

CENTRO UNIVERSITÁRIO DA FEI  
VALQUIRIA FENELON PEREIRA

**AUTOLOCALIZAÇÃO QUALITATIVA PARA UM ROBÔ MÓVEL UTILIZANDO  
INFORMAÇÃO DE SOMBRAS**

**São Bernardo do Campo, SP**

**2009**

VALQUIRIA FENELON PEREIRA

**AUTOLOCALIZAÇÃO QUALITATIVA PARA UM ROBÔ MÓVEL UTILIZANDO  
INFORMAÇÃO DAS SOMBRAS**

Dissertação de Mestrado apresentada ao Centro Uni-  
versitário da FEI para obtenção do título de Mestre  
em Engenharia Elétrica, orientada pelo Prof. Dr.  
Paulo Eduardo Santos

São Bernardo do Campo, SP  
2009

Pereira, Valquíria Fenelon

Autocalização qualitativa para um robô móvel utilizando  
informação de sombras Valquíria Fenelon Pereira. São Bernardo do  
Campo, 2009.

108 f. : il.

Trabalho de Conclusão de Curso - Centro Universitário da FEI.  
Orientador: Prof. Paulo Eduardo Santos

1. Raciocínio Espacial Qualitativo. 2. Robô Móvel. 3. Percepção  
Robótica. 4. Sombras. I. Santos, Paulo Eduardo, orient. II. Título.

CDU 658.52.011.56

Centro Universitário da FEI  
**(Original impresso assinado)**  
Candidata: Valquiria Fenelon Pereira

Dissertação: Autolocalização qualitativa para um robô móvel utilizando informação de  
sombras

A comissão Examinadora da Defesa da Dissertação de Mestrado , em sessão pública realizada  
06 de abril de 2009, considerou a candidata  
Aprovada  Reprovado  
para o obtenção do título de Mestre em Engenharia Elétrica

---

Prof. Dr. Paulo Eduardo Santos

---

Prof. Dr. Flávio Soares - USP

---

Prof. Dr. Márcio Rillo - FEI

São Bernardo do Campo, SP  
2009

*Aos meus pais, filho, e companheiro.  
A todos que me incentivaram a seguir  
o meus sonhos*

# Agradecimentos

Ao meu orientador Paulo Eduardo Santos, por não ter me deixado chorar. Aos Professores de mestrado, meu carinho especial a todos que me ajudaram muito além do que eu esperava. Por nem sempre facilitarem as coisas, fazendo me descobrir mais capaz do que imaginava. *No pain, no gain*

Aos colegas Junior, Murilo, Adeline, Eduardo, André por terem me ajudado com dicas, esclarecimento de dúvidas, revisões, com sugestões pertinentes, e cedendo materiais e ferramentas, afinal me dando uma mão quando necessário. A todos meus amigos de mestrado, principalmente Almanegra, Danilo e Giuliano que me ajudaram com o Opencv, Latex e retroprojetores. E aos demais colegas de pós-graduação, pelas críticas e sugestões. Aos amigos que sempre torcem por mim, essa torcida sempre me deu forças.

A minha família pelo apoio e incentivo para chegar até aqui e continuar seguindo em frente.

A instituição FEI e seus funcionários pelo o apoio com material e infraestrutura necessários para realização dos testes, e em especial ao pessoal do CLE que me ajudaram a consertar os equipamentos que deram problema.

À CAPES, pelo apoio financeiro.

Descendo ao Pireu, Platão para num canto do caminho e põe-se de costas para o Sol. A seus pés, Skia, sua sombra, se desenha com particular precisão. O vento revolve as copas das oliveiras.

SKIA (desesperada): Encerrada numa caverna. Usada como exemplo do conhecimento inferior. Apontada pelos séculos como espantinho da filosofia. Insisto: eu o acompanho o dia inteiro, do nascer ao pôr-do-sol, e você não para de me pisotear. Deveria me pedir desculpas.

PLATÃO: Mas que impertinência! Você também viu que, além de ser efêmera e obscura, você contém um ninho de contradições, cria confusão e espanto e deixa perplexos adultos e crianças. Parece-me que sua situação só piorou.

SKIA: Mas é exatamente esse o problema! Quero mostrar a você que, embora cause espanto e apesar de os humanos não saberem direito o que pensar de mim, posso ser útil a todos — inclusive aos cientistas e a filósofos como você.

PLATÃO: Essa é boa! Não sei para que você seria capaz de servir.

SKIA: Só para começar, se não fosse eu, não haveria alternância entre o dia e a noite, você não veria a forma das coisas, tudo pareceria sem relevo nem substância...

PLATÃO (exasperado): Concordo, mas você não passa de uma comparsa: quem faz o trabalho todo é a luz.

SKIA: Protesto! A luz não sabe fazer outra coisa senão percorrer o seu caminho. Essa desmiolada sai por aí sem pensar, quando encontra um corpo tropeça e sai andando por outro lugar. *Eu* é que conservo o indício desse encontro. A sombra é a memória da luz. Então, não está convencido? Ainda tenho muitas balas na agulha. Leia e verá. *Casati* (2001).

# Resumo

O raciocínio espacial qualitativo faz parte do senso comum humano. Percebemos o ambiente por nossos sentidos e o principal deles é a visão. Então é a partir do que vemos aliado ao nosso raciocínio espacial que somos capazes de lidar com nosso mundo. Também conseguimos interpretar esse mundo tridimensional em representações bidimensionais. Há tempos artistas notaram que a sombra tem um papel relevante na nossa percepção de formas, profundidade e movimento, e mais recentemente os psicólogos procuraram entender essa influência. Então a sombra, que muitas vezes é considerada um ruído, é também fonte de informações. Com isso, uma nova teoria surgiu, aliando raciocínio espacial qualitativo à informações sobre sombras, nesta teoria é possível determinar a posição do observador a partir da observação da sombra. Um robô que percebe o mundo somente através de imagens capturadas de uma única câmera utilizando um raciocínio sobre sombra, consegue obter resposta sobre sua localização relativa, dando o seu posicionamento qualitativo relacionado à fonte de luz e ao objeto que provocou a sombra. A visão computacional desempenha papel importante neste trabalho, contudo nos concentramos em desenvolver um sistema de raciocínio espacial qualitativo explorando a capacidade de utilizar sombra como fonte de informação para localização e profundidade. Foi encontrada também uma solução para corresponder sombra com o objeto que a provocou. Os resultados confirmaram que sombra pode ser fonte de informação útil para um robô móvel.

**Palavras-chave:** Raciocínio Espacial Qualitativo, Robô Móvel, Percepção Robótica, Sombras, Robótica Cognitiva, Raciocínio Perceptual Qualitativo sobre Sombras.

# Abstract

Qualitative spatial reasoning is part of the human Commonsense. We use our perception to sense the environment. Vision is the human main sensor. What we see, allied with our spatial reasoning, is what enable us for dealing with our world. Also, we interpret the three-dimensional world from two-dimensional scenes. Artists for long have discovered that shadows play an important role in our perception of shapes, depth, and motion. Psychologists recently started to be interested in understanding this influence. Although shadows are largely considered as noise by vision researchers, it carries information. Therewith, a new theory appears that associates a qualitative spatial reasoning framework with information about shadows. In this theory it is possible to reason about the observer's position from shadows of objects. A robot that perceives the world only through captured images from a single camera, by using reasoning about shadows it succeeds answering queries about its relative location, returning its qualitative position relative to a light source and the shadow's caster. Computer vision plays an important role in this work. However we focus on developing a qualitative spatial reasoning system that explores the information content of shadows for robot self-localization and depth perception. A partial solution to the shadow correspondence problem was also developed. The results suggest that shadow can be a useful source of information for a mobile robot.

**Keywords:** Qualitative Spatial Reasoning, Mobile Robot, Robotic Perception, Shadows, Cognitive Robotic, Perceptual Qualitative Reasoning about Shadows.

# Lista de Figuras

1.1	(a) <i>The spies of Jericho escape</i> de Belbello da Paiva, (b) rascunho. Veja a sombra no muro esquerdo, ela não apareceria deste modo em numa cena real, contudo não temos dificuldade de identificar qual objeto a provocou. . . . .	18
2.1	Robô <i>CiceRobot</i> com a representação do ambiente e sua própria representação em 3D. . . . .	22
2.2	Diagrama conceitual de vizinhança (CND) 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. . . . .	26
2.3	Representação de $R_1 = NTPP$ e $R_2 = TPP$ , note que a relação $R_3 = NTPP$ é a única relação possível neste caso (Tabela 2.2). . . . .	28
2.4	Dois sequências de imagens (da esquerda para direita) onde um cone passa em frente do cilindro(sequência superior) e por trás(sequência inferior), este é um exemplo da oclusão em ação. . . . .	28
2.5	Mudança na percepção da oclusão causada pelo movimento do observador (paralaxe). . . . .	29
2.6	Modelo gráfico das 20 relações básicas do cálculo sobre oclusão de região, o ROC-20. . . . .	30
2.7	Diagrama conceitual de vizinhança para ROC-6. . . . .	32
2.8	Objetos parcialmente oclusos, na imagem suas regiões estão sobrepostas parcialmente. . . . .	34
2.9	Sombras nos informam sobre a posição relativa dos objetos. Na imagem à esquerda não podemos determinar qual a posição do robô no entanto nas outras três imagens percebe-se que ele está cada vez mais longe do chão. . . . .	35
2.10	A sombra informa sobre a forma do objeto. Nesta figura podemos notar que o robô tem o braço ocluso e segura uma chaleira e um arco em duas das imagens. . . . .	36
2.11	A sombra informa sobre a superfície em que foi projetada. Na imagem da esquerda o chão é uma superfície ambígua, enquanto que a direita a sombra revela a geometria da superfície em que está projetada. . . . .	36
2.12	Estudos sobre sombra de Leonardo da Vinci. No esboço vê-se representados as sombras do objeto projetada em outra superfície (sombra-projetada) e em si próprio (sombra-própria) . . . . .	36
2.13	Movimento da sombra pode ser ocasionado pelo deslocamento da fonte de luz ( $A$ ), da superfície ( $B$ ), do objeto ( $C$ ) e do observador ( $D$ ). . . . .	37
2.14	Efeito na localização do objeto ocasionado pela posição da sombra. . . . .	38
2.15	Bola na caixa. Mesmo na mesma localização da bola em (b) e (c) percebemos posicionamentos diferentes devido a sombra. . . . .	39
2.16	As imagens do Taj Mahal foram tiradas em horários diferentes do dia, as fotos da esquerda para a direita são do amanhecer, meio do dia, e entardecer. O próprio monumento assume um tom diferentes em cada uma delas. . . . .	40
2.17	Representação gráfica do cubo de cores RGB, o cone da esquerda representa como uma cor(80,200,130) está localizada neste sistema cartesiano. . . . .	41
2.18	Representação gráfica do cone HSV. . . . .	42
2.19	Gráfico representando o filtro de limiar e o limiar $T$ . . . . .	43

2.20	Representação gráfica da erosão e dilatação utilizando uma matriz 3x3 como elemento estruturante <b>E</b> . . . . .	44
2.21	Representação gráfica da erosão e dilatação utilizando uma matriz 3x3 como elemento estruturante. . . . .	45
2.22	Aplicação do algoritmo de detecção de objetos em movimento. A esquerda nota-se uma detecção errada dos objetos, enquanto que a direita a localização é correta devido a supressão das sombras. . . . .	46
3.1	Representação gráfica das relações básicas do PQRS. . . . .	50
3.2	Imagens capturadas pelo <i>PeopleBot<sup>TM</sup></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). . . . .	51
3.3	(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
3.4	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 . . . . .	55
3.5	Imagens capturadas no laboratório, onde podemos aplicar o axioma de correspondência de sombra, note somente a sombra do objeto está externamente conectada a sua base e mais distante da fonte de luz. . . . .	59
3.6	Nesta figura podemos inferir que há outro objeto ocluído totalmente pelo balde. . . . .	59
4.1	Fluxograma representando o procedimento dos experimentos realizados com o <i>PeopleBot<sup>TM</sup></i> . . . . .	62
4.2	Fluxograma de detecção de sombra e objeto . . . . .	63
4.3	Fluxograma de correspondência entre objeto e sua sombra . . . . .	64
4.4	Fluxograma de autolocalização . . . . .	65
4.5	Fluxograma de proximidade . . . . .	66
4.6	Laboratório onde os experimentos preliminares foram realizados. . . . .	66
4.7	Representação da área em que o robô atuará neste trabalho, sua área de atuação ou região de atuação. . . . .	67
4.8	Representação espacial (a) das regiões considerando um ambiente real, e (b) vista superior de uma representação das regiões no limite da área de atuação do robô. . . . .	68
4.9	Imagens capturadas com o robô posicionado na região 1 à (a) esquerda, (b) direita, e (c) em frente do objeto. Note que o objeto não oclui a sombra. . . . .	69
4.10	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). Robô posicionado (a) à direita e (b) à esquerda do objeto . . . . .	70
4.11	Imagens capturadas com o robô na região 3. A sombra nesta região aparece com um dos lados totalmente ocluso pelo objeto. Existe uma conexão entre a parte superior do objeto e sua sombra projetada. Robô (a) à direita e (b) à esquerda do objeto. . . . .	71
4.12	Imagens capturadas com o robô na região 4. A sombra nesta região quase desaparece. Contudo ainda existe uma conexão entre a parte superior do objeto e a sombra. Robô (a) à direita e (b) à esquerda do objeto. . . . .	71
4.13	Imagem capturada com o robô na região 5. A sombra não aparece na imagem por estar totalmente oclusa. O observador está na região equivalente ao ponto de vista da luz. . . . .	72
4.14	Janela de controle do software de raciocínio <i>PeopleBot<sup>TM</sup></i> . . . . .	73
4.15	Janelas para calibração de limiar de corte para segmentar objeto e sombra, da esquerda para direita, corta altas intensidades, corta acima do limiar indicado fica objetos escuros, corta abaixo do limiar indicado permanece sombras. . . . .	73
4.16	Pseudocódigo do algoritmo principal utilizado no <i>PeopleBot<sup>TM</sup></i> . . . . .	74

4.17	Imagem do <i>PeopleBot<sup>TM</sup></i> mostrando o resultados encontrados, no canto superior esquerdo a região de posição do robô, retângulos indicando sombras e objetos, os números iguais nos retângulos demonstram a correspondência, a letra R e L indicam o objeto mais próximo do robô e da fonte da luz respectivamente. . . . .	75
4.18	Pseudocódigo do algoritmo de detecção de objeto e sombra utilizado no <i>PeopleBot<sup>TM</sup></i> . . . .	76
4.19	Exemplo de sequência da segmentação de um quadro na detecção ROI:(a)canal <b>V</b> , (b)canal <b>V</b> após limiar acima $\alpha$ , (c)filtro morfológico, (d)retângulo verde indica regiões candidatas a ROI, (e)canal <b>S</b> após limiar acima $\alpha_{sat}$ , (f)máscara ROI. . . . .	77
4.20	Exemplo de sequência da segmentação de um quadro na detecção de objetos:(a)Imagem objeto e sombra após 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)retângulo amarelo indica região com altas saturação, (f)máscara ROI. . . . .	79
4.21	Exemplo de sequência da segmentação de um quadro na detecção de sombras:(a)Imagem com objeto após aplicada a segunda máscara, (b)Imagem após filtros de limiar acima e morfológico de abertura, (c) retângulo vermelho indicando a sombra localizada. . . . .	80
4.22	Pseudocódigo do algoritmo de correspondência. . . . .	82
4.23	Imagem capturada com <i>PeopleBot<sup>TM</sup></i> . Com o objeto cilíndrico com duas sombras conectadas a ele, a qual o algoritmo de correspondência atual consegue resolver e o anterior não. . . . .	83
4.24	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 <i>EC</i> é verdade. . . . .	84
4.25	Pseudocódigo do algoritmo de localização do robô . . . . .	85
4.26	Pseudocódigo do algoritmo de proximidade . . . . .	86
5.1	Representação de um robô na região 3. . . . .	90
5.2	Imagem indicando a correspondência entre o objeto e a sombra; o objeto mais próximo da luz e do robô está marcado por L e R e a localização do robô é a região 2 (conforme marcado no canto superior esquerdo da imagem). . . . .	91

# Lista de Tabelas

2.1	Interpretações de algumas relações do RCC. . . . .	26
2.2	Tabela de composição para o RCC-8 com as possibilidades de relações $R_3(a, c)$ a partir das relações $R_1(a, b)$ e $R_2(b, c)$ . Fonte <i>Randell and Witkowski (2002)</i> . . . . .	27
5.1	Total de poses capturadas e analisadas no experimento preliminar. . . . .	93
5.2	Dados capturados nos experimentos preliminares. Colunas são as regiões reais e linhas as respostas dada pelo robô. . . . .	93
5.3	Resultados dos experimentos <i>on-line</i> de auto-localização, com calibração feita no momento do teste. . . . .	94
5.4	Resultados dos experimentos <i>off-line</i> de auto-localização com calibração $\alpha = 170$ . . . . .	94
5.5	Resultados dos experimentos <i>off-line</i> de auto-localização com calibração $\alpha = 140$ . . . . .	95
5.6	Resultados de poses analisadas pelo algoritmo de autolocalização onde pelo menos uma resposta correta foi encontrada entre os resultados das tabelas 5.3, 5.4 e 5.5, acertos pela (equação 5.1) . . . . .	95
5.7	Resultados dos experimentos <i>off-line</i> de auto-localização com calibração $\alpha = 120$ . . . . .	96
5.8	Acertos em porcentagem encontrados nas tabelas 5.3, 5.4 e 5.5 . . . . .	96
5.9	Resultados do teste de autolocalização realizado com imagens com mais de um objeto . . . . .	96
5.10	Tabela com resultados de proximidade com robô . . . . .	97
5.11	Tabela com resultados de correspondência, a coluna acerto somente considera como certo as imagens na qual todos os pares objeto-sombra foram corretamente identificados . . . . .	98

# Lista de Símbolos e Abreviaturas

$2D$	— bidimensional, com duas dimensões.
$3D$	— tridimensional, com três dimensões.
$\theta$	- Ângulo visual formado pela distância entre o objeto e sua sombra do ponto de vista de um observador.
$\nu$	— Ponto de vista de um observador.
$\Phi$	— Representa as relações de oclusão.
$\Psi$	— Representa as relações do RCC.
<i>ARIA</i>	— <i>ActivMedia Robotics Interface for Applications</i> ou Interface robótica <i>ActivMedia</i> para aplicações.
$D_k$	— Diferença da luminância de uma imagem do instante $k$ .
<i>CND</i>	— <i>Conceptual Neighbourhood Diagram</i> ou diagrama conceitual de vizinhança.
<i>HSV</i>	— <i>Hue, Saturation, Value</i> ou espaço de cores com três canais, matiz, saturação e valor.
$L$	— Constante que representa a fonte de luz.
$o$	— Objeto.
<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>RCC</i>	— <i>Region Connection Calculus</i> ou cálculo sobre conexão de região.
<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.
$s$	— Sombra.
$s_k$	— Luminância dos pontos na imagem do instante $k$ .
<i>Scr</i>	— Constante que representa a superfície onde uma sombra pode ser projetada.

# Sumário

<b>LISTA DE FIGURAS</b>	<b>9</b>
<b>LISTAS DE TABELAS</b>	<b>12</b>
<b>LISTA DE SÍMBOLOS E ABREVIATURAS</b>	<b>13</b>
<b>1 INTRODUÇÃO</b>	<b>16</b>
<b>2 REVISÃO DA LITERATURA</b>	<b>21</b>
2.1 <b>Percepção robótica</b> . . . . .	21
2.2 <b>Interpretação e bom senso</b> . . . . .	23
2.3 <b>Raciocínio Espacial Qualitativo</b> . . . . .	24
2.3.1 Cálculo sobre conexão de região . . . . .	25
2.3.2 Cálculo sobre oclusão de região . . . . .	28
2.4 <b>Percepção de sombras</b> . . . . .	35
2.5 <b>Visão computacional</b> . . . . .	38
2.5.1 Espaço de cores em imagens digitais . . . . .	40
2.5.2 Modelos de cores RGB e HSV . . . . .	40
2.5.3 Métodos para análise de imagens . . . . .	42
2.6 <b>Segmentação de sombras</b> . . . . .	45
2.7 <b>Discussão final</b> . . . . .	47
<b>3 RELAÇÕES PERCEPTUAIS QUALITATIVAS SOBRE SOMBRAS</b>	<b>48</b>
3.1 <b>Formalização das relações perceptuais qualitativas sobre sombras</b> . . . . .	48
3.2 <b>Localização relativa</b> . . . . .	52
3.3 <b>Sombras e percepção de profundidade</b> . . . . .	54
3.4 <b>Correspondência entre objeto e sombra</b> . . . . .	56
3.5 <b>Discussão final</b> . . . . .	58
<b>4 DESENVOLVIMENTO DO SISTEMA</b>	<b>61</b>
4.1 <b>Fluxo de dados do sistema</b> . . . . .	61
4.1.1 Fluxograma de visão computacional . . . . .	62

---

4.1.2	Fluxograma de correspondência . . . . .	63
4.1.3	Fluxograma da autolocalização . . . . .	64
4.1.4	Fluxograma de proximidade . . . . .	64
4.2	<b>Descrição do ambiente . . . . .</b>	<b>65</b>
4.3	<b>Definindo as regiões . . . . .</b>	<b>68</b>
4.4	<b>Funcionamento do software de autolocalização . . . . .</b>	<b>72</b>
4.4.1	Algoritmo de detecção de sombra e objeto. . . . .	74
4.4.2	Algoritmo para análise de correspondência . . . . .	81
4.4.3	Algoritmos de autolocalização e de proximidade . . . . .	85
4.5	<b>Discussão final . . . . .</b>	<b>87</b>
<b>5</b>	<b>PROCEDIMENTO EXPERIMENTAL E RESULTADOS</b>	<b>89</b>
5.1	<b>Método de validação . . . . .</b>	<b>89</b>
5.2	<b>Testes . . . . .</b>	<b>90</b>
5.3	<b>Resultados . . . . .</b>	<b>92</b>
5.3.1	Resultados preliminares . . . . .	93
5.3.2	Resultados de auto-localização . . . . .	94
5.3.3	Resultados de proximidade . . . . .	97
5.3.4	Resultados de correspondência . . . . .	97
5.4	<b>Discussão final . . . . .</b>	<b>98</b>
<b>6</b>	<b>CONCLUSÃO</b>	<b>100</b>
	<b>GLOSSÁRIO</b>	<b>107</b>
<b>A</b>	<b>TRABALHOS PUBLICADOS</b>	<b>108</b>

# Capítulo 1

## INTRODUÇÃO

A atual geração de robôs autônomos mostra um controle de movimento impressionante. Contudo sua programação ainda apresenta uma capacidade limitada para perceber e raciocinar sobre novos ambientes. Pesquisas tem sido feitas para que uma nova geração de robôs autônomos inteligentes sejam capazes de raciocinar e agir em qualquer ambiente, inclusive havendo interação com humanos (*Chella et al., 2005*).

As interações de um robô móvel com seu ambiente é o resultado de três componentes fundamentais: a tarefa programada para ser executada, o hardware do robô, e o próprio ambiente (*Nezmozow, 2006*). Do ambiente estaremos analisando um elemento pouco utilizado como fonte de informação: a sombra. Dos sensores acoplados no robô utilizado neste trabalho, somente a câmera é capaz de perceber uma sombra. A tarefa programada é interpretar os dados percebidos com aplicação de um raciocínio espacial qualitativo.

### **Objetivo**

O objetivo principal é desenvolver um sistema que utilize uma teoria sobre percepção de sombras, baseada em raciocínio espacial qualitativo, para que um agente robótico execute a tarefa de autolocalização. Esse robô móvel é equipado com uma câmera e utilizará a percepção de sombras para inferir sua localização de modo qualitativo.

### **Motivação**

A motivação surgiu na observação dos resultados de estudos sobre as influências causadas pelas sombras na percepção humana (*Mamassian et al., 1998*). Estes estudos psicofísicos sobre a percepção visual demonstraram que, ao olhar para cenas num espaço bidimensional, as sombras nos dão a impressão de profundidade e de movimento. Os testes realizados com hu-

---

manos evidenciaram que, dentre as características que indicam profundidade e movimento em imagens bidimensionais, utilizamos preferencialmente a sombra (*Kersten et al.*, 1994). Outra característica observada foi o tempo de resposta, com presença de sombra na imagem as respostas sobre localização eram mais rápidas do que somente com a presença de outro objeto (*Ni et al.*, 2004; *Rensink and Cavanagh*, 2004). A partir dos resultados verificados nestes estudos psicofísicos, a intenção é desenvolver ferramentas que utilizem as informações contidas em sombras, empregando conceitos de algumas áreas de inteligência artificial, possibilitando um robô a comportar-se como um agente inteligente (*Patnaik*, 2005).

No entanto, para que um robô utilize as informações das sombras é necessário, em primeiro lugar, um sistema de visão computacional capaz de segmentar e detectar sombras e os elementos geradores em cena (fonte de luz, objeto, superfície). A maior parte dos trabalhos de visão computacional considera sombra como um ruído, mas existem alguns trabalhos que exploram a sombra como informação (*Fitzpatrick and Torres-Jarra*, 2004; *Balan et al.*, 2007). Analisamos na literatura em visão computacional os trabalhos relacionados à detecção de sombras para montarmos um sistema de visão capaz de segmentar sombra e objetos. Além de segmentar a sombra é preciso relacioná-la com o objeto que a provocou, para que as informações sejam utilizadas de modo correto, esse relacionamento é conhecido como problema de correspondência de sombra (*Mamassian*, 2004).

*Casati* (2007) analisa o fato de alguns artistas (Figura 1.1) utilizarem a cópia da forma do objeto como recurso para solucionar o problema de correspondência. Esse tipo de correspondência funciona bem, pois nossa percepção qualitativa consegue relacionar o objeto com sua sombra, mesmo que a sombra pintada represente uma forma mais complicada que uma sombra real (*Casati*, 2007). Essa solução poderia ser utilizada para objetos em cenas reais, mas isso somente se uma sombra sempre se mostrasse com forma similar ao objeto que a formou. Relacionar a sombra ao objeto que a gerou é um das tarefas mais difíceis de serem resolvidas em visão computacional, pois não existe garantia de que, mesmo uma busca ponto a ponto, encontrará o objeto relacionado com a sombra de modo não ambíguo e único. Isso ocorre porque existe a possibilidade de alguns objetos estarem total ou parcialmente oclusos ou mesmo fora do foco de visão, tornando a comparação ponto a ponto inviável e com custo computacional elevado (*Mamassian*, 2004).

*Casati* (2007) sugere que uma solução seria utilizar um sistema computacional com maior tolerância a informações imprecisas ou parciais, assim como ocorre com a interpretação hu-



Figura 1.1: (a) *The spies of Jericho escape* de Belbello da Paiva, (b) rascunho. Veja a sombra no muro esquerdo, ela não apareceria deste modo em numa cena real, contudo não temos dificuldade de identificar qual objeto a provocou.

Fonte: Casati (2007)

mana que utiliza algumas poucas características da sombra, sem se importar com todas as regras físicas envolvida, ao interpretar imagens bidimensionais (Cavanagh, 2005). Observando uma pintura as pessoas relacionam uma sombra com o objeto que a provocou mesmo que a sombra desenhada seja impossível de ocorrer realmente (informação imprecisa ou parcial). Como exemplo visual tem-se a Figura 1.1 onde a sombra do muro esquerdo é uma cópia da forma dos personagens que descem pelo balanço. Porém em uma cena real essa forma de sombra não seria possível de ser observada, no entanto conseguimos entender as informações relacionadas de modo qualitativo e perceber o que provocou a sombra fazendo uma correspondência entre eles (Casati, 2007). Será proposta e aplicada uma solução para o problema de correspondência (seção 3.4), utilizando uma teoria sobre as relações perceptuais qualitativas sobre sombras (do inglês *Perceptual Qualitative Relations about Shadows - PQRS*) (Santos et al., 2008, 2009)

A teoria PQRS é uma extensão da teoria do cálculo sobre oclusão de região (Randell et al., 2001; Witkowski et al., 2001) e foi apresentada pela primeira vez em (Santos et al., 2008).

No Capítulo 3 veremos a teoria PQRS com mais detalhes. O cálculo sobre oclusão de região (Randell *et al.*, 2001) é um formalismo direcionado à representação da oclusão entre objetos, suas representações de interposição de objetos com formas diversas são uma extensão do cálculo sobre conexão de região (Randell *et al.*, 1992). Tanto o cálculo sobre conexão de região quanto o cálculo sobre oclusão de região serão abordados na seção 2.3.

## Proposta

Com base na teoria sobre as relações perceptuais qualitativas sobre sombras (Santos *et al.*, 2008), a proposta deste trabalho é demonstrar que um sistema robótico pode extrair informações de um sistema visual, e que a sombra pode ser utilizada para inferir posicionamento qualitativo dos objetos e do próprio robô. As interpretações e respostas do robô serão realizadas tendo como base um raciocínio espacial qualitativo, este irá inferir novas informações baseado nos dados das imagens e no conhecimento *a priori* do ambiente.

Assume-se neste trabalho que a fonte de luz será única e estará posicionada em um ponto mais alto que o robô e o objeto. As informações de qualquer outro sensor do *PeopleBot<sup>TM</sup>*, tais como sonares e sensores de toque, não serão considerados. A proposta deste trabalho é totalmente baseada no potencial de informação da sombra, deixando em aberto para trabalhos futuros a sua utilização conjunta com outras técnicas de percepção robótica .

## Contribuição

A contribuição deste trabalho é:

1. Demonstrar que é possível inferir posicionamento espacial qualitativo a partir de imagens de sombras.
2. Aplicar a teoria PQRS (apresentada em Santos *et al.* (2008)) na criação de um algoritmo de interpretação de cenas (Capítulo 4) e a realização de testes em ambiente real para avaliar a teoria (Capítulo 5).
3. Acrescentar na teoria PQRS novos axiomas capazes de solucionar o problema de correspondência em ambientes específicos (seção 3.4).

## **Estrutura do trabalho**

O conhecimento básico para o desenvolvimento deste trabalho está concentrado na revisão bibliográfica encontrada no Capítulo 2, onde também nos posicionaremos no estado da arte sobre áreas como, percepção de sombras, raciocínio com bom senso e algumas técnicas de visão computacional. Será feita uma breve descrição de algumas ferramentas necessárias para que um robô autônomo perceba, analise, interprete e se mova pelo ambiente. Dissertaremos sobre percepção robótica (seção 2.1), introduziremos a literatura sobre interpretação lógica e raciocínio com bom senso (seção 2.2), também veremos uma teoria básica para ajudar na interpretação do mundo utilizando teorias sobre raciocínio espacial qualitativo (seção 2.3). Descreve-se na seção 2.4 como a influência da sombra age na percepção humana sobre localização espacial em figuras 2D. Na seção 2.5 serão descritos alguns métodos de visão computacional utilizados como base para auxiliar o desenvolvimento deste trabalho.

A teoria das relações perceptuais qualitativas sobre sombras (PQRS) (*Santos et al.*, 2008) será apresentada no Capítulo 3, sendo ela a base teórica deste trabalho. Na seção 3.4 é apresentada uma extensão da PQRS, com axiomas que definem a correspondência entre sombra e objeto. A definição de como será desenvolvido o sistema, a partir da PQRS, é realizada no Capítulo 4 e ainda neste capítulo haverá uma descrição do ambiente e dos elementos utilizados nos experimentos.

No Capítulo 5 se encontra a descrição de como foram realizados os testes e os resultados obtidos, seguido pela discussão sobre os resultados e conclusão final. Como anexo encontra-se a reprodução dos trabalhos publicados referentes ao PQRS .

## Capítulo 2

# REVISÃO DA LITERATURA

Ao longo deste capítulo iremos realizar um apanhado de todas as ciências envolvidas de alguma forma com o trabalho desenvolvido, começaremos pela percepção robótica.

### 2.1 Percepção robótica

O processo de perceber como está o ambiente é útil para qualquer agente, é essa percepção que o capacita a extrair características do mundo. Saber como o mundo está caracterizado e mesmo as características do próprio agente é vital para execução de ações com algum propósito, seja a ação tomada de forma reativa ou mesmo em uma abordagem baseada em um conhecimento *a priori*, agentes artificiais percebem o mundo através de sensores (*Russel and Norvig, 2004*).

Existem vários tipos de sensores, tais como sensores de toque, sonares, radares, sensores de detecção de sons, sensores de captura de imagens etc. Dentre esses sensores, os sensores de visão, aliados com a aplicação de técnicas de processamento de imagens, podem fornecer informações complexas (*Russel and Norvig, 2004*). Contudo, a tecnologia atual ainda não permite desempenho tal como o de um ser humano na observação e interpretação do ambiente (*Patnaik, 2005*). Ainda que limitada, quando comparada a visão humana, com visão computacional somos capazes de obter vários tipos de informações. Essas informações podem ser somadas aos dados percebidos por outros tipos de sensores, esse conjunto de informações poderá alimentar um sistema de inteligência artificial, para auxiliar um robô a raciocinar sobre o estado do mundo e executar sua tarefa (*Gaskin, 2008*).

Uma tarefa primordial para um robô móvel é a navegação. Para que a navegação seja executada com eficiência, o robô deve ser capaz de se localizar em seu ambiente (*Frank Dellaert and*



Figura 2.1: Robô *CiceRobot* com a representação do ambiente e sua própria representação em 3D.  
Fonte: *Chella et al.* (2005).

*Thrun*, 1999). No entanto, há alguns problemas nesta localização. Problemas como a imprecisão com relação ao ponto e à direção inicial do robô, também existe incerteza na estimativa da posição atual, feita com base em dados capturados pelos sensores (como o hodômetro) (*Chella et al.*, 2005). Neste caso, a visão computacional pode auxiliar na atualização do posicionamento do robô no ambiente, como veremos a seguir.

Um modo de fazer uso da visão de máquina é utilizar objetos conhecidos e comparar com um mapa que o robô possua em seu banco de conhecimento. Um exemplo disso pode ser encontrado em (*Chella et al.*, 2005) que descreve o projeto do robô *CiceRobot* (Figura 2.1). Este robô executa a tarefa de guia turístico do museu arqueológico de *Agrigento*, Sicília. Marcos conhecidos pelo robô foram fixados nas paredes de modo a ficarem visíveis mesmo em presença de visitantes. Quando a autolocalização do robô é exigida, ele procura pelo marco mais próximo de sua posição atual com movimentos horizontais e verticais da sua câmera, aplicando as devidas transformações e comparando com sua localização no mapa do museu em sua base de conhecimento.

Outros exemplos de uso da percepção robótica são encontrados em (*Santos and Shanahan*, 2002), onde os dados coletados pelos sensores do robô móvel serviram de base para a construção de um sistema de raciocínio espacial, enquanto que em (*Scherl and Levesque*, 1993) esta percepção foi utilizada para produzir conhecimento a partir das ações executadas. Na esperança de desenvolver robôs autônomos, em (*Kanda et al.*, 2006) encontramos pesquisas realizadas

com robôs interagindo com pessoas, procurando a melhor maneira das máquinas ajudarem, de um modo aceitável e confiável, em tarefas humanas.

A intenção desses estudos é desenvolver melhores técnicas para capacitar um robô a executar tarefas diversas, mesmo que essas pesquisas ainda sejam feitas 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. Mesmo assim, alguns pesquisadores acreditam que poderemos elevar o nível cognitivo robótico agregando os vários algoritmos desenvolvidos para efetuar diversas tarefas (McCarthy, 2002).

Um desafio atual 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 *et al.*, 2003). Um robô móvel autônomo trabalha continuamente com percepção e ação (Nezmozow, 2006), 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. O robô deve decidir como agir dependendo do que percebe (Pirri *et al.*, 2003). Uma interpretação dos fatos com bom senso auxiliaria na tomada de decisão de qual ação executar (Pirri *et al.*, 2003). Interpretação e raciocínio com senso comum serão descritos na próxima seção.

## 2.2 Interpretação e bom senso

A filosofia pode contribuir com uma formalização do bom senso e raciocínio, esse formalismo vem sendo desenvolvido utilizando lógica de primeira ordem (McCarthy, 1990). O raciocínio com senso comum ou bom senso é a capacidade humana de interpretar e concluir qual será o resultado da ocorrência de eventos cotidianos. A automatização de um raciocínio com bom senso é um dos objetivos das pesquisas em inteligência artificial (Mueller, 2007; McCarthy, 1959). Desde o começo desta ciência, muitas abordagens procuravam uma lógica capaz de introduzir o bom senso nas decisões de um robô ou sistema. Programas que tiravam conclusões de uma lista de premissas foram primeiro sugeridos em (McCarthy, 1959) e eram chamados de *advice takers*. Essa automatização do pensamento põe em prática a noção de que o raciocínio humano sobre novas circunstâncias vem da analogia com situações que, por serem similares, são suas premissas. De fato, a lógica foi a teoria que abordou os problemas de

raciocínio com bom senso de modo bem focalizado, obtendo um alto grau de sucesso (*Mueller, 2007*). Uns dos formalismos baseados em lógica, que leva a compreender e facilitar a utilização do raciocínio com bom senso é o cálculo de situação (*Brachman and Levesque, 2004*) e o cálculo de eventos (*Muller, 1998; Shanahan, 1997*).

O cálculo de situação (*Brachman and Levesque, 2004*) é uma teoria em lógica de primeira ordem de tipos variados. Ela representa as mudanças ocorridas no mundo por causa de uma ação. O cálculo de situação contém axiomas que especificam o estado inicial dos fluentes, e para cada fluente um axioma de estado sucessor. Um fluente representa predicados ou funções cujos valores podem mudar de situação para situação, enquanto que uma situação significa o estado do mundo resultado de uma sequência de ações. Com o cálculo de situação, a partir de uma situação inicial (onde as ações ainda não ocorreram), infere-se uma sequência de ações que gere a situação final desejada (*Brachman and Levesque, 2004*).

O cálculo de eventos (*Muller, 1998; Shanahan, 1997*) é outro formalismo em lógica de primeira ordem para o raciocínio sobre ações e efeitos. Suas primitivas ontológicas consistem em eventos (ou ações), fluentes e instantes de tempo (*Pereira, 2002*). Diferente do cálculo de situação ele trabalha com tempo. O cálculo de evento é um mecanismo lógico que infere o que é verdade dada uma narrativa dos eventos e uma descrição dos efeitos das ações (*Shanahan, 1999*). Tanto no cálculo de eventos quanto no cálculo de situação existe a necessidade do agente ter alguma percepção espacial do ambiente no qual agirá. Para um robô móvel a navegação é uma tarefa importante, e para execução de tal tarefa de modo satisfatório, a percepção espacial torna-se imprescindível (*Nezmozow, 2006*). Vamos falar a seguir sobre o Raciocínio Espacial Qualitativo que é um campo da Inteligência Artificial que engloba teorias que representam e possuem métodos de raciocínios sobre os elementos espaciais, utilizando uma linguagem formal (*Stock, 1997*).

## 2.3 Raciocínio Espacial Qualitativo

O Raciocínio Espacial Qualitativo procura formalizar o conhecimento espacial, baseado em relações primitivas definidas sobre entidades espaciais elementares (*Cabalar and Santos, 2006*). As teorias de raciocínio espacial utilizam muitos modelos espaciais que, com primitivas ontológicas e métodos adequados de raciocínio, capacitam um computador a prever, explicar e diagnosticar comportamentos de um sistema físico de modo qualitativo sempre que uma avaliação quantitativa for intratável ou não avaliável (*Stock, 1997*). Uma das teorias básicas mais

conhecidas do raciocínio espacial qualitativo é o *Region Connection Calculus* (RCC) (Randell et al., 1992) que será descrito a seguir.

### 2.3.1 Cálculo sobre conexão de região

*Region Connection Calculus* (RCC) (Randell et al., 1992) ou cálculo sobre conexão de região é um sistema axiomático sobre regiões espaciais em lógica de primeira ordem, que tem como base a relação primitiva binária de conectividade ( $C/2$ ). Para um melhor entendimento, dado duas regiões  $x$  e  $y$ , a relação  $C(x, y)$ , que representa “ $x$  está conectado com  $y$ ”, é verdadeira se somente se  $x$  e  $y$  tem pelo menos um ponto em comum. A relação de conectividade obedece aos seguintes axiomas (Randell et al., 1992; Gotts, 1994);

$$\begin{aligned} &\forall x C(x, x); \\ &\forall x y C(x, y) \rightarrow C(y, x); \\ &\forall x y z [C(z, x) \leftrightarrow C(z, y)] \rightarrow x = y. \end{aligned}$$

Assumindo a relação binária  $C/2$  e variáveis  $x, y$  e  $z$  para regiões espaciais, pode-se definir um conjunto de relações básicas entre pares utilizando  $C(x, y)$ . Algumas dessas relações são definidas a seguir e suas interpretações podem ser vistas na Tabela 2.1 (Randell et al., 1992).

$$\begin{aligned} DC(x, y) &\equiv \neg C(x, y) \\ P(x, y) &\equiv \forall z [C(z, x) \rightarrow C(z, y)] \\ PP(x, y) &\equiv P(x, y) \wedge \neg P(y, x) \\ EQ(x, y) &\equiv P(x, y) \wedge P(y, x) \\ O(x, y) &\equiv \exists z [P(z, x) \wedge P(z, y)] \\ PO(x, y) &\equiv O(x, y) \wedge \neg P(x, y) \wedge \neg P(y, x) \\ DR(x, y) &\equiv \neg O(x, y) \\ TPP(x, y) &\equiv PP(x, y) \wedge \exists z [EC(z, x) \wedge EC(z, y)] \\ NTPP(x, y) &\equiv PP(x, y) \wedge \neg \exists z [EC(z, x) \wedge EC(z, y)] \end{aligned}$$

As representações de relações inversas são  $Pi$ ,  $PPi$ ,  $TPPi$ , e  $NTPPi$ . As relações  $DC(x, y)$ ,  $EQ(x, y)$ ,  $PO(x, y)$ ,  $EC(x, y)$ ,  $TPP(x, y)$ ,  $NTPP(x, y)$ ,  $TPPi/2$  e  $NTPPi/2$  formam um conjunto de relações referido como RCC-8 (Stock, 1997). No contexto do raciocínio espacial qualitativo assume-se que as mudanças ocorrem de forma contínua. Todas as transições topológicas entre regiões são contínuas, ou seja, quando uma relação muda, essa mudança segue

Tabela 2.1: Interpretações de algumas relações do RCC.

Relação	Interpretação
$DC(x, y)$	' $x$ está desconectado de $y$ '
$P(x, y)$	' $x$ é parte de $y$ '
$PP(x, y)$	' $x$ é parte própria de $y$ '
$EQ(x, y)$	' $x = y$ ': ' $x$ é idêntico a $y$ '
$O(x, y)$	' $x$ sobrepõe $y$ '
$PO(x, y)$	' $x$ sobrepõe parcialmente $y$ '
$DR(x, y)$	' $x$ é distinta de $y$ '
$TPP(x, y)$	' $x$ é uma parte própria tangencial de $y$ '
$NTPP(x, y)$	' $x$ é uma parte própria não tangencial de $y$ '

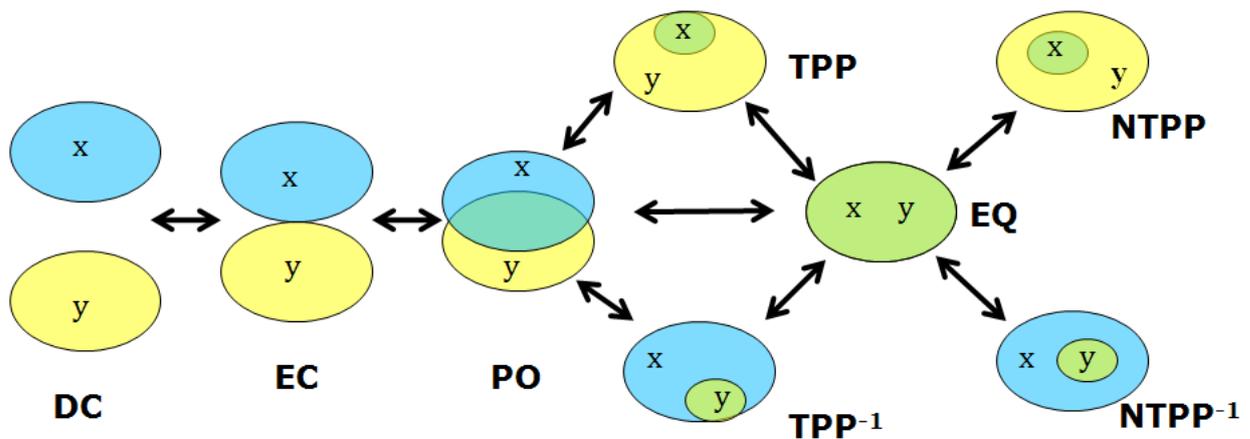


Figura 2.2: Diagrama conceitual de vizinhança (CND) das relações RCC-8. Note a continuidade das transições topológicas entre regiões (indicadas pelas setas), as mudanças de relações ocorrem sempre para uma relação vizinha.

Fonte: adaptada de (Cui *et al.*, 1992).

sempre para uma relação próxima, como é possível perceber as transições das 8 relações básicas do RCC (RCC-8) no diagrama conceitual de vizinhança da Figura 2.2 (Cui *et al.*, 1992; Randell *et al.*, 1992)

O diagrama conceitual de vizinhança (CND - *Conceptual Neighborhood Diagram*) é uma representação gráfica, tanto das relações entre duas regiões quanto das transições entre estas relações. Neste contexto, o conceito de transições contínuas significa que além das relações vizinhas, não há outra que possa acontecer (Santos, 2003). A continuidade nas transições das relações é um conceito importante (Stock, 1997), fazendo com que o CND seja uma importante ferramenta para raciocinar em várias situações (Cui *et al.*, 1992).

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 and Witkowski, 2002). Com ela tem-se um conjunto das relações possíveis entre duas regiões  $a$  e  $c$ , dadas as informações sobre as relações entre  $a$  e  $b$  e entre  $b$  e  $c$ . A Tabela 2.2 mostra a tabela

Tabela 2.2: Tabela de composição para o RCC-8 com as possibilidades de relações  $R_3(a, c)$  a partir das relações  $R_1(a, b)$  e  $R_2(b, c)$ . Fonte *Randell and Witkowski (2002)*

$R_1(a, b)$	$R_2(b, c)$							
	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	NTPPi
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

de composição para RCC-8, onde a primeira coluna representa as relações  $R_1(a, b)$  e a primeira linha  $R_2(b, c)$ . As possibilidades das relações  $R_3(a, c)$  são dadas nas células definidas pela intersecção entre a coluna  $R_1(a, b)$  e a linha  $R_2(b, c)$  (*Randell and Witkowski, 2002*). Um exemplo para visualização, na Figura 2.3 tem as três regiões representadas em que  $NTPP(a, b)$  e  $TPP(b, c)$ , na imagem vemos a relação  $NTPP(a, c)$ , essa é a única relação possível conforme indicado na Tabela 2.2.

O cálculo de conexão de região (RCC) representa as relações, entre regiões espaciais, independentes do ponto de vista do observador (*Santos, 2007*). Contudo existe uma teoria que é uma extensão do RCC (*Randell et al., 1992*) e depende do ponto de vista do observador, ela é conhecida como *Region Occlusion Calculus* (ROC) (*Randell et al., 2001*). O ROC foi

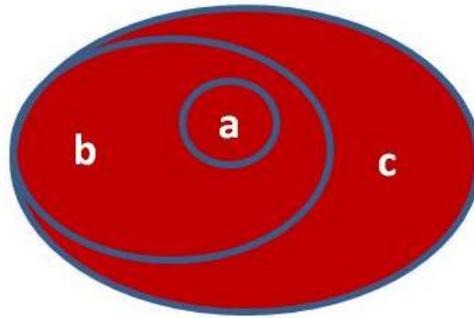


Figura 2.3: Representação de  $R_1 = NTPP$  e  $R_2 = TPP$ , note que a relação  $R_3 = NTPP$  é a única relação possível neste caso (Tabela 2.2).

proposto para representar as várias possibilidades de interposições entre quaisquer objetos. A seguir descreve-se brevemente esta teoria.

### 2.3.2 Cálculo sobre oclusão de região

O *Region Occlusion Calculus* (ROC) ou cálculo sobre oclusão de região representa as relações de oclusão espacial (Randell et al., 2001). Defini-se por oclusão espacial quando um corpo se interpõe a outro a partir de um ponto de vista. Na ocorrência de um evento de oclusão, o ROC (Randell et al., 2001) é uma teoria que auxilia a inferir o porquê um objeto não pode ser visualizado em uma sequência de imagens (Figura 2.4). Considerando, na Figura 2.4, as imagens dos objetos como regiões, quando ocorre uma oclusão as imagens dos objetos aparecem como uma única região. O ROC capacita a inferência de que existem (neste caso) dois objetos sob alguma relação de oclusão.

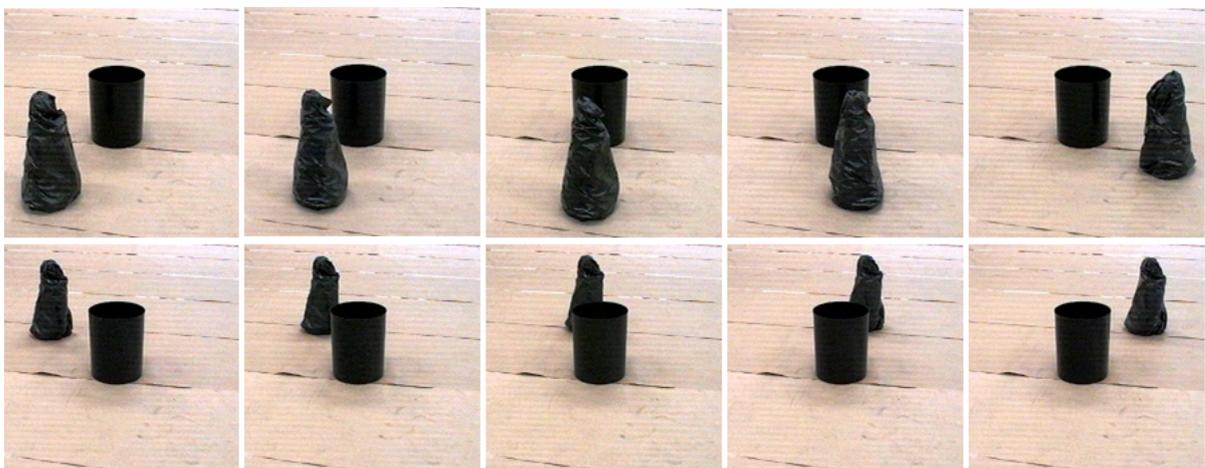


Figura 2.4: Duas seqüências de imagens (da esquerda para direita) onde um cone passa em frente do cilindro(seqüência superior) e por trás(seqüência inferior), este é um exemplo da oclusão em ação.

Vamos analisar a relação entre os objetos nas duas seqüências de cenas apresentadas na

Figura 2.4, da esquerda para a direita: pode-se inferir a informação das relações que podem ocorrer na cena seguinte considerando continuidade, assim sendo, pode-se concluir que um objeto está parcialmente ou totalmente ocluído pelo outro, nas cenas em que os objetos parecem ser uma região, pois nas anteriores existia mais de uma região, e a distância entre elas ficava menor a cada cena. Além de objetos escondidos, as transições das relações ROC podem ajudar a determinar os modelos de movimento, pois quando se muda o ponto de vista ocorre paralaxe (Figura 2.5), e ambos paralaxe e eventos de oclusão podem ser utilizados para derivar um modelo objetivo do mundo de um ponto de vista naturalmente restrito (Randell et al., 2001).

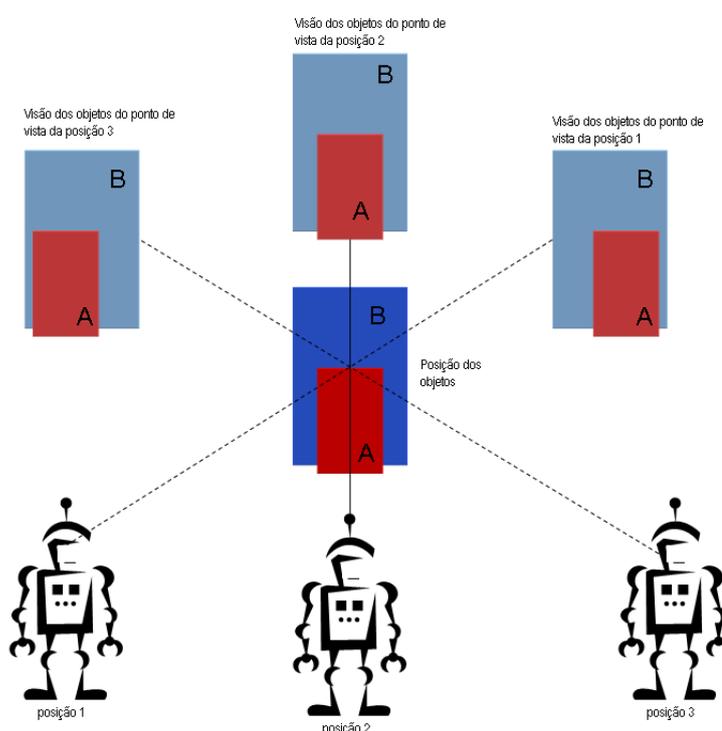


Figura 2.5: Mudança na percepção da oclusão causada pelo movimento do observador (paralaxe).  
Fonte: adaptado de (Randell et al., 2001).

Como exemplo da mudança de percepção relativa à oclusão e à paralaxe, considere dois objetos A e B no campo visual de um agente. Supondo que o agente mova-se para a esquerda, se a imagem do objeto A movimentar-se ao redor de B e este fica totalmente ocluído, o agente poderá inferir que A está na frente de B. Analogamente se o agente se movimentar para a direita e nenhuma mudança ocorrer, ele poderá inferir que possivelmente os objetos podem estar longe ou perto ou se movimentam na mesma direção que o observador. Inversamente, se A (quando visível) sempre aparecer em contraposição a B, o agente deve inferir que A e B estão fisicamente conectados (Randell et al., 2001). O cálculo sobre oclusão de região reduz a descrição de alinhamentos possíveis entre os objetos a 20 relações básicas, assim, este cálculo

é conhecido também por ROC-20 (Figura 2.6) (Witkowski *et al.*, 2001).

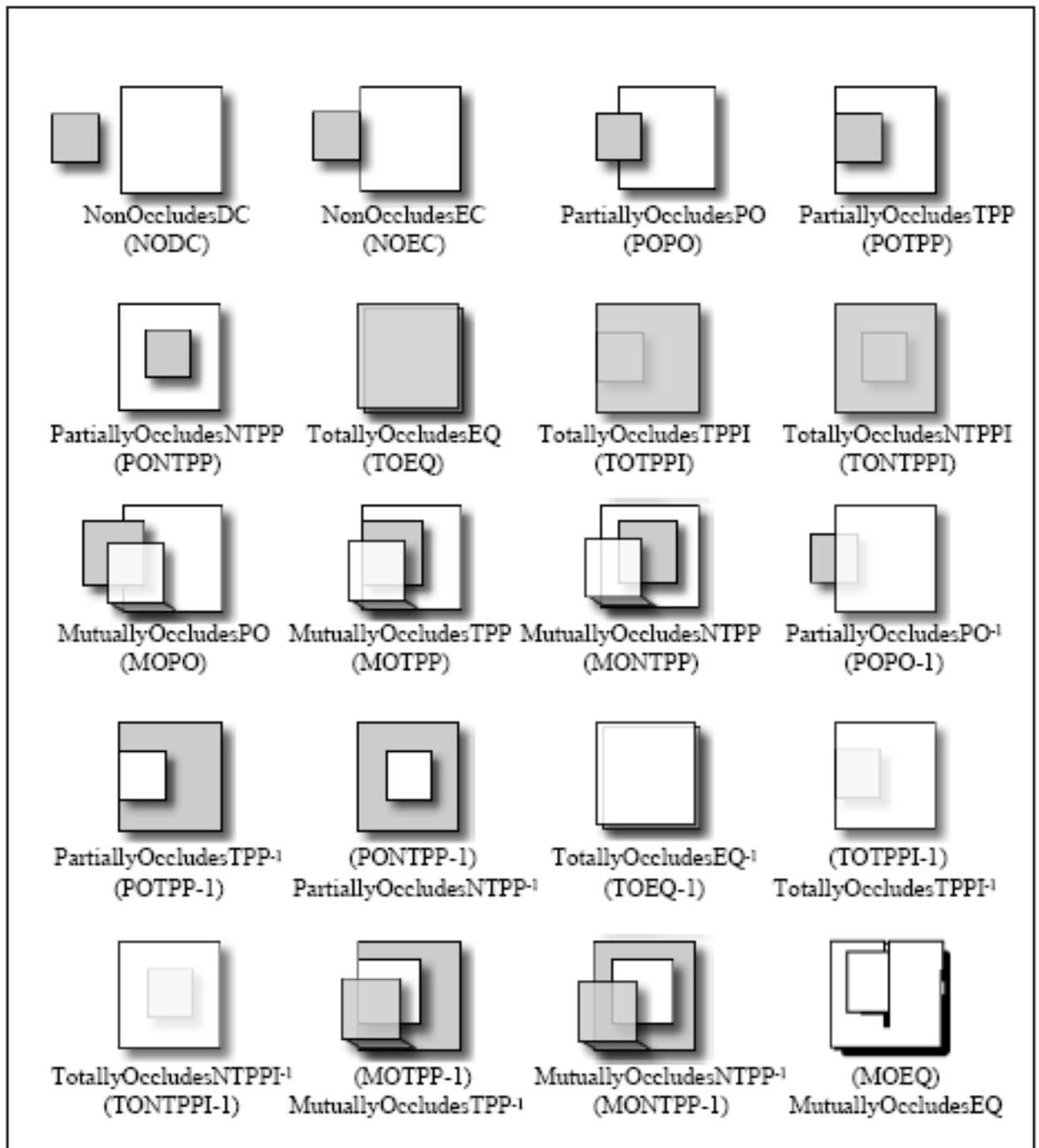


Figura 2.6: Modelo gráfico das 20 relações básicas do cálculo sobre oclusão de região, o ROC-20.

Fonte: (Randell and Witkowski, 2002).

O ROC-20 (Witkowski *et al.*, 2001) utiliza um conjunto de relações binárias baseadas no RCC-8 (Randell *et al.*, 1992), para modelar o relacionamento espacial entre corpos no espaço tridimensional e nas suas imagens bidimensionais, que foram obtidas a partir de um determinado ponto de vista. 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. Enquanto a função *image* representa o mapeamento relativo à uma projeção bidimensional do corpo físico, a partir de um determinado 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*”. Um conjunto de axiomas agindo como um conjunto de restrições espaciais entre os corpos de um dado ponto de vista e as imagens correspondentes deles são dados por (Randell and Witkowski, 2002) conforme descrito a seguir:

$$\forall x y [\Phi(\text{region}(x), \text{region}(y)) \rightarrow \forall \nu [\Phi(\text{image}(x, \nu), \text{image}(y, \nu))]]$$

**tipo** *region(Corpo)*: Região (significa que a função *region(Corpo)* é uma região espacial tridimensional [região 3D] pertencente ao tipo Região).

**tipo** *image(Corpo, Ponto)*: Região (significa que a função *image(Corpo, Ponto)* é uma região espacial bidimensional [região 2D] pertencente ao tipo Região).

**tipo**  $\Phi(\text{Região}, \text{Região})$  onde:

$\Phi \in \{C, O, P, NTPP, EQ\}$  A Região é o conjunto que contém a região 2D para imagem e região 3D para o mundo físico. Este axioma diz que, havendo alguma relação  $\Phi$  entre as regiões 3D dos corpos *x*, *y*, haverá uma relação do RCC  $\Phi$  na imagem de *x* e *y* do ponto de vista *ν*.

Portanto cada relação espacial descreve um alinhamento que pode ocorrer entre dois corpos em um campo visual, tornando possível modelar todas as transições contínuas por uma sequência de relações espaciais discretas. As mudanças no espaço visual e de posição dos corpos podem ser formuladas através de um diagrama conceitual de vizinhanças, esse gráfico é uma ferramenta para auxiliar na inferência de quais as possíveis sequências relativas a oclusão entre os objetos, na Figura 2.7 mostra um CND de 6 relações do ROC (ROC-6) (Randell and Witkowski, 2002).

### Oclusão

Para definir a oclusão entre dois corpos, uma segunda relação primitiva é definida: *TotallyOccludes(x, y, ν)*, lida como “*x* oclui totalmente *y* com relação ao ponto de vista *ν*”. *TotallyOccludes/3* é axiomatizada como transitiva e irreflexível (Randell et al., 2001).

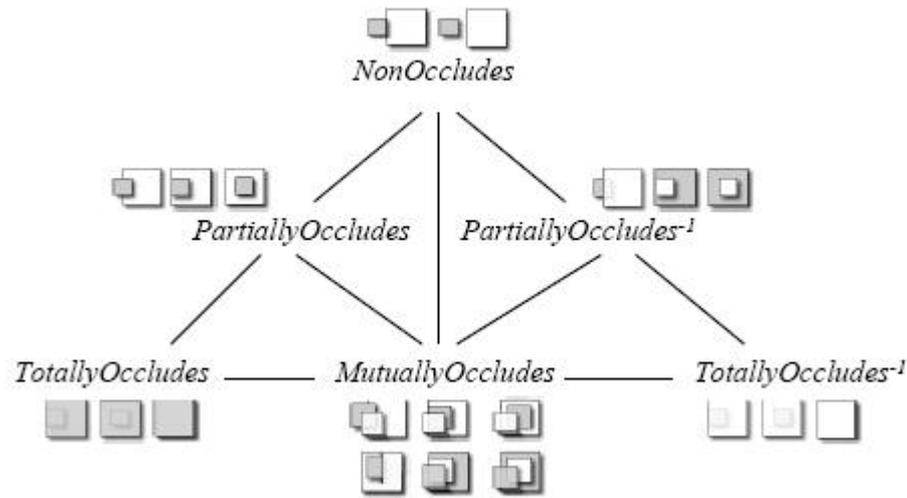


Figura 2.7: Diagrama conceitual de vizinhança para ROC-6.

Fonte:(Randell and Witkowski, 2002).

$\forall x \nu \neg \text{TotallyOccludes}(x, x, \nu)$  (irreflexível)

i.e. de qualquer ponto de vista  $\nu$  nenhum corpo se oclui totalmente.

$\forall x y z \nu [[\text{TotallyOccludes}(x, y, \nu) \wedge \text{TotallyOccludes}(y, z, \nu)] \rightarrow \text{TotallyOccludes}(x, z, \nu)]$  (transitividade)

i.e. dada a transitividade da oclusão total, um objeto  $x$  pode ocluir totalmente um objeto  $z$  mesmo se  $z$  estiver totalmente ocluído por outro objeto  $y$ .

Para demonstrar o significado geométrico da oclusão total, considere uma  $\text{linha}(p_1, p_2, p_3)$  definida tal 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,  $x$  oclui totalmente  $y$  do ponto de vista de  $\nu$  se e somente se para todo o ponto  $p$  em  $y$ , existir um ponto  $q$  em  $x$  tal que  $\text{linha}(\nu, q, p)$  seja verdade e nenhum outro ponto  $p'$  em  $y$  e  $q'$  em  $x$ , tal que  $\text{linha}(\nu, p', q')$  seja verdadeiro. Vários axiomas que constituem o ROC serão apresentados a seguir (Randell et al., 2001):

$\forall x y z \nu [[\text{TotallyOccludes}(x, y, \nu) \wedge P(\text{region}(z), \text{region}(y))] \rightarrow \text{TotallyOccludes}(x, z, \nu)]$

i.e. se  $x$  oclui totalmente  $y$ ,  $x$  oclui totalmente qualquer parte de  $y$ .

$\forall x y \nu [\text{TotallyOccludes}(x, y, \nu) \rightarrow \forall z [P(\text{region}(z), \text{region}(y))] \rightarrow \neg \text{TotallyOccludes}(z, x, \nu)]$

i.e. se  $x$  oclui totalmente  $y$  nenhuma parte de  $y$  oclui totalmente  $x$ .

$\forall x y \nu [\text{TotallyOccludes}(x, y, \nu) \rightarrow \forall z [[P(\text{region}(z), \text{region}(x)) \wedge P(\text{region}(u), \text{region}(y))] \rightarrow \neg \text{TotallyOccludes}(u, z, \nu)]]$

i.e. se  $x$  oclui totalmente  $y$  nenhuma parte de  $y$  oclui totalmente parte de  $x$ .

$$\forall x \nu \exists y z [P(\text{region}(y), \text{region}(x)) \wedge P(\text{region}(z), \text{region}(x)) \wedge \text{TotallyOccludes}(y, z, \nu)]$$

i.e. todo  $x$  tem uma parte que oclui totalmente outra parte de  $x$ . Este axioma garante que o corpo tem profundidade.

$$\forall x y \nu [\text{TotallyOccludes}(x, y, \nu) \rightarrow P(\text{image}(y, \nu), \text{image}(x, \nu))]$$

i.e. se  $x$  oclui totalmente  $y$ , a imagem de  $y$  é parte da imagem de  $x$ .

A seguir uma definição onde a relação de oclusão é enfraquecida para incluir, por exemplo, uma oclusão parcial: ‘ $\text{Occludes}(x, y, \nu)$ ’ lida como “ $x$  oclui  $y$  do ponto de vista de  $\nu$ ” (Randell et al., 2001) e outras relações de oclusão mais específicas tais como as oclusões parcial e mútua:

$$\text{Occludes}(x, y, \nu) \equiv \exists z u [P(\text{region}(z), \text{region}(x)) \wedge P(\text{region}(u), \text{region}(y)) \wedge \text{TotallyOccludes}(z, u, \nu)]$$

i.e.  $x$  oclui  $y$  se pelo menos existe uma parte de  $x$  que oclui totalmente pelo menos uma parte de  $y$ .

$$\text{PartiallyOccludes}(x, y, \nu) \equiv \text{Occludes}(x, y, \nu) \wedge \neg \text{TotallyOccludes}(x, y, \nu) \wedge \neg \text{Occludes}(y, x, \nu)$$

i.e  $x$  oclui parcialmente  $y$ , mas  $y$  não oclui  $x$ .

$$\text{MutuallyOccludes}(x, y, \nu) \equiv \text{Occludes}(x, y, \nu) \wedge \text{Occludes}(y, x, \nu)$$

i.e  $x$  e  $y$  ocluem um ao outro.

As relações inversas e nulas são encontradas em (Randell et al., 2001; Randell and Witkowski, 2002) e não serão listadas aqui. Refinando o conjunto de relações de oclusão temos um total de 20 relações de disjunções e conjunções de pares, essas relações tem sua representação gráfica apresentada na Figura 2.6 e são definidas pelo esquema de axiomas abaixo (Randell et al., 2001).

$$\Phi\Psi(x, y, \nu) \equiv \Phi(x, y, \nu) \wedge \Psi(\text{image}(x, \nu), \text{image}(y, \nu))$$

$$\Xi\Psi^{-1}(x, y, \nu) \equiv \Xi(y, x, \nu) \wedge \Psi(\text{image}(y, \nu), \text{image}(x, \nu))$$

onde:

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

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

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

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

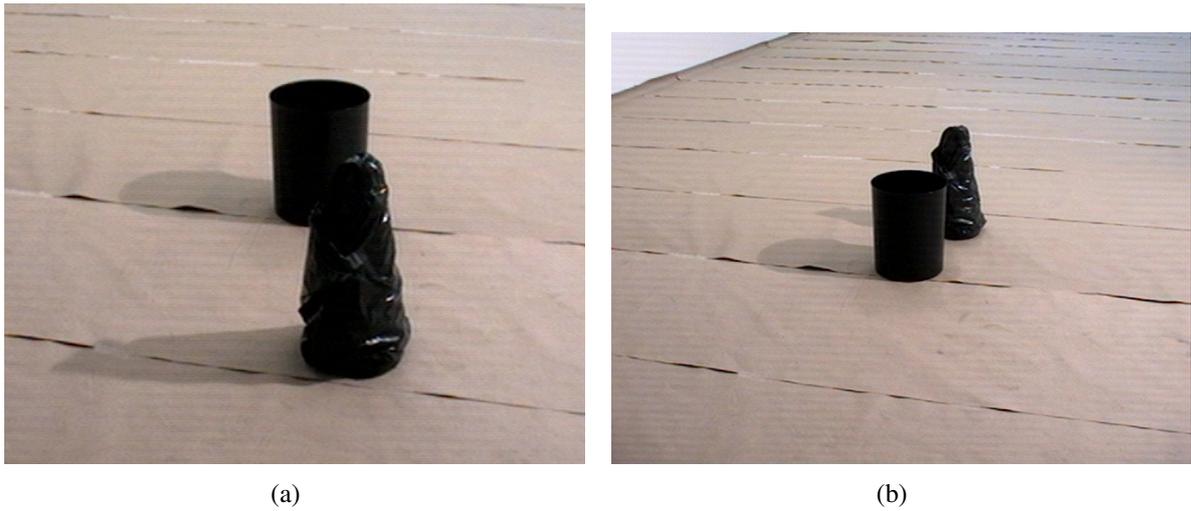


Figura 2.8: Objetos parcialmente oclusos, na imagem suas regiões estão sobrepostas parcialmente.

e onde:

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

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

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

Com os axiomas acima as relações de oclusão em ambiente tridimensional podem ser interpretadas num plano, como exemplos de definições a partir destes axiomas, seguem as fórmulas abaixo que podem representar a situação demonstrada na Figura 2.8, considerando que  $x$  seja o cone e  $y$  o cilindro:

**Figura 2.8(a)**  $PartiallyOccludesPO(x, y, \nu) \rightarrow PartiallyOccludes(x, y, \nu) \wedge PO(image(x, \nu), image(y, \nu))$

**Figura 2.8(b)**  $PartiallyOccludesPO^{-1}(x, y, \nu) \rightarrow PartiallyOccludesPO(y, x, \nu) \wedge PO(image(y, \nu), image(x, \nu))$

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. Ambas, profundidade e distância, são essenciais em teorias lógicas quando utilizadas em aplicações práticas em robótica (Witkowski *et al.*, 2001). Portanto a noção de oclusão pode ser utilizada para inferir profundidade a partir de cenas bidimensional. Deste modo a projeção bidimensional de uma sombra também pode produzir forte impressão na percepção tridimensional dos objetos (Kersten *et al.*, 1994). A seguir, veremos alguns estudos psicofísicos sobre a importância da sombra para a percepção humana.



Figura 2.9: Sombras nos informam sobre a posição relativa dos objetos. Na imagem à esquerda não podemos determinar qual a posição do robô no entanto nas outras três imagens percebe-se que ele está cada vez mais longe do chão.

Fonte: *Hasenfratz et al.* (2003).

## 2.4 Percepção de sombras

O ser humano tem uma habilidade extraordinária para perceber e analisar a estrutura tridimensional do mundo através do que vê mesmo a partir de imagens bidimensionais como fotografias, filmes etc. Apesar do processo de raciocínio utilizado não ser totalmente entendido, sabemos que vários fenômenos permitem a interpretação tridimensional de imagens bidimensionais (*Shapiro and Stockman*, 2001). Estudos sugerem que uma sombra é um forte indicativo sobre a informação tridimensional em cenas bidimensionais (*Imura et al.*, 2005), além de induzir a aparência de movimento em profundidade do objeto ao qual a sombra corresponde (*Kersten et al.*, 1994).

Pesquisadores tem desenvolvido experimentos para entender o impacto da sombra em nossa percepção de cenas. Através de testes psicofísicos foi possível confirmar que em imagens estáticas, sombras são um forte indicativo de profundidade até mesmo para crianças (*Imura et al.*, 2005), além de serem também fonte de informações sobre a posição do objeto (Figura 2.9). Na Figura 2.10 pode-se notar que as sombras contém informações em potencial sobre objetos oclusos do ponto de vista do observador, nesta imagem pode ser deduzido o que o robô tem na mão (que não é vista). A sombra pode ser utilizada também na percepção da forma de superfícies ambíguas (Figura 2.11) (*Mamassian et al.*, 1998).

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 (Figura 2.12), enquanto que na sombra-própria a superfície de projeção é o próprio objeto que a provocou



Figura 2.10: A sombra informa sobre a forma do objeto. Nesta figura podemos notar que o robô tem o braço ocluso e segura uma chaleira e um arco em duas das imagens.

Fonte: *Hasenfratz et al.* (2003).

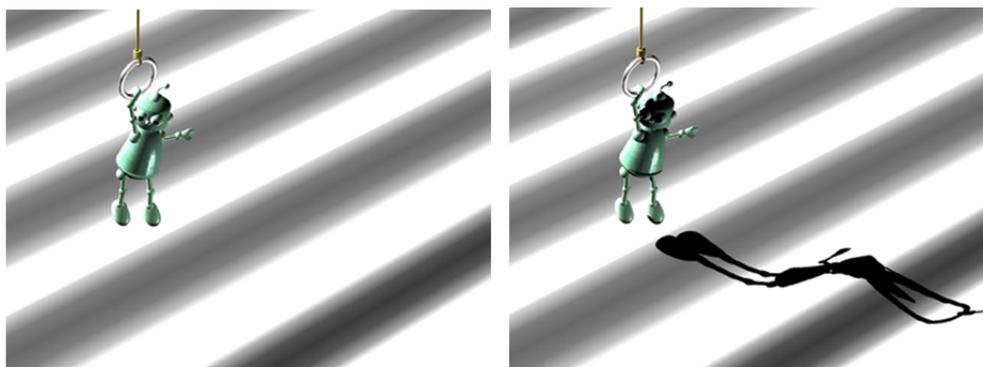


Figura 2.11: A sombra informa sobre a superfície em que foi projetada. Na imagem da esquerda o chão é uma superfície ambígua, enquanto que a direita a sombra revela a geometria da superfície em que está projetada.

Fonte: *Hasenfratz et al.* (2003)

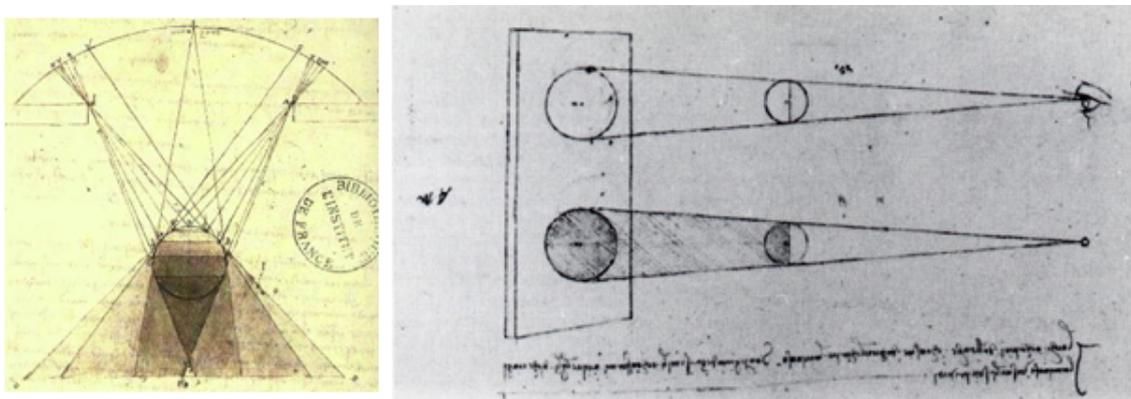


Figura 2.12: Estudos sobre sombra de Leonardo da Vinci. No esboço vê-se representados as sombras do objeto projetada em outra superfície (sombra-projetada) e em si próprio (sombra-própria)

Fonte: *da Vinci* (1970).

(*Knill et al.*, 1997), uma análise da estrutura geométrica das sombras pode ser encontrado em *Knill et al.* (1997). Nos 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.

O movimento causado pela alteração na localização da sombra é caracterizado pela mudança do ângulo visual  $\theta$ . O ângulo visual  $\theta$  é formado pela distância entre o objeto e a sua sombra

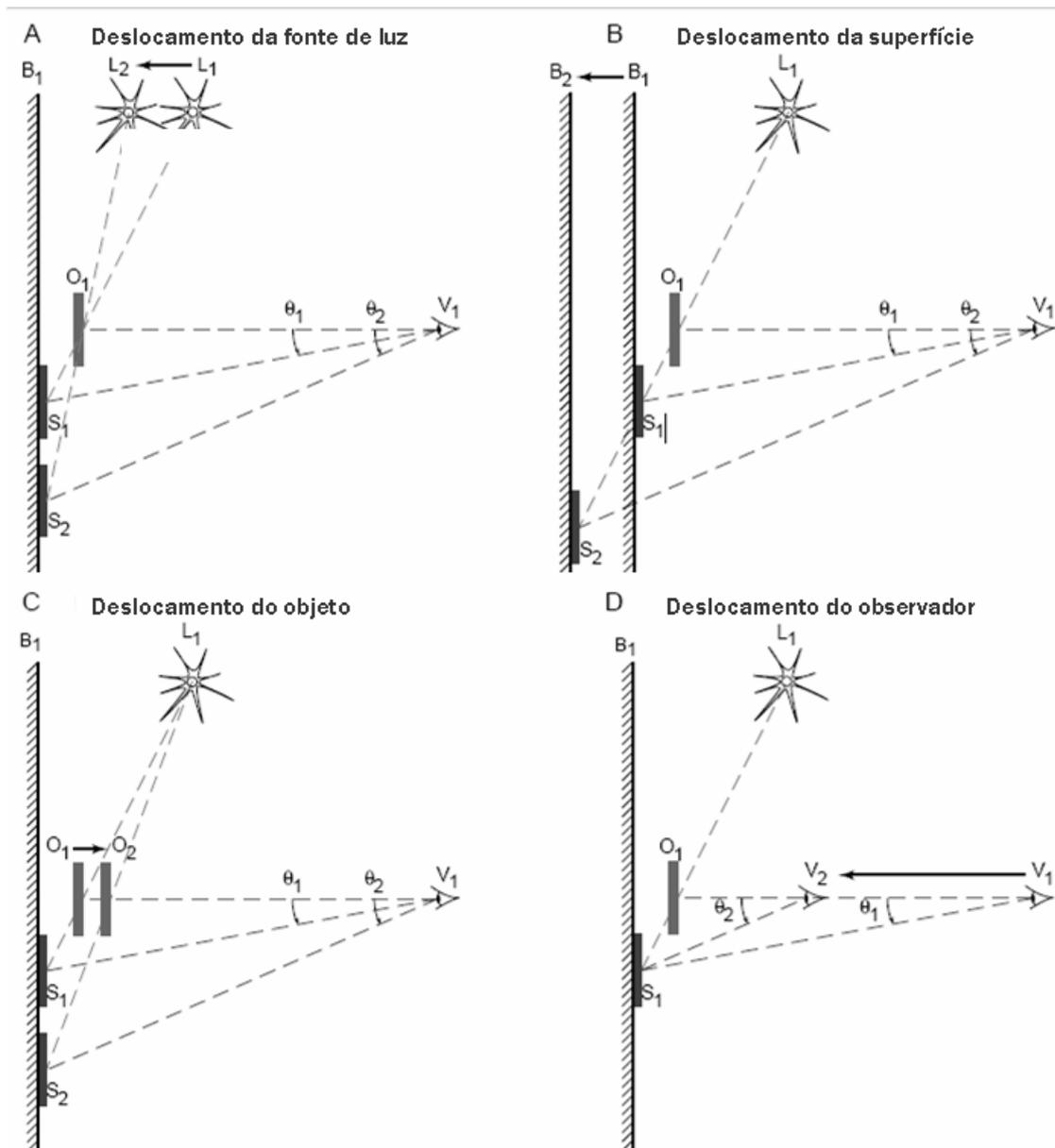


Figura 2.13: Movimento da sombra pode ser ocasionado pelo deslocamento da fonte de luz (A), da superfície (B), do objeto (C) e do observador (D).

Fonte: Mamassian et al. (1998).

do ponto de vista do observador. Em uma sequência de cenas como mostrado na Figura 2.13, podemos considerar os ângulos  $\theta_1$  e  $\theta_2$  como ângulos inicial e final, e com eles calcular as posições dos elementos envolvidos. Com isso, a sombra tem intrinsicamente informação sobre as propriedades 3D de uma cena tais como direção da iluminação, estrutura da superfície, posicionamento do observador e do objeto (Mamassian et al., 1998). Testes psicofísicos sugerem que a sombra ajuda na localização e na percepção do movimento em sequência de cenas (Rensink and Cavanagh, 2004; Ni et al., 2004), porém não são relevantes no reconhecimento de objetos (Bonfiglioli et al., 2004).

Na Figura 2.14 aparecem três imagens que (quando mostradas em sequência) dão a im-

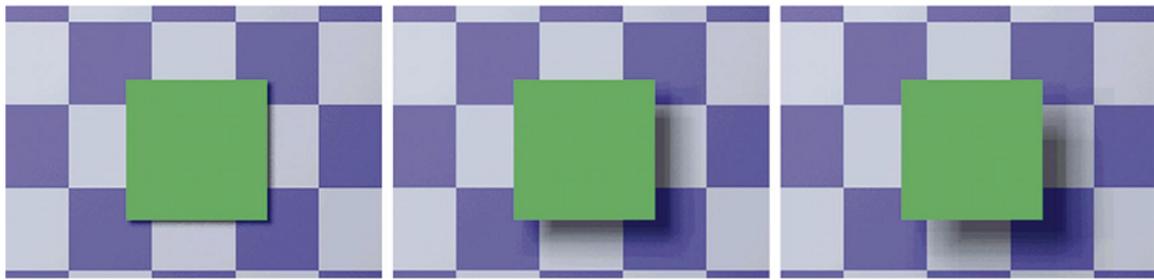


Figura 2.14: Efeito na localização do objeto ocasionado pela posição da sombra.  
Fonte: Kersten *et al.* (1994).

pressão 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.

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 *et al.*, 1998). De fato, nos estudos realizados por (Kersten *et al.*, 1994), relacionando objetos em movimento e a percepção relativa a 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 2.15. Observando a Figura 2.15 vemos em (a) a bola posicionada próxima a superfície a frente da caixa, enquanto que em (b) e (c) apesar da bola estar na mesma posição, devido a 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 seguir veremos algumas técnicas de visão computacional básicas para a detecção de sombras.

## 2.5 Visão computacional

A visão é um dos nossos sentidos mais desenvolvidos, portanto não é uma surpresa que a imagem exerça um papel importante na percepção humana. O objetivo da visão computacional

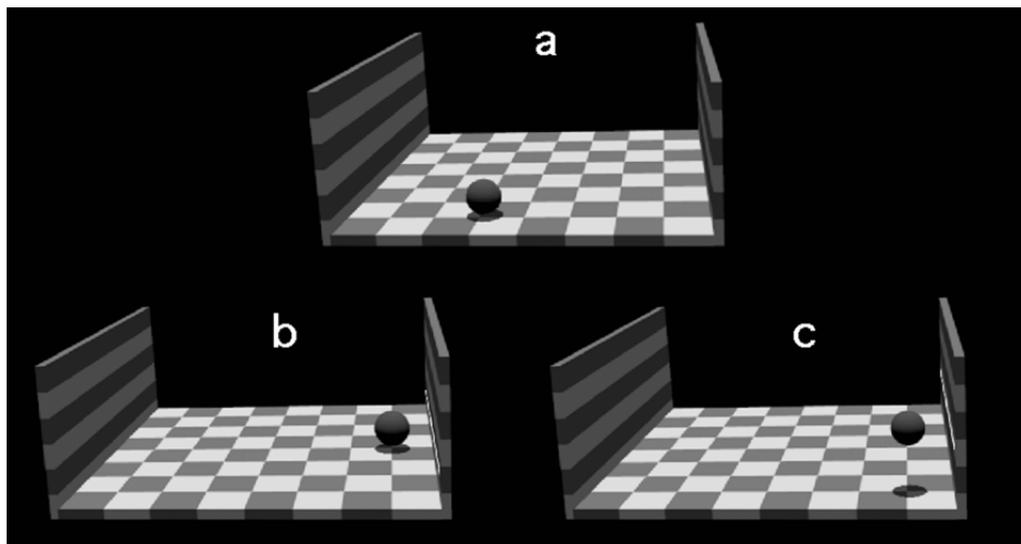


Figura 2.15: Bola na caixa. Mesmo na mesma localização da bola em (b) e (c) percebemos posicionamentos diferentes devido a sombra.

Fonte: Kersten et al. (1994).

(neste trabalho) é capacitar um robô móvel a obter informações úteis sobre objetos que estão em um mundo físico real através de imagens capturadas (Shapiro and Stockman, 2001).

A visão computacional pode ser definida por níveis de processo. Os processos de baixo nível de visão de baixo nível são caracterizados por melhorar a qualidade da imagem, 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 (i.e. separar a imagem em regiões ou objetos), 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 (Rafael C. Gonzalez, 2001). Pode-se perceber que cada nível é dependente do anterior, e todos são dependentes do processamento digital da imagem.

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 de oito bits com seu valor variando de zero a 255. Esse valor corresponde à intensidade de luminância discretizada em imagens monocromáticas, as imagens coloridas consistem de três

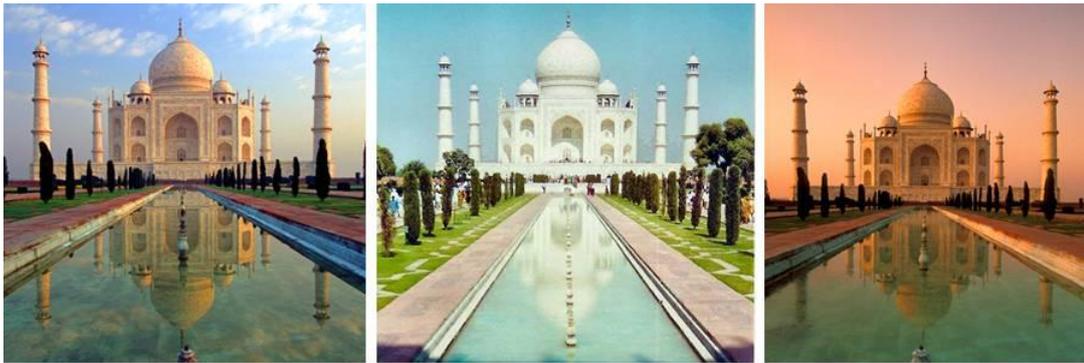


Figura 2.16: As imagens do Taj Mahal foram tiradas em horários diferentes do dia, as fotos da esquerda para a direita são do amanhecer, meio do dia, e entardecer. O próprio monumento assume um tom diferentes em cada uma delas.

Fonte: adaptado de vários sites <http://www.google.com.br> último acesso em 11/07/2008.

canais de informação de oito bits cada. Veremos a seguir a representação de cores em imagens digitais.

### 2.5.1 Espaço de cores em imagens digitais

As cores podem conter informações que possibilitam distinguir objetos, materiais, lugares e ou diferentes horários do dia (*Shapiro and Stockman, 2001*), a Figura 2.16 mostra o monumento *Taj Mahal* em diferentes horas do dia, as quais podem ser percebidas pelos detalhes de sombras e cores. As imagens digitais utilizam modelos de representação de cores conhecidos como espaços de cores. Há vários espaços de cores, dentre eles os mais conhecidos são o RGB (vermelho, verde, azul), CMY (ciano, magenta, amarelo) e HSV (hue, saturação, valor). Os padrões CMY e RGB são modelos geralmente ligados à hardware. O padrão CMY é o modelo utilizado por impressoras, e o RGB por monitores coloridos. O espaço de cores HSV é muito utilizados nos sistemas com visão computacional, por corresponder mais a descrição e interpretação de cores feitas por humanos (*Rafael C. Gonzalez, 2001*). Pelas características de hardware e dos tipos de dados analisados serão utilizados, neste trabalho, os espaços de cores RGB e HSV, dos quais falaremos a seguir.

### 2.5.2 Modelos de cores RGB e HSV

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 2.17). Esse cubo é um sólido composto de  $(2^8)^3 \approx 16.000.000$  de cores. Podemos simplificar dizendo que RGB é bom para

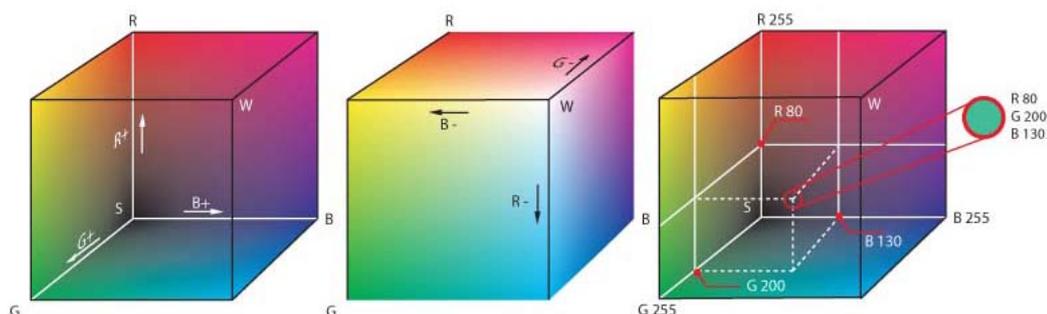


Figura 2.17: Representação gráfica do cubo de cores RGB, o cone da esquerda representa como uma cor(80,200,130) está localizada neste sistema cartesiano.

Fonte: [http://upload.wikimedia.org/wikipedia/commons/0/03/RGB\\_farbwuerfel.jpg](http://upload.wikimedia.org/wikipedia/commons/0/03/RGB_farbwuerfel.jpg) último acesso em 20/02/2009.

geração de imagens coloridas, porém é limitado para descrição de cores. Essa dificuldade de descrever cores é porque não pensamos em cores pelo resultado da combinação entre elas, mas por características como: cor, tonalidade, e sombreamento; essas são características intrínsecas aos canais de HSV (*Rafael C. Gonzalez, 2001*).

O modelo HSV ou HSI é baseado em um sistema de coordenadas cilíndricas, composto pelos componentes de intensidade ou brilho, pela cromaticidade representadas pelo “hue”, e pela saturação. As três propriedades deste modelo são representadas por um cone conforme Figura 2.18. O “hue” **H** ou matiz é definido pelo o ângulo entre 0 e  $2\pi$  relativo ao eixo do cone, matiz igual a 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 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.

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 (*Rafael C. Gonzalez, 2001; Shapiro and Stockman, 2001*). A transformação de um modelo de cor RGB para HSV e vice-versa é não linear, por brevidade não detalharemos sua matemática. Vamos ver a seguir alguns métodos de processamento de visão computacional de nível médio.

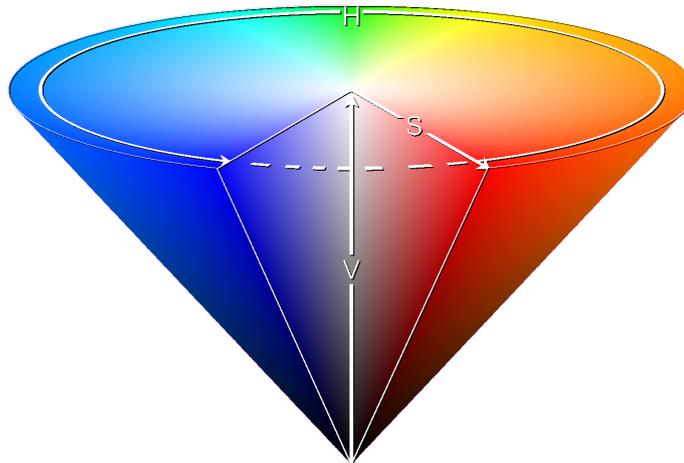


Figura 2.18: Representação gráfica do cone HSV.

Fonte: [http://upload.wikimedia.org/wikipedia/commons/e/ea/HSV\\_cone.png](http://upload.wikimedia.org/wikipedia/commons/e/ea/HSV_cone.png) último acesso em 20/02/2009.

### 2.5.3 Métodos para análise de imagens

A visão computacional é uma área da ciência ainda em desenvolvimento, mas possui uma grande variedade de métodos e técnicas para processamento de imagens, tais como: filtros, detectores de bordas, métodos de segmentação de imagem etc. Contudo, neste trabalho serão descritos somente aqueles que utilizaremos para a localização de objetos e sombras, o primeiro destes métodos é a limiarização, ou seja, aplicação do filtro de limiar.

#### Filtros 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 alvo (*foreground*), que é geralmente um objeto ou região de interesse, os outros *pixels* não escolhidos serão o fundo (*background*). Por exemplo, dado 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: limiar acima  $f(x, y) \geq T$ ; limiar abaixo  $f(x, y) \leq T$ ; limiar interno  $T_{min} \leq f(x, y) \leq T_{max}$ ; e limiar externo  $f(x, y) < T_{min}$  ou  $f(x, y) > T_{max}$  (Figura 2.19)(Rafael C. Gonzalez, 2001; Shapiro and 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 a zero mantendo

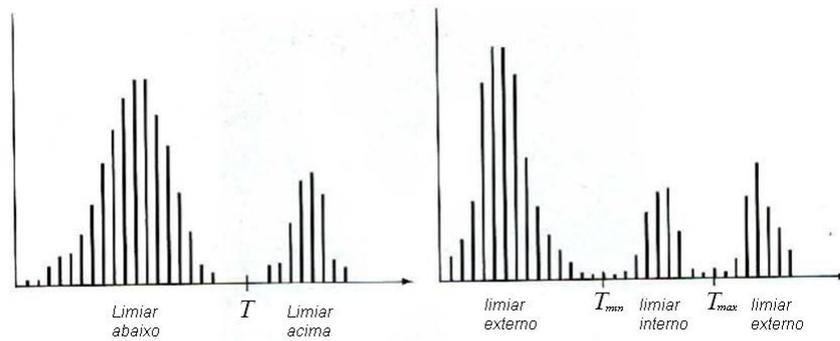


Figura 2.19: Gráfico representando o filtro de limiar e o limiar  $T$   
 Fonte: adaptada de *Rafael C. Gonzalez* (2001).

a intensidade do alvo. Esse é um filtro útil para segmentar imagens onde os objetos têm características bem definidas em um dos canais de informação (*Rafael C. Gonzalez*, 2001; *Shapiro and Stockman*, 2001).

### 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  $\mathbf{E}$ , forma-se 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 and Stockman*, 2001).

Procurando facilitar o entendimento tem-se uma representação da erosão e da dilatação na Figura 2.20. Pode-se utilizar um filtro morfológico de erosão para eliminar pequenas áreas e logo 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 ao operador morfológico de abertura, é utilizado para completar pequenos “buracos” (*Shapiro and Stockman*, 2001).

### Detectores de borda

Os detectores de bordas funcionam identificando as variações acentuadas de 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 proces-

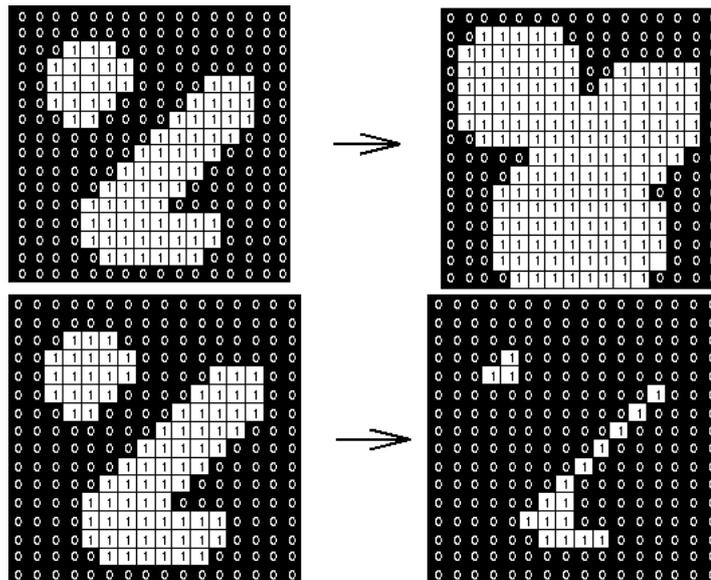


Figura 2.20: Representação gráfica da erosão e dilatação utilizando uma matriz 3x3 como elemento estruturante **E**.

Fonte: <http://homepages.inf.ed.ac.uk/rbf/HIPR2/> último acesso em 20/02/2009.

samento 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:

$$\nabla f = \begin{pmatrix} \frac{\delta f}{\delta x} \\ \frac{\delta f}{\delta y} \end{pmatrix} \quad (2.1)$$

e a magnitude deste vetor é dada por,

$$mag(\nabla f) = \left[ \left( \frac{\delta f}{\delta x} \right)^2 + \left( \frac{\delta f}{\delta y} \right)^2 \right]^{\frac{1}{2}}. \quad (2.2)$$

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 (Rafael C. Gonzalez, 2001; Shapiro and Stockman, 2001).

### Componentes conectados

A segmentação de imagens também exerce papel de auxiliar na diferenciação das diversas regiões numa imagem. Há vários tipos de segmentação, já foi descrita a segmentação por limiar, agora descreveremos um pouco a segmentação por componente conectado, pois a utilizamos para distinguir cada objeto e sombra. 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 2.21, da esquerda para direita, vê-se a imagem origi-

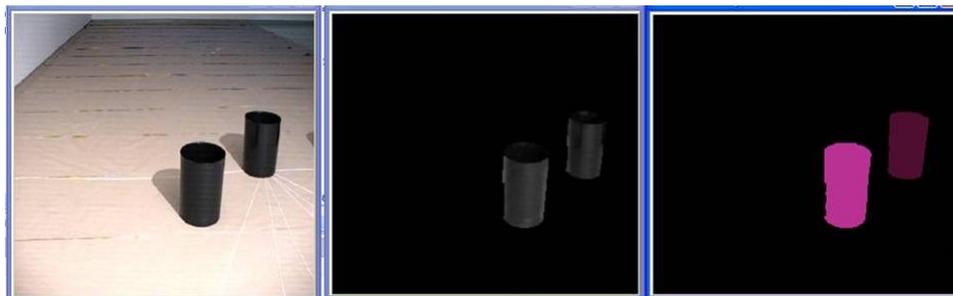


Figura 2.21: Representação gráfica da erosão e dilatação utilizando uma matriz 3x3 como elemento estruturante.

nal 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 podemos extrair informações como área, quantidade de elementos etc. As informações extraídas a partir deste filtro são utilizadas em etapas de nível alto, como classificação, decisão etc. A seguir veremos alguns trabalhos em visão computacional para detecção de sombras.

## 2.6 Segmentação de sombras

Vimos anteriormente que a sombra contém informação, porém não é nada por ela mesma. Um observador que deseje notá-la a fim de inferir localização, formas e a identidade dos objetos à sua volta, não vai querer confundi-la com objetos reais, e tão pouco querer que a sombra interfira na observação dos elementos de interesse na cena (*Elder et al.*, 2004). Todavia, o interesse em detectar sombra, por parte da comunidade de visão computacional, é para eliminá-la, isso porque freqüentemente pontos das sombras são classificados como pontos do objeto, tal engano causa erros em aplicações onde se utiliza o rastreamento visual ou o reconhecimento de objetos (Figura 2.22), portanto, a sombra é usualmente considerada um ruído e como tal, deve ser detectada e eliminada (*Joshi et al.*, 2007).

Em geral a detecção e filtragem das sombras são feitas tipicamente para melhorar a qualidade dos resultados da segmentação, e não para deduzir parâmetros do objeto ou dados do vídeo. Na literatura, os algoritmos de detecção de sombras são normalmente associados com técnicas de segmentação de objetos em movimento, uma avaliação das técnicas principais é encontrada em (*Prati et al.*, 2003). Os resultados em (*Prati et al.*, 2003) mostraram que o sistema de melhor desempenho geral foi o que utilizava uma abordagem de detecção de sombras baseada no espaço de cores HSV. Este sistema é detalhado em (*Chucchiara et al.*, 2001), onde as sombras são identificadas como regiões com baixa intensidade luminosa e baixa saturação. A abordagem com HSV também foi realizada em vários outros trabalhos, dentre eles o de de-



Figura 2.22: Aplicação do algoritmo de detecção de objetos em movimento. A esquerda nota-se uma detecção errada dos objetos, enquanto que a direita a localização é correta devido a supressão das sombras.

Fonte: *Chucchiara et al. (2001)*.

teção de sombras e nuvens em imagens de satélite, descrito em (*Polidorio et al., 2005*), que define uma região de interesse antes de aplicar limiar nos canais escolhidos.

Há trabalhos que seguem uma sequência de subtração de fundo, compara os *pixels* do fundo com o quadro atual, utilizando diferença na direção do gradiente, diferença de magnitude, proporção de intensidade, diferença de cores, achando assim sombras conforme a probabilidade de cada um desses parâmetros (*Joshi et al., 2007*). Outros trabalhos utilizam uma máscara de mudança de valores com valor 255 (no ponto onde ocorreu alteração no último quadro) e zero onde não ocorreu alteração (*Stauder et al., 1999*).

Em (*Chucchiara et al., 2001*), com uma câmera fixa, os objetos são detectados através da diferença da luminância da imagem  $D_k$  entre dois *frames* dada pela equação 2.3, onde  $s_k$  é a luminância dos pontos da imagem de coordenada  $(x, y)$  no instante  $k$ .

$$D_k(x, y) = s_{k+1}(x, y) - s_k(x, y) \quad (2.3)$$

Outro trabalho que tem sua detecção de sombra baseada em subtração de imagem, onde para imagem estática subtrai a imagem do ambiente sem sombra (fundo), e para imagens dinâmicas faz-se a diferença de quadros, só que neste caso ele utiliza a sombra como indicativo de profundidade para um robô, pela a distância na imagem entre a sombra do braço do robô e o seu objetivo (*Fitzpatrick and Torres-Jarra, 2004*). Foi encontrado também um trabalho, que utiliza a sombra para estimar pose e forma humana tridimensional, descrito em *Balan et al. (2007)*. Nele a segmentação é realizada separando o plano de fundo (*background*), o alvo (*foreground*) e a sombra, somente com as informações do canal de saturação. Neste canal, são aplicadas várias técnicas de visão computacional, incluindo filtro da mediana, dilatação e erosão de imagem

para remoção de pequenos componentes desconectados e para o preenchimento de lacunas.

## 2.7 Discussão final

Nesta revisão bibliográfica foi visto que a sombra influencia a percepção humana, induzindo a percepção de tridimensionalidade a partir de cena bidimensional. Um robô que perceba o mundo com sensor de visão monocular deve raciocinar sobre o mundo a partir de cenas bidimensionais. Sabendo que sombras são indicativos de movimento e profundidade para seres humanos, pode-se então criar um sistema que auxilie um robô a executar inferências sobre conhecimento espacial. Esse sistema utilizará as informações extraídas das imagens de objetos e suas sombras.

A partir de uma nova teoria espacial proposta em (Santos *et al.*, 2008) capacitamos um robô a interpretar o que ocorre no ambiente através das relações entre as regiões do objeto e da sombra ocorridas nas imagens. Essa nova teoria baseia-se nas teorias de raciocínio espacial qualitativo RCC (Randell *et al.*, 1992) e ROC (Randell *et al.*, 2001) relatadas na seção 2.3 e é chamada de Relações Perceptuais Qualitativas sobre Sombras (PQRS) a qual trataremos no Capítulo 3. Esse trabalho difere dos trabalhos realizados por (Fitzpatrick and Torres-Jarra, 2004; Balan *et al.*, 2007) por explorar outras informações espaciais, reforçando que sombra tem informações ainda a explorar.

## Capítulo 3

# RELAÇÕES PERCEPTUAIS QUALITATIVAS SOBRE SOMBRAS

Sendo a lógica uma teoria matemática que relaciona premissas ou conhecimentos anteriores com as características extraídas do mundo, ela é uma ferramenta que permite a dedução de novas informações relevantes para um sistema (McCarthy, 1990). Vamos utilizar atributos retirados das sombras de objetos, que são fonte de informações sobre movimento, profundidade e volume (Mamassian et al., 1998). Neste capítulo descreveremos uma nova teoria lógica proposta em (Santos et al., 2008, 2009) que tem como base as teorias do RCC e ROC (seção 2.3) e que chama-se relações perceptuais qualitativas sobre sombras (PQRS - *Perceptual Qualitative Relations about Shadows*). A maior parte do conteúdo deste capítulo foi adaptado dos artigos (Santos et al., 2008, 2009).

### 3.1 Formalização das relações perceptuais qualitativas sobre sombras

Nossa representação lógica deve conter os elementos do ambiente com o qual trabalhamos. 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  $S_{cr}$  representa a superfície onde a sombra é projetada. Essa superfície não precisa ser necessariamente lisa ou contínua, pois neste trabalho a segmentação utiliza a intensidade dos *pixels*, e não as formas dos objetos. O observador será representado com o símbolo  $\nu$ ;  $o$  representa objetos e  $s$  sombras. Os conjuntos de variáveis para estes elementos são representados como segue:

**objetos físicos:**  $o_1, \dots, o_n$ ;

**instantes de tempo:**  $t_1, \dots, t_n$ ;

**regiões espaciais:**  $r_1, \dots, r_n$ ;

**sombras:**  $s_1, \dots, s_n$ .

Assumiremos que as variáveis são universalmente quantificadas a menos que informado explicitamente. O conjunto de relações perceptuais qualitativas sobre sombras (PQRS) de um objeto  $o$  e sua sombra  $s$  utiliza o RCC e um subgrupo do ROC (seção 2.3), composto pelas relações representadas graficamente na Figura 3.1. A seguir as interpretações e definições das relações básicas do PQRS, são exemplificadas com imagens capturadas em nosso laboratório (Figura 3.2) (a conexão com a base do objeto é desconsiderada):

$$NonOccludesDC(o, s) \equiv NonOccludes(o, s) \wedge DC(image(o, \nu), image(s, \nu))$$

i.e. Figura 3.2(a), onde  $s$  não está ocluído e está desconectado de  $o$ ;

$$NonOccludesEC(o, s) \equiv NonOccludes(o, s) \wedge EC(image(o, \nu), image(s, \nu))$$

i.e. Figura 3.2(b),  $o$  não oclui  $s$  porém suas bordas estão conectadas;

$$PartiallyOccludesPO(o, s) \equiv$$

$$PartiallyOccludes(o, s) \wedge PO(image(o, \nu), image(s, \nu))$$

i.e. Figura 3.2(c),  $o$  oclui  $s$  de modo que este aparece parcialmente na imagem;

$$PartiallyOccludesTPP(o, s) \equiv$$

$$PartiallyOccludes(o, s) \wedge TPP(image(o, \nu), image(s, \nu))$$

i.e. Figura 3.2(d),  $o$  oclui parte de  $s$  tangenciando uma das bordas de  $s$  na imagem;

$$TotallyOccludesTPPI(o, s) \equiv$$

$$TotallyOccludes(o, s) \wedge TPPI(image(o, \nu), image(s, \nu))$$

i.e. Figura 3.2(e),  $o$  oclui totalmente  $s$  e  $s$  é parte tangencial a  $o$  na imagem;

$$TotallyOccludesEQ(o, s) \equiv$$

$$TotallyOccludes(o, s) \wedge EQ(image(o, \nu), image(s, \nu))$$

i.e. Figura 3.2(f),  $o$  oclui totalmente  $s$  e  $s$  é igual a  $o$  na imagem;

$$TotallyOccludesNTPPI(o, s) \equiv$$

$$TotallyOccludes(o, s) \wedge NTPPI(image(o, \nu), image(s, \nu))$$

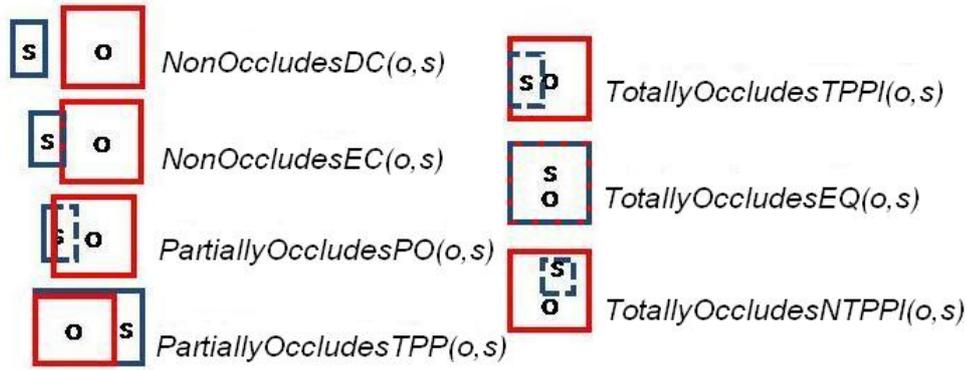


Figura 3.1: Representação gráfica das relações básicas do PQRS.

i.e. Figura 3.2(g),  $o$  oclui totalmente  $s$  e  $s$  não tangencia  $o$  na imagem.

O predicado  $Shadow(s, o, Scr, L)$  significa que  $s$  é a sombra projetada na superfície  $Scr$ , provocada pela obstrução da fonte de luz  $L$  pelo objeto  $o$ . Esse predicado diz que  $s$  é sombra se e somente se existir pelo menos um objeto  $o$ , uma superfície  $Scr$ , e uma fonte de luz  $L$  onde  $Shadow(s, o, Scr, L)$  é verdade. Portanto, pode-se declarar que a sombra de um objeto é a mesma região da superfície que está totalmente ocluída pelo objeto do ponto de vista da fonte de luz. Formalizando essa afirmação temos o seguinte axioma,

$$Shadow(s, o, Scr, L) \leftrightarrow PO(region(s), region(Scr)) \wedge \quad (3.1)$$

$$TotallyOccludes(o, s, L) \wedge \neg \exists o' TotallyOccludes(o', o, L).$$

Esse axioma diz que para  $Shadow(s, o, Scr, L)$  ser verdade, a região da sombra  $region(s)$  deve sobrepor parcialmente a região da superfície  $region(Scr)$ , e o objeto  $o$  deve ocluir totalmente a sombra  $s$  do ponto de vista da fonte de luz  $L$ , e não existe nenhum outro objeto  $o'$  que oclua totalmente o objeto  $o$  deste mesmo ponto de vista  $L$ . A conjunção  $\neg \exists o' TotallyOccludes(o', o, L)$  garante que a sombra é de  $o$  e de nenhum outro objeto  $o'$  ( $o \neq o'$ ). Portanto “ $s$  é uma sombra de  $o$ ” é definido pelo seguinte axioma (Santos et al., 2009):

$$Is\_a\_Shadow(s, o) \equiv \exists scr, l Shadow(s, o, scr, l) \quad (3.2)$$

A partir do Axioma 3.1 e dos axiomas do ROC (seção 2.3) pode-se provar os seguintes teoremas do PQRS propostos em (Santos et al., 2009):

**T1**  $Shadow(s, o, Scr, L) \longrightarrow \neg TotallyOccludes(s, o, L);$

**T2**  $Shadow(s, o, Scr, L) \longrightarrow \neg Shadow(s', s, Scr', L);$



Figura 3.2: Imagens capturadas pelo *PeopleBot*<sup>TM</sup>. 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).

**T3**  $Shadow(s, o, Scr, L) \wedge \exists o' \neg(o = o') \wedge O(region(s), region(o)) \rightarrow (Occludes(o, o', L) \vee Occludes(o', o, L))$ .

O Teorema T1 diz que do ponto de vista da luz, um objeto  $o$  não é ocluído por sua própria sombra  $s$ . O Teorema T2 formaliza o conhecimento de senso comum de que uma sombra não provoca outra sombra. E o teorema T3 diz que, se houver uma região de sombra, dois objetos distintos e a parte da sombra de um dos objetos está sobre parte do outro objeto, então, do ponto de vista da fonte de luz, um objeto oclui o outro. As provas podem ser encontradas em (Santos *et al.*, 2008). Até agora as descrições das relações de oclusão foram feitas do ponto de vista da fonte de luz, porém é possível raciocinar sobre sombras a partir do ponto de vista de um observador. Uma informação que pode ser inferida disso é a localização relativa do próprio observador. Na próxima seção descreve-se como efetuar essa autolocalização.

## 3.2 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. A partir da descrição de oclusão da sombra, cinco regiões foram observadas e definidas pelas linhas de visão entre a fonte de luz ( $L$ ), o objeto ( $O$ ) e sua sombra ( $S$ ), como representado na Figura 3.3(b).

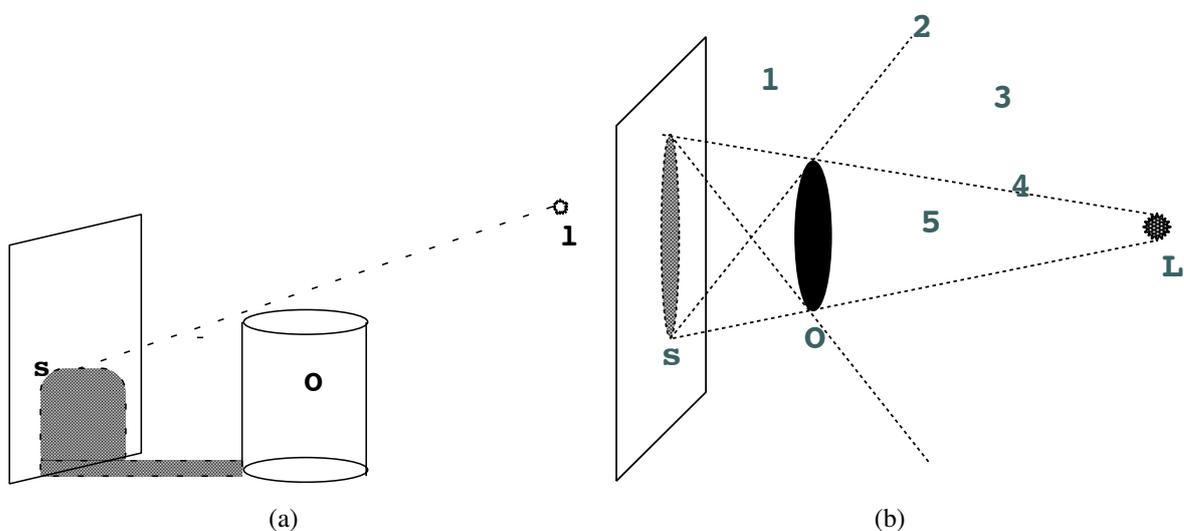


Figura 3.3: (a) Um objeto com sua sombra conectada e projetada numa superfície irregular, destaque para a parte da sombra relativa a parte superior do objeto; (b) Regiões distintas expressas pela observação do objeto e sua sombra, definidas pelas linhas de visão.

Fonte: Santos *et al.* (2008).

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 autolocalização do robô. Para definir formalmente esta localização, assumimos a relação  $located(r, \nu, o, s)$  que representa “um observador  $\nu$  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 (em destaque na Figura 3.3(a)), o termo *Region*  $i$  ( $i \in \{1, 2, 3, 4, 5\}$ ) representa as regiões mostradas na Figura 3.3(b), então temos que (Santos et al., 2009):

$$located(Region\ 1, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge \\ NonOccludesDC(o, s, \nu) \wedge \nu \neq o$$

i.e define a região 1 (Figura 3.4(a)) como a posição do espaço em que  $\nu$  observa que  $o$  não oclui e está desconectado de  $s$ , e  $s$  não é sombra de  $\nu$ ;

$$located(Region\ 2, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge \\ NonOccludesEC(o, s, \nu) \wedge \nu \neq o$$

i.e define a região 2 (Figura 3.4(b)) como a posição do espaço em que  $\nu$  observa que  $o$  não oclui e está externamente conectado de  $s$ , e  $s$  não é sombra de  $\nu$ ;

$$located(Region\ 3, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge \\ PartiallyOccludesPO(o, s, \nu) \wedge \nu \neq o$$

i.e define a região 3 (Figura 3.4(c)) como a posição do espaço em que  $\nu$  observa que  $o$  parcialmente oclui e sobrepõe  $s$ , e  $s$  não é sombra de  $\nu$ ;

$$located(Region\ 4, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge \\ TotallyOccludesTPPI(o, s, \nu) \wedge \nu \neq o$$

i.e define a região 4 (Figura 3.4(d)) como a posição do espaço em que  $\nu$  observa que  $o$  oclui totalmente  $s$  e este é parte tangencial de  $o$ , e  $s$  não é sombra de  $\nu$ ;

$$located(Region\ 5, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge \\ TotallyOccludesNTPPI(o, s, \nu) \wedge \nu \neq o$$

i.e define a região 5 (Figura 3.4(e)) como a posição do espaço em que  $\nu$  observa que  $o$  oclui totalmente  $s$  e este é parte não tangencial de  $o$ , e  $s$  não é sombra de  $\nu$ .

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, deslocando o ponto de vista do observador da região 1 até a região 5, com fotos capturadas pelo *PeopleBot<sup>TM</sup>* no laboratório, focalizando o objeto e a sua sombra projetada na parede.

Na Figura 3.4(a) 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 (Figura 3.4(b)). A região 2 é a superfície definida pelas linhas que conectam o lado do objeto com o lado oposto da sua sombra (Figura 3.3(b)), do ponto de vista de  $\nu$  vê-se a sombra  $s$  completa porém conectada ao objeto.

Na Figura 3.4(c) o observador está posicionado no que chamamos de região 3, pois a sombra está parcialmente oclusa pelo objeto, 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 desaparece. Note pela Figura 3.4(d) 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 3.4(e)), ou reaparecerá se estiver indo na direção da região 3 (Figura 3.4(c)), o que caracteriza a relação *TotallyOccludesTPPI*. Na região 5 estamos no campo de visão da luz, nesta região a sombra está projetada numa superfície totalmente ocluída pelo objeto e, portanto, não é possível vê-la (Figura 3.4(e)).

Fora demonstrado como as relações se apresentam dependendo da região, contudo devido a configuração de fronteira das regiões 2 e 4, sua observação em ambiente real é difícil. Uma adaptação foi necessária na aplicação prática, e será detalhada na seção 4.3. Até este ponto, a teoria considera somente um objeto em cena, vamos ver a seguir cenas com múltiplos objetos.

### 3.3 Sombras e percepção de profundidade

Sombras são importantes indicativos para profundidade assim como oclusão. A relação de vizinhança representada por  $N(x, y, z)$  lida como “ $x$  está mais próximo de  $y$  do que de  $z$ ” pode se integrar ao ROC (seção 2.3), relacionando a oclusão com distância relativa, conforme o axioma a seguir:

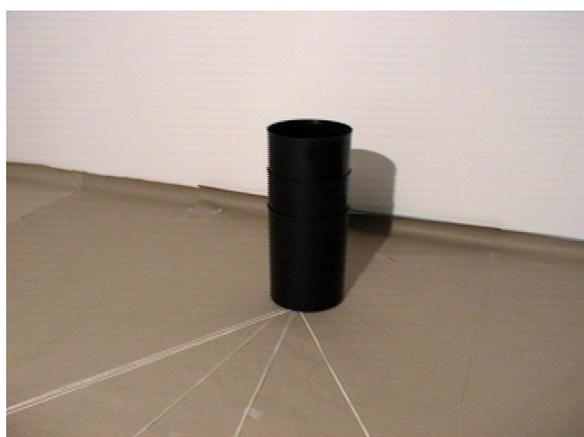
$$PartiallyOccludes(x, y, \nu) \longrightarrow N(\nu, x, y). \quad (3.3)$$



(a)



(b)



(c)



(d)



(e)

Figura 3.4: 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

Esse axioma representa que “se um corpo  $x$  oclui parcialmente  $y$  da perspectiva do observador  $\nu$ , então  $\nu$  está mais próximo de  $x$  do que está de  $y$ ”. Ao observar o axioma 3.1, o axioma 3.3 e partindo do fato de  $N(L, o, s)$  ter como consequência  $N(L, o, Scr)$ , tanto sombra quanto a superfície em que ela se projeta estão mais distantes da fonte de luz do que o objeto. Portanto pode-se dizer que se a sombra  $s$  do objeto  $o$  é projetada em um objeto  $o'$  (com  $o \neq o'$ ), então  $o$  estará mais próximo da fonte de luz do que  $o'$ , ainda que  $o'$  não seja diretamente percebido pelo observador. Formalmente temos:

$$Shadow(s, o, o', L) \longrightarrow N(L, o, o'). \quad (3.4)$$

Com a informação sobre qual objeto está mais próximo da luz, e a informação de localização relativa do observador, é possível inferir profundidade, ou seja, a distância relativa entre os objetos e o observador (*Santos et al.*, 2009), assim

$$N(\nu, o', o) \longrightarrow Located(Region1, \nu, o, s) \wedge Shadow(s, o, o', l);$$

$$N(\nu, o, o') \longrightarrow Located(Regioni, \nu, o, s) \wedge \\ Shadow(s, o, o', l) \wedge (i = 2 \vee i = 3 \vee i = 4)$$

Note que a relação de proximidade é uma relação ternária, a localização do observador é dada em relação ao objeto  $o$  que tem sua sombra projetada em outro objeto  $o'$ . Consequentemente quando o observador está localizado na região 1, ele encontra-se mais distante  $o$  do que de  $o'$ , enquanto que nas regiões 2, 3 e 4 o observador está mais próximo de  $o$  do que de  $o'$ .

Toda a teoria do PQRS é baseada no fato de que as sombras são corretamente relacionadas. Portanto relacionar a 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.

### 3.4 Correspondência entre objeto e sombra

Com base no PQRS será definida a correspondência entre o objeto e sua sombra. Uma sombra pode estar ou não conectada ao objeto que a projetou, um objeto suspenso no ar tem a sua sombra projetada desconectada do mesmo, contudo este caso será desconsiderado neste trabalho, onde o que se buscou fora a formalização de situações em um ambiente de laboratório.

Uma das hipóteses simplificadoras em nossos experimentos é que o objeto está imóvel sobre o chão. Logo, a sombra sempre estará conectada ao objeto que a provocou, assim sendo, nas imagens as regiões que representam o objeto e sua sombra estão externamente conectadas. Então, para formalizar a teoria, vamos definir antes quais as características que devem ser observadas para aplicá-la.

O axioma 3.2 garante que somente existe sombra  $s$  se houver o objeto  $o$ , uma superfície  $scr$  e a fonte de luz  $l$ , além destas restrições outras devem ser observadas, são elas:

- 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).

Já comentou-se que o objeto sempre está externamente conectado ( $EC$ ) com sua sombra. É, portanto, conveniente representar o objeto em base e topo. Definindo  $Base(x, z, \nu)$  lida como “ $x$  é base de  $z$  do ponto de vista de  $\nu$ ” e  $Top(y, z, \nu)$  que lê-se “ $y$  é topo de  $z$  do ponto de vista de  $\nu$ ”.  $Base/3$  obedece os seguintes axiomas:

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

$$\forall x y \nu Base(x, y, \nu) \rightarrow \exists z (x \neq z) TPP(region(x), region(y)) \wedge TPP(region(z), region(y)) \wedge EC(region(x), region(z)) \rightarrow Top(z, y, \nu)$$

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

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

$$\forall x y \nu Base(x, y, \nu) \rightarrow [\exists z [P(region(z), region(x)) \wedge Base(z, y, \nu)] \wedge \neg \exists u [P(region(u), region(x)) \wedge Top(u, y, \nu)]]$$

$$\forall x y z \nu Base(x, y, \nu) \wedge Top(z, y, \nu) \rightarrow P(image(x, \nu), image(y, \nu)) \wedge P(image(z, \nu), image(y, \nu)) \wedge EC(image(x), image(z))$$

Para definir que determinada sombra  $s$  corresponde ao objeto  $o$  utilizamos o axioma abaixo:

$$Shadow(s, o, Scr, L) \rightarrow N(L, o, s) \wedge \exists b \nu Base(b, o, \nu) \wedge NTPP(s, Scr) \wedge NTPP(b, Scr) \wedge EC(s, b) \quad (3.5)$$

O axioma 3.5 diz que, estando a sombra  $s$  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 tanto  $b$  quanto  $s$  estiverem sobre a mesma superfície, então  $s$  é sombra de  $o$ . Essa formalização é suficiente para corresponder a sombra com seu objeto nas situações consideradas neste trabalho.

Nas figuras 3.5 podem ser observadas as restrições do axioma 3.5, nelas a superfície onde pelo menos parte da sombra e a base do objeto estão, é a mesma (o piso). Na figura 3.5(a), cada objeto tem somente uma sombra externamente conectada à sua base e esta encontra-se mais distante da fonte de luz. Já na Figura 3.5(b) um dos objetos tem duas sombras conectadas à sua base, neste caso o axioma de proximidade  $N(L, o, s)$  servirá para confirmar qual a sombra é realmente do objeto. Na Figura 3.5(c) pode-se ver um outro tipo ambiguidade, nela um dos objetos tem três sombras conectadas a ele, destas somente a sombra que ele provocou é verdadeira para o axioma 3.5. 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 pode-se deduzir a oclusão (Figura 3.6).

### 3.5 Discussão final

O PQRS é uma teoria em desenvolvimento. Situações em que há mais de uma fonte de luz ou objetos flutuantes ainda não foram analisadas. Nossa solução para o problema de correspondência é uma solução parcial, ainda assim, para o ambiente de teste preparado em laboratório, essa solução mostrou-se satisfatória, abrangendo todas as possibilidades de correspondência que foram observadas. Acredito que essa solução possa ser utilizada em várias outras situações, nas quais o objeto esteja apoiado em uma superfície.

O PQRS foi a base teórica para o desenvolvimento do sistema utilizado pelo *PeopleBot<sup>TM</sup>*. Assim os algoritmos e funções foram desenvolvidos especificamente sobre os axiomas desta teoria lógica, apesar de não termos utilizados diretamente uma linguagem em lógica de primeira ordem, seguir a teoria serviu para deixar nosso sistema mais confiável em termos de respostas. Os experimentos terão como objetivo testar algumas hipóteses levantadas por esta teoria, todas relativas a relações percebidas entre sombra e objeto. As hipóteses testadas são:

- A identificação da posição relativa do observador, de modo qualitativo, a partir do observação das sombras, definindo em qual a região este se encontra.

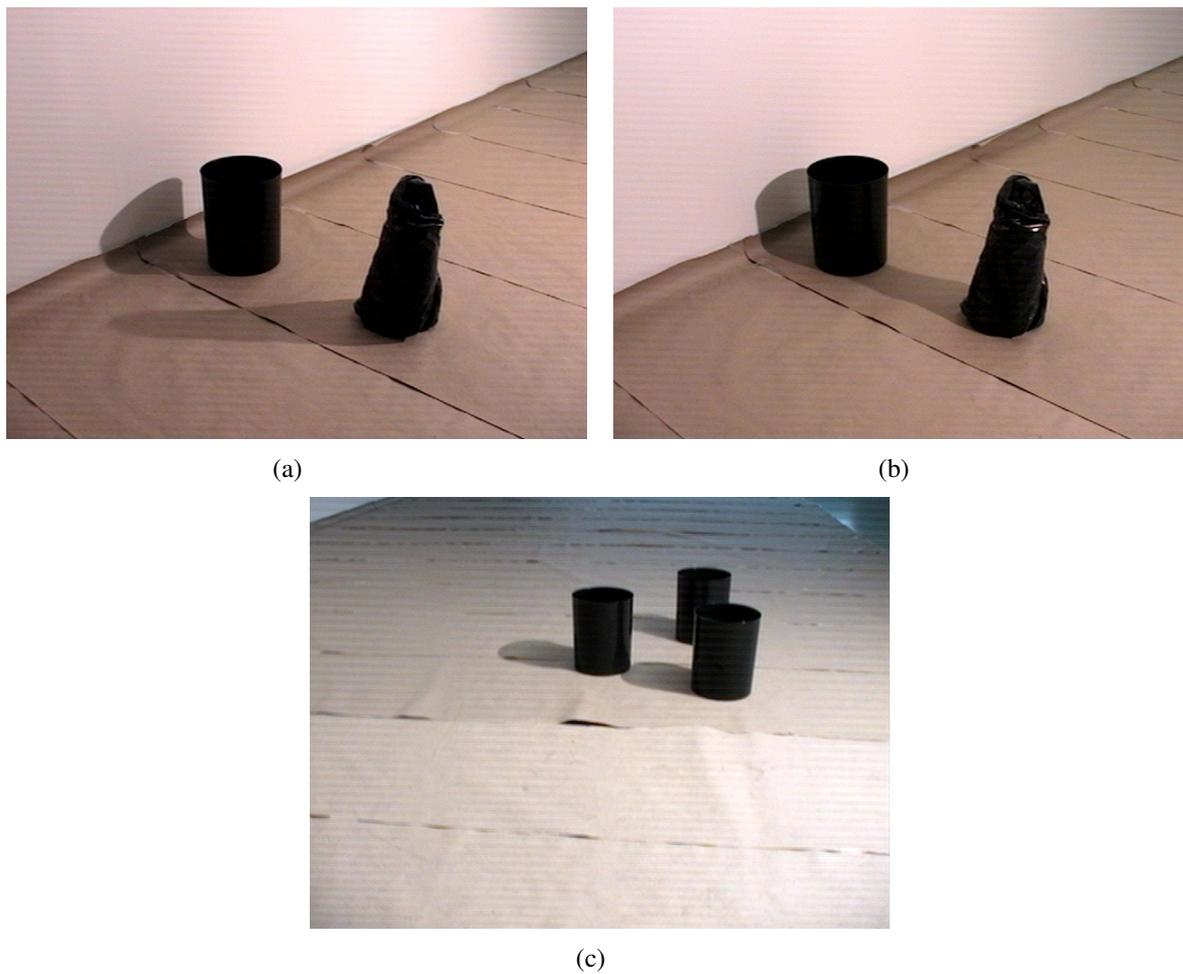


Figura 3.5: Imagens capturadas no laboratório, onde podemos aplicar o axioma de correspondência de sombra, note somente a sombra do objeto está externamente conectada a sua base e mais distante da fonte de luz.



Figura 3.6: Nesta figura podemos inferir que há outro objeto ocluído totalmente pelo balde.  
Fonte: Santos *et al.* (2008).

- A determinação da profundidade relativa entre pares de objetos, ou seja, qual objeto está mais próximo do observador.

Nos próximos capítulos descreveremos como serão realizados os experimentos e os resultados obtidos.

## Capítulo 4

# DESENVOLVIMENTO DO SISTEMA

Neste trabalho foi desenvolvido um sistema que utiliza sombra como fonte de informação qualitativa, provendo um agente robótico com a capacidade de raciocinar sobre a percepção de sombras no ambiente. Esse sistema abrange visão computacional, controle robótico e raciocínio espacial qualitativo. Tanto no controle robótico quanto na visão computacional foram utilizadas técnicas e métodos de bibliotecas de programas de livre acesso. O foco principal do presente desenvolvimento foi o raciocínio espacial qualitativo. Assim, a contribuição deste trabalho é validar a teoria lógica sobre sombras, então o raciocínio espacial qualitativo terá como base o conhecimento representado nos axiomas do PQRS. Neste capítulo será descrito, com auxílio de fluxogramas, como todo o sistema foi desenvolvido (seção 4.1), terá uma descrição do ambiente e dos componentes utilizados na seção 4.2. Na seção 4.3 está descrito como as relações, entre a sombra e objeto, se apresentam para cada região utilizada nos testes. Na seção 4.4 encontra-se uma descrição do software desenvolvido e os pseudocódigos das funções principais.

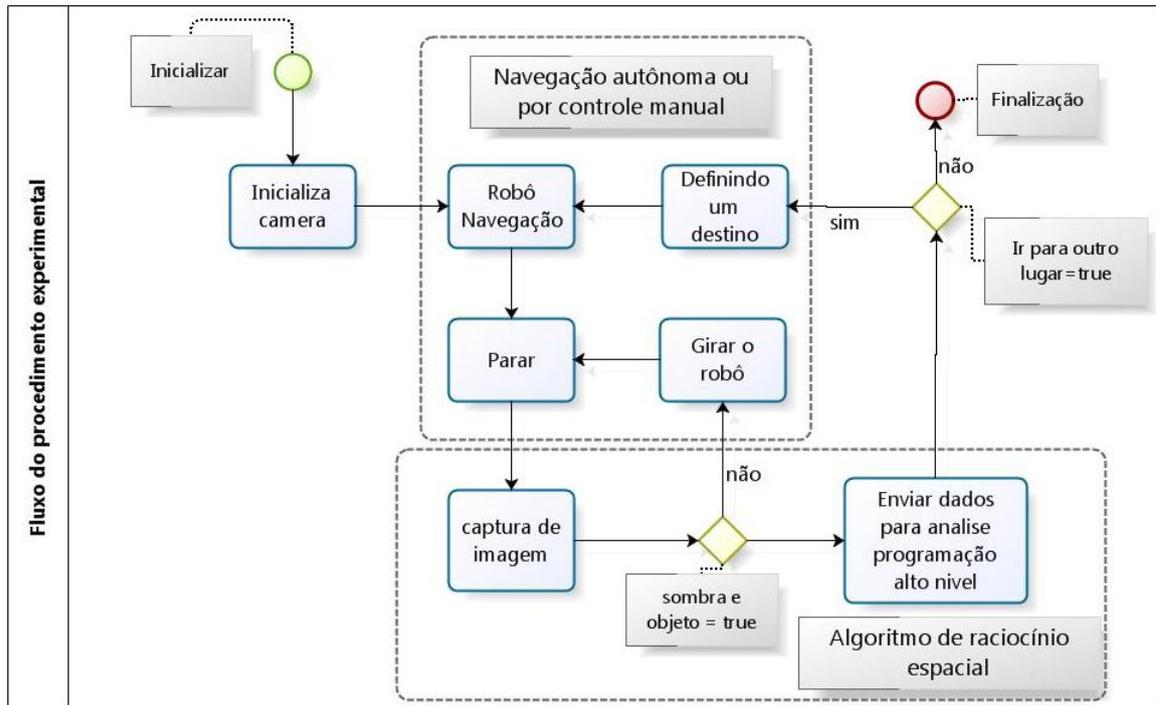
### 4.1 Fluxo de dados do sistema

Em linha gerais o *PeopleBot<sup>TM</sup>* está equipado com uma câmera. Ele irá navegar pelo ambiente, conforme representado pelo fluxograma <sup>1</sup> da Figura 4.1, capturando e analisando as imagens obtidas, utilizando o PQRS como base para o algoritmo de raciocínio espacial, e a sua tarefa será retornar uma resposta sobre os seguintes itens:

- correspondência entre objeto e sua sombra;
- autolocalização;

---

<sup>1</sup>*BizAgi Process Modeler* software livre utilizado para confecção dos fluxogramas.



powered by  
BizAgj  
Process Modeler

Figura 4.1: Fluxograma representando o procedimento dos experimentos realizados com o *PeopleBot<sup>TM</sup>*.

- proximidade relativa a outros objetos.

A seguir o sistema será descrito em partes, começando pelo sistema de visão computacional.

#### 4.1.1 Fluxograma de visão computacional

No fluxograma da Figura 4.2 é apresentado o algoritmo de detecção de sombra e objeto, executados pelo sistema em visão computacional. Ele começa pela captura da imagem, depois essa imagem é transformada de RGB para HSV. Nos canais de **S** e **V** são aplicados filtros de limiares, separando os *pixels* com alta intensidade daqueles com baixa. As altas e baixas intensidades são substituídos por zero em **V** e **S** respectivamente. Note que com este filtro de limiar permanecem regiões que contém sombras, mas não necessariamente objetos, essas passam a ser as regiões de interesse (ROI).

A região do objeto na imagem tem nível de saturação alta, então nas mesmas coordenadas de regiões ROI em **V**, verifica-se em **S**, a proporção de *pixel* com alta saturação em cada região. As regiões que apresentam uma saturação muito baixa são descartadas. Assim, a imagem de intensidade **V** somente manterá as regiões que contém objeto e sombra. A partir desta imagem

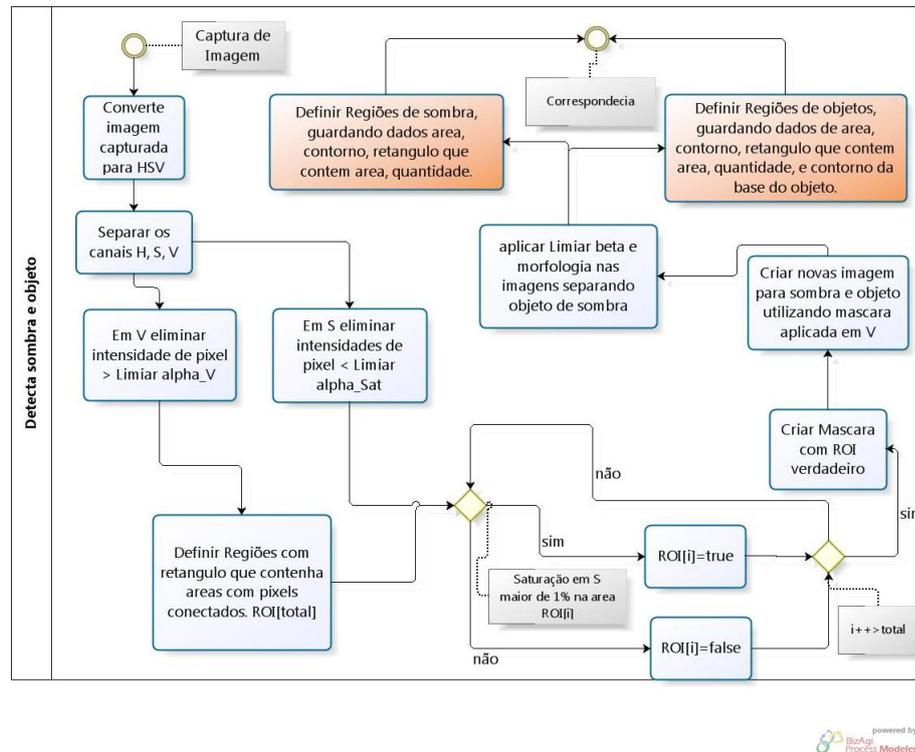


Figura 4.2: Fluxograma de detecção de sombra e objeto

aplica-se um novo limiar para criar novas imagens separando sombras e objetos.

Das imagens de objetos e de sombras são extraídas características, como: área, coordenadas do menor retângulo que contém toda a área, coordenadas do contorno (bordas), quantidade de sombras e quantidade de objetos encontrados, esses atributos serão utilizados para analisar correspondência entre sombra e objeto.

#### 4.1.2 Fluxograma de correspondência

Após localizadas as sombras e os objetos, a próxima etapa do algoritmo de raciocínio é decidir a correspondência entre eles, ou seja, fazer a ligação entre o objeto e a sua sombra. Com base no axioma 3.5 é retirada uma nova informação de base de todos os objetos. Então, verifica-se para cada sombra, a qual base de objeto ela está conectada, confirmando com a proximidade do objeto da luz ( $N(L, o, s)$ ) a correspondência entre eles. O fluxograma do processo de correspondência é mostrado na Figura 4.3. Quando ocorre a conexão de uma sombra com mais de um objeto, as informações destas conexões serão guardadas para que sejam avaliadas posteriormente pelo algoritmo de proximidade. Após completar a correspondência segue-se para a etapa de autolocalização.

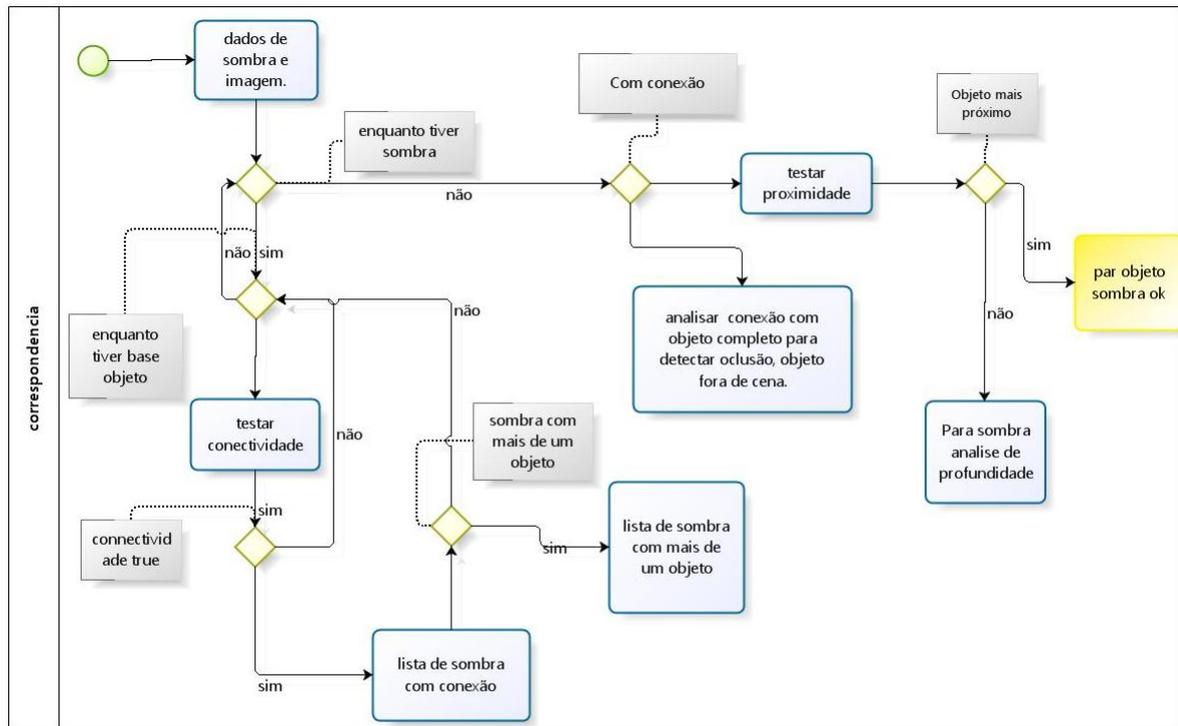


Figura 4.3: Fluxograma de correspondência entre objeto e sua sombra

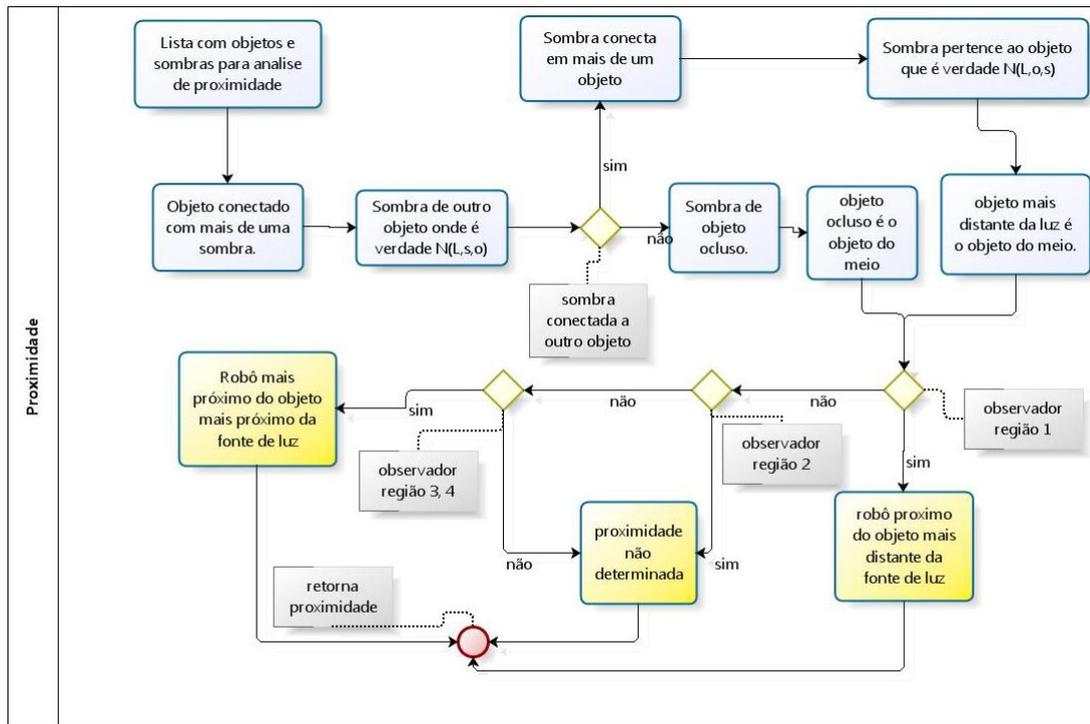
### 4.1.3 Fluxograma da autolocalização

Depois de encontrada a sombra correspondente de cada objeto, o próximo passo é determinar qual tipo de relação existe entre eles, e conseqüentemente a região onde está localizado o observador. O fluxograma da Figura 4.4 mostra como se procede na autolocalização, primeiro verifica-se a existência de sombra e objeto, na falta do objeto ou de ambos ao mesmo tempo nada é determinado, havendo somente objeto sem sombra o observador encontra-se na região 5, com ambos, sombra e objeto presentes na cena, deve ser realizada uma verificação de suas relações, para então determinar em qual região o observador se encontra. Após determinar a posição do *PeopleBot<sup>TM</sup>* (observador), verifica-se a proximidade dos objetos.

### 4.1.4 Fluxograma de proximidade

A proximidade do objeto só pode ser determinada em casos onde: a sombra de um objeto é projetada sobre outro objeto, ou seja, quando essa sombra aparece na imagem conectada a outro objeto. O fluxograma da Figura 4.5 analisa as sombras e objetos que apresentam (na imagem) mais de uma conexão externa. Neste sistema, para os casos em que a imagem apresenta todos





powered by  
BizAgil  
Process Modeler

Figura 4.5: Fluxograma de proximidade



Figura 4.6: Laboratório onde os experimentos preliminares foram realizados.

Fonte: Santos et al. (2008).

importante neste ambiente é a iluminação.

Para que a sombra seja bem definida, o ambiente terá uma fonte de luz predominante. Como fonte de luz utilizamos uma lâmpada halógena de 1000W localizada em tripé, o que garante

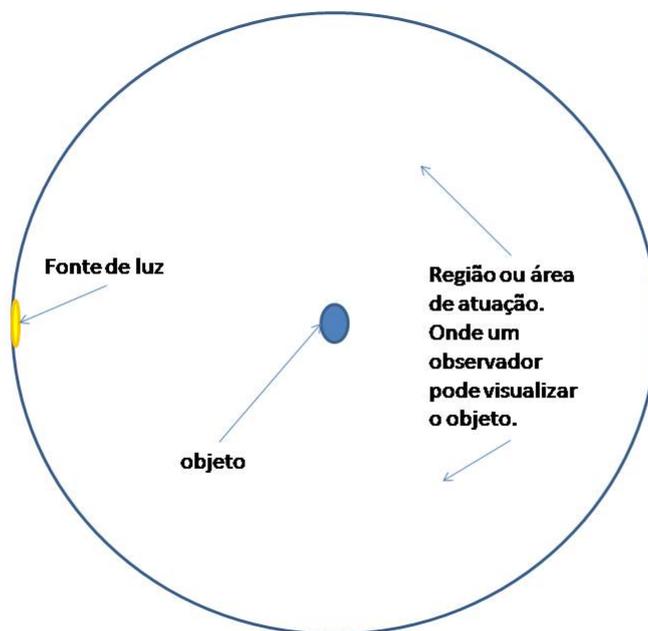


Figura 4.7: Representação da área em que o robô atuará neste trabalho, sua área de atuação ou região de atuação.

colocá-la em posições distintas a nossa escolha, entretanto a fonte de luz estará sempre acima do objeto, tanto quanto do observador (para garantir um ambiente cognitivamente relevante (Mamassian and Goutcher, 2001)).

Dentro desse ambiente vamos definir a região na qual o robô atuará, a Figura 4.7 é uma representação da vista superior desta área de atuação, note que o objeto está posicionado ao centro da área circular, e que a distância entre a fonte de luz e o objeto é o raio do círculo, todos os experimentos foram realizados com o robô circulando dentro desta área. Veremos agora uma descrição do observador.

No papel de observador iremos utilizar um robô *ActivMedia* tipo *PeopleBot<sup>TM</sup>* com um processador *Pentium III*, 512MB de memória, velocidade 533hz e com uma única câmera acoplada, a qual é responsável pela captura de imagem do ambiente. Esse robô pode ser utilizado com uma programação de alto nível, pois possui uma plataforma *ActivMedia Robotics Interface for Applications* (ARIA) que é uma biblioteca de controle para o *PeopleBot<sup>TM</sup>*, portanto, com essa biblioteca, o controle necessário para as ações do robô já estão desenvolvidas. O ARIA é um ambiente de desenvolvimento com código aberto baseado em C++, essa característica foi importante para o desenvolvimento do sistema deste trabalho, pois os comandos relacionados às ações foram todos baseados e adaptados do ARIA.

O nosso algoritmo está sendo desenvolvido em C++, pois além do ARIA que controla o

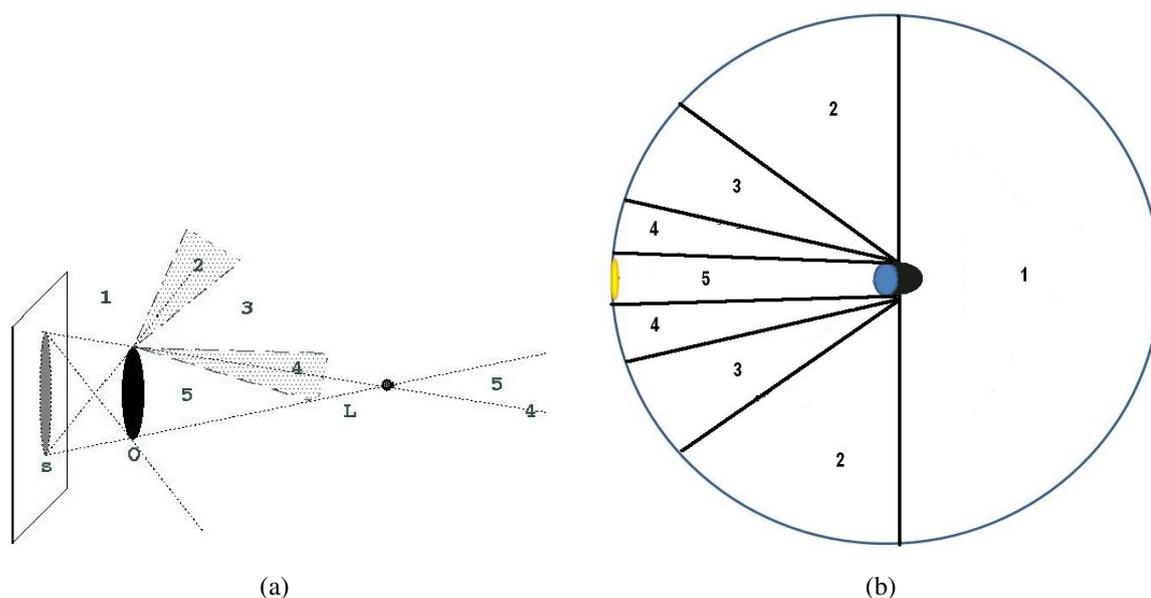


Figura 4.8: Representação espacial (a) das regiões considerando um ambiente real, e (b) vista superior de uma representação das regiões no limite da área de atuação do robô.

Fonte:(a) Santos et al. (2008).

robô, o sistema de visão computacional foi todo desenvolvido com comandos da biblioteca *Open Source Computer Vision Library (OpenCV)*, que é baseado na mesma plataforma C++. A seguir descreveremos como as regiões estão definidas dentro da área de atuação do robô.

### 4.3 Definindo as regiões

A teoria demonstrou que, em relação à fonte de luz, o espaço em volta do objeto pode ser dividido em cinco regiões (Capítulo 3). Para utilizar o conceito de regiões na autolocalização do robô em ambiente real, precisamos executar algumas modificações (Figura 4.8). As regiões 2 e 4 são regiões de fronteira conforme (Santos et al., 2008), por causa desta característica fica difícil observá-las com um robô em ambiente real. Uma solução seria desconsiderar tais fronteiras, contudo preferimos assumir uma “região de incerteza” em um entorno de 2 e 4 conforme representado pelas áreas hachuradas na Figura 4.8(a). Outra característica a ressaltar é que as relações sombra-objeto do ponto de vista do observador permanecem existindo mesmo fora do círculo de atuação, ou seja as regiões se estendem para além da fonte de luz (Figura 4.8(a)). A seguir será utilizado as imagens, de um objeto fixo, capturadas pelo *PeopleBot<sup>TM</sup>* para demonstrar como cada região foi definida em nosso laboratório

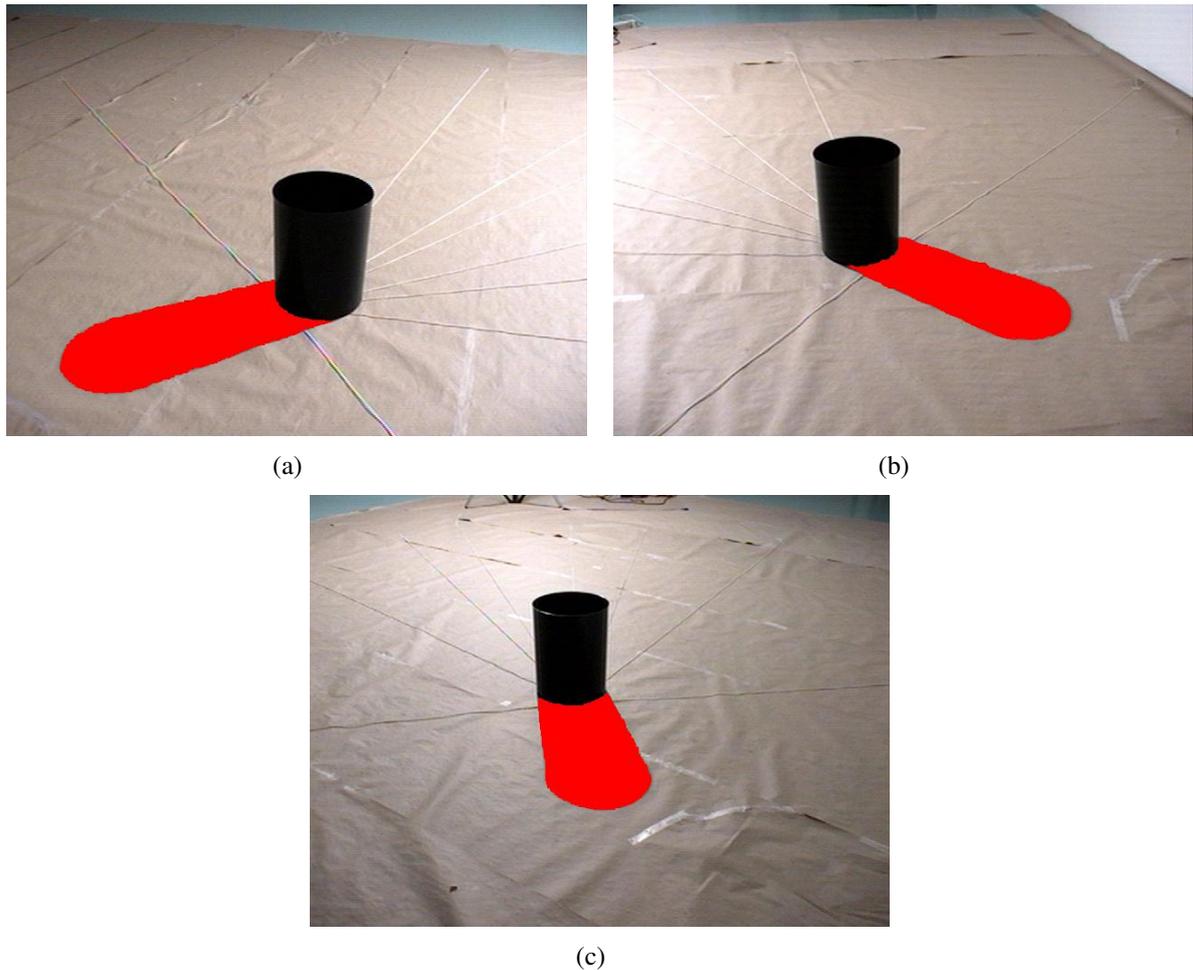


Figura 4.9: Imagens capturadas com o robô posicionado na região 1 à (a) esquerda, (b) direita, e (c) em frente do objeto. Note que o objeto não oclui a sombra.

### Região 1

A região 1 é definida na teoria pelo axioma  $located(Region\ 1, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge NonOccludesDC(o, s, \nu) \wedge \nu \neq o$ . Esta região consiste da maior parte da área de atuação do robô (Figura 4.8(b)), nela o objeto está sempre entre a fonte de luz e o robô. Do ponto de vista do observador a sombra nunca é ocluída pelo objeto que a projetou, não importando se a observação é realizada pelo lado esquerdo (Figura 4.9(a)), direito (Figura 4.9(b)) ou na frente (Figura 4.9(c)) do objeto. Na imagem (Figura 4.9) existe uma conexão externa entre as regiões da base do objeto e a sua sombra. A área da região da sombra é maior que 90% da área do objeto, enquanto que no eixo  $y$  da imagem, a região da sombra está sempre posicionada abaixo da região do objeto.

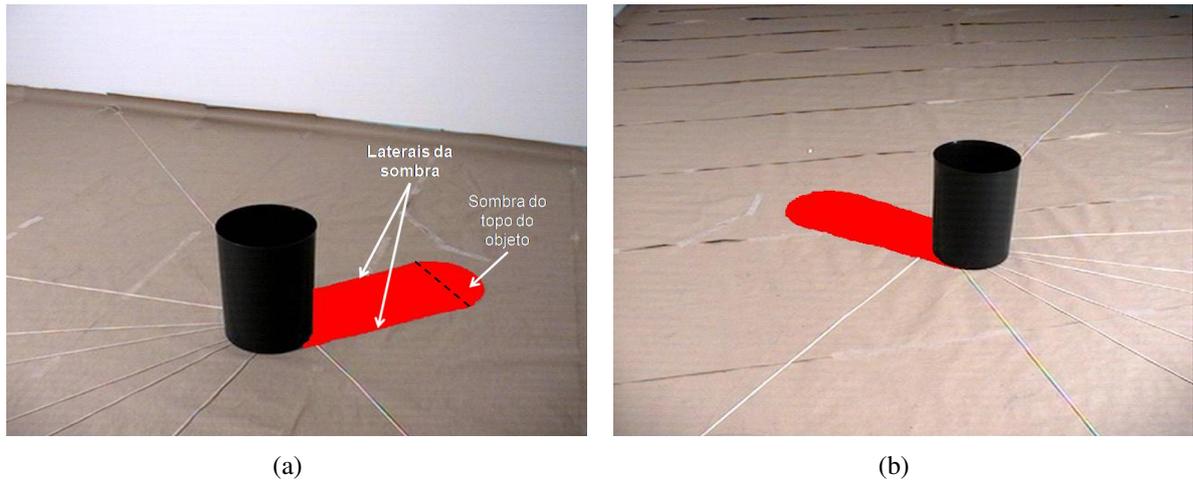


Figura 4.10: Imagens capturadas com o robô na região 2. Note que o objeto começa a ocluir a sombra. Todavia ainda se vê a sombra do topo do objetivo e as laterais da sombra do objeto indicados em (a). Robô posicionado (a) à direita e (b) à esquerda do objeto

### Região 2

A região 2 é definida na teoria pelo axioma  $located(Region\ 2, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge NonOccludesEC(o, s, \nu) \wedge \nu \neq o$ . Esta região é vizinha da região 1 e representa (na região de atuação) a segunda maior região, considerando-se a soma das regiões 2 esquerda e direita (Figura 4.8(b)), nela o robô está posicionado entre o objeto e a fonte de luz. Do ponto de vista do observador a maior parte da sombra pode ser percebida, porém a oclusão desta pelo seu objeto começa a ocorrer, enquanto o observador estiver na região 2 ambas as laterais do objeto podem ser observadas na sombra projetada (uma das laterais aparece parcialmente). Na imagem (Figura 4.10) também se percebe as duas laterais da sombra, e o tamanho da sua área fica entre 50% e 90% do tamanho da área do objeto, isso devido à oclusão que começa a ocorrer. Quanto as posições em relação ao eixo  $y$  da imagem ambos base do objeto e sombra encontram-se aproximadamente no mesmo nível, deste modo a região da sombra que representa a sombra do topo do objeto encontra-se totalmente desconectada do objeto (Figura 4.10(a)). Na região 2 a única conexão externa entre objeto e sombra ocorre na base do objeto.

### Região 3

A região 3 é definida na teoria pelo axioma  $located(Region\ 3, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge PartiallyOccludesPO(o, s, \nu) \wedge \nu \neq o$ . Na região de atuação a região 3 fica localizada entre as regiões 2 e 4. O robô nesta região está entre o objeto e a fonte de luz, assim como a região 2, a região 3 aparece de ambos os lados da linha reta objeto-fonte de luz. Do ponto de vista do observador, a sombra está ocluída parcialmente pelo objeto e somente uma das laterais da

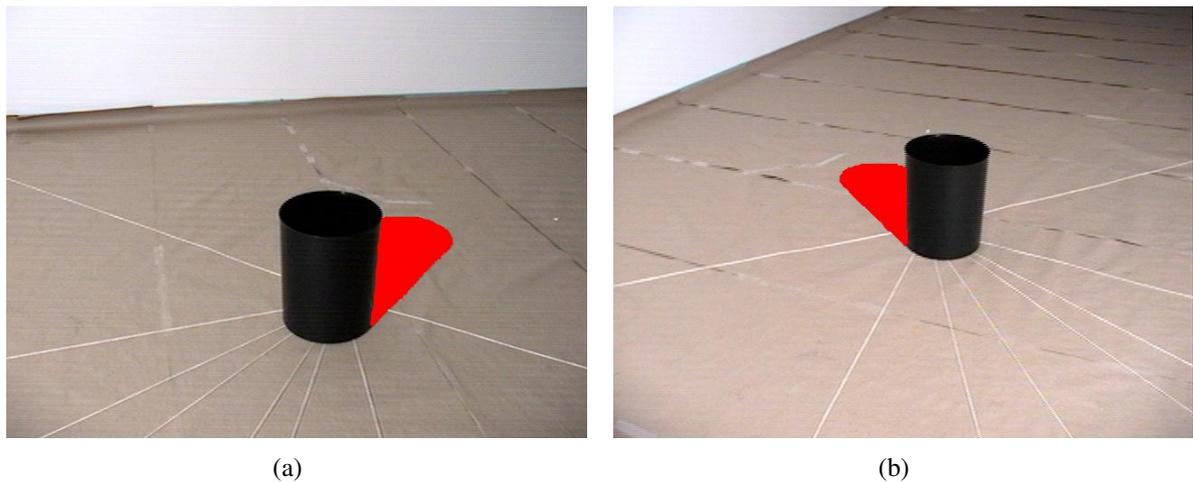


Figura 4.11: Imagens capturadas com o robô na região 3. A sombra nesta região aparece com um dos lados totalmente ocluído pelo objeto. Existe uma conexão entre a parte superior do objeto e sua sombra projetada. Robô (a) à direita e (b) à esquerda do objeto.

sombra não pode ser observada. Na imagem (Figura 4.11) a área da sombra é menor que 50% da área do objeto. Nota-se que, devido ao posicionamento da sombra, a conexão externa entre sombra e objeto aumentou, agora além da base do objeto, o topo e a sombra do topo do objeto também estão externamente conectados.

#### Região 4

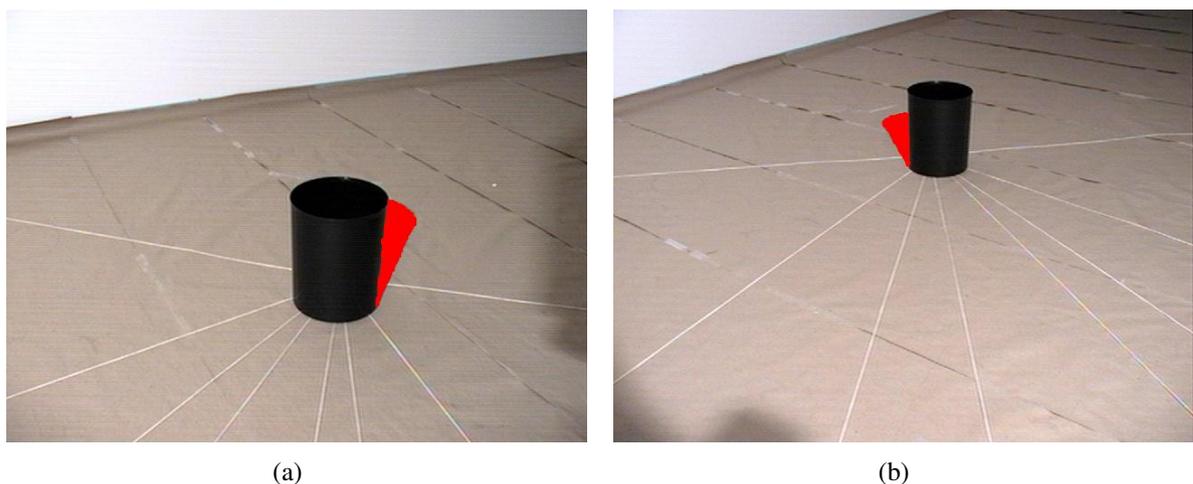


Figura 4.12: Imagens capturadas com o robô na região 4. A sombra nesta região quase desaparece. Contudo ainda existe uma conexão entre a parte superior do objeto e a sombra. Robô (a) à direita e (b) à esquerda do objeto.

A região 4 é definida na teoria pelo axioma  $located(Region\ 4, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge TotallyOccludesTPPI(o, s, \nu) \wedge \nu \neq o$ . Esta região é uma das menores regiões da área de atuação (Figura 4.8(b)) e está localizada nos dois lados da reta objeto-fonte de luz, quando na região 4 o robô está posicionado entre o objeto e a sombra. Do ponto de vista do observador,

pequena parte da sombra do objeto é percebida. Na imagem (Figura 4.12) a conexão entre as áreas da sombra e do objeto é semelhante a que ocorre quando na região 3, porém com a área da sombra menor, como pode ser visto nas imagens da Figura 4.12.

### Região 5



Figura 4.13: Imagem capturada com o robô na região 5. A sombra não aparece na imagem por estar totalmente oclusa. O observador está na região equivalente ao ponto de vista da luz.

A região 5 é definida na teoria pelo axioma  $located(Region\ 5, \nu, o, s) \leftarrow Is\_a\_Shadow(s, o) \wedge TotallyOccludesNTPPI(o, s, \nu) \wedge \nu \neq o$ . Esta região está disposta em uma única região entre o objeto e a fonte da luz (Figura 4.8(b)), onde a sombra está totalmente oclusa pelo objeto. Do ponto de observador, a sombra existe mas não pode ser percebida, pois o ponto de vista do observador é semelhante ao ponto de vista da fonte luz. Na imagem (Figura 4.13) vemos somente o objeto, embora dependendo da posição do observador, a única sombra que ele percebe é a projeção da sua própria sombra, esses casos não serão tratados neste trabalho (Figura 4.13). Com a definição das regiões no laboratório, a descrição do software de raciocínio espacial desenvolvido para a autocalização, o qual será utilizado no *PeopleBot<sup>TM</sup>*, será apresentada na próxima seção.

## 4.4 Funcionamento do software de autocalização

O software foi desenvolvido em C++ utilizando a biblioteca *OpenCV(Intel)* para os processos de visão computacional e a biblioteca *ARIA* para comandos de controle do *PeopleBot<sup>TM</sup>*. Ao carregar o software, a janela vista na Figura 4.14 é aberta. Antes de inicializar a autocalização, os parâmetros de visão computacional devem ser calibrados. Quando o botão de

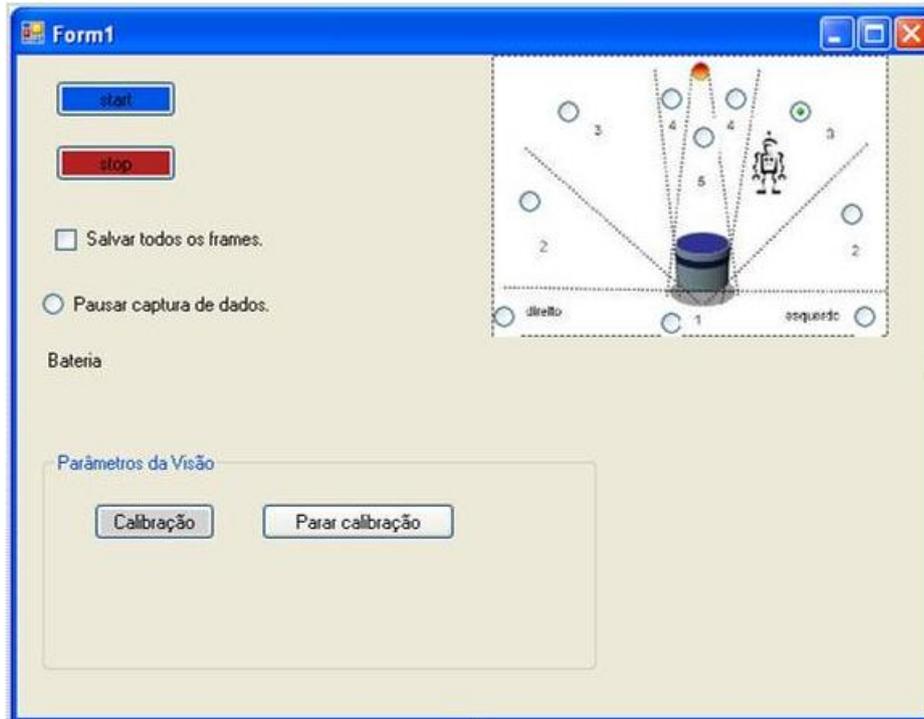


Figura 4.14: Janela de controle do software de raciocínio *PeopleBot<sup>TM</sup>*.

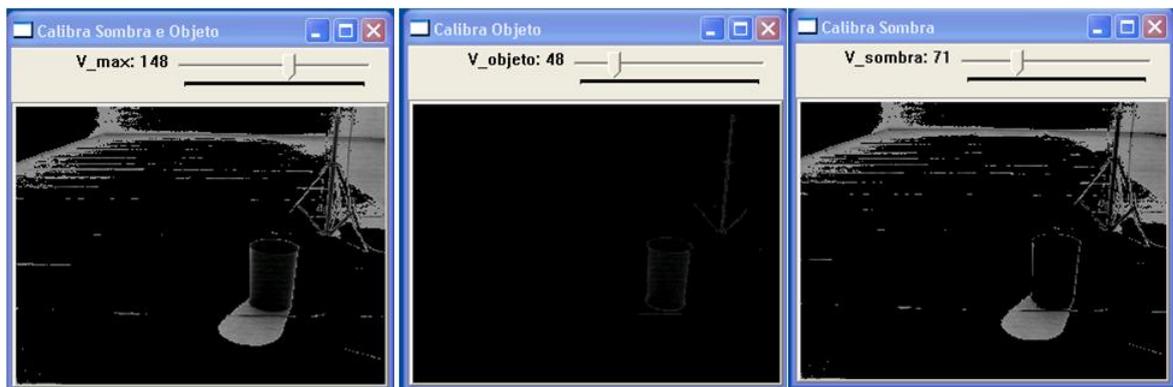


Figura 4.15: Janelas para calibração de limiar de corte para segmentar objeto e sombra, da esquerda para direita, corta altas intensidades, corta acima do limiar indicado fica objetos escuros, corta abaixo do limiar indicado permanece sombras.

calibração é acionado, três janelas aparecem e nelas a imagem do canal  $V$  capturadas pelo robô, como na Figura 4.15. Cada uma destas janelas tem uma barra, onde é possível variar os valores de 0 a 255, com elas alteram-se os limiares que separam sombra e objeto, depois de calibrado a autocalibração pode ser inicializada clicando no botão de início (start).

Quando o botão de início é acionado, a câmera do *PeopleBot<sup>TM</sup>* é ativada. Se em modo autônomo o *PeopleBot<sup>TM</sup>* navega aleatoriamente pelo o ambiente por alguns segundos e ao parar captura a imagem e faz a análise de sua localização. Quando o controle de navegação é realizado por um operador, este usando um *joystick* faz com que o robô movimente-se pela região da atuação, enquanto que a autocalibração é realizada simultaneamente. Para compara-

```
Peoplebot_Sistema  
Variaveis:  
  ImgCap: Imagem;  
  Dados: Estrutura de informação;  
begin  
  while (Ligado) do  
    ImgCap ← Imagem capturada pelo sensor (camera)  
    if (habilitado para testes de análise de resultados) then  
      Salva cópia de ImgCap;  
    end if  
    Detecta sombra e objeto(ImgCap);  
    // com os dados retirados da Visão  
    Análise de Correspondencia (Dados);  
    Autocalização(Dados de correspondência)  
    Proximidade(Lista de proximidade);  
    Salva dados em arquivo  
    return Imagem mostrando segmentação e localização  
  end while  
end
```

Figura 4.16: Pseudocódigo do algoritmo principal utilizado no *PeopleBot<sup>TM</sup>*

ções posteriores com a resposta de autocalização dada pelo robô, ao lado direito superior da janela (Figura 4.14) existe uma figura onde deve ser marcada, pelo operador, a região real em que o robô está posicionado. Este dado e outros relativos a cada quadro são guardados para análise de resultados, para testes posteriores em modo *off-line*. Os quadros capturados são arquivados como imagens. Na Figura 4.16 apresenta-se o pseudocódigo geral do sistema.

A título de visualização dos resultados em tempo real, enquanto o software está sendo executada, duas janelas se abrem, uma mostra a imagem capturada, na outra aparece uma imagem semelhante, na qual os objetos e sombras encontrados na cena são indicados por retângulos. A correspondência entre eles é indicada pelo número escrito no canto superior esquerdo destes retângulos, são escritos “L” e “R” nos objetos que respectivamente estiverem mais próximo da fonte de luz e do robô, no canto superior esquerdo da imagem vê-se a resposta dada pelo software, quanto a região onde o *PeopleBot<sup>TM</sup>* está localizado (Figura 4.17).

#### 4.4.1 Algoritmo de detecção de sombra e objeto.

O pseudocódigo mostrado na Figura 4.18 refere-se ao sistema de visão computacional responsável pela detecção de sombras e objetos. A função de detecção foi desenvolvida especialmente para este trabalho, levando em consideração as restrições impostas nos experimentos.

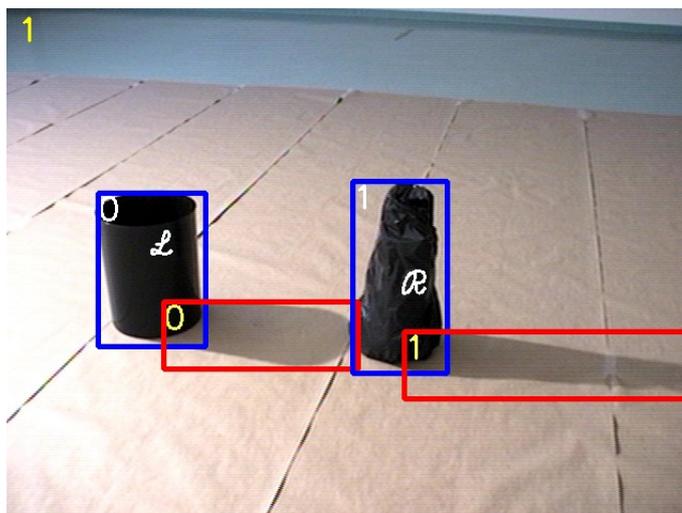


Figura 4.17: Imagem do *PeopleBot™* mostrando o resultados encontrados, no canto superior esquerdo a região de posição do robô, retângulos indicando sombras e objetos, os números iguais nos retângulos demonstram a correspondência, a letra R e L indicam o objeto mais próximo do robô e da fonte da luz respectivamente.

As restrições são: fonte de luz única e acima do objeto e observador, os objetos escuros e sem buracos, piso opaco. Com conhecimento dessas características do ambiente, segue a descrição do algoritmo de visão computacional.

Após capturada, a imagem é convertida para o espaço de cores HSV. O canal de cor a ser primeiro trabalhado é o canal de intensidade  $V$  (Figura 4.19(a)). Nele aplicamos um filtro de limiar abaixo, com valor de limiar  $\alpha$  para separar o plano de fundo do alvo (objeto e sombra) utilizando a seguinte equação:

$$Imagem\ com\ objeto\ e\ sombra\ (x, y) = \begin{cases} V(x, y), & se\ V(x, y) < \alpha ; \\ 0 & em\ outro\ caso. \end{cases} \quad (4.1)$$

Na imagem resultante ficaram os objetos, as sombras e alguns elementos não desejados, tal como penumbra (Figura 4.19(b)). O próximo passo é encontrar as regiões conectadas, e assim começar a delimitar as regiões de interesse, só que antes é aplicado na imagem um filtro morfológico de abertura, esse procedimento serve para diminuir a quantidade de grupos de componentes conectados na imagem. Por exemplo, a imagem da Figura 4.19(b) tem inicialmente 322 grupos de componentes conectados, este número é reduzido para dez depois de aplicado o filtro morfológico. Nesta mesma imagem (Figura 4.19(c)) somente dois grupos entre os dez são candidatos a região de interesse ROI (Figura 4.19(d)), isso ocorre porque existe outra restrição para considerar a região como região de interesse: o tamanho da área deve ser sempre igual ou maior ao tamanho da área de um objeto.

Para confirmar se as regiões encontradas são realmente regiões de interesse (ROI), indicadas

```

Detecta sombra e objeto (ImgCap)
  ImgHSV ← Converte RGB para HSV(ImgCap);
  (Filtro de limiar na intensidade V)
    if ( $V(x, y) > \alpha_V$ )
       $V(x, y) \leftarrow zero$ ;
    else
       $V(x, y) \leftarrow V(x, y)$ ;
    end if
  Definir Região(V, verde); // Cor Verde definida para
  desenhar retângulo na região com probabilidade de conter sombra e objeto
  (Filtro de limiar na saturação S)
    if ( $S(x, y) > \alpha_{Sat}$ )
       $S(x, y) \leftarrow ImgS(x, y)$ ;
    else
       $S(x, y) \leftarrow zero$ ;
    end if
  for i=0 até i= total de ROI then
    if (ROI em S com mais de 1% de saturação)then
      ROI[i] ← true;
    else
      ROI[i] ← false;
    end if
  end for
  Criar máscara com ROI[True];
  Imagem de objeto ← máscara em V e Limiar  $\beta$  para separar objeto de sombra.
  morfologia abertura(Imagem de objeto);
  Definir Região (Imagem de objeto, azul);
  for each ROI → Confirmar ROI com objeto → máscara;
  Imagem de sombra ← máscara em V;
  Imagem de sombra ← Limiar abaixo  $\beta$  para separar objeto de sombra;
  morfologia abertura(Imagem de sombra);
  Definir Região (Imagem de sombra, vermelho);

```

---

```

Definir Região (Imagem, Cor)
informações ( área, contorno da área, retângulo que contém área, quantidade encontrada);
Segmentação da Imagem por componentes conectados ;
  if Cor = vermelho then
    Guarda informações() das regiões de sombra segmentadas;
    Descarta sombra encontrada dentro da região do objeto;
  elseif (Cor = azul) then
    Guarda informações() das regiões de objetos segmentadas;
    Salvar os contornos das bases dos objetos.
  elseif (Cor = azul) then
    Guarda informações() das regiões de interesse ROI segmentadas;
  end if
end if
end if
end

```

Figura 4.18: Pseudocódigo do algoritmo de detecção de objeto e sombra utilizado no *PeopleBot*<sup>TM</sup>

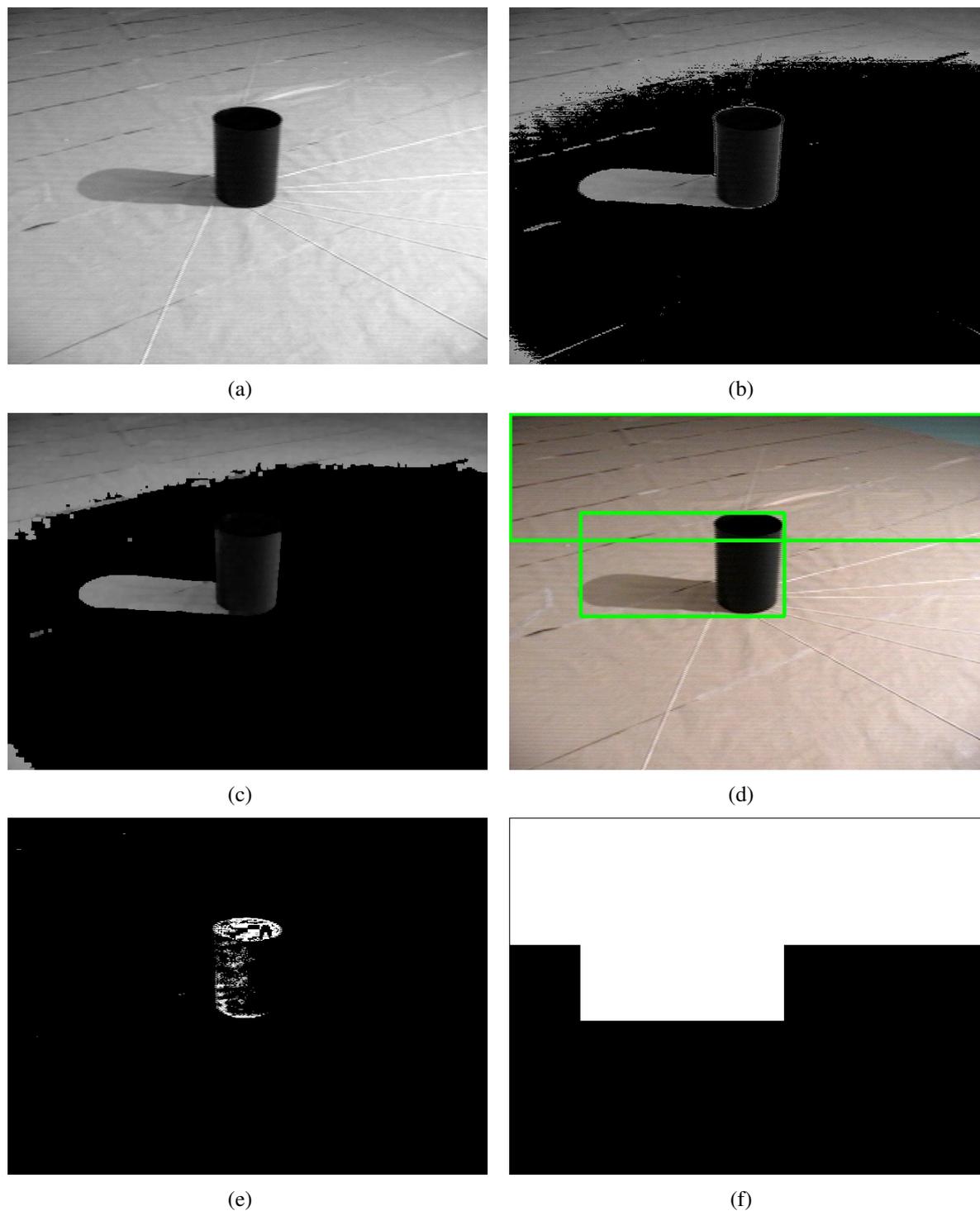


Figura 4.19: Exemplo de sequência da segmentação de um quadro na detecção ROI: (a) canal V, (b) canal V após limiar acima  $\alpha$ , (c) filtro morfológico, (d) retângulo verde indica regiões candidatas a ROI, (e) canal S após limiar acima  $\alpha_{sat}$ , (f) máscara ROI.

por retângulos na Figura 4.19(d), utiliza-se o canal de saturação, no qual foi aplicado um filtro de limiar acima (limiar  $\alpha_{sat}$ ) (Figura 4.19(e)). Neste trabalho, uma região de interesse contém pelo menos um objeto. Como a saturação da região do objeto é alta, analisamos a saturação nas regiões ROI encontradas, caso em alguma dessas regiões a saturação seja menor que 1%, esta região será desconsiderada, com as regiões que permanecerem cria-se uma máscara para ser utilizada na detecção de objetos (Figura 4.19(f)).

Para detectar os objetos, a imagem com objeto e sombra é multiplicada pela máscara de ROI, o resultado é uma imagem na qual permanece as áreas que tem maior probabilidade de haver objetos e sombras (Figura 4.20(a)). Então nesta imagem é aplicado o filtro de limiar abaixo, com limiar  $\beta$ , como o da equação 4.2:

$$Imagem\ com\ objeto\ e\ sombra\ (x, y) = \begin{cases} V(x, y), & se\ V(x, y) < \beta; \\ 0 & em\ outro\ caso. \end{cases} \quad (4.2)$$

Seguido do filtro morfológico de fechamento (Figura 4.20(b)), na imagem resultante são encontrados os componentes conectados, dos quais são guardadas as informações de cada objeto referentes: aos dados de área, ao menor retângulo que contenha a área do objeto, ao contorno, se objeto está encostado nas bordas da imagem, à base do objeto (Figura 4.20(c)) e o total de objetos encontrados, na Figura 4.20(d) tem o exemplo da imagem com o retângulo desenhado indicando o objeto encontrado.

A saturação na região do objeto deve ser alta, e é verificado o grau de saturação de todas as regiões classificadas como objetos, aquelas com baixa saturação deixam de ser considerada objetos. Por exemplo, a Figura 4.20(e) é uma cena onde foram encontradas cinco regiões como objetos, os objetos alvos com alta saturação estão indicados pelos retângulos amarelos, as regiões em azul, por terem baixa saturação, passam a fazer parte do fundo (*background*). Após localizados os objetos, um último procedimento será adotado para auxiliar na localização das sombras, mantendo o mesmo conhecimento *a priori* de que a região de interesse deva conter o objeto, uma nova máscara será criada somente com ROI que contenham os objetos localizados (Figura 4.20(f)).

Aplicando a nova máscara na imagem com sombra e objeto obtém-se uma imagem que mantém somente as regiões que tenham objetos (Figurar 4.21(a)). Quando comparamos a imagem obtida com a primeira máscara (Figura 4.20(a)) pode-se notar que regiões sem objetos são descartadas (Figurar 4.21(a)), este procedimento equivale a uma subtração de fundo, deixando

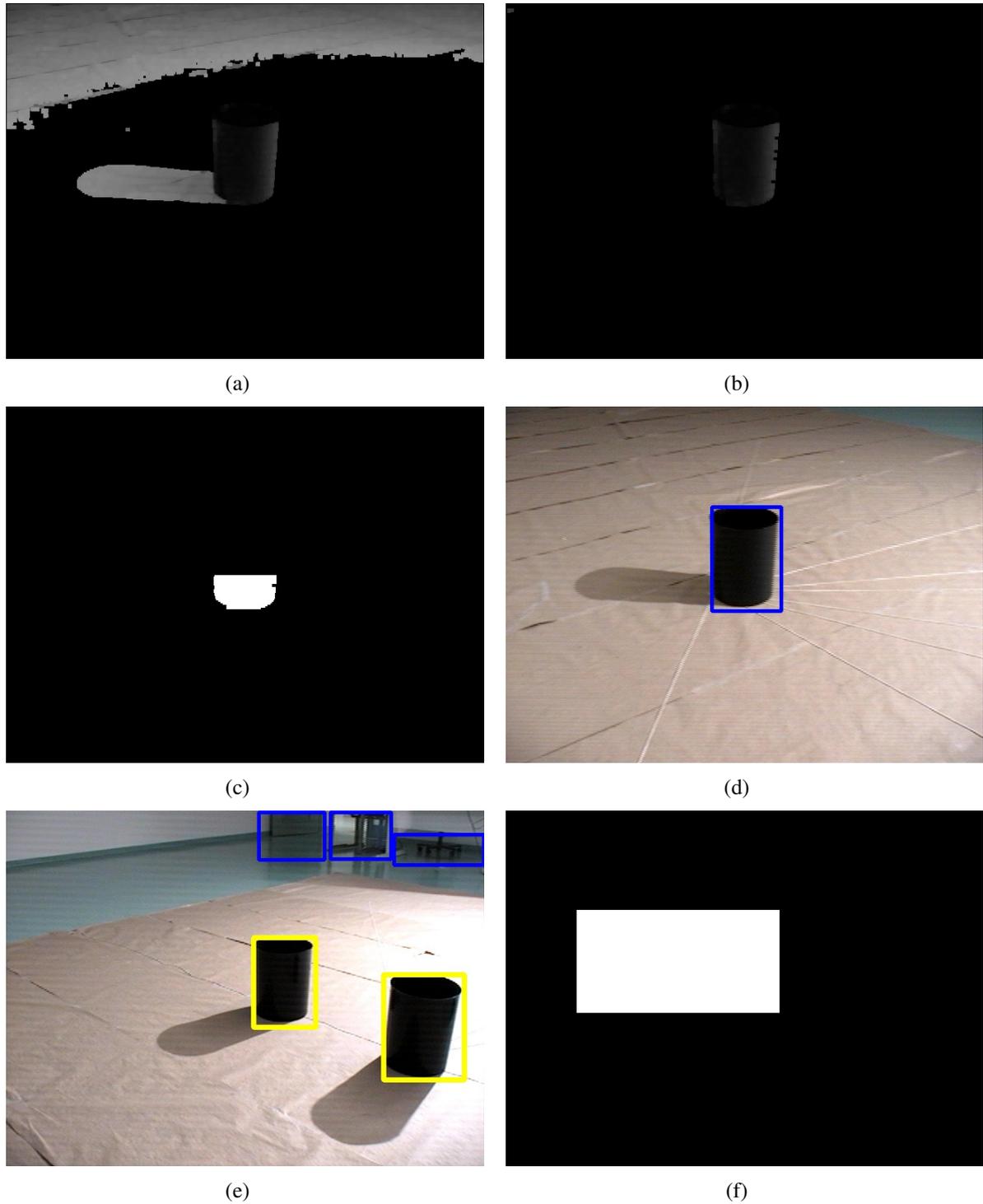


Figura 4.20: Exemplo de seqüência da segmentação de um quadro na detecção de objetos:(a)Imagem objeto e sombra após aplicado a máscara, (b) filtro de limiar abaixo e filtro morfológico, (c)imagem binária da base do objeto, (d)retângulo azul indica objetos encontrados na segmentação, (e)retângulo amarelo indica região com altas saturação, (f)máscara ROI.

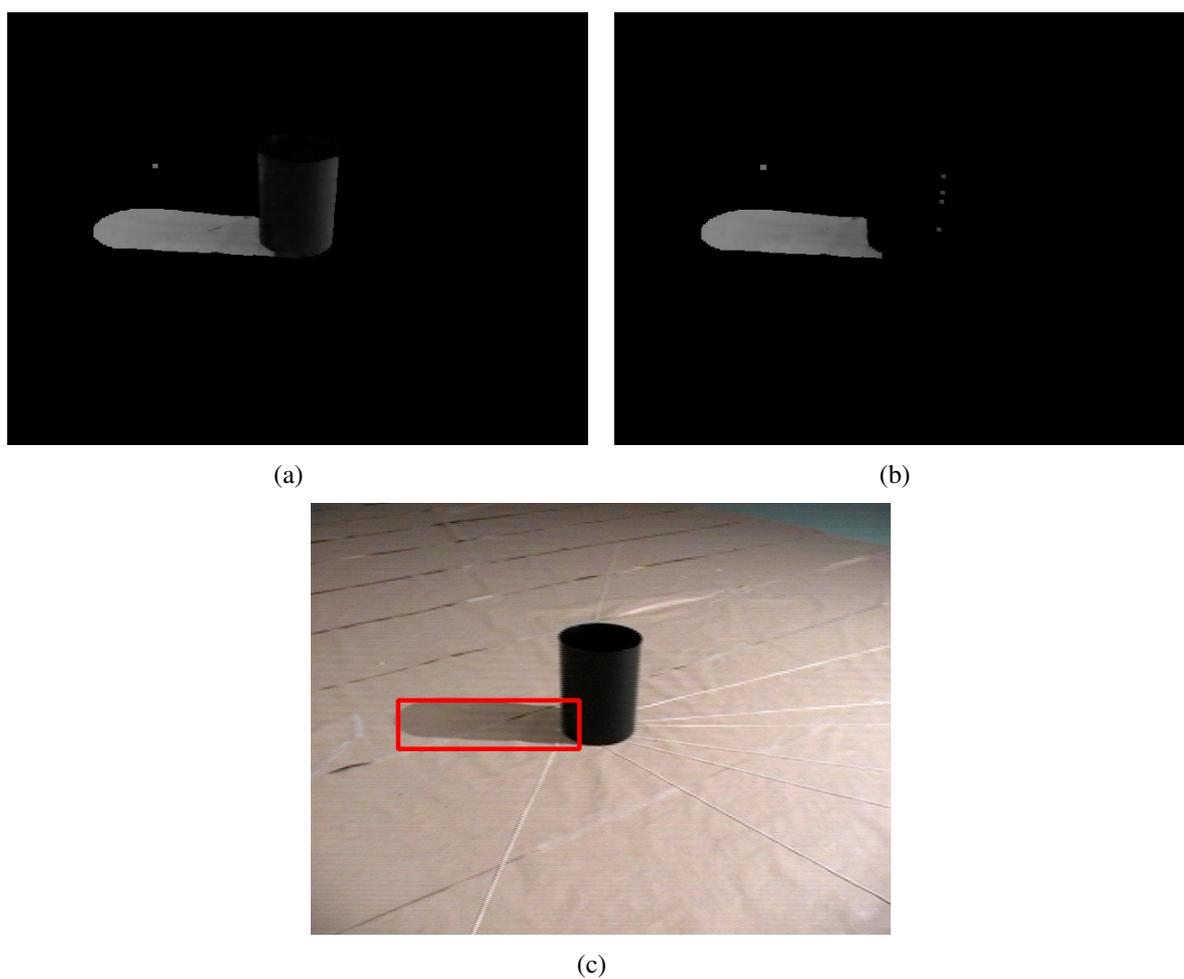


Figura 4.21: Exemplo de sequência da segmentação de um quadro na detecção de sombras:(a)Imagem com objeto após aplicada a segunda máscara, (b)Imagem após filtros de limiar acima e morfológico de abertura, (c) retângulo vermelho indicando a sombra localizada.

somente o alvo (*foreground*). Em seguida, aplicamos o filtro de limiar acima  $\beta$ , conforme segue:

$$Imagem\ sombra(x, y) = \begin{cases} V(x, y), & \text{se } V(x, y) > \beta; \\ 0 & \text{em outro caso.} \end{cases} \quad (4.3)$$

Como resultado desta equação tem-se uma imagem somente com as sombras. Os limiares  $\alpha$  e  $\beta$  são determinados na calibração, ao iniciar o software. Após aplicar o filtro de limiar acima  $\beta$ , um filtro morfológico de abertura é aplicado (Figura 4.21(b)). O passo seguinte é determinar as sombras. Utilizando componentes conectados, as sombras que estiverem totalmente dentro do retângulo de algum objeto geralmente são reflexos no objeto, então não são mais consideradas como sombra. Para os demais casos, guardam-se as características das regiões de sombra, tal como área, 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 Figura 4.21(c) mostra a sombra encontrada indicada pelo retângulo. Após encontrar as sombras e os objetos, estes devem ser relacionados entre si, para isso a próxima função a ser executada é a de análise de correspondência, que será mostrada a seguir.

#### 4.4.2 Algoritmo para análise de correspondência

Em experimentos preliminares foram realizados somente com um único objeto no ambiente. Assim achando uma sombra, esta seria diretamente relacionada ao objeto, para os casos no qual existisse mais de uma sombra verificava-se a vizinhança do objeto para encontrar qual provavelmente seria a sua. Além disso havia outra restrição em relação ao eixo  $y$  da matriz da imagem onde, a sombra de um objeto obrigatoriamente aparece na imagem com coordenadas de  $y$  máximo menor que o  $y$  máximo do objeto, isso devido ao fato de que a fonte de luz estar posicionada sempre acima do objeto. Contudo essa heurística não é suficiente para solucionar o problema de correspondência para casos onde exista mais de uma sombra na vizinhança do objeto, tal como na Figura 4.23 onde pela correspondência anterior o objeto cilíndrico teria duas sombras. Portanto houve a necessidade de buscar uma nova solução que apresentasse uma base mais sólida, apoiada em um formalismo teórico.

Na seção 3.4 foi sugerida uma solução para o problema de correspondência baseada no PQRS, em que é sombra do objeto aquela que está ligada à base deste objeto, caso a fonte de luz esteja mais próxima do objeto do que da sombra, e tanto objeto quanto pelo menos parte da sombra estejam sobre a mesma superfície. Na Figura 4.22 encontra-se o algoritmo utilizado

**Análise de Correspondência** (*Dados*)

Variáveis:

Booleano Conectado;

**while** objeto[i] **do**  **while** sombras[j] **do**

Conectado ← Conectividade(objeto[i], sombra[j]);

**if** Conectado = True **then**

sombra conectada ao objeto[i] ← sombra[j];

objeto conectado à sombra[j] ← objeto[i];

incrementa informação de total de conectado em sombra e objeto;

incrementa j;

**end if**  **end while**

incrementa i;

**end while****while** objeto[i] **do**  **if** total de conectado=1 **then**    **if** Conectado com base = True **then**      **while** sombras[j] **do**

sombra correspondente do objeto[i] ← sombra[j];

correspondência ← true;

**end while**    **end if**  **elseif** total de conectado > 1 **then**

lista de proximidade, dúvida ← objeto[i];

**while** lista dúvida **do**      **if** objeto mais próximo da luz do que sombra **then**

sombra correspondente do objeto[i] ← sombra[j];

**end if**    **end while**  **end if****end while**

---

**Conectividade** (Objeto, Sombra)

Imagem 1 ← contorno Sombra ← 100;

Imagem 2 ← contorno Objeto ← 100;

Imagem Conectado ← Soma(Imagem 1, Imagem 2);

Imagem Conectado contorno - 100;

**if** Conta (pixel ≠ zero) > zero **then**

retorna True;

**else**

retorna False;

**end if**

Figura 4.22: Pseudocódigo do algoritmo de correspondência.



Figura 4.23: Imagem capturada com *PeopleBot™*. Com o objeto cilíndrico com duas sombras conectadas a ele, a qual o algoritmo de correspondência atual consegue resolver e o anterior não.

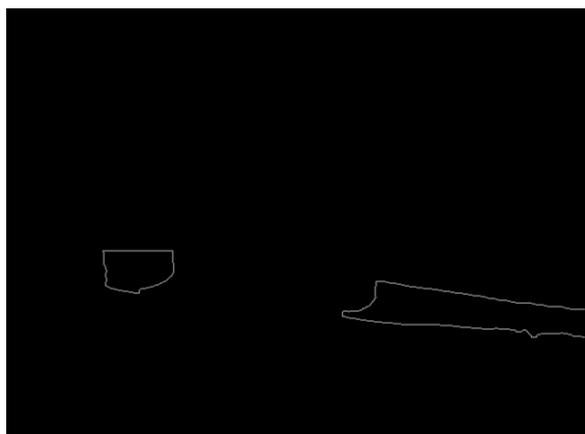
para correspondência.

O algoritmo de correspondência funciona da seguinte modo: com os dados armazenados pelo algoritmo de detecção de sombra e objeto (Figura 4.18) são analisadas as conexões externas existentes entre cada sombra e a base de todos os objetos. Em seguida, nos casos afirmativos de conexão externa, confirma-se se a fonte de luz está mais próxima do objeto do que da sombra, tanto as informações de pares objeto-sombra encontrados quanto as conexões com objetos não correspondentes são armazenadas, essas informações serão utilizadas nas próximas etapas do software. Na Figura 4.24 encontramos um exemplo de como é verificada a conexão externa das sombras com a base dos objetos, na imagem com dois objetos e duas sombras (Figura 4.24(a)) nota-se que um dos objetos tem duas sombras conectadas externamente a ele (Figuras 4.24(c) e 4.24(e)), porém somente uma delas está mais distante da fonte de luz do que o objeto. A sombra que está conectada a base de dois objetos (Figuras 4.24(d) e 4.24(e)) será utilizada para determinar qual objeto está mais próximo do robô.

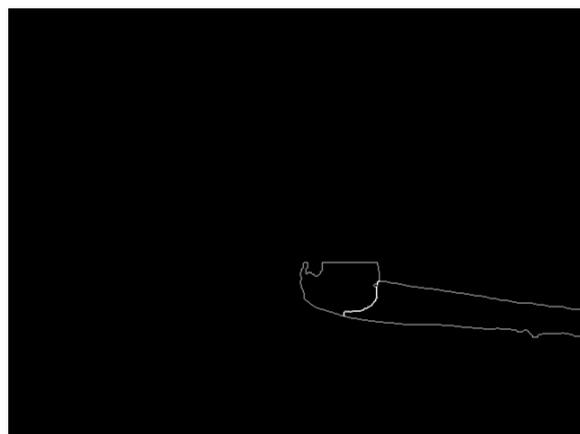
Se alguma sombra não tiver nenhuma base de objeto conectada, então será realizado um novo teste de conexão para esta sombra, porém agora com o objeto inteiro. Se encontrar uma conexão, então essa sombra deve ser de um objeto oculto, com informações sobre a proximidade com a luz pode-se determinar se a sombra é de um objeto ocluído pelo objeto ao qual ela está conectada (sombra mais distante da luz), ou de um objeto que está fora de cena (sombra mais próxima da luz). Todos os pares objeto-sombra que não tiverem restrições serão analisados no algoritmo de autocalibração. As restrições são: área da sombra ou do objeto encostada na borda da imagem e objeto ocluído do ponto de vista do observador. No primeiro caso o fato de



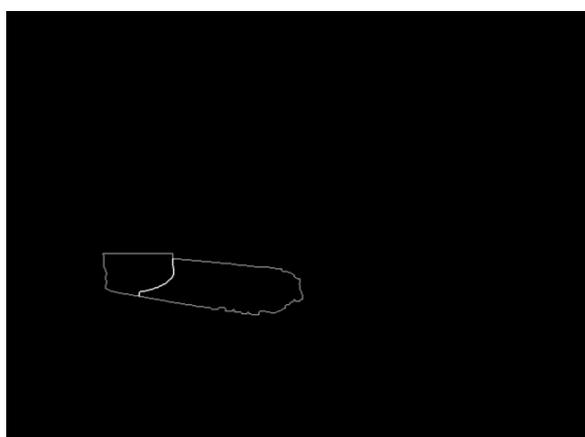
(a)



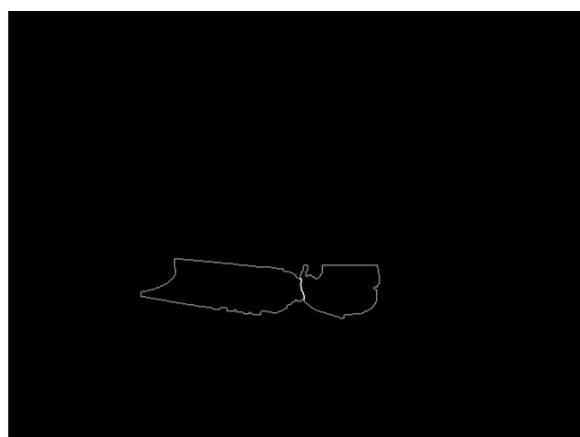
(b)



(c)



(d)



(e)

Figura 4.24: 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.

```

Autocalização (Dados de correspondencia)
  if  $\exists$  sombra then
    if  $\exists$  objeto then
      // pode estar nas regiões 1, 2, 3 ou 4
      if (area da sombra > 0.8*area do objeto) then
        // pode estar nas regiões 1 ou 2
        if  $DC(o, s)$  then // verificado utilizando
          a distancia entre a base do retângulo que
          determina objeto e sombra
          retorna região 1;
        else
          retorna região2;
        end if
      else // pode estar nas regiões 3 ou 4
        if  $PO(o, s)$  then
          // verificado pela largura do retângulo que
          delimita a sombra
          retorna região 3;
        else
          retorna região 4;
        end if
      end if
    else
      região indeterminada;
    else
      if  $\exists$  objeto then
        retorna região 5;
      else
        retorna indeterminado;
      end if
    end if
  end if
end if

```

Figura 4.25: Pseudocódigo do algoritmo de localização do robô

uma área estar encostada na borda da imagem pode significar que parte das informações não foram coletadas. No segundo caso, a oclusão ocasiona a segmentação de dois objetos como um, esses casos não são tratados neste trabalho, somente são detectados. Tanto a autocalização quanto a proximidade serão discutidas na próxima seção.

#### 4.4.3 Algoritmos de autocalização e de proximidade

Na Figura 4.25 encontra-se o algoritmo de autocalização, nele a resposta da localização do robô é dada a partir das relações percebidas na imagem entre as áreas dos pares objeto-sombra. Em experimentos preliminares foi utilizado o posicionamento e tamanho dos retângulos que

```
Proximidade(Lista de proximidade)
Lista de proximidade ← sombra com mais de um objeto conectado
if (Lista de proximidade == true)
  for each (sombra da lista de proximidade)
     $o$  ← objeto que provocou a sombra;
     $o'$  ← objeto no qual a sombra se projeta;
    if (região==1) then
       $N(o', o, rob)$ ;
    elseif (região==2 ou região==3 ou região==4)
       $N(o, o', rob)$ ;
    end if
  end for
end if
```

Figura 4.26: Pseudocódigo do algoritmo de proximidade

indicavam as áreas de sombra e de objeto na imagem, entretanto essa era uma solução provisória pois a área de um retângulo não retrata de modo inequívoco o tamanho da região do elemento contido dentro dele. Assim, procurando melhorar o algoritmo, a análise da relação entre sombra-objeto é realizada a partir da posição e tamanho das áreas destas regiões, esse algoritmo depende dos dados guardados no algoritmo de detecção de sombra e objeto (Figura 4.18) e no algoritmo de correspondência (Figura 4.22).

A primeira observação realizada pelo algoritmo de autolocalização (Figura 4.25) é sobre a presença na imagem de sombra e objeto, então se:

- existem ambos (sombra e objeto), para cada par encontrado na correspondência como apto para ser utilizado neste algoritmo, a primeira característica analisada é o tamanho da área da sombra em relação ao objeto. Devido à oclusão da sombra as áreas observadas a partir das regiões 3 e 4 são bem menores que aquelas observadas a partir das regiões 1 e 2. Das duas regiões separadas pelo tamanho, a posição da área da sombra em relação à área do objeto é a característica que diferencia entre uma ou outra região. Com a posição do robô relativa a cada objeto, por conveniência, retorna-se a resposta relativa ao objeto mais próximo do centro da imagem.
- existem somente objetos, retorna região 5;
- existem somente sombras ou ambos (sombra e objeto) não aparecem, não é possível determinar a região.

A partir das informações encontradas nos algoritmos anteriores, o algoritmo de proximidade (Figura 4.26) analisa cada sombra que esteja conectada a mais de uma base de objeto, e utiliza a

teoria apresentada na seção 3.3 para determinar qual dos objetos conectados está mais próximo do robô. O algoritmo verifica objetos par a par para determinar o mais próximo, em alguns casos o algoritmo pode encontrar mais de um objeto próximo ao robô, se existir uma sombra que esteja conectada nestes objetos (par objeto próximo), será escolhido o mais próximo entre eles. Além da conexão da sombra com o par de objetos, o que determina a distância relativa entre o robô e os objetos, é a região de localização do robô em relação ao objeto que provocou a sombra. Assim, se na região 1 então o robô está mais próximo do objeto no qual a sombra se projeta sobre ele. De outro modo, se nas regiões 2, 3 ou 4, o robô estará mais próximo do objeto que provocou a sombra. Esse algoritmo não trata a oclusão como informação de profundidade, deixando em aberto para trabalhos futuros. Todos os algoritmos propostos acima foram testados conforme procedimento descrito no próximo capítulo.

## 4.5 Discussão final

Este foi um capítulo onde algoritmos foram definidos, pelo fluxo de dados, pela descrição do ambiente e dos materiais dentro dele, e pelos pseudo-códigos do software desenvolvido. Os fluxogramas apresentados demonstraram como os dados são tratados em cada etapa relevante deste trabalho. Com a descrição do ambiente delimitou-se a área de atuação do robô, as restrições existentes no ambiente e como as relações entre sombra e objetos são percebidas nas imagens para cada posição do observador, padronizando os dados que são analisados nos experimentos.

Os pseudocódigos foram baseados nas soluções encontradas e utilizadas neste trabalho. A opção por descrever o código utilizado desta maneira é porque, caso haja interesse em utilizá-los, pode-se fazê-lo em qualquer outro tipo de linguagem de programação. A escolha de fazer uma sistema baseado em C++ é justificada pelo fato de que já existiam bibliotecas prontas na área de visão e controle do robô baseados neste sistema, além de ser uma linguagem na qual já existia um conhecimento prévio, o que nos garantiu uma maior eficiência em ir direto ao objetivo. Deixamos para trabalhos futuros, a melhoria da sistema de visão, que tem influência direta no desempenho do software de modo geral. Um dos algoritmos mais influenciados pela visão é a autolocalização, pois resolve a relação entre sombra e objeto com as informações de área, assim qualquer perda de dados ocorrida no algoritmo de detecção acarreta numa resposta errada da localização do robô.

A proporção entre as áreas de sombra e objeto utilizadas para definir as regiões é uma

solução pontual para este trabalho, visto que esta proporção muda com a mudança da altura ou distância da luz em relação ao objeto, por esse motivo para trabalhos futuros uma nova maneira de analisar relações entre objeto e sombra deva ser levada em consideração. O software não analisa objetos ocultos do ponto de vista do observador. Todos esses casos são problemas que ficaram em aberto para investigações futuras.

## Capítulo 5

# PROCEDIMENTO EXPERIMENTAL E RESULTADOS

Neste capítulo estão descritos como os procedimentos experimentais foram realizados, como se deu a coleta de dados, como estes foram avaliados e os resultados obtidos nos vários testes. Os experimentos foram realizados em um laboratório, onde o robô móvel *PeopleBot<sup>TM</sup>* captura imagens do ambiente para extrair dados que são informações de regiões na imagem relativas às sombras e objetos. Com essas informações o sistema retorna como resposta o posicionamento qualitativo do robô, a correspondência entre objeto e sombra, e a proximidade do robô com objeto. Nos primeiros experimentos, as imagens com objeto e sombra foram capturadas de modo não sequencial assim o robô foi posicionado de modo que a imagem capturada teria o objeto bem centralizado, no entanto com a evolução do software todos os quadros começaram a ser analisados. Os casos indeterminados não são analisados para autolocalização, são eles as imagens em que não existem objetos, ou objeto ou sombra aparecem na extremidade da imagem, quando isso ocorre é provável que o elemento não esteja por completo no campo de visão do *PeopleBot<sup>TM</sup>*. As respostas do sistema e os quadros são armazenados para avaliação posterior.

### 5.1 Método de validação

O objetivo principal deste trabalho é a autolocalização qualitativa. O robô captura imagens do ambiente enquanto navega pela sua região de atuação (Figura 4.7). A partir da observação e processamento destas imagens, realiza-se uma análise dos objetos e sombras a partir da qual é retornada como resposta a região em que o robô se encontra, além de respostas de proximidade entre objetos e robô, objeto e luz, e a correspondência entre objeto e sombra. Para análise de

autolocalização será utilizada a comparação entre a resposta do sistema e a real posição do robô marcada pelo operador diretamente no software no momento que os testes foram realizados. Por exemplo, um robô que pare na região 3 (Figura 5.1)  $P$  vezes, terá sua porcentagem de

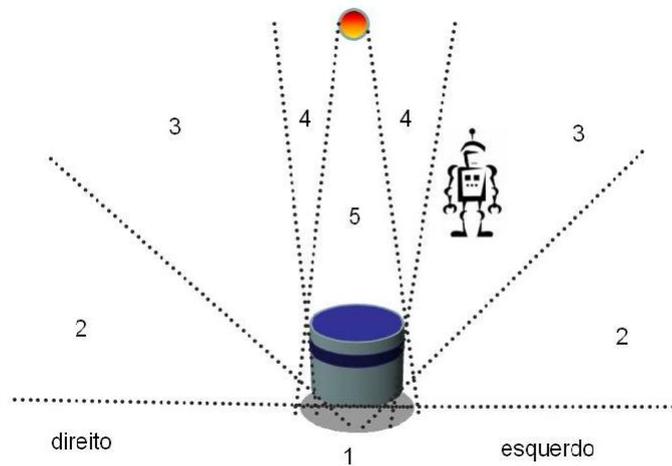


Figura 5.1: Representação de um robô na região 3.

acerto dada pela fórmula usual:

$$\%acerto = \frac{P - total\ de\ erro}{P}. \quad (5.1)$$

As proximidades objeto-robô, objeto-luz e a correspondência serão analisadas por amostragem em testes *off-line*. A seguir serão descritos os procedimentos para realização dos testes e coletas de dados.

## 5.2 Testes

Na preparação do ambiente a fonte de luz é colocada em uma posição escolhida pelo operador, e o objeto à uma distância adequada, ou seja, essa distância é tal que o robô possa passar pela região 5 sem obstruir a luz. Após definida as posições do objeto e da luz, as regiões serão demarcadas a fim de prover a base de conhecimento verdadeira (*Gold standard*) sobre os eventos, que será contrastada com a resposta automática.

Para auxiliar na demarcação das regiões em relação a pelo menos um dos objetos, foi utilizado um barbante que ajuda o operador a reconhecer as regiões. Pela cor do barbante o sistema de visão irá descartá-lo (não influenciando portanto, no resposta do sistema). O operador deve marcar em qual região o robô está posicionado, esse dado é o *gold standard*.

Com o ambiente preparado, posiciona-se o robô em um ponto escolhido aleatoriamente den-

tro da sua área de atuação (Figura 4.7). Neste ponto deve-se determinar qual tipo de controle será utilizado para a navegação do robô. A navegação do *PeopleBot<sup>TM</sup>* pode ser realizada pelo controle manual ou autônomo. Visando acelerar o processo de coleta de dados, foi escolhido o controle manual. Os parâmetros de visão são determinados neste momento por meio da calibração da imagem. Após a calibração inicializa-se o software (seção 4.4) para autolocalização e proximidade. O operador deve marcar a região em que se encontra o *PeopleBot<sup>TM</sup>* e, assim que ele estiver pronto para movimentar-se, deve também marcar a opção de salvar imagens e dados. Todos os quadros são armazenados para novos testes *off-line*.

Os testes *off-line* serviram para quantificar os totais de acertos em correspondência e proximidade. Esses testes são úteis também para refazer os testes de auto-localização na mesma imagem com parâmetros de visão diferentes, e assim analisar se os erros ocorridos foram causados pelo sistema de visão, ou pelo algoritmo de raciocínio. As respostas para correspondência e proximidade foram analisadas diretamente na imagem. Por exemplo na Figura 5.2, como resposta para correspondência, cada objeto recebe um número e a sua sombra deve receber um número igual. Para mostrar qual objeto está mais próximo da luz e aquele que está mais próximo do robô foram utilizadas as letras *L* e *R* respectivamente.

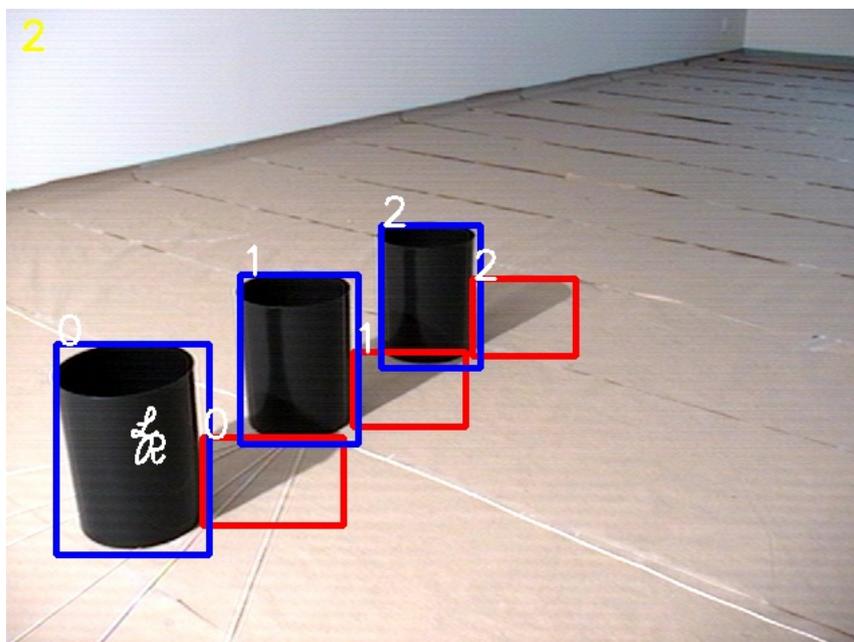


Figura 5.2: Imagem indicando a correspondência entre o objeto e a sombra; o objeto mais próximo da luz e do robô está marcado por *L* e *R* e a localização do robô é a região 2 (conforme marcado no canto superior esquerdo da imagem).

Para os testes de autolocalização todos os quadros foram salvos em arquivo de imagem e os resultados em arquivo texto. As informações arquivadas foram: a região real marcada pelo

operador, a região localizada pelo robô e o nome do arquivo de imagem correspondente. Outros testes *off-line* de autolocalização foram realizados com as mesmas imagens do teste *on-line*, as respostas de localização foram salvas em arquivo e comparadas com a região anotada pelo operador para a contagem de acertos.

A contagem dos dados para autolocalização é realizada do seguinte modo, para cada teste com um conjunto de imagens é contabilizado o número de respostas dadas em cada região real. As respostas correspondentes à região real são consideradas corretas, para autolocalização somente a resposta de um objeto em cena foi considerada.

Para os resultados sobre a correspondência objeto/sombra foram selecionadas amostras de imagens com três, dois, um e nenhum objeto. Na contagem geral de acertos, nesse caso, desconsideramos cenas em que algum erro em correspondência ocorreu. Ou seja, por exemplo, uma cena com três objetos, em que somente dois destes foram corretamente associados às suas sombras, foi considerada como um resultado negativo de correspondência.

Tanto para os testes de correspondência quanto para os de proximidade procuramos atenuar a influência do sistema de visão, retirando da análise todas as imagens onde a segmentação foi diretamente responsável pelo erro. Isto é, os erros apresentados não foram causados pela segmentação. Ainda para a proximidade foram analisadas imagens em que uma sombra estava projetada na superfície de outro objeto, isso pode ser observado quando os objetos estão oclusos do ponto de vista da luz, e o observador encontra-se nas regiões 1, 2, 3 e 4.

Outro fator que teve influência nos resultados são as imagens onde a oclusão dos objetos (do ponto de vista do robô) é observada, a oclusão em nosso sistema é identificada porém ainda não foi explorada como fonte de informação neste trabalho. Assim, em correspondência, o objeto ocluso também é considerado. Nos testes de autolocalização, nos casos em que há oclusão de objetos, o sistema responde “indeterminado”. Para a proximidade de luz nenhuma restrição foi acrescida ao algoritmo. A seguir vamos apresentar os resultados obtidos.

### 5.3 Resultados

Os resultados aqui apresentados foram obtidos a partir dos testes descritos na seção anterior. Serão transcritos os resultados obtidos com os experimentos nas fases preliminares (encontrados em Santos *et al.* (2008, 2009)), e com os testes posteriores realizados para validação das soluções para autolocalização e proximidade (apresentadas nas seções 3.2 e 3.3 respectivamente).

### 5.3.1 Resultados preliminares

Na tabela 5.1 têm-se os resultado dos experimentos preliminares de autolocalização. As linhas desta tabela são as regiões do ambiente onde o robô estava posicionado. A segunda coluna (Total de poses) representa o total de quadros capturados na região correspondente, a terceira e quarta coluna são respectivamente, o total de imagens na qual a análise sobre a localização foi correta e sua porcentagem de acerto. A porcentagem de acerto para cada região foi calculada utilizando a equação 5.1. Assim para a região 1, foram analisadas 56 poses das quais 40 foram corretamente percebidas, equivalendo a 71% de acerto. Na região 2, de 35 poses capturadas, a localização foi correta em 27 delas, correspondendo a 77% de acerto. Para a região 3, de 21 poses 15 tiveram respostas positiva com 71% de acerto. A região 4 acertou 6 das 7 poses capturadas ou seja 86% de respostas corretas. Já a região 5 obteve 100% de acerto de um total de 9 poses. Foram analisadas 128 imagens.

Tabela 5.1: Total de poses capturadas e analisadas no experimento preliminar.

	Total de poses	Localização correta	acertos (%)
Região 1	56	40	71
Região 2	35	27	77
Região 3	21	15	71
Região 4	7	6	86
Região 5	9	9	100

Tabela 5.2: Dados capturados nos experimentos preliminares. Colunas são as regiões reais e linhas as respostas dada pelo robô.

	Região 1	Região 2	Região 3	Região 4	Região 5
Região 1	40	0	0	0	0
Região 2	11	27	2	0	0
Região 3	0	2	15	0	0
Região 4	0	5	4	6	0
Região 5	0	1	0	1	9
sem localização	5	0	0	0	0
Acerto (%)	71,4	77,1	71,4	85,7	100

A tabela 5.2 apresenta os mesmos resultados encontrados da tabela 5.1, entretanto de modo mais detalhado. Esta tabela mostra também qual a resposta dada pelo algoritmo de autolocalização para cada região em que se encontrava. Cada coluna representa a região em que o robô estava posicionado e as linhas são as respostas. Assim pode-se ver pela tabela 5.2 que, quando na região 1, a resposta dada foi de que estava na região 2 onze vezes e não se localizou cinco

Tabela 5.3: Resultados dos experimentos *on-line* de auto-localização, com calibração feita no momento do teste.

	Região 1	Região 2	Região 3	Região 4	Região 5	imagem inválida
Região 1	32	7	19	8	6	79
Região 2	20	85	14	19	8	66
Região 3	3	8	7	8	0	0
Região 4	14	12	7	7	0	25
Região 5	7	20	8	9	8	10
indeterminado	14	20	0	0	0	13
Acerto %	35,6	55,9	12,7	13,7	36,4	6,7

Tabela 5.4: Resultados dos experimentos *off-line* de auto-localização com calibração  $\alpha = 170$ .

	Região 1	Região 2	Região 3	Região 4	Região 5	imagem inválida
Região 1	6	2	0	0	0	5
Região 2	0	110	0	0	0	10
Região 3	0	0	12	1	0	2
Região 4	4	5	2	11	0	1
Região 5	79	35	27	18	0	54
indeterminado	1	0	14	21	22	121
Acerto %	6,7	72,4	21,8	21,6	0,0	62,7

vezes. Quando na região 2, respondeu que estava na região 3 duas vezes, e nas regiões mais distantes 4 e 5, cinco e uma vezes respectivamente. Para imagens capturadas na região 3 errou seis vezes considerando-a nas regiões vizinhas sendo, duas vezes na região 2 e quatro na região 4.

### 5.3.2 Resultados de auto-localização

Os testes de auto-localização foram realizados em horários distintos, sendo que para cada teste uma nova calibração era feita em região diferente. No total foram analisados 563 poses de imagens capturadas em três testes. Os resultados dos testes *on-line* estão na tabela 5.3, onde cada coluna representa a posição no ambiente em que o robô se encontrava. A coluna imagem inválida refere-se a cenas na qual os dados não seriam suficientes para determinar a região corretamente. Com o mesmo grupo de imagens outros testes de autolocalização *off-line* foram realizados, modificando-se o parâmetro de limiar para intensidade luminosa  $\alpha$  (equação 4.1). Na tabela 5.4 estão os resultados para um parâmetro  $\alpha = 170$  e a tabela 5.5 estão os resultados com  $\alpha = 140$ . Nestas tabelas estão contabilizadas as respostas dadas para cada região.

Os resultados na tabela 5.6 assumem um sistema capaz de escolher automaticamente o melhor limiar em cada situação. Ou seja, esta tabela coleta os melhores resultados dos três testes (apresentados nas tabelas 5.3, 5.4 e 5.5). Assim de 90 poses capturadas na região 1, 81 tiveram uma resposta positiva (em algum dos testes), para a região 2 foram 152 imagens das quais 146

Tabela 5.5: Resultados dos experimentos *off-line* de auto-localização com calibração  $\alpha = 140$ .

	Região 1	Região 2	Região 3	Região 4	Região 5	imagem inválida
Região 1	60	9	0	0	0	3
Região 2	0	12	0	0	0	0
Região 3	0	97	39	0	0	3
Região 4	0	15	7	39	0	5
Região 5	23	11	3	4	18	5
indeterminado	7	8	6	8	4	177
Acerto %	66,7	7,9	70,9	76,5	81,8	91,7

Tabela 5.6: Resultados de poses analisadas pelo algoritmo de autolocalização onde pelo menos uma resposta correta foi encontrada entre os resultados das tabelas 5.3, 5.4 e 5.5, acertos pela (equação 5.1)

	Total de poses	Localização correta	acertos (%)
Região 1	90	81	90
Região 2	152	146	96
Região 3	55	37	67
Região 4	51	43	84
Região 5	22	20	91
inválida	193	180	93

tiveram sua localização correta, para a região 3 de 55 poses 37 estavam certas, enquanto 43 imagens de 51 capturadas na região 4 estavam corretas, e da região 5 vinte poses foram identificadas corretamente de 22. Nesta as imagem inválidas foram identificadas 180 vezes das 193 existentes.

Após realizados os testes acima, outro teste de autolocalização foi realizado utilizando o limiar  $\alpha = 120$ . Os resultados deste teste foram transcritos na tabela 5.7. Acrescentando os resultados deste teste aos resultados apresentados na tabela 5.6, todas imagens capturadas nas regiões 2, 3, 4 e 5 tiveram uma resposta positiva com pelo menos um dos parâmetros. Região 1 e imagens inválidas tiveram um rendimento positivo em 96% e 99% respectivamente. Dos quatro testes realizados (Tabelas 5.3, 5.4, 5.5 e 5.7) sobraram seis imagens com resposta errada, com as quais foi realizado um teste com um novo limiar cada uma e uma resposta positiva de autolocalização foi encontrada. As 563 poses analisadas foram coletadas em ambiente com somente um objeto. Deste modo, o sistema de visão foi diretamente responsável pelo erro na distinção das regiões.

A tabela 5.8 mostra um resumo dos resultados obtidos nos testes que estão transcritos nas tabelas 5.4, 5.5, 5.6 e 5.7. Os resultados em porcentagem foram concentrados para a tabela 5.8, note que cada região teve um desempenho maior que 70% em pelo menos um dos testes. O problema reside portanto na calibração do sistema de visão. Isso seria facilmente resolvido

Tabela 5.7: Resultados dos experimentos *off-line* de auto-localização com calibração  $\alpha = 120$ .

	Região 1	Região 2	Região 3	Região 4	Região 5	imagem inválida
Região 1	75	0	0	0	0	3
Região 2	0	35	0	0	0	0
Região 3	7	7	21	0	0	6
Região 4	0	92	24	29	0	9
Região 5	1	18	6	22	21	8
indeterminado	7	0	2	0	1	169
Acerto %	83,3	23,0	39,6	56,9	95,5	86,7

Tabela 5.8: Acertos em porcentagem encontrados nas tabelas 5.3, 5.4 e 5.5

Localização do robô	total de acertos (%)			
	teste <i>on-line</i>	$\alpha = 170$	$\alpha = 140$	$\alpha = 120$
Região 1	35,6	6,7	66,7	<b>83,3</b>
Região 2	55,9	<b>72,4</b>	7,9	23,0
Região 3	12,7	21,8	<b>70,9</b>	39,6
Região 4	13,7	21,6	<b>76,5</b>	56,9
Região 5	36,4	0,0	81,8	<b>95,5</b>
inválida	6,7	62,7	<b>91,7</b>	86,7

Tabela 5.9: Resultados do teste de autolocalização realizado com imagens com mais de um objeto

	Total de poses	Localização correta	acertos (%)
Região 1	25	24	96
Região 2	81	81	100
Região 3	35	34	97
Região 4	15	13	87
Região 5	5	5	100
Total geral	161	157	97

com uma calibração adaptativa (que é um problema para trabalhos futuros). A partir desses resultados assumimos o problema de visão como resolvido para os testes a seguir, os quais foram realizados com mais de um objeto no ambiente.

Para testar a autolocalização em ambiente com mais de um objeto, foram selecionadas 161 imagens onde os erros não foram causados diretamente pela segmentação. Essa pré-seleção considera que o problema de visão foi comprovado e delimitado nos testes anteriores. Na tabela 5.9 vê-se que foi analisada: 25 imagens da região 1; 81 da região 2; 35 da região 3; quinze da região 4 e cinco da região 5. Desta amostragem 4 imagens tiveram uma resposta errada de autolocalização, o erro foi ocasionado por uma pequena perda de informação quanto ao tamanho de área na segmentação da sombra, o tamanho da área é uma característica utilizada neste algoritmo.

### 5.3.3 Resultados de proximidade

A proximidade do robô em relação ao objeto foi analisada nos casos em que pelo menos parte do objeto  $o'$  é superfície onde se projeta a sombra de um objeto  $o$ . Assim, foram analisadas 24 imagens capturas na região 1, 72 imagens na região 2, 28 na região 3 e 2 na região 4, em um total de 126 imagens. Essas imagens foram selecionadas de modo a representar o problema de segmentação solucionado. Na tabela 5.10, a primeira coluna refere-se ao total de poses analisadas para cada região, na segunda coluna está a quantidade de acerto ao determinar qual o objeto mais próximo do robô e esse resultado em porcentagem está na terceira coluna.

O motivo do erro ocorrido em duas cenas capturadas na região 1 foi a resposta errada da autolocalização. Quando o erro de autolocalização ocorre entre as regiões 2, 3 ou 4 não existe influência na resposta final de proximidade. Entretanto quando existe oclusão do ponto de vista do observador a autolocalização não é realizada, assim não há uma resposta positiva nestes casos, os erros ocorridos nas regiões 3 e 4 tiveram a oclusão como causa. Quando a sombra de um objeto se projeta sobre outro, eles estão oclusos do ponto de vista da luz, essa oclusão é percebida pelo observador quando este está na região 5, 4 e na área da região 3 próxima da fronteira com a região 4. A solução para estes casos é começar a trabalhar com as informações de oclusão.

Determinar qual é o objeto mais próximo da fonte da luz, é um requisito para confirmar a correspondência entre sombra e objeto e por consequência para todos os algoritmos desenvolvidos. Nos experimentos realizados percebeu-se que a proximidade da luz teve uma resposta positiva em todos os casos testados quando o algoritmo de detecção conseguia extrair pelo menos parcialmente as sombras e objetos existentes em cena.

Tabela 5.10: Tabela com resultados de proximidade com robô

	Total de poses	Proximidade com robô	Acertos %
Região 1	24	22	92
Região 2	72	72	100
Região 3	28	17	61
Região 4	2	0	0
Total	126	111	88

### 5.3.4 Resultados de correspondência

Os resultados sobre a correspondência estão relacionados na tabela 5.11. A coluna total analisadas, refere-se aos 326 quadros onde 114 tem um objeto, 38 com dois objetos, 129 com

três objetos e 45 sem nenhum objeto em cena. Os resultados foram contados do seguinte modo, nas colunas cor 0 significa que não foi encontrado nenhum objeto no quadro, cor 1 que na imagem somente um objeto teve sua correspondência positiva, em cor 2 houve pelo menos dois objetos na imagem com correspondência correta e cor 3 três objetos corresponderam corretamente com sua sombra. Assim pode-se notar que, das 129 imagens com três objetos, sete não corresponderam, em quinze pelo menos um objeto teve seu resultado positivo de correspondência, em 34 quadros dois objetos corresponderam corretamente com suas sombras, e em 73 imagens todos foram corretamente relacionados com suas sombras. Todas as outras imagens são analisadas de modo análogo, e na coluna de acertos somente são consideradas as imagens nas quais todos os objetos na imagem foram corretamente ligados às suas sombras.

Tabela 5.11: Tabela com resultados de correspondência, a coluna acerto somente considera como certo as imagens na qual todos os pares objeto-sombra foram corretamente identificados

Imagens com	total analisadas	cor 0	cor 1	cor 2	cor 3	Acerto %
sem objeto	45	45	—	—	—	100
1 Objeto	114	9	105	—	—	92
2 Objetos	38	0	2	36	—	95
3 Objetos	129	7	15	34	73	56

## 5.4 Discussão final

O experimento preliminar apresentado (*Santos et al.*, 2008) foi realizado com um esboço de sistema. Inicialmente o filtro de sombra era feito somente pela luminância. O ambiente era controlado para que a sombra e o objeto ficassem bem destacados em cena. Neste cenário o robô navegava aleatoriamente, ao parar a função de captura de imagem e o algoritmo de auto-localização eram acionados, contudo esse método tem um custo alto de tempo. A opção foi utilizar o controle manual para navegação visto que, assim, os testes seriam em menor tempo. O controle manual não interfere nas tarefas testadas: a capacidade do robô em raciocinar espacialmente de modo qualitativo, inferindo sua localização relativa ao objeto e a fonte luz, a partir da observação da relação sombra-objeto.

A segmentação de sombras e objetos é uma parte essencial quanto ao sucesso da autolocalização. A visão tem uma grande influência na localização, isso porque a informação utilizada para determinar as relações entre objeto e sombra é a área de regiões na imagem. Assim, partes das informações que sejam alteradas pelos filtros irão alterar os resultados do raciocínio. Nesse trabalho a segmentação é baseada em filtros de limiares. Para os primeiros experimentos, os

limiaries eram alterados direto no código, e para cada teste eram determinados no modo de tentativa e erro. Nos novos testes a calibração é realizada através do software no início dos testes, ainda assim, calibrar os parâmetros de visão no momento de realizar os experimento não foi a melhor solução, como ficou evidente nos resultados obtidos no teste *on-line* (Tabela 5.3)

Nos testes realizados visando analisar a correspondência, percebemos que a segmentação tem uma influência menor nos resultados, pois a correspondência está relacionada diretamente com a conexão externa entre a sombra e a base do objeto. Mesmo que um objeto ou uma sombra esteja parcialmente fora de cena, ou que tenha perdido parte de sua área na segmentação, ainda assim, isso não prejudica a análise da correspondência.

Os resultados obtidos, quando minimizamos a influência da segmentação na resposta final do sistema, deixam evidente que: as informações espaciais obtidas podem servir como heurísticas aliadas a outros sistemas de navegação. Isso só será realmente viável quando a segmentação tiver sua influência negativa diminuída.

## Capítulo 6

# CONCLUSÃO

As sombras são consideradas como um ruído na maioria dos trabalhos que envolvem visão computacional. Contudo as sombras são úteis para percepção humana, para inferir movimento, profundidade e posicionamento de elementos em imagens bidimensionais. A teoria PQRS é baseada em lógica e consiste em formalizar o bom senso sobre o raciocínio espacial utilizando as sombras, ela permite a inferência de novos dados sobre o ambiente. Em trabalhos futuros esta informação poderá ser utilizada como fluente em cálculo de situação (*Brachman and Levesque, 2004*) ou no cálculo de eventos (*Muller, 1998; Shanahan, 1997*).

O software de autolocalização do robô móvel foi desenvolvido tendo como base os axiomas da teoria PQRS em relação a: localização, proximidade e correspondência sombra-objeto. Com esse software o robô é capaz de responder em qual região está localizado a partir da relação entre par sombra-objeto, em casos especiais o robô pode inferir qual objeto em cena está mais próximo. As informações utilizadas foram encontradas a partir dos dados coletados nas imagens capturadas pelo *PeopleBot<sup>TM</sup>*.

Nos experimentos realizados o modo de controle manual foi escolhido para agilizar a coleta de imagens, e não interferiu no alcance do objetivo principal de dar à um robô móvel a capacidade de inferir a sua localização, utilizando a teoria de relações perceptuais qualitativa sobre sombras. A contribuição deste trabalho consistiu em aplicar a teoria PQRS no desenvolvimento do software que infere, a partir da imagem de uma câmera, a localização espacial qualitativa, profundidade, oclusão e uma solução parcial para o problema de correspondência.

Para que as informações sobre sombras possam ser utilizadas de maneira eficaz é necessário um sistema de segmentação de sombras mais robusto ou um outro modo de interpretar os dados de informações incompletas. A segmentação de sombra deve ser melhorada em trabalhos futuros, existe na literatura de visão computacional outras técnicas que podem ser utilizadas

---

para isso, como por exemplo segmentação por campos aleatórios de Markov (*Markov random fields*) (Pérez, 1998), segmentação por entropia de Tsallis (*Tsallis entropy*) (Rodrigues et al., 2006) e segmentação com redes neurais (Simões and Reali-Costa, 2000). Contudo, a partir dos resultados dos testes de autolocalização, no qual obtivemos uma resposta positiva alterando o limiar, fica como sugestão uma solução de utilizar o diagrama de localização relativa no ambiente, criando assim uma expectativa no robô de sua localização. Então a resposta do sistema será confrontada com a localização esperada, caso não sejam iguais o limiar utilizado na visão computacional será alterado até que a comparação seja positiva, ou todos os limiares possíveis dentro de um range, tenham sido testados. Outra sugestão é alterar o modo como as relações entre sombra e objetos são analisadas no algoritmo de autolocalização, deixando de utilizar a informação de tamanho de área, pois esse dado exige uma segmentação mais robusta.

Uma melhora do sistema de detecção de sombra irá auxiliar a agregar as informações que podem serem extraídas da mesma a outros tipo de sistema e técnicas já existentes, como uma informação que ajudará a resolver, confirmar ou aumentar a dúvida em relação ao ambiente. Mas especificamente no caso de um sistema baseado em PQRS, pode ser aplicado como auxiliar em sistema de mapeamento robótico, aliado a um sistema de reconhecimento de objetos, auxiliando no posicionamento qualitativo em relação a objeto reconhecidos neste ambiente ou mesmo no reconhecimento de localização utilizando técnicas estatísticas descartando a análise de regiões nas quais o raciocínio lógico sobre sombras já tenha descartado. Enfim a sombra é um componente que pode nos dar de modo qualitativo informações que podem vir a reduzir a complexidade de controle de sistemas robótico ou de análise de cenas.

# Referências Bibliográficas

- Balan, A. O., M. Black, H. Haussecker, and L. Sigal, Shining a light on human pose: on shadows, shading and the estimation of pose and shape, *IEEE 11th International Conference on Computer Vision - ICCV*, pp. 1–8, 2007.
- Bonfiglioli, C., F. Pavani, and U. Castiello, Differential effects of cast shadows on perception and action, pp. 1291–1304, doi:10.1068/5325, 2004.
- Brachman, R. J., and H. J. Levesque, *Knowledge Representation and Reasoning*, Cap 14, 285 - 304 pp., Morgan Kaufmann, San Francisco, CA, 2004.
- Cabalar, P., and P. Santos, Strings and holes: An exercise on spatial reasoning, in *Advances in Artificial Intelligence - IBERAMIA-SBIA 2006, Lecture Notes in Computer Science*, vol. 4140, pp. 419–429, Springer, 2006.
- Casati, R., *A descoberta da sombra - de Platão a Galileu, a história de um enigma que fascina a humanidade*, São Paulo, tradução Eduardo Brandão, 2001.
- Casati, R., The copycat solution to the shadow correspondence problem, *Perception*, 37 (4), 495 – 503, 2007.
- Cavanagh, P., The artist as neuroscientist, *Nature*, 434, 301–307, 2005.
- Chella, A., M. Liotta, and I. Macaluso, Cicerobot: a cognitive robot for museum tours, *IASTED International Conference on Robotics and Applications*, pp. 318–323, doi:10.1108/01439910710832101, 2005.
- Chucchiara, R., C. Grana, M. Piccardi, A. Prati, and S. Sirotti, Improving shadow suppression in moving object detection with HSV color information, in *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, pp. 334–339, 2001.

- Cui, Z., A. G. Cohn, and D. A. Randell, Qualitative simulation based on a logic of space and time, in *QR-92*, Heriot-Watt University, Scotland, 1992.
- da Vinci, L., *The Notebooks of Leonardo da Vinci*, vol. 1, Dover, 1970.
- Elder, J. H., S. Trithart, G. Pintilie, and D. MacLean, Rapid processing of cast and attached shadows, *Perception*, 33, 1319–1338, 2004.
- Fitzpatrick, P. M., and E. R. Torres-Jarra, The power of the dark side: Using cast shadows for visually-guided touching, *WSPC/Instruction file*, 2004.
- Frank Dellaert, W. B., Dieter Fox, and S. Thrun, Monte carlo localization for mobile robots, in *IEEE International Conference on Robotics and Automation (ICRA99)*, 1999.
- Gaskin, J. E., Whatever happened to artificial intelligence?: The grand promise of intelligent machines underestimated the complexity of reproducing human cognition., *Network World*, 2008.
- Gotts, N. M., How far can we ‘C’? defining a doughnut using connection alone, in *KR’94: Principles of Knowledge Representation and Reasoning*, edited by J. Doyle, E. Sandewall, and P. Torasso, pp. 246–257, Morgan Kaufmann, San Francisco, California, 1994.
- Hasenfratz, J. M., M. Lapierre, N. Holzschuch, and F. X. Sillion, A survey of real-time soft shadows algorithms, *Eurographics Star - State of The Art Report*, 2003.
- Imura, T., M. K. Yamaguchi, M. Tomonaga, and A. Yagi, Perception of motion trajectory from the moving cast shadow in human infants, *J. Vis.*, 5(8), 329–329, 2005.
- Intel, Intel®open source computer vision library.
- Joshi, A. J., S. Atev, O. Masoud, and N. Papanikolopoulos, Moving shadow detection with low- and mid- level reasoning, *IEEE Int. Conf. Robotics and Automation (ICRA 2007)*, 2007.
- Kanda, T., M. Kamasima, M. Imai, T. Ono, D. Sakamoto, H. Ishiguro, and Y. Anzai, A humanoid robot that pretends to listen to route guidance from a human, *Springer Science + Business Media, LLC 2006, Auton Robot*, doi:10.1007/s10514-006-9007-6, 2006.
- Kersten, D., P. Mamassian, and D. C. Knill, Moving cast shadows and the perception of relative depth., *Max-Planck-Institut fuer biologische Kybernetik, Technical Report No.6*, 1994.

- Knill, D. C., P. Mamassian, and D. Kersten, Geometry of shadows, *Journal of the Optical Society of America A*, 14, 3216–3232, 1997.
- Mamassian, P., Impossible shadows and the shadow correspondence problem, *Perception*, 33, 1279–1290, doi:10.1068/p5280, 2004.
- Mamassian, P., and R. Goutcher, Prior knowledge on the illumination position, in *Cognition*, vol. 81, pp. 1–9(9), Elsevier, doi:10.1016/S0010-0277(01)00116-0, 2001.
- Mamassian, P., D. C. Knill, and D. Kersten, The perception of cast shadows, *Trends in cognitive Sciences*, 2(8), 288–295, 1998.
- McCarthy, J., Programs with common sense, in *Proceedings of the Teddington Conference on the Mechanization of Thought Processes*, pp. 75–91, Her Majesty’s Stationary Office, London, 1959.
- McCarthy, J., Artificial intelligence, logic and formalizing common sense, in *Philosophical Logic and Artificial Intelligence*, pp. 161–190, Kluwer Academic, 1990.
- McCarthy, J., What is artificial intelligence, *Computer Science Department, Stanford University, Stanford, CA*, 2002.
- Mueller, E. T., Automating commonsense reasoning using the event calculus., *Communications of the ACM*, 2007.
- Muller, P., A qualitative theory of motion based on spatio-temporal primitives, in *KR’98: Principles of Knowledge Representation and Reasoning*, edited by A. G. Cohn, L. Schubert, and S. C. Shapiro, pp. 131–141, Morgan Kaufmann, San Francisco, California, 1998.
- Nezmzow, U., *Scientific Methods in Mobile Robotics*, 3, 8-9, 22 pp., 2006.
- Ni, R., M. L. Braustein, and G. J. Andersen, Perception of scene layout from optical contact, shadows, and motion, *Perception*, 33, 1305–1318, doi:10.1068/p5288, 2004.
- Patnaik, S., *Robot Cognition and Navigation: An Experiment with Mobile Robots (Cognitive Technologies)*, 6 pp., Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.
- Pereira, S. L., Planejamento abduativo no cálculo de eventos, Master’s thesis, Universidade de São Paulo, USP, orientadora: Leliane Nunes de Barros, 2002.

- Pirri, F., I. Mentuccia, and S. Storri, The domestic robot - a friendly cognitive system takes care of your home, in *Ambient Intelligence: Impact on Embedded System Design.*, edited by T. Basten, M. C. W. Geilen, and H. W. H. de Groot, Kluwer Academic Publishers, 2003.
- Polidorio, A. M., F. C. Flores, C. Franco, N. N. Imai, and A. M. G. Tommaselli, Detecção automática de sombras e nuvens em imagens CBERS e Landsat 7 ETM, *Anais XII Simpósio Brasileiro de Sensoriamento Remoto*, pp. 4233–4240, 2005.
- Prati, A., I. Mikic, M. M. Trivedi, and R. Cucchiara, Detecting moving shadows: Algorithms and evaluation, *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(7), 918–923, doi:<http://dx.doi.org/10.1109/TPAMI.2003.1206520>, 2003.
- Pérez, P., Markov random fields and images, in *CWI Quarterly*, 4, vol. 11, pp. 413–437, 1998.
- Rafael C. Gonzalez, R. E. W., *Digital Image Processing*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.
- Randell, D., Z. Cui, and A. G. Cohn, A spatial logic based on regions and connection, in *Proc. 3rd Int. Conf. on Knowledge Representation and Reasoning*, pp. 165–176, Morgan Kaufmann, San Mateo, 1992.
- Randell, D. A., and M. Witkowski, Building large composition tables via axiomatic theories, in *Proc KR*, pp. 26–36, 2002.
- Randell, D. A., M. Witkowski, and M. Shanahan, From images to bodies: Modelling and exploiting spatial occlusion and motion parallax, in *IJCAI*, edited by B. Nebel, pp. 57–66, Morgan Kaufmann, 2001.
- Rensink, R. A., and P. Cavanagh, The influence of cast shadows on visual search, *Perception*, 33, 1339–1358, doi:10:1068/5322, 2004.
- Rodrigues, P. S., G. A. Giraldo, R.-F. Chang, and J. S. Suri, *Object tracking in image sequence combining Hausdorff distance, nonextensive entropy in levelset formulation*, chap. 15, 2006.
- Russel, S., and P. Norvig, *Inteligência Artificial, segunda edição traduzida*, 835-867 pp., Ed. Campus, 2004.
- Santos, P., and M. Shanahan, Hypothesising object relations from image transitions, in *Proc. of ECAI-02*, edited by F. van Harmelen, pp. 292–296, Lyon, France, 2002.

- Santos, P., H. M. Dee, and V. Fenelon, Notes on a qualitative theory of shadows, *Cognitive Robotics workshop at ECAI 2008*, pp. 47–54, university of Patras - Patras, Greece, 2008.
- Santos, P., H. M. Dee, and V. Fenelon, Qualitative robot localisation using information from cast shadows, *ICRA-09 IEEE International Conference on Robotics and Automation*, 2009.
- Santos, P. E., Spatial reasoning and abductive interpretation of sensor data obtained by a mobile robot in a dynamic environment, Ph.D. thesis, Imperial College, 2003.
- Santos, P. E., Reasoning about depth and motion from an observer’s viewpoint, in *Spatial Cognition and Computation*, vol. 7, edited by L. Earlbbaum, pp. 133–178, doi:10.1080/13875860701418198, 2007.
- Scherl, R. B., and H. J. Levesque, The frame problem and knowledge-producing actions, in *Proceedings of the Eleventh National Conference on Artificial Intelligence (AAAI-93)*, pp. 689–697, AAAI Press/MIT Press, Washington, D.C., USA, 1993.
- Shanahan, M., *Solving the frame problem: a mathematical investigation of the common sense law of inertia*, MIT Press, Cambridge, MA, USA, 1997.
- Shanahan, M. P., The event calculus explained, in *Artificial Intelligence Today, Lecture Notes in AI no. 1600*, vol. 1600, edited by M. J. Woolridge and M. Veloso, pp. 409–431, Springer, 1999.
- Shapiro, L. G., and G. C. Stockman, *Computer Vision*, 371-409 pp., Tom Robbins, 2001.
- Simões, A. S., and A. H. Reali-Costa, Classificação de cores por redes neurais usando a representação hsv., in *International Joint Conference: Iberoamerican conference on artificial intelligence (IBERAMA) and Brazilian symposium on artificial intelligence (SBIA)*, vol. 1, pp. 19–22, Atibaia - SP, 2000.
- Stauder, J., R. Mech, and J. Ostermann, Detection of moving cast shadows for object segmentation, *IEEE Transactions on Multimedia*, 1(1), 65–76, 1999.
- Stock, O., *Spatial and Temporal Reasoning*, Kluwer Academic Publishers, Norwell, MA, USA, 1997.
- Witkowski, M., M. Shanahan, P. Santos, and D. Randell, Cognitive robotics: On the semantic knife-edge, in *Towards Intelligent Mobile Robots*, 2001.

- Boleano — relativo a lógica matemática criada por George Boole (1815 - 1864), uma variável do tipo boleano dá a resposta como verdadeiro ou falso.
- estado — do dicionário Houaiss - “...o conjunto de qualidades ou características com que as coisas se apresentam ou o conjunto de condições em que se encontram em determinado momento.”
- Heurística — do dicionário Houaiss - “...método de investigação baseado na aproximação progressiva de um dado problema...”
- Hardware — do dicionário Houaiss - “... conjunto dos componentes físicos (material eletrônico, placas, monitor, equipamentos periféricos etc) de um computador...”
- Hodômetro — do dicionário Houaiss - “... instrumento que indica distância percorridas por pedestres ou por veículos. Observação: forma não preferencial: odômetro.”
- Inteligência artificial — do dicionário Houaiss - “ ramo da informática que visa dotar os computadores da capacidade de simular certos aspectos da inteligência humana, tais como aprender com a experiência, inferir a partir de dados incompletos, tomar decisões em condições de incerteza e compreender a linguagem falada, entre outros.”
- Joystick* — é um periférico de computador e videogame pessoal ou um dispositivo geral de controle que consistem em uma vara vertical na qual os pivôs se aproximam de uma extremidade e transmitem seu ângulo em duas ou três dimensões a um computador.
- OpenCV* — É uma biblioteca que reúne funções em C e algumas Classe em C++, que executam vários algoritmos de visão computacional *Intel*.
- Paralaxe — do dicionário Houaiss “...deslocamento aparente de um objeto quando se muda o ponto de observação...”
- Penumbra — meia-luz, sombra parcial.
- PQRS — *Perceptual Qualitative Reasoning about Shadows* é uma teoria baseada em raciocínio espacial qualitativo sobre as relações de sombra e seu objeto. Apresenta pela primeira vez em Santos *et al.* (2008).
- Psicofísicos — do dicionário Houaiss “ Relativo a psicofísica. Psicofísica é o ramo da psicologia que estuda a relação entre estímulos físicos e as respectivas sensações.”
- Segmentação — ato de separar por característica.
- Software — do dicionário Houaiss - “...conjunto de componentes lógicos de um computador ou sistema de processamento de dados; programa, rotina ou conjunto de instruções que controlam o funcionamento de um computador; suporte lógico...”

# Apêndice A

## TRABALHOS PUBLICADOS

A seguir anexo cópia dos trabalhos publicados em congressos internacionais (anexo somente na versão final impressa):

**ECAI 2008** *Notes on a qualitative theory of shadows*

**ICRA 2009** *Qualitative robot localisation using information from cast shadows*