações numéricas não estão disponíveis.

Estamos dando pequenos passos na direção de robôs que possam conviver e interagir de modo mais natural e intuitivo com humanos. A abordagem com raciocínio espacial qualitativo probabilístico aqui proposta não é a solução final, mas espero que seja parte de um mecanismo que permita no futuro uma boa interação entre homens e máquinas próxima ou igual aquelas vistas na ficção. Existem vários possíveis trabalhos futuros derivados desta pesquisa.

Seria importante realizar os experimentos da seção de auto localização com múltiplos objetos, com um modelo de imagem mais completo, com dados anotados durante a captura de imagens. Analisar a complexidade dos métodos de raciocínio espacial qualitativo probabilístico aqui utilizados. E aplicar métodos de programação que melhorem o desempenho em velocidade de processamento dos algoritmos utilizados pelo robô autônomo. Passar esse código para um sistema que opere com o robô em tempo real.

Desenvolver o sistema de modo que o filtro Bayesiano combinado com os métodos de raciocínio espacial qualitativo seja utilizado para modificar parâmetros de controle que determinem a atuação do robô no ambiente tornando esse autônomo. Para uma navegação além do sistema de localização outros pontos importantes ainda devem ser desenvolvidos. A navegação autônoma exige tomada de decisão, quando em determinada região é necessário ir para outra região, a percepção espacial indicará se as mudanças estão ocorrendo na direção certa.

Uma próxima aplicação do sistema de raciocínio espacial qualitativo probabilístico é integrar com outros métodos de localização robótica para formar um sistema mais robusto e eficiente para a tarefa de localização. A ideia é que o método de auto localização qualitativa sirva como heurística para os métodos probabilísticos atuais. Também desenvolver um algoritmo de localização e mapeamento qualitativo simultâneo com nosso atual algoritmo de localização. Esses mapas qualitativos poderiam ser utilizados em sistema com dois robôs, onde um robô tem acesso visual a região que o outro não tem. Utilizar as ferramentas que fazem o mapeamento

qualitativo também para traduzir para um significado semântico que torne o sistema capaz de realizar uma comunicação homem-robô.

A percepção de movimento através da imagem agrega informação. Devemos utilizar os sensores do robô que indicam o movimento para deixar os modelos de transição mais robustos. Além de que novas inferências podem ser formalizadas a partir das mudanças de relações percebidas relativas ao ambiente e aquelas do próprio robô. Prevendo a ação de outros elementos que estejam inseridos no ambiente, ou seja, começar a expandir o sistema para ser utilizados em ambientes dinâmicos. Como o ambiente de tráfego utilizado neste trabalho utilizando *CRALLC* para inferir a localização do veículo.

Para a aplicação do *CRALLC* em robótica móvel seria necessário fazer algumas alterações: a criação de uma regra para que ao tratar com a transitividade de uma composição de relações, para que essas pudessem ser definidas no banco de conhecimento; ou utilizar uma lógica de descrição mais completa que permita a composição. Uma dificuldade inicial para a formalização com o *CRALLC* foi evitar ciclos, em que cada conceito e papel deveriam ser definidos em uma sequência tal, que um conceito que esteja à direita tenha sido definido anteriormente. Um trabalho futuro poderia estender a linguagem de modo tal que ela mesma procuraria a melhor sequência, mostrando de forma mais intuitiva quando um ciclo ocorreu.

REFERÊNCIAS

BAADER, F.; CALVANESE, D.; MCGUINNESS, D.; NARDI, D.; PATEL-SCHNEIDER, P. Description logic handbook. Cambridge University Press, 2003. Citado na página 62.

BAADER, F.; NUTT, W. Basic description logics. In: *Description Logic Handbook*. [S.l.]: Cambridge University Press, 2002. p. 47–100. Citado na página 65.

BACCHUS, F.; GROVE, A.; HALPERN, J. Y.; KOLLER, D. From statistical knowledge bases to degrees of belief. *Artificial Intelligence*, v. 87, p. 75–143, 1996. Citado na página 31.

BENNETT, B.; COHN, A.; MAGEE, D. Enforcing global spatio-temporal consistency to enhance reliability of moving object tracking and classification. *Künstliche Intelligenz*, v. 2, p. 32–35, 2005. Citado na página 71.

BONNER, A. J.; KIFER, M. An overview of transaction logic. *Theoretical Computer Science*, Elsevier, v. 133, n. 2, p. 205–265, 1994. Citado na página 71.

CAMPOS, C. P. de; COZMAN, F. G.; LUNA, J. E. O. Assessing a consistent set of sentences in relational probabilistic logic with stochastic indepenence. *Journal of Applied Logic: Combining Probability and Logic*, v. 7, p. 137–154, 2009. Citado na página 31.

CHEN, J.; COHN, A. G.; LIU, D.; WANG, S.; OUYANG, J.; YU, Q. A survey of qualitative spatial representations. *The Knowledge Engineering Review*, Cambridge Univ Press, p. 1–31, 2013. Citado 3 vezes nas páginas 30, 37 e 38.

CHEN, S. Kalman filter for robot vision: a survey. *Industrial Electronics, IEEE Transactions on*, IEEE, v. 59, n. 11, p. 4409–4420, 2012. Citado na página 70.

COHN, A. G.; HAZARIKA, S. M. Qualitative spatial representation and reasoning: An overview. *Fundamenta Informaticae*, v. 46, n. 1-2, p. 1–29, 2001. Citado 2 vezes nas páginas 30 e 38.

COZMAN, F. G.; CAMPOS, C. P. D.; IDE, J. S.; ROCHA, J. C. F. da. Propositional and relational Bayesian networks associated with imprecise and qualitative probabilistic assessments. In: AUAI PRESS. *Proceedings of the 20th conference on Uncertainty in artificial intelligence*. [S.l.], 2004. p. 104–111. Citado na página 62.

COZMAN, F. G.; POLASTRO, R. B. Loopy propagation in a probabilistic description logic. In: *SUM '08: Proceedings of the 2nd international conference on Scalable Uncertainty Management*. Berlin, Heidelberg: Springer-Verlag, 2008. p. 120–133. ISBN 978-3-540-87992-3. Citado 7 vezes nas páginas 31, 32, 37, 62, 66, 69 e 79.

CRISTANI, M.; GABRIELLI, N. Practical issues of description logics for spatial reasoning. In: *AAAI Spring Symposium: Benchmarking of Qualitative Spatial and Temporal Reasoning Systems*. [S.l.: s.n.], 2009. p. 5–10. Citado na página 34.

CUI, Z.; COHN, A. G.; RANDELL, D. A. Qualitative simulation based on a logical formalism of space and time. In: *AAAI*. [S.l.: s.n.], 1992. p. 679–684. Citado 2 vezes nas páginas 39 e 118.

DÁMATO, C.; FANIZZI, N.; LUKASIEWICZ, T. Tractable reasoning with bayesian description logics. In: *Scalable Uncertainty Management*. [S.l.]: Springer, 2008. p. 146–159. Citado na página 31.

DELLAERT, F.; FOX, D.; BURGARD, W.; THRUN, S. Monte carlo localization for mobile robots. In: IEEE. *IEEE International Conference on Robotics and Automation*. [S.l.], 1999. v. 2, p. 1322–1328. Citado na página 70.

DING, Z.; PENG, Y. A probabilistic extension to ontology language OWL. In: IEEE. *Proceedings of the 37th Annual Hawaii international conference on System Sciences*. [S.l.], 2004. p. 10–pp. Citado na página 64.

DING, Z.; PENG, Y.; PAN, R. BayesOWL: Uncertainty modeling in semantic web ontologies. In: *Soft Computing in Ontologies and Semantic Web*. Berlin/Heidelberg: Springer, 2006, (Studies in Fuzziness and Soft Computing, v. 204). p. 3–29. ISBN 978-3-540-33472-9. Citado na página 64.

DÜRIG, M.; STUDER, T. Probabilistic ABox reasoning: preliminary results. In: *Description Logics*. [S.l.: s.n.], 2005. p. 104–111. Citado 2 vezes nas páginas 31 e 63.

FELICE, G. D.; FOGLIARONI, P.; WALLGRÜN, J. O. A hybrid geometric-qualitative spatial reasoning system and its application in gis. In: *Spatial Information Theory*. [S.l.]: Springer, 2011. p. 188–209. Citado na página 30.

FENELON-PEREIRA, V.; SANTOS, P. E.; COZMAN, F. G. Reconhecimento de objetos baseado em contexto utilizando a lógica de descrição probabilística. *IX Encontro em Inteligência Artificial (ENIA)*, 2012. Citado 3 vezes nas páginas 33, 69 e 174.

FENELON, V.; HUMMEL, B.; SANTOS, P. E.; COZMAN, F. G. Encoding spatial domains with relational bayesian networks. In: *UAI 2010*. [S.l.: s.n.], 2010. p. 49–54. Citado 18 vezes nas páginas 31, 32, 33, 34, 40, 66, 72, 80, 81, 82, 86, 87, 92, 93, 94, 95, 138 e 173.

FENELON, V.; SANTOS, P.; DEE, H.; COZMAN, F. Reasoning about shadows in a mobile robot environment. *Applied Intelligence*, Springer US, p. 1–13, 2012. ISSN 0924-669X. Citado 9 vezes nas páginas 33, 34, 51, 99, 100, 101, 108, 146 e 173.

FENELON, V.; SANTOS, P. E.; MARTINS, M. F.; COZMAN, F. G. A qualitative-probabilistic approach to autonomous mobile robot self localisation and self vision calibration. *The 2nd Brazilian Conference on Intelligent Systems (BRACIS-13)*, 2013. Citado 9 vezes nas páginas 33, 34, 40, 71, 119, 130, 146, 147 e 174.

FERNYHOUGH, J.; COHN, A. G.; HOGG, D. C. Constructing qualitative event models automatically from video input. *Image and Vision Computing*, v. 18, p. 81–103, 2000. Citado na página 71.

FOGLIARONI, P.; WALLGRÜN, J. O.; CLEMENTINI, E.; TARQUINI, F.; WOLTER, D. A qualitative approach to localization and navigation based on visibility information. In: *Spatial Information Theory*. [S.l.]: Springer, 2009. p. 312–329. Citado 2 vezes nas páginas 34 e 70.

FOX, D.; BURGARD, W.; THRUN, S. Markov localization for mobile robots in dynamic environments. *Journal of Artificial Intelligence Research*, v. 11, p. 391–427, 1999. Citado na página 70.

FRANK, A. U. Qualitative spatial reasoning: Cardinal directions as an example. *International Journal of Geographical Information Science*, v. 10, n. 3, p. 269–290, 1996. Citado 3 vezes nas páginas 37, 56 e 79.

FREKSA, C. Using orientation information for qualitative spatial reasoning. In: *Theories and Methods of Spatial-Temporal Reasoning in Geographic Space*. [S.l.]: Springer-Verlag, 1992, (LNCS, v. 629). Citado na página 38.

FRIEDMAN, N.; KOLLER, L. G. ad D.; PFEFFER, A. Learning probabilistic relational models. In: *International Joint Conference on Artificial Intelligence*. [S.l.: s.n.], 1999. p. 1300–1309. Citado 2 vezes nas páginas 63 e 64.

FUHR, N. Probabilistic datalog — a logic for powerful retrieval methods. In: *Conference on Research and Development in Information Retrieval*. Seattle, Washington: [s.n.], 1995. p. 282–290. Citado 2 vezes nas páginas 63 e 64.

FUKUSHIGE, Y. Representing probabilistic knowledge in the semantic web. In: *W3C Workshop on Semantic Web for Life Sciences*. Cambridge, MA: [s.n.], 2004. Citado na página 63.

GALTON, A. Lines of sight. In: *Proc. of the Seventh Annual Conference of AI and Cognitive Science*. Dublin, Ireland: [s.n.], 1994. p. 103–113. Citado 2 vezes nas páginas 40 e 45.

GETOOR, L.; TASKAR, B. *Introduction to Statistical Relational Learning*. [S.l.]: MIT Press, 2007. ISBN 0262072882. Citado na página 64.

GIUGNO, R.; LUKASIEWICZ, T. P-SHOQ(D): A probabilistic extension of SHOQ(D) for probabilistic ontologies in the semantic web. In: *European Conference on Logics in Artificial Intelligence*. Cosenza, Italy: Springer, 2002. v. 2424, p. 86–97. Citado na página 63.

GONZALEZ, R. E. W. R. C. *Digital Image Processing*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2001. ISBN 0201180758. Citado 4 vezes nas páginas 72, 73, 74 e 75.

GOTTS, N. M. How far can we ‘C’? defining a doughnut using connection alone. In: DOYLE, J.; SANDEWALL, E.; TORASSO, P. (Ed.). *KR’94: Principles of Knowledge Representation and Reasoning*. San Francisco, California: Morgan Kaufmann, 1994. p. 246–257. Citado na página 38.

GUO, P.; ZHANG, R.-x. Spatial topology relationships concrete domain and consistency determination. In: *IEEE. Information Engineering, 2009. ICIE’09. WASE International Conference on*. [S.l.], 2009. v. 1, p. 611–614. Citado na página 34.

HALPERN, J. Y. *Reasoning about Uncertainty*. Cambridge, Massachusetts: MIT Press, 2003. Citado na página 63.

HANSEN, P.; JAUMARD, B. *Probabilistic satisfiability*. [S.l.], 1996. Citado na página 63.

HASENFRATZ, J. M.; LAPIERRE, M.; HOLZSCHUCH, N.; SILLION, F. X. A survey of real-time soft shadows algorithms. *Eurographics Star - State of The Art Report*, 2003. Citado na página 47.

HAZARIKA, S. M.; COHN, A. G. Abducing qualitative spatio-temporal histories from partial observations. In: *Proc. of KR*. Toulouse, France: [s.n.], 2002. p. 14–25. Citado na página 71.

- HEINSOHN, J. Probabilistic description logics. In: *Conference on Uncertainty in Artificial Intelligence*. [S.l.: s.n.], 1994. p. 311–318. Citado na página 63.
- HERMAN, D. Spatial cognition in natural-language narratives. In: *Proceedings of the AAAI Fall Symposium on Narrative Intelligence*. [S.l.: s.n.], 1999. p. 21–25. Citado na página 30.
- HOLI, M.; HYVÖNEN, E. Modeling degrees of conceptual overlap in semantic web ontologies. In: *Uncertainty Reasoning for the Semantic Web*. [S.l.: s.n.], 2005. p. 98–99. Citado na página 63.
- HUMMEL, B. Dynamic and mobile gis: Investigating changes in space and time. In: _____. [S.l.]: CRC Press, 2006. cap. Map Matching for Vehicle Guidance. Citado 2 vezes nas páginas 79 e 80.
- HUMMEL, B. *Description Logic for Intersection Understanding at the Example of Urban Road Intersections*. Tese (Doutorado) — University of Karlsruhe, April 2009. Citado 4 vezes nas páginas 32, 64, 71 e 81.
- HUMMEL, B.; THIEMANN, W.; LULCHEVA, I. Scene understanding of urban road intersections with description logic. *Logic and Probability for Scene Interpretation*, n. 08091, 2008. Citado na página 33.
- HUNG, E.; GETOOR, L.; SUBRAHMANIAN, V. S. Probabilistic interval XML. *ACM Transactions on Computational Logic*, v. 8, n. 4, 2007. Citado na página 63.
- JAEGER, M. *A Probabilistic Extension of Terminological Logics*. [S.l.]: Max-Planck-Institut für Informatik, 1994. Citado na página 63.
- JAEGER, M. Probabilistic reasoning in terminological logics. In: *Principles of Knowledge Representation (KR)*. [S.l.: s.n.], 1994. p. 461–472. Citado 3 vezes nas páginas 31, 62 e 63.
- JAEGER, M. Relational Bayesian networks. In: *Conference on Uncertainty in Artificial Intelligence*. San Francisco, California: Morgan Kaufmann, 1997. p. 266–273. Citado 3 vezes nas páginas 62, 63 e 64.
- JAEGER, M.; RAEDT, L. D.; DIETTERICH, T.; GETOOR, L.; MUGGLETON, S. Importance sampling on relational Bayesian networks. *Probabilistic, Logical and Relational Learning-Towards a Synthesis. Internationales Begegnungs-und Forschungszentrum fuer Informatik (IBFI)*, 2006. Citado na página 62.
- KANDA, T.; KAMASIMA, M.; IMAI, M.; ONO, T.; SAKAMOTO, D.; ISHIGURO, H.; ANZAI, Y. A humanoid robot that pretends to listen to route guidance from a human. *Springer Science + Business Media, LLC 2006*, Auton Robot, September 2006. Citado na página 38.
- KERSTEN, D.; MAMASSIAN, P.; KNILL, D. C. Moving cast shadows and the perception of relative depth. *Max-Planck-Institut fuer biologische Kybernetik*, Technical Report No.6, JUNE 1994. Disponível em: <ftp://ftp.mpik-tueb.mpg.de/pub/mpi-memos/TR-006.ps.Z>. Citado 2 vezes nas páginas 47 e 48.
- KNILL, D. C.; MAMASSIAN, P.; KERSTEN, D. Geometry of shadows. *Journal of the Optical Society of America A*, v. 14, p. 3216–3232, 1997. Citado na página 46.
- KOLLER, D.; PFEFFER, A. Object-oriented Bayesian networks. In: *Conference on Uncertainty in Artificial Intelligence*. [S.l.: s.n.], 1997. p. 302–313. Citado na página 64.

- KOLLER, D.; PFEFFER, A. Probabilistic frame-based systems. In: *National Conference on Artificial Intelligence (AAAI)*. [S.l.: s.n.], 1998. p. 580–587. Citado 2 vezes nas páginas 63 e 64.
- KORB, K.; NICHOLSON, A. E. *Bayesian Artificial Intelligence*. Boca Raton, FL, USA: CRC Press, Inc., 2003. ISBN 1584883871. Citado na página 62.
- KRUSE, T.; PANDEY, A. K.; ALAMI, R.; KIRSCH, A. Human-aware robot navigation: A survey. *Robotics and Autonomous Systems*, Elsevier, v. 61, n. 12, p. 1726–1743, 2013. Citado na página 69.
- LAKSHMANAN, L. V. S.; SADRI, F. Probabilistic deductive databases. In: *Symposium on Logic Programming*. [S.l.: s.n.], 1994. p. 254–268. Citado na página 63.
- LAKSHMANAN, L. V. S.; SHIRI, N. A parametric approach to deductive databases with uncertainty. *IEEE Transactions on Knowledge and Data Engineering*, v. 13, n. 4, p. 554–570, 2001. Citado na página 63.
- LASKEY, K. B.; COSTA, P. C. G. da. Of starships and klingons: Bayesian logic for the 23rd century. *CoRR*, abs/1207.1354, p. abs/1207.1354, 2012. Citado 2 vezes nas páginas 31 e 64.
- LEVITT, T. S.; LAWTON, D. T. Qualitative navigation for mobile robots. *Artificial intelligence*, Elsevier, v. 44, n. 3, p. 305–360, 1990. Citado na página 69.
- LIGOZAT, G. Reasoning about cardinal directions. *J. Vis. Lang. Comput.*, v. 9, n. 1, p. 23–44, 1998. Citado 4 vezes nas páginas 38, 58, 79 e 83.
- LUKASIEWICZ, T. Expressive probabilistic description logics. *Artif. Intell.*, Elsevier Science Publishers Ltd., Essex, UK, v. 172, n. 6-7, p. 852–883, 2008. ISSN 0004-3702. Citado na página 63.
- LUNA, J. E. O.; CAMPOS, C. P. de; COZMAN, F. G. Independence in relational languages with finite domains. In: *Encontro Nacional de Inteligência Artificial*. Rio de Janeiro, RJ: [s.n.], 2007. p. 1420–1429. Citado na página 34.
- MAMASSIAN, P.; GOUTCHER, R. Prior knowledge on the illumination position. In: *Cognition*. [S.l.]: Elsevier, 2001. v. 81, n. 1, p. 1–9(9). Citado na página 139.
- MAMASSIAN, P.; KNILL, D. C.; KERSTEN, D. The perception of cast shadows. *Trends in cognitive Sciences*, v. 2, n. 8, p. 288–295, August 1998. Citado na página 47.
- MATSUYAMA, T.; HWANG, V. S. *SIGMA: A Knowledge-Based Image Understanding System*. New York, U.S.: Plenum Press, 1990. Citado na página 70.
- MCCALL, J. C.; TRIVEDI, M. M. Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation. *IEEE Transactions on Intelligent Transportation Systems*, v. 7:1, p. 20–37, 2006. Citado na página 71.
- MCCARTHY, J. What is artificial intelligence. URL: <http://www-formal.stanford.edu/jmc/whatisai.html>, 2007. Citado na página 38.
- MCCLELLAND, M.; CAMPBELL, M.; ESTLIN, T. Qualitative relational mapping for planetary rovers. In: *Workshops at the Twenty-Seventh AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2013. Citado na página 70.

- MILCH, B.; RUSSELL, S. First-order probabilistic languages: into the unknown. In: *International Conference on Inductive Logic Programming*. [S.l.: s.n.], 2007. Citado 2 vezes nas páginas 31 e 64.
- MITRA, P.; NOY, N. F.; JAISWAL, A. OMEN: A probabilistic ontology mapping tool. In: *ISWC*. [S.l.: s.n.], 2005. p. 537–547. Citado na página 63.
- MORATZ, R.; RENZ, J.; WOLTER, D. Qualitative spatial reasoning about line segments. In: *ECAI*. [S.l.: s.n.], 2000. p. 234–238. Citado na página 38.
- NAGEL, H.-H. Image sequence evaluation: 30 years and still going strong. In: *Proc. of ICPR*. Barcelona, Spain: [s.n.], 2000. p. 1149–1158. Citado na página 71.
- NEHMZOW, U. *Mobile robotics: a practical introduction*. [S.l.]: Springer Verlag, 2003. Citado 2 vezes nas páginas 37 e 69.
- NEUMANN, B.; MÖLLER, R. On scene interpretation with description logics. *Image and Vision Computing, Special Issue on Cognitive Vision*, Elsevier, v. 26, n. 1, p. 82–101, 2008. Citado na página 71.
- NGO, L.; HADDAWY, P. Answering queries from context-sensitive probabilistic knowledge bases. *Theoretical Computer Science*, v. 171, n. 1–2, p. 147–177, 1997. Citado 2 vezes nas páginas 63 e 64.
- NIKOLOV, A.; UREN, V.; MOTTA, E.; ROECK, A. de. Using the Dempster-Shafer theory of evidence to resolve ABox inconsistencies. In: *Workshop on Uncertainty Reasoning for the Semantic Web*. [S.l.: s.n.], 2007. Citado 2 vezes nas páginas 31 e 63.
- NILSSON, N. J. Probabilistic logic. *Artif. Intell.*, v. 28, n. 1, p. 71–87, 1986. Citado na página 64.
- NOTTELMANN, H.; FUHR, N. pDAML+OIL: A probabilistic extension to DAML+OIL based on probabilistic datalog. In: *Proceedings IPMU*. [S.l.: s.n.], 2004. Citado na página 64.
- NOTTELMANN, H.; FUHR, N. Adding probabilities and rules to OWL lite subsets based on probabilistic datalog. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, v. 14, n. 1, p. 17–42, 2006. Citado na página 64.
- PARSONS, S. Current approaches to handling imperfect information in data and knowledge bases. *IEEE Transactions on Knowledge and Data Engineering*, v. 8, n. 3, p. 353–372, June 1996. ISSN 1041-4347. Citado na página 63.
- PEARL, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. [S.l.]: Morgan Kaufmann, 1988. Citado na página 61.
- PIRRI, F.; MENTUCCIA, I.; STORRI, S. The domestic robot: A friendly cognitive system takes care of your home. In: BASTEN, T.; GEILEN, M. C. W.; GROOT, H. W. H. de (Ed.). *Ambient Intelligence: Impact on Embedded System Design*. [S.l.]: Kluwer Academic Publishers, 2003. Citado na página 37.
- POLASTRO, R.; CORRÊA, F.; COZMAN, F.; JR, J. O. Semantic mapping with a probabilistic description logic. *SBIA*, 2010. Citado na página 69.

POLASTRO, R. B. *Lógica probabilística baseada em redes Bayesianas relacionais com inferência em primeira ordem*. Tese (Tese (Doutorado em Engenharia de Controle e Automação Mecânica)) — Escola Politécnica, Universidade de São Paulo, Escola Politécnica, Universidade de São Paulo, São Paulo,, 2012. Citado 3 vezes nas páginas 31, 67 e 68.

POOL, M.; AIKIN, J. KEEPER and protege: An elicitation environment for Bayesian inference tools. In: *Workshop on Protege and Reasoning (at the 7th International Protege Conference)*. [S.l.: s.n.], 2004. Citado na página 63.

POOLE, D. Probabilistic Horn abduction and Bayesian networks. *Artificial Intelligence*, v. 64, p. 81–129, 1993. Citado 2 vezes nas páginas 63 e 64.

POOLE, D.; GOEBEL, R.; ALELIUNAS, R. Theorist: A logical reasoning system for defaults and diagnosis. In: CERCONE, N.; MCCALLA, G. (Ed.). *The Knowledge Frontier – Essays in the Representation of Knowledge*. [S.l.]: Springer-Verlag, 1987. p. 331–352. Citado na página 70.

PRATI, A.; MIKIC, I.; TRIVEDI, M. M.; CUCCHIARA, R. Detecting moving shadows: Algorithms and evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.*, IEEE Computer Society, Washington, DC, USA, v. 25, n. 7, p. 918–923, 2003. ISSN 0162-8828. Citado na página 47.

RANDELL, D.; CUI, Z.; COHN, A. G. A spatial logic based on regions and connection. In: *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning*. San Mateo: Morgan Kaufmann, 1992. p. 165–176. Citado 5 vezes nas páginas 37, 38, 39, 40 e 45.

RANDELL, D. A.; WITKOWSKI, M. Building large composition tables via axiomatic theories. In: *Proc KR*. [S.l.: s.n.], 2002. p. 26–36. Citado 7 vezes nas páginas 39, 40, 41, 43, 44, 45 e 49.

RANDELL, D. A.; WITKOWSKI, M.; SHANAHAN, M. From images to bodies: Modelling and exploiting spatial occlusion and motion parallax. In: NEBEL, B. (Ed.). *IJCAI*. [S.l.]: Morgan Kaufmann, 2001. p. 57–66. ISBN 1-55860-777-3. Citado 5 vezes nas páginas 37, 40, 42, 43 e 45.

REITER, R.; MACKWORTH, A. A logical framework for depiction and image interpretation. *Artificial Intelligence*, v. 41, n. 2, p. 125–155, 1989. Citado na página 70.

RENZ, J. *Qualitative Spatial Reasoning with Topological Information*. [S.l.]: Springer-Verlag, 2002. Citado 2 vezes nas páginas 38 e 39.

SANTOS, M. V.; BRITO, R. C. de; PARK, H.-H.; SANTOS, P. Logic-based interpretation of geometrically observable changes occurring in dynamic scenes. *Applied Intelligence*, v. 1, 2008. In print. Citado na página 71.

SANTOS, P.; DEE, H. M.; FENELON, V. Notes on a qualitative theory of shadows. *Cognitive Robotics workshop at ECAI 2008*, p. 47–54, 2008. University of Patras - Patras, Greece. Citado 13 vezes nas páginas 32, 37, 46, 47, 50, 51, 52, 56, 57, 70, 99, 101 e 140.

SANTOS, P.; DEE, H. M.; FENELON, V. Qualitative robot localisation using information from cast shadows. *ICRA-09 IEEE International Conference on Robotics and Automation*, 2009. Citado 5 vezes nas páginas 32, 46, 51, 56 e 99.

SANTOS, P.; HUMMEL, B.; FENELON, V.; COZMAN, F. G. Probabilistic logic encoding of spatial domains. In: *UniDL 2010*. [S.l.: s.n.], 2010. Citado 15 vezes nas páginas 31, 32, 33, 34, 66, 72, 81, 82, 86, 87, 89, 90, 91, 138 e 173.

SANTOS, P.; SHANAHAN, M. Hypothesising object relations from image transitions. In: HARMELEN, F. van (Ed.). *Proc. of ECAI-02*. Lyon, France: [s.n.], 2002. p. 292–296. Disponível em: <citeseer.ist.psu.edu/532993.html>. Citado na página 38.

SANTOS, P.; SHANAHAN, M. Hypothesising object relations from image transitions. In: HARMELEN, F. van (Ed.). *Proc. of ECAI*. Lyon, France: [s.n.], 2002. p. 292–296. Citado na página 71.

SANTOS, P.; SHANAHAN, M. A logic-based algorithm for image sequence interpretation and anchoring. In: *Proc. of IJCAI*. Acapulco, Mexico: [s.n.], 2003. p. 1408–1410. Citado na página 71.

SANTOS, P. E. Reasoning about depth and motion from an observer's viewpoint. *Spatial Cognition and Computation*, v. 7, n. 2, p. 133–178, 2007. Citado na página 71.

SCHERL, R. B.; LEVESQUE, H. J. The frame problem and knowledge-producing actions. In: *Proceedings of the Eleventh National Conference on Artificial Intelligence (AAAI-93)*. Washington, D.C., USA: AAAI Press/MIT Press, 1993. p. 689–697. Disponível em: <citeseer.ist.psu.edu/scherl93frame.html>. Citado na página 38.

SCHLIEDER, C. Qualitative shape representation. In: BURROUGH, P. A.; FRANK, A. U. (Ed.). *Geographic Objects with Indeterminate Boundaries*. [S.l.]: Taylor & Francis Inc., 1996. p. 123–140. Citado 2 vezes nas páginas 38 e 69.

SCHMIDT-SCHAUSS, M.; SMOLKA, G. Attributive concept descriptions with complements. *Artificial Intelligence*, v. 48, n. 1, p. 1–26, 1991. Citado 2 vezes nas páginas 62 e 65.

SCHROEDER, C.; NEUMANN, B. On the logics of image interpretation: Model construction in a formal knowledge representation framework. In: *International Conference on Image Processing*. Switzerland: [s.n.], 1996. v. 2, p. 785–788. Citado na página 71.

SCHULZ, R.; GLOVER, A.; MILFORD, M. J.; WYETH, G.; WILES, J. Lingodroids: Studies in spatial cognition and language. In: IEEE. *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. [S.l.], 2011. p. 178–183. Citado na página 30.

SEBASTIANI, F. A probabilistic terminological logic for modelling information retrieval. 1994. Citado na página 31.

SEBASTIANI, F. A probabilistic terminological logic for modelling information retrieval. In: *International ACM Conference on Research and Development in Information Retrieval*. Dublin, Ireland: Springer-Verlag, 1994. p. 122–130. Citado na página 63.

SHANAHAN, M. Robotics and the common sense informatic situation. In: *Proc. of ECAI*. Budapest, Hungary: [s.n.], 1996. p. 684–688. Citado na página 71.

SHAPIRO, L. G.; STOCKMAN, G. C. *Computer Vision*. [S.l.]: Tom Robbins, 2001. 371-409 p. Citado 3 vezes nas páginas 73, 74 e 75.

SHAPIRO, S. C. Knowledge representation. *Encyclopedia of Cognitive Science*, Macmillan Publishers Ltd, v. 2, p. 671–680, 2003. Citado na página 29.

SHIRAZI, A.; AMIR, E. Probabilistic modal logic. In: *AAAI'07: Proceedings of the 22nd national conference on Artificial intelligence*. [S.l.]: AAAI Press, 2007. p. 489–495. ISBN 978-1-57735-323-2. Citado na página 30.

SIEGWART, R.; NOURBAKHSH, I. R. *Introduction to Autonomous Mobile Robots*. [S.l.]: The MIT press, 2004. Citado na página 69.

SJÖÖ, K.; AYDEMIR, A.; JENSFELT, P. Topological spatial relations for active visual search. *Robotics and Autonomous Systems*, Elsevier, v. 60, n. 9, p. 1093–1107, 2012. Citado 2 vezes nas páginas 34 e 35.

SKIADOPOULOS, S.; KOUBARAKIS, M. Composing cardinal direction relations. In: *Advances in Spatial and Temporal Databases*. [S.l.]: Springer, 2001. p. 299–317. Citado na página 93.

STAKER, R. Reasoning in expressive description logics using belief networks. In: *International Conference on Information and Knowledge Engineering*. Las Vegas, USA: [s.n.], 2002. p. 489–495. Citado na página 64.

STOCK, O. *Spatial and Temporal Reasoning*. Norwell, MA, USA: Kluwer Academic Publishers, 1997. ISBN 0792346440. Citado 4 vezes nas páginas 30, 37, 38 e 39.

STRACCIA, U. Description logics with fuzzy concrete domains. In: *Conference on Uncertainty in Artificial Intelligence*. [S.l.: s.n.], 2005. Citado na página 63.

STUART, R.; PETER, N. *Artificial Intelligence: a Modern Approach*. New Jersey, USA: Prentice-Hall, 2010. Citado 2 vezes nas páginas 60 e 61.

THRUN, S. Robotic mapping: A survey. *Exploring artificial intelligence in the new millennium*, Morgan Kaufmann San Mateo, CA, p. 1–35, 2002. Citado 2 vezes nas páginas 31 e 70.

THRUN, S.; BURGARD, W.; FOX, D. et al. *Probabilistic robotics*. [S.l.]: MIT press Cambridge, 2005. Citado 5 vezes nas páginas 29, 31, 59, 60 e 70.

UDREA, O.; SUBRAHMANYAN, V. S.; MAJKIC, Z. Probabilistic RDF. In: *IEEE International Conference on Information Reuse and Integration*. [S.l.: s.n.], 2006. p. 172–177. Citado na página 63.

VINCI, L. da. *The Notebooks of Leonardo da Vinci*. [S.l.]: Dover, 1970. Citado na página 46.

WAGNER, T.; VISSER, U.; HERZOG, O. Egocentric qualitative spatial knowledge representation for physical robots. *Robotics and Autonomous Systems*, Elsevier, v. 49, n. 1, p. 25–42, 2004. Citado na página 70.

WELLMAN, M. P.; BREESE, J. S.; GOLDMAN, R. P. From knowledge bases to decision models. *Knowledge Engineering Review*, v. 7, n. 1, p. 35–53, 1992. Citado 2 vezes nas páginas 63 e 64.

WITKOWSKI, M.; SHANAHAN, M.; SANTOS, P.; RANDELL, D. Cognitive robotics: On the semantic knife-edge. In: *Towards Intelligent Mobile Robots*. [S.l.: s.n.], 2001. Citado 2 vezes nas páginas 45 e 46.

WOLTER, D.; FREKSA, C.; LATECKI, L. J. Towards a generalization of self-localization. In: *Robotics and cognitive approaches to spatial mapping*. [S.l.]: Springer, 2008. p. 105–134. Citado na página 70.

WOLTER, D.; WALLGRÜN, J. O. Qualitative spatial reasoning for applications: New challenges and the sparq toolbox. *Qualitative spatio-temporal representation and reasoning: Trends and future directions. IGI Global, Hershey. doi*, v. 10, p. 978–1, 2012. Citado na página 29.

YANG, Y.; CALMET, J. OntoBayes: An ontology-driven uncertainty model. In: *Proceedings IAWTIC*. [S.l.: s.n.], 2005. Citado na página 63.

YELLAND, P. M. *Market analysis using a combination of Bayesian networks and description logics*. [S.l.], 1999. Citado na página 64.

GLOSSÁRIO

Axioma	— "Na lógica tradicional, um axioma ou postulado é uma sentença ou proposição que não é provada ou demonstrada e é considerada como óbvia ou como um consenso inicial necessário para a construção ou aceitação de uma teoria. Por essa razão, é aceito como verdade e serve como ponto inicial para dedução e inferências de outras verdades (dependentes de teoria)". wikipedia
Fórmula	— "Em Lógica, uma fórmula é uma frase, construída segundo as regras gramaticais de uma determinada linguagem formal, a respeito de objetos do universo de discurso." fonte: wikipedia
Heurística	— do dicionário Houaiss - "...método de investigação baseado na aproximação progressiva de um dado problema..."
Hodômetro	— do dicionário Houaiss - "... instrumento que indica distância percorridas por pedestres ou por veículos. Observação: forma não preferencial: odômetro."
mereologia	— do Wikcionário "(Lógica) teoria ou estudo lógico-matemático das relações entre as partes e o todo e das relações entre as partes no interior de um todo."
mereotopologia	— tradução livre de mereotopology
mereotopology	— is a first-order theory, embodying mereological and topological concepts, of the relations among wholes, parts, parts of parts, and the boundaries between parts.
Paralaxe	— do dicionário Houaiss "...deslocamento aparente de um objeto quando se muda o ponto de observação..."
Penumbra	— meia-luz, sombra parcial.
Psicofísicos	— do dicionário Houaiss "Relativo a psicofísica. Psicofísica é o ramo da psicologia que estuda a relação entre estímulos físicos e as respectivas sensações."
Satisfazibilidade	— termo utilizado neste trabalho, para substituir o termo Satisfatibilidade, por ter como raiz a palavra satisfazer.
Satisfatibilidade	— do inglês <i>Satisfiability</i> é um conceito em lógica matemática ...do wikipédia "Uma fórmula é satisfatível se é possível achar uma interpretação (modelo) que torne a fórmula verdadeira."
Segmentação	— ato de separar por característica.
Topológico	— do dicionário Priberam da Língua Portuguesa "adjetivo - relativo à topologia." http://www.priberam.pt/dlpo/topol%c3%b3gico [consultado em 12-01-2014].

ANEXO A – TRABALHOS PUBLICADOS

Os artigos publicados durante o doutorado estão relacionados abaixo. Os trabalhos relatados nestes artigos fazem parte dessa tese. Em ordem cronológica:

- Probabilistic logic encoding of spatial domains (SANTOS et al., 2010).

This paper presents a formalisation of a spatial domain in terms of a qualitative spatial reasoning formalism, encoded in a probabilistic description logic. The QSR formalism chosen is a subset of a cardinal direction calculus and the probabilistic description logic used has the relational structures of the well-known $\mathcal{ALC}$ language, allied with the inference methods of Bayesian Networks. We consider a scenario consisting of a road navigated by an experimental vehicle equipped with three on-board sensors: a digital map, a GPS and a video camera. This paper presents experiments where the proposed formalism is used to answer queries about driving directions, lanes and vehicles.

- Encoding spatial domains with relational Bayesian networks (FENELON et al., 2010).

This paper investigates an encoding of a probabilistic TBox using relational Bayesian networks that are specified through a probabilistic description logic. The probabilistic description logic extends the popular $\mathcal{ALC}$ language; on top of this language we add a few operations that are needed to represent the cardinal direction calculus. Using such resources we model roads containing lanes, and vehicles containing digital maps, GPS and video cameras.

- Reasoning about shadows in a mobile robot environment (FENELON et al., 2012).

This paper describes a logic-based formalism for qualitative spatial reasoning with cast shadows (Perceptual Qualitative Relations on Shadows, or PQRS) and presents results of a mobile robot qualitative self-localisation experiment using this formalism. Shadow detection was accomplished by mapping the images from the robot's monocular colour camera into a HSV colour space and then thresholding on V. We present results of self-localisation using two methods for obtaining the threshold automatically: in one method the images are segmented according to their grey-scale histograms, in the other, the threshold is set according to a prediction about the robot's location, given a shadow-based map defined using a qualitative spatial reasoning theory. This map-related threshold

search is the main contribution of the present work, and to the best of our knowledge this is the first work that uses qualitative spatial representations both to perform egolocation and to calibrate a robot's interpretation of its perceptual input.

- Reconhecimento de Objetos Baseado em Contexto Utilizando a Lógica de Descrição Probabilística (FENELON-PEREIRA; SANTOS; COZMAN, 2012).

Agentes robóticos ainda não têm a habilidade de raciocinar sobre objetos em cenas reais. Neste artigo propomos uma abordagem em que o contexto de objetos é modelado por meio de uma lógica de descrição probabilística (CRACC), e onde os dados extraídos por visão computacional são fontes de informação para inferências sobre objetos. Em nossos experimentos usamos o algoritmo SURF (*Speed Up Robust Features*) para obter descritores de objetos em imagens capturadas de dois ambientes distintos. Os resultados apresentam um aumento nas respostas positivas devido à modelagem de contexto por meio de CRACC.

- A Qualitative-Probabilistic Approach to Autonomous Mobile Robot Self Localisation and Self Vision Calibration (FENELON et al., 2013)

Typically, the spatial features of a robot's environment are specified using metric coordinates, and well-known mobile robot localisation techniques are used to track the exact robot position. In this paper, a qualitative-probabilistic approach is proposed to address the problem of mobile robot localisation. This approach combines a recently proposed logic theory called Perceptual Qualitative Reasoning about Shadows (PQRS) with a Bayesian filter. The approach herein proposed was systematically evaluated through experiments using a mobile robot in a real environment, where the sequential prediction and measurement steps of the Bayesian filter are used to both self-localisation and self-calibration of the robot's vision system from the observation of object's and their shadows. The results demonstrate that the qualitative-probabilistic approach effectively improves the accuracy of robot localisation, keeping the vision system well calibrated so that shadows can be properly detected.