



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
DO CEARÁ
CAMPUS DE FORTALEZA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
TELECOMUNICAÇÕES**

Jéssyca Almeida Bessa

**LOCALIZAÇÃO DE ROBÔS MÓVEIS ATRAVÉS DE
RECONHECIMENTO DE PADRÕES EM IMAGENS
OMNIDIRECIONAIS**

**Fortaleza
2015**

Jéssyca Almeida Bessa

LOCALIZAÇÃO DE ROBÔS MÓVEIS ATRAVÉS DE
RECONHECIMENTO DE PADRÕES EM IMAGENS
OMNIDIRECIONAIS

Dissertação submetida à
Coordenação do Programa de
Pós-Graduação em Engenharia de
Telecomunicações do Instituto Federal
do Ceará, como requisito parcial
para obtenção do grau de Mestre em
Engenharia de Telecomunicações.

Área de concentração:
Processamento Digital de Sinais

Orientador: Auzuir Ripardo de
Alexandria

Fortaleza
2015

Instituto Federal de Educação, Ciência e Tecnologia do Ceará
Biblioteca Waldyr Diogo de Siqueira
Catalogação na fonte

B557l Bessa, Jéssyca Almeida

Localização de robôs móveis através de reconhecimento de padrões em imagens omnidirecionais. / Jéssyca Almeida Bessa. - Fortaleza: IFCE, 2015.
150 p.

Dissertação (Engenharia de Telecomunicações).

Orientador: Prof. Dr. Auzuir Ripardo de Alexandria.

1. Telecomunicações - Engenharia 2. Robôs móveis. 3. Imagens omnidirecionais
- Técnicas de reconhecimento de padrões. I. Título.
CDD 621.382



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELECOMUNICAÇÕES

Jéssyca Almeida Bessa

Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia de Telecomunicações, sendo aprovada pelo Programa de Pós-Graduação em Engenharia de Telecomunicações do Instituto Federal de Educação, Ciência e Tecnologia do Ceará e pela banca examinadora:

Orientador: Prof. Dr. Auzuir Ripardo de
Alexandria
Instituto Federal do Ceará - IFCE

Co-orientador: Prof. Dr. Ajalmar Rêgo da
Rocha Neto
Instituto Federal do Ceará - IFCE

Prof. Dr. Guilherme de Alencar Barreto
Universidade Federal do Ceará - UFC

Prof. Dr. Glendo de Freitas Guimarães
Instituto Federal do Ceará - IFCE

Fortaleza, 11 de agosto de 2015

Dedico este trabalho, primeiramente, a Deus e aos meus pais, pelo total apoio, dedicação e incentivo em todas as etapas da vida.

Agradecimentos

Primeiro agradeço a Deus por todas as oportunidades a mim dadas;
Aos meus pais, Farnésio e Vanda, pelo contínuo incentivo;
A minha irmã, Laryssa, por todo companheirismo e paciência;
Ao meu namorado, Darlan, por sempre me fazer olhar para frente, e seguir um passo de cada vez;
Ao meu orientador, Auzuir, pela confiança, ensinamentos e apoio;
Ao IFCE, Instituto Federal de Educação, Ciência e Tecnologia do Ceará, por todo o repasse de conhecimento;
Aos amigos e professores do curso de Pós Graduação em Engenharia de Telecomunicações, por todas as lições aprendidas;
Aos meus alunos e amigos que no IFCE cultivei;
A FUNCAP e ao CTI Renato Archer pelo apoio financeiro prestado ao longo do curso de Pós Graduação em Engenharia de Telecomunicações.

"Noventa por cento do sucesso se baseia simplesmente em insistir".

Woody Allen

Resumo

Localizar-se é uma das necessidades básicas de um ser vivo. Tal capacidade também é muito importante quando se deseja dar autonomia e mobilidade aos robôs, fato este evidenciado pelo crescimento de pesquisas nesta área e a grande variedade de métodos desenvolvidos. Dentre estes métodos, estão os métodos que utilizam técnicas de Reconhecimento de Padrões. Esta dissertação descreve um estudo comparativo para resolver o problema de localização de robôs móveis através de técnicas de Reconhecimento de Padrões em imagens omnidirecionais. Este trabalho é motivado pela implementação e comparação de técnicas de extração de características que possam ser utilizadas em imagens omnidirecionais, buscando a invariância à rotação do robô e formando descritores que possam ser utilizados em diferentes algoritmos de Reconhecimento de Padrões. No total, cinco técnicas de extração de características com suas respectivas adaptações para imagens omnidirecionais e cinco classificadores de padrões são apresentados e comparados. As taxas de acerto médio e o desvio padrão dos resultados para os conjunto de teste são calculados e avaliados. Os resultados são apresentados com intuito de escolher qual o extrator de característica e o classificador mais adequado para se utilizar na localização de robôs móveis. Os extratores de características são avaliados segundo os critérios: tempo de processamento e qualidade de descrição da cena que é mensurada através das taxas de acerto médio dos classificadores. Os resultados são satisfatórios e elegem os descritores GIST e PCA e a rede MLP como os algoritmos mais adequados à aplicação. As principais contribuições deste trabalho são as adaptações realizadas nos extratores de características a fim de se obter a invariância rotacional do robô para as imagens omnidirecionais e a utilização destas na obtenção de descritores utilizados em técnicas de Reconhecimento de Padrões para localização de robôs móveis. Outra contribuição é a formação de bancos de imagem criados a partir de um ambiente sintético e um ambiente real que podem ser utilizados em outras pesquisas.

Abstract

Knowing your location is a basic necessity of any living being. The growing number of researches and developed methods in this area prove that such necessity is also very important when you wish to give autonomy and mobility to robots. Among these methods are the ones that utilize Pattern Recognition techniques. This dissertation describes a comparison on the proposal to locate mobile robots by Pattern Recognition in omnidirectional images. This work is motivated by the implementation of comparison and feature extraction techniques that can be used in omnidirectional images, seeking invariance to rotation and building descriptors that can be used in different pattern recognition algorithms. Five feature extraction methods with their respective adaptations for omnidirectional images and five Pattern Recognition algorithms are presented and compared. The average hit rates and variance of the results for the test set are calculated. The results are presented in order to choose which feature extractor and the most suitable classifier to use the location of mobile robots. The feature extractors are evaluated on criteria: processing time and quality of description of the scene. These metrics are evaluated by the average hit rates of the classifiers. The results are satisfactory and elect the GIST descriptor, the PCA descriptor and the MLP network as the most suitable for implementation algorithms. The main contribution of this work is performed in the adjustments feature extractors in order to achieve rotational invariance for omnidirectional images and to use these to obtain descriptors used in pattern recognition techniques for locating mobile robots. Another contribution is the building of image databases created from a synthetic environment and a real environment that can be used in other researches.

Lista de Figuras

1	Diagrama de módulos componentes de um robô móvel autônomo. Adaptada de (SIEGWART; NOURBAKHS; SCARAMUZZA, 2011). . .	31
2	Entrada e saída de um sistema de visão computacional.	35
3	Sistema tradicional de Visão Computacional.	36
4	“Onde estou?”	38
5	Localização do robô utilizando odometria.	40
6	Esquema de um sistema de localização baseado em mapas.	42
7	Representação de um mapa topológico em um ambiente interno. . . .	44
8	Interação entre o mapa e a posição do robô.	46
9	O olho de um trilobite fóssil (SRINIVASAN; VENKATESH, 1997). . . .	51
10	Um esboço do olho composto de um inseto diurno. A cada faceta tem um fotorreceptor associado (BENOSMAN; KANG, 2001).	51
11	Um esboço do olho composto de um inseto noturno. Os raios de luz que entram em diferentes facetas estão focados no mesmo ponto na retina com uma refração (BENOSMAN; KANG, 2001).	52
12	Um esboço do olho composto de um crustáceo. Observe que os raios de luz de entrada em diferentes facetas estão focados no mesmo ponto na retina com uma reflexão (BENOSMAN; KANG, 2001).	52
13	Crustáceo <i>Gigantocypris</i> (APPELTANS <i>et al.</i> , 2012).	53
14	Representação em pinturas: (a) “O casamento de Giovanni Arnolfini” de Jan van Eyck e (b) o zoom no espelho bruxa na parede atrás do casal de noivos mostrando o resto da sala e do pintor (DRUCKER, 2013). . .	54
15	Uma litografia de Escher: <i>Hand with Reflecting Sphere</i> (ESCHER, 1935). .	54
16	Um exemplo de anamorfose por Awtar Singh Viridi, século XVI.	55

17	Panorama da Filadélfia em 1987.	55
18	Um dos primeiros modelos de câmera panorâmica.	56
19	Sistema omnidirecional com múltiplas câmeras.	59
20	Sistemas com múltiplas câmeras: (a) DodecaTM 1000 (MENEGATTI, 2002), (b) S.O.S. (Stereo Omnidirectional System)(SHIMIZU <i>et al.</i> , 2004), (c) RingCam (NANDA; CUTLER, 2001), (d) Visão N-estéreo (KANADE, 1999) (e) <i>Ladybug2 spherical digital video camera</i>	60
21	Sistema omnidirecional com câmera giratória	61
22	Imagem obtida através de câmera com lente “olho de peixe”.	62
23	Sistema catadióptrico com espelho esférico.	63
24	Centro de projeção: (a) espelho esférico (não há centro único de projeção), (b) espelho hiperbólico (os raios se interceptam no foco da hipérbole e (c) espelho parabólico (projeção ortográfica dos raios refletidos). A figura foi adaptada de (GRASSI, 2002).	64
25	Perfis de espelho: (a) espelho cônico, (b) espelho esférico, (c) espelho hiperbólico e (d) espelho parabólico.	66
26	Imagem obtida com um espelho esférico à esquerda e com um espelho projetado por Hicks e Bajcsy (2000) à direita.	67
27	Classificação de trabalhos sobre Visão Omnidirecional.	68
28	Levantamento de trabalhos cadastrados na plataforma <i>Google Scholar</i> na área de estudo.	68
29	Fases do do processo de reconhecimento de padrões.	70
30	Imagem colorida e matrizes RGB.	71
31	Imagem colorida com os respectivos histogramas de cada banda. . . .	72
32	Descritor HOG: (a)orientações dos gradientes e (b) áreas chamadas de células e suas respectivas direções das orientações de gradiente. . . .	73
33	Imagens de teste utilizadas: (a)imagem original, (b)imagem espelhada, (c)imagem transladada e reduzida e (d)imagem rotacionada e reduzida.	75

34	Análise de Componentes Principais: (a) Dados de entrada e (b) Dados reduzidos com PCA.	77
35	Descritor Gist: (a) imagem original, (b) transformada Fourier 2D calculada em várias regiões.	79
36	Arquitetura típica de uma rede neural supervisionada.	80
37	Modelo de uma rede SOM: (a) 2d e (b) 1d.	83
38	Representação do algoritmo K-vizinhos mais próximos.	85
39	Superfícies de decisão geradas por análise de discriminante quadrático em um conjunto de dados com 3 classes distintas.	87
40	Fluxo de informação do método de localização implementado.	89
41	Sistema catadióptrico montado para o sistema: (a) sistema catadióptrico e (b) vista do robô com o sistema catadióptrico.	90
42	Imagens sintéticas: (a) quadrado, (b) triângulo, (c) círculo, (d) estrela e (e) seta.	91
43	Plataforma <i>POV-Ray</i>	92
44	Ambiente sintético: (a) vista superior do ambiente e (b) vista omnidirecional.	92
45	Ambiente real: (a) vista superior do ambiente e (b) vista omnidirecional.	93
46	Sistema de navegação: (a) aplicativo portátil e (b) <i>software</i> de navegação.	94
47	Sistema catadióptrico: (a) câmera utilizada e (b) geometria do espelho usado.	94
48	Imagens do ambiente sintético: (a) omnidirecional e (b) panorâmica.	96
49	Cálculo do centróide de uma banda qualquer de uma imagem colorida.	96
50	Distribuição das células horizontais em uma imagem panorâmica para construir um descritor invariante a rotação do robô.	98
51	Descritor formado pelos momentos invariantes.	98
52	Imagens rotacionadas em cinco subespaços.	99
53	Exemplo de imagem filtrada com filtro de Gabor com $k_4 = 4$ e orientações em 0, 45, 90 e 135 graus em duas escalas.	100

54	Formação do descritor GIST.	101
55	Novas imagens de teste: (a) imagem com iluminação alterada e (b) imagem com iluminação alterada e submetida a ruído com variância igual a 0,02.	105
56	Histograma das novas imagens de teste: (a) iluminação original e (b) iluminação alterada.	105
57	Representação da metodologia utilizada neste trabalho.	108
58	Suporte para o sistema catadióptrico.	146
59	Dimensões do espelho usinado para o sistema.	147
60	Exemplos de imagens omnidirecionais típicas dos compartimentos do ambiente real.	148
61	Vista do robô com o sistema catadióptrico.	149

Lista de Tabelas

1	Classificação dos sensores utilizados em aplicações de robôs móveis. Baseada em (SIEGWART; NOURBAKHSH; SCARAMUZZA, 2011) . . .	34
2	Momentos invariantes para as imagens da Figura 33.	65
3	Momentos invariantes para as imagens da Figura 33.	76
4	Grupos de imagens sintéticas.	91
5	Grupos de imagens do ambiente sintético.	93
6	Grupos de imagens do ambiente real.	95
7	Quantidade de entradas para cada extrator de características.	102
8	Grupos de imagens do ambiente sintético.	102
9	Quantidade de imagens utilizadas nos testes com diferentes níveis de iluminação e ruído.	105
10	Tempo médio de execução dos algoritmos de extração de atributos. . .	110
11	Melhor configuração para cada extrator de características $\times$ grupo de imagens.	111
12	Resultados de treinamento e teste da MLP para imagens sintéticas com 80% dos dados para treinamento e 20% para teste.	111
13	Resultados de treinamento e teste da MLP para imagens de ambiente sintético com 80% dos dados para treinamento e 20% para teste. . . .	111
14	Resultados de treinamento e teste da MLP para imagens de ambiente real com 80% dos dados para treinamento e 20% para teste.	111
15	Acerto médio no teste e desvio padrão da rede SOM para cada extrator de características com 80% dos dados para treinamento e 20% para teste.	113

16	Melhor valor de K, acerto médio no teste e desvio padrão para o algoritmo KNN para cada extrator de características com 80% dos dados para treinamento e 20% para teste.	114
17	Acerto médio no teste e desvio padrão do algoritmo LDA para cada extrator de características com 80% dos dados para treinamento e 20% para teste.	114
18	Acerto médio no teste e desvio padrão do algoritmo QDA para cada extrator de características com 80% dos dados para treinamento e 20% para teste.	115
19	Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) e utilizando a melhor configuração da rede MLP obtida.	116
20	Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) utilizando a melhor configuração da rede SOM obtida.	116
21	Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) utilizando a melhor configuração do algoritmo KNN obtido.	116
22	Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) utilizando a melhor configuração do algoritmo LDA obtido.	117
23	Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) utilizando a melhor configuração do algoritmo QDA obtido.	117

Lista de Símbolos

$\max$	Máximo
$\min$	Mínimo
x,y	Coordenadas espaciais de uma imagem
M,N	Dimensões de uma imagem
$f(x,y)$	Função bidimensional denominada imagem
$\vec{f}$	Imagem vetorizada
p	Padrão ou Vetor Característico
n	Dimensão do Vetor
Q	Quantidade de padrões
D	Conjunto com Q padrões
r	Níveis de cinza
np	Quantidade de <i>pixels</i>
b	Número de <i>bits</i> da imagem
ϕ	Momentos invariantes
τ	Momentos invariantes centrais
γ	Fator de normalização dos momentos invariantes
η	Momentos invariantes normalizados
e	Número de imagens
$\vec{p}$	Projeção de imagem utilizadas no método PCA
C	Número de características PCA

P	Matriz de projeção PCA
V	Matriz de auto-vetores
F	Matriz com imagens dispostas em colunas
u	Ativação dos neurônios da camada oculta
a	Ativação dos neurônios da camada de saída
W, M	Pesos dos neurônios
q	Quantidade de neurônios da camada oculta
m	Quantidade de neurônios da camada de saída
ϕ	Função de ativação
v	Saída dos neurônios da camada oculta
s	Saída dos neurônios da camada de saída
d	Saída desejada
λ	Taxa de aprendizagem
ϑ	Raio de influência
i^*	Índice do neurônio vencedor
h	Função de vizinhança
K	Número de vizinhos mais próximos
ω	Classe ou agrupamento
μ	Média
$\hat{\Sigma}$	Matriz de covariância
$D_x \text{ e } D_y$	Máscaras para convolução
G	Gradiente
θ	Orientação de gradiente
k	Número de máscaras de Gabor

h	Número de células horizontais da imagem
acc	Acurácia
R	Número de realizações
σ	Desvio padrão
$f(z)$	função densidade de probabilidade
$F(z)$	função de distribuição acumulada
σ	desvio padrão
μ	média ou esperança matemática
$ \cdot $	operador magnitude
∇	operador gradiente

Lista de Abreviações

1D	Um dimensão, unidimensional
2D	Duas dimensões, bidimensional
CCD	<i>Charge-Coupled Device</i>
CMOS	<i>Complementary Metal-Oxide Semiconductor</i>
GPS	<i>Global Positioning System</i>
DR	<i>Dead-reckoning</i>
DGPS	<i>Differential GPS</i>
LM	Localização de Markov
LMC	Localização de Monte Carlo
SLAM	<i>Simultaneous Localisation and Mapping</i>
GVG	Gráfico de Voronoi Generalizado
EKF	<i>Extended Kalman Filter</i>
COPIS	<i>conic projection image sensor</i>
RNA	Redes Neurais Artificiais
SVP	<i>Single View Point</i> [NON-SVP] <i>Non-Single View Point</i>
RGB	<i>Red, Green and Blue</i>
PDF	<i>Probability Density Function</i>
HOG	<i>Histogram of Oriented Gradients</i>
PCA	<i>Principal Component Analysis</i>
MLP	<i>Multilayer Perceptron</i>

SOM	<i>Self Organizing Maps</i>
QDA	<i>Quadratic Discriminant Analysis</i>
LDA	<i>Linear Discriminant Analysis</i>

Sumário

1	Introdução	23
1.1	Objetivos	26
1.2	Motivação	26
1.3	Contribuições	26
1.4	Produção científica	27
1.5	Organização da dissertação	28
2	Robótica Móvel	30
2.1	Conceitos	30
2.2	Percepção do ambiente	32
2.2.1	Sistema de Visão Computacional	34
2.3	Localização e Mapeamento	37
2.3.1	Localização local e global	38
2.3.2	Métodos de localização	39
2.4	Conclusão do Capítulo	49
3	Imagens Omnidirecionais	50
3.1	Visão omnidirecional	50
3.1.1	Visão omnidirecional e sua inspiração em animais	50
3.1.2	Visão Omnidirecional nas artes	53
3.1.3	A arte panorâmica	55
3.1.4	Visão Omnidirecional na Ciência	56

3.2	Sistemas Catadióptricos	62
3.2.1	Propriedades dos espelhos	64
3.3	Conclusão do Capítulo	67
4	Reconhecimento de Padrões	69
4.1	Extração de Características	70
4.1.1	Histograma RGB	71
4.1.2	Histograma de Gradientes Orientados	72
4.1.3	Os momentos invariantes de Hu	74
4.1.4	Análise de componentes principais (PCA)	76
4.1.5	Essência da cena	78
4.2	Classificadores	79
4.2.1	Classificadores baseados em Redes Neurais	79
4.2.2	Classificadores Estatísticos	84
4.3	Conclusão do Capítulo	87
5	Materiais e métodos	88
5.1	Descrição do fluxo de informação para comparação dos métodos . . .	88
5.2	Aquisição de imagens	90
5.2.1	Imagens Sintéticas	90
5.2.2	Imagens do ambiente sintético	91
5.2.3	Imagens do ambiente real	93
5.3	Extração de características	95
5.4	Classificadores	101
5.5	Testes com imagens em diferentes níveis de iluminação e de ruído . . .	104
5.6	Métricas de avaliação	106
5.7	Conclusão do Capítulo	107

6	Resultados e discussões	109
6.1	Resultados para os métodos de extração de características	109
6.2	Resultados para os classificadores de padrões	110
6.2.1	Classificador MLP	110
6.2.2	Classificador SOM	112
6.2.3	Classificador KNN	113
6.2.4	Classificador LDA	114
6.2.5	Classificador QDA	115
6.3	Resultados para imagens com diferentes níveis de iluminação e de ruído	115
6.4	Conclusão do Capítulo	116
7	Conclusão e Trabalhos Futuros	118
	Referências Bibliográficas	122
	Apêndice A – Apêndice	146

INTRODUÇÃO

Os meios de comunicação veiculam informações sobre a inserção de robôs no cotidiano diariamente. O estudo da Robótica é um tema bastante relevante e atual, onde esta área de pesquisa e desenvolvimento apresentou um grande salto nas últimas duas décadas. A aplicação de robôs junto a diferentes atividades em nossa sociedade vem demonstrando o quão promissor é esta área.

Um robô é um conjunto de dispositivos mecânicos, elétricos e eletrônicos (BEKEY; YUH, 2007). Estes são cada vez mais utilizados para substituir atividades repetitivas, insalubres e perigosas realizadas pelos seres humanos como, por exemplo, em situações que possam provocar algum dano à saúde humana que é o caso da exploração de ambientes hostis com a presença de gases tóxicos (HAYES; MARTINOLI; GOODMAN, 2002), inspeção de tubulações para detecção de vazamentos e a medição da concentração de gases (KAWAGUCHI *et al.*, 1995; DUCKETT; AXELSSON; SAFFIOTTI, 2001; Di Rocco; REGGENTE; SAFFIOTTI, 2011), locais com temperatura elevada ou expostos à radiação (MONKMAN; TAYLOR, 1993; RUSSELL, 1997), detecção de bombas em zonas de guerra e limpeza de reatores nucleares (WAGNER; LINDENBAUM; BRUCKSTEIN, 1999).

No campo educacional, podem-se observar algumas organizações que realizam competições de futebol de robôs com objetivo de promover o desenvolvimento da Robótica, Visão Computacional e da Inteligência Artificial (COSTA; PEGORARO, 2000; KUO; YANG; YANG, 2005; WANG *et al.*, 2013). Essa grande variedade de tarefas que podem ser realizadas por robôs tem impulsionado o interesse por esse campo do conhecimento, principalmente robôs móveis.

Robôs móveis são programados para realizar tarefas de maneira autônoma (MCGILL *et al.*, 2012). A autonomia de cada robô é condicionada pela capacidade de percepção do ambiente e interação coerente com o meio externo, sendo capaz de obter informações úteis, usando-as para realização de alguma tarefa de interesse. Um problema importante em robótica móvel é a estimação da localização (SARIFF;

BUNIYAMIN, 2006; OBDZALEK, 2009; KUO; CHEN, 2010).

Para este problema, faz-se necessária a extração de informações do meio e o fornecimento de dados para o robô se localizar no ambiente e navegar de forma segura. Dada a sua importância, o desenvolvimento de técnicas para localização de robôs móveis constitui uma das maiores tendências em pesquisas na área da Robótica (MENGLONG; XIAOLEI, 2011; WEI; LEE, 2011; ZHAO; WANG, 2012; HERNANDEZ *et al.*, 2014; SHEIKH; JAMIL; AYAZ, 2014). A autonomia dos robôs móveis é importante e necessária para que os mesmos possam localizar-se e planejar sua trajetória. Esta também é essencial quando o robô precisa reagir com rapidez às mudanças do ambiente ou a algum estímulo recebido (ANTONELO *et al.*, 2005; ZHANG, 2011).

Além disso, é indispensável a constante obtenção de informações do mundo externo por meio de sensores, possibilitando uma adequada interação entre o robô e seu ambiente de navegação. Dessa maneira, é interessante para o sistema de localização obter a maior quantidade de informações sobre o ambiente. Existem inúmeros tipos de sensores utilizados em aplicações de Robótica Móvel, entre estes os que são baseados em ultrassom, laser e Visão Computacional (SORIA *et al.*, 2003). A vantagem do uso de um sistema de Visão Computacional é a possibilidade de captura das imagens de forma contínua e o fornecimento da posição dos objetos no ambiente.

Os sistemas de Visão Computacional visam emular a visão biológica. Tais sistemas são responsáveis pela aquisição e extração de pontos de interesse, provendo ao computador a capacidade de percepção do ambiente sobre o qual atua. A área de Visão Computacional está em crescente ascensão, pois ela é essencial para o desenvolvimento de outras tecnologias, tal como a Robótica Móvel. Nesta, sistemas de visão são fundamentais para o reconhecimento do ambiente no qual o robô atua, sendo responsáveis por extrair dados para que se possa realizar as tomadas de decisões.

Os métodos que utilizam Visão Computacional para estimar a localização podem ser divididos segundo o critério do posicionamento da câmera, a saber: *on-board*, quando esta é posicionada a bordo do robô, ou *off-board*, quando a câmera é posicionada fora do robô (CARVALHO *et al.*, 2006).

Os métodos *off-board* podem estimar a posição do robô de forma direta, ou seja, estes métodos não precisam medir um outro atributo do robô, a exemplo da velocidade, e estimar a posição a partir dele (CARVALHO *et al.*, 2006). Esses métodos

têm a sua utilização restrita a ambientes pequenos e sem objetos que possam encobrir o robô, devido ao campo de visão limitado da câmera (CARVALHO *et al.*, 2006).

Os métodos *on-board* são usados em várias aplicações, sobretudo naquelas em ambientes externos, podendo fornecer uma boa estimativa da posição nestas aplicações, onde o uso de uma câmera *off-board* é inviável (CAMPBELL *et al.*, 2005).

Indo mais além, observa-se que algumas aplicações necessitam de um campo visual mais abrangente do que o obtido por uma câmera convencional. Sistemas de visão omnidirecional (do latim “*omnis*”, em português, “*oni*”, significa tudo ou todo) tem sido amplamente utilizados por aumentar o número de informações disponíveis. Um sistema de visão panorâmica ou omnidirecional fornece uma imagem de 360° do ambiente, que pode ser necessário, por exemplo, quando a tarefa a ser executada pelo robô necessita de informações visuais constantes de objetos localizados em diferentes (NAYAR, 1997).

A visão omnidirecional tem sido utilizada em inúmeras aplicações, dentre elas, destacam-se a vigilância remota (FLUSSER; SUK, 1994), a reconstrução tridimensional do ambiente (BUNSCHOTEN; KROSE, 2003), o rastreamento visual (BONARINI; ALIVERTI; LUCIONI, 1999) e a navegação de robôs móveis (WAHLGREN; DUCKETT, 2005). Os sistemas de visão omnidirecional tem sido propostos e construídos de formas diferentes, oferecendo uma alternativa à visão tradicional (YASUSHI, 1999a), entre estes, estão os sistemas catadióptricos. Tais sistemas catadióptricos são compostos por câmera e um espelho convexo, ou seja, se baseiam no princípio da refração(lente) e reflexão(espelho) (NAYAR; BAKER, 1997). Sistemas catadióptricos podem ser utilizados para calcular as distâncias entre o robô e os objetos no ambiente, ou até mesmo a localização de um robô no ambiente.

Localizar-se é a mais básica necessidade de um ser vivo. Tal capacidade também é muito importante quando se deseja dar autonomia e mobilidade aos robôs, fato este evidenciado pelo crescimento de pesquisas nesta área e a grande variedade de métodos desenvolvidos. Entretanto, dentre os muitos modelos e estratégias de localização que os robôs podem usar para este fim, este trabalho é focado na localização de robôs móveis baseada em Reconhecimento de Padrões, os quais são representações internas do *layout* espacial do ambiente explorado pelo robô.

1.1 Objetivos

O principal objetivo deste trabalho consiste em propor um método de localização de robôs móveis em um mapa topológico utilizando técnicas de Reconhecimento de Padrões aplicadas em descritores extraídos de imagens omnidirecionais. Os descritores são obtidos através de técnicas de extração características que também são comparadas entre si. Desta forma, os objetivos específicos deste trabalho são os seguintes:

- criar um banco de imagens omnidirecionais a partir de um ambiente virtual;
- criar um banco de imagens omnidirecionais a partir de um ambiente real;
- implementar e avaliar algoritmos de extração de características de imagens omnidirecionais, propiciando a estes algoritmos invariância à rotação do robô;
- implementar e avaliar classificadores de padrões para serem utilizados na localização de robôs móveis em um mapa topológico;
- comparar e identificar os melhores classificadores de padrões em imagens omnidirecionais afim de localizar um dispositivo robótico. A comparação será realizada levando em consideração o tempo de processamento, poder de descrição (taxa de acerto médio) e desvio padrão dos resultados.

1.2 Motivação

O crescente uso e atribuição de tarefas a robôs móveis tem motivado pesquisas que objetivam aumentar o grau de autonomia com que estes agentes realizam suas tarefas. Tendo em vista que a localização é a tarefa mais básica e necessária para a realização das demais tarefas, reconhecer padrões é uma ferramenta extremamente útil, tanto para propósitos futuros de navegação, quanto para planejamento e otimização de trajetórias.

1.3 Contribuições

As principais contribuições deste trabalho são, do ponto de vista tecnológico, a originalidade do uso de técnicas de Reconhecimento de Padrões em imagens

omnidirecionais para localização de robôs móveis em ambientes não estruturados; e, do ponto de vista científico, as modificações dos algoritmos de extração de características para imagens omnidirecionais, conferindo invariância à rotação do robô aos mesmos. A avaliação dos extratores de características e dos classificadores para localização de robôs móveis se deu usando imagens obtidas a partir de um ambiente sintético e de um ambiente real.

Por fim, mas não menos importante, outra contribuição desse trabalho é a criação de dois bancos de imagens omnidirecionais a partir de dois tipos de ambientes: sintético e real (BESSA, 2015b) (BESSA, 2015a).

Assim, a partir desta pesquisa, outros trabalhos podem ser desenvolvidos, tanto para melhorar o que foi proposto, como para utilizar a ideia em outras aplicações como mapeamento e navegação de robôs móveis. Em resumo, as contribuições deste trabalho são listadas a seguir.

- criação de bancos de imagens omnidirecionais a partir de um ambiente virtual e de um ambiente real;
- estudo comparativo dos algoritmos de extração de características com suas respectivas adaptações que propiciam invariância à rotação do robô;
- estudo comparativo dos classificadores de padrões que são utilizados na localização de robôs móveis em um mapa topológico; e
- identificação da viabilidade do uso das técnicas utilizadas para aplicação em tempo real.

1.4 Produção científica

Em termos de produção científica, durante a realização deste trabalho, foram publicados os artigos listados a seguir.

1. **BESSA, Jéssyca Almeida**; CORTEZ, Paulo César; FÉLIX, John Hebert da Silva; NETO, Ajalmar Rêgo da Rocha; ALEXANDRIA, Auzuir Ripardo de. *Radial snakes: Comparison of segmentation methods in synthetic noisy images*. *Expert Systems with Applications*, Elsevier, v. 42, n. 6, p. 3079-3088, 2015.

2. ALEXANDRIA, Auzuir Ripardo De; CORTEZ, Paulo César; **BESSA, Jessyca Almeida**; FÉLIX, John Hebert da Silva; ABREU, José Sebastião De; ALBUQUERQUE, Victor Hugo C De. *pSnakes: A new radial active contour model and its application in the segmentation of the left ventricle from echocardiographic images. Computer methods and programs in biomedicine*, Elsevier, v. 116, n. 3, p. 260-273, 2014.
3. CARNEIRO, Romulo Ferrer L; **BESSA, Jessyca Almeida**; MORAES, Jermana Lopes de; NETO, Edson Cavalcanti; ALEXANDRIA, Auzuir Ripardo de. *Techniques of binarization, thinning and feature extraction applied to a fingerprint system. International Journal of Computer Applications*, v. 103, n. 10, 2014.
4. ALEXANDRIA, Auzuir Ripardo de; CORTEZ, Paulo Cesar; FELIX, John Hebert da Silva; OLIVEIRA, Tiberio Menezes de; GIRAO, Anaxagoras Maia; FROTA, Joao Batista Bezerra; **ALMEIDA, Jessyca. An OCR system for numerals applied to energy meters. Latin America Transactions, IEEE (Revista IEEE America Latina)**, IEEE, v. 12, n. 6, p. 957-964, 2014.
5. MORAES, Jermana Lopes; **BESSA, Jéssyca Almeida**; NETO, Edson Cavalcanti; BESSA, Renan; MOURA, Lorena Braga; ALEXANDRIA, Auzuir Ripardo de. Sistema de aquisição de dados para a máquina de impacto charpy. *HOLOS*, v. 4, p. 224-240, 2014.
6. **Jéssyca A. Bessa**, Ajalmar R. Rocha Neto, Auzuir R. de Alexandria (2015), “*Pattern recognition applied to omnidirectional images towards mobile robot localization in unstructured environment.*”, artigo submetido para a revista *Applied Soft Computing* e aguardando avaliação.

1.5 Organização da dissertação

No **Capítulo 2** é realizada uma revisão sobre Robótica Móvel, no qual conceitos relativos a percepção do ambiente, localização e mapeamento para robôs móveis é discutida. Um sistema de Visão Computacional típico é apresentado com seus elementos e funções, bem como aplicações. As principais técnicas de localização são revisadas e uma discussão acerca da evolução dos trabalhos sobre Reconhecimento de Padrões para localização de robôs móveis é iniciada.

No **Capítulo 3** é realizada uma revisão sobre Visão Omnidirecional, apresentando suas principais vertentes: nos animais, nas artes e na ciência. Uma breve discussão sobre os modos de captura de imagens omnidirecionais é realizada. Assim, os sistemas catadióptricos são apresentados, dando-se destaque às propriedades dos espelhos. O capítulo é concluído com uma classificação geral de todos os trabalhos que envolvem imagens omnidirecionais e um levantamento de trabalhos encontrados sobre o assunto.

Descreve-se, no **Capítulo 4**, as técnicas de extração de características e técnicas de Reconhecimento de Padrões utilizados no trabalho. Cria-se um vínculo com o **Capítulo 5**, onde estão descritos, de maneira quantitativa e qualitativa, cada algoritmo e cada técnica aplicado no trabalho. Os resultados obtidos em termos de tempo de processamento, qualidade de descrição de cena e acurácia são apresentados no **Capítulo 6**.

No **Capítulo 7**, por fim, são apresentadas as conclusões e contribuições finais deste trabalho e as indicações de trabalhos que podem dar continuidade a esta pesquisa e desenvolvimento.

ROBÓTICA MÓVEL

Neste capítulo são apresentados alguns conceitos sobre Robótica Móvel e as principais técnicas de percepção e localização utilizadas em robôs móveis.

2.1 Conceitos

A Federação Internacional de Robótica prevê que até 2020, mais de 15 milhões de robôs de serviços serão vendidos para uso pessoal. Tecnologias de sensores de última geração, tais como LIDAR (da sigla inglesa “*Light Detection And Ranging*”) e visão estéreo estão permitindo que esses dispositivos complexos tornem-se cada vez mais comuns (ROBOTICS; ASSOCIATION *et al.*, 2015).

Os robôs móveis contribuem em contextos de realização de tarefas insalubres e perigosas para o homem. Os militares usam sistemas robóticos para tarefas perigosas, como desativar bombas ou caminhar em ambientes hostis (WEISBIN *et al.*, 1999; VOTH, 2004; SINGER, 2009; MARCHANT *et al.*, 2011). O Departamento de Defesa dos Estados Unidos tem feito vários investimentos para diversos projetos de robôs autônomos militares como os robôs militares TALON (VOTH, 2004) que são veículos potentes, leves, controlados e que são amplamente utilizados para eliminação de explosivo, reconhecimento, comunicações, segurança, defesa e salvamento. Eles possuem capacidades anfíbias e podem navegar virtualmente em qualquer terreno. Eles são prestadores de ajuda na ocorrência de desastres naturais e/ou participantes em ações de reforço da paz (LEYVA, 2012).

Os agricultores usam robôs móveis para realizar tarefas pesadas, incluindo a colheita, a coleta de dados de culturas, capina e micro aspersão (SISTLER, 1987; NOGUCHI; TERAOKA, 1997; HOLLINGUM, 1999; HAGRAS *et al.*, 2002; SOUSA *et al.*, 2003). Na área da saúde, hospitais usam robôs móveis no auxílio em atividades laboratoriais (KREBS *et al.*, 1998, 1999; PASIC *et al.*, 2010).

Os robôs móveis também são usados para executar tarefas rotineiras ao redor da casa, como aspiração e limpeza de piscinas e calhas (YOU *et al.*, 2003; LEE *et al.*, 2005; FORLIZZI; DISALVO, 2006; PARK *et al.*, 2007; TIAN *et al.*, 2007). No entanto, não são somente essas as possíveis aplicações, há uma gama enorme de possibilidades de aplicação.

Dentre a grande variedade de robôs móveis, estão os robôs móveis autônomos. Segundo Bekey (2005), um robô móvel autônomo é um dispositivo mecânico montado sobre uma base não fixa que age sob o controle de um sistema computacional, equipado com sensores e atuadores que o permitem interagir (sentindo e reagindo) com o ambiente com certo nível de autonomia. A interação com o ambiente se dá através de ciclos de percepção-ação. Um diagrama ilustrando os principais blocos que constituem um robô móvel autônomo de uso geral é exibido na Figura 1.

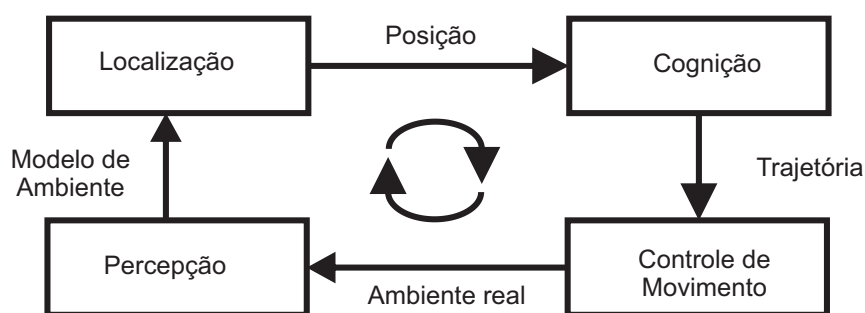


Figura 1: Diagrama de módulos componentes de um robô móvel autônomo. Adaptada de (SIEGWART; NOURBAKHS; SCARAMUZZA, 2011).

Assim, um robô móvel autônomo deve ter a habilidade de estimar sua posição o tempo todo. Dada uma posição final, baseado na cadeia de informações adquiridas, um robô deve modular a saída do motor para realizar a trajetória desejada. Essas informações conferem a necessidade de percepção do ambiente, auto-localização do robô e formação do mapa do ambiente de navegação. Assim, a característica principal que define um robô móvel autônomo é a capacidade de agir com base nas suas próprias decisões, e não por meio do controle de um ser humano (MATARIĆ, 2007). Leonard e Durrant-Whyte (1991), de maneira sucinta, subdividem o problema geral de navegação de robôs móveis autônomos em três questões, cada uma referente a uma subcategoria. São elas:

- Onde estou?: Essa questão constitui o problema de localização, que pode ser

definido como: Um robô está em uma posição desconhecida no ambiente. Ele “olha” sobre a sua posição e, com base nessas observações, deve perceber o lugar onde poderia estar.

- Para onde vou?: O problema consiste em mapear de um ambiente desconhecido pelo robô. Assim, navegando por este ambiente, há a criação de um mapa.
- Como faço para chegar lá?: O problema de planejamento de trajetória indica uma preocupação geral de como encontrar caminhos para conectar diferentes locais em um ambiente.

A navegação robótica é um problema difícil e complexo, com muitas soluções e abordagens diferentes. Por que navegação de robôs móveis e autônomos de maneira robusta e confiável continua a ser um problema tão difícil? Em Leonard e Durrant-Whyte (1991), é comentado que a dificuldade não surge do processo de navegação por si só, mas que o problema reside na falta de confiança na aquisição ou extração de informação, e a correlação ou correspondência automática dessa informação com a localização. Isto é o que faz com que o problema de localização e navegação autônoma seja um tema bastante ativo atualmente.

2.2 Percepção do ambiente

A percepção é o processo pelo qual os robôs mapeiam as medições dos sensores em representações internas do ambiente. A percepção é difícil porque, em geral, os sensores são ruidosos e o ambiente é frequentemente dinâmico. Como uma regra prática, as boas representações internas tem três propriedades: (i) contêm informações suficientes para o robô tomar as decisões corretas, (ii) são estruturadas de tal modo que possam ser atualizadas com eficiência, e (iii) são naturais, no sentido de que as variáveis internas correspondem as variáveis de estados naturais no mundo físico (KURFESS, 2004).

Um sensor é um dispositivo que pode registrar algum aspecto do ambiente e repassá-lo como entrada a um programa de agente. O sensor poderia ser bastante simples, tal como um sensor de um *bit* que detecta se um interruptor está ligado ou desligado, ou complexo como a retina do olho humano, que contém mais de cem milhões de células. Assim, tem-se a necessidade de se citar os sensores mais

comumente utilizados por robôs móveis e como extrair informações (NICHOLLS; LEE, 1989).

Há uma ampla variedade de sensores utilizados em robôs móveis. Alguns sensores são usados para medir valores simples, tais como a temperatura interna ou a velocidade de rotação dos motores. Outros sensores, tais como os baseados em Visão Computacional, mais sofisticados, podem ser utilizados para adquirir informações sobre o ambiente do robô ou mesmo para medir diretamente a posição global do mesmo (SIEGWART; NOURBAKHSH; SCARAMUZZA, 2011).

Os sensores podem ser classificados em passivos e ativos. Os sensores passivos não possuem fonte própria de energia, apenas detectam a energia do ambiente. Exemplos de sensores passivos incluem sondas de temperatura, microfones, câmeras CCD (*Charge-Coupled Device*) ou CMOS (*Complementary Metal-Oxide Semiconductor*). Já os sensores ativos emitem energia para o meio ambiente, em seguida medem a reação do meio ambiente. Como os sensores ativos podem gerenciar e controlar as interações com o meio ambiente, muitas vezes obtêm um desempenho superior. No entanto, a detecção ativa introduz vários riscos como a interferência entre o seu sinal e aqueles que estão além de seu controle. Por exemplo, os sinais emitidos por outros robôs próximos, ou semelhante sensores no mesmo robô, podem influenciar as medições resultantes. Exemplos de sensores ativos incluem codificadores de quadratura (utilizados em rodas) e sensores ultrassônicos (SIEGWART; NOURBAKHSH; SCARAMUZZA, 2011). Uma taxonomia dos sensores mais utilizados em aplicações robóticas em grupos e a classificação entre ativos e passivos é apresentada na Tabela 1. Nesta tabela, mostra-se também um exemplo de aplicação para cada tipo de sensor mencionado.

A visão é um sentido bastante poderoso (RUSSELL, 1997). Ela nos fornece uma enorme quantidade de informações sobre o ambiente e permite a interação rica e inteligente em ambientes dinâmicos. Portanto, não é surpreendente que um grande esforço tenha sido dedicado a máquinas fornecendo sensores que imitam as capacidades do sistema de visão humana. O primeiro passo neste processo é a criação de dispositivos sensores que capturam a mesma informação de luz que o sistema de visão humana.

Tabela 1: Classificação dos sensores utilizados em aplicações de robôs móveis. Baseada em (SIEGWART; NOURBAKHS; SCARAMUZZA, 2011)

Classificação Geral (uso típico)	Sensor	Passivo(P) ou Ativo(A)
Sensores táteis	Interruptores (MORISHITA; FUKUI; SATO, 2002)	P
	Barreiras óticas (NOLFI; PARISI, 1995)	A
	Sensores de proximidade (SAGÜES; MONTANO; ARAGÓN,)	A
Sensores de roda ou motor	Potenciômetros (EDSINGER-GONZALES; WEBER, 2004)	P
	Resolvers (LYNCH; PESHKIN, 2014)	A
	Enconders ópticos (BORENSTEIN <i>et al.</i> , 1996)	A
	Enconders magnéticos (BORENSTEIN <i>et al.</i> , 1996)	A
	Enconders indutivos (BORENSTEIN <i>et al.</i> , 1996)	A
	Enconders capacitivos (ZHENG <i>et al.</i> , 2015)	A
Sensores de direção	Bússola (XUE <i>et al.</i> , 2014)	P
	Giroscópio (FULLER <i>et al.</i> , 2014)	P
	Inclinômetro (CHEN <i>et al.</i> , 2015)	P/A
Sensores baseados em um "Feixe"	GPS (TAKAI; YANG; NOGUCHI, 2014)	A
	"Feixe" Ativo Óptico ou RF (PIRKL <i>et al.</i> , 2012)	A
	"Feixe" Ativo Ultra-sônico (ECKERT; GERMAN; DRESSLER, 2011)	A
	"Feixe" Reflexivo (ARRAS; VESTLI, 1998)	A
Sensores de alcance ativo	Sensores Reflexivos (RUAN; LI, 2014)	A
	Sensores Ultra-sônicos (OHYA; KOSAKA; KAK, 1998)	A
	Lasers (HIREMATH <i>et al.</i> , 2014)	A
	Triangulação Óptica 1D (MARQUES; NUNES; ALMEIDA, 1998)	A
	Luz Estruturada 2D (LIU <i>et al.</i> , 2015)	A
Sensores de som ou velocidade	Radar Doppler (KIM; SEKHON, 2015)	A
	Som Doppler (GUDELOGLU <i>et al.</i> , 2015)	A
Sensores baseados em visão	Câmera CCD (LI; ZHU; HUA, 2015)	P
	Câmera CMOS (GUZEL; ERWIN; ZAKARIA, 2015)	P

2.2.1 Sistema de Visão Computacional

Grande parte dos seres vivos em nosso planeta possuem a capacidade de visão. Esta é uma capacidade de grande importância no processo evolutivo destes seres e é usada no monitoramento da realidade que os rodeia, contribuindo assim, para a análise do ambiente e para a tomada de decisões que auxiliem na sobrevivência da sua espécie (HEIJDEN, 1994).

Homens e animais possuem um sofisticado e complexo sistema de visão que utiliza luz visível como modalidade (MORITA *et al.*, 1989). Em particular, sendo um ser inteligente e sabendo da importância da visão para a sua sobrevivência, o homem desenvolveu, com o passar do tempo, várias formas de representação das cenas que o rodeiam, denominadas de imagem. Estas estão dispostas das mais diversas formas: pintura, escrita, fotografia, cinema, televisão, entre outras e a utilização destas imagens estão nos mais diversos setores, desde a representação artística até aplicações bélicas (ANDRADE; MATOS; FREIRE, 2006).

De maneira simplificada, a aquisição de imagem se dá quando determinado objeto é iluminado por uma fonte de luz. Assim, a imagem é captada por um elemento

sensor que é capaz de quantificar o quanto de luz é refletida por esse objeto em uma superfície plana, como por exemplo, um filme fotográfico (CORSON; MOSS, 1986). Assim, o termo imagem faz referência a função bidimensional de intensidade luminosa $f(x,y)$, em que x e y são as coordenadas espaciais e o valor de f em um ponto (x,y) qualquer, é conhecido como nível de cinza da imagem, para imagens monocromáticas (GONZALEZ, 2009).

Imagens digitais são objetos de estudo de uma grande área da Engenharia conhecida como visão artificial (ou visão computacional). Sistemas de visão computacional, normalmente dispõem das imagens digitais como entrada e fornecem atributos da cena correspondente, conforme ilustrado na Figura 2 (MEKKAOU; BENAMEUR; OURIACHI, 1995).

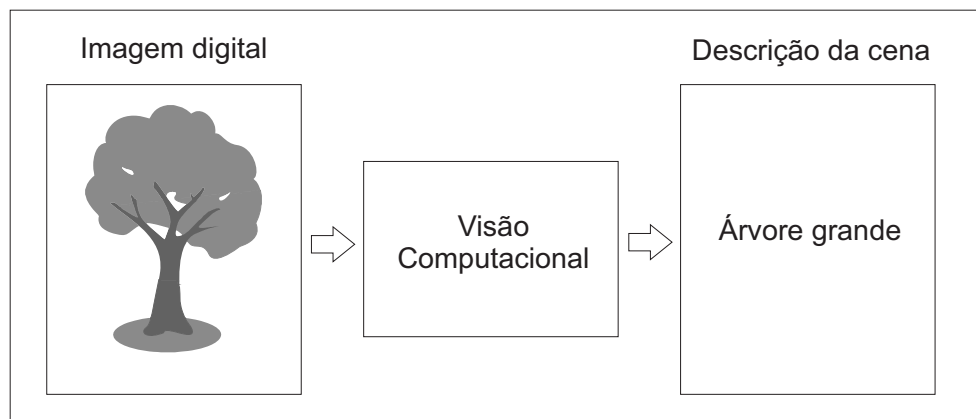


Figura 2: Entrada e saída de um sistema de visão computacional.

Com o avanço dos computadores digitais, surge uma nova forma de representação de uma cena: a imagem digital. Neste caso, a matriz bidimensional $f(x,y)$ tem cada elemento conhecido como *pixel* (*picture element* ou elemento de figura). Em cada *pixel*, localiza-se a intensidade da radiação refletida pelos objetos da cena em questão. Uma imagem digital é formada por um grupo de pontos discretos de tons de cinza que são geralmente organizados em uma matriz de *pixels* (GONZALEZ, 2009).

A necessidade de fazer com que máquinas processem imagens e reconheçam padrões nelas vem se tornando uma aplicação cada vez mais frequente (AYACHE, 1991). A visão computacional ganha espaço e se insere no cotidiano de diferentes formas. Sistemas de segurança que reconhecem os usuários através de imagens (WEN; ZHONGRU-WU; SU, 2004; RACHEL; CUGNASCA, 2006; NETO *et al.*, 2015), processamento de imagens SAR (FISHER *et al.*, 1999), sistemas de navegação

autônoma (AHRARY *et al.*, 2005; SILVA *et al.*, 2008; DIOSI *et al.*, 2011) e sistemas de soldagem (NANDHITHA *et al.*, 2007) são algumas das aplicações que envolvem Visão Computacional.

Segundo Marques Filho e Vieira Neto (1999), um sistema de visão artificial é definido como um sistema computadorizado capaz de adquirir, processar e interpretar imagens correspondentes a cenas reais. Considerando a descrição anterior, um sistema típico de visão computacional pode ser representado conforme a Figura 3, em que se encontra subdividido em vários processos.

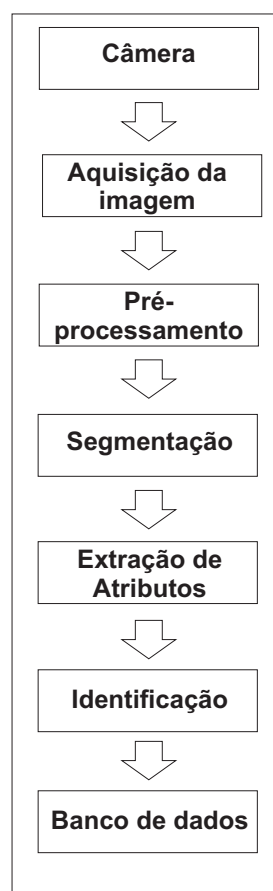


Figura 3: Sistema tradicional de Visão Computacional.

Conforme já mencionado, o primeiro processo em um sistema de visão é a aquisição de imagens. Nenhum algoritmo de condicionamento de imagem ou classificação de cena seria aplicado em um sistema sem a captura da imagem. Um sistema de visão computacional pode ser composto por diferentes formas de se capturar imagens (TORFEH-ISFAHANI *et al.*, 1989). Em aplicações de tempo real, é necessário o uso de uma câmera para que se possa capturar imagens que servem de fonte de informações para alimentar um determinado sistema (CHIU; DEPOMMIER;

SPINDLER, 2004)

Um problema inerente ao uso de câmera para captura de imagens é a limitação do campo de visão da mesma. Assim sendo, faz-se necessário o estudo de formas para contornar este problema, tornando possível o aumento do campo de visão, para que estas consigam melhorar seu desempenho e do sistema como um todo (ISHIGURO; YAMAMOTO; TSUJI, 1992). O desenvolvimento de sistemas que utilizam imagens panorâmicas (imagens que capturam ou tentam capturar informações em 360° de um ambiente, também chamadas de imagens omnidirecionais) vem sendo usado como uma resposta a estas limitações (CAO; OH; HALL, 1985). O Capítulo 3 foi direcionado inteiramente à discussão desse tipo de sensoriamento (visão omnidirecional).

2.3 Localização e Mapeamento

A localização é um exemplo genérico de percepção dos robôs. Trata-se do problema de determinar onde o robô está ou a localização dos objetos que compõem a cena. Como dito anteriormente, a localização é um dos problemas clássicos em robótica móvel, porque o conhecimento da localização é o núcleo de qualquer interação física bem-sucedida. A localização é o processo de estimação de onde o robô está relativamente a algum modelo do ambiente, usando qualquer sensor que esteja disponível. Como o robô continua se movendo, a estimativa da sua posição sofre mudanças (MATARIĆ, 2007) e atualizações sobre a posição do robô precisam ser realizadas. Essas atualizações são realizadas com base no reconhecimento de características em pontos de referência, nos dados do sensor e em modelos probabilísticos.

Localização de robôs é o problema-chave para tornar robôs verdadeiramente autônomos (COX, 1991; THRUN *et al.*, 2001; NEGENBORN, 2003). Se um robô não sabe onde está em relação ao ambiente, é difícil decidir o que fazer. O robô precisará pelo menos ter alguma ideia de onde ele está para ser capaz de atuar e operar com sucesso (BORENSTEIN *et al.*, 1997; KORTENKAMP; BONASSO; MURPHY, 1998). Assim, percebe-se que o problema da localização do robô consiste em responder a pergunta “Onde estou?” do ponto de vista de um robô conforme a Figura 4. Isso significa que o robô tem de descobrir a sua localização no ambiente.

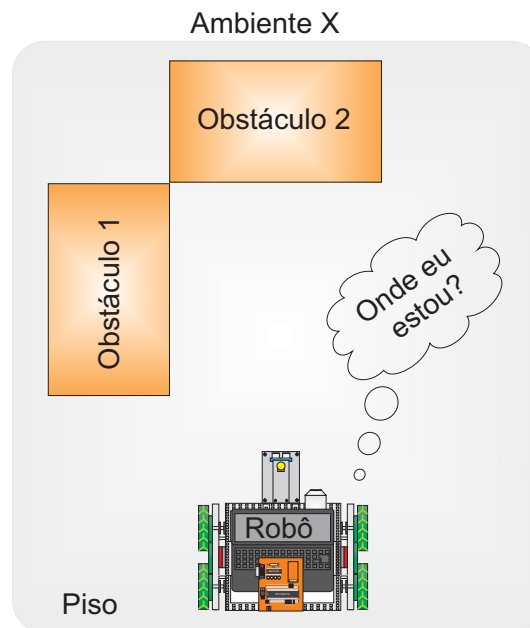


Figura 4: “Onde estou?”

2.3.1 Localização local e global

O nível de dificuldade do problema de localização é definido pelo tipo de informação que está disponível e quanto tempo se tem para executar essa informação. Pode-se dividir o problema de localização em dois tipos: localização local e localização global. A localização local descreve os métodos de localização quando a posição inicial do robô é conhecida. É conhecido como problema local por causa da probabilidade do robô estar próximo da posição dita verdadeira. Resolver esse tipo de problema é focar apenas no ruído produzido pelo movimento do robô, ou seja, compensar os erros do cálculo da posição durante a navegação (SE; LOWE; LITTLE, 2001).

A localização global ocorre quando a posição inicial do robô é desconhecida. Isso significa que um robô poderia ser inicialmente colocado em qualquer lugar em um ambiente e ele não saberia onde está. Esse tipo de problema é considerado mais difícil que a localização local (SE; LOWE; LITTLE, 2001). Se, Lowe e Little (2005) afirmam que uma versão mais difícil do problema de localização global é o problema do “sequestro” do robô. Isso descreve quando um robô é retirado do seu ambiente natural e levado a outro local. A dificuldade por trás desse problema é que o robô pode pensar que ele sabe onde está, mas na verdade não sabe.

2.3.2 Métodos de localização

Sendo a localização de robôs móveis uma área que tem sido bastante estudada ultimamente, há uma grande variedade de métodos de localização. Os métodos mais amplamente utilizados serão descritos a seguir.

Odometria

O cálculo de posição ou *Dead-reckoning* (DR) é o método de localização mais comum e simples. Na maioria dos casos, o DR é conhecido como odometria. Odometria é descrita por Thrun, Burgard e Fox (2005) como um deslocamento do veículo determinado por um odômetro a bordo. O odômetro, neste caso, é geralmente um conjunto de leitores ópticos sobre as rodas. Alternativamente, existem outros sensores baseados em atração magnética, que são habitualmente utilizados na indústria automobilística (NAST *et al.*, 2014).

Para exemplificar esse método, pode-se pensar em um robô seguindo uma certa trajetória como mostrado na Figura 5. No instante t_0 , o robô se encontra na posição P_0 e inicia o seu deslocamento. Quando t_1 , o robô se deslocou D_1 , e sua nova posição P_1 pode ser calculada como a soma entre a posição anterior do robô P_0 e o seu deslocamento D_1 . No instante seguinte t_2 , a posição atual P_2 é novamente calculada com base na sua posição anterior P_1 e o deslocamento realizado D_2 . Assim, a posição atual pode ser calculada como a soma dos deslocamentos efetuados pelo robô com relação à sua posição inicial. Assim, nota-se que a localização por odometria é um método de localização local.

A questão principal com relação à odometria como meio de localização é sua propensão à acumulação de erros. Esses erros aumentam proporcionalmente com a distância percorrida. Como é sabido, a odometria baseia-se no pressuposto de que as revoluções da roda podem ser traduzidas em deslocamento linear em relação ao chão. Esta suposição tem validade limitada. Um exemplo extremo é a patinagem das rodas: se uma roda deslizar sobre, por exemplo, óleo derramado, então o *encoder* associado iria registrar as rotações das rodas mesmo que essas revoluções não correspondessem a um deslocamento linear da roda (BORENSTEIN; FENG, 1996).

Assim, a odometria é uma forma aceitável de localização a curto prazo que pode ser combinadas com medições de posição absoluta para fornecer estimativas de posição mais confiáveis (CHENAVIER; CROWLEY, 1992).

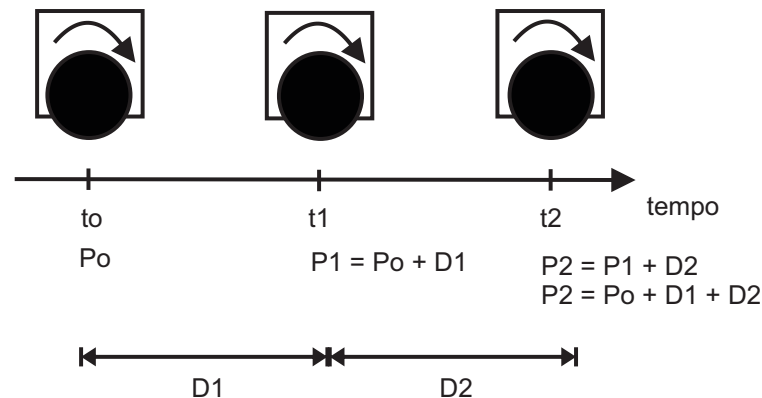


Figura 5: Localização do robô utilizando odometria.

Borenstein e Feng (1996) apresentam uma forma muito simples e eficaz de combinação de medições a partir de um giroscópio com as medidas de *encoders* nas rodas de um robô (odometria). Já Ohno *et al.* (2004) descrevem a navegação de um robô móvel usando GPS diferencial (DGPS) e a odometria em um ambiente aberto. A posição do robô é estimada pela fusão dos dados do DGPS e da odometria.

Bezerra (2004) propôs um sistema híbrido para localização de robôs móveis com rodas em ambientes internos, baseado em odometria e marcos naturais. A odometria foi utilizada para reduzir o tempo de busca dos marcos. Rotineiramente, a odometria é utilizada para preencher as lacunas entre as estimativas de posição com pontos de referência, uma vez que não seria prático ter uma marca visível para cada intervalo de tempo do algoritmo utilizado para localização (BORENSTEIN; FENG, 1996; WANG *et al.*, 2015).

Localização baseada em marcos

O reconhecimento de marcos pode auxiliar na correção de dados obtidos através da odometria. Marcos são características distinguíveis que um robô pode reconhecer em seu ambiente, a partir de dados coletados por seus sensores (BORENSTEIN *et al.*, 1996). As formas geométricas como linhas, retângulos e círculos são marcos artificiais clássicos utilizados para localização de robôs móveis. Normalmente, quando os marcos são escolhidos, eles estão em posições fixas, e as coordenadas são registradas. Conhecendo a posição dos marcos com relação ao referencial do ambiente e com relação ao referencial do robô, pode-se determinar a localização do

robô (BEZERRA, 2004).

Os marcos ou pontos de referência precisam ser facilmente identificáveis, e o grau de dificuldade de identificação depende inteiramente do sensor a ser utilizado. Se for usado um scanner a laser, por exemplo, os marcos precisariam ser geométricos, enquanto que, se for uma câmera, um código de barras pode ser utilizado. Quando o sensor é uma câmera digital, a quantidade de informação a respeito do ambiente faz com que o método necessite de um tempo maior para processamento da imagem.

Jang, Lee e Kweon (2002) utilizaram imagens de quadrados como marcos. Cada quadrado foi dividido horizontalmente em dois retângulos coloridos (vermelho e azul) para auxiliar na detecção. Após detectar o marco, a posição do robô era determinada com base na posição da câmera em relação ao marco.

De maneira semelhante, Fairfield (2002) utilizou anéis verdes de plásticos espalhados pelo ambiente como marcos artificiais. Já Launay e Ohya (2001) utilizaram as lâmpadas fluorescentes localizadas no teto do ambiente como uma espécie de marco natural. Alguns trabalhos utilizam as arestas presentes no ambiente provenientes de encontro de paredes, portas, janelas, teto ou piso como marcos naturais (HOWARD; KITCHEN, 1998; BRANCA; STELLA; DISTANTE, 2000; CLERENTIN *et al.*, 2002; ROUS; LUPSCHEN; KRAISS, 2005; NÚÑEZ *et al.*, 2008).

Localização baseada em mapas

A localização baseada em mapas ou “à base de mapa de posicionamento” é uma técnica em que um robô automaticamente gera um mapa de seu ambiente através de sensores a bordo. O mapa gerado, neste caso, é de área local, este é então comparado com um mapa pré-armazenado. Quando uma correspondência é encontrada, então, o posicionamento do robô pode ser calculado (SIEGWART; NOURBAKHS; SCARAMUZZA, 2011). As etapas de construção de mapa e análise de correspondência em um sistema de navegação de robôs móveis baseado em mapas são destacadas na Figura 6.

A seguir, são listadas algumas vantagens do mapa com base em posicionamento (SIEGWART; NOURBAKHS; SCARAMUZZA, 2011):

- faz uso do ambiente interno sem modificação para estimar a localização;
- rotineiramente, os mapas de um ambiente precisam ser atualizados, pois

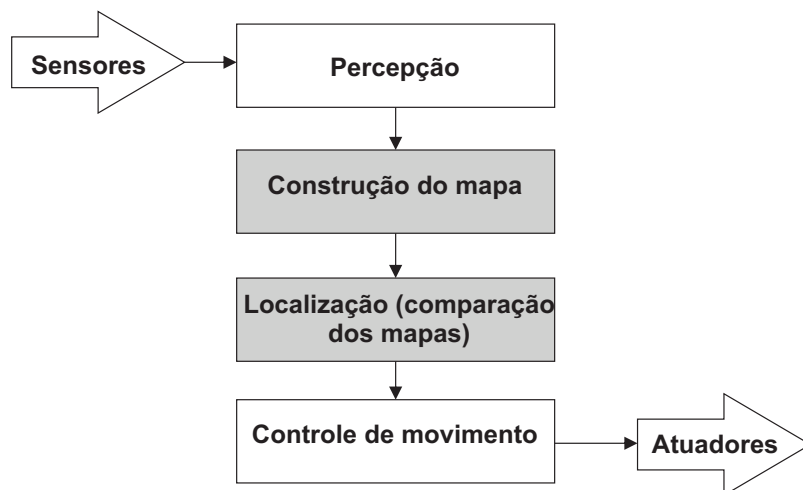


Figura 6: Esquema de um sistema de localização baseado em mapas.

ambientes internos são tipicamente dinâmicos;

- o mapa criado para estimar a posição pode ser utilizado como um recurso para atualizar um mapa global armazenado; e
- mesmo se um ambiente já for mapeado, novas áreas podem se tornar visíveis e, assim, novos mapas podem ser gerados.

Porém, esse tipo de método de localização possui desvantagens (SIEGWART; NOURBAKHS; SCARAMUZZA, 2011):

- precisa de um número satisfatório de pontos de referência fixos e facilmente reconhecíveis para executar uma correspondência entre os mapas;
- o mapa global, que é usado para a correspondência, tem que ser preciso para atender a demanda de tarefas; e
- tem-se a necessidade de sensores de qualidade e poder computacional suficiente para processar dados adquiridos.

A abordagem de localização baseada em mapas exige mais esforço na navegação do robô móvel. Porém, este esforço de desenvolvimento resulta em uma arquitetura que pode mapear com sucesso e navegar em uma variedade de ambientes, amortizando o custo do projeto ao longo do tempo (ORIOLO; ULIVI; VENDITTELLI, 1998). A questão fundamental que diferencia vários sistemas de localização baseados

em mapas é a representação. Esta é a postulação mais direta possível da posição do robô móvel (SIEGWART; NOURBAKHSH; SCARAMUZZA, 2011).

As representações de mapas têm uma longa história. Na década de 1980 e início de 1990, o campo de mapeamento foi amplamente dividido em abordagens métricas e topológicas. Mapas métricos capturam as propriedades geométricas do ambiente, enquanto que mapas topológicos descrevem a conectividade de lugares diferentes (THRUN *et al.*, 2002).

Representantes iniciais dessa abordagem foram Elfes (1987), Elfes (1989) e Moravec (1988) que representaram mapas por grades de granulação fina que modelam o espaço ocupado e livre do meio ambiente. Um algoritmo alternativo de mapeamento métrico foi proposto por Chatila e Laumond (1985) e utiliza conjuntos de poliedros para descrever a geometria de ambientes.

Uma das representações de mapas mais conhecidas dentro da Robótica Móvel é a decomposição topológica. Abordagens topológicas evitam medições diretas da geometria do ambiente, concentrando-se em características do ambiente que são mais relevantes para o robô. Formalmente, uma representação topológica é um grafo que indica duas coisas: nós e a conectividade entre os nós (arcos) (THRUN *et al.*, 2002). Na medida em que uma representação topológica destina-se à utilização de um robô móvel, os nós são usadas para denotar as áreas do ambiente e os arcos são usados para denotar a adjacência de pares de nós. Quando um arco conecta dois nós, o robô pode navegar de um nó para o outro sem a necessidade de uma travessia por qualquer outro nó intermediário. Nota-se, assim, que os nós documentam uma área com base em qualquer discriminante de tal forma que o robô pode reconhecer a entrada e saída do nó.

Uma representação topológica de um conjunto de cômodos em um ambiente interno é descrita na Figura 7. Nota-se que os nós capturam uma visão geométrica do ambiente, e arcos simplesmente representam as conectividades. Um exemplo de representação topológica é o trabalho de Simhon e Dudek (1998), no qual descrevem um passo importante para a criação de mapas de grande escala que combinam, ambos, métrica e conhecimento topológico. Alguns outros autores também realizam essa integração de topologia e de métrica para localização de robôs: com uma proposta estatística (THRUN *et al.*, 1998), com *SLAM* (*Simultaneous Localization and Mapping*) (TOMATIS; NOURBAKHSH; SIEGWART, 2001, 2003), com mapas compactos (KO; SONG; LEE, 2004), com algoritmos de otimização

(KONOLIGE; MARDER-EPPSTEIN; MARTHI, 2011). Outras abordagens utilizando mapas topológicos são apresentadas com base em um modelo distribuído do ambiente (MATARIC, 1990), em uma representação espacial (KUIPERS; BYUN, 1991), em correção de erro no aprendizado de mapas (ENGELSON; MCDERMOTT, 1992), em gráfico de Voronoi generalizado (CHOSSET; BURDICK, 2000), em Visão Computacional (SE; LOWE; LITTLE, 2005), em imagens panorâmicas (KONOLIGE; MARDER-EPPSTEIN; MARTHI, 2011) e em robôs domésticos (GERSTMAYR-HILLEN *et al.*, 2013).

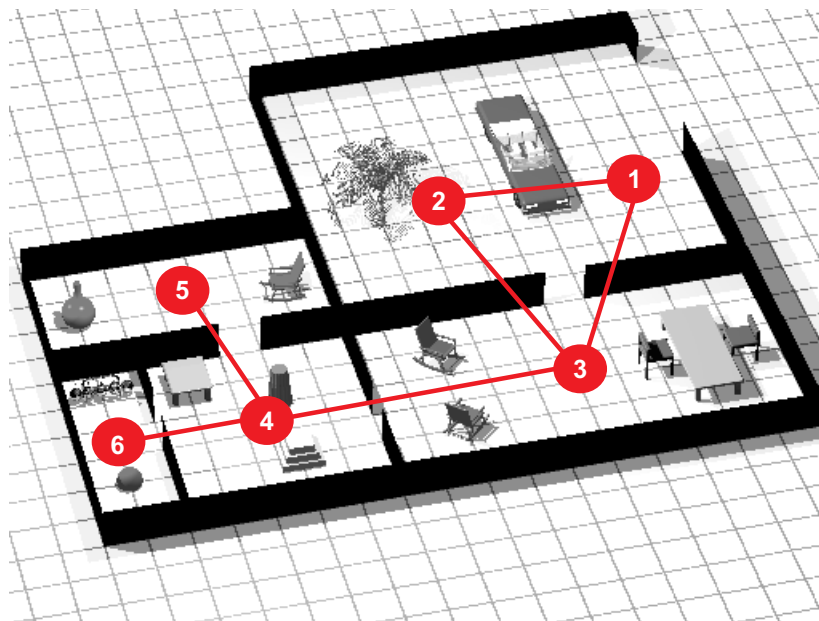


Figura 7: Representação de um mapa topológico em um ambiente interno.

A fim de navegar usando um mapa topológico de forma robusta, deve existir um meio para detectar a sua posição atual em termos dos nós do gráfico topológico (método de localização). As dimensões das conexões entre os nós devem ser otimizadas para coincidir com a discriminação sensorial do *hardware* do robô móvel (THRUN *et al.*, 2002).

Porém, o mapeamento topológico por si só não fornece qualquer indicação das chances relativas (probabilidades) entre as várias posições possíveis do robô (SIEGWART; NOURBAKHSH; SCARAMUZZA, 2011). Técnicas probabilísticas diferem do mapeamento topológico porque estas identificam explicitamente as probabilidades das possíveis posições do robô, e por este motivo estes métodos tem recentemente crescido em pesquisa.

Localização de Markov

O teorema de *Bayes* é a base dos algoritmos de localização baseados em probabilidades (DURRANT-WHYTE; BAILEY, 2006). Como visto pela primeira vez no trabalho de Fox, Burgard e Thrun (1998), a localização de Markov (LM) refere-se à aplicação direta do teorema de Bayes para localização. Com a LM, é possível a realização de estimação do estado a partir de dados do sensor. A LM é descrita como um algoritmo probabilístico, ou seja, para que se tenha uma única hipótese quanto à localização do robô no ambiente. Na LM, existe uma distribuição de probabilidade para cada localização (BURGARD *et al.*, 1996; ROY *et al.*, 1999).

A aplicação direta da localização de Markov é possível quando a representação do ambiente do robô já fornece uma decomposição adequada. Este é o caso quando a representação do ambiente é puramente topológica. Powers e Birchfield (1995) descrevem essa situação empregando a localização probabilística de Markov e a crença de várias hipóteses sobre a representação do ambiente topológico.

Kosecká e Li (2004) estudam um modelo topológico do ambiente dentro de uma casa por meio de detecção visual, onde cada localização no modelo é representado por um conjunto de pontos de vista e seus descritores representativos e mostram como modelar as relações espaciais entre locais individuais por um modelo de Markov oculto.

Outros trabalhos incluem o modelo probabilístico da localização de Markov em robôs colaborativos (FOX *et al.*, 1999b), em robôs com mapas imprecisos (TOMONO; YUTA, 2001), na localização de clientes em uma rede LAN (WALLBAUM; WASCH, 2004), na busca do CEP no bloco do endereço em envelopes postais (BORGES, 2005), em robôs em ambientes externos (MIURA; YAMAMOTO, 2008) e na localização de robôs autônomos em ambiente industrial (MOREIRA, 2014).

Localização de Monte Carlo

Após a localização de Markov, Fox *et al.* (1999a) propõem uma outra abordagem chamada localização de Monte Carlo (LMC). Assim com a LM, o algoritmo de LMC se aplica tanto ao nível global, quanto aos problemas locais de localização. Ele funciona basicamente através da utilização de filtros de partículas para estimar a localização do robô.

O método de filtro de partículas aproxima a sequência de distribuições de probabilidade de interesse ao longo do tempo usando um grande conjunto de amostras aleatórias, nomeadas de partículas. Estas partículas tem um peso associado que descreve a qualidade da partícula específica (GORDON; SALMOND; SMITH, 1993). Em contraste com os filtros de Kalman (KALMAN, 1960) amplamente difundidos que utilizam distribuições gaussianas, os filtros de partículas podem aproximar uma grande variedade de distribuições de probabilidade.

Essa abordagem é muito popular por ser relativamente fácil de implementar e adaptar-se à disposição dos recursos computacionais e a necessidade de localização, uma vez que o número de partículas pode ser dinamicamente ajustado para cada situação. Filtros de partículas tem sido muito usados para resolver problemas de localização de robôs móveis (DELLAERT *et al.*, 1999; THRUN *et al.*, 2001; BANSAL, 2002; WOLF; BURGARD; BURKHARDT, 2005; ZHANG; ZAPATA; LÉPINAY, 2012; PÉREZ; CABALLERO; MERINO, 2015).

Localização e Mapeamento simultâneo

Localização e Mapeamento simultâneo (*Simultaneous Localization and Mapping*, SLAM) é essencial para navegação de robôs móveis, principalmente para planejamento de uma trajetória autônoma. SLAM pode ser entendido como um processo para construção de um mapa de um ambiente desconhecido e, ao mesmo tempo, estimar da posição atual do robô. Ambas as etapas dependem uma da outra, ou seja, um bom mapa é necessário para calcular a posição do robô e a estimativa da posição precisa fornecer a possibilidade de construção de um mapa correto como mostrado na Figura 8 (RÜCKERT, 2009).

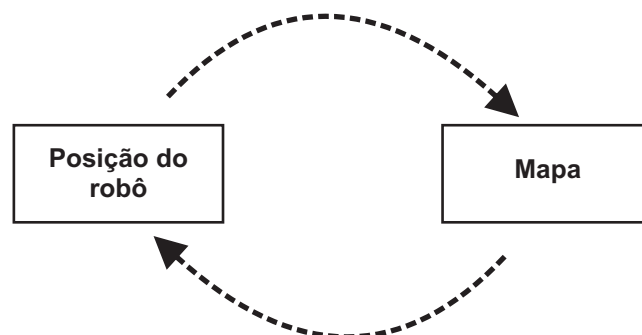


Figura 8: Interação entre o mapa e a posição do robô.

Como já visto, é necessária uma estimativa simultânea de ambos os locais do robô.

Para isso, o veículo deve ser equipado com um sistema sensorial capaz de tomar as medidas da localização relativa entre os marcos de maneira autônoma (THRUN; LEONARD, 2008). Vários grupos de pesquisa e pesquisadores tem trabalhado com SLAM e os sensores mais comumente usados podem ser categorizados em laser, sonar e sistemas de visão. Adicionalmente, algumas fontes sensoriais são usadas para perceber melhor o ambiente, tais como bússolas e Sistema de Posicionamento Global (GPS) (THRUN *et al.*, 2002).

Durrant-Whyte (1988) afirma que a “solução” do problema SLAM tem sido um dos sucessos notáveis da comunidade que estuda Robótica. A base estatística do problema de localização e mapeamento simultâneo foi descrito pela primeira vez por Smith e Cheeseman (1986) e Durrant-Whyte (1988). Estes trabalhos fornecem técnicas de estimativa de probabilidades para correlações entre as características e as posições robô.

Ao mesmo tempo, Ayache e Faugeras (1988) iniciaram o estudo de SLAM com navegação visual. Crowley (1989) e Leonard e Durrant-Whyte (1991) propõem métodos de SLAM utilizando segmentos de linha extraídos a partir de dados ultrassônicos. Outros pesquisadores como Vandorpe, Brussel e Xu (1996) e Gonzalez, Ollero e Reina (1994) usam dados de laser, por sua vez.

Choset e Nagatani (2001) apresentaram um novo método para SLAM que explora a topologia de espaço livre do robô para localizar o robô em um mapa parcialmente construído. A topologia do ambiente é codificada em um mapa topológico. Em tempo, o mapa topológico particular utilizado nesse trabalho é o gráfico de Voronoi generalizado (GVG), que também codifica algumas informações métrica sobre o ambiente do robô.

Montemerlo *et al.* () descrevem algoritmo eficiente de SLAM chamado *FastSLAM*. O algoritmo *FastSLAM* decompõe o problema SLAM em um problema de localização do robô e em uma coleção de problemas de estimação de posição. Os resultados experimentais, utilizando um robô físico e um robô simulado, ilustram que o algoritmo pode lidar com mais pontos de referência em menos tempo. Alguns outros trabalhos apresentaram algoritmos que seguem a fundamentação do *FastSLAM* (MONTEMERLO; THRUN, 2003; HAHNEL *et al.*, 2003; MONTEMERLO; THRUN, 2007; ÖZKUCUR; AKIN, 2010), inclusive utilizando, como sistema sensorial, as câmeras omnidirecionais (WONGPHATI; NIPARNAN; SUDSANG, 2009; GUTIERREZ *et al.*, 2011).

Outra representação do SLAM é a adotada por Bailey *et al.* (2006) denominada de EKF-SLAM (*Extended Kalman Filter formulation of Simultaneous Localization and Mapping*). Esse algoritmo é considerado um algoritmo de SLAM não-linear e é predominantemente implementado como um Filtro de Kalman Estendido (*Extended Kalman filter*, EKF).

Os desafios cruciais para o SLAM estão maiores e mais persuasivos. Enquanto o progresso tem sido substancial, a escala e a estrutura de muitos ambientes são limitadas. Tais métodos estão abrindo novos caminhos e fazendo ligações de volta aos princípios fundamentais na percepção e localização do robô, criando assim a necessidade de se estudar outras perspectivas.

Localização com técnicas de Reconhecimento de Padrões

Dentro da comunidade Robótica, a localização é considerada um tema aberto à investigação e em constante evolução. O foco de pesquisa tem se centrado no desenvolvimento de abordagens de baixo custo, aquelas que não exigem o cálculo pesado ou sensores sofisticados. De maneira ainda acanhada, percebe-se que técnicas de reconhecimento de padrões, incluindo Redes Neurais Artificiais (RNAs) tem sido aplicada à área com sucesso variado.

O reconhecimento de padrões é uma capacidade extremamente desenvolvida nos seres humanos e em alguns animais. O ser humano é hábil em reconhecer rostos, vozes e, até mesmo, o humor de pessoas conhecidas. O grau de requinte do reconhecimento de padrões, por parte do ser humano, pode chegar a ponto de diferenciar uma pintura de um grande mestre daquela feita por um exímio falsário ou, ainda mais, pode estabelecer uma tomada de decisão por parte de um operador em um dia de grande movimento em uma bolsa de valores (de Castro; do Prado, 2001).

Algumas áreas que podem ser destacadas na aplicação do reconhecimento de padrões são: identificação através de impressões digitais e análise da íris (KAWAGOE; TOJO, 1984), diagnósticos médicos (STEINER, 1995), visão computacional (PERELMUTER *et al.*, 1995), diagnósticos pré e pós-natal e certos diagnósticos de câncer (AGUIAR, 2000), reconhecimento de voz (FERREIRA, 1994), processamento de imagens (RUIDIAS; DE, 2005), análise de peças para manutenção preventiva (LIMA *et al.*, 2010), identificação e controle de sistemas dinâmicos (BARRETO; ARAUJO, 2004), reconhecimento de caracteres (ALEXANDRIA *et al.*,

2014), análise de caracteres manuscritos (PATEL; PATEL; PATEL, 2012) , análise de eletrocardiogramas (GONZÁLEZ *et al.*, 2014), reconhecimento de patologias na coluna vertebral (NETO AJALMAR, 2011), reconhecimento e identificação de cromossomos (RITTER; GALLEGOS, 1997), dentre outras.

Com o avanço e a disponibilidade de recursos computacionais, tem popularizado a utilização de elaborados métodos de classificação de padrões em várias áreas do conhecimento. Contudo, a utilização dessas técnicas na localização global de robôs móveis é um assunto relativamente novo. Assim, ao se observar uma área com uma indisponibilidade de resultados, cria-se a oportunidade de uma abordagem baseada em Reconhecimento de Padrões para localização de um robô móvel.

2.4 Conclusão do Capítulo

Ciente a importância da percepção e da localização para robôs móveis, realizou-se uma revisão dos principais métodos de percepção e localização. No próximo Capítulo, uma revisão sobre imagens omnidirecionais é realizada. Os métodos de reconhecimento de padrões e os métodos de como o ambiente é representado no modelo computacional (extração de características) que são utilizados neste trabalho são abordados no Capítulo 4.

IMAGENS OMNIDIRECIONAIS

Neste Capítulo é apresentada base teórica para o desenvolvimento desta dissertação. Os sistemas de visão omnidirecional são apresentados, juntamente com uma breve explicação sobre suas vantagens para servir como embasamento e justificativa para a metodologia apresentada no Capítulo 5.

3.1 Visão omnidirecional

Os sistemas de visão omnidirecional conseguem representar todo o entorno de uma localidade em uma única imagem, apresentando um amplo campo visual. Imagens omnidirecionais representam a cena com um ângulo de visão de 360° ao redor do sensor, expandindo assim, a capacidade dos sensores convencionais de gerar apenas imagens restritas a uma pequena região. O resultado disso é o aumento na quantidade de informações disponíveis na imagem (ISHIGURO; YAMAMOTO; TSUJI, 1992).

3.1.1 Visão omnidirecional e sua inspiração em animais

Inúmeras vezes, o progresso da ciência foi inspirado pela observação da natureza. Um dos primeiros exemplos que inspirou o desenvolvimento de sensores de visão omnidirecional apareceu no planeta Terra há cerca de 50 milhões de anos. O olho de um fóssil de tribolite, uma das primeiras inspirações para sistemas de visão omnidirecional, este que é semelhante a dos insetos atuais, é mostrado na Figura 9 (BENOSMAN; KANG, 2001; SRINIVASAN; VENKATESH, 1997).

Quase todos os olhos dos animais que podem ser considerados primórdios de visão omnidirecional são olhos compostos. O olho composto é um conjunto de elementos ópticos aptos a recolher os raios de luz vindos de diferentes direções. Estes elementos ópticos estão densamente compactados para formar o olho e são



Figura 9: O olho de um trilobite fóssil (SRINIVASAN; VENKATESH, 1997).

chamados facetas (LAND; FERNALD, 1992).

Pode-se dividir os animais de olhos compostos em três principais categorias. Esta divisão depende da técnica utilizada para focar a luz na retina. O primeiro conjunto é composto de olhos, onde a cada faceta está associada a uma célula fotorreceptora, como esboçado na Figura 10. Este tipo de olho é frequente em insetos diurnos (BENOSMAN; KANG, 2001).

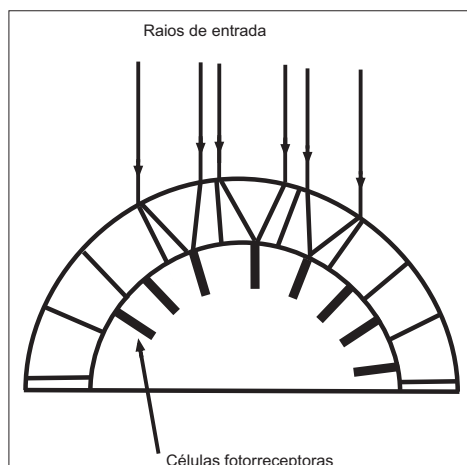


Figura 10: Um esboço do olho composto de um inseto diurno. A cada faceta tem um fotorreceptor associado (BENOSMAN; KANG, 2001).

O segundo grupo é composto de olhos com uma única retina. O índice de refração de cada faceta é formado a fim de dobrar progressivamente o caminho dos raios de luz que chegam. A dobração faz com que os raios paralelos provenientes da mesma fonte se concentrem no mesmo ponto na retina, conforme mostrado na Figura 11. Tal tipo de olho está apto a trabalhar em condições de baixa iluminação. Este olho composto é encontrado em insetos noturnos (BENOSMAN; KANG, 2001).

O terceiro grupo é composto de olhos semelhantes ao dos insetos noturnos, mas

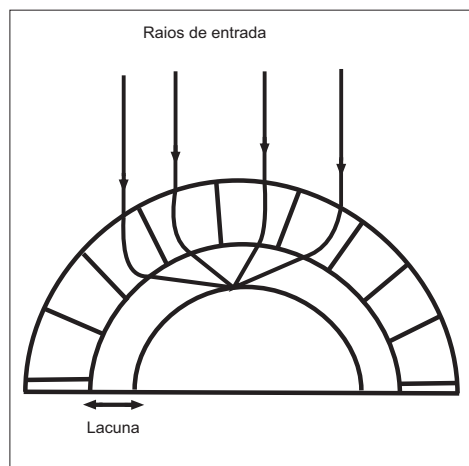


Figura 11: Um esboço do olho composto de um inseto noturno. Os raios de luz que entram em diferentes facetas estão focados no mesmo ponto na retina com uma refração (BENOSMAN; KANG, 2001).

este grupo usa reflexões em vez de refrações, a fim de concentrar a luz sobre a retina. Cada faceta está associado a um elemento óptico com uma superfície lateral refletora. Uma vez que os raios de luz paralelos provenientes do mesmo objeto entram nas diferentes facetas e são refletidos na superfície lateral dos elementos ópticos e focalizado na retina (BENOSMAN; KANG, 2001). Um esboço deste tipo de olho composto é mostrado na Figura 12.

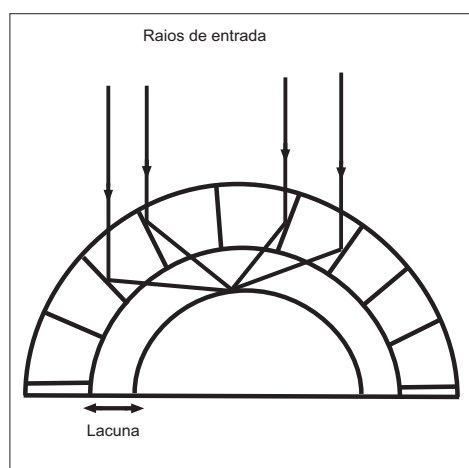


Figura 12: Um esboço do olho composto de um crustáceo. Observe que os raios de luz de entrada em diferentes facetas estão focados no mesmo ponto na retina com uma reflexão (BENOSMAN; KANG, 2001).

Este último tipo de olho pode ser encontrada em crustáceos. Um caso particularmente interessante é crustáceo de mar profundo *Gigantocypris*, que tem grandes olhos com refletores que são semelhantes aos das lentes dos telescópios

como mostrado na Figura 13. Essa característica permite que ele veja em condições muito adversas em profundidades da ordem dos milhares de metros (BENOSMAN; KANG, 2001).



Figura 13: Crustáceo *Gigantocypris* (APPELTANS *et al.*, 2012).

3.1.2 Visão Omnidirecional nas artes

No início do século XV, depois da intuição brilhante de Giotto em meados do século XIV, Filippo Brunelleschi não se entregou às regras geométricas subjacentes à percepção visual e criou a *perspectiva*. A perspectiva permitiu ao artista dar uma representação mais fidedigna da realidade, seguindo algumas regras geométricas simples. Ao mesmo tempo, os artistas descobriram que podiam brincar com essas regras geométricas para alterar a percepção da realidade. Um dos primeiros exemplos é a representação em pinturas de espelhos hemisféricos chamados espelhos de bruxa conforme a Figura 14.

Na pintura "*O casamento de Giovanni Arnolfini*" de Jan van Eyck, o espelho de bruxa representa um artifício cênico para estender a percepção do espectador. Apresentando o espelho bruxa, ele pode expandir o horizonte sensorial por trás da fronteira do habitual e representar todo um compartimento (DRUCKER, 2013).

O significado filosófico da extensão da realidade e do conhecimento é atingido também na famosa impressão de Escher, *Hand with Reflecting Sphere* (1935), conforme a Figura 15 (ESCHER, 1935). O pintor é observado pelo espectador por meio de uma reflexão sobre a esfera e o observador tem uma visão completa do mundo do pintor.

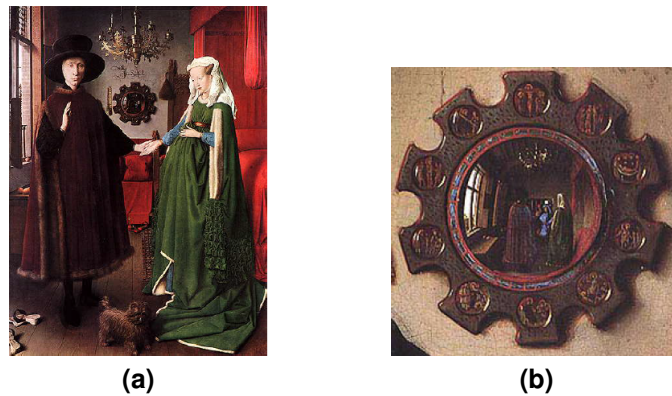


Figura 14: Representação em pinturas: (a) "*O casamento de Giovanni Arnolfini*" de Jan van Eyck e (b) o zoom no espelho bruxa na parede atrás do casal de noivos mostrando o resto da sala e do pintor (DRUCKER, 2013).

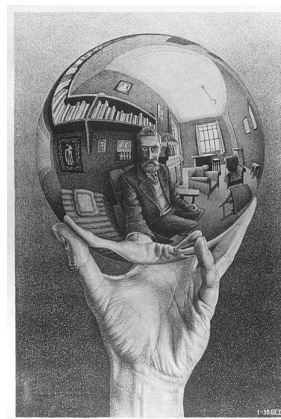


Figura 15: Uma litografia de Escher: *Hand with Reflecting Sphere* (ESCHER, 1935).

Outro exemplo de como a superfície refletora pode ampliar o sensorial humanos é o caso da anamorfose. Um exemplo simples de anamorfose está descrita na Figura 16. A pessoa não está representada seguindo a regra geométrica da perspectiva, mas geometricamente distorcida a um grau em que não é reconhecível. A distorção geométrica pode ser removida apenas com uma superfície refletora com uma forma e posição apropriada. A imagem refletida pela superfície do espelho aparece agora seguindo as regras da perspectiva (BENOSMAN; KANG, 2001).

Benosman e Kang (2001) disse que a anamorfose é fascinante, não só porque revela o que está oculto, mas também por que oferece duas maneiras contraditórias de perceber a realidade. Na verdade, o olho percebe a sua versão distorcida e versão corrigida no espelho, e compreende, simultaneamente, tanto a ilusão e o mecanismo da ilusão.

A distorção da perspectiva, os espelhos de bruxas e a anamorfose, contribuíram para criar o entendimento de que o que ser humano percebe através dos olhos



Figura 16: Um exemplo de anamorfose por Awtar Singh Viridi, século XVI.

é apenas uma das possíveis representações da realidade. Pode-se mudar a maneira como a realidade é percebida e, assim, expandir o conhecimento extraindo mais informações a partir do aparecimento de outra realidade como será mostrado posteriormente.

3.1.3 A arte panorâmica

Já viu-se como os artistas brincam com a percepção a fim de ampliar o horizonte sensorial do observador. Agora, tem-se a necessidade de relatar os esforços realizados pelos artistas para capturar uma visão completa em 360° de uma cena, de modo a reproduzir o panorama do ambiente.

O panorama foi introduzido no final do século XVIII, e é resultado da concatenação de duas palavras gregas: *pan* (= tudo) e *horama* (=visão). A maioria das pinturas realizadas em panoramas foram perdidas. Poucos delas restaram. O amplo panorama da Filadélfia em 1987 é reproduzido na Figura 17.



Figura 17: Panorama da Filadélfia em 1987.

As imagens panorâmicas foram construídas em todas as grandes cidades da Europa e dos Estados Unidos da América. Foi um grande sucesso comercial, mas no final do século XIX, essa arte declinou rapidamente. Neste mesmo século, a fotografia foi inventada e, assim, a fotografia panorâmica. As câmeras usadas para capturar

essas primeiras fotos panorâmicas são as ancestrais das modernas câmeras de visão panorâmica. Uma das câmeras panorâmicas de maior sucesso, a Kodak Cirkut 10, é mostrada na Figura 18.

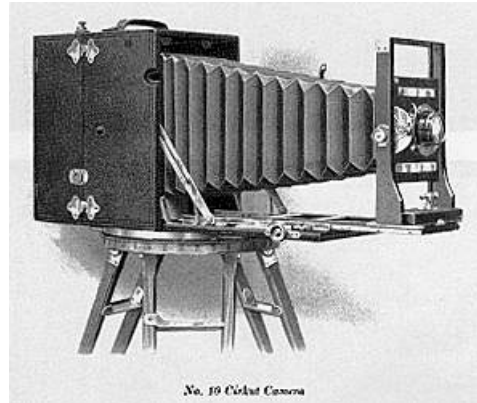


Figura 18: Um dos primeiros modelos de câmera panorâmica.

3.1.4 Visão Omnidirecional na Ciência

A flexibilidade da arte panorâmica inspirou a comunidade científica voltada para visão de máquina a explorar uma nova maneira de capturar imagens da cena. A ideia original do sensor de visão omnidirecional foi proposta por Rees (1970). Na patente, Rees (1970) propõe a utilização de um espelho hiperboloidal para capturar a imagem omnidirecional que pode ser transformado para imagens normais em perspectiva.

Cao, Oh e Hall (1985) deram um passo importante na concepção de um robô móvel, descrevendo um novo algoritmo para a navegação por visão omnidireccional. Eles montaram um protótipo de um sistema de visão omnidirecional e implementaram técnicas de navegação que utilizam este sensor.

Yagi e Kawato (1990) construíram um sensor de visão omnidirecional de projeção cônica (*conic projection image sensor*, COPIS). COPIS foi proposto como um novo sistema de posicionamento global, que é capaz de obter informações por visão omnidirecional de um quadro de uma imagem panorâmica. Os autores relatam a configuração do COPIS e sua aplicação para gerar um mapa do meio ambiente de um corredor interno a partir de um robô móvel. O método utiliza as propriedades comuns a muitos ambientes internos ao invés do modelo geométrico do mundo.

Um ano depois, a construção de um sensor de visão omnidirecional foi proposto de forma diferente. Hong *et al.* (1992) utilizaram um espelho esférico em seu trabalho e relataram sobre um sistema que permite a um robô utilizar um modelo do seu ambiente

para navegar. Eles mostraram que a navegação baseada na imagem omnidirecional é uma alternativa viável para navegação de robôs móveis.

Yamazawa, Yagi e Yachida (1993) conseguiram obter imagens com ângulo de 360° graus através de outro tipo de espelho. Eles descrevem um sensor de imagem com um espelho hiperboloidal para navegação baseada na visão de um robô móvel. Seu nome é *HyperOmni Vision*. Este sistema de detecção pode adquirir uma vista omnidirecional em torno do robô, em tempo real, com o uso de um espelho hiperboloidal.

Nayar e Baker (1997) desenvolveram um sensor de visão omnidirecional usando um espelho parabólico e uma lente telecêntrica. Eles relataram que as câmeras de vídeo convencionais possuíam campos de visão limitado, que as tornavam restritiva em uma variedade de aplicações de visão computacional. Assim, uma câmera com um campo hemisférico de vista é apresentado. Os resultados são apresentados na geração de imagens em perspectiva por *software* a partir de uma imagem omnidirecional. O artigo termina com uma discussão sobre a resolução espacial da câmera proposta. O sistema de imagem desenvolvido por Nayar e Baker (1997) é ideal, mas é difícil a construção em menor escala, visto que o sistema usa uma lente telecêntrica.

Ishiguro, Yamamoto e Tsuji (1992) desenvolveram sensores omnidirecionais compactos usando espelhos esféricos e hiperbólicos de baixo custo.

As primeiras aplicações reais com esses tipos de sensores apareceram com a navegação baseada em mapas (YAGI; NISHIZAWA; YACHIDA, 1992), projeção hiperboloidal (YAMAZAWA; YAGI; YACHIDA, 1993), a detecção de obstáculos (YAMAZAWA; YAGI; YACHIDA, 1995), o uso de sensores de visão omnidirecional de baixo custo (ISHIGURO, 1998), o uso de visão omnidirecional estéreo (GLUCKMAN; NAYAR, 1998), soluções com representação de ambiente (GASPAR; NIALI; SANTOS-VICTOR, 2000) e com o sistema SYCLOP (BRASSART *et al.*, 2000).

Os sensores de visão omnidirecional ganham a cena, principalmente em aplicações robóticas como a navegação com reconstrução de cena (BUNSCHOTEN; KROSE, 2003), a navegação em ambientes dinâmicos (MENEGATTI *et al.*, 2006), a navegação topológica (GOEDEME *et al.*, 2007), a navegação com estimação de movimento (SCARAMUZZA, 2008), o futebol de robôs (LU *et al.*, 2011) e a navegação *indoor* (YIP *et al.*, 2014).

Os modernos sensores de visão omnidirecional utilizados em visão computacional

se distinguem quanto à técnica de obtenção da imagem e podem ser divididos em quatro grupos:

1. sistemas com múltiplas câmeras;
2. câmeras giratórias ou panorâmicas;
3. lentes especiais;
4. sistemas catadióptricos.

Nesse ponto, é necessária uma breve nota sobre a terminologia. Quanto a isso, a comunidade científica costuma se manifestar com muito rigor. Os adjetivos panorâmica e omnidirecional são utilizados como sinônimos, sendo o termo omnidirecional ligeiramente mais técnico. Nesta dissertação, utilizar-se-á o termo “panorâmica” para se referir a um dispositivo capaz de gerar uma imagem em perspectiva 360° com um campo vertical de visão limitado. O termo “omnidirecional” será usado para referir-se a sensores que produzem imagens “não-perspectiva” com um campo de visão horizontal de 360° e com um campo de visão vertical que pode ser limitado ou não. (BENOSMAN; KANG, 2001)

Sistemas com múltiplas câmeras

Sistemas de múltiplas câmeras utilizam mais de uma câmera para a captura de imagens e, então, gerar a imagem panorâmica do ambiente, assim como mostrado na Figura 19. Os eixos ópticos pertencentes a cada câmera são dispostos em diferentes direções, formando assim um aglomerado de imagens que quando processadas geram a imagem panorâmica (KINOSHITA *et al.*, 2005). No entanto, além de ser uma solução computacionalmente cara, torna-se economicamente inviável dependendo dos requisitos do projeto em questão (KRISHNASAMY *et al.*, 2004). Yasushi (1999a) também afirma que é necessário realizar a calibração de todas as câmeras envolvidas no sistema de captura.

Um das câmeras nesse viés é a a DodecaTM 1000 Camera que é composta por 11 câmeras capazes de capturar imagem quase esférica. DodecaTM 1000 Camera é produzida pela companhia *Immersive Mídia* e é mostrada na Figura 20(a) (MENEGATTI, 2002). A partir de 2009, uma versão mais recente chamada de Dodeca 2360 foi lançada pela mesma empresa.

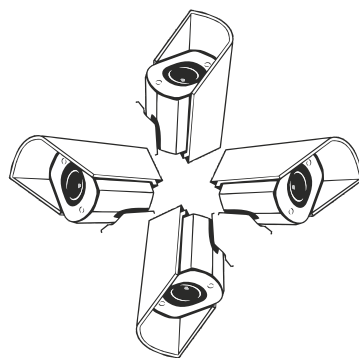


Figura 19: Sistema omnidirecional com múltiplas câmeras.

O sistema omnidirecional estéreo (*Stereo Omnidirectional System*, S.O.S.) tem quase o mesmo campo de visão esférico da câmera DodecaTM 1000, além de contemplar visão estéreo. Este pode ser visto na Figura 20(b). Isto é possível porque o sistema é composto por 20 trios de câmeras CMOS olhando para direções diferentes. O S.O.S. sensor foi idealizado por *ViewPlus Inc.* Shimizu *et al.* (2004) propuseram um novo método para resolver o problema da detecção de objetos em movimento para robôs móveis usando o S.O.S.

Apesar destes dois protótipos, os sistemas de múltiplas câmeras são mais comumente compostos de algumas câmeras olhando para o horizonte em direções diferentes que disparam ao mesmo tempo, como a RingCam da Microsoft, representada na Figura 20(c). As imagens das diferentes câmeras são fundidas para produzir uma imagem panorâmica, também chamada de cilindro panorâmico (NANDA; CUTLER, 2001).

Apresenta-se um quarto tipo de sensor que geralmente não é considerado um sensor omnidireccional, mas pode ser considerado um peculiar tipo de câmera de olho composto. Este sensor não está olhando ao redor, como os outros sensores apresentados, ele olha para dentro, conforme Figura 20(d). Esse sistema não gera uma imagem omnidirecional única. Ele executa visão N-estéreo e é usado para digitalizar performances de movimento como dança ou eventos esportivos (KANADE, 1999).

O quinto sensor apresentado é o sistema esférico *Ladybug2* que possui 0.8 *Megapixels* de resolução e que permite que a imagem capture mais do que 80% do total de uma esfera, veja a Figura 20(e). Este é utilizado em projetos de navegação autônoma de robôs desenvolvidos pelo Centro de Tecnologia da Informação Renato

Archer¹.

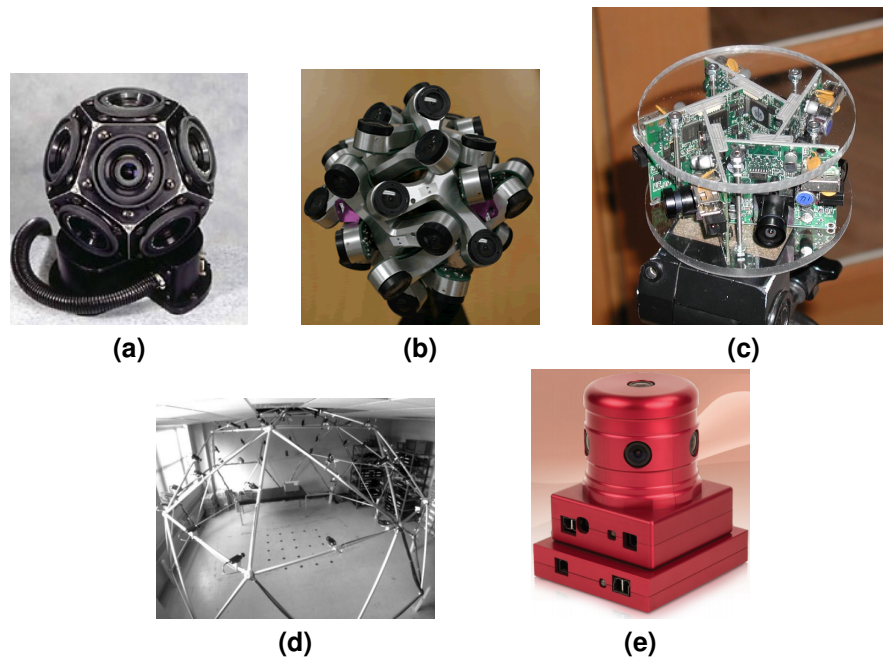


Figura 20: Sistemas com múltiplas câmeras: (a) DodecaTM 1000 (MENEGATTI, 2002), (b) S.O.S. (Stereo Omnidirectional System)(SHIMIZU *et al.*, 2004), (c) RingCam (NANDA; CUTLER, 2001), (d) Visão N-estéreo (KANADE, 1999) (e) *Ladybug2 spherical digital video camera*.

Câmeras giratórias ou panorâmicas

Trata-se de sistema com câmera giratória que consiste em fixar a câmera em um eixo giratório por um determinado período de tempo, conforme a Figura 21 (CHEN, 1995). Ishiguro, Yamamoto e Tsuji (1990) construíram uma única câmera para girar sobre o eixo vertical, capturando consecutivas imagens e as organizando em uma representação panorâmica, obtendo informações ricas de todo ponto de observação. Browitt (1997) afirma que a câmera gira em torno de um eixo vertical a uma certa distância do seu ponto focal, tem-se visão omnidirecional estéreo baseado no trabalho de Ishiguro, Yamamoto e Tsuji (1992).

Krishnan e Ahuja (1996) afirmam que nesse tipo de aplicação, as imagens são então combinadas de forma a gerar uma única imagem, que possui uma abrangência de 360° do ambiente em questão. Esta técnica, no entanto, além de ser computacionalmente cara, compromete o desempenho de um sistema que precisa capturar informações do ambiente em tempo real e de ambientes dinâmicos, como é

¹ <http://www.cti.gov.br/>

o caso de aplicações em navegação para robôs autônomos, pois a rotação da câmera insere uma carga de tempo extra ao sistema (YASUSHI, 1999a).

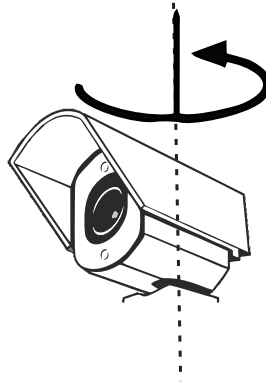


Figura 21: Sistema omnidirecional com câmera giratória

Lentes especiais

Uma opção é usar lentes especiais que melhoram a capacidade de captura das câmeras (CAO; RONG; HALL, 1987; KRISHNAN; NAYAR, 2008; JIA; LI, 2011; COURBON *et al.*, 2007). Pode-se também utilizar lentes com perfis diferentes. Uma lente bastante utilizada para o propósito de se aumentar o campo de visão de uma câmera é a lente "olho de peixe". Câmeras com o uso de lente olho de peixe podem adquirir quase uma visão hemisférica completa (SHAH; AGGARWAL, 1994, 1996).

A desvantagem é que a resolução das imagens é boa no centro mas reduzida na periferia. Isso não é bom para navegação de robôs, pois objetos que representam a cena podem aparecer tanto no centro quanto na periferia. Uma imagem obtida através dessa lente pode ser observada na Figura 22. O alto custo de comercialização desta lente pode tornar o seu uso inviável também.

Sistemas catadióptricos

Mesmo que câmeras omnidirecionais tenham sido obtidas com diferentes técnicas, a ideia básica é sempre a mesma. Câmeras que usam apenas o efeito de refração das lentes para curvar a luz são chamados de câmeras dióptricas. Dioptria é uma unidade de medida que afere o poder de vergência, ou refração, de um sistema óptico (COLENBRANDER, 2009). Câmeras que usam os efeitos combinados de reflexão a partir de um espelho e de refração de uma lente são chamadas catadióptricas. Um



Figura 22: Imagem obtida através de câmera com lente “olho de peixe”.

sistema catadióptrico é todo sistema óptico em que se combinam os efeitos da luz refletida e da refratada, isto é, que compreende lentes e espelho (COLENBRANDER, 2009).

3.2 Sistemas Catadióptricos

Sistemas com o uso de espelhos convexos também são chamados de sistemas catadióptricos. A ideia é usar uma câmera apontada verticalmente para um espelho convexo com o eixo óptico da câmera alinhado com o do espelho. Podem ser utilizadas várias formas de espelhos, tais como esféricos, cônicos, parabólicos e hiperbólicos.

Como mostrado na Figura 23, um sistema catadióptrico utiliza uma superfície refletora para aumentar o campo de visão. O espelho retrovisor de um carro é usado exatamente no desta forma. No entanto, a forma, posição, e orientação da superfície refletora estão relacionadas ao campo de visão de uma forma complexa. Exemplos de sensores catadióptricos podem ser encontrado em diferentes formas como: projeções cônicas (YAGI; KAWATO, 1990), projeções hiperboloidais (YAMAZAWA; YAGI; YACHIDA, 1995) e projeções esféricas (NAYAR; BAKER, 1997).

Quanto à estrutura utilizada para obter um sistema omnidirecional, uma outra propriedade que caracteriza esse tipo de sistema é a propriedade do centro único de projeção (YASUSHI, 1999a). O centro único de projeção (*single view point*) é uma característica importante em visão omnidirecional dependendo do tipo de aplicação.

Dentre os sistemas catadióptricos, se faz uma distinção entre sistemas com um único ponto de projeção (*Single View Point*, SVP) e com pontos de projeção não-individuais (*Non-Single View Point*, NON-SVP). Baker e Nayar (1999) afirmaram que sistemas SVP utilizam formas de espelhos que assegurem que o sistema tenha

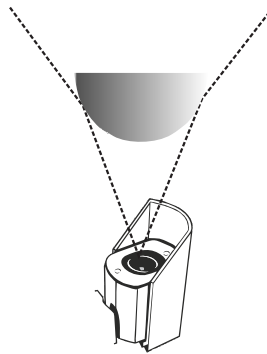


Figura 23: Sistema catadióptico com espelho esférico.

um único centro de projeção eficaz. Esse centro garante a unicidade entre um ponto do espaço tridimensional e a sua projeção em único ponto da imagem omnidirecional (BAKER; NAYAR, 1999). A razão de um ponto de projeção único é tão desejável que é um requisito para a geração de uma imagem em perspectiva pura a partir das imagens detectadas, pois é livre de borramento. Como observado em (REES, 1970; YAMAZAWA; YAGI; YACHIDA, 1995; NAYAR; BAKER, 1997; STRAUSS; COMBY, 2007; MEI; RIVES, 2007) é altamente desejável que um sistema catadióptico tenha um único ponto de projeção.

A maioria dos sistemas de imagens tem a necessidade de um único ponto de vista. Porém, os sistemas NON-SVP podem ser projetados para tarefas específicas ou características da imagem como campo de visão específico (SWAMINATHAN; GROSSBERG; NAYAR, 2006). Algumas formas de obtenção de imagens omnidirecionais não garantem o centro único de projeção, tal como no caso das lentes do tipo “olho de peixe”, que dificulta a geração de uma imagem perspectiva livre de distorções.

Assim, a obtenção de um centro único de projeção é relativamente simples e se faz importante em um sistema catadióptico. O espelho deve ser escolhido de forma tal a garantir que somente um único ponto no mundo irá refletir na superfície do espelho e incidir em um único ponto na imagem.

Existem algumas maneiras para se obter o centro único de projeção nos sistemas catadióptricos. Para o espelho hiperbólico, todo raio incidente no espelho é refletido passando pelo foco e, assim, as leis da reflexão garantem o centro único de projeção (SVOBODA; PAJDLA, 1997). No caso do espelho parabólico, a característica de centro único de projeção único é satisfeita ao se modelar o sistema com uma câmera de projeção ortográfica, onde, após a incidência no espelho, todos os raios são refletidos

paralelamente ao eixo do espelho (NAYAR; BAKER, 1997). No caso do espelho esférico, os raios incidentes são refletidos após passarem por diferentes focos, o que não garante um centro único de projeção e, por esse motivo, imagens adquiridas por espelhos esféricos são mais distorcidas (DERRIEN; KONOLIGE, 2000). Contudo, ele tem a vantagem de ter uma curvatura constante, sendo assim fácil de ser fabricado. Os espelhos cônicos tem as mesmas características dos espelhos esféricos.

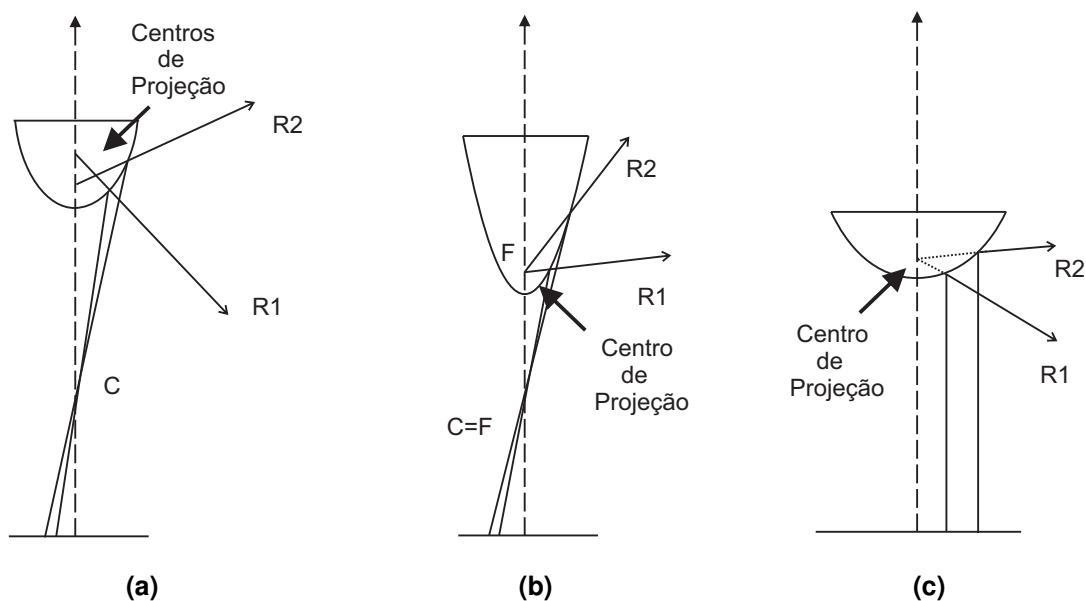


Figura 24: Centro de projeção: (a) espelho esférico (não há centro único de projeção), (b) espelho hiperbólico (os raios se interceptam no foco da hipérbole e (c) espelho parabólico (projeção ortográfica dos raios refletidos). A figura foi adaptada de (GRASSI, 2002).

As projeções em três tipos de espelhos: esférico, hiperbólico e parabólico são mostradas na Figura 24. Observa-se que o espelho esférico não possui centro único de projeção.

3.2.1 Propriedades dos espelhos

A flexibilidade dos sistemas catadióptricos deriva da possibilidade de facilmente alterar a função de transferência do sensor. A função de transferência do sensor descreve como pontos do mundo tridimensional são mapeados em pontos na imagem bidimensional. Em um sistema catadióptico, o formato do espelho determina a função de transferência do sensor.

O formato do espelho determina a formação da imagem de um sistema de visão catadióptico. Muitas formas de espelho têm sido utilizadas em diversos trabalhos. O

projeto de um perfil espelho personalizado permite superar as limitações dos sensores omnidirecionais clássicos e expandir a sua gama de aplicações.

Essas formas tem sido tão amplamente utilizadas em Robótica Móvel. Assim, pode-se discutir as principais características de cada tipo de espelho.

1. **Espelhos cônicos** tem uma boa resolução no periférico, mas eles produzem uma singularidade na ponta cone (YAGI; YACHIDA, 1991; YAGI; KAWATO, 1994; YAGI; NISHIZAWA; YACHIDA, 1995).
2. **Espelhos esféricos** tem boa resolução na área central, mas baixa resolução na região periférica. Esses espelhos apresentam o ângulo de visão mais amplo entre os espelhos convexos (GLUCKMAN; NAYAR, 1998; GASPARI; SANTOS-VICTOR, 1999; WINTERS; SANTOS-VICTOR, 1999a, 1999b; LI, 2006).
3. **Espelhos hiperbólicos** possuem uma boa resolução tanto no centro quanto na periferia. O ângulo de visão é quase tão grande quanto a do espelho esférico, mas a vantagem mais importante é a presença de um centro de projeção único localizado no foco da hipérbole. A presença de um único centro de projeção permite a reconstrução de imagens livres de distorção (YAMAZAWA; YAGI; YACHIDA, 1993, 1995; ONOE *et al.*, 1998; ZIVKOVIC; BOOIJ, 2005).
4. **Espelhos parabólicos** possuem a mesma resolução que o espelho hiperbólico ainda que com menor ângulo de visão. Eles apresentam um único centro de projeção quando usado com uma câmera de projeção ortográfica (NAYAR; BAKER, 1997; YASUSHI, 1999b; MATSUMOTO *et al.*, 2000; GEYER; DANIILIDIS, 2003).

De maneira resumida, as propriedades dos principais perfis de espelhos são mostrados na Tabela 2.

Tabela 2: Momentos invariantes para as imagens da Figura 33.

Geometria do espelho	Construção	SVP / NON-SVP	Qualidade de imagem
Cônico	Fácil	NON-SVP	Boa resolução na periferia
Esférico	Fácil	NON-SVP	Boa resolução na área central
Hiperbólico	Difícil	SVP	Boa resolução na área central e na periferia
Parabólico	Difícil	SVP (projeção ortográfica)	Boa resolução na área central e na periferia

Na Figura 25, as formas de espelho mais utilizadas são apresentadas. Como já é sabido, cada uma das formas de espelho realiza um mapeamento diferente entre

os pontos no mundo e os pontos da imagem. Esses mapeamentos determinam a resolução e o campo de visão.

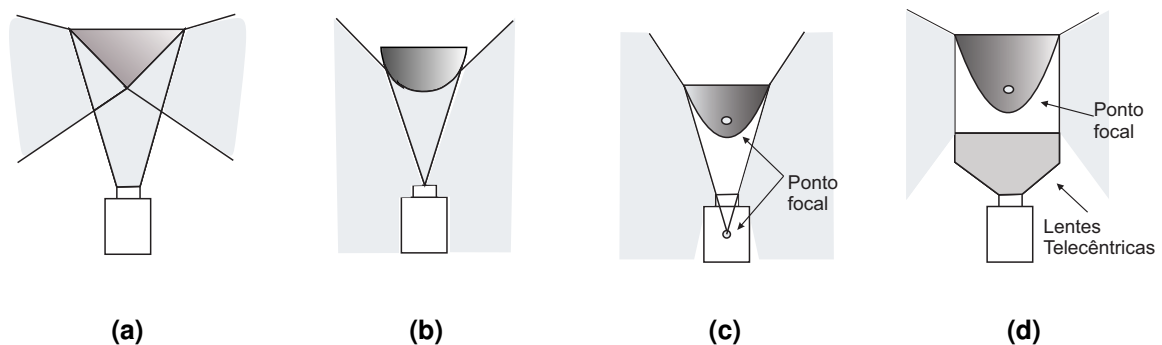


Figura 25: Perfis de espelho: (a) espelho cônico, (b) espelho esférico, (c) espelho hiperbólico e (d) espelho parabólico.

Essas topologias podem ser misturadas para explorar os benefícios dos diferentes perfis de espelhos. Ultimamente, os pesquisadores começaram a usar superfícies mais complexas, a fim de realizar funções de mapeamento personalizados entre o mundo tridimensional e a imagem.

Por exemplo, Bonarini, Aliverti e Lucioni (1999) propuseram o uso de um espelho composto por uma esfera intersectando um cone, a fim de evitar a excessiva distorção introduzida pela ponta do cone e garantir a boa resolução do espelho esférico no centro da imagem.

Hicks e Bajcsy (2000) desenvolveram um perfil de espelho que faz um mapeamento linear do plano horizontal do espaço tridimensional. O perfil foi desenvolvido com o propósito de auxiliar a navegação de robôs móveis. Como nota-se na Figura 26, todas as retas desenhadas no chão continuam retas na imagem omnidirecional, sem a necessidade de qualquer eventual processamento da imagem captada pelo sistema de visão. Esse tipo de visão ficou conhecido como “*birds eyes view*” (VASSALLO *et al.*, 2004; EHLGEN; THOM; GLASER, 2007; LIU; LIN; CHEN, 2008).

Outro exemplo é o espelho da Marchese e Sorrenti (2001), que propuseram um espelho de múltiplas partes para a tarefa específica das competições *Robocup*². Nesse espelho, cada parte é dedicada à observação de uma área diferente da área do jogo.

Tan, Hua e Ahuja (2004) propuseram um espelho piramidal que consiste em um

²www.robocup.org

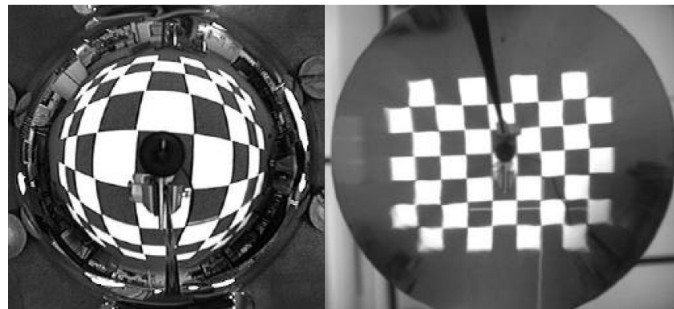


Figura 26: Imagem obtida com um espelho esférico à esquerda e com um espelho projetado por Hicks e Bajcsy (2000) à direita.

conjunto de espelhos planos dispostos em torno de um eixo de simetria e inclinados, para formar uma pirâmide. Esse perfil tem propriedades atraentes, incluindo geração de imagens com centro de projeção único e alta resolução. As imagens adquiridas usando o protótipo experimental são exibidas no trabalho.

Spacek (2005) desenvolveram uma solução que combina os benefícios dos espelhos planos (sem distorção radial, sem perda radial de resolução), com as vantagens do sensor catadióptico e sua simetria de rotação através de perfis cônicos diferenciados.

De forma a sintetizar como os sistemas catadióptricos avançaram com o passar do tempo, um quadro resumo com a classificação dos principais trabalhos sobre visão omnidirecional é exibido na Figura 27 e um levantamento da quantidade de trabalhos que envolvem visão omnidirecional e robótica móvel é mostrado na Figura 28. O gráfico apresentado na Figura 28 foi obtido utilizando, na busca, os termos *Mobile Robot* e *omnidirectional vision*.

3.3 Conclusão do Capítulo

Neste Capítulo, foi realizada uma revisão sobre imagens omnidirecionais e suas aplicações. Essa categoria de imagem merece destaque pelo ângulo de visão superior. Sabido isso, o próximo Capítulo é dedicado a apresentar como os métodos de extração de características podem ser aplicados a esse tipo de imagem e métodos de reconhecimento de padrões.

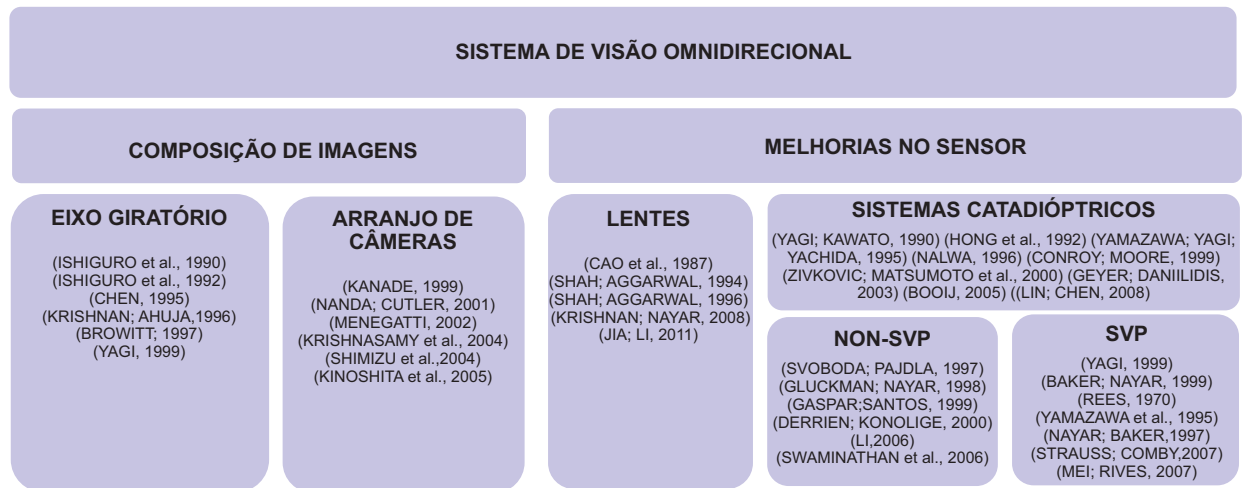


Figura 27: Classificação de trabalhos sobre Visão Omnidirecional.

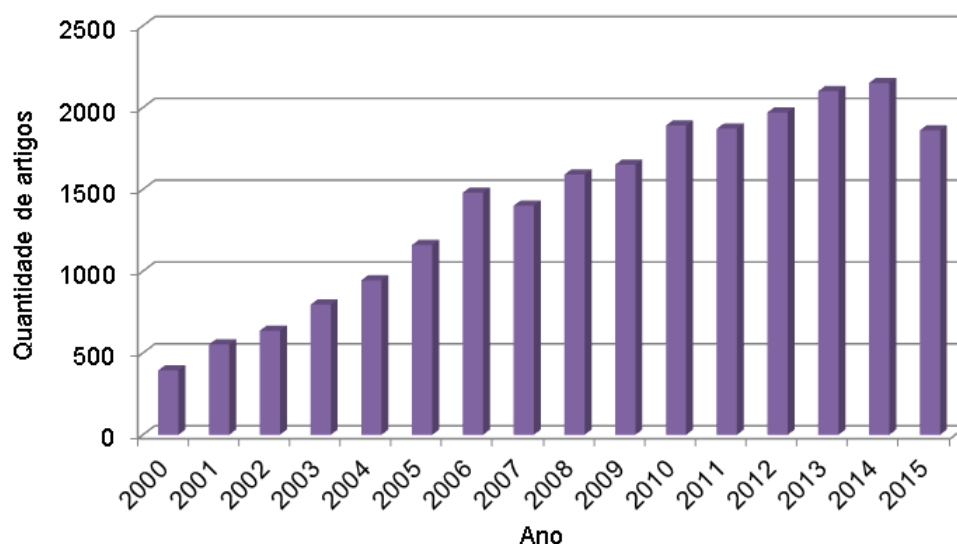


Figura 28: Levantamento de trabalhos cadastrados na plataforma *Google Scholar* na área de estudo.

RECONHECIMENTO DE PADRÕES

Um padrão é a descrição de um objeto e este padrão pode ser qualquer característica perceptível a um dos cinco sentidos, quando humano. Uma imagem de um rosto, uma impressão digital, uma palavra falada, um personagem escrito à mão, e uma forma de onda biológica são, dependendo da aplicação, alguns exemplos de padrões. Padrões são representados por um conjunto de características (atributos). Formalmente, um padrão $\mathbf{p}$ é um vetor n -dimensional, tal que $\mathbf{p} = [p_1 \ p_2 \ \dots \ p_n]^T$. Um conjunto com uma quantidade de Q padrões é representado por $D = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \dots, \mathbf{p}_Q\}$.

Reconhecimento de padrões pode ser definido como um procedimento em que se busca a identificação de certas estruturas nos padrões de entrada em comparação a padrões conhecidos e sua posterior categorização em classes, de modo que o grau de associação seja maior entre padrões de mesma classe e menor entre os padrões de classes diferentes. No caso de um sistema de localização baseado em visão computacional, os dados de entradas (padrões) são informações extraídas da representação do ambiente onde o robô se encontra. Essas informações devem ser facilmente observáveis, distinguíveis entre si e abundantes no ambiente para que o classificador funcione da melhor forma (RIISGARD; BLAS, 2005).

Neste tipo de sistema, a tarefa de reconhecimento de padrões engloba três grandes etapas: representação dos dados de entrada, extração das características e finalmente identificação e classificação do objeto em estudo (HAYKIN *et al.*, 2009). A primeira etapa refere-se à representação dos dados de entrada que podem ser mensurados a partir do objeto a ser estudado. Essa mensuração deverá descrever informações características do objeto, possibilitando a sua posterior classificação numa determinada classe. A segunda etapa consiste na extração de características ou atributos do objeto e consequente redução da dimensionalidade do vetor padrão. A escolha das características é de fundamental importância para um bom desempenho do classificador. Exige-se, portanto, um conhecimento específico sobre o problema em estudo.

A terceira etapa em reconhecimento de padrões envolve a determinação de procedimentos que possibilitem a identificação e classificação do objeto em uma classe de objetos. De modo diferente da segunda etapa, aqui a concepção do classificador pode ser abordada de forma abstrata e independente da natureza do problema. Nesta etapa, o classificador infere dentre as classes, aquela à qual o objeto pertence. Um esquema de como essas três etapas se comportam dentro de um sistema de localização baseado em visão computacional é mostrado na Figura 29.

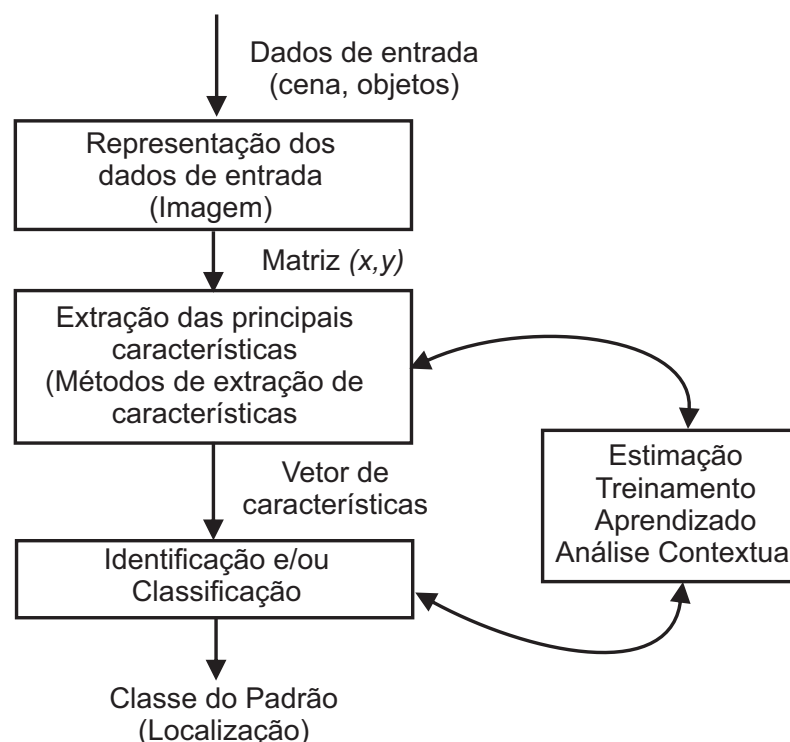


Figura 29: Fases do do processo de reconhecimento de padrões.

4.1 Extração de Características

Em alguns problemas de reconhecimento de padrões, como o deste trabalho, os dados de entrada são imagens; ou seja, os objetos do mundo real são representados por imagens digitais. Uma grande quantidade de dados é usada para representar uma imagem, de modo que a análise de imagens necessita de uma grande quantidade de memória e um custo computacional elevado. A fim de reduzir a quantidade de dados, uma imagem é representada através de um conjunto de características. Essa etapa se baseia em extrair algumas características, tais como forma e textura. De uma forma geral, buscam-se características que contenham as informações relevantes de uma

imagem. Nas subseções a seguir, alguns métodos de extração de características são descritos.

4.1.1 Histograma RGB

Uma imagem digital é formada por um grupo de pontos discretos de tons de cinza ou por combinação das cores vermelho, verde e azul (*red, green and blue, RGB*) que são geralmente organizados em uma matriz de *pixels* (SOUTO, 2000). A representação de um *pixel* é um ponto (x,y) que corresponde à sua localização na imagem. A localização $(0,0)$, normalmente, corresponde ao canto superior esquerdo da imagem, e $f(x,y)$ corresponde ao tom de cinza (que, no caso das imagens de 8 *bits*, varia de 0 a 255) ou à cor da imagem na posição (x,y) .

Uma imagem colorida é uma combinação de algumas cores básicas (vermelho, verde e azul), onde cada cor básica é representada por uma matriz diferente de tons de cinza. A representação de uma imagem colorida é de três matrizes organizadas em sequência, criando uma nova função $g(x, y, 3)$ conforme a Figura 30. Uma das formas de se extrair informações de uma imagem, seja colorida ou em tons de cinza, é o histograma.

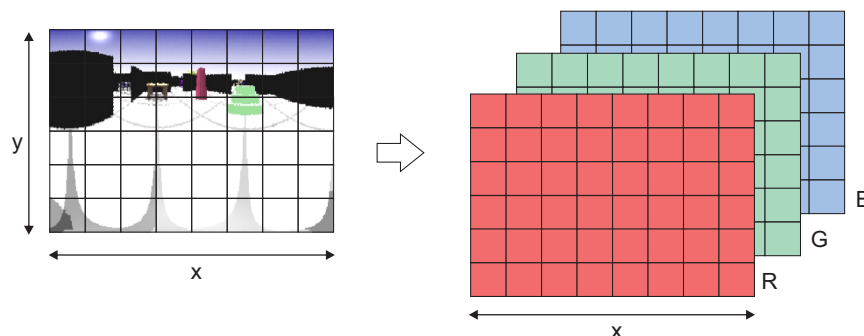


Figura 30: Imagem colorida e matrizes RGB.

Conforme Souto (2000), histograma é uma das formas mais comuns de se representar a distribuição dos valores dos *pixels* de uma imagem. Esse fornece a informação sobre quantos *pixels* na imagem possuem cada valor possível de tom de cinza ou, de forma equivalente, qual a proporção da imagem que corresponde a cada valor de tom de cinza. O histograma de uma imagem digital com níveis de cinza no intervalo $[0, L - 1]$ é uma função discreta $p(r_k) = np_k$, em que r_k é o k -ésimo nível de cinza, np_k é o número de *pixels* na imagem com esse valor de intensidade e $L = 2^b$, sendo b o número de *bits* da imagem.

Ao se observar o histograma de uma imagem, tem-se uma noção sobre as características da mesma. A forma do histograma fornece informações de grande importância no caso das imagens de sensoriamento remoto, tais como intensidade média e espalhamento dos valores de tom de cinza. Este, por sua vez, dá a medida do contraste de uma imagem: quanto maior o espalhamento ao longo do eixo dos tons de cinza, maior o contraste da imagem. Um exemplo dos três histogramas das bandas vermelha, verde e azul de uma imagem colorida é mostrado na Figura 31.

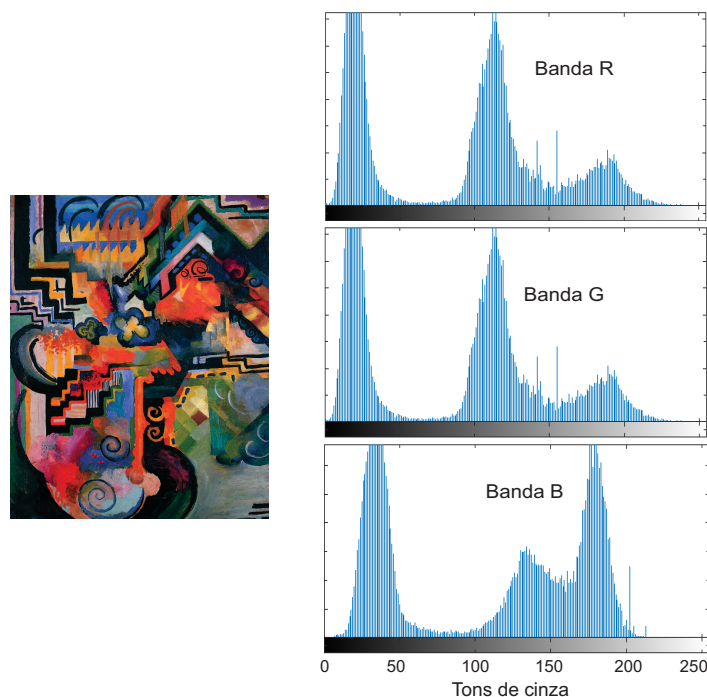


Figura 31: Imagem colorida com os respectivos histogramas de cada banda.

4.1.2 Histograma de Gradientes Orientados

O Histograma de Gradientes Orientados (*Histogram of Oriented Gradients*, HOG) foi primeiramente descrito em 2005 em Dalal e Triggs (2005). A idéia principal deste descritor é que a aparência e forma de objetos em uma imagem podem ser descritas através da distribuição dos gradientes de intensidade dos *pixels* ou pelas direções das bordas (DALAL; TRIGGS, 2005).

A implementação básica deste extrator consiste em dividir a imagem em pequenas regiões conectadas, nomeadas células, e a construção de um histograma de orientações de gradiente para cada célula. O descritor final é composto por esse conjunto de histogramas dispostos em um único vetor conforme a Figura 32(a).

Na Figura 32(b), são mostradas as células onde o gradiente se apresenta de maneira mais forte. Esses descritores foram usados tradicionalmente na detecção de objetos em cenas. Outros autores desenvolveram versões melhoradas tanto em custo computacional e eficácia para a detecção de pessoas (ZHU *et al.*, 2006).

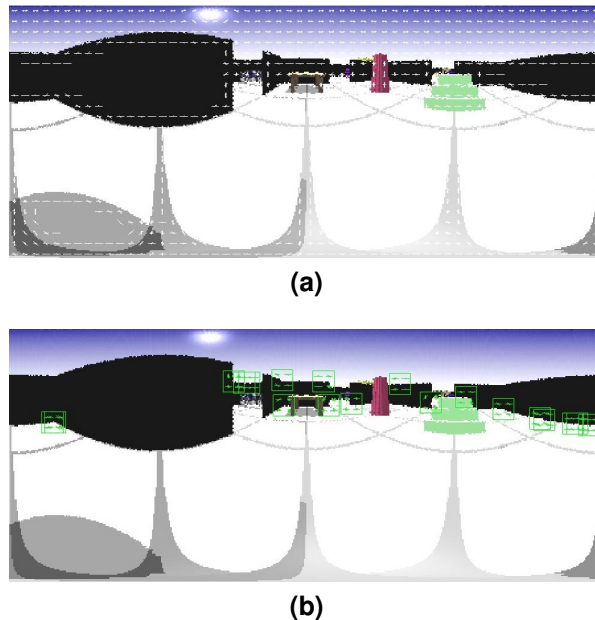


Figura 32: Descritor HOG: (a)orientações dos gradientes e (b) áreas chamadas de células e suas respectivas direções das orientações de gradiente.

O histórico de aplicação deste tipo de descritor em localização e mapeamento de robôs é muito limitado. Hofmeister, Liebsch e Zell (2009) usaram um histograma ponderado dos gradientes orientados em pequenos robôs móveis que se deslocam em um ambiente pequeno e controlado. O algoritmo desenvolvido pelos autores funciona bem sob essas condições restritivas e apenas se a orientação do robô é semelhante à orientação das imagens no mapa. Os mesmos autores desenvolveram um estudo comparativo entre HOG e outras técnicas de extração de características para a localização de robôs móveis em ambientes pequenos, com resultados semelhantes (HOFMEISTER; VORST; ZELL, 2010).

Inspirado neste método de extração de características, neste trabalho, propõe-se um descritor HOG aplicável a imagens omnidirecionais que oferece invariância à rotação do robô e permite estimar a posição deste. Mais detalhes podem ser vistos no Capítulo 5.

4.1.3 Os momentos invariantes de Hu

Os momentos invariantes de Hu têm sido frequentemente utilizados como recursos para processamento de imagem, sensoriamento remoto, reconhecimento de formas e classificação. Momentos podem fornecer características de uma cena que representam unicamente a sua forma. Hu (1962) introduz a base matemática para momentos invariantes bidimensionais e demonstrou as suas aplicações para reconhecimento de formas. Eles foram primeiramente aplicados para formas de aeronaves e se mostraram rápidos e confiáveis (DUDANI; BREEDING; MCGHEE, 1977). Wong e Hall (1978) e Flusser e Suk (1994) utilizaram os momentos invariantes em comparação de imagens de satélites. Em outros trabalhos, Belkasim, Shridhar e Ahmadi (1991) utilizaram momentos para reconhecimento de caracteres. Os momentos invariantes também tem sido aplicados com sucesso no reconhecimento de faces em fotografias (FLUSSER; SUK; SAIC, 1996), bem como para normalizar as imagens tremidas nas formas canônicas (ZHANG; WEN; ZHANG, 2000) e em imagens SAR (KUAN *et al.*, 2015).

Os momentos utilizados para construir os momentos invariantes são definidos de forma contínua, mas para implementação prática eles são computados na forma discreta (HU, 1962; PROKOP; REEVES, 1992). Os momentos invariantes podem ser utilizados para descrever uma imagem. Dada uma imagem digital $f(x,y)$ de tamanho $M \times N$, estes momentos são definidos por

$$m_{pq} = \int \int x^p y^q f(x,y) dx dy, \quad (4.1)$$

em que M_{pq} é o momento bidimensional da função $f(x,y)$. A ordem do momento é $(p+q)$, onde p e q são ambos números naturais. A versão discreta da Equação 4.1 é dada por

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x,y). \quad (4.2)$$

Para normalizar e se traduzir no plano da imagem, os centróides de imagem são usados para definir os momentos centrais. As coordenadas do centro de gravidade da imagem são calculadas utilizando as seguintes equações (GONZALEZ, 2009):

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \bar{y} = \frac{m_{01}}{m_{00}}. \quad (4.3)$$

Os momentos centrais podem, então, ser definidos em sua representação discreta como

$$\tau_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y). \quad (4.4)$$

Os momentos são normalizados para os efeitos de mudança de escala, na forma

$$\eta_{pq} = \frac{\tau_{pq}}{\tau_{00}^\gamma}, \quad (4.5)$$

em que o fator de normalização $\gamma = \frac{p+q}{2} + 1$. A partir disso, um conjunto de sete valores $(\phi_1, \dots, \phi_7)$ podem ser calculados e são definidos por

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \\ \phi_5 &= (3\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &\quad (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (4.6)$$

Para verificar a eficácia do extrator, três imagens foram utilizadas para calcular e comparar os momentos invariantes apresentados anteriormente e estão dispostas na Figura 33.

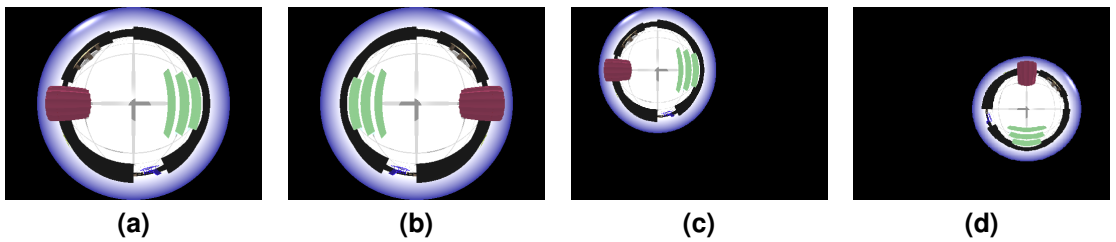


Figura 33: Imagens de teste utilizadas: (a) imagem original, (b) imagem espelhada, (c) imagem transladada e reduzida e (d) imagem rotacionada e reduzida.

Tabela 3: Momentos invariantes para as imagens da Figura 33.

Hu	Imagem original	Imagem espelhada	Imagem translada e reduzida	Imagem rotacionada e reduzida
ϕ_1	0,7946	0,7946	0,7980	0,7952
ϕ_2	0,0099	0,0099	0,0079	0,0068
ϕ_3	0,4821	0,4821	0,5244	0,4798
ϕ_4	0,1481	0,1481	0,1566	0,1566
ϕ_5	0,0021	0,0021	0,0054	0,0060
ϕ_6	-0,0338	-0,0338	-0,0494	-0,0415
ϕ_7	-0,0395	0,0395	-0,0445	-0,0425

O contorno preto foi introduzido para fazer com que todas as imagens fossem das mesmas dimensões. A imagem original é mostrada na Figura 33(a). A Figura 33(b) é uma representação da imagem original transladada e reduzida em escala e na Figura 33(c), a imagem original é rotacionada e reduzida. Os valores dos setes momentos são resumidos na Tabela 3. Pode-se notar que os valores dos momentos são semelhantes independente da rotação, translação e mudança de escala e observe que o sinal de ϕ_7 da imagem espelhada é diferente da original; essa propriedade é utilizada para detectar espelhamento.

4.1.4 Análise de componentes principais (PCA)

Em 1987, Sirovich e Kirby (1987) apresentaram trabalho empregando uma técnica de estatística multivariada, conhecida como *Principal Components Analysis* (PCA), para representar imagens de face em um subespaço reduzido preservando o máximo de informação original.

Análise de Componentes Principais é uma poderosa técnica de estatística multivariada para a extração de características (componentes) de um conjunto de dados de alta-dimensão, que correspondem à extração dos autovetores associados aos auto-valores a partir de um conjunto de dados de entrada. Em particular, permite identificar as principais direções nas quais os dados variam.

Por exemplo, na Figura 34, os círculos representam um conjunto de dados, onde duas variáveis são medidas e dispostas no plano X-Y do sistema de coordenadas. A direção principal na qual os dados variam é mostrada pelo eixo U e a segunda direção pelo eixo ortogonal V. Para um determinado conjunto de dados, a análise de componentes principais encontra o sistema de eixos definidos pelas direções principais de variância (ou seja, o sistema U-V). As direções U e V são chamadas de componentes principais. Em termos computacionais as componentes principais são encontradas através do cálculo dos autovetores e autovalores da matriz de covariância

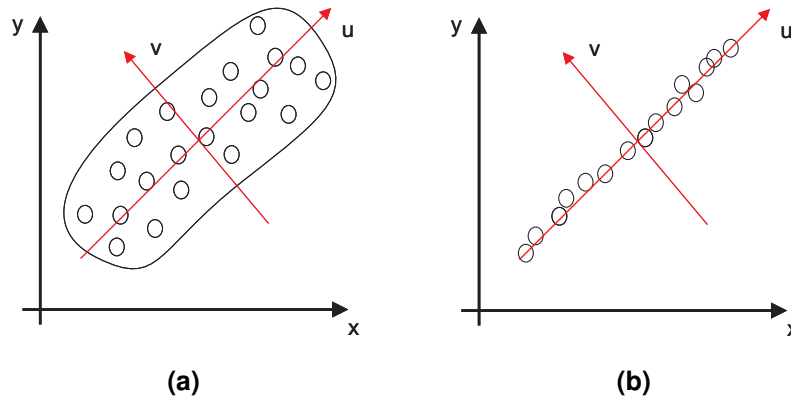


Figura 34: Análise de Componentes Principais: (a) Dados de entrada e (b) Dados reduzidos com PCA.

dos dados. Este processo é equivalente a encontrar o sistema de eixos no qual a matriz de covariância é diagonal.

Quando se dispõe de um conjunto de imagens $f_j(x, y), \in \mathbb{R}^{M \cdot N}, j = 1, \dots, e$, onde cada imagem pode ser considerada como um vetor de dados que cai em um espaço com dimensão $M \cdot N$. No entanto, esses vetores foram gerados com um grande número de informações. Como estes dados estão altamente correlacionados, a filosofia da PCA consiste na realização de uma redução de dimensionalidade com o objetivo de reter as informações mais relevantes de cada cena.

Usando a formulação clássica de PCA, como exposto em (NAYAR; NENE; MURASE, 1996) (KIRBY, 2000), pode-se projetar o conjunto de vetores de dados (imagens) $\vec{f}_j(i) \in \mathbb{R}^{M \cdot N} \quad j = 1, \dots, e$, sendo e o número de imagens e $M \cdot N$ o número de *pixels* em cada imagem, $e \ll M \cdot N$, dentro de um conjunto de vetores de características chamado projeções das imagens $\vec{p}_j(i) \in \mathbb{R}^{C \times 1} \quad j = 1, \dots, e$, sendo C as características PCA que contêm as informações mais relevantes da a imagem ($C \leq e$). Essa projeção é calculada por

$$\mathbf{P} = \mathbf{V}^T \cdot \mathbf{F}, \quad (4.7)$$

em que $\mathbf{F}$ é a matriz de dados constituída pelos vetores de dados dispostos em colunas e $\mathbf{P}$ é a matriz de projeção, contendo o as projeções dos dados no novo espaço. $\mathbf{V}$ contém os principais autovetores de C da matriz de covariância dispostos em colunas. Após esse processo, cada imagem com $M \cdot N$ *pixels* foi transformada em um vetor com C componentes (KIM; JUNG; KIM, 2002).

Alguns autores têm aplicado PCA em localização de robôs móveis (LEONARDIS;

BISCHOF, 2000) (KROSE *et al.*, 2004). No entanto, essa formulação clássica de PCA apresenta duas desvantagens. Em primeiro lugar, as projeções não podem ser calculadas de maneira incremental, ou seja, todos os dados devem estar dispostos para o processamento inicial; e em segundo lugar, os descritores obtidos não são invariáveis a mudanças da orientação do robô. Para superar esta limitação segunda, esse trabalho faz uso da formulação proposta por (JOGAN; LEONARDIS, 2000), onde o banco de dados é criado tendo em conta todas as possíveis orientações do robô durante a captura de cada imagem de referência.

4.1.5 Essência da cena

A essência da cena (*GIST of the scene*, GIST) tenta imitar a capacidade humana de reconhecer imediatamente uma cena por meio de algumas regiões que possuem uma cor e/ou uma textura em destaque no que diz respeito ao resto da imagem. O recurso GIST é um recurso de imagem global, o que caracteriza vários dados importantes sobre uma cena (OLIVA; TORRALBA, 2006). Além disso, pode codificar a quantidade ou a força de linhas verticais ou horizontais em uma imagem, o que, por exemplo, pode ajudar a combinar cenas com linhas semelhantes.

Esse descritor foi introduzido por Oliva e Torralba (2001) e Oliva e Torralba (2006) como uma representação holística do espaço. Matematicamente, eles codificam a informação espacial através da Transformada de Fourier 2D calculada em várias regiões distribuídas igualmente por toda a imagem conforme a Figura 35. Esse conjunto de dados é, então, dimensionalmente reduzido por meio da aplicação de PCA.

Mais recentemente, os mesmos autores fizeram uso de pirâmides Wavelet em vez de Transformadas de Fourier (TORRALBA *et al.*, 2003). Os autores mostram como esses tipos de descritores são capazes de classificar conjuntos de imagens de acordo com características subjetivas, como grau de naturalidade, abertura, rugosidade, etc.

Pode-se encontrar na literatura algumas aplicações do descritor GIST em robótica móvel. Por exemplo, Chang, Siagian e Itti (2011) apresentam um sistema de localização com base no descritor GIST e Murillo *et al.* (2013) fazem uso de descritores GIST para resolver o problema de localização nas zonas urbanas. No entanto, estas duas obras calculam descritores para determinadas áreas da imagem. O objetivo desse trabalho é testar o desempenho do descritor GIST inspirado em (SIAGIAN;

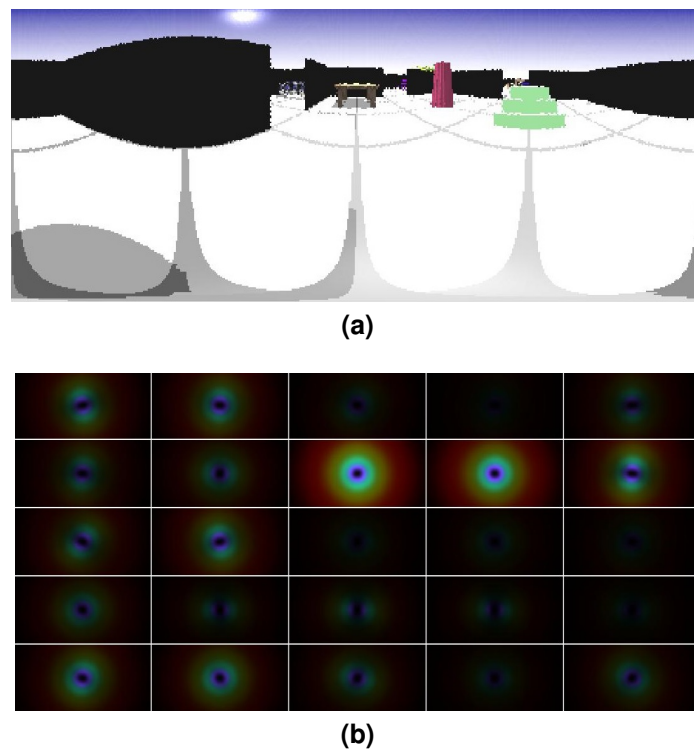


Figura 35: Descritor Gist: (a) imagem original, (b) transformada Fourier 2D calculada em várias regiões.

ITTI, 2007) onde se cria um único descritor por cena que, quando calculado a partir de uma imagem panorâmica, apresenta invariância à rotação do robô. Esse modelo será descrito, em detalhes, no Capítulo 5.

4.2 Classificadores

Como já dito anteriormente, a terceira etapa em reconhecimento de padrões envolve a determinação de procedimentos que possibilitem a identificação e classificação do objeto em uma classe de objetos. Aqui, o classificador é abordado independente da natureza do problema. Nas subseções a seguir, os classificadores utilizados nesse trabalho são discutidos.

4.2.1 Classificadores baseados em Redes Neurais

As Redes Neurais Artificiais (RNA) foram originalmente concebidas com intuito de modelar o processamento da informação no sistema nervoso biológico para se obter soluções para problemas computacionais mais complexos, em nível compatível ao

dos humanos. A partir desta subseção, algumas técnicas baseadas em redes neurais artificiais serão descritos.

MLP (*MultiLayer Perceptron*)

O perceptron ou neurônio artificial é um modelo do processamento da informação em uma célula nervosa (neurônio). É a forma mais básica, ou seja, menor unidade de uma rede neural. O primeiro modelo matemático para um neurônio biológico foi proposto por McCulloch e Pitts (1943). A saída de um neurônio depende das entradas e pesos associados a ele e pode servir de entrada para outros, formando uma rede disposta em camadas. As redes perceptron de uma única camada não resolvem problemas de classificação não linearmente separáveis. Neste caso, necessita-se de uma abordagem sobre redes de múltiplas camadas (GARDNER; DORLING, 1998).

As redes de múltiplas camadas, sob algumas condições, resolvem problemas cujos os padrões não são linearmente separáveis. Essas redes possuem em uma camada de entrada que recebe os vetores de atributos, uma ou mais camadas intermediárias compostas de neurônios com função de ativação não linear e uma camada de saída, também composta por neurônios (que podem ser lineares ou não) (ROCHA NETO *et al.*, 2006). As camadas intermediárias (escondidas) são as que tornam a rede capaz de extrair as características mais significativas do espaço de atributos de entrada. A arquitetura típica das redes de múltiplas camadas com uma única camada escondida pode ser vista na Figura 36.

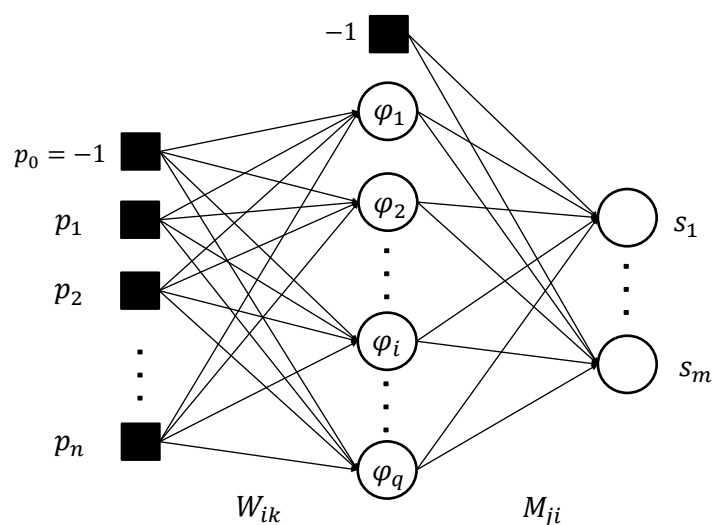


Figura 36: Arquitetura típica de uma rede neural supervisionada.

As redes MLP são treinadas com base na propagação do sinal em dois sentidos: direto e inverso. O algoritmo utilizado na fase de treinamento que permite a atualização dos pesos da camada oculta é conhecido como *backpropagation* (HAYKIN, 2001). Este algoritmo é descrito por Haykin *et al.* (2009) e pode ser resumido nos passos a seguir.

1. Os pesos são inicializados de forma aleatória.
2. No sentido direto, um padrão é apresentado à rede como entrada e as ativações dos neurônios da camada escondida são calculadas por

$$u_i(t) = \sum_{k=0}^n W_{ik}(t)p_k(t), \quad i = 1, \dots, q, \quad (4.8)$$

em que q indica o número de neurônios da camada escondida, W_{ik} é o k -ésimo peso do i -ésimo neurônio da camada oculta.

3. Propaga-se o sinal da camada escondida de tal maneira que a saída dos neurônios da camada oculta é calculada por

$$v_i(t) = \varphi_i[u_i(t)], \quad (4.9)$$

em que φ_i assume geralmente as formas da função logística ou função tangente hiperbólica (HAYKIN *et al.*, 2009).

4. Após isso, deve-se calcular as ativações e as saídas para os neurônios da camada de saída, conforme

$$a_j(t) = \sum_{i=0}^q M_{ji}(t)v_i(t), \quad j = 1, \dots, m, \quad (4.10)$$

$$s_j(t) = \varphi_j[a_j(t)]. \quad (4.11)$$

em que m é o número de neurônios da saída, M_{ji} é o i -ésimo peso do j -ésimo neurônio de saída e s_j é a j -ésima saída.

5. No sentido reverso deve-se calcular o erro e realizar a atualização do vetor de pesos para cada neurônio da rede. Para tal, o erro para o j -ésimo neurônio da camada de saída pode ser obtido diretamente por

$$erro_j(t) = d_j(t) - s_j(t), \quad (4.12)$$

uma vez que se conhece a saída desejada d_j . Assim, a regra de aprendizagem descrita por (HAYKIN *et al.*, 2009) pode ser utilizada para atualizar os pesos dos neurônios da saída

$$M_{ji}(t+1) = M_{ji}(t) + \lambda s'_j(t) erro_j(t) v_i(t), \quad (4.13)$$

em que $s'_j(t)$ é a derivada da função sigmóide utilizada (função logística ou tangente hiperbólica) e λ é a taxa de aprendizagem.

6. Para calcular o erro para os neurônios da camada oculta, uma vez que não se possui a saída desejada, usa-se o algoritmo *backpropagation* (RUMELHART; HINTON; WILLIAMS, 1988). Assim, a regra de aprendizagem para atualização dos pesos dos neurônios da camada oculta é dada por:

$$W_{ik}(t+1) = W_{ik}(t) + \lambda v'_i(t) s_i(t) p_k(t), \quad (4.14)$$

em que o erro retropropagado para i -ésimo neurônio da camada de oculta deve ser obtido com base na contribuição relativa dos erros $erro_j(t)$ obtidos para os neurônios da camada de saída conforme a seguinte equação

$$erro_i(t) = \sum_{k=1}^m M_{ji} s'_j(t) erro_j(t). \quad (4.15)$$

7. Os passos 2 a 7 são repetidos até que a rede convirja ou que um número de épocas seja alcançado.

No campo da Robótica, as redes MLP's tem ampla utilização, uma vez que são aplicadas em navegação (ORTEGA; CAMACHO, 1996), no controle de trajetória (BUGEJA; FABRI, 2006), em tarefas de mapeamento (VIDAL *et al.*, 2013) e em veículos auto-guiados (GHIASEDDIN; SERYASAT; HADDADNIA, 2013).

SOM (Self-Organizing Map)

A rede neural *Self-Organizing Map* (SOM) possui um aprendizado não-supervisionado. De uma maneira geral, redes neurais não-supervisionadas, tal como a SOM, extraem características estatísticas dos dados de entrada e constroem, de forma auto-organizada (sem auxílio externo e sem conhecimento prévio), uma representação do espaço de entrada (KOHONEN, 2001).

Esta rede difere de outras redes não-supervisionadas pelo fato de seus neurônios estarem dispostos em uma grade, de modo que possa existir uma relação de vizinhança espacial entre esses neurônios. A função de vizinhança funciona como uma espécie de janela de ponderação, que faz com que os neurônios mais próximos do neurônio vencedor atual tenham seus vetores de pesos atualizados mais intensamente do que aqueles neurônios que estão mais distantes do neurônio vencedor (KOHONEN, 2001). Na Figura 37(a) é apresentada uma rede bidimensional em uma topologia quadrada formada por nove neurônios arranjados numa disposição 3×3 , assim como a rede SOM como 9 neurônios com uma topologia unidimensional é mostrada na Figura 37(b). No exemplo, tem-se ainda que o primeiro e o último neurônios são definidos como vizinhos.

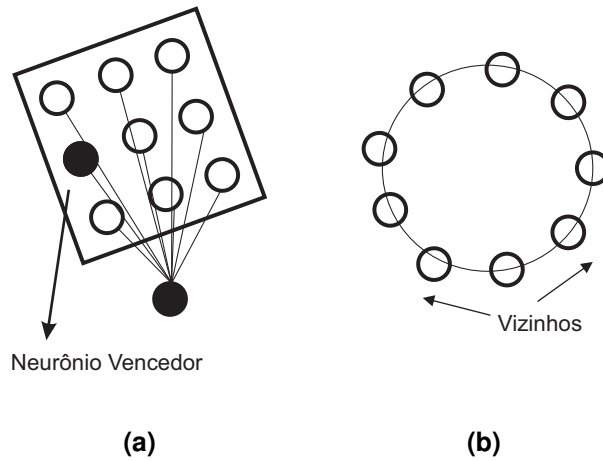


Figura 37: Modelo de uma rede SOM: (a)2d e (b)1d.

O treinamento da rede SOM é resumido a seguir pode ser resumido nos passos a seguir.

1. Seleção aleatória de um padrão de treinamento $\mathbf{p}(t)$ como vetor de entrada atual.
2. Quando um padrão é apresentado a rede SOM, os neurônios competem entre si, e a rede determina qual neurônio é mais semelhante ao padrão de entrada. A busca pelo neurônio vencedor para o padrão de entrada $\mathbf{p}(t)$ é dado por

$$i^*(t) = \arg \min_{\forall i} \|\mathbf{p}(t) - \mathbf{W}_i(t)\|. \quad (4.16)$$

3. Após a busca pelo neurônio vencedor, atualiza-se o vetor de pesos do neurônio vencedor $\mathbf{W}_{i^*}(t)$ pela equação

$$\mathbf{W}_i(t+1) = \mathbf{W}_i(t) + \lambda(t)h(i^*, i; t)[\mathbf{p}(t) - \mathbf{W}_i(t)], \quad (4.17)$$

em que $h(i^*, i; t)$ é função de vizinhança, geralmente do tipo gaussiana, ou seja,

$$h(i^*, i; t) = e^{-\frac{\|p_i(t) - p_{i^*}(t)\|^2}{\vartheta^2(t)}}, \quad (4.18)$$

tal que $\vartheta(t)$ é o raio de influência da função de vizinhança, enquanto $p_i(t)$ e $p_{i^*}(t)$ são, respectivamente, as posições dos neurônios i e i^* no arranjo geométrico da grade.

A função de vizinhança funciona como uma espécie de janela de ponderação e faz com que os neurônios mais próximos do neurônio vencedor no instante t , tenham seus vetores de pesos atualizados mais intensamente do que aqueles neurônios que estão mais distantes do neurônio vencedor. Desta forma, em cada interação, os neurônios vencedores, bem como os neurônios que fazem parte da sua vizinhança têm seus pesos ajustados de forma a responder melhor a este padrão de entrada.

A rede neural SOM geralmente não é utilizada para classificação de padrões e, sim, para agrupamento de dados (ROCHA NETO *et al.*, 2006). Porém, em tarefas de classificação, após a rede SOM treinada, é possível efetuar a rotulação dos neurônios. Uma estratégia possível a ser seguida é descrita por ROCHA NETO *et al.* (2006), onde uma associação para cada padrão de treinamento e o seu respectivo neurônio (vencedor) é realizado. Após isso, k padrões estão associados a um determinado neurônio, em seguida atribui-se ao neurônio a classe com mais ocorrências dentre os k padrões associados. Assim, quando um novo padrão chega a rede, esta verifica qual o neurônio vencedor e atribui a esse padrão o rótulo do neurônio vencedor.

4.2.2 Classificadores Estatísticos

Alguns classificadores estatísticos clássicos são apresentados, a saber: K-Vizinhos mais Próximos e Discriminante Linear e Quadrático.

K-Vizinhos mais Próximos

O algoritmo K-vizinhos mais próximos (*K-nearest neighbors*, KNN) é um classificador que categoriza um padrão com base nos rótulos dos K vizinhos mais próximos deste padrão (FUKUNAGA; HUMMELS, 1987). Considere um conjunto de dados com Q amostras, todas rotuladas, $D = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \dots, \mathbf{p}_Q\}$, onde $\mathbf{p}$ é um padrão que deve ser classificado e $\mathbf{p}$ não pertence ao conjunto D . A idéia do algoritmo KNN

é atribuir a mesma classe ao padrão $\mathbf{p}$ dos K vizinhos mais próximos (ROCHA NETO *et al.*, 2006).

Na Figura 38, a idéia do algoritmo é apresentada, onde o padrão “quadrado” é desconhecido e deve ser classificado dentre as classes “círculo” e “triângulo”. Assim, as K menores distâncias entre o “quadrado” e os demais padrões são calculadas. Observa-se o rótulos desses K padrões mais próximos, e atribui-se a $\mathbf{p}$ a moda do vetor de K rótulos. Então, o rótulo desse é comparado com o rótulo de $\mathbf{p}$. Se igual, um acerto é computado; caso não, um erro.

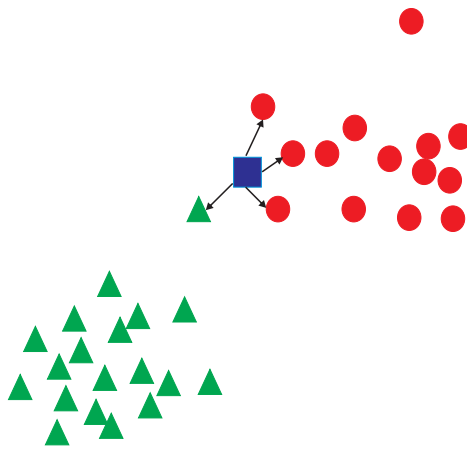


Figura 38: Representação do algoritmo K-vizinhos mais próximos.

Este algoritmo baseia-se em uma função de distância entre dois padrões, quando $K = 1$, dada por

$$\min_j [dist(\mathbf{p} - \mathbf{p}_{(j)}^{(i)})] \leq \min_j [dist(\mathbf{p} - \mathbf{p}_o^{(i)})], \forall j \neq o, \quad (4.19)$$

onde $p_{(j)}^{(i)}$ é o j -ésimo padrão pertencente à classe ω_i e $dist(.,.)$ é a medida de distância, que pode ser efetuada de várias maneiras, porém a forma mais comum, também conhecida como distância euclidiana, é dada por

$$dist_{eucl}(a, b) = \sqrt{\sum_{t=1}^n (a_t - b_t)^2}, \quad (4.20)$$

em que a_t e b_t são as componentes dos padrões a e b .

Resumidamente, o algoritmo KNN é decrito pelos passos a seguir.

1. O conjunto de amostras é dividido em dois conjuntos: treinamento e teste.

2. Os padrões do conjunto de teste são apresentados, um a um, e procura-se os K padrões mais próximos do conjunto de treinamento através do cálculo da distância.
3. O rótulo mais presente nos K padrões mais próximos é ofertado ao padrão de teste que foi apresentado.
4. Se o rótulo atribuído for igual ao rótulo do padrão de teste, computa-se um acerto.

O classificador *KNN* foi utilizado nas mais diversas aplicações, entre elas: diagnóstico de patologias da coluna vertebral (ROCHA NETO *et al.*, 2006), reconhecimento de caracteres (SANTOS, 2010), seleção de características e classificação de abelhas (BUANI *et al.*, 2002), identificação de falhas de manipuladores robóticos (PINA; PINA FILHO,) e visão robótica (PIZO, 2014).

Discriminante Gaussiano Linear e Quadrático

A Análise Discriminante Quadrática (*Quadratic Discriminant Analysis*, QDA) se diferencia da Análise Discriminante Linear (*Linear Discriminant Analysis*, LDA) (LIU; WECHSLER, 2002) no que diz respeito a separação das regiões distintas, no espaço multiespectral, correspondente a cada classe, pelas suas fronteiras quadráticas. De uma forma geral, tanto QDA quanto LDA podem fornecer resultados satisfatórios. Dependendo das características do problema que poder ser linearmente separável ou não.

No caso do classificador quadrático, estima-se para cada classe um vetor média e uma matriz covariância distintos, resultando em funções de decisão quadráticas (THEODORIDIS; KOUTROUMBAS, 2009).

A disposição dos dados das classes é um fator importante. Quando as classes apresentam distribuições onde suas médias são bem distintas seria a situação ideal, para fins de classificação, especialmente se a dispersão interna de cada classe (variância) é pequena. A matriz covariância associada a uma classe contém informações a respeito do formato da distribuição dos dados, ou seja, classes que compartilham o mesmo vetor média podem ainda assim serem separadas com certa precisão, quando as matrizes covariância, são suficientemente distintas. Fukunaga e Hummels (1987) mostram que esta capacidade de separação cresce com a dimensionalidade dos dados. O método linear (LDA) pressupõe que a diferença entre

as classes está relacionada apenas com o vetor média das classes, assumindo que a matriz covariância é igual para todas as classes.

Sendo Q o número total de amostras de treinamento. O classificador QDA aplica a função decisão apresentada na Equação 4.21, e classifica o padrão teste $\mathbf{p}$ de entrada na classe que maximiza

$$G_i(\mathbf{p}) = -\frac{1}{2}(\mathbf{p} - \hat{\mu}_i)^T \hat{\Sigma}_i^{-1}(\mathbf{p} - \hat{\mu}_i) + \ln(P(\omega_i)) + c_i \quad (4.21)$$

em que μ_i é a média para a i -ésima classe, $\hat{\Sigma}_i^{-1}$ é a inversa da matriz de covariância para a i -ésima classe, $P(\omega_i)$ é a probabilidade *a priori* e c_i é uma constante que pode ser negligenciada (THEODORIDIS; KOUTROUMBAS, 2009). Esta função de decisão claramente significa um classificador quadrático, isto é, inclui termos quadráticos, resultantes do termo $(\mathbf{p} - \hat{\mu}_i)^T \hat{\Sigma}_i^{-1}(\mathbf{p} - \hat{\mu}_i)$. Assim, as fronteiras de decisão assumem a forma de superfícies (ou hiper-superfícies) quadráticas, como elipsóides (ou hiper-elipsóides). Na Figura 39, são apresentadas duas superfícies de decisão geradas pelo classificador QDA.

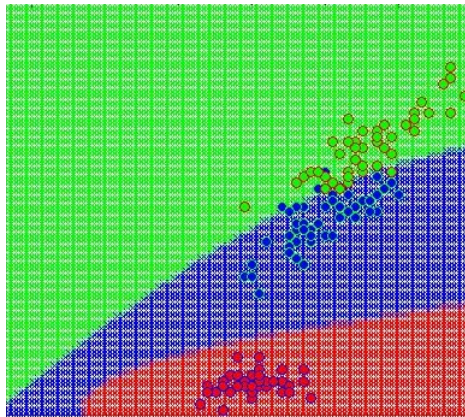


Figura 39: Superfícies de decisão geradas por análise de discriminante quadrático em um conjunto de dados com 3 classes distintas.

4.3 Conclusão do Capítulo

Conhecidos os métodos de extração de características e os classificadores utilizados nesse trabalho, no próximo Capítulo é apresentada a implementação do sistema para localização de robôs móveis através de reconhecimento de padrões em imagens omnidirecionais.

MATERIAIS E MÉTODOS

Neste capítulo, os métodos de aquisição das imagens omnidirecionais e de extração de informações das mesmas são apresentados. Trata-se também do modo de aplicação dos métodos de classificação utilizados para localização de robôs móveis.

5.1 Descrição do fluxo de informação para comparação dos métodos

Seguindo o fluxo de informação, pode-se iniciar a descrição do trabalho realizado pelo sistema catadióptrico formado pela câmera apontada para o espelho esférico de tal maneira a formar a imagem omnidirecional. A câmera é focalizada na direção do espelho. Uma vez obtida a imagem digital, esta é submetida à análise pelo sistema de localização propriamente dito, que inicia pela extração de características. Nesta aplicação, cinco métodos de extração são implementados e testados. Os vetores de características são, assim, apresentados a um classificador para que se possa realizar o reconhecimento do ambiente. Nessa etapa, os testes foram realizados através de cinco classificadores diferentes. Após classificados, pode-se analisar as vantagens e desvantagens de cada extrator e classificador. O fluxo de informação é ilustrado na Figura 40. Nota-se que a Figura 40 é uma expansão da Figura 1.

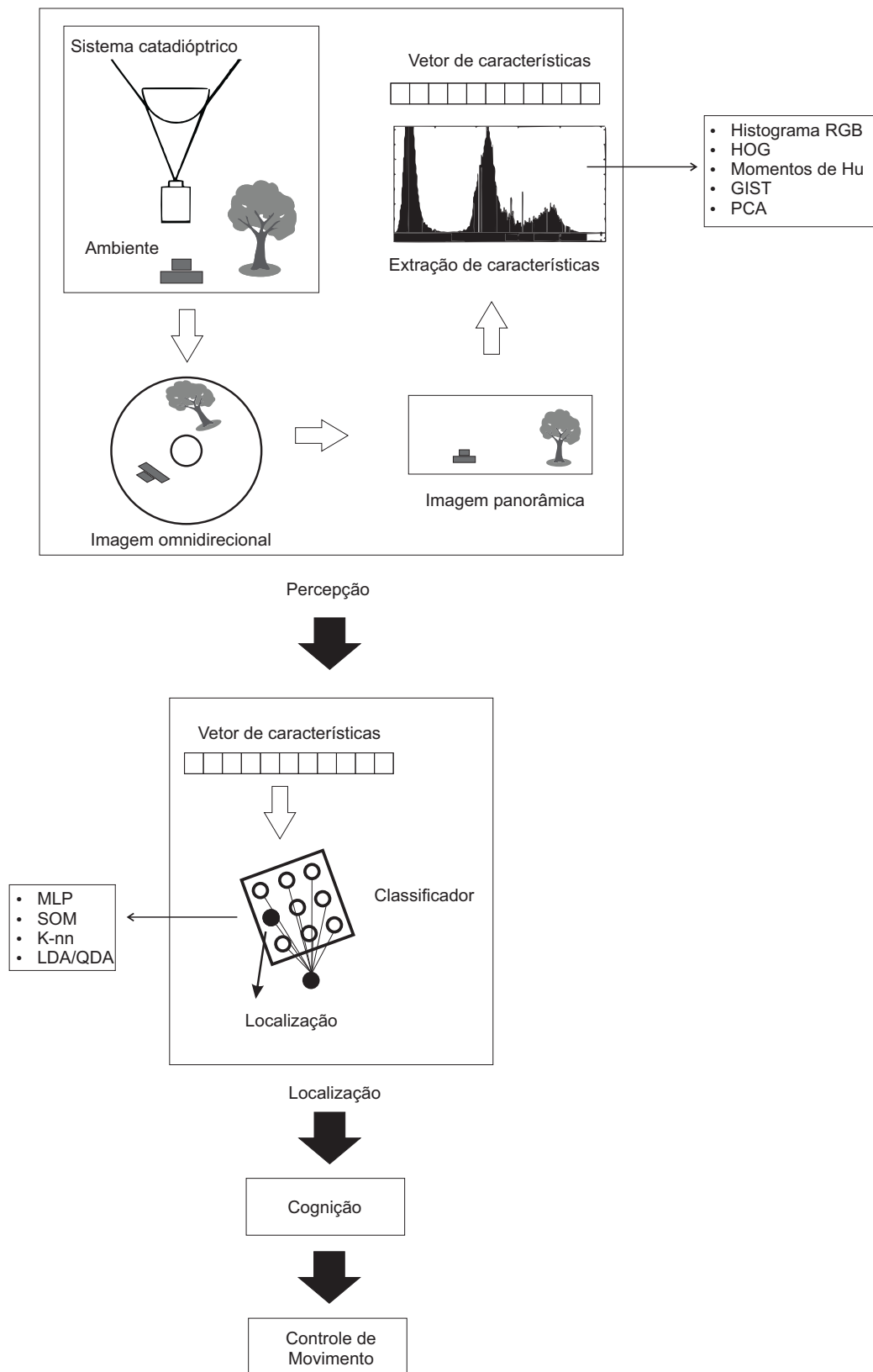


Figura 40: Fluxo de informação do método de localização implementado.

5.2 Aquisição de imagens

Como é sabido, as imagens omnidirecionais fornecem um campo de visão de 360° em uma única imagem. Quando um espelho convexo é colocado em frente a uma câmera convencional para essa finalidade, tem-se um sistema de aquisição de imagens chamado catadióptrico conforme a Figura 41(a). Na Figura 41(b), é mostrado o robô utilizado neste trabalho com o sistema catadióptrico montado.

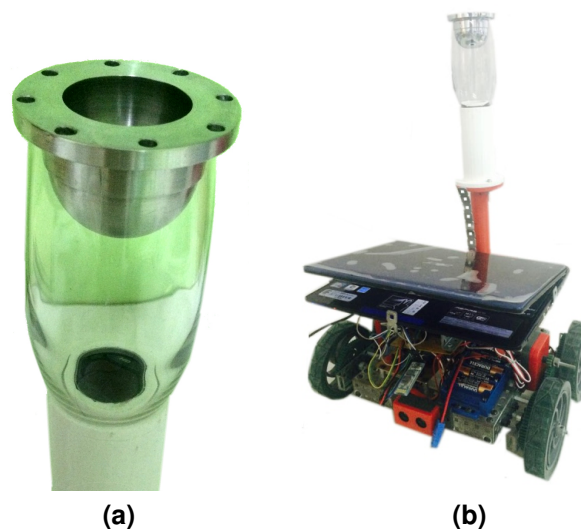


Figura 41: Sistema catadióptrico montado para o sistema: (a) sistema catadióptrico e (b) vista do robô com o sistema catadióptrico.

Os bancos de imagem utilizadas nesse trabalho estão divididas em três grupos:

1. imagens sintéticas (512 x 384 *pixels*);
2. imagens do ambiente sintético (512 x 384 *pixels*); e
3. imagens do ambiente real (640 x 480 *pixels*).

5.2.1 Imagens Sintéticas

Em relação às imagens sintéticas, a escolha das imagens foi norteadada pelo formato das imagens omnidirecionais reais. Formas como quadrados, círculos, triângulos, estrelas e setas de cores distintas foram adicionadas às imagens conforme a Figura 42. Assim, criou-se cinco grupos de imagens distintos descritos na Tabela 4.

Algumas dessas imagens foram expostas à presença de ruído; este foi adicionado às imagens através da função *speckle* (BESSA *et al.*, 2015). O *speckle* é um ruído

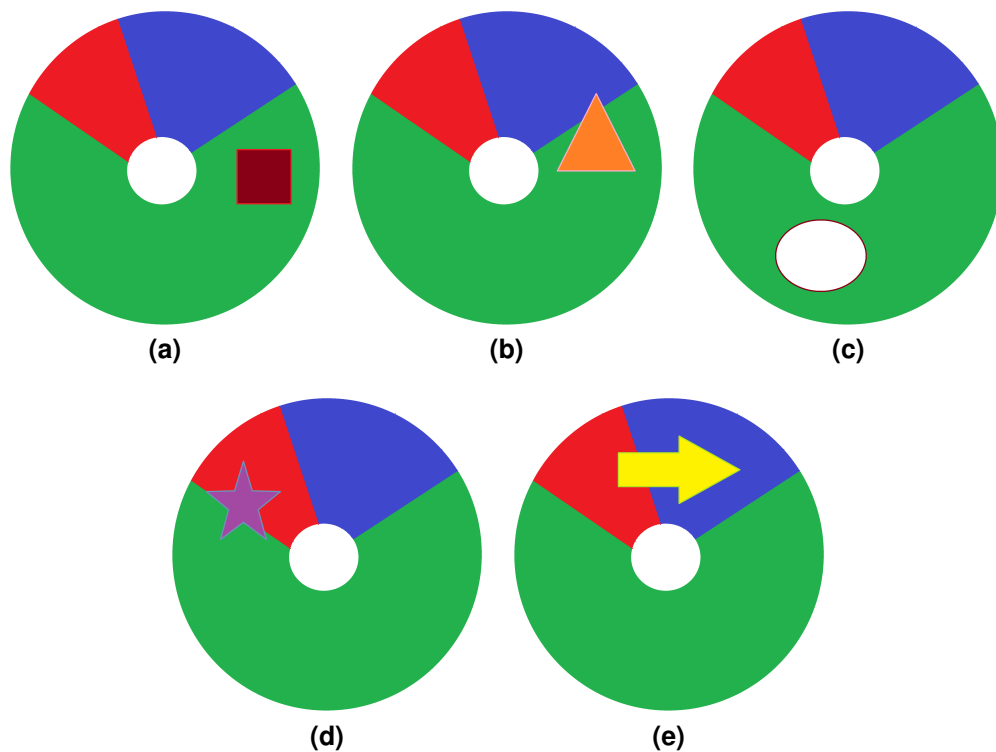


Figura 42: Imagens sintéticas: (a) quadrado, (b) triângulo, (c) círculo, (d) estrela e (e) seta.

multiplicativo proporcional à intensidade do sinal recebido. O efeito visual desse ruído proporciona uma textura granulosa que pode dificultar a interpretação das imagens, reduzindo o desempenho dos classificadores de padrões.

Vale ressaltar que as imagens sintéticas foram utilizadas inicialmente de forma a encorajar o uso do método em imagens reais. Somente após o sucesso em imagens sintéticas, outros grupos de imagens foram testados.

Tabela 4: Grupos de imagens sintéticas.

Grupo	Forma	Quantidade de imagens
1	Quadrado	25
2	Triângulo	28
3	Círculo	23
4	Estrela	24
5	Seta	29

5.2.2 Imagens do ambiente sintético

O ambiente sintético utilizado neste trabalho foi construído no ambiente de programação *POV-Ray*. O *POV-Ray*, é um programa de *Ray Tracing* disponível para variadas plataformas de computação. Foi originalmente baseado no *DKBTrace*

(PLACHETKA, 1998). O *POV-Ray* é livre, com seu código-fonte disponível e é mostrado na Figura 43.

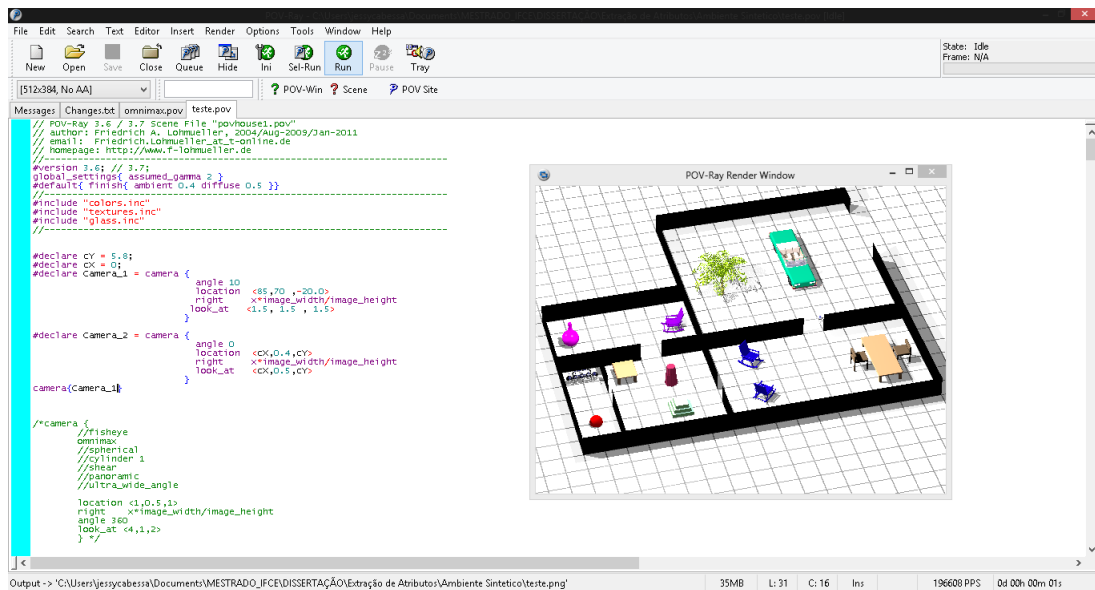


Figura 43: Plataforma *POV-Ray*.

O ambiente construído utilizando o *POV-Ray* pode ser caracterizado como um ambiente *indoor* composto por cinco espaços distintos em formas e cores. A escolha dos objetos que compõem a cena da Figura 44(a) foi realizada de maneira aleatória. Na Figura 44(b), uma vista omnidirecional em um ponto de um dos espaços do ambiente emulado é mostrada. O *POV-Ray* não possui a funcionalidade de criar uma vista obtida a partir de um sistema catadióptrico. Assim, para a realização deste trabalho, um arranjo de objetos dentro do ambiente de simulação foi realizado. A vista omnidirecional foi obtida através da emulação de uma câmera comum apontada para uma esfera com superfície reflexiva.

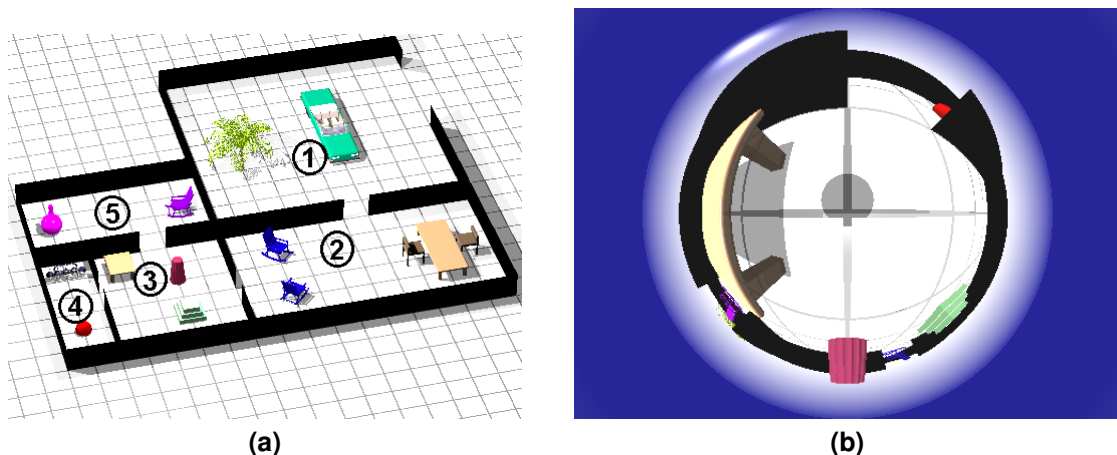


Figura 44: Ambiente sintético: (a) vista superior do ambiente e (b) vista omnidirecional.

O banco de imagens obtidos através da emulação desse ambiente é composto por 132 imagens subdivididas em 5 classes (compartimentos) conforme a Tabela 5. Vale ressaltar, que as 132 imagens foram obtidas a partir de 66 imagens diferentes que sofreram mudança na orientação apenas. Esse banco de imagens é uma das contribuições desse trabalho e pode ser visto em (BESSA, 2015b). O código que gera o ambiente sintético e as vistas omnidirecionais pode ser consultado em (BESSA, 2015c).

Tabela 5: Grupos de imagens do ambiente sintético.

Espaço	Quantidade de imagens
1	54
2	32
3	22
4	6
5	18

5.2.3 Imagens do ambiente real

O ambiente real utilizado nesse trabalho é uma residência com seis cômodos, não estruturada (sem *landmarks*) e dividida em nove espaços diferentes conforme a Figura 45 (a). O ambiente também pode ser caracterizado como *indoor*. Na Figura 45(b), a vista omnidirecional de uma das imagens capturada no compartimento 3 é exibida. Nota-se que todos os detalhes do espaço servirão como características para posterior classificação do cômodo.

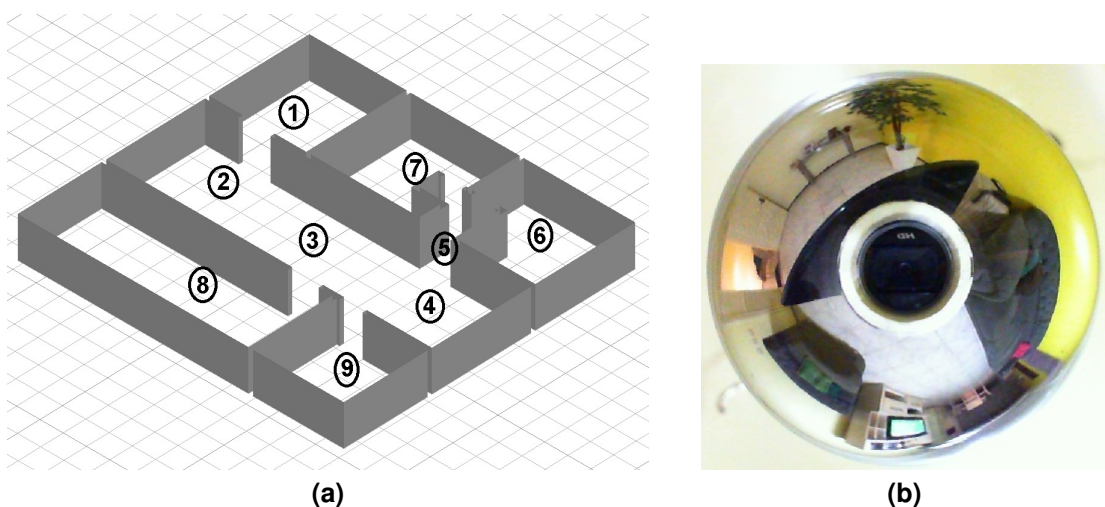


Figura 45: Ambiente real: (a) vista superior do ambiente e (b) vista omnidirecional.

A captura das imagens foi realizada de forma automática através da locomoção do robô da Figura 41(a). Esta foi realizada através de um aplicativo portátil utilizado para

esse fim mostrado na Figura 46(a), onde pode-se controlar movimentos de avanço, de recuo e laterais. O *software* de navegação desenvolvido por Almeida, Nascimento e Filho (2014), mostrado na Figura 46(b), auxiliou a captura das imagens de maneira intermitente (de 5 em 5 segundos).

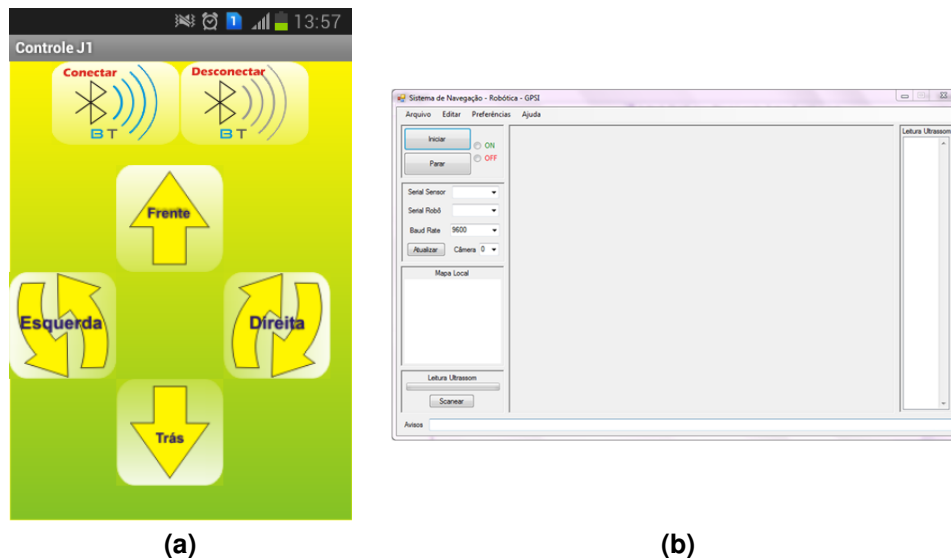


Figura 46: Sistema de navegação: (a) aplicativo portátil e (b) *software* de navegação.

A câmera utilizada na montagem do sistema catadióptrico em questão é a *LifeCam Cinema HD* da *Microsoft* mostrada na Figura 47(a). Esta é apontada para um espelho esférico usinado em aço e banhado em cromo que possui a geometria conforme a Figura 47(b). A escolha da câmera e das dimensões do espelho foram baseadas no trabalho de Grassi *et al.* (2001). O espelho foi desenhado no *software* AutoCAD, que é um *software* para projeto auxiliado por computador (*computer aided design*, CAD) criado e comercializado pela Autodesk, Inc. desde 1982. É utilizado principalmente para a elaboração de peças de desenho técnico em duas dimensões (2D) e para criação de modelos tridimensionais (3D). As dimensões dos desenhos são mostradas no Apêndice B.



Figura 47: Sistema catadióptrico: (a) câmera utilizada e (b) geometria do espelho usinado.

O banco de imagens formado para localização do robô nesse ambiente é composto por 678 imagens, divididas nos nove espaços já mencionados conforme a Tabela 6. O banco de imagens formado a partir do ambiente real é uma contribuição importante desse trabalho e pode ser visualizado em (BESSA, 2015a).

Tabela 6: Grupos de imagens do ambiente real.

Espaço	Quantidade de imagens
1	94
2	48
3	34
4	48
5	38
6	54
7	130
8	58
9	174

5.3 Extração de características

Todos os métodos descritos no Capítulo 4 foram utilizados para extrair características dos três bancos de imagens descritos na seção anterior. Estes são:

- RGB;
- HOG;
- momentos Hu;
- PCA; e
- GIST.

No entanto, o uso de alguns extratores de características só foi possível com a formação de uma imagem panorâmica a partir das imagens omnidirecionais clássicas. Essa transformação é realizada conforme Grassi *et al.* (2001). Na Figura 48, a imagem omnidirecional e sua respectiva formação panorâmica é exibida.

Outro fato, porém, é que as formulações clássicas dos extratores de características fornecerem como resultado descritores que não são invariáveis às mudanças da orientação do robô (rotação). Para superar esta limitação, este trabalho faz uso de formulações propostas pelo autor que serão descritas posteriormente, juntamente com todos os parâmetros utilizados para obtenção dos resultados.

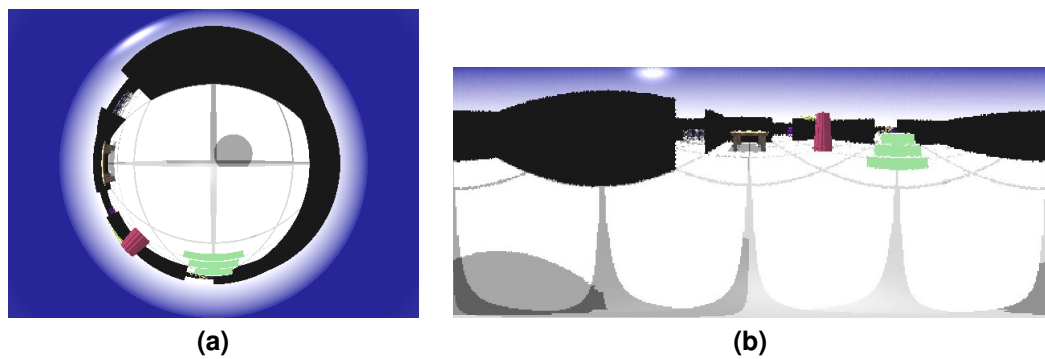


Figura 48: Imagens do ambiente sintético: (a) omnidirecional e (b) panorâmica.

Histograma RGB

Como já dito anteriormente, ao se observar o histograma de uma imagem, tem-se uma noção sobre as características da mesma. Nas imagens utilizadas neste trabalho, os histogramas são de imagens coloridas, ou seja, tem-se três histogramas das bandas vermelha, verde e azul.

Após o cálculo dos três histogramas da imagem, o cálculo de um centróide de cada histograma é realizado. O cálculo do centróide é descrito pela Figura 49 e realizado através dos passos abaixo:

1. Calcula-se os histograma das bandas vermelha, verde e azul.
2. Os histogramas são, então, vetorizados.
3. A mediana de cada vetor é calculada, resultando em três centróides que representam cada banda.



Figura 49: Cálculo do centróide de uma banda qualquer de uma imagem colorida.

Histograma de gradientes orientados (HOG)

A implementação básica do histograma de gradientes orientados (HOG) consiste em dividir a imagem em pequenas regiões conectadas, nomeadas células, e a

construção de um histograma de orientações de gradiente para cada célula conforme dito anteriormente. Porém, esse vetor não é invariante a rotação do robô. Para conseguir aplicar esse método à imagens omnidirecionais, algumas adaptações são necessárias; estas são descritas abaixo.

1. O cálculo do gradiente de cada cena é realizado. Realiza-se a convolução com as máscaras $D_x = [-101]$ e $D_y = [-101]^T$ para extrair as componentes horizontal e vertical $i_x = i * D_x$ e $i_y = i * D_y$, respectivamente. Assim, pode-se calcular a magnitude do gradiente $|G| = \sqrt{i_x^2 + i_y^2}$ e sua orientação $\theta = \arctan \frac{i_y}{i_x}$. $|G|$, como θ são matrizes $M \times N$.
2. A imagem é dividida em células e o histograma de cada célula é calculado. Decide-se trabalhar com oito barras do histograma distribuídos uniformemente entre 0 e 180 graus. Cada *pixel* na célula contribui para uma barra do histograma que contém a orientação daquele *pixel* com um fator de ponderação igual à magnitude do gradiente do *pixel* α .
3. Neste ponto, introduz-se uma mudança para o descritor HOG clássico. A cena panorâmica é dividida em cinco células horizontais, para obter-se um descritor h invariante a rotação do robô conforme a Figura 50.

Os momentos invariantes de Hu

Os momentos invariantes de HU são frequentemente utilizados como recursos para processamento de imagem, principalmente, no que diz respeito a reconhecimento de formas e classificação. Os momentos invariantes foram utilizados na sua descrição clássica conforme descritos no Capítulo 4. Na Figura 51, o descritor formado pelos momentos invariantes é exibido.

Análise de componentes principais (PCA)

O descritor PCA utilizado nesta dissertação é descrito no trabalho de Jogan e Leonardis (2000) . Esse modelo utiliza propriedades específicas de imagens panorâmicas e pode ser visto nos tópicos abaixo:

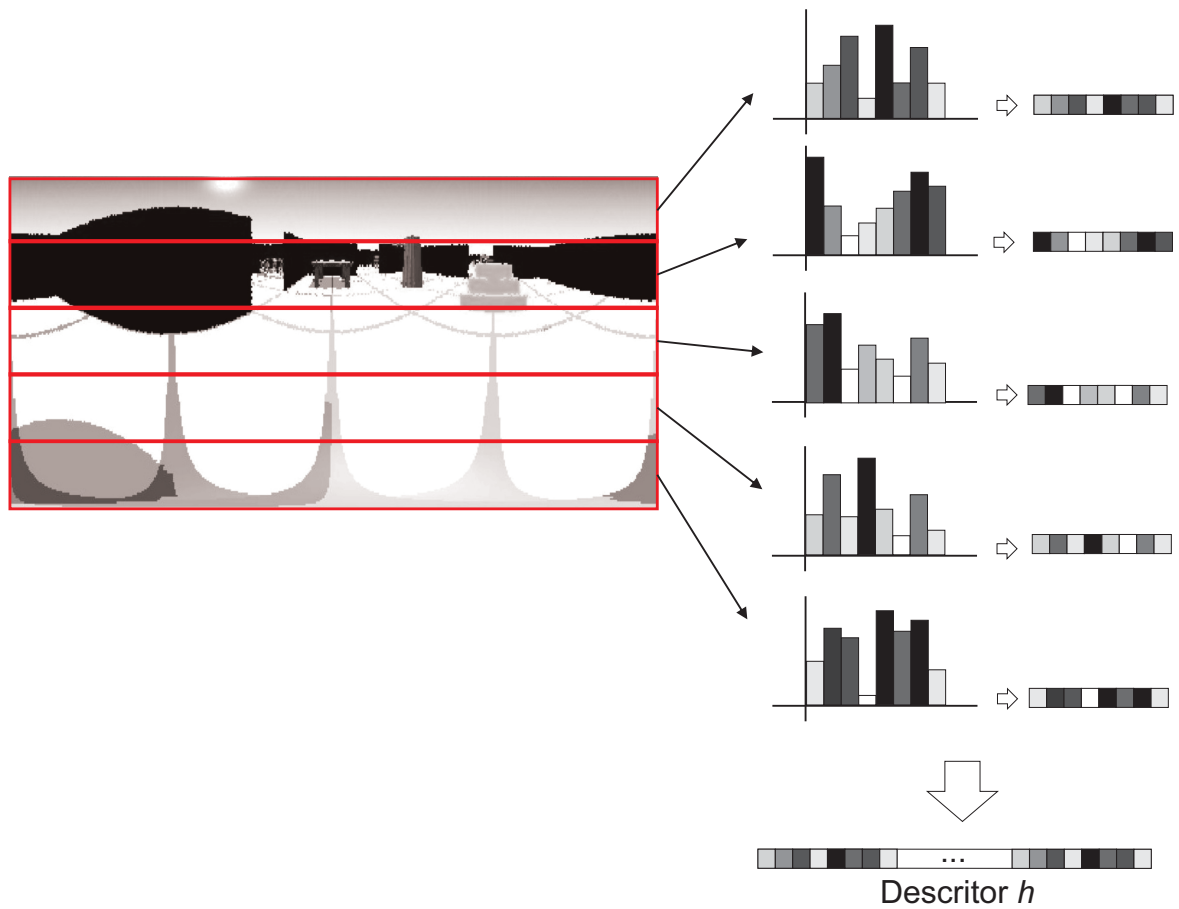


Figura 50: Distribuição das células horizontais em uma imagem panorâmica para construir um descritor invariante a rotação do robô.

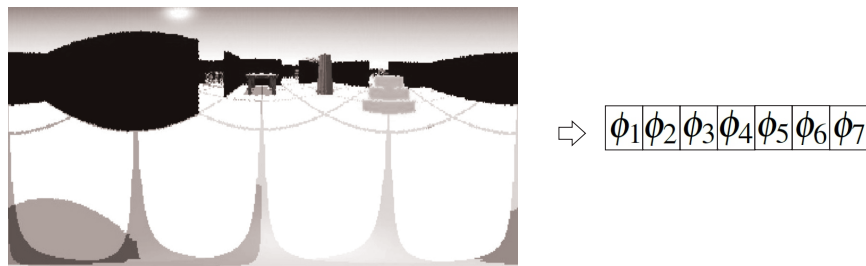


Figura 51: Decritor formado pelos momentos invariantes.

1. Cria-se um conjunto de e imagens transladadas da imagem panorâmica original. Então, tem-se e vetores de dados da imagem original. E as une em uma única imagem.
2. A partir dessa imagem composta, calcula-se a matriz de covariância e, em seguida, obtém-se, através da decomposição em valores singulares, os autovalores não-nulos e a matriz de autovetores, semelhante a formulação clássica.

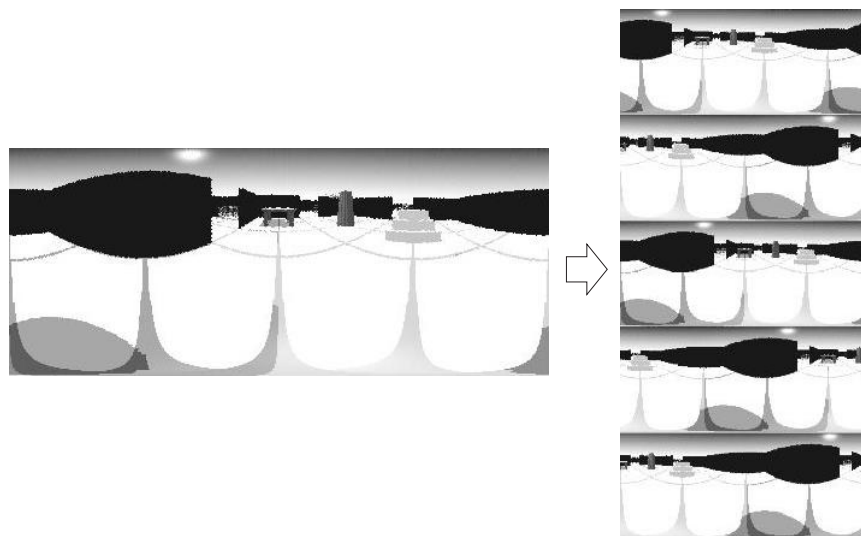


Figura 52: Imagens rotacionadas em cinco subespaços.

3. Considera-se apenas os C autovetores com maior autovalores relacionados. Neste trabalho, procura-se apenas o autovetor com maior autovalor equivalente. Assim, obtém-se um vetor de descritores.

Essência de cena

Conforme o Capítulo 4, o descritor GIST tenta imitar a capacidade humana de reconhecer imediatamente uma cena por meio de algumas regiões que tem uma cor e/ou uma textura em destaque no que diz respeito ao resto da imagem. O descritor GIST implementado neste trabalho é calculado a partir de uma imagem panorâmica conforme os passos a seguir:

1. Deve-se construir uma pirâmide de imagens com o objetivo de descrever as propriedades de imagem em diferentes escalas e entre escalas. O primeiro nível é a imagem original. Para obter cada novo nível, aplica-se um filtro passa-baixa (Gaussiano) e a imagem é sub-amostrada para se obter uma nova imagem com tamanho $0,5 \cdot M \times 0,5 \cdot N$.
2. Para incluir informação de orientação, cada nível da pirâmide é filtrado com um banco de filtros de Gabor (WELDON; HIGGINS; DUNN, 1996) com k_4 orientações distribuídas uniformemente entre 0 e 180 graus. Como resultado, tem-se k_4 matrizes por nível da pirâmide com informações sobre os orientações analisadas. Aplicou-se para as duas primeiras imagens da pirâmide então, obteve-se $2 \cdot k_4$ matrizes conforme a Figura 53.

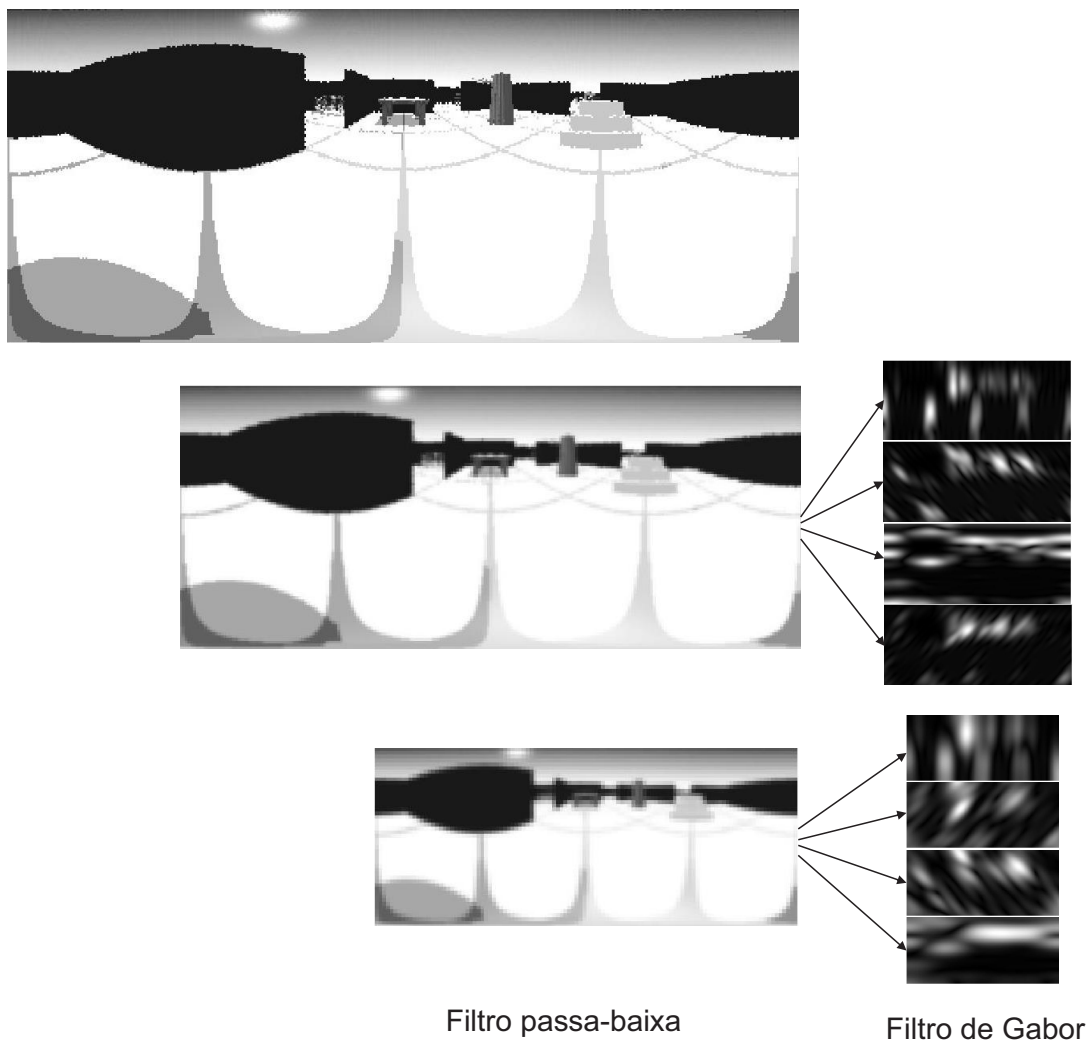


Figura 53: Exemplo de imagem filtrada com filtro de Gabor com $k_4 = 4$ e orientações em 0, 45, 90 e 135 graus em duas escalas.

3. Para reduzir a quantidade de informação, agrupa-se os *pixels* de cada matriz resultante em blocos para o cálculo do valor médio de intensidade dos *pixels* em cada bloco. Decidiu-se fazer a divisão em cinco células de uma forma semelhante à do HOG e, assim, computa-se um descritor com as células horizontais. Esse processo fornece um descritor invariante a rotação do robô conforme a Figura 54.
4. Os parâmetros configuráveis do descritor são: (a) o número de máscaras de Gabor, $k_4 = 4$ e (b) o número de células horizontais $h = 5$.

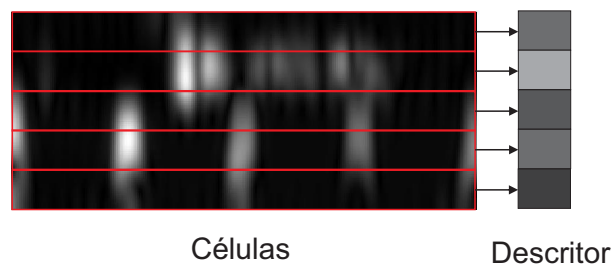


Figura 54: Formação do descritor GIST.

5.4 Classificadores

Nesse trabalho, cinco classificadores de padrões são implementados e testados. Os experimentos com todos os algoritmos foram realizados com diferentes proporções de dados para treinamento-teste (*hold-out*), entre essas: 60 – 40%, 70 – 30%, 80 – 20% e 90 – 10%. Como dito anteriormente, os classificadores avaliados neste trabalho são:

1. MLP;
2. SOM;
3. KNN;
4. LDA; e
5. QDA.

Abaixo, a descrição de como os algoritmos acima citados foram avaliados.

MLP (*MultiLayer Perceptron*)

Neste trabalho, um dos classificadores utilizados para efetuar a localização do robô é uma rede perceptron multicamada (MLP). Para todos os grupos de imagens utilizados, as diferentes proporções de dados para treinamento-teste já mencionadas foram testadas.

A taxa de aprendizagem (η) utilizada é de 0,5 (decrecendo exponencialmente a cada época) com condição de parada por número de épocas de treinamento.

Quanto à topologia da RNA, a quantidade de entradas é dependente do extrator de características utilizado, sabido que cada extrator tem como resultado um descritor

de tamanho diferente. Cada componente do descritor representa uma entrada da RNA conforme a Tabela 7. Por exemplo, caso o extrator de característica utilizado seja os momentos invariantes, a RNA terá 7 entradas.

Tabela 7: Quantidade de entradas para cada extrator de características.

Grupo de imagens	Extrator de característica	Quantidade de entrada
Imagens Sintéticas	Histograma RGB	3
	Histograma de gradientes orientados (HOG)	40
	Os momentos invariantes de Hu	7
	Descritor GIST	40
	Análise de componentes principais (PCA)	960
Imagens do ambiente sintético	Histograma RGB	3
	Histograma de gradientes orientados (HOG)	40
	Os momentos invariantes de Hu	7
	Descritor GIST	40
	Análise de componentes principais (PCA)	960
Imagens do ambiente real	Histograma RGB	3
	Histograma de gradientes orientados (HOG)	40
	Os momentos invariantes de Hu	7
	Descritor GIST	40
	Análise de componentes principais (PCA)	980

Quanto à camada de saída, a quantidade de neurônios depende do ambiente (quantidade de compartimentos) que está sendo investigado conforme a Tabela 8.

Tabela 8: Grupos de imagens do ambiente sintético.

Grupo de imagens	Quantidade neurônios na camada de saída
Imagens sintéticas	5
Imagens do ambiente sintético	5
Imagens do ambiente real	9

A determinação da quantidade de neurônios na camada oculta foi realizada de maneira empírica. Porém, a quantidade de neurônios da camada oculta tiveram como base a Equação descrita por Yin, Liu e Han (2005):

$$HN = \max(M_1, M_2) \times \frac{2}{3}, \quad (5.1)$$

em que HN é a quantidade de neurônios a ser utilizada na camada escondida, M_1 é a quantidade de entradas e M_2 é a quantidade de neurônios na saída da RNA. Variações em torno desse valor foram realizadas e são mostradas no Capítulo 6.

Vale ressaltar que para cada topologia descrita, o treinamento é repetido por 100 vezes (realizações) para escolha da melhor configuração.

SOM (*Self Organizing Map*)

Quando o mapeamento da rede SOM é feito para um espaço unidimensional. O classificador SOM é projetado com o número de neurônios de acordo com o número de espaços a ser reconhecidos, pois a rede auto-organizável foi utilizada para localização do robô. A cada neurônio é atribuído as possíveis classificações (espaços) no ambiente. Como entrada para a rede neural, os descritores obtidos na extração de características da imagem são utilizados.

O algoritmo de treinamento de Kohonen (1990) foi utilizado para treinar a rede e o número de iterações é calculado em 100 vezes o valor do tamanho do conjunto de treinamento. O valor da taxa de aprendizagem no treinamento da rede SOM decai exponencialmente com o tempo, iniciando com um valor igual a 0,5 e chegando ao final com valor em torno de 0,01.

Como dito anteriormente, a função de vizinhança funciona como uma espécie de janela de ponderação, fazendo com que os neurônios mais próximos do neurônio vencedor atual, tenham seus vetores de pesos atualizados mais intensamente do que aqueles neurônios que estão mais distantes do neurônio vencedor. Por questões de convergência e estabilização do aprendizado, a função de vizinhança deve decrescer no tempo, ou seja, o raio de influência ($\vartheta(t)$) decai com o decorrer do treinamento.

Vale ressaltar que todas as configurações de *hold-out* mencionadas anteriormente foram testadas também para rede SOM e repetidas por 100 vezes (realizações).

K-vizinhos mais próximos

Como já dito anteriormente, o K-vizinhos mais próximos é um classificador que faz previsões de acordo com os rótulos de K vizinhos mais próximos da instância de teste. O projeto do classificador baseia-se na distância euclidiana como métrica, e varia-se o parâmetro K entre 4 e 10, assim obtêm-se os melhores resultados de generalização do classificador (ZOUHAL; DENOEU, 1998). Nota-se que para esse algoritmo de classificação, o único parâmetro configurável é K . As taxas médias de acerto e desvio padrão são obtidas com base em 100 realizações.

Discriminante Linear (LDA - *Linear Discriminant Analysis*)

A Análise Discriminante Linear diz respeito a separação das regiões distintas, correspondente a cada classe, através de fronteiras lineares. O parâmetro (média) desse classificador depende inteiramente dos dados de entrada. Todas as configurações de *hold-out* foram testadas. O treinamento é repetido por 100 vezes (realizações).

Discriminante Quadrático (QDA - *Quadratic Discriminant Analysis*)

Como já dito anteriormente, a Análise Discriminante Quadrática diz respeito a separação das regiões distintas no espaço multiespectral, correspondente a cada classe, através de fronteiras quadráticas. Os parâmetros (média, matriz de covariância e probabilidade *a priori*) desse classificador dependem inteiramente dos dados de entrada. As configurações de *hold-out* mencionadas, assim como em todos os classificadores já descritos, foram testadas também para esse classificador. O treinamento é repetido por 100 vezes (realizações).

5.5 Testes com imagens em diferentes níveis de iluminação e de ruído

Com o intuito de testar a robustez do treinamento dos algoritmos de classificação, novas imagens, com diferentes níveis de iluminação (imagens em diferentes horários do dia), foram capturadas, conforme Figura 55 (a). Quanto às imagens do *POV-Ray*, o parâmetro iluminação foi modificado com o intuito de obter novas imagens com diferentes níveis de iluminação. Como forma de dimensionar a diferença de iluminação entre as imagens, pode-se observar o histograma de uma imagem com iluminação original e de outra imagem com iluminação alterada conforme a Figura 56. Nota-se que existe uma compressão ou expansão do histograma, no que se refere aos níveis de cinza, com a variação da iluminação.

Algumas destas imagens também foram expostas a ruído *speckle* com variância igual a 0,02, 0,05 e 0,10 conforme Figura 55 (b). O ruído *speckle* é um ruído multiplicativo proporcional à intensidade do sinal recebido (BESSA *et al.*, 2015). Como já dito anteriormente, o efeito visual desse ruído proporciona uma textura granulosa que pode dificultar a interpretação das imagens, reduzindo o desempenho

dos classificadores de padrões. Apenas imagens do ambiente sintético e imagens do ambiente real foram utilizadas para os testes nesta seção. As imagens com diferentes níveis de iluminação e de ruído são testadas juntas. As quantidades de imagens incluídas neste teste são mostradas na Tabela 9.



Figura 55: Novas imagens de teste: (a) imagem com iluminação alterada e (b) imagem com iluminação alterada e submetida a ruído com variância igual a 0,02.

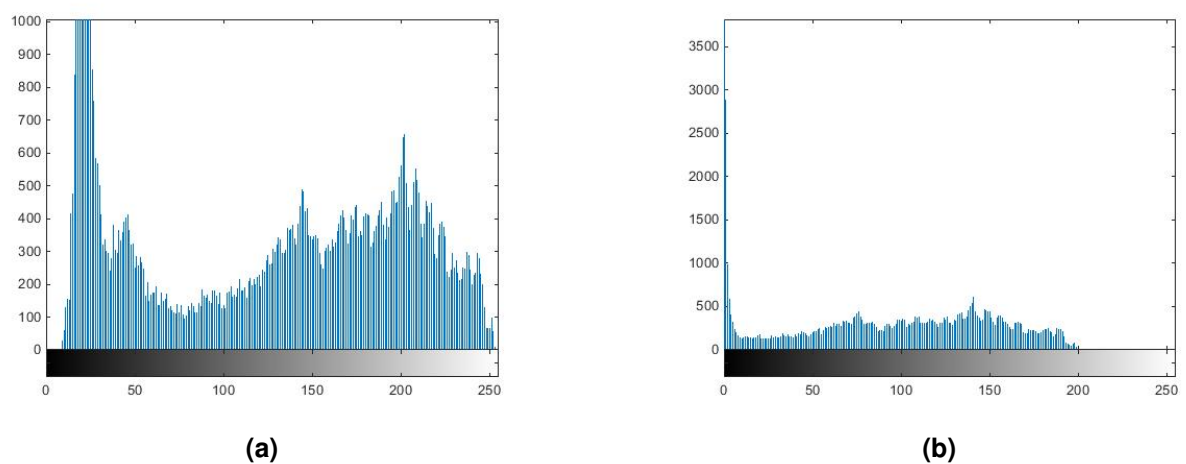


Figura 56: Histograma das novas imagens de teste: (a) iluminação original e (b) iluminação alterada.

Tabela 9: Quantidade de imagens utilizadas nos testes com diferentes níveis de iluminação e ruído.

Grupo de imagens	Quantidade de novas imagens
Imagens do ambiente sintético	42
Imagens do ambiente real	97

5.6 Métricas de avaliação

As métricas utilizadas para avaliar os algoritmos de classificação implementados neste trabalho são:

- taxa média de acerto (acurácia);
- desvio padrão dos resultados.

O desempenho dos classificadores foi avaliado a partir das suas séries de acertos de classificação utilizando taxas médias de acerto e desvio padrão. A taxa média de acerto é a principal métrica utilizada para comparação dos classificadores e da qualidade de descrição de cena por parte dos extratores. Essa taxa mensura a percentagem das predições nas quais os padrões desejados e padrões que os classificadores retornam coincidem. Uma média das taxas de acerto nos testes de todas as realizações é calculada e dada percentualmente por

$$\overline{acc}(\%) = \frac{1}{R} \sum_{i=1}^R acc_i \times 100, \quad (5.2)$$

em que $\overline{acc}$ é a média dada percentualmente, R é o número de realizações e acc_i é a taxa de acerto de cada realização.

Como forma de tratar os valores extremos (máximos e mínimos) de taxa de acerto, usa-se o conceito de média aparada. A média aparada é calculada excluindo-se valores extremos no cálculo da média, ou seja, uma dada percentagem dos valores mais baixos e mais altos e, então, calcula-se a média com os valores restantes. Neste trabalho, exclui-se os 2,5% dos valores inferiores e superiores de taxa de acerto e usa-se os remanescentes, ou seja, uma média aparada de 5%. A vantagem dessa metodologia é que a média aparada não é afetada por valores discrepantes (*outliers*) como a média aritmética simples, ou seja, sem julgamentos tendenciosos.

O desvio padrão mede dispersão estatística dos resultados de taxas de acerto em todas as realizações, indicando “o quão longe”, em geral, os valores se encontram da média dos valores (na mesma ordem de grandeza) e é dado por

$$\sigma = \sqrt{\frac{1}{R} \sum_{i=1}^R (acc_i - \overline{acc})^2}, \quad (5.3)$$

em que σ é o desvio padrão das taxas de acerto.

5.7 Conclusão do Capítulo

A metodologia apresentada neste Capítulo é representada, resumidamente, pela Figura 57. O fluxo de informação deste trabalho segue os seguintes itens:

- aquisição de imagens;
- extração de características dos três grupos de imagens visando a formações de descritores;
- classificação de padrões visando a localização do robô móvel em um mapa topológico pré-existente;
- avaliação da classificação a partir das métricas já mencionadas.

Assim, no próximo Capítulo, os resultados obtidos com as respectivas discussões são apresentados.

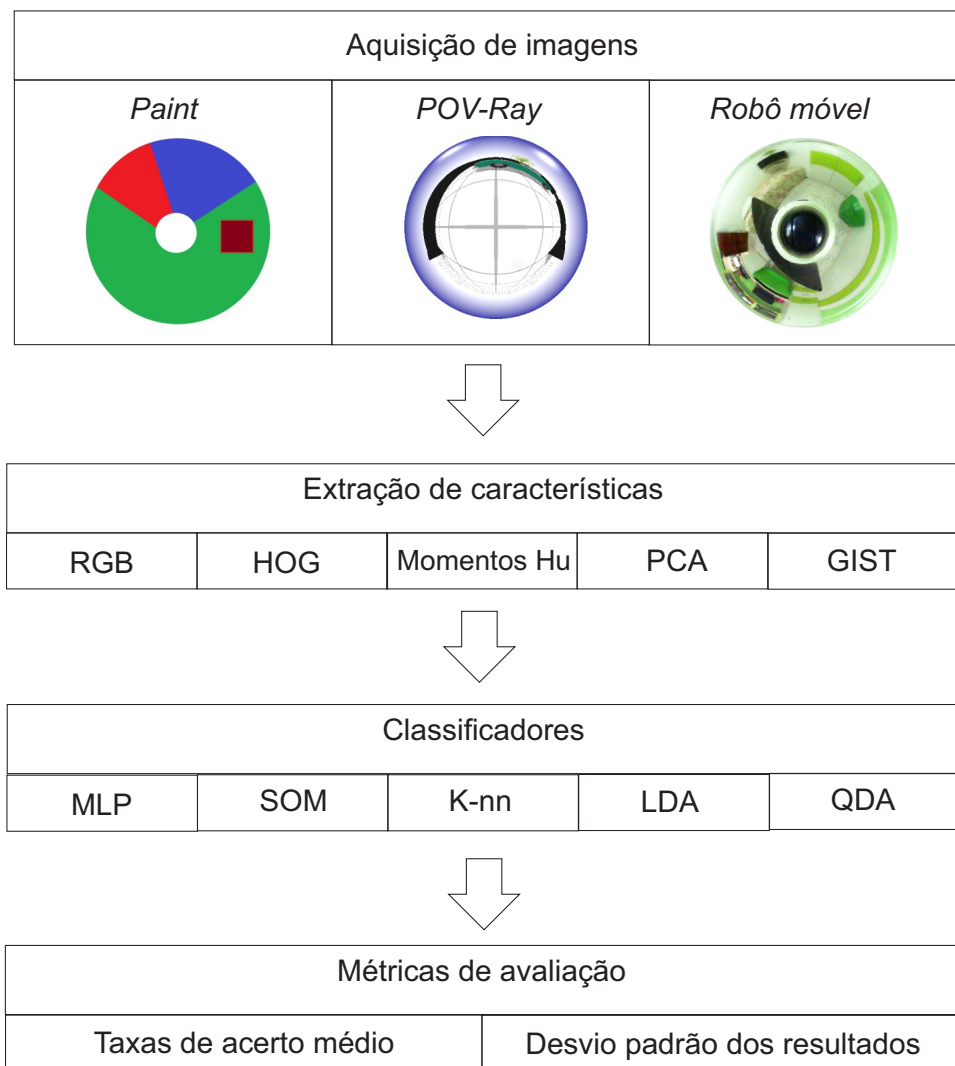


Figura 57: Representação da metodologia utilizada neste trabalho.

RESULTADOS E DISCUSSÕES

Neste Capítulo são apresentados e discutidos os resultados obtidos com imagens sintéticas e reais. As imagens sintéticas são utilizadas com o intuito de verificar a viabilidades dos métodos. Já as imagens reais são utilizadas no sentido de verificar o comportamento dos classificadores perante aplicações de Visão Computacional mais desafiadoras. Todos os métodos de extração de características e de classificação são comparados entre si.

6.1 Resultados para os métodos de extração de características

Todos os métodos de extração de características são implementados de forma a conhecer o tempo de execução de cada algoritmo, levando em consideração o tipo de imagem a qual está se aplicando o método e a necessidade ou não de obter a vista panorâmica da imagem omnidirecional em questão. Os diversos métodos são aplicados a todo o grupo de imagens, sendo este processo repetido 10 vezes e uma média dos tempos é calculada. Os métodos são implementados e as simulações realizadas utilizando o ambiente de desenvolvimento e simulação MATLAB, versão R2014b. Os tempos de execução são mostrados na Tabela 10.

Nota-se que para todos os algoritmos, o tempo de execução para as imagens do ambiente real é maior, explicado pelo fato das imagens desse ambiente ter maiores dimensões e uma quantidade maior de informação. Nota-se que o método mais aplicável em tempo real, ao se levar em consideração apenas o tempo de processamento, é o algoritmo dos momentos invariantes de Hu; a velocidade na execução é explicado pelo fato do algoritmo ser composto apenas por cálculos com operações básicas. Já o algoritmo mais lento para todos os grupos de imagens é o PCA, onde o tempo de execução grande pode ser explicado pelo cálculo de matrizes inversas. Porém, para um confronto mais efetivo, tem-se a necessidade de quantificar

Tabela 10: Tempo médio de execução dos algoritmos de extração de atributos.

Método de extração	Grupo de imagens	Tempo de Execução (s)
Imagens Sintéticas	RGB	0,0421
	HOG	0,0472
	Momentos de Hu	0,0109
	PCA	0,3214
	GIST	0,1851
Ambiente Sintético	RGB	0,0443
	HOG	0,0497
	Momentos de Hu	0,0112
	PCA	0,3392
	GIST	0,1972
Ambiente Real	RGB	0,0591
	HOG	0,0640
	Momentos de Hu	0,0123
	PCA	0,3399
	GIST	0,2106

a qualidade de descrição da cena de todos os métodos.

6.2 Resultados para os classificadores de padrões

Nas simulações computacionais realizadas para avaliação dos classificadores em conjuntos com os extratores de atributos utiliza-se medidas de taxa de acerto e desvio padrão para diferentes configurações. Os valores máximo (max) e mínimo (min) das taxas de acerto no teste de cada classificador também são apresentados. Sabido o classificador, é necessário uma verificação do quanto este consegue generalizar e resolver o problema proposto para dados não-apresentados (ou seja, pertencentes ao conjunto de teste) e, por isso, não participam do treinamento.

6.2.1 Classificador MLP

A fim de se obter dados para uma análise quantitativa, informações referentes a fase de treinamento e a fase de teste são armazenados: melhor configuração conforme a Tabela 11 (levando em consideração a metodologia *hold-out* 80%-20%, pois este apresentou menor desvio padrão nas taxas de acerto para o conjunto de teste), taxa média de acerto para o conjunto de treinamento, taxa média de acerto para o conjunto de teste e desvio padrão das taxas de acerto no teste conforme as Tabelas 12, 13 e 14. A média de todos os valores tem como base 100 realizações.

Como já foi dito, as Tabelas 12, 13 e 14 apresentam os resultados para a metodologia *hold-out* com 80% do conjunto de dados para treinamento e 20% para

Tabela 11: Melhor configuração para cada extrator de características × grupo de imagens.

Extrator de característica	Grupo de imagens	Nº de épocas de treinamento	Nº de neurônios camada oculta
Histograma RGB	Imagens sintéticas	200	4
	Ambiente sintético	1000	8
	Ambiente real	1000	10
HOG	Imagens sintéticas	200	4
	Ambiente sintético	1000	10
	Ambiente real	1000	12
Momentos de Hu	Imagens sintéticas	200	6
	Ambiente sintético	1000	10
	Ambiente real	1000	10
PCA	Imagens sintéticas	200	5
	Ambiente sintético	1000	10
	Ambiente real	1000	10
GIST	Imagens sintéticas	200	5
	Ambiente sintético	1000	7
	Ambiente real	1000	12

Tabela 12: Resultados de treinamento e teste da MLP para imagens sintéticas com 80% dos dados para treinamento e 20% para teste.

Extrator de característica	Acerto médio - treinamento (%)	Acerto médio - teste (%)	Desvio padrão
Histograma RGB	99,3	100,0	0,0
HOG	99,5	100,0	0,0
Momentos de Hu	98,6	100,0	0,0
PCA	99,4	100,0	0,0
GIST	99,7	100,0	0,0

Tabela 13: Resultados de treinamento e teste da MLP para imagens de ambiente sintético com 80% dos dados para treinamento e 20% para teste.

Extrator de característica	Acerto médio - treinamento (%)	Acerto médio - teste (%)	Desvio padrão
Histograma RGB	76,8	84,2	2,9
HOG	97,7	92,1	1,2
Momentos de Hu	76,0	77,2	2,0
PCA	98,7	95,1	1,1
GIST	99,6	98,0	1,7

Tabela 14: Resultados de treinamento e teste da MLP para imagens de ambiente real com 80% dos dados para treinamento e 20% para teste.

Extrator de característica	Acerto médio - treinamento (%)	Acerto médio - teste (%)	Desvio padrão
Histograma RGB	92,7	92,4	0,8
HOG	95,2	89,6	0,8
Momentos de Hu	83,6	89,0	0,6
PCA	99,7	99,3	0,3
GIST	99,5	99,7	0,2

teste. A quantidade de épocas e neurônios da camada oculta apresentados na Tabela 11 foram escolhidos através de comparação de resultados, buscando a maior taxa média de acerto com menor desvio padrão no conjunto de teste.

Observando os resultados para as imagens sintéticas (veja Tabela 12), observa-se que todos os extratores promovem uma boa descrição de cena, levando o classificador

a uma taxa de acerto média próxima a 100%, fato esse devido à facilidade de extração de atributos que representem bem as imagens.

Quanto aos resultados das imagens do ambiente sintético, descritos na Tabela 13), observa-se que os extratores RGB e Momentos Hu alcançam taxas médias de acerto menores que os outros três (HOG, GIST e PCA) que apresentaram taxas médias de acerto superiores a 90%. Uma possível explicação para isso é o fato desses dois extratores formarem descritores pequenos, provavelmente, sem uma qualidade de descrição de cena. A maior taxa de acerto média para esse grupo de imagens é utilizando descritor GIST (98,08%), ou seja, pode-se reafirmar que esse descritor tenta imitar a capacidade humana de reconhecer imediatamente uma cena por meio de algumas regiões que tem uma cor e/ou uma textura em destaque. Percebe-se, também, que o desvio padrão dos resultados cresce com o decaimento da quantidade de dados para o conjunto de teste, e isso pode ser explicado pela quantidade de imagens do espaço 4 do ambiente sintético, conforme a Tabela 13.

Os resultados das imagens do ambiente real (Tabela 14) são semelhantes aos das imagens do ambiente sintético, com uma única diferença: o extrator RGB gerou uma taxa de acerto média acima de 90%. Esse fato pode ser explicado pela grande diferença de cores entre os espaços do ambiente real, favorecendo a descrição de cena por parte desse descritor. Nota-se, mais uma vez, que o descritor Momentos Hu é desfavorecido devido a sua pequena quantidade de informação. Os extratores GIST e PCA tem taxas médias de acerto próximas a 100%, fortalecendo a ideia de uma boa descrição de cena.

6.2.2 Classificador SOM

A respeito da rede SOM, para obtenção de uma análise quantitativa, informações como taxa média de acerto para o conjunto de teste e desvio padrão das taxas de acerto no teste conforme a Tabela 15 são guardadas. A média de todos os valores tem como base 100 realizações levando em consideração o *hold-out* 80%-20%.

Os melhores resultados no que se refere a descrição de cena se repetem. Os descritores propostos GIST e PCA se destacam diante dos outros descritores. A quantidade de informação contida nesses dois descritores possibilitam uma melhor classificação dos padrões. Um fato que também se repete é que as taxas médias de acerto são maiores para o ambiente real, que pode ser explicado pela maior

quantidade de informação nas imagens no que se refere a formas e texturas.

Tabela 15: Acerto médio no teste e desvio padrão da rede SOM para cada extrator de características com 80% dos dados para treinamento e 20% para teste.

Extrator de característica	Grupo de imagens	Acerto médio - teste (%)	Desvio padrão
Imagens sintéticas	RGB	98,1	0,8
	HOG	98,5	0,8
	Momentos de Hu	96,3	0,9
	PCA	99,6	0,2
	GIST	99,8	0,2
Ambiente sintético	RGB	80,4	2,8
	HOG	91,2	1,7
	Momentos de Hu	79,9	4,2
	PCA	94,9	1,4
	GIST	95,2	1,1
Ambiente real	RGB	91,2	3,4
	HOG	90,1	2,7
	Momentos de Hu	84,2	3,5
	PCA	99,1	0,5
	GIST	99,4	0,5

6.2.3 Classificador KNN

Como já dito anteriormente, o projeto do classificador KNN baseia-se na distância euclidiana como métrica de proximidade. Sabido isso, variou-se o valor de K entre 4 e 10 e na Tabela 16 é mostrado qual o melhor valor de K para cada situação, ou seja, onde obtêm-se os melhores resultados, levando em consideração o *hold-out* 80%-20%, juntamente com as taxas médias de acerto para o conjunto de teste e desvio padrão das taxas de acerto. Vale ressaltar que para o algoritmo KNN, o *hold-out* 80%-20% significa que 20% das amostras vão ser comparadas aos outros 80%, e daí se tira as K amostras mais próximas.

Os melhores resultados dos classificadores baseados em redes neurais se repetem para o classificador estatístico KNN, porém com taxas médias de acerto menores e desvio padrão maiores. Dos resultados, pode-se concluir que o classificador tem baixo desempenho de classificação, principalmente se comparado a outros classificadores apresentados anteriormente. Os descritores GIST e PCA implementados nesse trabalho repetem a sua eficácia quanto a descrição de cena e a possibilidade de reconhecimento de padrões.

Tabela 16: Melhor valor de K, acerto médio no teste e desvio padrão para o algoritmo KNN para cada extrator de características com 80% dos dados para treinamento e 20% para teste.

Extrator de característica	Grupo de imagens	K	Acerto médio - teste (%)	Desvio padrão
Imagens sintéticas	RGB	4	98,2	1,5
	HOG	5	97,4	2,1
	Momentos de Hu	5	95,3	3,8
	PCA	5	98,1	1,7
	GIST	5	99,1	0,4
Ambiente sintético	RGB	6	78,7	5,6
	HOG	5	61,3	5,9
	Momentos de Hu	6	67,1	4,5
	PCA	5	84,1	3,4
	GIST	6	87,21	3,1
Ambiente real	RGB	6	88,2	4,1
	HOG	5	69,1	5,2
	Momentos de Hu	5	88,1	4,9
	PCA	5	96,2	2,5
	GIST	5	97,4	2,1

6.2.4 Classificador LDA

Os resultados exibidos na Tabela 17 tem como base o *hold-out* 80%-20% e as taxas médias de acerto e desvio padrão tem como base 100 realizações. Os resultados presentes nessa tabela mostram que o problema não é linear. As taxas de acerto médio são baixas quando comparadas com o classificadores não-lineares. As taxas de acerto médio para as imagens do ambiente sintético e imagens do ambiente real são inferiores a 90 %.

Tabela 17: Acerto médio no teste e desvio padrão do algoritmo LDA para cada extrator de características com 80% dos dados para treinamento e 20% para teste.

Extrator de característica	Grupo de imagens	Acerto médio - teste (%)	Desvio padrão
Imagens sintéticas	RGB	82,6	3,2
	HOG	76,7	4,1
	Momentos de Hu	70,7	3,9
	PCA	90,1	3,6
	GIST	93,1	3,2
Ambiente sintético	RGB	62,1	4,1
	HOG	45,4	4,8
	Momentos de Hu	43,1	5,0
	PCA	79,9	4,0
	GIST	81,9	3,8
Ambiente real	RGB	69,1	4,0
	HOG	62,0	5,2
	Momentos de Hu	60,3	6,1
	PCA	89,1	3,2
	GIST	91,1	3,0

6.2.5 Classificador QDA

A vantagem do classificador QDA é a inexistência de parâmetros configuráveis. A desvantagem é a necessidade de cálculo de matrizes de covariância, médias e probabilidades a priori, que demandam alto custo computacional. Os resultados exibidos na Tabela 18 tem como base o *hold-out* 80%-20% e as taxas médias de acerto e desvio padrão tem como base 100 realizações, como já dito anteriormente. Os acertos médios não apresentam melhora significativa e apresentam desvio padrão maiores quando comparados a MLP. Vale ressaltar que os resultados reafirmar os descritores GIST e PCA como bons extratores de características.

Tabela 18: Acerto médio no teste e desvio padrão do algoritmo QDA para cada extrator de características com 80% dos dados para treinamento e 20% para teste.

Extrator de característica	Grupo de imagens	Acerto médio - teste (%)	Desvio padrão
Imagens sintéticas	RGB	94,8	0,9
	HOG	93,9	1,0
	Momentos de Hu	98,4	0,8
	PCA	99,5	0,4
	GIST	99,8	0,1
Ambiente sintético	RGB	79,1	3,9
	HOG	89,2	3,1
	Momentos de Hu	75,2	4,9
	PCA	94,5	2,0
	GIST	95,2	1,9
Ambiente real	RGB	90,3	3,1
	HOG	89,8	3,9
	Momentos de Hu	85,9	4,0
	PCA	99,4	0,1
	GIST	99,0	0,8

6.3 Resultados para imagens com diferentes níveis de iluminação e de ruído

Após o treinamento dos classificadores e testes realizados com a metodologia *hold-out* 80%-20%, escolheu-se a melhor configuração para cada classificador com o objetivo de testá-los em um grupo de imagens com diferentes níveis de iluminação e de ruído e que não foram utilizadas para obtenção dos resultados anteriores. Os resultados obtidos são exibidos nas Tabelas 19, 20, 21, 22 e 23

Tabela 19: Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) e utilizando a melhor configuração da rede MLP obtida.

Extrator de característica	Grupo de imagens	Acerto no teste (%)
Ambiente sintético	RGB	76,2
	HOG	88,1
	Momentos de Hu	71,4
	PCA	92,9
	GIST	92,9
Ambiente real	RGB	87,6
	HOG	85,6
	Momentos de Hu	74,2
	PCA	94,8
	GIST	96,9

Tabela 20: Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) utilizando a melhor configuração da rede SOM obtida.

Extrator de característica	Grupo de imagens	Acerto no teste (%)
Ambiente sintético	RGB	73,8
	HOG	85,7
	Momentos de Hu	69,0
	PCA	90,5
	GIST	92,9
Ambiente real	RGB	86,6
	HOG	84,5
	Momentos de Hu	74,2
	PCA	93,8
	GIST	95,9

Tabela 21: Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) utilizando a melhor configuração do algoritmo KNN obtido.

Extrator de característica	Grupo de imagens	Acerto no teste (%)
Ambiente sintético	RGB	69,0
	HOG	80,9
	Momentos de Hu	61,9
	PCA	88,1
	GIST	90,5
Ambiente real	RGB	86,6
	HOG	83,5
	Momentos de Hu	73,2
	PCA	93,8
	GIST	94,8

6.4 Conclusão do Capítulo

Neste capítulo foram apresentados os resultados obtidos para as diversas simulações computacionais realizadas utilizando os extratores de características RGB, HOG, Momentos Hu, PCA e GIST e os classificadores MLP, SOM, KNN, LDA e QDA.

Tabela 22: Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) utilizando a melhor configuração do algoritmo LDA obtido.

Extrator de característica	Grupo de imagens	Acerto no teste (%)
Ambiente sintético	RGB	52,3
	HOG	35,7
	Momentos de Hu	35,7
	PCA	73,8
	GIST	76,2
Ambiente real	RGB	60,8
	HOG	43,3
	Momentos de Hu	47,4
	PCA	83,5
	GIST	86,6

Tabela 23: Acerto no teste para imagens novas com diferentes níveis de iluminação e de ruído (variância igual a 0,02, 0,05 e 0,10) utilizando a melhor configuração do algoritmo QDA obtido.

Extrator de característica	Grupo de imagens	Acerto no teste (%)
Ambiente sintético	RGB	73,8
	HOG	88,1
	Momentos de Hu	69,0
	PCA	92,9
	GIST	92,9
Ambiente real	RGB	87,6
	HOG	84,5
	Momentos de Hu	75,3
	PCA	94,8
	GIST	95,9

CONCLUSÃO E TRABALHOS FUTUROS

Uma premissa fundamental para que um robô móvel consiga executar suas tarefas é que ele possua conhecimento de sua localização. Basta pensar em um ser humano sem a capacidade de saber onde está e se enxerga a necessidade de se localizar. Diante esse problema, essa dissertação propõe um método de localização global de robôs móveis em um mapa topológico pré-existente através de técnicas de reconhecimento de padrões em imagens omnidirecionais. Uma comparação entre cinco métodos de reconhecimento de padrões é realizada. Para isto, uma comparação é efetuada: métodos de extração de características adaptados a esse tipo de imagem (omnidirecional). Os resultados obtidos são aplicados a localização de robôs móveis em ambientes *indoor* não estruturados, ou seja, sem utilização de *landmarks*.

Os descritores de cena adaptados para se trabalhar com esse tipo de imagem que são confrontados nesse trabalho são: Histograma RGB, Histograma de Gradiente Orientados, Momentos invariantes de Hu, Descritor GIST e Análise de Componentes Principais (PCA). Quando comparados em relação a tempo de processamento, os descritores GIST e PCA possuem certa desvantagem em relação aos outros descritores. O descritor GIST implementado neste trabalho possui tempo de processamento superior a 0,18s para todos os grupos de imagens e o descritor PCA, conforme implementado nesta dissertação, é capaz de extrair atributos em tempo superior a 0,3s. Os outros descritores comparados executam suas respectivas tarefas em tempos inferiores a 0,06s.

Porém, quando esses descritores são comparados levando em questão a qualidade de descrição de cena, os descritores GIST e PCA são mais vantajosos e se mostram bastante eficazes na tarefa em questão. Para três dos classificadores confrontados e todos grupos de imagens, estes dois descritores alcançam taxas médias de acerto superiores a 94%. Os resultados divergentes aparecem quando os descritores são utilizados juntamente ao classificador KNN em imagens de ambiente sintético e ao algoritmo LDA em todos os grupos de imagens

Um acontecimento que vale ser lembrado é que as taxas médias de acerto para imagens de ambiente real foram superiores às taxas médias de acerto para imagens de ambiente sintético. Isto pode ser explicado por dois motivos: a pequena quantidade de dados do espaço 4 do ambiente sintético, o que dificulta o aprendizado e a pequena quantidade de informação contida nesse ambiente no que se refere a textura e a formas. Essa pequena quantidade de imagens se deve ao tamanho do compartimento, impossibilitando outras zonas de captura de imagens.

Quanto aos classificadores, inicialmente, pode-se comentar sobre os resultados obtidos para o algoritmo LDA. Esse classificador possui superfícies de decisão lineares e, assim, alcança valores de taxa de acerto médio bem inferiores quando comparado a outros algoritmos de classificação não-lineares. Assim, conclui-se que o problema em questão não será bem resolvido com classificadores lineares. Pode-se, assim, descartar o algoritmo LDA de futuras análises.

Quanto a comparação dos classificadores não-lineares, o estudo comparativo foi realizado com quatro técnicas diferentes (dois baseados em redes neurais artificiais e dois baseados em estatística): MLP, SOM, KNN e QDA. Diante dos resultados apresentados, as taxas médias de acerto da rede MLP são superiores a rede SOM, enquanto rede neural. Os resultados utilizando o descritor GIST (maiores taxas de acerto) são superiores a 98% para todos os grupos de imagens, enquanto os resultados para rede SOM são superiores a 95%. Quando a comparação é realizada com os classificadores estatísticos, o QDA tem resultados mais satisfatórios do que o KNN. Todos os resultados utilizando o descritor GIST e o classificador QDA são superiores a 95% enquanto a taxa de acerto do algoritmo KNN para as imagens de ambiente sintético é 84,1%. O descritor PCA alcança resultados semelhantes com os resultados obtidos através da utilização do descritor GIST.

A inclusão de imagens de teste com diferentes níveis de iluminação e de ruído confirmou a superioridade dos descritores PCA e GIST. Estes se mostram mais imunes, diante dos outros descritores, a esse tipo de mudança. Nota-se que as taxas de acerto médio para o descritor RGB e descritor HOG reduziram de maneira mais significativa quando comparado aos outros descritores. Isto se deve a vulnerabilidade do descritor RGB à mudança de cor na imagem e a sensibilidade do descritor HOG ao acréscimo de ruído.

De uma forma geral, pode-se eleger os descritores GIST e PCA adaptados para imagens omnidirecionais como melhores descritores de cena para o fim em questão.

Dentre os classificadores, a rede MLP e o Discriminante Quadrático obtiveram resultados mais satisfatórios, sugerindo que essas duas técnicas podem ser utilizadas com sucesso na localização de robôs móveis utilizando imagens omnidirecionais.

As principais contribuições deste trabalho são as adaptações realizadas nos extratores de características a fim de se obter a invariância à rotação do robô para as imagens omnidirecionais e a utilização destas para localização de robôs móveis. Notou-se também que utilização de técnicas de reconhecimento de padrões a fim de localizar o robô ainda não é uma área em grande ascensão, acenando, assim, para uma nova linha de pesquisa. Outra grande contribuição dessa dissertação, é a formação de dois bancos de imagens: o primeiro utilizando o ambiente *POV-Ray* composto por 132 imagens omnidirecionais sintéticas e o segundo um banco de imagens omnidirecionais reais. Ambas podem ser utilizadas em outros trabalhos de maneira a comparar outras técnicas de localização de robôs móveis, ou até, implementar técnicas de mapeamento e navegação nesse ambiente.

Como trabalhos futuros, sugerem-se:

- criação do mapa topológico do ambiente a partir dos resultados de classificação obtidos através de redes de Petri;
- avaliação da complexidade dos algoritmos de extração utilizados nesse trabalho;
- avaliação da complexidade dos classificadores utilizados nesse trabalho;
- investigação de aplicações e o desempenho dos métodos utilizados;
- investigação da possibilidade de extração de características das imagens omnidirecionais sem transformá-las em imagens panorâmicas.
- comparação de outros descritores, inclusive descritores locais como: SIFT e SURF, pois estes podem ter resultados satisfatórios na descrição de cena;
- combinação dos descritores para gerar padrões e posterior avaliação de técnicas de Reconhecimento de Padrões;
- comparação de outros métodos de classificação, que possam vir a ser mais eficazes, como máquina de vetores de suporte (*support vector machine*, SVM) e/ou *optimum-path forest* (OPF);

- navegação de robôs móveis através das imagens omnidirecionais atreladas ao algoritmos utilizados simultaneamente a formação do mapa topológico do ambiente;
- investigação do uso de visão omnidirecional estéreo para localização, mapeamento topológico e navegação de robôs móveis.

Referências Bibliográficas

AGUIAR, Ricardo G. **Segmentação de Imagens em cores de Himunohistoquímica**. Tese (Doutorado) — Dissertação de Mestrado. Universidade Federal de São Carlos, 2000.

AHRARY, Alireza; TIAN, Li; KAMATA, Sei-ichiro; ISHIKAWA, Masumi. An autonomous sewer robots navigation based on stereo camera information. In: IEEE. **Tools with Artificial Intelligence, 2005. ICTAI 05. 17th IEEE International Conference on**. 2005. p. 6–pp.

ALEXANDRIA, Auzuir Ripardo de; CORTEZ, Paulo Cesar; FELIX, John Hebert da Silva; OLIVEIRA, Tiberio Menezes de; GIRAO, Anaxagoras Maia; FROTA, Joao Batista Bezerra; ALMEIDA, Jessyca. An ocr system for numerals applied to energy meters. **Latin America Transactions, IEEE (Revista IEEE America Latina)**, IEEE, v. 12, n. 6, p. 957–964, 2014.

ALMEIDA, Jefferson S; NASCIMENTO, Rafael S; FILHO, Pedro P Rebouças. Development of navigation system for an autonomous mobile robot via the construction of a local map. **Journal INNOVER**, v. 1, n. 4, p. 37–47, 2014.

ANDRADE, Beatriz Trinchão; MATOS, Leonardo Nogueira; FREIRE, Eduardo Oliveira. Um sistema de visão computacional para robôs móveis. In: **XXVI Congresso da SBC, Campo Grande, MS**. 2006.

ANTONELO, Eric Aislan; FIGUEIREDO, Mauricio; BAERVELDT, A-J; CALVO, Rodrigo. Intelligent autonomous navigation for mobile robots: spatial concept acquisition and object discrimination. In: **Computational Intelligence in Robotics and Automation, 2005. CIRA 2005. Proceedings. 2005 IEEE International Symposium on**. 2005. p. 553–557.

APPELTANS, Ward; AHYONG, Shane T.; ANDERSON, Gary; WILSON, Simon P. The magnitude of global marine species diversity. **Current Biology**, v. 22, n. 23, p. 2189 – 2202, 2012. ISSN 0960-9822.

ARRAS, Kai Oliver; VESTLI, Sjur J. Hybrid, high-precision localisation for the mail distributing mobile robot system mops. In: IEEE. **Robotics and Automation, 1998. Proceedings. 1998 IEEE International Conference on**. 1998. v. 4, p. 3129–3134.

AYACHE, Nicholas. **Artificial vision for mobile robots: stereo vision and multisensory perception**. : Mit Press, 1991.

AYACHE, Nicholas; FAUGERAS, Olivier D. Building, registrating, and fusing noisy visual maps. **The International Journal of Robotics Research**, Sage Publications, v. 7, n. 6, p. 45–65, 1988.

BAILEY, Tim; NIETO, Juan; GUIVANT, Jose; STEVENS, Michael; NEBOT, Eduardo. Consistency of the EKF-SLAM algorithm. In: IEEE. **Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on**. 2006. p. 3562–3568.

BAKER, Simon; NAYAR, Shree K. A theory of single-viewpoint catadioptric image formation. **International Journal of Computer Vision**, Springer, v. 35, n. 2, p. 175–196, 1999.

BANSAL, Ajay. **Monte Carlo localization for mobile robots**. Tese (Doutorado) — Texas Tech University, 2002.

BARRETO, Guilherme A; ARAUJO, Aluizio FR. Identification and control of dynamical systems using the self-organizing map. **Neural Networks, IEEE Transactions on**, IEEE, v. 15, n. 5, p. 1244–1259, 2004.

BEKEY, George A. **Autonomous robots: from biological inspiration to implementation and control**. : MIT press, 2005.

BEKEY, George A.; YUH, Junku. The status of robotics. **IEEE Robot. Automat. Mag.**, p. 76–81, 2007.

BELKASIM, Saeid O; SHRIDHAR, Malayappan; AHMADI, Majid. Pattern recognition with moment invariants: a comparative study and new results. **Pattern recognition**, Elsevier, v. 24, n. 12, p. 1117–1138, 1991.

BENOSMAN, Ryad; KANG, Sing Bing. **Panoramic Vision: Sensors, Theory, and Applications**. : Springer, 2001. (Monographs in Computer Science).

BESSA, Jessyca Almeida. **Banco de imagens omnidirecionais a partir de um ambiente real**. 2015. Disponível em: <<http://migre.me/pOtvL>>.

BESSA, Jessyca Almeida. **Banco de imagens omnidirecionais a partir de um ambiente sintético**. 2015. Disponível em: <<http://migre.me/pW1UO>>.

BESSA, Jessyca Almeida. **Código para gerar imagens omnidirecionais a partir de um ambiente sintético**. 2015. Disponível em: <<http://migre.me/pOtep>>.

BESSA, Jéssyca Almeida; CORTEZ, Paulo César; FÉLIX, John Hebert da Silva; NETO, Ajalmar Rêgo da Rocha; ALEXANDRIA, Auzuir Ripardo de. Radial snakes: Comparison of segmentation methods in synthetic noisy images. **Expert Systems with Applications**, Elsevier, v. 42, n. 6, p. 3079–3088, 2015.

BEZERRA, Clauber Gomes. **Localização de um robô móvel usando odometria e marcos naturais**. Tese (Doutorado) — Universidade Federal do Rio Grande do Norte, 2004.

BONARINI, Andrea; ALIVERTI, Paolo; LUCIONI, Michele. An omnidirectional vision sensor for fast tracking for mobile robots. In: IEEE. **Instrumentation and Measurement Technology Conference, 1999. IMTC/99. Proceedings of the 16th IEEE**. 1999. v. 1, p. 151–155.

- BORENSTEIN, Johann; EVERETT, HR; FENG, Liqiang *et al.* Where am i? sensors and methods for mobile robot positioning. **University of Michigan**, v. 119, n. 120, p. 27, 1996.
- BORENSTEIN, Johann; EVERETT, Hobart R; FENG, Liqiang; WEHE, David. **Mobile robot positioning-sensors and techniques**. 1997.
- BORENSTEIN, Johann; FENG, Liqiang. Gyrodometry: A new method for combining data from gyros and odometry in mobile robots. In: IEEE. **Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on**. 1996. v. 1, p. 423–428.
- BORGES, Aderly Terezinha Stresser. **Localização do CEP no bloco do endereço em envelopes postais**. Tese (Doutorado) — Pontifícia Universidade Católica do Paraná, 2005.
- BRANCA, A; STELLA, E; DISTANTE, A. Matching of line segments for indoor navigation using projective invariants. **Proceedings of Intelligent Autonomous Systems**, v. 6, p. 439–446, 2000.
- BRASSART, Eric; DELAHOCHÉ, Laurent; CAUCHOIS, Cyril; DROCOURT, Cyril; PEGARD, Claude; MOUADDIB, El Mustapha. Experimental results got with the omnidirectional vision sensor: Syclop. In: IEEE. **Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on**. 2000. p. 145–152.
- BROWITT, Geoff. **Virtual Reality Telepresence Using a Mobile Robot**. : University of Edinburgh, 1997.
- BUANI, Bruna Ez; HIRAKAWA, André R; SARAIVA, Antonio M; BUENO, Jesus F. Aplicação do algoritmo dos k-vizinhos mais próximos estendido com análise estatística e lógica fuzzy knn para seleção de características e classificação de abelhas. 2002.
- BUGEJA, Marvin K; FABRI, Simon G. Multilayer perceptron adaptive dynamic control for trajectory tracking of mobile robots. In: IEEE. **IEEE Industrial Electronics, IECON 2006-32nd Annual Conference on**. 2006. p. 3798–3803.
- BUNSCHOTEN, Roland; KROSE, Ben. Robust scene reconstruction from an omnidirectional vision system. **Robotics and Automation, IEEE Transactions on**, IEEE, v. 19, n. 2, p. 351–357, 2003.
- BURGARD, Wolfram; FOX, Dieter; HENNIG, Daniel; SCHMIDT, Timo. Estimating the absolute position of a mobile robot using position probability grids. In: **Proceedings of the national conference on artificial intelligence**. 1996. p. 896–901.
- CAMPBELL, Jason; SUKTHANKAR, Rahul; NOURBAKHSH, Illah; PAHWA, Aroon. A robust visual odometry and precipice detection system using consumer-grade monocular vision. In: **Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on**. 2005. p. 3421–3427.

- CAO, ZL; RONING, JJ; HALL, EL. Omnidirectional vision navigation integrating bancon recognition with positioning. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Cambridge Symposium_Intelligent Robotics Systems**. 1987. p. 213–220.
- CAO, Zuo L.; OH, Sung J.; HALL, Ernest L. Dynamic omnidirectional vision for mobile robots. **Proc. SPIE**, v. 0579, p. 405–414, 1985.
- CARVALHO, E. A. N.; SÁ, Alan; ALVES, E.; FREIRE, E. O.; MATOS, L. N.; LUCIANO, B. Estimação de parâmetros para navegação de um robô móvel a partir de dados obtidos por um sistema de visão computacional. In: **XVI Congresso Brasileiro de Automática**. 2006. p. 3457–3462.
- CHANG, Chin-Kai; SIAGIAN, Christian; ITTI, Laurent. Mobile robot vision navigation and obstacle avoidance based on gist and saliency algorithms. **Journal of Vision**, Association for Research in Vision and Ophthalmology, v. 11, n. 11, p. 927–927, 2011.
- CHATILA, Raja; LAUMOND, Jean-Paul. Position referencing and consistent world modeling for mobile robots. In: IEEE. **Robotics and Automation. Proceedings. 1985 IEEE International Conference on**. 1985. v. 2, p. 138–145.
- CHEN, Ching-Pei; CHEN, Jing-Yi; HUANG, Chun-Kai; LU, Jau-Ching; LIN, Pei-Chun. Sensor data fusion for body state estimation in a bipedal robot and its feedback control application for stable walking. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 15, n. 3, p. 4925–4946, 2015.
- CHEN, Shenchang Eric. Quicktime vr: An image-based approach to virtual environment navigation. In: **Proceedings of the 22Nd Annual Conference on Computer Graphics and Interactive Techniques**. New York, NY, USA: ACM, 1995. (SIGGRAPH '95), p. 29–38. ISBN 0-89791-701-4.
- CHENAVIER, Frédéric; CROWLEY, James L. Position estimation for a mobile robot using vision and odometry. In: IEEE. **Robotics and Automation, 1992. Proceedings., 1992 IEEE International Conference on**. 1992. p. 2588–2593.
- CHIU, Ming-Yee; DEPOMMIER, R.; SPINDLER, T. An embedded real-time vision system for 24-hour indoor/outdoor car-counting applications. In: **Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on**. 2004. v. 3, p. 338–341 Vol.3. ISSN 1051-4651.
- CHOSSET, Howie; BURDICK, Joel. Sensor-based exploration: The hierarchical generalized voronoi graph. **The International Journal of Robotics Research**, SAGE Publications, v. 19, n. 2, p. 96–125, 2000.
- CHOSSET, Howie; NAGATANI, Keiji. Topological simultaneous localization and mapping (slam): toward exact localization without explicit localization. **Robotics and Automation, IEEE Transactions on**, IEEE, v. 17, n. 2, p. 125–137, 2001.
- CLERENTIN, Arnaud; DELAHOUCHE, Laurent; BRASSART, Eric; CAUCHOIS, Cyril. Mobile robot localization based on multi target tracking. In: IEEE. **Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on**. 2002. v. 1, p. 13–18.

COLENBRANDER, August. Measuring vision and vision loss. **Duane's Ophthalmology**. 15th ed. Philadelphia, Pa: Lippincott Williams & Wilkins, 2009.

CORSON, M.S.; MOSS, R.H. Computer vision: An in-depth look at the technique and applications of digital image processing. **Potentials, IEEE**, v. 5, n. 4, p. 31–34, Dec 1986. ISSN 0278-6648.

COSTA, Anna Helena Reali; PEGORARO, R. Construindo robôs autônomos para partidas de futebol: o time guaranÁ. sba. In: **Controle&Automação**. 2000. p. 141–149.

COURBON, Jonathan; MEZOUAR, Youcef; ECK, Laurent; MARTINET, Philippe. A generic fisheye camera model for robotic applications. In: IEEE. **Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on**. 2007. p. 1683–1688.

COX, Ingemar J. Blanche-an experiment in guidance and navigation of an autonomous robot vehicle. **Robotics and Automation, IEEE Transactions on**, IEEE, v. 7, n. 2, p. 193–204, 1991.

CROWLEY, James L. World modeling and position estimation for a mobile robot using ultrasonic ranging. In: IEEE. **Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on**. 1989. p. 674–680.

DALAL, Navneet; TRIGGS, Bill. Histograms of oriented gradients for human detection. In: IEEE. **Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on**. 2005. v. 1, p. 886–893.

de Castro, Armando Antonio Monteiro; do Prado, Pedro Paulo Leite. Algoritmos para reconhecimento de padrões. **Revista Ciências Exatas**, v. 8, n. 2002, 2001.

DELLAERT, Frank; FOX, Dieter; BURGARD, Wolfram; THRUN, Sebastian. Monte carlo localization for mobile robots. In: IEEE. **Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on**. 1999. v. 2, p. 1322–1328.

DERRIEN, Steven; KONOLIGE, Kurt. Approximating a single viewpoint in panoramic imaging devices. In: IEEE. **Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on**. 2000. v. 4, p. 3931–3938.

Di Rocco, Maurizio; REGGENTE, Matteo; SAFFIOTTI, Alessandro. Gas source localization in indoor environments using multiple inexpensive robots and stigmergy. In: **Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on**. 2011. p. 5007–5014.

DIOSI, Albert; SEGVIC, Sinisa; REMAZEILLES, Anthony; CHAUMETTE, François. Experimental evaluation of autonomous driving based on visual memory and image-based visual servoing. **Intelligent Transportation Systems, IEEE Transactions on**, v. 12, n. 3, p. 870–883, Sept 2011. ISSN 1524-9050.

DRUCKER, Johanna. Is there a “digital” art history? **Visual Resources**, Taylor & Francis, v. 29, n. 1-2, p. 5–13, 2013.

DUCKETT, Tom; AXELSSON, Mikael; SAFFIOTTI, Alessandro. Learning to locate an odour source with a mobile robot. In: **Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on**. 2001. v. 4, p. 4017–4022 vol.4. ISSN 1050-4729.

DUDANI, Sahibsingh A; BREEDING, Kenneth J; MCGHEE, Robert B. Aircraft identification by moment invariants. **Computers, IEEE Transactions on**, IEEE, v. 100, n. 1, p. 39–46, 1977.

DURRANT-WHYTE, Hugh; BAILEY, Tim. Simultaneous localization and mapping: part i. **Robotics & Automation Magazine, IEEE**, IEEE, v. 13, n. 2, p. 99–110, 2006.

DURRANT-WHYTE, Hugh F. Uncertain geometry in robotics. **Robotics and Automation, IEEE Journal of**, IEEE, v. 4, n. 1, p. 23–31, 1988.

ECKERT, Juergen; GERMAN, Reinhard; DRESSLER, Falko. An indoor localization framework for four-rotor flying robots using low-power sensor nodes. **IEEE Transactions on Instrumentation and Measurement**, v. 60, n. 2, p. 336–344, 2011.

EDSINGER-GONZALES, Aaron; WEBER, Jeff. Domo: a force sensing humanoid robot for manipulation research. In: IEEE. **Humanoid Robots, 2004 4th IEEE/RAS International Conference on**. 2004. v. 1, p. 273–291.

EHLGEN, Tobias; THOM, Markus; GLASER, Markus. Omnidirectional cameras as backing-up aid. In: IEEE. **Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on**. 2007. p. 1–5.

ELFES, Alberto. Sonar-based real-world mapping and navigation. **Robotics and Automation, IEEE Journal of**, IEEE, v. 3, n. 3, p. 249–265, 1987.

ELFES, Alberto. **Occupancy grids: a probabilistic framework for robot perception and navigation**. Tese (Doutorado) — Carnegie Mellon University, 1989.

ENGELSON, Sean P; MCDERMOTT, Drew V. Error correction in mobile robot map learning. In: IEEE. **Robotics and Automation, 1992. Proceedings., 1992 IEEE International Conference on**. 1992. p. 2555–2560.

ESCHER, Maurits Cornelis. **Hand with reflecting sphere**. 1935.

FAIRFIELD, N. Simple landmark navigation on a three-layer mobile robot. In: . 2002.

FERREIRA, MF. **O Reconhecimento de Padrões**. Tese (Doutorado) — Universidade de Brasília, 1994.

FISHER, Forest; CHIEN, Steve; LO, Edisanter; GREELEY, Ronald. Sar image processing using artificial intelligence planning. In: **Image Analysis and Processing, 1999. Proceedings. International Conference on**. 1999. p. 856–861.

FLUSSER, Jan; SUK, Tomaš. A moment-based approach to registration of images with affine geometric distortion. **Geoscience and Remote Sensing, IEEE Transactions on**, IEEE, v. 32, n. 2, p. 382–387, 1994.

- FLUSSER, Jan; SUK, Tomas; SAIC, Stanislav. Recognition of blurred images by the method of moments. **Image Processing, IEEE Transactions on**, IEEE, v. 5, n. 3, p. 533–538, 1996.
- FORLIZZI, Jodi; DISALVO, Carl. Service robots in the domestic environment: a study of the roomba vacuum in the home. In: ACM. **Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction**. 2006. p. 258–265.
- FOX, Dieter; BURGARD, Wolfram; DELLAERT, Frank; THRUN, Sebastian. Monte carlo localization: Efficient position estimation for mobile robots. **AAAI/IAAI**, v. 1999, p. 343–349, 1999.
- FOX, Dieter; BURGARD, Wolfram; KRUPPA, Hannes; THRUN, Sebastian. Collaborative multi-robot localization. In: **Mustererkennung 1999**. : Springer, 1999. p. 15–26.
- FOX, Dieter; BURGARD, Wolfram; THRUN, Sebastian. Active markov localization for mobile robots. **Robotics and Autonomous Systems**, Elsevier, v. 25, n. 3, p. 195–207, 1998.
- FUKUNAGA, Keinosuke; HUMMELS, Donald M. Bayes error estimation using parzen and k-nn procedures. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, n. 5, p. 634–643, 1987.
- FULLER, Sawyer B; HELBLING, E Farrell; CHIRARATTANANON, Pakpong; WOOD, Robert J. Using a mems gyroscope to stabilize the attitude of a fly-sized hovering robot. In: DELFT UNIVERSITY OF TECHNOLOGY. **IMAV 2014: International Micro Air Vehicle Conference and Competition 2014, Delft, The Netherlands, August 12-15, 2014**. 2014.
- GARDNER, MW; DORLING, SR. Artificial neural networks (the multilayer perceptron) - a review of applications in the atmospheric sciences. **Atmospheric environment**, Elsevier, v. 32, n. 14, p. 2627–2636, 1998.
- GASPAR, José; NIALI, Winters; SANTOS-VICTOR, José. Vision-based navigation and environmental representations with an omnidirectional camera. **Robotics and Automation, IEEE Transactions on**, v. 16, n. 6, p. 890–898, December 2000. ISSN 1042-296X.
- GASPAR, José; SANTOS-VICTOR, Jose. Visual path following with a catadioptric panoramic camera. In: CITESEER. **7th Int. Symp. on Intelligent Robotic Systems**. 1999. p. 139–147.
- GERSTMAYR-HILLEN, Lorenz; RÖBEN, Frank; KRZYKAWSKI, Martin; KREFT, Sven; VENJAKOB, Daniel; MÖLLER, Ralf. Dense topological maps and partial pose estimation for visual control of an autonomous cleaning robot. **Robotics and Autonomous Systems**, Elsevier, v. 61, n. 5, p. 497–516, 2013.
- GEYER, Christopher; DANIILIDIS, Kostas. Omnidirectional video. **The Visual Computer**, Springer, v. 19, n. 6, p. 405–416, 2003.

- GHIASEDDIN, Behnam; SERYASAT, Omid Rahmani; HADDADNIA, Javad. Multilayer perceptron network for automatic driving vehicle. **Life Science Journal**, v. 10, n. 5s, 2013.
- GLUCKMAN, Joshua; NAYAR, Shree K. Ego-motion and omnidirectional cameras. In: **Computer Vision, 1998. Sixth International Conference on**. 1998. p. 999–1005.
- GOEDEMÉ, Toon; NUTTIN, Marnix; TUYTELAARS, Tinne; GOOL, Luc Van. Omnidirectional vision based topological navigation. **International Journal of Computer Vision**, Springer, v. 74, n. 3, p. 219–236, 2007.
- GONZÁLEZ, Beatriz; VALDEZ, Fevrier; MELIN, Patricia; PRADO-ARECHIGA, German. Echocardiogram image recognition using neural networks. In: **Recent Advances on Hybrid Approaches for Designing Intelligent Systems**. : Springer, 2014. p. 427–435.
- GONZALEZ, Javier; OLLERO, Anibal; REINA, Antonio. Map building for a mobile robot equipped with a 2d laser rangefinder. In: IEEE. **Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on**. 1994. p. 1904–1909.
- GONZALEZ, Rafael C. **Digital image processing**. : Pearson Education India, 2009.
- GORDON, Neil J; SALMOND, David J; SMITH, Adrian FM. Novel approach to nonlinear/non-gaussian bayesian state estimation. In: IET. **IEE Proceedings F (Radar and Signal Processing)**. 1993. v. 140, n. 2, p. 107–113.
- GRASSI, Valdir. **Sistema de visão omnidirecional aplicado no controle de robôs móveis**. Tese (Doutorado) — Universidade de São Paulo, 2002.
- GRASSI, Valdir; DECCÓ, Cláudia CG; JUNIOR, Jun Okamoto; PORTO, Arthur José Vieira. Desenvolvimento de um sistema de visão omnidirecional. In: **XVI Congresso Brasileiro de Engenharia Mecânica**. 2001.
- GUDELOGLU, Ahmet; BRAHMBHATT, Jamin V; MULANEY, Annika; PAREKATTIL, Sijo. Intraoperative doppler ultrasound during robotic surgery. In: **Advances in Image-Guided Urologic Surgery**. : Springer, 2015. p. 91–101.
- GUTIERREZ, Daniel; RITUERTO, Alejandro; MONTIEL, JMM; GUERRERO, José Jesús. Adapting a real-time monocular visual slam from conventional to omnidirectional cameras. In: IEEE. **Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on**. 2011. p. 343–350.
- GUZEL, Mehmet S; ERWIN, John; ZAKARIA, Wan Nurshazwani Wan. Vision based object recognition of e-puck mobile robot for warehouse application. **International Journal of Integrated Engineering**, v. 6, n. 3, 2015.
- HAGRAS, Hani; COLLEY, Martin; CALLAGHAN, Victor; CARR-WEST, Malcolm. Online learning and adaptation of autonomous mobile robots for sustainable agriculture. **Autonomous Robots**, Springer, v. 13, n. 1, p. 37–52, 2002.

- HAHNEL, Dirk; BURGARD, Wolfram; FOX, Dieter; THRUN, Sebastian. An efficient fastslam algorithm for generating maps of large-scale cyclic environments from raw laser range measurements. In: IEEE. **Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on**. 2003. v. 1, p. 206–211.
- HAYES, Adam T; MARTINOLI, Alcherio; GOODMAN, Rodney M. Distributed odor source localization. **Sensors Journal, IEEE**, v. 2, n. 3, p. 260–271, Jun 2002. ISSN 1530-437X.
- HAYKIN, Simon S. **Redes neurais**. : Bookman, 2001.
- HAYKIN, Simon S; HAYKIN, Simon S; HAYKIN, Simon S; HAYKIN, Simon S. **Neural networks and learning machines**. : Pearson Education Upper Saddle River, 2009.
- HEIJDEN, F. Van Der. **Image based measurement systems: object recognition and parameter estimation**. Chichester: John Wiley & Sons Ltd, 1994.
- HERNANDEZ, Danilo Caceres; HOANG, Van-Dung; FILONENKO, Alexander; JO, Kang-Hyun. Vision-based heading angle estimation for an autonomous mobile robots navigation. In: **Industrial Electronics (ISIE), 2014 IEEE 23rd International Symposium on**. 2014. p. 1967–1972.
- HICKS, R Andrew; BAJCSY, Ruzena. Catadioptric sensors that approximate wide-angle perspective projections. In: IEEE. **Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on**. 2000. v. 1, p. 545–551.
- HIREMATH, Santosh A; HEIJDEN, Gerie WAM Van Der; EVERT, Frits K Van; STEIN, Alfred; BRAAK, Cajo JF Ter. Laser range finder model for autonomous navigation of a robot in a maize field using a particle filter. **Computers and Electronics in Agriculture**, Elsevier, v. 100, p. 41–50, 2014.
- HOFMEISTER, Marius; LIEBSCH, Maria; ZELL, Andreas. Visual self-localization for small mobile robots with weighted gradient orientation histograms. In: **40th International Symposium on Robotics (ISR)**. 2009. p. 87–91.
- HOFMEISTER, Marius; VORST, Philipp; ZELL, Andreas. A comparison of efficient global image features for localizing small mobile robots. In: VDE. **Robotics (ISR), 2010 41st International Symposium on and 2010 6th German Conference on Robotics (ROBOTIK)**. 2010. p. 1–8.
- HOLLINGUM, Jack. Robots in agriculture. **Industrial Robot: An International Journal**, MCB UP Ltd, v. 26, n. 6, p. 438–446, 1999.
- HONG, Jia-Wei; TAN, Xiaonan; PINETTE, Brian; WEISS, Richard; RISEMAN, Edward M. Image-based homing. **Control Systems, IEEE**, IEEE, v. 12, n. 1, p. 38–45, 1992.
- HOWARD, Andrew; KITCHEN, Les. Vision-based navigation using natural landmarks. In: SPRINGER. **Field and Service Robotics**. 1998. p. 336–343.
- HU, Ming-Kuei. Visual pattern recognition by moment invariants. **Information Theory, IRE Transactions on**, IEEE, v. 8, n. 2, p. 179–187, 1962.

ISHIGURO, Hiroshi. Development of low-cost compact omnidirectional vision sensors and their applications. In: **Proc. Int. Conf. Information systems, analysis and synthesis**. 1998. p. 433–439.

ISHIGURO, Hiroshi; YAMAMOTO, Masashi; TSUJI, Saburo. Omni-directional stereo for making global map. In: **Computer Vision, 1990. Proceedings, Third International Conference on**. 1990. p. 540–547.

ISHIGURO, Hiroshi; YAMAMOTO, Masashi; TSUJI, Saburo. Omni-directional stereo. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, v. 14, n. 2, p. 257–262, Feb 1992. ISSN 0162-8828.

JANG, Gijeong; LEE, Sungho; KWEON, Inso. Color landmark based self-localization for indoor mobile robots. In: IEEE. **Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on**. 2002. v. 1, p. 1037–1042.

JIA, Hanchao; LI, Shigang. Scene analysis based on horse vision system. In: **MVA**. 2011. p. 267–270.

JOGAN, Matjaz; LEONARDIS, Ales. Robust localization using eigenspace of spinning-images. In: IEEE. **Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on**. 2000. p. 37–44.

KALMAN, Rudolph Emil. A new approach to linear filtering and prediction problems. **Journal of Fluids Engineering**, American Society of Mechanical Engineers, v. 82, n. 1, p. 35–45, 1960.

KANADE, Takeo. Virtualized reality: Digitizing a 3-d time varying event as is an in real time. **Mixed Reality**, p. 41–57, 1999.

KAWAGOE, Masahiro; TOJO, Akio. Fingerprint pattern classification. **Pattern Recognition**, Elsevier, v. 17, n. 3, p. 295–303, 1984.

KAWAGUCHI, Yoshida; YOSHIDA, I; KURUMATANI, H; KIKUTA, TAKT; YAMADA, YAYY. Internal pipe inspection robot. In: **Robotics and Automation, 1995. Proceedings., 1995 IEEE International Conference on**. 1995. v. 1, p. 857–862 vol.1. ISSN 1050-4729.

KIM, Kwang In; JUNG, Keechul; KIM, Hang Joon. Face recognition using kernel principal component analysis. **Signal Processing Letters, IEEE**, IEEE, v. 9, n. 2, p. 40–42, 2002.

KIM, Youngwook; SEKHON, Satbir S. Detection of moving target and localization of clutter using doppler radar on mobile platform. IEEE, 2015.

KINOSHITA, Hideyuki; HOSHINO, Kazunori; MATSUMOTO, Kiyoshi; SHIMOYAMA, Isao. Thin compound eye camera with a zooming function by reflective optics. In: **Micro Electro Mechanical Systems, 2005. MEMS 2005. 18th IEEE International Conference on**. 2005. p. 235–238. ISSN 1084-6999.

KIRBY, Michael. **Geometric data analysis: an empirical approach to dimensionality reduction and the study of patterns**. : John Wiley & Sons, Inc., 2000.

KO, Bang-Yun; SONG, Jae-Bok; LEE, Sooyong. Real-time building of a thinning-based topological map with metric features. In: IEEE. **Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on.** 2004. v. 2, p. 1524–1529.

KOHONEN, Teuvo. The self-organizing map. **Proceedings of the IEEE**, IEEE, v. 78, n. 9, p. 1464–1480, 1990.

KOHONEN, Teuvo. **Self-organizing maps.** : Springer Science & Business Media, 2001.

KONOLIGE, Kurt; MARDER-EPPSTEIN, Eitan; MARTHI, Bhaskara. Navigation in hybrid metric-topological maps. In: IEEE. **Robotics and Automation (ICRA), 2011 IEEE International Conference on.** 2011. p. 3041–3047.

KORTENKAMP, David; BONASSO, RP; MURPHY, R. **AI-based Mobile Robots: Case studies of successful robot systems.** : Cambridge, MA, 1998.

KOSECKÁ, Jana; LI, Fayin. Vision based topological markov localization. In: IEEE. **Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on.** 2004. v. 2, p. 1481–1486.

KREBS, HI; HOGAN, N; VOLPE, BT; AISEN, ML; EDELSTEIN, L; DIELS, C. Overview of clinical trials with mit-manus: a robot-aided neuro-rehabilitation facility. **Technology and Health Care**, IOS Press, v. 7, n. 6, p. 419–423, 1999.

KREBS, H Igo; HOGAN, Neville; AISEN, Mindy L; VOLPE, Bruce T. Robot-aided neurorehabilitation. **Rehabilitation Engineering, IEEE Transactions on**, IEEE, v. 6, n. 1, p. 75–87, 1998.

KRISHNAN, A.; AHUJA, N. Panoramic image acquisition. In: **Computer Vision and Pattern Recognition, 1996. Proceedings CVPR '96, 1996 IEEE Computer Society Conference on.** 1996. p. 379–384. ISSN 1063-6919.

KRISHNAN, Gurunandan; NAYAR, Shree K. Cata-fisheye camera for panoramic imaging. In: IEEE. **Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on.** 2008. p. 1–8.

KRISHNASAMY, R.; WONG, W.; SHEN, E.; PEPIC, S.; HORNSEY, R.; THOMAS, P.J. High precision target tracking with a compound-eye image sensor. In: **Electrical and Computer Engineering, 2004. Canadian Conference on.** 2004. v. 4, p. 2319–2323 Vol.4. ISSN 0840-7789.

KROSE, Ben; BUNSCHOTEN, Roland; HAGEN, ST; TERWIJN, B; VLASSIS, Nikos. Household robots look and learn: environment modeling and localization from an omnidirectional vision system. **Robotics & Automation Magazine, IEEE**, IEEE, v. 11, n. 4, p. 45–52, 2004.

KUAN, Lin; YI, Zhang; YAWEN, Dang; JUNHENG, Huang. Sar image matching algorithm base on improved hu invariant moments. **International Journal of Signal Processing, Image Processing and Pattern Recognition**, v. 8, n. 1, p. 1–8, 2015.

- KUIPERS, Benjamin; BYUN, Yung-Tai. A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations. **Robotics and autonomous systems**, Elsevier, v. 8, n. 1, p. 47–63, 1991.
- KUO, Chung-Hsien; CHEN, Ting-Shuo. Modeling and control of autonomous soccer robots using high-level petri nets. In: **SICE Annual Conference 2010, Proceedings of**. 2010. p. 2226–2231.
- KUO, Chung-Hsien; YANG, Chun-Ming; YANG, Fang-Chung. Development of intelligent vision fusion based autonomous soccer robot. In: **Mechatronics, 2005. ICM '05. IEEE International Conference on**. 2005. p. 124–129.
- KURFESS, Thomas R. **Robotics and automation handbook**. : CRC press, 2004.
- LAND, Michael F; FERNALD, Russell D. The evolution of eyes. **Annual Review of Neuroscience**, v. 15, n. 1, p. 1–29, 1992. PMID: 1575438.
- LAUNAY, Fabien; OHYA, Akihisa. Autonomous indoor mobile robot navigation by detecting fluorescent tubes. In: **In Proceedings of the 10th International Conference on Advanced Robotics (ICAR'01)**. 2001.
- LEE, KangWoo; KIM, Hyoung-Rock; YOON, Wan Chul; YOON, Young-Sik; KWON, Dong-Soo. Designing a human-robot interaction framework for home service robot. In: IEEE. **Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on**. 2005. p. 286–293.
- LEONARD, John J; DURRANT-WHYTE, Hugh F. Simultaneous map building and localization for an autonomous mobile robot. In: IEEE. **Intelligent Robots and Systems' 91. Intelligence for Mechanical Systems, Proceedings IROS'91. IEEE/RSJ International Workshop on**. 1991. p. 1442–1447.
- LEONARDIS, Aleš; BISCHOF, Horst. Robust recognition using eigenimages. **Computer Vision and Image Understanding**, Elsevier, v. 78, n. 1, p. 99–118, 2000.
- LEYVA, F Hugo Ramírez. Robótica 1. introducción. 2012.
- LI, Shigang. Full-view spherical image camera. In: IEEE. **Pattern Recognition, 2006. ICPR 2006. 18th International Conference on**. 2006. v. 4, p. 386–390.
- LI, Xing Ze; ZHU, Ling; HUA, Yi. Embedded robot vision system based on dsp. In: TRANS TECH PUBL. **Applied Mechanics and Materials**. 2015. v. 734, p. 168–171.
- LIMA, SL; SAAVEDRA, OR; SILVA, AC; SOUSA, AN de; FILHO, RN Rosa. Rnb aplicada ao reconhecimento de padrões de falhas em imagens térmicas de equipamentos de subestações. **III Simpósio Brasileiro de Sistemas Elétricos (SBSE/2010)**, 2010.
- LIU, Chengjun; WECHSLER, Harry. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. **Image processing, IEEE Transactions on**, IEEE, v. 11, n. 4, p. 467–476, 2002.
- LIU, Yu-Chih; LIN, Kai-Ying; CHEN, Yong-Sheng. Bird's-eye view vision system for vehicle surrounding monitoring. In: **Robot Vision**. : Springer, 2008. p. 207–218.

- LIU, Zhen; LI, Xiaojing; LI, Fengjiao; ZHANG, Guangjun. Calibration method for line-structured light vision sensor based on a single ball target. **Optics and Lasers in Engineering**, Elsevier, v. 69, p. 20–28, 2015.
- LU, Huimin; YANG, Shaowu; ZHANG, Hui; ZHENG, Zhiqiang. A robust omnidirectional vision sensor for soccer robots. **Mechatronics**, Elsevier, v. 21, n. 2, p. 373–389, 2011.
- LYNCH, Kevin M; PESHKIN, Michael A. 19.1 linear and rotational sensors. **The Mechatronics Handbook, -2 Volume Set**, CRC Press, v. 19, p. 1, 2014.
- MARCHANT, Gary E; ALLENBY, Braden; ARKIN, Ronald; BARRETT, Edward T; BORENSTEIN, Jason; GAUDET, Lyn M; KITTRIE, Orde; LIN, Patrick; LUCAS, George R; O'MEARA, Richard *et al.* International governance of autonomous military robots. **Colum. Sci. & Tech. L. Rev.**, v. 12, p. 272–276, 2011.
- MARCHESE, Fabio M; SORRENTI, Domenico G. Omni-directional vision with a multi-part mirror. In: **RoboCup 2000: Robot Soccer World Cup IV.** : Springer, 2001. p. 179–188.
- Marques Filho, Ogê; Vieira Neto, Hugo. **Processamento Digital de Imagens**. 1st. ed. Rio de Janeiro: Brasport Livros e Multimídia Ltda, 1999.
- MARQUES, Lino; NUNES, Urbano; ALMEIDA, Anibal T de. A new 3d optical triangulation sensor for robotics. In: IEEE. **Advanced Motion Control, 1998. AMC'98-Coimbra., 1998 5th International Workshop on.** 1998. p. 512–517.
- MATARIC, Maja J. **A distributed model for mobile robot environment-learning and navigation.** 1990.
- MATARIĆ, Maja J. **The robotics primer.** : Mit Press, 2007.
- MATSUMOTO, Yoshio; IKEDA, Kazunori; INABA, Masayuki; INOUE, Hirochika. Exploration and navigation in corridor environment based on omni-view sequence. In: IEEE. **Intelligent Robots and Systems, 2000.(IROS 2000). Proceedings. 2000 IEEE/RSJ International Conference on.** 2000. v. 2, p. 1505–1510.
- MCCULLOCH, Warren S; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943.
- MCGILL, Matthew; SALLEH, Rudino; WILEY, Timothy; RATTER, Adrian; FARID, Reza; SAMMUT, Claude; MILSTEIN, Adam. Virtual reconstruction using an autonomous robot. In: **Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on.** 2012. p. 1–8.
- MEI, Christopher; RIVES, Patrick. Single view point omnidirectional camera calibration from planar grids. In: IEEE. **Robotics and Automation, 2007 IEEE International Conference on.** 2007. p. 3945–3950.
- MEKKAOUI, M.; BENAMEUR, N.; OURIACHI, K. A description approach of information in a robotics-vision system. In: **Computer Architectures for Machine Perception, 1995. Proceedings. CAMP '95.** 1995. p. 114–119.

- MENEGATTI, Emanuele. **Omnidirectional Vision for Mobile Robots**. Tese (Doutorado) — Citeseer, 2002.
- MENEGATTI, Emanuele; PRETTO, Alberto; SCARPA, Alberto; PAGELLO, Enrico. Omnidirectional vision scan matching for robot localization in dynamic environments. **Robotics, IEEE Transactions on**, IEEE, v. 22, n. 3, p. 523–535, 2006.
- MENGLONG, Cao; XIAOLEI, Zuo. Collaborative pseudolites autonomous navigation for multi-vehicle. In: **Control Conference (CCC), 2011 30th Chinese**. 2011. p. 4797–4801. ISSN 1934-1768.
- MIURA, Jun; YAMAMOTO, Koshiro. Robust view matching-based markov localization in outdoor environments. In: IEEE. **Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on**. 2008. p. 2970–2976.
- MONKMAN, Gareth J; TAYLOR, PM. Thermal tactile sensing. **Robotics and Automation, IEEE Transactions on**, v. 9, n. 3, p. 313–318, Jun 1993. ISSN 1042-296X.
- MONTEMERLO, Michael; THRUN, Sebastian. Simultaneous localization and mapping with unknown data association using fastslam. In: IEEE. **Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on**. 2003. v. 2, p. 1985–1991.
- MONTEMERLO, Michael; THRUN, Sebastian. Fastslam 2.0. **FastSLAM: A Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics**, Springer, p. 63–90, 2007.
- MONTEMERLO, Michael; THRUN, Sebastian; KOLLER, Daphne; WEGBREIT, Ben *et al.* Fastslam: A factored solution to the simultaneous localization and mapping problem. In: **Proceedings of the AAAI National Conference on Artificial Intelligence**. Edmonton, Canada: AAAI.
- MORAVEC, Hans P. Sensor fusion in certainty grids for mobile robots. **AI magazine**, v. 9, n. 2, p. 61, 1988.
- MOREIRA, Nuno Miguel da Silva. **Localização de Robôs Autônomos em Ambiente Industrial**. Tese (Doutorado) — UNIVERSIDADE DO PORTO, 2014.
- MORISHITA, Hiroshi; FUKUI, Rui; SATO, Tomomasa. High resolution pressure sensor distributed floor for future human-robot symbiosis environments. In: IEEE. **Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on**. 2002. v. 2, p. 1246–1251.
- MORITA, Toshihiko; YASUKAWA, Yusuke; INAMOTO, Yasushi; TAKASHI, U; KAWAKAMI, Susumu. Measurement in three dimensions by motion stereo and spherical mapping. In: **Computer Vision and Pattern Recognition, 1989. Proceedings CVPR '89., IEEE Computer Society Conference on**. 1989. p. 422–428. ISSN 1063-6919.
- MURILLO, Ana Cristina; SINGH, Gautam; KOSECKA, Jana; GUERRERO, José Jesús. Localization in urban environments using a panoramic gist descriptor. **Robotics, IEEE Transactions on**, IEEE, v. 29, n. 1, p. 146–160, 2013.

NANDA, Harsh; CUTLER, Ross. Practical calibrations for a real-time digital omnidirectional camera. **CVPR Technical Sketch**, Citeseer, v. 20, 2001.

NANDHITHA, N.M.; MANOHARAN, N.; RANI, B.S.; VENKATARAMAN, B.; SUNDARAM, P.K.; RAJ, Baldev. Detection and quantification of tungsten inclusion in weld thermographs for on-line weld monitoring by region growing and morphological image processing algorithms. In: **Conference on Computational Intelligence and Multimedia Applications, 2007. International Conference on**. 2007. v. 3, p. 513–518.

NAST, Rafael Reis; RABELO, Francisco Felipe Lacerda; CASTRO, Rodrigo Lopes Cabral de; SILVA, Edson do Socorro Cardoso da; BEZERRA, Johelden Campos; SOBRINHO, Elionai Gomes de Almeida. Transformação e conservação de energia em um circuito automobilístico. **Engenharia de Computação em Revista**, v. 1, n. 1, 2014.

NAYAR, Shree K. Catadioptric omnidirectional camera. In: **Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on**. 1997. p. 482–488. ISSN 1063-6919.

NAYAR, Shree K; BAKER, Simon. Catadioptric image formation. In: **Proceedings of the 1997 DARPA Image Understanding Workshop**. 1997. p. 1431–1437.

NAYAR, Shree K; NENE, Sameer A; MURASE, Hiroshi. Subspace methods for robot vision. **Robotics and Automation, IEEE Transactions on**, IEEE, v. 12, n. 5, p. 750–758, 1996.

NEGENBORN, Rudy. **Robot localization and kalman filters**. Tese (Doutorado) — Utrecht University, 2003.

NETO AJALMAR, Ricardo Sousa Guilherme de A. Barreto Jaime S. Cardoso Rocha. Diagnostic of pathology on the vertebral column with embedded reject option. In: **Proceedings of the 5th Iberian conference on Pattern recognition and image analysis**. 2011.

NETO, Edson Cavalcanti; REBOUCAS, Elizangela Souza; MORAES, Jermana Lopes de; GOMES, Samuel Luz; FILHO, Pedro Pedrosa Reboucas. Development control parking access using techniques digital image processing and applied computational intelligence. **Latin America Transactions, IEEE (Revista IEEE America Latina)**, v. 13, n. 1, p. 272–276, Jan 2015. ISSN 1548-0992.

NICHOLLS, Howard R; LEE, Mark H. A survey of robot tactile sensing technology. **The International Journal of Robotics Research**, Sage Publications, v. 8, n. 3, p. 3–30, 1989.

NOGUCHI, Noboru; TERAOKA, Hideo. Path planning of an agricultural mobile robot by neural network and genetic algorithm. **Computers and electronics in agriculture**, Elsevier, v. 18, n. 2, p. 187–204, 1997.

NOLFI, Stefano; PARISI, Domenico. Evolving non-trivial behaviors on real robots: an autonomous robot that picks up objects. In: **Topics in artificial intelligence**. : Springer, 1995. p. 243–254.

- NÚÑEZ, Pedro; VÁZQUEZ-MARTÍN, R; TORO, JC Del; BANDERA, Antonio; SANDOVAL, F. Natural landmark extraction for mobile robot navigation based on an adaptive curvature estimation. **Robotics and Autonomous Systems**, Elsevier, v. 56, n. 3, p. 247–264, 2008.
- OBDRZALEK, David. Small autonomous robot localization system. In: **Research and Development (SCoReD), 2009 IEEE Student Conference on**. 2009. p. 461–463.
- OHNO, Kazunori; TSUBOUCHI, Takashi; SHIGEMATSU, Bunji; YUTA, Shin'ichi. Differential gps and odometry-based outdoor navigation of a mobile robot. **Advanced Robotics**, Taylor & Francis, v. 18, n. 6, p. 611–635, 2004.
- OHYA, Akihisa; KOSAKA, Akio; KAK, Avi. Vision-based navigation by a mobile robot with obstacle avoidance using single-camera vision and ultrasonic sensing. **Robotics and Automation, IEEE Transactions on**, IEEE, v. 14, n. 6, p. 969–978, 1998.
- OLIVA, Aude; TORRALBA, Antonio. Modeling the shape of the scene: A holistic representation of the spatial envelope. **International journal of computer vision**, Springer, v. 42, n. 3, p. 145–175, 2001.
- OLIVA, Aude; TORRALBA, Antonio. Building the gist of a scene: The role of global image features in recognition. **Progress in brain research**, Elsevier, v. 155, p. 23–36, 2006.
- ONOE, Yoshio; YOKOYA, Naokazu; YAMAZAWA, Kazumasa; TAKEMURA, Haruo. Visual surveillance and monitoring system using an omnidirectional video camera. In: **IEEE. Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on**. 1998. v. 1, p. 588–592.
- ORIOLO, Giuseppe; ULIVI, Giovanni; VENDITTELLI, Marilena. Real-time map building and navigation for autonomous robots in unknown environments. **Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on**, IEEE, v. 28, n. 3, p. 316–333, 1998.
- ORTEGA, J Gomez; CAMACHO, EF. Mobile robot navigation in a partially structured static environment, using neural predictive control. **Control Engineering Practice**, Elsevier, v. 4, n. 12, p. 1669–1679, 1996.
- ÖZKUCUR, N Ergin; AKIN, H Levent. Cooperative multi-robot map merging using fast-slam. In: **RoboCup 2009: Robot Soccer World Cup XIII**. : Springer, 2010. p. 449–460.
- PARK, Shinyoung; LEE, Changsu; HAN, Kyunghee; WON, Seongho; JUNG, Youngbae; KIM, Younggil; YANG, Okryul; KANG, Wook; YOON, Songye. **Home service robot**. : Google Patents, ago. 28 2007. US Patent D549,756.
- PASIC, Resad P; RIZZO, John A; FANG, Hai; ROSS, Susan; MOORE, Matt; GUNNARSSON, Candace. Comparing robot-assisted with conventional laparoscopic hysterectomy: impact on cost and clinical outcomes. **Journal of minimally invasive gynecology**, Elsevier, v. 17, n. 6, p. 730–738, 2010.

- PATEL, Chirag; PATEL, Atul; PATEL, Dharmendra. Optical character recognition by open source ocr tool tesseract: A case study. **International Journal of Computer Applications**, Foundation of Computer Science, 244 5 th Avenue, # 1526, New York, NY 10001, USA India, v. 55, n. 10, p. 50–56, 2012.
- PERELMUTER, Guy; CARRERA, EV; VELLASCO, Marley; PACHECO, A. Reconhecimento de imagens bidimensionais utilizando redes neurais artificiais. **Anais do VIII SIBGRAPI**, p. 197–203, 1995.
- PÉREZ, Javier; CABALLERO, Fernando; MERINO, Luis. Enhanced monte carlo localization with visual place recognition for robust robot localization. **Journal of Intelligent & Robotic Systems**, Springer, p. 1–16, 2015.
- PINA, Aloísio Carlos de; PINA FILHO, Armando Carlos de. Identificação de falhas de um manipulador robótico pela aplicação de métodos de inteligência computacional.
- PIRKL, Gerald; MUNARETTO, Daniele; FISCHER, Carl; AN, Chunlei; LUKOWICZ, Paul; KLEPAL, Martin; TIMM-GIEL, Andreas; WIDMER, Joerg; PESCH, Dirk; GELLERSEN, Hans *et al.* Virtual lifeline: Multimodal sensor data fusion for robust navigation in unknown environments. **Pervasive and Mobile Computing**, Elsevier, v. 8, n. 3, p. 388–401, 2012.
- PIZO, Gerardo Antonio Idrobo. Projeto de um descritor para o alinhamento de imagens de profundidade de superfícies com aplicação em visão robótica. 2014.
- PLACHETKA, Tomas. Pov ray: Persistence of vision parallel raytracer. In: **Proc. of Spring Conf. on Computer Graphics, Budmerice, Slovakia**. 1998. p. 123–129.
- POWERS, I Nourbakhsk R; BIRCHFIELD, S. Dervish: An office navigation robot. **AI Magazine**, v. 16, n. 2, p. 53–60, 1995.
- PROKOP, Richard J; REEVES, Anthony P. A survey of moment-based techniques for unoccluded object representation and recognition. **CVGIP: Graphical Models and Image Processing**, Elsevier, v. 54, n. 5, p. 438–460, 1992.
- RACHEL, F.M.; CUGNASCA, P.S. Using artificial intelligence techniques in critical safety systems. In: **System Safety, 2006. The 1st Institution of Engineering and Technology International Conference on**. 2006. p. 64–70. ISSN 0537-9989.
- REES, Donald W. **Panoramic television viewing system**. : Google Patents, abr. 7 1970. US Patent 3,505,465.
- RIISGARD, S; BLAS, Morten Rufus. A tutorial approach to simultaneous localization and mapping. **Massachusetts Institute of Technology (MIT), Mobile Robotics course materials**, 2005.
- RITTER, Gunter; GALLEGOS, María Teresa. Outliers in statistical pattern recognition and an application to automatic chromosome classification. **Pattern Recognition Letters**, Elsevier, v. 18, n. 6, p. 525–539, 1997.
- ROBOTICS, International Federation of; ASSOCIATION, CLAWAR *et al.* **Robotica an official journal of the International Federation of Robotics and of the CLAWAR Association**. : Cambridge University Press, 2015.

- ROCHA NETO, AR; BARRETO, GA; CORTEZ, PC; MOTA, H da. Sinpatco: Sistema inteligente para diagnóstico de patologias da coluna vertebral. In: **Congresso Brasileiro de Automática**. 2006. p. 929–934.
- ROUS, Martin; LUPSCHEN, H; KRAISS, Karl-Friedrich. Vision-based indoor scene analysis for natural landmark detection. In: IEEE. **Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on**. 2005. p. 4642–4647.
- ROY, Nicholas; BURGARD, Wolfram; FOX, Dieter; THRUN, Sebastian. Coastal navigation-mobile robot navigation with uncertainty in dynamic environments. In: IEEE. **Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on**. 1999. v. 1, p. 35–40.
- RUAN, Xiaogang; LI, Wangbo. Ultrasonic sensor based two-wheeled self-balancing robot obstacle avoidance control system. In: IEEE. **Mechatronics and Automation (ICMA), 2014 IEEE International Conference on**. 2014. p. 896–900.
- RÜCKERT, Elmar A. **Simultaneous localisation and mapping for mobile robots with recent sensor technologies**. : na, 2009.
- RUIDIAS, Sandro Mossmann; DE, EXTRAÇÃO E ANÁLISE DE CARACTERÍSTICAS. **Extração e análise de características de imagens digitais para construção de um modelo de reconhecimento de padrões**. Tese (Doutorado) — Dissertação apresentada para conclusão de licenciatura em Sistemas de informação. Novo Hamburgo: Universidade de Feevale, 2005.
- RUMELHART, David E; HINTON, Geoffrey E; WILLIAMS, Ronald J. Learning representations by back-propagating errors. **Cognitive modeling**, v. 5, 1988.
- RUSSELL, R. Andrew. Heat trails as short-lived navigational markers for mobile robots. In: **Robotics and Automation, 1997. Proceedings., 1997 IEEE International Conference on**. 1997. v. 4, p. 3534–3539 vol.4.
- SAGÜÉS, C; MONTANO, L; ARAGÓN, Diputación General de. Dynamic behaviour analysis of non-contact compliant robot motions: Application to edge and plane localization.
- SANTOS, FERNANDO CHAGAS. Variações do método knn e suas aplicações na classificação automática de textos. Universidade Federal de Goiás, 2010.
- SARIFF, N; BUNIYAMIN, Norlida. An overview of autonomous mobile robot path planning algorithms. In: **Research and Development, 2006. SCORed 2006. 4th Student Conference on**. 2006. p. 183–188.
- SCARAMUZZA, Davide. **Omnidirectional vision: from calibration to robot motion estimation**. Tese (Doutorado) — Citeseer, 2008.
- SE, Stephen; LOWE, David; LITTLE, Jim. Local and global localization for mobile robots using visual landmarks. In: IEEE. **Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on**. 2001. v. 1, p. 414–420.

SE, Stephen; LOWE, David G; LITTLE, James J. Vision-based global localization and mapping for mobile robots. **Robotics, IEEE Transactions on**, IEEE, v. 21, n. 3, p. 364–375, 2005.

SHAH, Shishir; AGGARWAL, JK. A simple calibration procedure for fish-eye (high distortion) lens camera. In: IEEE. **Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on**. 1994. p. 3422–3427.

SHAH, Shishir; AGGARWAL, JK. Intrinsic parameter calibration procedure for a (high-distortion) fish-eye lens camera with distortion model and accuracy estimation*. **Pattern Recognition**, Elsevier, v. 29, n. 11, p. 1775–1788, 1996.

SHEIKH, Usman Ayub; JAMIL, Mohsin; AYAZ, Yasar. A comparison of various robotic control architectures for autonomous navigation of mobile robots. In: **Robotics and Emerging Allied Technologies in Engineering (iCREATE), 2014 International Conference on**. 2014. p. 239–243.

SHIMIZU, Sanae; YAMAMOTO, Kazuhiko; WANG, Caihua; SATO, Yutaka; TANAHASHI, Hideki; NIWA, Yoshinori. Moving object detection with mobile stereo omni-directional system (sos) based on motion compensatory inter-frame depth subtraction. In: **Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on**. 2004. v. 3, p. 248–251 Vol.3. ISSN 1051-4651.

SIAGIAN, Christian; ITTI, Laurent. Rapid biologically-inspired scene classification using features shared with visual attention. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 29, n. 2, p. 300–312, 2007.

SIEGWART, Roland; NOURBAKHSH, Illah Reza; SCARAMUZZA, Davide. **Introduction to autonomous mobile robots**. : MIT press, 2011.

SILVA, Luciana L; TRONCO, Mario Luiz; VIAN, Henrique A; PELLINSON, Giovana; PORTO, Arthur JV. Environment mapping for mobile robots navigation using hierarchical neural network and omnivision. In: **Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on**. 2008. p. 3292–3297. ISSN 1098-7576.

SIMHON, Saul; DUDEK, Gregory. A global topological map formed by local metric maps. In: IEEE. **Intelligent Robots and Systems, 1998. Proceedings., 1998 IEEE/RSJ International Conference on**. 1998. v. 3, p. 1708–1714.

SINGER, Peter W. Military robots and the laws of war. **The New Atlantis**, v. 27, p. 27–47, 2009.

SIROVICH, Lawrence; KIRBY, Michael. Low-dimensional procedure for the characterization of human faces. **JOSA A**, Optical Society of America, v. 4, n. 3, p. 519–524, 1987.

SISTLER, F. Robotics and intelligent machines in agriculture. **Robotics and Automation, IEEE Journal of**, IEEE, v. 3, n. 1, p. 3–6, 1987.

SMITH, Randall C; CHEESEMAN, Peter. On the representation and estimation of spatial uncertainty. **The international journal of Robotics Research**, Sage Publications, v. 5, n. 4, p. 56–68, 1986.

SORIA, C.; CARELLI, R.; Bastos Filho, T.; Sarcinelli Filho, M. Control de robots celulares en base a visión artificial. 2003.

SOUSA, Rafael Vieira de; SILVA, Andrea Ribari Yoshizawa da; INAMASU, Ricardo Yassushi; PORTO, Arthur José Vieira. Robô agrícola móvel (ram): uma revisão das pesquisas recentes sobre sistemas de navegação autônoma de robôs e veículos agrícolas. In: **IV Congresso Brasileiro da Sociedade Brasileira de Informática Aplicada à Agropecuária e à Agroindústria. Anais.** 2003.

SOUTO, Roberto Pinto. Segmentação de imagem multiespectral utilizando-se o atributo matiz. **São José dos Campos: INPE**, 2000.

SPACEK, Libor. A catadioptric sensor with multiple viewpoints. **Robotics and Autonomous Systems**, Elsevier, v. 51, n. 1, p. 3–15, 2005.

SRINIVASAN, Venkatesh V; VENKATESH, Svetha. **From Living Eyes to Seeing Machines.** : OUP Oxford, 1997. ISBN 9780198577850.

STEINER, Maria Teresinha Arns. **Uma metodologia para o reconhecimento de padrões multivariados com resposta dicotômica.** Tese (Doutorado) — Universidade Federal de Santa Catarina, Centro Tecnológico., 1995.

STRAUSS, Olivier; COMBY, Frédéric. Variable structuring element based fuzzy morphological operations for single viewpoint omnidirectional images. **Pattern Recognition**, Elsevier, v. 40, n. 12, p. 3578–3596, 2007.

SVOBODA, Tom as; PAJDLA, Tom as. Central panoramic cameras: Geometry and design. Citeseer, 1997.

SWAMINATHAN, Rahul; GROSSBERG, Michael D; NAYAR, Shree K. Non-single viewpoint catadioptric cameras: Geometry and analysis. **International Journal of Computer Vision**, Springer, v. 66, n. 3, p. 211–229, 2006.

TAKAI, Ryosuke; YANG, Liangliang; NOGUCHI, Noboru. Development of a crawler-type robot tractor using rtk-gps and imu. **Engineering in Agriculture, Environment and Food**, Elsevier, v. 7, n. 4, p. 143–147, 2014.

TAN, Kar-Han; HUA, Hong; AHUJA, Narendra. Multiview panoramic cameras using mirror pyramids. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 26, n. 7, p. 941–946, 2004.

THEODORIDIS, Sergios; KOUTROUMBAS, Konstantinos. **Pattern recognition.** 4st. ed. : Springer, 2009. 13-39 p.

THRUN, Sebastian; BURGARD, Wolfram; FOX, Dieter. Probabilistic robotics (intelligent robotics and autonomous agents series). **Intelligent robotics and autonomous agents**, 2005.

THRUN, Sebastian; FOX, Dieter; BURGARD, Wolfram; DELLAERT, Frank. Robust monte carlo localization for mobile robots. **Artificial intelligence**, Elsevier, v. 128, n. 1, p. 99–141, 2001.

- THRUN, Sebastian; GUTMANN, Jens-Steffen; FOX, Dieter; BURGARD, Wolfram; KUIPERS, Benjamin. Integrating topological and metric maps for mobile robot navigation: A statistical approach. In: **AAAI/IAAI**. 1998. p. 989–995.
- THRUN, Sebastian; LEONARD, John J. Simultaneous localization and mapping. In: **Springer handbook of robotics**. : Springer, 2008. p. 871–889.
- THRUN, Sebastian *et al.* Robotic mapping: A survey. **Exploring artificial intelligence in the new millennium**, Morgan Kaufmann San Francisco, CA, p. 1–35, 2002.
- TIAN, Guo-hui; LI, Xiao-lei; ZHAO, Shou-peng; LU, Fei. Research and development of intelligent space technology for home service robot. **Journal of Shandong University (Engineering Science)**, v. 37, n. 5, p. 53–59, 2007.
- TOMATIS, Nicola; NOURBAKHSH, Illah; SIEGWART, Roland. Simultaneous localization and map building: A global topological model with local metric maps. In: IEEE. **Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on**. 2001. v. 1, p. 421–426.
- TOMATIS, Nicola; NOURBAKHSH, Illah; SIEGWART, Roland. Hybrid simultaneous localization and map building: a natural integration of topological and metric. **Robotics and Autonomous systems**, Elsevier, v. 44, n. 1, p. 3–14, 2003.
- TOMONO, Masahiro; YUTA, Shinichi. Mobile robot localization based on an inaccurate map. In: IEEE. **Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on**. 2001. v. 1, p. 399–404.
- TORFEH-ISFAHANI, M.; AKIMOTO, Y.; TANAKA, H.; OGI, H. Development of an intelligent vision system. In: **Industrial Electronics Society, 1989. IECON '89., 15th Annual Conference of IEEE**. 1989. p. 681–684 vol.3.
- TORRALBA, Antonio; MURPHY, Kevin P; FREEMAN, William T; RUBIN, Mark A. Context-based vision system for place and object recognition. In: IEEE. **Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on**. 2003. p. 273–280.
- VANDORPE, Jurgen; BRUSSEL, Hendrik Van; XU, Hong. Exact dynamic map building for a mobile robot using geometrical primitives produced by a 2d range finder. In: IEEE. **Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on**. 1996. v. 1, p. 901–908.
- VASSALLO, Raquel Frizera; ENCARNACAO, Lucas Frizera; VICTOR, JS; SCHNEEBELI, Hans Jörg. Bird's eye view remapping and path following based on omnidirectional vision. In: **Processing of the XV Automatic Brazilian Conference**. 2004.
- VIDAL, Fábio Silveira; ROSA, Paulo Fernando Ferreira; NETO, Adão de Melo; OLIVEIRA, Thiago Esutaquio Alves de. Multilayer perceptron use in a mapping task by cooperating robots. In: IEEE. **Computational Intelligence and 11th Brazilian Congress on Computational Intelligence (BRICS-CCI & CBIC), 2013 BRICS Congress on**. 2013. p. 580–585.

- VOTH, Danna. A new generation of military robots. **Intelligent Systems, IEEE, IEEE**, v. 19, n. 4, p. 2–3, 2004.
- WAGNER, Israel A; LINDENBAUM, Michael; BRUCKSTEIN, Alfred MF. Distributed covering by ant-robots using evaporating traces. **Robotics and Automation, IEEE Transactions on**, v. 15, n. 5, p. 918–933, Oct 1999. ISSN 1042-296X.
- WAHLGREN, Christoffer; DUCKETT, Tom. Topological map building for mobile robots using omnidirectional vision. In: CITESEER. **Proceedings of SWAR, 3th Swedish Workshop on Autonomous Robotics**. 2005. p. 28–39.
- WALLBAUM, Michael; WASCH, Torsten. Markov localization of wireless local area network clients. In: **Wireless on-Demand Network Systems**. : Springer, 2004. p. 1–15.
- WANG, Jian; HU, Andong; LIU, Chunyan; LI, Xin. A floor-map-aided wifi/pseudo-odometry integration algorithm for an indoor positioning system. **Sensors**, v. 15, n. 4, p. 7096–7124, 2015. ISSN 1424-8220.
- WANG, Min-Liang; WU, Jing-Ren; KAO, Li-Wei; LIN, Huei-Yung. Development of a vision system and a strategy simulator for middle size soccer robot. In: **Advanced Robotics and Intelligent Systems (ARIS), 2013 International Conference on**. 2013. p. 54–58.
- WEI, Yuan Long; LEE, Min Cheol. Mobile robot autonomous navigation using mems gyro north finding method in global urban system. In: **Mechatronics and Automation (ICMA), 2011 International Conference on**. 2011. p. 91–96. ISSN 2152-7431.
- WEISBIN, Charles R *et al.* Miniature robots for space and military missions. 1999.
- WELDON, Thomas P; HIGGINS, William E; DUNN, Dennis F. Efficient gabor filter design for texture segmentation. **Pattern Recognition**, Elsevier, v. 29, n. 12, p. 2005–2015, 1996.
- WEN, Zhi-Ping; ZHONGRU-WU; SU, Huai-Zhi. Safety monitoring system of dam based on bionics. In: **Machine Learning and Cybernetics, 2004. Proceedings of 2004 International Conference on**. 2004. v. 2, p. 1099–1104 vol.2.
- WINTERS, Niall; SANTOS-VICTOR, José. Mobile robot navigation using omni-directional vision. In: **Proc. 3rd Irish Machine Vision and Image Processing Conference (IMVIP99)**. 1999. p. 151–166.
- WINTERS, Niall; SANTOS-VICTOR, Jose. Omni-directional visual navigation. In: CITESEER. **Proc. of the 7th Int. Symp. on Intelligent Robotic Systems**. 1999. p. 109–118.
- WOLF, Jürgen; BURGARD, Wolfram; BURKHARDT, Hans. Robust vision-based localization by combining an image-retrieval system with monte carlo localization. **Robotics, IEEE Transactions on**, IEEE, v. 21, n. 2, p. 208–216, 2005.
- WONG, Robert Y; HALL, Ernest L. Scene matching with invariant moments. **Computer Graphics and Image Processing**, Elsevier, v. 8, n. 1, p. 16–24, 1978.

WONGPHATI, Mahisorn; NIPARNAN, Nattee; SUDSANG, Attawith. Bearing only fastslam using vertical line information from an omnidirectional camera. In: IEEE. **Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference on**. 2009. p. 1188–1193.

XUE, Yang; YANG, Jun Tao; DONG, Ya Ling; SHEN, Jia Li; PENG, Ru; HUANG, Ri Bo; DING, Jie. A new approach for autonomous robot obstacle avoidance using psd infrared sensor combined with digital compass. **Applied Mechanics and Materials**, Trans Tech Publ, v. 511, p. 101–104, 2014.

YAGI, Yasushi; KAWATO, Shinjiro. Panorama scene analysis with conic projection. In: IEEE. **Intelligent Robots and Systems' 90. Towards a New Frontier of Applications', Proceedings. IROS'90. IEEE International Workshop on**. 1990. p. 181–187.

YAGI, Yasushi; KAWATO, Shinjiro. Real-time omnidirectional image sensor (copis) for vision-guided navigation. **Robotics and Automation, IEEE Transactions on**, v. 10, n. 1, p. 11–22, Feb 1994. ISSN 1042-296X.

YAGI, Yasushi; NISHIZAWA, Yoshimitsu; YACHIDA, Masahiko. Map based navigation of the mobile robot using omnidirectional image sensor copis. In: **Robotics and Automation, 1992. Proceedings., 1992 IEEE International Conference on**. 1992. p. 47–52 vol.1.

YAGI, Yasushi; NISHIZAWA, Yoshimitsu; YACHIDA, Masahiko. Map-based navigation for a mobile robot with omnidirectional image sensor copis. **Robotics and Automation, IEEE Transactions on**, IEEE, v. 11, n. 5, p. 634–648, 1995.

YAGI, YASHIDA; YACHIDA, Masahiko. Real-time generation of environmental map and obstacle avoidance using omnidirectional image sensor with conic mirror. In: IEEE. **Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on**. 1991. p. 160–165.

YAMAZAWA, Kazumasa; YAGI, Yasushi; YACHIDA, Masahiko. Omnidirectional imaging with hyperboloidal projection. In: IEEE. **Intelligent Robots and Systems' 93, IROS'93. Proceedings of the 1993 IEEE/RSJ International Conference on**. 1993. v. 2, p. 1029–1034.

YAMAZAWA, Kazumasa; YAGI, Yasushi; YACHIDA, Masahiko. Obstacle detection with omnidirectional image sensor hyperomni vision. In: IEEE. **Robotics and Automation, 1995. Proceedings., 1995 IEEE International Conference on**. 1995. v. 1, p. 1062–1067.

YASUSHI, Yagi. **Omnidirectional Sensing and Its Applications**. 1999.

YASUSHI, Yagi. Omnidirectional sensing and its applications. **IEICE Transactions on Information and Systems**, The Institute of Electronics, Information and Communication Engineers, v. 82, n. 3, p. 568–579, 1999.

YIN, Xu-Cheng; LIU, Chang-Ping; HAN, Zhi. Feature combination using boosting. **Pattern Recognition Letters**, Elsevier, v. 26, n. 14, p. 2195–2205, 2005.

YIP, Hon Ming; HO, Kin Kwan; CHU, Man Hin Anson; LAI, King WC. Development of an omnidirectional mobile robot using a rgb-d sensor for indoor navigation. In: IEEE. **Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), 2014 IEEE 4th Annual International Conference on**. 2014. p. 162–167.

YOU, Bum-Jae; HWANGBO, Myung; LEE, Sung-On; OH, Sang-Rok; KWON, Young Do; LIM, San. Development of a home service robot 'issac'. In: IEEE. **Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on**. 2003. v. 3, p. 2630–2635.

ZHANG, Jisheng. Autonomous navigation for an unmanned mobile robot in urban areas. In: **Mechatronics and Automation (ICMA), 2011 International Conference on**. 2011. p. 2243–2248. ISSN 2152-7431.

ZHANG, Lei; ZAPATA, René; LÉPINAY, Pascal. Self-adaptive monte carlo localization for mobile robots using range finders. **Robotica**, Cambridge Univ Press, v. 30, n. 02, p. 229–244, 2012.

ZHANG, Yani; WEN, Changyun; ZHANG, Ying. Estimation of motion parameters from blurred images. **Pattern Recognition Letters**, Elsevier, v. 21, n. 5, p. 425–433, 2000.

ZHAO, Teng; WANG, Ying. A neural-network based autonomous navigation system using mobile robots. In: **Control Automation Robotics Vision (ICARCV), 2012 12th International Conference on**. 2012. p. 1101–1106.

ZHENG, Dezhi; ZHANG, Shaobo; WANG, Shuai; HU, Chun; ZHAO, Xiaomeng. A capacitive rotary encoder based on quadrature modulation and demodulation. **Instrumentation and Measurement, IEEE Transactions on**, IEEE, v. 64, n. 1, p. 143–153, 2015.

ZHU, Qiang; YEH, M-C; CHENG, Kwang-Ting; AVIDAN, Shai. Fast human detection using a cascade of histograms of oriented gradients. In: IEEE. **Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on**. 2006. v. 2, p. 1491–1498.

ZIVKOVIC, Zoran; BOOIJ, Olaf. How did we built our hyperbolic mirror omnidirectional camera-practical issues and basic geometry. **University of Amsterdam, Tech. Rep. IAS-UVA-05-04**, 2005.

ZOUHAL, Lalla Merieme; DENOEU, Thierry. An evidence-theoretic k-nn rule with parameter optimization. **Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on**, IEEE, v. 28, n. 2, p. 263–271, 1998.

APÊNDICE A – Apêndice

APÊNDICE A - Ambiente sintético

O código utilizado para gerar o ambiente sintético no simulador *POV-Ray* pode ser visualizado em: <http://migre.me/p0tep> e o conjunto de imagens do ambiente sintético pode ser visto em: <http://migre.me/pW1U0>.

APÊNDICE B - Ambiente real

Alguns detalhes construtivos acerca do sistema catadióptrico montado para o sistema são mostrados nas Figuras 58 e 59: suporte utilizado para o espelho e representação do espelho respectivamente. As imagens capturadas do ambiente real são representações de nove espaços em uma residência. Um exemplo de cada um desses espaços é mostrado na Figura 60.

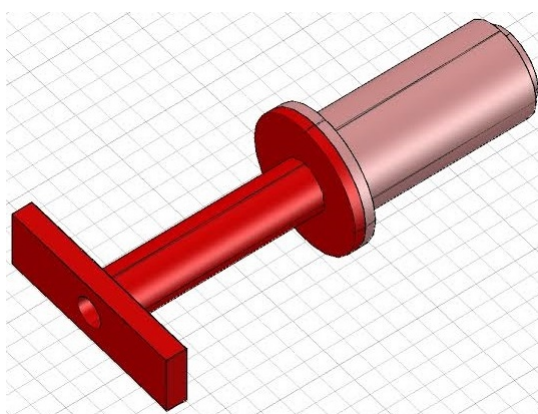


Figura 58: Suporte para o sistema catadióptrico.

O banco de imagens do ambiente real, obtido a partir da navegação do robô mostrado na Figura 61, pode ser visualizado em: <http://migre.me/p0tvL>.

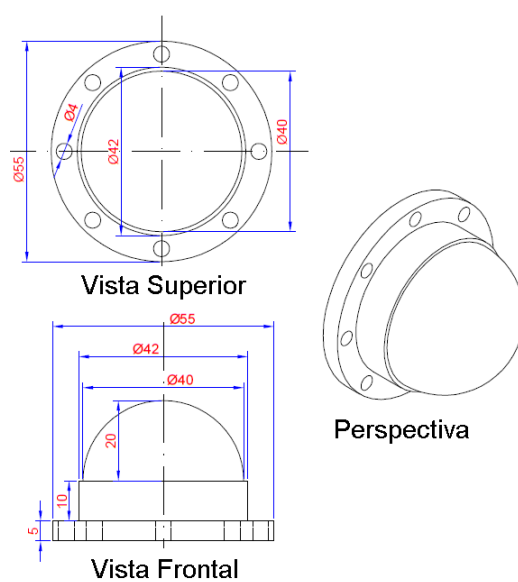


Figura 59: Dimensões do espelho usinado para o sistema.

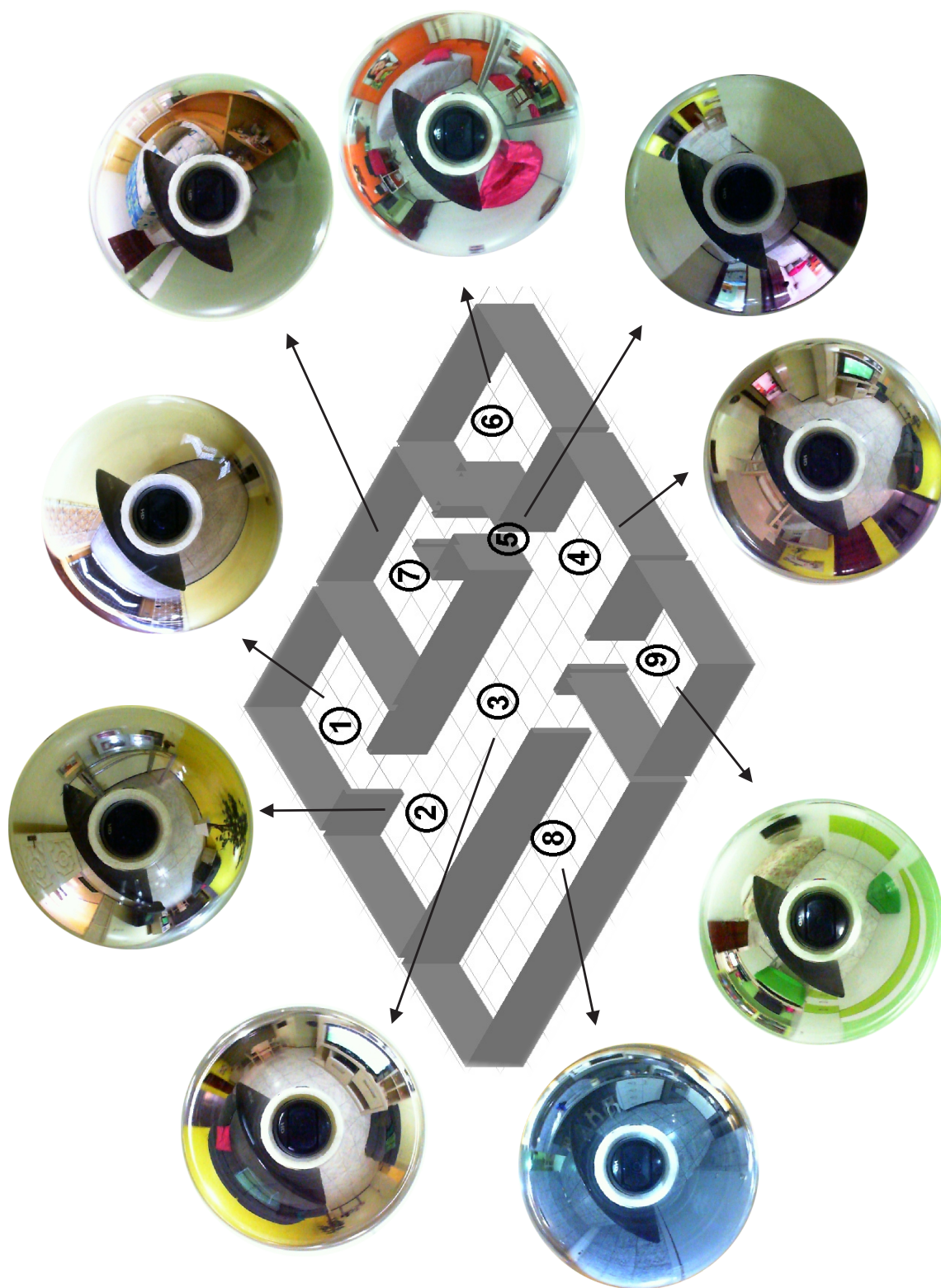


Figura 60: Exemplos de imagens omnidirecionais típicas dos compartimentos do ambiente real.

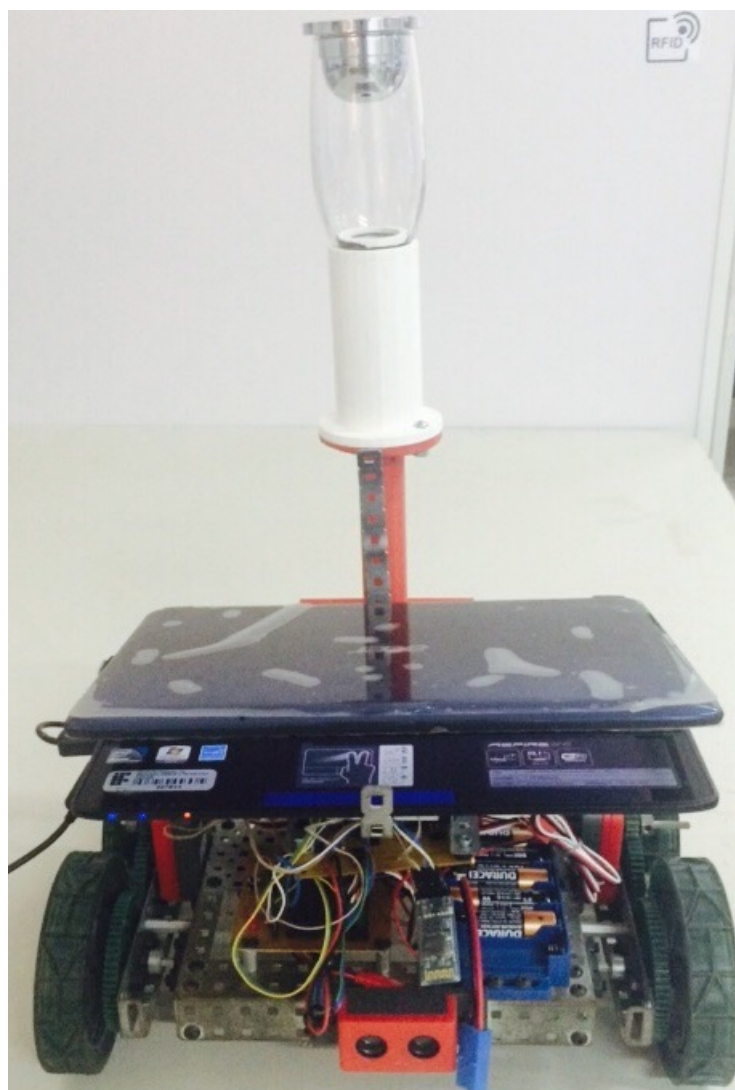


Figura 61: Vista do robô com o sistema catadióptrico.