

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL
FACULDADE DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**SISTEMA HÍBRIDO DE LOCALIZAÇÃO INDOOR
BASEADO EM RFID E ANÁLISE VISUAL**

EVERTON LUÍS BERZ

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Ciência da Computação na Pontifícia Universidade Católica do Rio Grande do Sul.

Orientador: Prof. Dr. Fabiano Passuelo Hessel

Porto Alegre
2015

Dados Internacionais de Catalogação na Publicação (CIP)

B553s Berz, Everton Luís

Sistema híbrido de localização indoor baseado em RFID e análise visual / Everton Luís Berz. – Porto Alegre, 2015.
155 p.

Tese (Doutorado) – Faculdade de Informática, PUCRS.
Orientador: Prof. Dr. Fabiano Passuelo Hessel.

1. Informática. 2. Sistemas Automáticos de Identificação. 3. RFID.
I. Hessel, Fabiano Passuelo. II. Título.

CDD 006.31

**Ficha Catalográfica elaborada pelo
Setor de Tratamento da Informação da BC-PUCRS**



TERMO DE APRESENTAÇÃO DE TESE DE DOUTORADO

Tese intitulada "Sistema Híbrido de Localização Indoor Baseado em RFID e Análise Visual" apresentada por Everton Luís Berz como parte dos requisitos para obtenção do grau de Doutor em Ciência da Computação, aprovada em 30/03/2015 pela Comissão Examinadora:

Prof. Dr. Fabiano Passuelo Hessel-
Orientador

PPGCC/PUCRS

Prof. Dr. César Augusto Missio Marcon-

PPGCC/PUCRS

Prof. Dr. Dario Francisco Guimarães de Azevedo-

PPGEE/PUCRS

Prof. Dr. Carlos Eduardo Pereira-

UFRGS

Homologada em 18/06/2015, conforme Ata No. 010 pela Comissão Coordenadora.

Prof. Dr. Luiz Gustavo Leão Fernandes
Coordenador.

PUCRS

Campus Central

Av. Ipiranga, 6681 - P32- sala 507 - CEP: 90619-900

Fone: (51) 3320-3611 - Fax (51) 3320-3621

E-mail: ppgcc@pucrs.br

www.pucrs.br/facin/pos

DEDICATÓRIA

Dedico este trabalho a minha mãe, Erica Reny Berz (*in memoriam*).

AGRADECIMENTOS

Agradeço a todos que participaram e se envolveram de alguma forma com este trabalho.

Ao meu orientador Prof. Dr. Fabiano Passuelo Hessel pela confiança, compreensão e direcionamento nesta jornada, além da oportunidade oferecida para o desenvolvimento deste trabalho.

Aos professores/colegas/amigos Marcelo Azambuja e Carlos Fernando Jung pelas discussões, resolução de dúvidas e o incentivo dado ao longo destes anos.

Ao colega de equipe Deivid Tesch pela parceria na escrita de artigos, no compartilhamento de dúvidas e conhecimentos e por toda a ajuda nos ensaios realizados.

Aos amigos Julio Ody e Thomas Müller pelas colaborações durante o período de doutoramento.

Aos colegas da FACCAT, TJ-RS e TRT-4 pelo apoio e suporte fornecidos durante esta jornada.

Aos meus pais, Erica (*in memoriam*) e Valdir, por toda a dedicação, pelos ensinamentos e educação concebidos, que certamente foram de fundamental importância para eu chegar onde estou hoje. Infelizmente, mãe Erica só acompanhou parte da minha jornada no doutorado, dedico este trabalho a toda sua simplicidade, humildade, carisma e cuidado que tinha com todos, o que me ensinou a dar importância às pessoas e às pequenas coisas da vida.

Por fim, à Daniela, minha esposa, por estar sempre ao meu lado, pelo apoio incondicional, por todo o carinho e amor. Foram quatro anos de companheirismo, onde algumas coisas tiveram que ser deixadas em segundo plano, mas todas as dificuldades foram enfrentadas e superadas. Mais uma vez foste uma guerreira. Obrigado e espero retribuir da melhor forma possível!

SISTEMA HÍBRIDO DE LOCALIZAÇÃO INDOOR BASEADO EM RFID E ANÁLISE VISUAL

RESUMO

Com o objetivo de otimizar a experiência do usuário e resolver problemas de logística e segurança, aplicativos necessitam conhecer a localização física de objetos e pessoas. As propostas de sistemas de localização para ambientes internos (*indoor*) baseados em somente uma tecnologia não vêm obtendo bom desempenho na localização, principalmente devido a limitações em relação às características interferentes presentes em ambientes internos. Este trabalho tem como objetivo o desenvolvimento de um sistema híbrido de localização *indoor*, capaz de obter a localização bidimensional e tridimensional de objetos estáticos, com precisão de poucos centímetros, utilizando equipamentos já estabelecidos no mercado. A partir do uso da tecnologia RFID, a localização é estimada através de modelos de aprendizado de máquina. São avaliados modelos baseados em Redes Neurais Artificiais (RNA) e *Support Vector Regression* (SVR). Um subsistema de visão computacional detecta marcadores visuais no cenário a fim de aprimorar a localização. Visando combinar as tecnologias RFID e óptica, é definido um método de fusão de sensores baseado em múltiplas regiões de interesse a partir da técnica k-means. Além disso, um método de multi-frequência é proposto com o objetivo de permitir e otimizar o uso do sistema em equipamentos *off-the-shelf*. O sistema foi implementado e avaliado em experimentos de laboratório. Inicialmente, foi analisado o desempenho do sistema em relação ao posicionamento de antenas e etiquetas RFID presentes no ambiente. Na avaliação geral do sistema sob escalabilidade 2D, o erro de localização se manteve entre 9 e 33 cm. Nesta dimensão, o modelo RNA superou o SVR em 30%. Na comparação entre o sistema híbrido e o emprego somente da tecnologia RFID, a fusão de sensores melhorou a precisão em aproximadamente 32%. Em escalabilidade tridimensional, a precisão foi de 63 cm no melhor caso.

Palavras Chave: sistemas de localização; RFID; fusão de sensores; análise visual.

A HYBRID INDOOR POSITIONING SYSTEM BASED ON RFID AND VISUAL ANALYSIS

ABSTRACT

Systems need to know the physical location of objects and people in order to improve the user experience and solve logistic and security problems. This work proposes a hybrid indoor positioning system based on passive RFID and visual analysis. The system focuses on bidimensional and tridimensional space localization of stationary objects with centimeter level accuracy. Also, off-the-shelf equipment are employed as sensor devices. By using RFID technology, machine learning approaches based on Artificial Neural Networks (ANN) and Support Vector Regression (SVR) are proposed. A Computer Vision (CV) subsystem detects visual markers in the scenario to enhance RFID localization. In order to combine RFID and optical technologies, we propose a novel sensor fusion method based on Multiple Region of Interest (MROI) and k-means technique. A multi-frequency method is proposed aimed to allow and improve the localization when using off-the-shelf equipment. We have implemented our system and evaluated it using real experiments. First, a performance benchmark was made in order to evaluate RFID antennas and tags positioning. Regarding to overall system performance, the localization error was between 9 and 33 cm under a 2D scenario. On this dimension, ANN performed 30% better than RNA approach. In comparison to RFID-only approach results, the hybrid system had improved by 32%. Three-dimensional space localization had a 63 cm accuracy in best case scenario.

Keywords: indoor positioning systems; RFID; sensor fusion; visual analysis.

LISTA DE FIGURAS

Figura 1 – Arquitetura em camadas associado a um diagrama de blocos contendo os componentes fundamentais de um IPS. Fonte: [35], [70].	30
Figura 2 – Para obter a posição 2D utilizando lateração, é necessário medir a distância entre o objeto “X” e três pontos não-colineares. Fonte: [38].	34
Figura 3 – Exemplo de angulação que localiza o objeto X. Fonte: [38].	36
Figura 4 – Técnica de localização por proximidade. Fonte: [35].	37
Figura 5 – Passos fundamentais em um sistema de visão computacional. Fonte: [32].	39
Figura 6 – Imagem original (150 x 135 <i>pixels</i>) na esquerda, zoom (3200%) de determinada região (7 x 7 <i>pixels</i>) ao centro e níveis de intensidade de cada <i>pixel</i> na direita. Fonte: o autor.	40
Figura 7 – Imagens resultantes em cada etapa do detector de bordas Canny. Fonte: [31], [58].	42
Figura 8 – Componentes de um sistema RFID. Fonte: [28].	45
Figura 9 – Funcionamento do RFID em espaço livre com variáveis do modelo. Fonte: [6].	48
Figura 10 – Gráfico de dispersão e a linha de regressão linear. Fonte: [49].	53
Figura 11 – Modelo não-linear de um neurônio artificial. Fonte: [36].	54
Figura 12 – Rede <i>feedforward</i> totalmente conectada com uma camada oculta e uma camada de saída. Fonte: [36].	56
Figura 13 – Possibilidades de separação através de hiperplanos distintos. Fonte: [2].	60
Figura 14 – Separação ótima de dois conjuntos linearmente separáveis (a) e vetores de suporte (b). Fonte: [2], [8].	60
Figura 15 – Violação das margens de classificação no lado correto (a) e no lado incorreto da superfície de decisão (b). Fonte: [8].	61
Figura 16 – Mapeamento não-linear do espaço de entrada para o espaço de características. Fonte: [10].	62
Figura 17 – Função de perda ϵ -insensitive. Fonte: [81].	63
Figura 18 – Arquitetura de uma SVR. Fonte: [19], [81].	64
Figura 19 – Resultado ilustrado do SVR com duas funções <i>kernel</i> distintas. Fonte: [8].	64

Figura 20 – Resultado de um agrupamento k-means. Os pontos com formas geométricas iguais estão no mesmo <i>cluster</i> . O centro dos <i>clusters</i> é indicado com uma cruz. Fonte: o autor.....	65
Figura 21 - Ambiente de experimentos do Landmarc: dimensões do cenário e organização dos componentes. Fonte: [66].	66
Figura 22 – Rede neural proposta (a) e configuração de ambiente dos experimentos (b). Fonte: [12].....	67
Figura 23 – Desenho do recipiente plástico e posicionamento das antenas RFID (a) e o cubo (verde) representando as amostras coletadas nos experimentos (b). Fonte: [96].	68
Figura 24 – Ambiente interno dividido em regiões e os pesos (W) das etiquetas afixadas. Fonte: [11].	70
Figura 25 – Na parte superior, fotografias do ambiente destacando o objeto selecionado pelo usuário. Na parte inferior, imagens geradas a partir do nível de intensidade do sinal recebido pela antena RFID. Fonte: [18].	72
Figura 26 – Leiaute do ambiente na técnica empregada na localização RFID. Fonte: [93].....	74
Figura 27 – Diagrama esquemático da localização RFID (a), visual (b) e o método de fusão proposto (c). Fonte: [93].	75
Figura 28 – Padrão triangular das etiquetas de referência, cobertura do leitor RFID e extração da região de interesse. Fonte: [7].	76
Figura 29 – Arquitetura do sistema de localização baseado em WISPs. Fonte: [76].	77
Figura 30 – Padrão de irradiação (a) e respectiva aproximação Gaussiana (b) do nível de intensidade de sinal medido. Fonte: [30].	79
Figura 31 – Arquitetura em três camadas associada a um diagrama de blocos contendo os componentes fundamentais do IPS proposto. Fonte: o autor.	85
Figura 32 – Diagrama de blocos da arquitetura do sistema. As entradas são as leituras RFID e a imagem capturada. A saída é a posição do objeto alvo. Fonte: o autor.	86
Figura 33 – Exemplo de posicionamento das etiquetas de referência e antenas (a) e grade virtual sobre uma fotografia capturada do cenário (b). Fonte: o autor.	88
Figura 34 – Método de multi-frequência: criação de instâncias do modelo de aprendizado para cada faixa de frequência detectada na coleta de dados. Fonte: o autor.	89

Figura 35 – Amostras coletadas individualmente pelo leitor RFID (esquerda) e agrupamento de dados (direita). Frequência: 923,25 MHz. Fonte: o autor.....	90
Figura 36 – Estrutura da rede neural artificial (escalabilidade 2D). Fonte: o autor..	92
Figura 37 – Simplificação da modelagem SVR aplicada ao problema de localização. Fonte: o autor.	93
Figura 38 – Técnica k-means aplicada às posições estimadas sobre as coordenadas (x, y) da grade virtual. Fonte: o autor.	97
Figura 39 – Algoritmo contendo a lógica de funcionamento do subsistema visual. Fonte: o autor.	99
Figura 40 – Exemplo de códigos de barras Code128 e Code39. Fonte: [84].....	101
Figura 41 – Etapas do subsistema visual alinhadas aos passos fundamentais de um sistema de visão computacional. Fonte: o autor.	103
Figura 42 – Subimagem recortada da fotografia original do cenário e a detecção do marcador visual. Fonte: o autor.	104
Figura 43 – Captura de tela do processo de configuração das posições conhecidas e resolução da grade virtual. Os ícones amarelos representam os marcadores conhecidos e seus respectivos EPCs. Fonte: o autor.....	107
Figura 44 - Probabilidade de sucesso para cada tipo de marcador visual entre 20 e 220 centímetros. Fonte: o autor.....	109
Figura 45 – Imagem do código de barras capturado. Zoom: 1.200%. Distância: 1 m. Fonte: o autor.	110
Figura 46 - Antenas e suas posições em relação ao quadro onde os marcadores são afixados. Fonte: o autor.	111
Figura 47 – Ambiente de testes e todos os componentes do sistema. Formato de posicionamento das antenas: EIXO. Fonte: o autor.	112
Figura 48 – Ilustração do cenário contemplando as etiquetas de referência (quadrados cinzas) e os marcadores alvo da fase <i>online</i> (triângulos vermelhos). Fonte: o autor.....	113
Figura 49 – MSE do treinamento para todas as instâncias de modelo RNA na avaliação do posicionamento das antenas. Fonte: o autor.	114
Figura 50 – Desempenho do treinamento das redes neurais (MSE em coordenadas da grade virtual) em relação às épocas do algoritmo de retro-propagação. Fonte: o autor.	115
Figura 51 – Erro de localização acumulado (CDF) para os três posicionamentos avaliados. Fonte: o autor.	117

Figura 52 – Níveis de intensidade das etiquetas de referência para cada antena e formato de posicionamento. Os eixos dos gráficos representam as coordenadas na grade virtual. Fonte: o autor.	118
Figura 53 – Formas de organização das etiquetas de referência contendo 5, 9 e 13 elementos, respectivamente. Fonte: o autor.	120
Figura 54 - Erro de localização acumulado (CDF) das três formas de organização de etiquetas. Fonte: o autor.	122
Figura 55 – <i>Fingerprint</i> RSSI das etiquetas de referência para cada antena e estratégia de organização. Frequência: 924,25 MHz. Fonte: o autor.	122
Figura 56 - MSE do treinamento para todas as instâncias de modelo RNA nos experimentos de tempo de leitura das etiquetas. Fonte: o autor.	124
Figura 57 – Capturas de tela do sistema após tentativa de localização nos testes MU (a) e MF (b). Modo: “somente RFID”; modelo: RNA; marcador alvo: 15; precisão: 43 cm (a) e 8 cm (b). Fonte: o autor.	127
Figura 58 - Cenários avaliados e o deslocamento utilizado no treinamento e nas estimativas do sistema. Fonte: o autor.	128
Figura 59 - Capturas de tela dos quatro cenários e amostras das dimensões do ROI. Fonte: o autor.	129
Figura 60 - MSE do treinamento para as instâncias RNA na avaliação geral do sistema híbrido. Fonte: o autor.	130
Figura 61 - Desempenho do treinamento da rede neural em relação às épocas do algoritmo de retro-propagação. Cenário: C3; frequência: 923,75 MHz. Fonte: o autor. ..	131
Figura 62 – Captura de tela do sistema em execução. Cenário: C2; teste: denso; modelo: RNA; marcador alvo: 15; precisão: 0 cm. Fonte: o autor.	131
Figura 63 – Detalhe das estimativas de localização do subsistema RFID (a) e do teste “denso” (b). Cenário: C4; modelo: RNA; marcador: 16; precisão: 24 cm (a) e 29 cm (b). Fonte: o autor.	133
Figura 64 - Erro de localização acumulado (CDF) para ambos os modelos nos três testes de validação. Escalabilidade: 2D. Fonte: o autor.	133
Figura 65 – Erro acumulado (CDF) na localização individual de cada marcador alvo. Medições contemplam todos os cenários, modelos e testes de validação. Fonte: o autor.	134
Figura 66 – <i>Fingerprint</i> RSSI das etiquetas de referência em ambas as antenas. A posição centroide de cada <i>cluster</i> k-means é exibida nas caixas de texto (etiqueta 15).	

<i>Clusters</i> com maior peso possuem cores escuras. Freq.: 923,75 MHz; cenário: C3. Fonte: o autor.....	135
Figura 67 - Cenário avaliado sob escalabilidade tridimensional. Antena em destaque foi adicionada para contemplar o eixo z. Fonte: o autor.....	136
Figura 68 - Exemplo de experimento em escalabilidade 3D. Teste “limpo”; modelo: SVR; distância z real: 220 cm; marcador alvo: 5; precisão: 144,7 cm. Tela do sistema exibindo o resultado (a) e estimativas RFID e regiões de interesse na visão 2D e 3D em (b) e (c), respectivamente. Fonte: o autor.....	138
Figura 69 - Erro de localização acumulado (CDF) para ambos os modelos de aprendizado nos três testes de validação. Escalabilidade: 3D. Fonte: o autor.	139
Figura 70 – Erro acumulado (CDF) para cada eixo no espaço de localização. Medições contemplam todos os cenários, modelos de aprendizado e testes de validação. Fonte: o autor.	140
Figura 71 - Regiões de vizinhança (máscaras) de tamanhos diferentes, porém, todos com lado ímpar. Fonte: [55].	153
Figura 72 - Esquema da operação de correlação. Níveis de intensidade dos <i>pixels</i> na imagem original (a); máscara utilizada (b); produtos dos valores da máscara pelos valores dos <i>pixels</i> da imagem, ponto a ponto (c); somatória dos valores multiplicados (d); e, finalmente, o novo valor do <i>pixel</i> na imagem após a correlação (e). Fonte: [55].	154
Figura 73 - Máscara gaussiana para um filtro do tipo passa-baixa. Fonte: [55]. ...	155

LISTA DE TABELAS

Tabela 1 – Tecnologias de localização <i>indoor</i> e a precisão típica correspondente [54].....	30
Tabela 2 – Características dos tipos de etiquetas RFID [4].	46
Tabela 3 – Visão geral das técnicas de localização utilizadas nos trabalhos correlatos.	81
Tabela 4 – Características dos sistemas propostos nos trabalhos correlatos analisados.....	82
Tabela 5 – Exemplo de lista de MROI fornecida pela fase k-means para o subsistema visual.	98
Tabela 6 – Componentes e respectivas propriedades que permaneceram fixos em todas as seções de experimentos.	106
Tabela 7 – Características da câmera digital utilizada no subsistema visual.	106
Tabela 8 – Propriedades das configurações primárias do sistema empregados nos testes iniciais.	113
Tabela 9 – Desempenho da localização (RMSE em cm) para cada formato de posicionamento.....	116
Tabela 10 – Média do RSSI das etiquetas alvo em determinadas frequências para cada antena e estratégia de posicionamento.	119
Tabela 11 – Desempenho da localização (RMSE em cm) para cada estratégia de organização das etiquetas de referência.	121
Tabela 12 – Quantidade de amostras, desempenho do treinamento (MSE) e da localização (RMSE em cm) para a avaliação do tempo de leitura das etiquetas nas fases <i>offline</i> e <i>online</i>	124
Tabela 13 – Propriedades das configurações primárias definidas após série de experimentos.	126
Tabela 14 – Desempenho da localização (RMSE em cm) na avaliação do método de multi-frequência desenvolvido.	126
Tabela 15 – Desempenho da localização (RMSE em cm) para cada cenário, teste de validação e abordagem de aprendizado de máquina.	132
Tabela 16 - Desempenho da localização (RMSE em cm) sob escalabilidade tridimensional.....	137

LISTA DE SIGLAS

ANATEL	Agência Nacional de Telecomunicações
AOA	<i>Angle of Arrival</i>
CDF	<i>Cumulative Distribution Function</i>
CUKF	<i>Constrained Unscented Kalman Filter</i>
CV	<i>Computer Vision</i>
EPC	<i>Electronic Product Code</i>
FCC	<i>Federal Communications Commission</i>
GPS	<i>Global Positioning System</i>
GUI	<i>Graphical User Interface</i>
IPS	<i>Indoor Positioning System</i>
IQR	<i>Interquartile Range</i>
kNN	<i>k-Nearest Neighbors</i>
LLRP	<i>Low Level Reader Protocol</i>
LOS	<i>Line-of-Sight</i>
MROI	<i>Multiple Region of Interest</i>
MSE	<i>Mean Square Error</i>
POA	<i>Phase of Arrival</i>
RFID	<i>Radio Frequency Identification</i>
RMSE	<i>Root Mean Square Error</i>
RNA	Redes Neurais Artificiais
ROI	<i>Region of Interest</i>
RSP	<i>Received Signal Phase</i>
RSSI	<i>Received Signal Strength Indicator</i>
RTOF	<i>Round-Trip Time of Flight</i>
SIFT	<i>Scale-Invariant Feature Transform</i>
SVM	<i>Support Vector Machines</i>
SVR	<i>Support Vector Regression</i>
TDOA	<i>Time Difference of Arrival</i>
TPS	<i>Thin-Plate Spline</i>
UHF	<i>Ultra High Frequency</i>
UTM	<i>Universal Transverse Mercator</i>
WISP	<i>Wireless Identification and Sensing Platform</i>

SUMÁRIO

1	INTRODUÇÃO	23
1.1	Objetivos	26
1.2	Motivações	26
1.3	Contribuição Original	28
1.4	Organização	28
2	CONCEITOS E DEFINIÇÕES	29
2.1	Sistemas de Localização <i>Indoor</i>	29
2.1.1	Técnicas de Localização	33
2.1.1.1	Triangulação	33
2.1.1.1.1	Lateração.....	34
2.1.1.1.2	Angulação.....	35
2.1.1.2	Proximidade.....	36
2.1.1.3	Análise de Padrões.....	37
2.1.2	Sistemas Ópticos.....	38
2.1.2.1	Visão Computacional.....	39
2.1.2.1.1	Segmentação de Imagens.....	40
2.1.2.2	Marcadores Visuais	43
2.1.2.3	<i>Markerless</i>	44
2.1.3	Identificação por Radiofrequência (RFID).....	44
2.1.3.1	Modelagem do Canal de Comunicação	47
2.1.3.2	Modelo Tradicional de IPS RFID	49
2.1.4	Fusão de Sensores.....	50
2.2	Aprendizado de Máquina	51
2.2.1	Regressão Linear Múltipla	53
2.2.2	Redes Neurais Artificiais (RNA).....	54
2.2.2.1	Arquiteturas de RNA.....	56
2.2.2.2	Aprendizagem da Rede Neural	57
2.2.3	<i>Support Vector Machines</i> (SVM)	59
2.2.3.1	Support Vector Regression (SVR).....	62
2.2.4	K-Means	64
3	TRABALHOS CORRELATOS	66
3.1	IPSS baseados em RFID e Aprendizado de Máquina	66
3.2	Sistemas Híbridos de Localização <i>Indoor</i>	69

3.2.1	Navegação de Robôs Móveis Baseada em RFID e Sensor Visual	69
3.2.2	RF Vision	71
3.2.3	Sistema de Identificação e Localização Baseado em RFID e Análise Visual	73
3.2.4	Otimização de IPS RFID através do Padrão de Movimentação de Objetos ...	75
3.2.5	Localização Óptica Utilizando WISP	76
3.2.6	Localização de Etiquetas RFID Auxiliada por Câmeras	78
3.2.7	Localização Probabilística via Fusão de RFID e Visão Computacional	79
3.2.8	Discussão	80
4	DESCRIÇÃO DO SISTEMA	84
4.1	Arquitetura do Sistema	84
4.2	Fase <i>Offline</i>	87
4.2.1	Multi-Frequência	88
4.2.2	Preparação dos Dados Coletados	90
4.2.3	Modelagem RLM	91
4.2.4	Modelagem RNA	91
4.2.5	Modelagem SVR	93
4.3	Fase <i>Online</i>	94
4.3.1	Eliminação de Dados Espúrios	95
4.3.2	K-Means	96
4.4	Integração com Análise Visual	98
4.4.1	Marcador Visual Codificado	100
4.4.2	Marcador Visual Não-Codificado	102
5	EXPERIMENTOS E RESULTADOS	105
5.1	Tipo de Marcador Visual	108
5.1.1	Resultados	109
5.2	Posicionamento das Antenas	111
5.2.1	Desempenho da Fase <i>Offline</i>	114
5.2.2	Resultados da Fase <i>Online</i>	116
5.3	Organização das Etiquetas de Referência	120
5.3.1	Resultados	121
5.4	Tempo de Leitura das Etiquetas	123
5.4.1	Resultados	123
5.5	Multi-Frequência	125
5.5.1	Resultados	126
5.6	Avaliação Geral do Sistema Híbrido em Escalabilidade 2D	128

5.6.1	Resultados.....	130
5.7	Avaliação do Sistema Híbrido em Escalabilidade Tridimensional	135
5.7.1	Resultados.....	137
6	CONCLUSÃO E TRABALHOS FUTUROS	141
6.1	Contribuições do Trabalho	141
6.2	Publicações.....	142
6.3	Trabalhos Futuros	143
	REFERÊNCIAS BIBLIOGRÁFICAS	145
	APÊNDICE A – EXEMPLO DE AMOSTRAS COLETADAS PELO LEITOR RFID	152
	ANEXO A – VISÃO COMPUTACIONAL: CONVOLUÇÃO E FILTRO GAUSSIANO.	153

1 INTRODUÇÃO

A localização de objetos e pessoas em ambientes internos (*indoor*) vem sendo amplamente pesquisada, principalmente em virtude dos problemas de segurança e também dos benefícios que um sistema de localização pode fornecer. No que tange à localização em ambientes externos (*outdoor*), a tecnologia GPS (*Global Positioning System*) tem crescido cada vez mais, tornando-se o padrão *de facto* no que concerne a localização e navegação de veículos, pessoas e outros objetos. Entretanto, esta tecnologia não é adequada para ambientes internos, pois necessita uma comunicação com visada direta para os satélites. Ambientes internos são mais complexos quando comparados a ambientes externos, devido à grande quantidade de obstáculos e fenômenos interferentes existentes em uma densidade de espaço reduzida. Considerando estes problemas, sistemas de localização focados em ambientes internos trazem novos desafios para o futuro dos sistemas de comunicação [35], [78].

Um sistema de localização para ambientes internos (*Indoor Positioning System – IPS*) é definido como um sistema que pode determinar, continuamente e em tempo real, a posição de alguma coisa ou alguém em um espaço físico como um hospital, uma escola, um ginásio, etc. Um IPS deve oferecer informações de localização atualizadas do alvo, estimar tais posições dentro de um atraso temporal máximo a ser respeitado, e cobrir a área esperada e exigida pelo usuário do sistema [35].

Há uma demanda crescente por aplicações comerciais para residências e organizações, onde IPSs são utilizados para rastrear e guiar pessoas com necessidades especiais, idosos e crianças; localizar equipamentos e medicamentos em hospitais e encontrar itens específicos em centros de distribuição. Na segurança pública e no uso militar, tais sistemas são necessários para rastrear detentos e auxiliar na navegação de policiais, bombeiros e soldados em suas missões dentro de prédios [70].

No que se refere às redes de comunicação, aplicações de telecomunicação necessitam diversos tipos de informações sobre o ambiente, pessoas e dispositivos, visando oferecer serviços cada vez mais flexíveis e adaptativos. No futuro dos sistemas de comunicação, a informação de localização pode trazer diversos benefícios, como, por exemplo, a organização autônoma de sensores ou dispositivos em redes *ad hoc* [35].

Bibliotecas e museus são outros cenários potenciais para utilização de IPSs. Geralmente, sistemas de gestão de bibliotecas somente informam se o livro está

emprestado ou não. Outros sistemas mantêm em uma base de dados estática o número da estante em que o livro deveria estar presente. Entretanto, tais sistemas não conseguem fornecer a localização precisa e dinâmica do livro no acervo. Já em museus, turistas podem se beneficiar de serviços de navegação para consultarem a localização de peças e criarem uma sequência mais atrativa para a visita.

Em relação aos critérios de adoção de IPSs, diversas características destes sistemas podem ser levadas em consideração no processo de avaliação. Tais critérios são descritos com foco nas preferências e experiências dos usuários, verificando se o sistema desejado atende as suas necessidades.

O principal critério de um IPS é o desempenho. A precisão e a probabilidade de sucesso são as duas propriedades mais importantes de desempenho na avaliação deste tipo de sistema. Precisão é o termo que representa o erro médio da distância estimada pelo sistema, enquanto a probabilidade de sucesso é definida como o percentual de estimativas que localizam o objeto alvo sob determinada precisão. Por exemplo, um IPS localiza objetos com erro de 5 metros (precisão) para aproximadamente 95% das tentativas de medição (probabilidade de sucesso) [35], [37], [52].

O custo de um IPS também recebe destaque entre os critérios levados em consideração na adoção da tecnologia, podendo ser classificado em três partes [35], [37], [52]:

- i. Componentes de infraestrutura: alguns sistemas de localização possuem uma infraestrutura bastante grande, o que se torna caro e complexo. Outros sistemas conseguem reutilizar uma infraestrutura já existente, como redes sem-fio já presentes no ambiente, sendo mais atrativos, pois nenhum gasto extra é necessário nesta parte da aplicação.
- ii. Dispositivos de localização: diz respeito ao valor de fabricação (ou de compra, no caso de produto já disponível) do dispositivo utilizado pelo usuário para realizar a localização dos objetos. Neste quesito, também são observados os gastos necessários nos componentes que são colocados nos objetos que se deseja localizar. Além disso, deve ser considerado o custo de manutenção e vida útil dos equipamentos, principalmente em relação à duração da bateria e consumo de energia.
- iii. Instalação e manutenção do sistema: trata-se do custo associado ao processo de fixação de componentes do sistema no ambiente, bem como a

manutenção que tais componentes podem gerar após o sistema estar em funcionamento. O tempo dispendido e o nível de especialização dos profissionais necessários para a realização destas tarefas também devem ser considerados.

Outro aspecto presente em um IPS é a escalabilidade. O conceito de escalabilidade aplicado a sistemas de localização pode ser definido como o número de objetos que o sistema pode localizar a partir de certa quantidade de dispositivos de localização presentes na infraestrutura e contemplando um determinado requisito de tempo. A escalabilidade assegura o funcionamento estável do sistema à medida que seu escopo aumenta. Geralmente, o desempenho do sistema sofre degradação conforme aumenta a distância entre o transmissor e o receptor. Um sistema de localização pode ser escalável sob dois atributos: geográfico e densidade. A escala geográfica representa a área ou volume de cobertura do sistema. A densidade significa o número de dispositivos alocados por área geográfica em determinado período de tempo. Quanto maior for a área de cobertura, ou mais dispositivos forem implantados na área, maior será a emissão de sinais necessários pelos sistemas (podendo até congestionar o tráfego), mais cálculos serão necessários, ou uma infraestrutura maior de comunicação pode ser necessária. Outra medida de escalabilidade é o espaço dimensional do sistema. Os IPSs atuais são capazes de localizar objetos em espaço 2D ou 3D, sendo que alguns trabalham sob ambas as dimensões [35], [52].

Muitos sistemas têm sido desenvolvidos no decorrer dos anos a fim de fornecer estimativas de localização *indoor*. Da mesma forma, diversas tecnologias são empregadas na criação destes sistemas, cada uma com diferentes características físicas e operacionais. São exemplos de tecnologias estabelecidas no uso de IPSs: sistemas ópticos, redes sem fio e identificação por radiofrequência (*Radio Frequency Identification* - RFID), ultrassom, infravermelho e sistemas magnéticos. Em grande parte dos casos, estas tecnologias atuam como sensores do sistema, realizando a percepção do ambiente e sendo integrados a técnicas e algoritmos de localização.

No que diz respeito à localização de objetos em movimento *versus* estáticos, a localização em cenários dinâmicos geralmente possui mais recursos, inerentes ao ambiente, que contribuem no processo de localização. Toma-se como exemplo um cenário dinâmico em que somente o objeto alvo está em movimento e todo o resto permaneça estático. A análise de cena nesse ambiente pode utilizar atributos do objeto

alvo que se alteram em função do tempo, sendo que estas características não existem em cenários estacionários.

Atualmente, carecem pesquisas relativas a IPSs com baixo custo, precisão no nível de centímetros (até 15 cm), escalabilidade tridimensional e aplicados a cenários estáticos. Assim, este trabalho de doutoramento propõe um novo IPS que atenda esses critérios. A proposta também visa definir um método original de fusão de sensores¹ (sistema híbrido), a fim de obter melhores resultados e contemplar os requisitos funcionais necessários na adoção de um IPS.

1.1 Objetivos

O trabalho tem como objetivo principal o desenvolvimento de um sistema híbrido de localização *indoor*, capaz de obter a localização 2D e 3D de objetos estáticos, com precisão de poucos centímetros (até 15 cm), utilizando equipamentos disponíveis no mercado (*off-the-shelf*).

A fim de atingir o objetivo principal, foram traçados os seguintes objetivos específicos:

- i. Buscar e dominar o funcionamento de dispositivos RFID e sensores ópticos disponíveis comercialmente, a fim de integrá-los ao sistema;
- ii. Propor um modelo de localização *indoor* baseado em RFID;
- iii. Propor um modelo de localização *indoor* baseado em sistemas ópticos;
- iv. Definir e criar um método inovador de fusão de sensores, com o intuito de integrar os modelos de localização propostos.

O sistema proposto não está direcionado a um caso de uso específico, permitindo a sua adequação a mais de um tipo de aplicação. São exemplos de aplicações que o sistema pode contemplar: (i) localização de itens em centros de distribuição; (ii) complementação de mecanismos robóticos a fim de coletar itens automaticamente (*object grasping*); (iii) garantia de localização de objetos, sendo capaz de verificar se os objetos estão onde deveriam estar.

1.2 Motivações

Levando em consideração os requisitos de tempo real e desempenho de um IPS, se torna cada vez mais necessário o desenvolvimento de algoritmos com alta precisão e

¹ Fusão de sensores é a combinação de sensores distintos em qualquer tipo de implementação que receba dados de diferentes tecnologias sensoriais simultaneamente [37]. Este método é detalhado na Seção 2.1.4.

baixo tempo de resposta. Além disso, muitas organizações necessitam de sistemas de localização em todos os seus centros de distribuição. Assim, um sistema de localização que reutilize uma infraestrutura já existente, ou empregue equipamentos *off-the-shelf* e distribuídos em larga escala, pode reduzir os custos associados à implantação de um IPS.

Muitas aplicações necessitam de informações mais precisas da localização dos objetos em cenários estacionários. Por exemplo, a coleta de itens é uma extensão natural dos tradicionais sistemas de inventário. Nesta aplicação, um humano ou robô necessita encontrar a localização de determinado item, sendo que tal localização deve ser precisa o suficiente para que o item seja coletado corretamente, mesmo se existirem outros objetos por perto ou qualquer outro tipo de fenômeno interferente. Outra aplicação, considerada uma variante da coleta de itens, é a garantia de localização de objetos. Nesta aplicação, o sistema deve verificar se produtos, materiais ou equipamentos médicos estão de fato nos locais em que deveriam estar [76].

Atualmente existe uma grande quantidade de pesquisas acerca de IPSs, mantendo inclusive conferências destinadas exclusivamente ao assunto [42], [71]. Embora a pesquisa e o desenvolvimento de tais sistemas venham aumentando, poucos trabalhos utilizam uma abordagem híbrida em relação ao problema. Um sistema *wireless* combinado com outras tecnologias, tais como sistemas ópticos, inerciais e ultrassônicos são tendências na área. Desenvolver um sistema funcional que combine estas tecnologias é outro desafio existente, sendo tal combinação conhecida como fusão de sensores. Esta proposta aborda este tópico de maneira distinta aos poucos trabalhos encontrados sobre o assunto.

Embora as pesquisas já desenvolvidas demonstrem uma localização cada vez mais precisa, grande parte dos sistemas propostos são dependentes de novos componentes que devem ser fabricados em sua totalidade. Geralmente tal fabricação é laborosa e dependente de diversos outros materiais e recursos. Esta proposta de trabalho está baseada somente na utilização de dispositivos já existentes no mercado, tanto para a infraestrutura necessária quanto para os próprios dispositivos de localização. Além disso, os componentes utilizados possuem um valor mais acessível e parte deles já fazem parte do dia-a-dia das organizações em geral.

Por fim, este trabalho também visa dar continuidade a pesquisas relacionadas à tecnologia RFID, realizadas pelo autor e o grupo de pesquisa ao qual está inserido no programa acadêmico.

1.3 Contribuição Original

A originalidade do trabalho consiste na proposta de um método inovador de localização de objetos em ambientes estacionários, empregando dispositivos já existentes na infraestrutura das organizações, ou disponíveis em larga escala no mercado. O trabalho procura explorar uma lacuna existente através da combinação de tecnologias e técnicas de localização, visando otimizar a precisão dos sistemas de localização. A proposta também tem como contribuição a exploração do espaço tridimensional dos ambientes, onde tal escalabilidade possui baixa expressividade nas pesquisas em desenvolvimento.

1.4 Organização

O presente trabalho está organizado da seguinte maneira. O Capítulo 2 apresenta os principais conceitos, técnicas e tecnologias utilizadas em sistemas de localização *indoor*. O Capítulo 3 apresenta uma revisão dos trabalhos relacionados aos assuntos abordados na presente tese. O Capítulo 4 descreve o sistema desenvolvido, fornecendo detalhes dos métodos e modelos propostos. No Capítulo 5, são apresentados os experimentos realizados envolvendo o sistema proposto, além dos resultados obtidos. Por fim, o Capítulo 6 traz a conclusão e sugestões de trabalhos futuros.

2 CONCEITOS E DEFINIÇÕES

Este capítulo apresenta uma revisão dos conceitos, técnicas e tecnologias relacionados a sistemas de localização *indoor*. Na Seção 2.1 as definições e características de sistemas de localização são apresentadas, contemplando detalhes referentes às técnicas e tecnologias que podem ser empregadas neste tipo de sistema. A Seção 2.2 apresenta uma revisão conceitual sobre algoritmos de aprendizado de máquina, fornecendo detalhes dos modelos de algoritmos utilizados no IPS proposto neste trabalho.

2.1 Sistemas de Localização *Indoor*

Um sistema de localização *indoor* pode ser definido como um conjunto de dispositivos, técnicas, algoritmos e aplicações que estimam automaticamente as coordenadas relativas ou absolutas de objetos ou pessoas em ambientes internos. Um IPS deve determinar, continuamente e em tempo real, a posição de alguma coisa ou alguém em um espaço físico limitado e fechado, tal como uma casa, escola, laboratório, etc. Um IPS também deve ser capaz de considerar restrições de espaço e um alto número de obstáculos [35], [80].

Os IPSs são geralmente classificados pelas tecnologias empregadas nos dispositivos sensoriais utilizados na solução, tais como sistemas ópticos, radiofrequência, ultrassom, infravermelho, sistemas magnéticos, entre outros. Baseado nestas tecnologias, diversos IPSs foram desenvolvidos por diferentes empresas, centros de pesquisa e universidades. Os sistemas desenvolvidos geralmente adotam uma tecnologia em particular, ou a fusão destas tecnologias, sendo que a combinação destas não deixa de trazer para o sistema as limitações individuais de cada uma.

Uma localização cada vez mais precisa pode ser realizada com diversas tecnologias que possuem diferentes características físicas, operacionais e de custo [54]. A Tabela 1 fornece uma lista de tecnologias e a precisão típica alcançada.

Duas tecnologias que se destacam e possuem maior aplicabilidade são os sistemas ópticos (análise visual) e RFID, sendo abordados com maior ênfase nas Seções 2.1.2 e 2.1.3. Além destes sistemas, outros podem ser encontrados na literatura, tais como sensores magnéticos, ultrassom e infravermelho. Maiores detalhes sobre estes sistemas são descritos em [35], [40], [68].

Tabela 1 – Tecnologias de localização *indoor* e a precisão típica correspondente [54]

Tecnologia	Precisão típica
Dead reckoning ¹	Poucos centímetros
Wi-Fi	1-5 m
Bluetooth	2 m
RFID (etiquetas passivas)	1 m
RFID (etiquetas ativas)	1 m – 3 m
Ultra Wideband	30 cm – 1 m
ZigBee	1 m
Ópticos	Poucos centímetros
Localização acústica	25 m

Os principais elementos de um IPS são: (i) os dispositivos sensoriais, cuja função é obter as métricas referentes à posição relativa do objeto alvo; (ii) um algoritmo ou técnica de localização que processa as métricas oriundas dos dispositivos sensoriais; (iii) uma interface gráfica que ilustra a localização do objeto alvo para os usuários [70]. Em [35], estes elementos são associados a uma arquitetura em três camadas (Figura 1), definindo que os estímulos de localização, detectados por dispositivos sensoriais, fazem parte da primeira camada. Os algoritmos e técnicas de localização fazem parte da camada de abstração de software, convertendo os dados informados pela primeira camada para o formato exigido na camada mais alta, cuja função é ilustrar e fornecer controles geográficos e de navegação ao usuário.

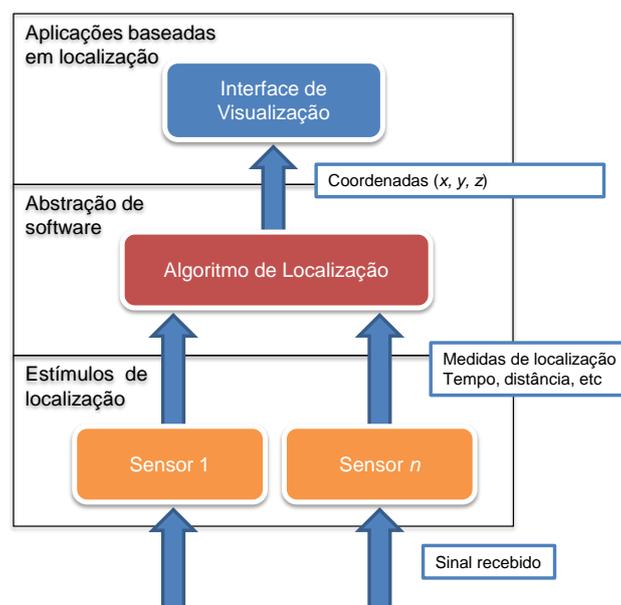


Figura 1 – Arquitetura em camadas associado a um diagrama de blocos contendo os componentes fundamentais de um IPS. Fonte: [35], [70].

¹ Processo no qual um objeto se auto-localiza através de estimativas em relação a uma posição sua anteriormente conhecida, de acordo com a distância, direção e velocidade percorrida.

Além de propriedades como desempenho, custo e escalabilidade, abordadas no Capítulo 1, IPSs também possuem outros critérios de avaliação, que podem ser definidos como [35], [37], [52]:

- i. **Segurança e privacidade:** é constante a preocupação dos usuários em relação a quem está realizando o seu rastreamento ou tem acesso ao histórico de locais percorridos. A privacidade dos sistemas de localização pode ser melhorada mediante o controle de acesso à informação de localização. Algumas implementações de sistemas de localização assumem que somente o próprio dispositivo alvo calcule e processe suas informações de localização. Neste modelo de sistema, é garantido que nenhuma outra entidade irá conhecer sua localização, ao menos que o dispositivo publique esta informação explicitamente.
- ii. **Atraso ou tempo de resposta:** o atraso existente em um IPS é considerado outro aspecto referente ao seu desempenho, isto porque um objeto rastreado pode estar se movendo rapidamente e também devido ao fato de ambientes internos estarem sob constante mudança. Este atraso pode estar presente na obtenção dos valores medidos, no processamento dos algoritmos de localização e até na publicação do resultado para as partes interessadas.
- iii. **Robustez e tolerância a falhas:** um IPS robusto deve se manter em operação mesmo quando determinados dispositivos do sistema não estão disponíveis ou não estão funcionando adequadamente (falta de energia, por exemplo). Por exemplo, em um sistema que utilize radiofrequência, o sinal da unidade transmissora pode estar completamente bloqueado. Neste caso, a única informação disponível para estimar a localização é o sinal de outros dispositivos. Assim, mesmo em situações inesperadas do ambiente, ou em falhas no sistema, o sistema de localização deve continuamente oferecer o serviço de localização.
- iv. **Complexidade:** este critério diz respeito ao esforço e intervenção humana na implantação e manutenção do IPS. Um cenário de baixa complexidade exige uma fase de configuração rápida, com um baixo número de componentes a serem fixados na infraestrutura, e uma interface de software de fácil uso. Outra característica da complexidade é o tempo de processamento que o dispositivo leva para determinar a localização. É desejável que um sistema de localização utilize uma metodologia com baixa complexidade de

algoritmos e cálculos. Relativo à manutenção do sistema, um IPS deve ter a capacidade de se adequar a infraestrutura de acordo com o crescimento do ambiente de trabalho.

- v. Limitações: embora a constante melhoria dos IPSs, eles ainda possuem algumas limitações, principalmente devido a tecnologia utilizada. Por exemplo, o uso de redes wireless (IEEE 802.11) em sistemas de localização é de grande interesse dos projetistas, pois o sistema pode reutilizar a infraestrutura de rede existente e reduzir o custo dos serviços de localização. Entretanto, sistemas baseados em radiofrequência sofrem com os efeitos de multi-caminho e reflexão dos sinais, resultando em uma alta taxa de erros. Outro tipo de limitação bastante relevante é a área de escopo do sistema. Alguns IPSs cobrem uma área pequena e não são escaláveis para ambientes maiores. Além disso, a limitação que diz respeito à quantidade de objetos/pessoas a serem localizados simultaneamente é outro fator que pode ser decisivo na escolha do sistema.

No que concerne às informações de localização dos sistemas *indoor*, diferentes aplicações possuem necessidades distintas quanto à exibição e ao tratamento da informação de localização. A informação de localização fornecida pelos sistemas pode ser de dois tipos: física ou simbólica. A localização física é expressa na forma de coordenadas, que identificam um ponto em um mapa 2D ou 3D. O GPS é um exemplo de sistema que fornece informação física, exibida em graus/minutos/segundos, graus/decimais/minutos, ou no formato UTM (*Universal Transverse Mercator*). A localização simbólica expressa as informações em linguagem natural, fornecendo ideias abstratas de aonde determinado objeto se encontra, tais como: no escritório, no terceiro andar, próximo ao armário [37], [52].

Outra propriedade dos sistemas de localização diz respeito ao sistema de coordenadas empregado. Esse tipo de informação pode ser implementado através de duas abordagens: absoluta ou relativa. Sistemas de localização com coordenadas absolutas são aqueles que usam grade de referência compartilhada para todos os objetos. Por exemplo, todos os receptores GPS usam latitude, longitude e altitude (ou equivalentes) para informar a localização. Já em sistemas de localização relativa cada objeto possui seu próprio elemento de referência, sendo que informações deste tipo são geralmente baseadas na proximidade a pontos de referência conhecidos. Por exemplo, uma equipe de bombeiros pode utilizar *smartphones* para localizar vítimas de um

terremoto. Neste caso, o sistema de localização exibe a posição da vítima relativa à localização do bombeiro portador do *smartphone* [37].

As seções seguintes deste capítulo descrevem as técnicas, tecnologias e métodos de fusão utilizados em IPSs. O capítulo aborda com maior detalhamento os assuntos determinantes para o desenvolvimento da proposta de sistema deste trabalho, sendo explanadas as tecnologias ópticas, radiofrequência e os métodos de fusão de sensores comumente empregados.

2.1.1 Técnicas de Localização

Na tentativa de determinar a localização *indoor* de objetos e pessoas, três principais técnicas¹ podem ser escolhidas: triangulação, proximidade e análise de padrões. Estas técnicas podem ser utilizadas individualmente ou combinadas entre si. Para tanto, torna-se necessária a criação de modelos que sejam capazes de processar as medições efetuadas. Entretanto, a definição de tais modelos não é uma tarefa trivial. Isto porque em ambientes internos pode ocorrer com frequência o efeito de multi-caminho, além de existir uma baixa probabilidade de linha de visada (*Line-of-Sight* – LOS) entre os dispositivos. Neste cenário, também se encontram variáveis que são inerentes ao ambiente, como o desenho e a estrutura do piso e paredes, além de objetos móveis e diversas superfícies reflexivas. Atualmente, são desconhecidos bons modelos para localização em ambientes internos com características de multi-caminho. Além de empregar a triangulação tradicional, algoritmos de localização que utilizam análise de padrões ou proximidade são desenvolvidos para mitigar os erros. Direcionados para diferentes aplicações ou serviços, estas três técnicas possuem vantagens e desvantagens. O uso de mais de uma técnica simultaneamente pode contribuir com o desempenho geral do modelo [52].

2.1.1.1 Triangulação

A técnica de triangulação utiliza as propriedades geométricas dos triângulos a fim de encontrar a localização de um objeto. A triangulação pode ser realizada através da lateração, que estima a posição do objeto medindo sua distância até múltiplos pontos de referência conhecidos, ou via angulação, que mede o ângulo relativo a pontos de referência cuja distância entre eles seja conhecida [37], [52].

¹ Algumas referências bibliográficas também identificam tais técnicas como algoritmos de localização. Este trabalho manteve a nomenclatura original de uma das primeiras pesquisas em relação ao assunto [37].

2.1.1.1.1 Lateração

A lateração calcula a posição de um objeto através de medições da sua distância até a posição de múltiplos pontos de referência. O cálculo da localização do objeto em duas dimensões necessita do valor de distância para até três pontos de referências não-colineares¹, conforme ilustrado na Figura 2. A lateração circular consiste em definir a distância como o raio da circunferência, centralizado em um respectivo ponto de referência. Assim, a posição do objeto é estimada através da intersecção das circunferências calculada em relação aos pontos de referência [38].

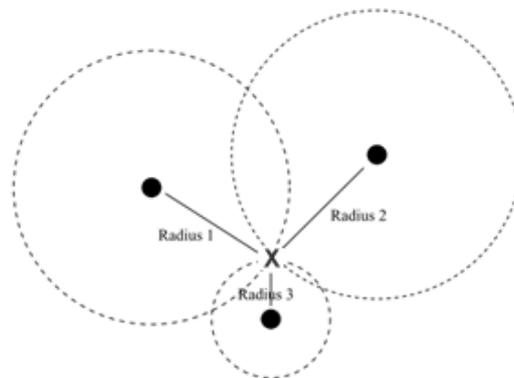


Figura 2 – Para obter a posição 2D utilizando lateração, é necessário medir a distância entre o objeto “X” e três pontos não-colineares. Fonte: [38].

Com o intuito de medir as distâncias necessárias pela técnica de lateração, três abordagens distintas podem ser aplicadas [9], [38], [52]:

- i. Direta: usa um movimento ou ação física para medir as distâncias necessárias no processo de triangulação, ou seja, a distância é medida explicitamente por uma entidade física que se move no ambiente. Por exemplo, um robô pode estender um dispositivo de medição até tocar em algum obstáculo (no ponto de referência, por exemplo). Esta abordagem é de fácil entendimento, mas muito difícil de ser automatizada.
- ii. Tempo de chegada (*Time of Arrival* – TOA): está baseado na medição do tempo que algum tipo de sinal leva para ir do objeto (transmissor) até cada um dos pontos de referência (P). Por exemplo, a velocidade aproximada das ondas sonoras a 21° C é de 344 m/s. Assim, se um pulso ultrassônico enviado por um objeto leva 14,5 ms até o ponto P , nos permite concluir que a distância entre P e o objeto é de 5 m. Sinais de radiofrequência também

¹ Três ou mais pontos são colineares se estiverem posicionados de tal forma que se pode traçar uma reta sobre todos eles.

são muito utilizados em sistemas TOA. Uma das limitações desta abordagem é a sincronização de todas as estações de referência, assim como do dispositivo transmissor que está junto ao objeto. Devido às unidades de tempo utilizadas, um pequeno erro de sincronização resulta em um grande erro de localização. Por exemplo, um erro de 100 ns pode resultar em um erro de 30m. Esta é a técnica utilizada no GPS.

- iii. Atenuação ou intensidade do sinal recebido (*Received Signal Strength Indicator* – RSSI): a intensidade de um sinal diminui conforme a distância aumenta. Atenuação é a perda gradual de intensidade em relação ao valor de origem. Com uma função que correlacione atenuação e distância, dado um tipo de emissão e a intensidade na origem, é possível estimar a distância de um objeto até algum ponto de referência P , medindo a força do sinal quando este chegar a P . Diversos modelos teóricos e empíricos têm sido propostos a fim de obter a estimativa de distância a partir da diferença entre a intensidade de sinal transmitida e recebida. Entretanto, efeitos de multi-caminho, refração e reflexão, presentes em ambientes internos, fazem com que os modelos de atenuação não funcionem adequadamente na maioria das situações. A estimativa da distância através desta técnica é geralmente imprecisa, levando os modelos a empregarem parâmetros *site-specific*. A precisão desta técnica pode ser melhorada utilizando dados de medições do RSSI coletados previamente no ambiente.

Atualmente, existem diversas pesquisas acerca de novas técnicas de lateração, que, em sua maioria, são variantes das abordagens descritas. São exemplos de técnicas mais recentes de lateração: tempo de chegada relativo (*Time Difference of Arrival* - TDOA), *Round-Trip Time of Flight* (RTOF) e *Received Signal Phase* (RSP), também conhecido como *Phase of Arrival* (POA) [9], [52].

2.1.1.1.2 Angulação

A angulação emprega a análise de ângulos, ao invés de distâncias, para determinar a posição de um objeto. Em cenários 2D, a angulação necessita, pelo menos, dos ângulos de dois pontos de referência e também da distância entre estes dois pontos, conforme ilustrado na Figura 3. A partir destas informações, a posição do objeto é estimada utilizando funções trigonométricas [38].

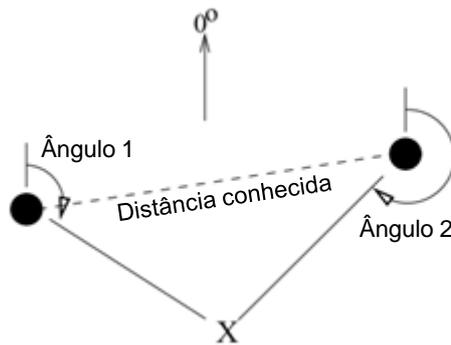


Figura 3 – Exemplo de angulação que localiza o objeto X. Fonte: [38].

O ângulo de chegada (*Angle of Arrival* – AOA) é um dos métodos tipicamente empregados na angulação. O AOA consiste na verificação do ângulo de incidência do sinal quando este chega a um receptor de referência. O uso de uma antena giratória é uma das formas mais simples de detectar o ângulo de maior incidência do sinal, entretanto é difícil de ser implementada na prática. Assim, o uso de *array* de antenas é o mais indicado. Um *array* de antenas consiste em múltiplas antenas com distância conhecida entre si e que medem o tempo de chegada de um sinal. Dada a diferença no tempo de chegada e a geometria do *array*, é possível calcular o ângulo de onde se originou o sinal [38].

As vantagens do AOA referem-se ao baixo número de pontos de referência necessários na infraestrutura, além de não ser necessária a sincronização dos relógios entre os componentes do sistema. Como desvantagem, este método requer equipamentos geralmente caros e de maior complexidade. Além disso, a precisão na estimativa de localização diminui conforme a distância entre os componentes aumenta, soma-se a isto o fato das tecnologias utilizadas neste método também sofrerem com efeitos de multi-caminho e reflexões [9].

2.1.1.2 Proximidade

Proximidade é a técnica que determina o quão próximo um objeto está de uma localização conhecida. Os algoritmos de proximidade informam a localização simbólica e relativa dos objetos. Esta técnica necessita de sensores afixados em posições conhecidas do ambiente. Quando um objeto alvo é detectado pelo sensor, considera-se que a posição do objeto está próxima à área de atuação do sensor [35], [52], [80].

Conforme ilustrado pela Figura 4, E_2 e E_3 são os objetos rastreados. A área de atuação do sensor D é delimitada pelo retângulo pontilhado. O sensor D monitora se E_2 e

E_3 estão posicionados ou não sob sua área de atuação. No caso, E_2 está na área de D e E_3 não está.

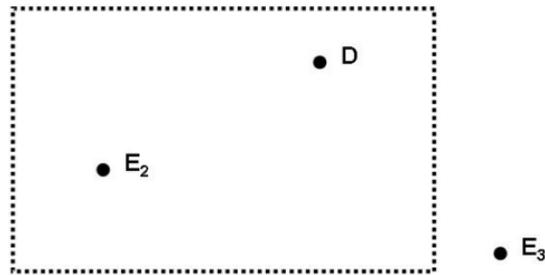


Figura 4 – Técnica de localização por proximidade. Fonte: [35].

A técnica é relativamente simples de implementar, podendo ser utilizados diferentes tipos de equipamentos. Atualmente, sistemas de localização baseados em proximidade geralmente utilizam radiação infravermelha ou RFID. Quando empregada a tecnologia RFID, geralmente um *grid* de antenas é arranjado no ambiente, cada uma em uma posição conhecida pelo sistema. No momento em que um objeto alvo (etiqueta) é detectado por uma única antena, considera-se que o objeto esteja dentro da área de proximidade daquela antena. Caso mais de uma antena detecte o objeto, considera-se a área da antena que recebeu o sinal de resposta mais intenso [52].

2.1.1.3 Análise de Padrões

Este método utiliza as características de uma cena observada a partir de um ou mais pontos, a fim de obter respostas sobre a localização do observador ou dos objetos da cena. Geralmente, a cena observada deve ser o mais simples possível, visando uma fácil captura das características que serão empregadas e comparadas pelo sistema de localização. Este método também é conhecido como análise de cena [38].

A vantagem da análise de padrões está no fato da localização ser estimada utilizando observações e características que não dizem respeito a ângulos geométricos ou distâncias. A desvantagem desta técnica é a necessidade de conhecer as diversas características do ambiente em que as cenas serão capturadas. As mudanças no ambiente, de maneira que alterem as características observadas, podem fazer com que seja necessária a reconstrução dos dados pré-definidos ou até a captura de todo um novo conjunto de dados [38].

O cenário observado pode ser obtido a partir da técnica denominada *fingerprinting*, que consiste na análise de valores medidos pelos sensores, ou através da técnica de

análise visual, baseado na análise de imagens provenientes de uma ou mais câmeras. Geralmente, a técnica *fingerprinting* emprega sensores baseados em radiofrequência, utilizando as impressões eletromagnéticas durante a análise.

A técnica *fingerprinting* utiliza dados de localização medidos previamente para tentar melhorar a precisão dos sistemas de localização *indoor*. O *fingerprinting* consiste em duas fases: a fase de treinamento *offline* e a fase de estimativa *online*. Na fase *offline*, determinadas características dos componentes da infraestrutura e as informações de localização são medidas e coletadas em diferentes locais da área de abrangência do sistema. Durante a fase *online*, as características de um objeto alvo são medidas e comparadas aos dados coletados na fase *offline*. Assim, ao encontrar determinadas características idênticas ou semelhantes na base *offline*, é possível estimar as informações de localização do objeto. Tipicamente, a técnica *fingerprinting* emprega um dos seguintes modelos a fim de realizar a estimativa da localização: métodos probabilísticos, *k-Nearest Neighbors* (kNN), Redes Neurais Artificiais (RNA) e *Support Vector Regression* (SVR).

A análise visual estima a localização a partir de imagens capturadas por uma ou mais câmeras. A localização visual pode trazer maior conforto e eficiência aos usuários, pois geralmente não é necessária carregar ou afixar nenhum dispositivo extra aos objetos. Tipicamente, uma ou mais câmeras são posicionadas na área de atuação do IPS para cobrir todo o local e capturar imagens em tempo real. A partir das imagens, os objetos alvo são detectados utilizando algoritmos de visão computacional. As imagens dos objetos capturados podem ser buscadas em uma base de dados pré-definida, a fim de estimar a posição do objeto em relação à câmera [35]. Maiores detalhes sobre os dispositivos e elementos presentes nesta técnica são explorados na Seção 2.1.2.

2.1.2 Sistemas Ópticos

Sistemas ópticos de localização *indoor* são aqueles que empregam câmeras como tecnologia sensorial. Geralmente são utilizadas câmeras de vídeo ou câmeras digitais no processo de captura das imagens. Grande parte dos sistemas ópticos de localização é baseada na técnica de análise visual e empregam algoritmos de visão computacional. Referente à infraestrutura do ambiente e os elementos presentes no cenário, esta tecnologia pode ser do tipo *marker* [46], [83] ou *markerless* [45], [93].

Em alguns sistemas, as imagens capturadas na análise visual são provenientes de visão estéreo (duas câmeras, no mínimo), com o objetivo de possibilitar uma localização com escalabilidade de espaço tridimensional.

2.1.2.1 Visão Computacional

Visão computacional (*Computer Vision – CV*) é uma área que inclui métodos para capturar, processar, analisar e entender imagens do mundo real visando a exploração técnica e científica [44]. O objetivo da visão computacional é criar decisões úteis sobre objetos físicos e cenas reais baseadas em imagens capturadas [79].

Não é clara a fronteira entre o processamento de imagens e a visão computacional. Processamento de imagens é um processo onde a entrada do sistema é uma imagem e a saída é um conjunto de valores numéricos, que podem ou não compor uma outra imagem. A visão computacional procura emular a visão humana, portanto também possui como entrada uma imagem, porém, a saída é uma interpretação da imagem como um todo, ou parcialmente [55].

O diagrama da Figura 5 ilustra um modelo representativo das fases existentes em um sistema de visão computacional. O sistema é composto de várias etapas e delas é obtido um resultado final específico para cada aplicação desejada.



Figura 5 – Passos fundamentais em um sistema de visão computacional. Fonte: [32].

No primeiro passo – aquisição de imagens - as imagens são capturadas por dispositivos imageadores, tais como câmera de vídeo, *scanner* e outros. Na maioria das câmeras, o sinal original é analógico e deve ser convertido para um sinal digital através de um digitalizador, processo pelo qual os sinais elétricos se transformam em uma matriz de bits ou *pixels*. Uma imagem digital é uma imagem $f(x, y)$ discretizada tanto em coordenadas espaciais quanto em brilho. Uma imagem digital pode ser considerada como sendo uma matriz cujos índices de linhas e de colunas identificam um ponto na imagem, e o correspondente valor do elemento da matriz identifica o nível de cinza (no caso de uma imagem monocromática) naquele ponto. Os elementos dessa matriz são mais conhecidos como *pixels* (abreviação de *picture elements*) [32]. A Figura 6 ilustra a ampliação da região de uma imagem onde são exibidos os respectivos *pixels*.

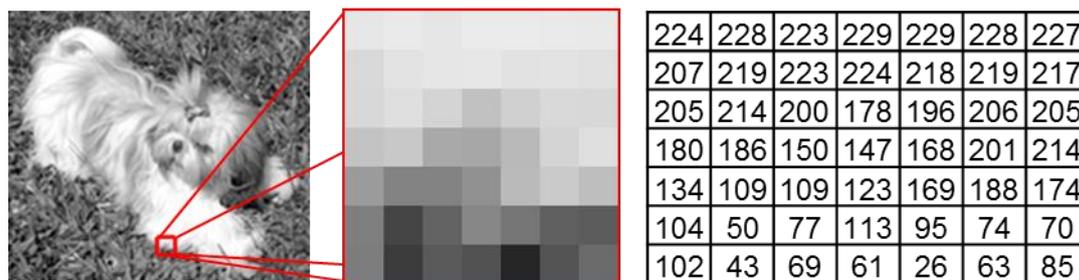


Figura 6 – Imagem original (150 x 135 *pixels*) na esquerda, zoom (3200%) de determinada região (7 x 7 *pixels*) ao centro e níveis de intensidade de cada *pixel* na direita. Fonte: o autor.

O próximo passo trata de pré-processar aquela imagem. O papel principal desta etapa é melhorar a imagem de forma a aumentar as chances para o sucesso dos processos seguintes [32]. As imagens de onde são extraídas informações podem necessitar a conversão para um determinado formato ou tamanho e ainda serem filtradas para remoção de ruídos provenientes do processo de aquisição da imagem [55].

A segmentação de imagens é uma das etapas fundamentais em um sistema de visão computacional. A segmentação divide uma imagem de entrada em partes ou objetos constituintes. Em geral, a segmentação automática é uma das tarefas mais difíceis no processamento de imagens digitais. No caso de reconhecimento de caracteres, o papel básico da segmentação é extrair caracteres individuais e palavras do fundo da imagem [32]. Mais detalhes sobre o processo de segmentação e algoritmos utilizados são apresentados na Seção 2.1.2.1.1.

Em relação ao quarto passo, o processo de descrição procura extrair características que resultem em alguma informação quantitativa de interesse ou que sejam básicas para discriminação entre classes de objetos. O último estágio na Figura 5 envolve reconhecimento e interpretação. Reconhecimento é o processo que atribui um rótulo a um objeto, baseado na informação fornecida pelo seu descritor. A interpretação envolve a atribuição de significado a um conjunto de objetos reconhecidos [32].

2.1.2.1.1 Segmentação de Imagens

O nível de detalhamento em um processo de segmentação depende da tarefa a ser executada e da resolução da imagem que se tem. Por exemplo, se procuramos por uma casa em imagens que foram obtidas do nível da rua, estamos procurando regiões que ocupam uma boa porcentagem da imagem (regiões grandes), porém, se procuramos por casas a partir de imagens de satélite, estamos procurando regiões pequenas. Embora a

tarefa seja a mesma, a resolução das imagens é diferente e o tratamento utilizado no processo de segmentação pode ser diferente também [55].

Existem diversas técnicas que podem ser utilizadas para a segmentação, podendo ser apresentadas em três grupos: (i) segmentação por detecção de borda; (ii) segmentação por corte; (iii) segmentação por crescimento de região. Devido ao seu desempenho e disponibilidade, a segmentação por detecção de borda foi empregada neste trabalho.

Uma borda em uma imagem é caracterizada por uma mudança, normalmente abrupta, no nível de intensidade dos *pixels*. Os detectores de borda são definidos para encontrar este tipo de variação nos *pixels* e quando estes *pixels* estão próximos eles podem ser conectados formando uma borda ou um contorno e assim definindo uma região ou objeto. Existem várias técnicas de detecção de bordas, tais como: Canny, Prewitt e Sobel [55].

A técnica de detecção de borda desenvolvida por Canny, embora seja a mais complexa, é também a que se destaca pelo desempenho. O detector de bordas de Canny possui três objetivos básicos: baixa taxa de erro (todas as bordas devem ser encontradas), os pontos da borda devem ser bem localizados (as bordas localizadas devem ser próximas das bordas reais) e resposta única para os pontos de uma borda (o operador deve retornar apenas um ponto para cada ponto sobre a borda) [55]. A seguir são descritas as etapas do algoritmo Canny [16], [58].

Redução de ruído: inicia-se convoluindo a imagem com uma máscara Gaussiana (detalhes no Anexo A). A saída do primeiro passo é uma versão ligeiramente mais desfocada da imagem original. Isto conseqüentemente reduz o efeito de *pixels* ruidosos na imagem (Figura 7b).

Busca por gradientes de intensidade: uma borda em uma imagem pode estar direcionada em vários sentidos. Assim, quatro máscaras são usadas para detectar bordas horizontais, verticais e diagonais (0, 45, 90 e 135 graus). Os resultados do processamento da imagem original com cada uma destas máscaras são armazenados. Para cada *pixel*, marca-se então o maior resultado do gradiente nesse *pixel*, e o sentido da máscara que produziu essa borda. Da imagem original, cria-se dois mapas: (i) dos gradientes da intensidade em cada ponto na imagem; (ii) do sentido do gradiente no ponto. A Figura 7c ilustra o resultado desta etapa.

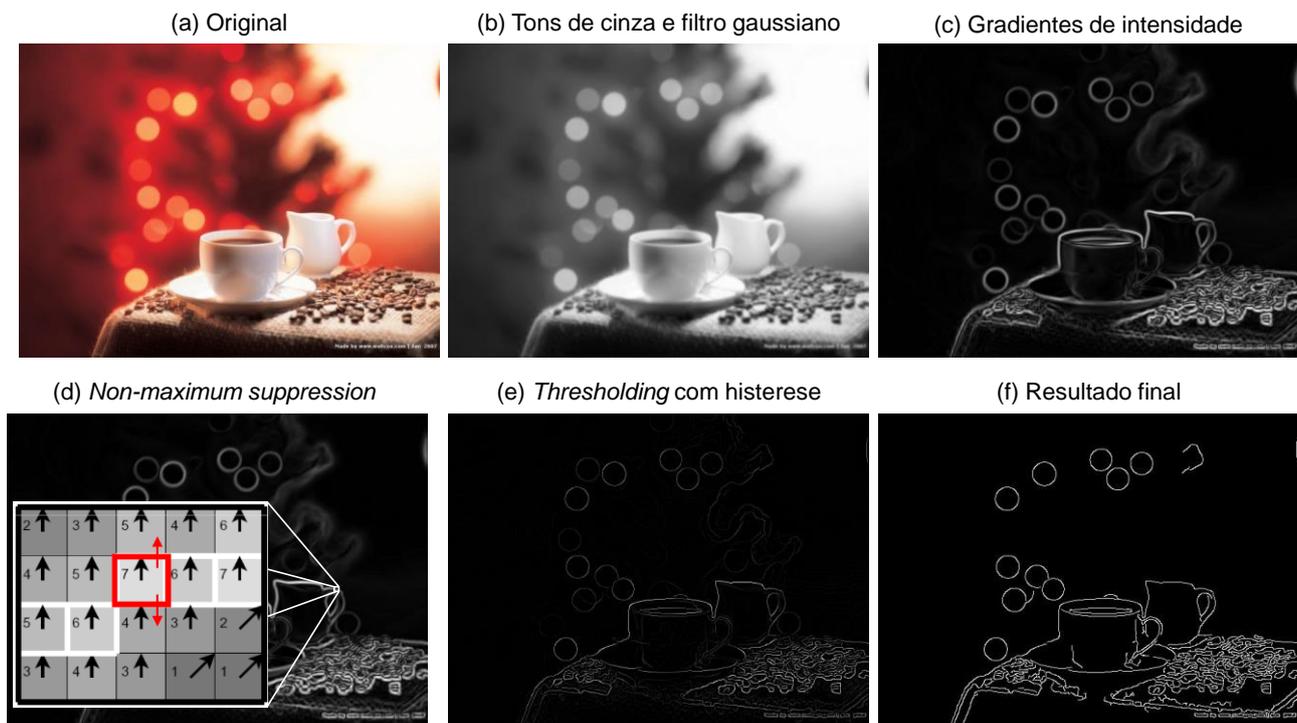


Figura 7 – Imagens resultantes em cada etapa do detector de bordas Canny. Fonte: [31], [58].

Non-maximum supression: o propósito deste passo é converter arestas suavizadas na imagem do passo anterior em arestas mais “sharp”. Isto é realizado preservando os máximos locais na imagem do gradiente e removendo o restante. A comparação é feita na direção do ângulo. Por exemplo, se a direção for de 0° , compara o elemento anterior e posterior da mesma linha. Na Figura 7d, os *pixels* apontando para o norte são comparados com *pixels* abaixo e acima. Aqueles que são preservados estão nos quadrados brancos, pois são os maiores (naquela direção), os demais são suprimidos. No exemplo, o valor 7 é comparado com as intensidades dos *pixels* acima e abaixo, como ele é maior, permanece.

Thresholding com histerese: os gradientes de intensidade mais elevada são mais prováveis de serem bordas. Mas não há um valor absoluto pré-definido de quanto uma dada intensidade passa a ser uma borda. A ideia de Canny é usar limiares para isto. A ideia é chamada de *thresholding* com histerese e requer dois limiares iniciais: um superior e outro inferior. Supondo que as bordas devem ser linhas contínuas, mesmo linhas de pouca intensidade são investigadas, mas evita-se identificar *pixels* que não constituem uma linha. Assim, um *pixel* (x, y) é chamado forte e convertido para branco se $cor(x, y) > Th_{superior}$, é denominado fraco se $cor(x, y) \leq Th_{inferior}$, e todos os outros *pixels* são candidatos. Os *pixels* ao redor dos *pixels* fortes são analisados, caso seus valores sejam maiores do que $Th_{inferior}$, eles também são convertidos para branco.

Uma vez que estas etapas estejam completas, tem-se uma imagem binária onde cada *pixel* é marcado como um *pixel* de borda ou de não-borda (Figura 7f).

2.1.2.2 Marcadores Visuais

Algoritmos do tipo *marker* são baseados em marcas com padrões únicos em seu conteúdo. Geralmente, elementos são criados para receberem marcas possuindo tais características, recebendo assim o nome de marcadores visuais ou fiduciais. Os marcadores visuais são arranjados em um ambiente físico a fim de dar suporte à identificação, localização e rastreamento de objetos [57]. Um marcador visual é projetado para resolver o seguinte problema: para uma determinada imagem de entrada (imagem estática ou quadro de *streaming* de vídeo), forneça a lista de marcadores encontrados na imagem [25].

Em um sistema deste tipo, os padrões contidos nos marcadores devem ser distintos o suficiente para não serem confundidos com o restante do ambiente. O processamento de imagens deve ser suficientemente robusto a ponto de encontrar marcadores em situações com variação na iluminação, ruído na imagem, escala desconhecida, imagem embaçada e ocultação parcial. Preferencialmente, os marcadores devem ser passivos (que não necessitam energia) e possuírem uma superfície plana. Desta forma, além de possuir um custo mais baixo, o marcador visual pode ser impresso e alocado de forma mais fácil no cenário. Além disso, é desejável que o marcador ocupe o mínimo possível de *pixels* na imagem capturada, visando maximizar o espaço de uso [25].

Sistemas ópticos de localização baseados em marcadores costumam ser bastante precisos, entretanto exigem um esforço maior na configuração da infraestrutura do ambiente, necessitando a afiação ou arranjo dos marcadores nos objetos ou no cenário. Marcadores visuais podem ser associados e afixados a objetos que se deseja localizar. Com a relação de associação mantida em uma base de dados do sistema, a partir do momento que a análise visual detecta, localiza e identifica o marcador, a posição do objeto respectivo também é conhecida.

Tipicamente, marcadores visuais possuem padrões de cores, formatos e conteúdo distintos do restante da imagem. No decorrer dos anos, algumas propostas de marcadores visuais foram desenvolvidas.

Um marcador visual muito popular é o código de barras. No código de barras emprega-se a codificação de dados em imagens utilizando uma abordagem binária. O padrão é tipicamente bitonal, reduzindo o processo de decisão *pixel* a *pixel* para um

processo de decisão binária. Assim, os requisitos de sensibilidade da câmera (como iluminação, por exemplo) são reduzidos [25].

Usualmente, diversos padrões de código de barras podem ser utilizados para codificar informações. No entanto, tais padrões não são tão úteis em sistemas de marcadores visuais, pois: (i) eles não foram criados para este objetivo e podem não funcionar muito bem em cenas que introduzem distorção na perspectiva; (ii) estes padrões tipicamente necessitam de uma grande área da imagem, limitando o espaço em que os marcadores podem ser empregados. Embora possuindo restrições, os padrões de bitonalidade e a robustez dos métodos digitais empregados por estes padrões podem ser aplicados a projetos de marcadores visuais [25].

2.1.2.3 *Markerless*

Nos algoritmos *markerless*, as características visuais do objeto que se deseja localizar não dependem de marcas artificiais. Em sistemas deste tipo, os padrões existentes nos objetos são inerentes a sua própria natureza, cabendo ao sistema a função de encontrar os padrões respectivos na imagem.

Toma-se como exemplo um IPS cujo objetivo seja localizar uma onça-pintada ao ar-livre. Ou seja, ao analisar a imagem, o algoritmo deve ser capaz de encontrar um padrão que identifique um animal de médio porte, cor amarelada e com pintas pretas. Apesar da pele da onça ter certo padrão de cor e formato, não é necessário afixar um marcador visual em seu corpo para realizar a localização. Pesquisas recentes têm utilizado técnicas de análise visual *markerless* para realizar o rastreamento de animais em zoológicos [45].

Outros exemplos típicos de algoritmos *markerless* são os detectores de pessoas, gestos e face [93]. Tipicamente, algoritmos *markerless* exigem um maior poder de processamento computacional em relação a algoritmos que empregam marcadores visuais.

2.1.3 Identificação por Radiofrequência (RFID)

A sigla RFID significa *Radio Frequency Identification*, ou identificação por radiofrequência, sendo o termo que representa qualquer sistema de identificação onde existe um dispositivo eletrônico que usa radiofrequência ou variações de campo magnético para efetuar a comunicação entre os elementos do sistema [28].

A tecnologia RFID oferece benefícios para pessoas e organizações que precisam manter a rastreabilidade de bens físicos. Implantando tecnologias RFID, os fabricantes melhoram o planejamento e a execução de suas cadeias de suprimento. No varejo, o RFID é utilizado para prevenir roubos, aumentar a eficiência das vendas, e melhorar o planejamento do estoque. Governos de determinados países estão emitindo passaportes utilizando a tecnologia RFID para identificação dos passageiros nos aeroportos. O interesse das organizações em utilizar o sistema é cada vez maior, devido ao seu potencial que torna mais simples e eficiente a identificação e o rastreamento automático de pessoas e objetos.

Em um sistema RFID sempre estão presentes dois componentes:

- Etiqueta (também chamada de *tag* ou *transponder*): dispositivo de identificação afixado ao objeto que se deseja rastrear;
- Leitor (ou interrogador): é o dispositivo que pode detectar a presença de etiquetas RFID e ler as informações nelas armazenadas. Dependendo da tecnologia ou do projeto, um leitor também pode escrever informações nas etiquetas.

Geralmente o leitor está conectado em algum sistema que possui um software cuja função é intermediar a transmissão de dados entre o leitor e a aplicação. O computador que executa o software geralmente é conhecido como servidor do *middleware* RFID. O leitor gera sinais eletromagnéticos que são transmitidos por uma ou mais antenas. Como segunda função, o leitor também monitora sinais eletromagnéticos oriundos das etiquetas, usualmente utilizando o mesmo conjunto de antenas [26], [28]. A Figura 8 ilustra como se relacionam os componentes do sistema RFID.

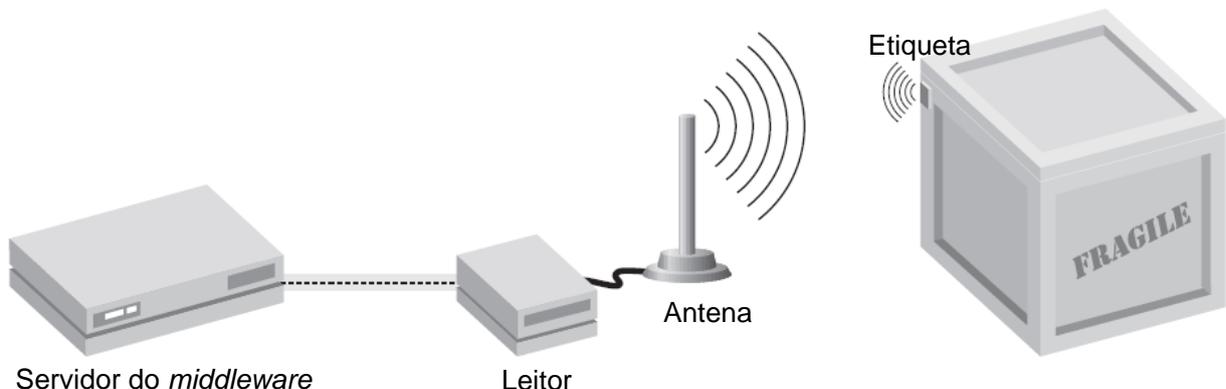


Figura 8 – Componentes de um sistema RFID. Fonte: [28].

As etiquetas de identificação são os dispositivos que, anexados ao que se deseja identificar, respondem aos sinais emitidos pelo equipamento leitor, que por sua vez consegue detectar a presença de etiquetas RFID e ler as informações armazenadas nessas etiquetas. A informação mais comum armazenada em uma etiqueta é o EPC (*Electronic Product Code*). Esse código é um valor único para identificações, sendo um padrão estabelecido pela EPCglobal¹ [34]. Um leitor obtém dados das etiquetas através da transmissão de um sinal de onda contínuo; as etiquetas respondem retornando um sinal modulado com o coeficiente de reflexão da sua antena. O canal de comunicação é compartilhado entre todos os dispositivos do sistema [5].

As etiquetas de um sistema RFID estão enquadradas em três categorias: ativas, passivas e semi-passivas. As etiquetas ativas possuem uma fonte própria de energia, tal como uma bateria, e podem tomar a iniciativa de iniciar a transmissão de dados. As etiquetas passivas obtêm sua energia do sinal transmitido pelo leitor, emitindo informações somente após terem sido interrogadas por um equipamento leitor. As etiquetas semi-passivas possuem fonte própria de energia para alimentar o circuito da etiqueta, mas não possuem funcionalidades de transmissão ativa, apenas respondendo os seus dados quando questionadas pelo leitor. A Tabela 2 relaciona as características entre os dois principais tipos de etiquetas.

Tabela 2 – Características dos tipos de etiquetas RFID [4].

Característica	Etiqueta ativa	Etiqueta passiva
Fonte de energia	Bateria	Fornecida pelo leitor
Força do sinal necessário	Baixo	Alto
Área de alcance	Longa distância (~100 m)	Curta distância (~3 m)
Armazenamento de dados	Até 128 Kb Leitura e escrita	128 bytes Leitura e escrita
Custo por etiqueta	Em média: \$15 a \$100	Em média: \$0,15 a \$5,00
Tamanho da etiqueta	Depende da aplicação	De adesivos até o tamanho de um cartão de crédito

Devido à combinação de tamanho da etiqueta, capacidade de leitura a distâncias próximas a 3 metros e controle da área de leitura através do direcionamento das antenas e configurações do leitor, grande parte dos esforços para a utilização de RFID no controle

¹ A organização EPCglobal (criada em 2003) é a responsável pela padronização dos códigos numéricos únicos EPC para RFID. Esta organização nasceu a partir dos trabalhos iniciados em 1999 pelo laboratório de pesquisa Auto-ID Center do MIT (Massachusetts Institute of Technology), sendo, portanto, o EPCglobal uma entidade organizada visando permitir a profissionalização dos resultados dos trabalhos do MIT/Auto-ID Center.

de cadeias de suprimentos e bens de consumo são direcionados para as etiquetas passivas UHF, padronizadas pela norma ISO 18000-6.

Levando em consideração que sistemas RFID estão cada vez mais presentes no cotidiano das pessoas e empresas, existe uma alta demanda por sistemas de localização baseados nesta tecnologia. Outras características que tornam a tecnologia atrativa dizem respeito ao tamanho reduzido dos seus componentes e o baixo consumo de energia, principalmente em etiquetas passivas que não utilizam baterias. Entretanto, obstáculos como paredes, equipamentos e seres humanos influenciam na propagação de ondas eletromagnéticas, gerando o chamado efeito de multi-caminho. Interferências e ruídos oriundos de outras redes (com e sem fio) também podem degradar a precisão de um IPS. Como os sinais destes sistemas não podem ser medidos com muita precisão, diversos algoritmos de localização *indoor* tem sido propostos na literatura [35].

2.1.3.1 Modelagem do Canal de Comunicação

Uma dos sinais presentes na tecnologia RFID que é muito utilizado em sistemas de localização é o nível de intensidade do sinal recebido (*Received Signal Strength Indicator* – RSSI). O RSSI é medido através da potência e é geralmente representado em dBm¹. Para descrever o valor do RSSI, uma modelagem do canal de comunicação aplicado a sistemas RFID é apresentada [6].

Um modelo de propagação em espaço livre é utilizado para estimar a intensidade do sinal recebido quando o transmissor e o receptor possuem um caminho de linha de visada limpo e sem obstruções entre eles. Em espaço livre, a potência de chegada na antena receptora – que é separada da antena transmissora pela distância r – é obtida pela equação Friis,

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi r} \right)^2 \quad (1)$$

onde P_r é a potência recebida, P_t é a potência transmitida, G_t e G_r são os ganhos² do transmissor e receptor, respectivamente. O lambda (λ) é o comprimento de onda da frequência de transmissão, e r é a distância em metros entre o transmissor e o receptor.

¹ dBm: dB miliwatt, medida mais comum para expressar a potência de um equipamento de transmissão de rádio.

² A propriedade de ganho de potência (G) é uma medida relacional entre a diretividade do sinal emitido por uma dada antena em relação a uma antena isotrópica, a qual irradia uniformemente em todas as direções [5].

Partindo deste modelo básico é possível aplicar a equação Friis em um modelo de propagação RFID em espaço livre, tendo como transmissor o equipamento interrogador (leitor), e como receptor a etiqueta de identificação (*tag*). A equação Friis adaptada para o caminho do leitor até a etiqueta (canal de ida) pode ser descrito como

$$P_{2r} = P_{1t} G_1 G_2 \left(\frac{\lambda}{4\pi r} \right)^2 \quad (2)$$

onde P_{2r} é a potência recebida pela etiqueta, P_{1t} é a potência transmitida pelo leitor, e G_1 e G_2 são os ganhos do leitor e da etiqueta, respectivamente. Para o canal reverso (da etiqueta para o leitor) o resultado da equação deve ser a potência recebida pelo leitor:

$$P_{1r} = P_{2t} G_1 G_2 \left(\frac{\lambda}{4\pi r} \right)^2 \quad (3)$$

No entanto, existe uma perda de intensidade entre a potência recebida pela etiqueta e a potência refletida, isto porque certa quantidade de energia é utilizada pelo chip da etiqueta para ativá-la e poder efetuar a resposta. Esta perda na transmissão refletida é conhecida como *backscattered loss*, ou B_L [27].

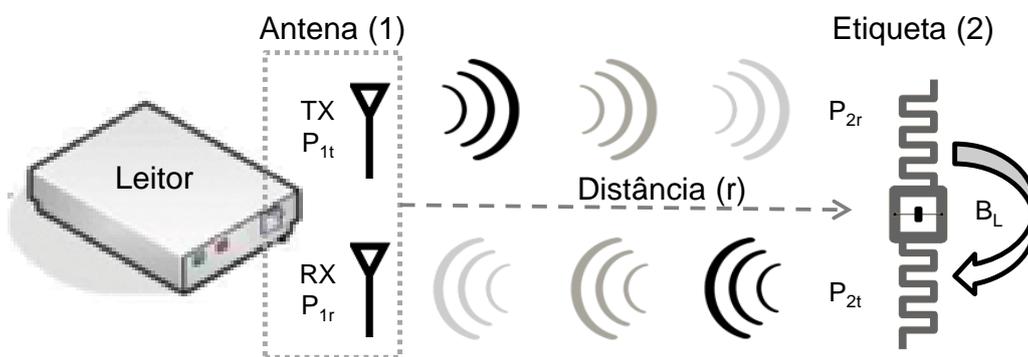


Figura 9 – Funcionamento do RFID em espaço livre com variáveis do modelo. Fonte: [6].

Na equação que contempla todo o cenário RFID em espaço livre, a potência recebida pelo leitor diminui com o inverso da distância na quarta potência. Além disso, como agora ambas as antenas são utilizadas duas vezes (na transmissão e na recepção), o ganho de cada uma é elevado na potência 2. A fórmula completa com a potência recebida pelo leitor está definida em (4), e o ambiente é ilustrado na Figura 9.

$$P_{1r} = P_{1t} (G_1 G_2)^2 \left(\frac{\lambda}{4\pi r} \right)^4 B_L \quad (4)$$

O valor de RSSI empregado em sistemas de localização baseia-se na potência recebida pelo leitor, representada pela variável P_{1r} na modelagem descrita. Atualmente,

leitores em conformidade com o padrão LLRP¹ opcionalmente implementam o parâmetro “PeakRSSI”, que fornece o valor de RSSI para as leituras realizadas [22].

2.1.3.2 Modelo Tradicional de IPS RFID

Em sistemas de localização *indoor*, a tecnologia RFID pode utilizar etiquetas passivas ou ativas, cada uma com suas vantagens e desvantagens. Um exemplo de modelo convencional de uso da tecnologia RFID em IPSs pode ser descrito através dos seguintes passos [54]:

- i. Um ou mais leitores RFID são configurados no cenário. Cada leitor possui uma ou mais antenas. As antenas são instaladas em determinado local, de forma que uma etiqueta possa ser lida por múltiplas antenas simultaneamente.
- ii. O sistema de localização ativa o leitor RFID para realizar a leitura de etiquetas específicas ou um conjunto de etiquetas.
- iii. Os leitores RFID efetuam a leitura das etiquetas e armazenam o RSSI, na recepção do leitor, para cada etiqueta lida. Os leitores encaminham os dados da etiqueta e o RSSI para o sistema de localização.
- iv. O sistema de localização calcula a localização das etiquetas utilizando alguma técnica de triangulação baseada na localização das antenas e nos valores de RSSI.

Referente a IPSs que adotam a tecnologia RFID com etiquetas ativas, um modelo típico de localização pode ser descrito como [54]:

- i. Etiquetas RFID ativas enviam sinais continuamente, por exemplo, a cada 10 segundos.
- ii. Estes sinais são recebidos pelos leitores RFID instalados no ambiente.
- iii. O leitor RFID realiza medições destes sinais a partir do tempo de chegada, intensidade do sinal ou ângulo de chegada. As medições são então encaminhadas ao sistema de localização.
- iv. O sistema de localização processa as medições recebidas, aplicando alguma técnica de localização, tais como lateração, angulação ou análise de padrões.

¹ LLRP (*Low Level Reader Protocol*) é uma especificação da EPCglobal para a interface entre clientes e leitores RFID. Esta interface é responsável por, entre outras coisas, oferecer de forma mais robusta as informações de *status* durante as operações de acesso às etiquetas.

Os seguintes itens podem ser listados como vantagens no uso da tecnologia RFID em IPSs: baixo custo da etiqueta, fácil afixação aos objetos, alta taxa de leitura (alta vazão), e precisão próxima a 1 m. Como desvantagens, podem ser listadas as seguintes características: falta de padronização, intolerância a determinados tipos de material (água e metais, principalmente) e baixa escalabilidade, ou seja, quanto maior a área, mais leitores e etiquetas serão necessários [54]. Além disso, as técnicas descritas nesta seção dependem diretamente de um ambiente de radiofrequência bem controlado, ou seja, onde exista garantia de baixa probabilidade de fenômenos interferentes e alterações das propriedades eletromagnéticas dos dispositivos e objetos. Assim, os efeitos de multicaminho são minimizados, ou, pelo menos, bem definidos e conhecidos. A garantia de um ambiente de radiofrequência bem controlado pode ser almejada através do uso de uma câmara anecóica¹ ou de um cenário bem estruturado. Entretanto, estas possibilidades não são viáveis na maioria dos cenários do mundo real [76].

2.1.4 Fusão de Sensores

Fusão de sensores é o termo técnico empregado para a combinação de diferentes sensores em qualquer tipo de implementação que receba dados de diferentes tecnologias sensoriais simultaneamente. A fusão de sensores busca melhorar a precisão através da integração de diversos sistemas de localização, formando níveis de solução hierárquicos que possam ser combinados [37].

Métodos de fusão de sensores estão permitindo o uso de múltiplas tecnologias sensoriais em sistemas de localização. Na maioria das vezes, uma tecnologia de localização isolada não possui a capacidade adequada (por exemplo, precisão e escalabilidade) para satisfazer toda a demanda de requisitos das aplicações de IPSs. No entanto, um sistema baseado em fusão de sensores pode empregar sensores independentes para fornecer características agregadas e flexibilidade, não contempladas quando se utiliza uma tecnologia isolada [39].

Não foi encontrada na literatura uma lista taxativa de métodos de fusão de sensores. Assim, podem ser listados os métodos de fusão baseados no entendimento obtido durante o estudo dos trabalhos correlatos, são eles: extração da região de interesse (*Region of Interest* – ROI) e o filtro de Kalman.

¹ Câmaras anecóicas são salas blindadas que têm como objetivo reduzir o fenômeno da reflexão e outras interferências externas que possam influenciar o funcionamento de sistemas baseados em radiofrequência. A blindagem objetiva que se capte apenas o sinal emitido pelo equipamento em teste. A característica de ser anecóica significa que, as paredes da câmara, revestidas por materiais apropriados, não permitem a reflexão das ondas eletromagnéticas, livrando de reflexões o sinal principal emitido pelo equipamento em teste [17]

A extração do ROI utiliza diferentes sensores a fim de dividir o problema em partes menores. Este método exige que seja definida uma hierarquia aos sistemas envolvidos e que possuem sensores distintos. Neste método, primeiramente emprega-se uma determinada tecnologia sensorial cujo objetivo é extrair uma região estimada de localização do objeto alvo. A partir da determinação da região estimada de localização do objeto, um segundo sistema que utiliza um tipo de sensor distinto do primeiro tem seu escopo de atuação reduzido, abrangendo somente a dimensão do ROI. Este método de fusão é empregado em diversos trabalhos da área, cada um com abordagens distintas em relação à tecnologia sensorial, dimensão do ROI e níveis hierárquicos empregados.

O filtro de Kalman é um conjunto de equações matemáticas que implementam uma função de estimativa, minimizando a covariância do erro estimado. Este método tem sido objeto de diversas pesquisas e aplicado nas mais variadas áreas. O filtro de Kalman é empregado no controle de sistemas dinâmicos complexos, tais como processos de manufatura, aviação, e navegação assistida e autônoma. No controle de um sistema dinâmico, nem sempre é possível ou desejável medir cada variável que se deseja controlar. O filtro de Kalman fornece meios para inferir informações perdidas que podem ocorrer em medições com interferências (ruídos). O filtro de Kalman também é utilizado em sistemas dinâmicos não passíveis de controle humano, podendo realizar a predição do provável estado futuro do sistema. São exemplos de sistemas deste tipo o fluxo de rios durante enchentes e a trajetória de corpos celestes [33], [94].

2.2 Aprendizado de Máquina

Aprendizado de máquina é uma área da Inteligência Artificial dedicada ao desenvolvimento de técnicas computacionais sobre o aprendizado bem como a construção de sistemas capazes de adquirir conhecimento de forma automática. Um sistema de aprendizado é um programa de computador que toma decisões baseado em experiências acumuladas através da solução bem sucedida de problemas anteriores [61].

Murphy [62] define aprendizado de máquina como um conjunto de métodos que podem automaticamente detectar padrões nos dados, e assim utilizar padrões desconhecidos para estimar dados futuros, ou para executar outros tipos de tomada de decisão sob algum grau de incerteza.

O aprendizado de máquina é geralmente dividido em dois tipos: supervisionado e não-supervisionado. No aprendizado supervisionado é fornecido ao algoritmo de aprendizado um conjunto de amostras de treinamento para os quais o rótulo da classe

associada é conhecido. Em geral, cada amostra é descrita por um vetor de valores de características, ou atributos, e o rótulo da classe associada. O objetivo do algoritmo é construir um classificador que possa determinar corretamente a classe de novos exemplos ainda não rotulados, ou seja, amostras que não tenham o rótulo da classe. Para rótulos de classe discretos, esse problema é conhecido como classificação e para valores contínuos como regressão [61]. Por exemplo, em sistemas de localização *indoor*, um algoritmo de aprendizado de máquina pode ser treinado utilizando características conhecidas e as posições dos objetos. Posteriormente, o algoritmo pode inferir, com certo grau de certeza, a localização de um objeto alvo desconhecido. São exemplos de algoritmos de aprendizado supervisionado amplamente utilizados: redes neurais artificiais, *Support Vector Machines* (SVM), *k-Nearest Neighbors* (k-NN), regressão linear, regressão logística e árvores de decisão.

Quando um algoritmo de aprendizado supervisionado realiza uma estimativa, através das amostras disponíveis, é possível que a predição seja muito específica para o conjunto de treinamento utilizado. Como o conjunto de treinamento é apenas uma parte de todas as amostras do domínio, é possível induzir estimativas que melhorem seu desempenho no conjunto de treinamento, enquanto pioram o desempenho em amostras diferentes daquelas pertencentes ao conjunto de treinamento. Nesta situação, o erro (MSE) em um conjunto de validação independente evidencia um desempenho ruim da estimativa. Neste caso, diz-se que a predição ajustou-se em excesso ao conjunto de treinamento ou que houve um sobreajuste (*overfitting*) [61].

Referente ao tipo de aprendizado não-supervisionado, os dados consistem somente de amostras que não possuem um rótulo correspondente. O algoritmo analisa as amostras fornecidas e tenta determinar se algumas delas podem ser agrupadas de alguma maneira, formando agrupamentos ou *clusters*. Após a determinação dos agrupamentos, normalmente, é necessária uma análise para determinar o que cada agrupamento significa no contexto do problema que está sendo analisado [61]. Em sistemas de localização, este tipo de algoritmo pode auxiliar no agrupamento de regiões de interesse que serão analisados posteriormente por um sistema secundário. Alguns exemplos de algoritmos de aprendizado não-supervisionado são: k-means, agrupamento hierárquico e modelo oculto de Markov.

As próximas seções descrevem os algoritmos de aprendizado de máquina pesquisados e agregados ao desenvolvimento do IPS deste trabalho.

2.2.1 Regressão Linear Múltipla

A regressão linear é uma técnica que envolve a análise de dados amostrais para investigar e modelar a relação entre variáveis, sendo utilizada principalmente com o propósito de previsão. O modelo obtido através da regressão linear é denominado equação de regressão. Uma equação de regressão é uma expressão utilizada para estimar uma variável dependente y (resposta), em função das variáveis independentes (explanatórias), x_1, x_2, \dots , e é denominada equação de regressão de y para x_1, x_2, \dots [82].

A equação de regressão é uma equação para a linha de regressão (ou linha de melhor ajuste) que melhor modela os dados na correlação entre variáveis. Considerando o gráfico de dispersão e a linha de regressão traçada na Figura 10. Para cada ponto dado, d_i representa a diferença entre o valor y observado e o valor y previsto para um dado valor x na linha. Essas diferenças são chamadas de resíduos. Se o valor y observado é igual ao valor y previsto, então $d_i = 0$. De todas as linhas possíveis que podem ser desenhadas através de um conjunto de dados, a linha de regressão é a linha para a qual a soma dos quadrados de todos os resíduos ($\sum d_i^2$) é um mínimo [49].

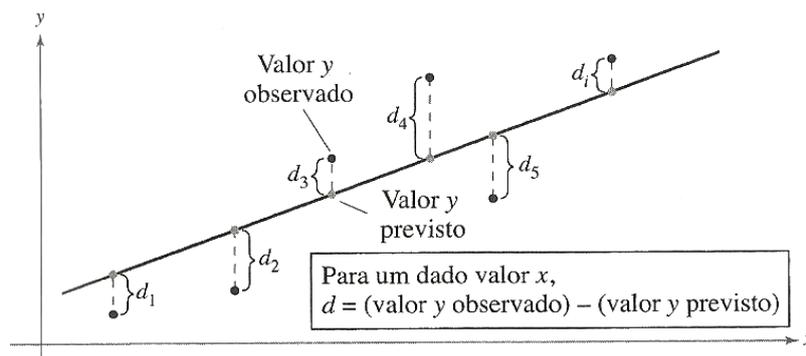


Figura 10 – Gráfico de dispersão e a linha de regressão linear. Fonte: [49].

Em muitos casos, um modelo de previsão necessita mais do que uma variável independente. A quantidade de variáveis experimentais medidas distingue os dois métodos de regressão existentes: a regressão linear simples e a regressão linear múltipla (RLM). Para uma regressão simples há somente uma variável independente x e uma dependente y . A regressão linear múltipla acontece quando o modelo regredido conta com uma variável dependente y e várias variáveis independentes x .

Um modelo básico de IPS baseado em RLM poderia empregar características do objeto e do ambiente (níveis de intensidade do sinal de rádio-frequência, por exemplo) como variáveis independentes, e a posição geográfica do objeto como variável dependente.

Uma das medidas de regressão mais encontradas na literatura é o R^2 (coeficiente de determinação). O R^2 representa o percentual de variação dos dados utilizados pelo modelo linear, ou seja, o percentual da variação de y que pode ser explicado pela relação entre x e y . Por exemplo, um R^2 igual a 0,81 significa que 81% da variação de y pode ser explicada pela relação entre x e y ; 19% restante da variação não é explicada, em razão de outros fatores ou erro da amostra. Um R^2 de 1,0 (100%) indica que a linha de regressão cruzou perfeitamente o conjunto de dados amostrados.

2.2.2 Redes Neurais Artificiais (RNA)

Redes Neurais Artificiais (RNA) são ferramentas de aproximação de funções que aprendem a relação entre variáveis dependentes e independentes, à semelhança da regressão linear ou outras abordagens mais tradicionais. A principal vantagem de uma rede neural é a sua capacidade de aproximar relações funcionais, particularmente quando as relações não são bem definidas ou não são lineares, o que torna difícil a utilização de métodos convencionais para a tentativa de prever as variações futuras dessas relações [36], [47].

Uma RNA é um modelo computacional que compartilha algumas das propriedades do cérebro: (i) o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem; (ii) forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido [36].

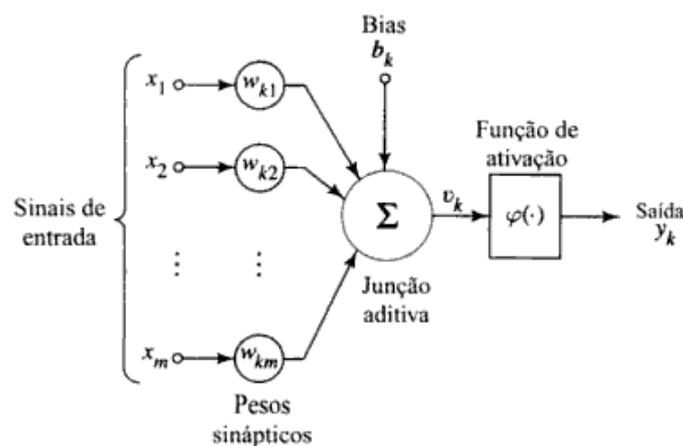


Figura 11 – Modelo não-linear de um neurônio artificial. Fonte: [36].

A Figura 11 exibe um modelo onde podem ser vistos: sinapses representadas pelas entradas e pesos sinápticos, somatório e função de ativação. Cada sinapse é caracterizada por um estímulo de entrada (x_j) multiplicado pelo seu peso sináptico (w_{kj}) correspondente. Depois desta multiplicação, cada sinal de entrada é somado e o

resultado é então aplicado a uma função de ativação (φ) que define a saída (y_k) do neurônio [73]. Tipicamente, o intervalo de saída de uma função de ativação está em um intervalo fechado $[0, 1]$ ou alternativamente $[-1, +1]$. O modelo também tem um bias (b_k) aplicado externamente, o bias tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativo, respectivamente [36].

Em termos matemáticos, o neurônio artificial pode ser descrito pelo seguinte par de equações:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (5)$$

$$y_k = \varphi(u_k + b_k) \quad (6)$$

onde x_1, x_2, \dots, x_m são sinais de entrada; $w_{k1}, w_{k2}, \dots, w_{km}$ são os pesos sinápticos do neurônio k ; u_k é a saída do combinador linear; b_k é o bias (tendência); φ é a função de ativação; e y_k é o sinal de saída do neurônio. Para o resultado da soma $u_k + b_k$ é dado o nome de “campo local induzido” [36].

A função de ativação define a saída de um neurônio em termos do campo local induzido, tal função deve simular as características não lineares do neurônio biológico. As funções de ativação mais utilizadas são [36], [73]:

- Função linear: é uma equação linear.
 - $f(x) = a.x$
- Função degrau: utilizada para valores binários.
 - se $x > 0$ $f(x) = 1$
 - se $x \leq 0$ $f(x) = 0$
- Função sigmóide: é a forma mais comum de ativação utilizada na construção de redes neurais. É definida como uma função estritamente crescente que exhibe um balanceamento adequado entre o comportamento linear e não-linear. A função logística é um exemplo de função sigmóide, definida por:

$$\varphi(v) = \frac{1}{1 + \exp(-\alpha v)} \quad (7)$$

onde v é o campo local induzido do neurônio e α é o parâmetro de inclinação da função sigmóide.

2.2.2.1 Arquiteturas de RNA

Basicamente, três classes de arquitetura de RNA podem ser classificadas: redes *feedforward* com camada única, redes *feedforward* com múltiplas camadas e redes recorrentes. As redes *feedforward* são geralmente representadas com várias camadas e, neste caso, são chamadas redes *feedforward* multicamadas. Os neurônios de saída constituem a camada de saída da rede. Os neurônios restantes constituem as camadas ocultas da rede. As unidades ocultas não são parte da saída ou da entrada da rede. A primeira camada oculta é alimentada pela camada de entrada (nós de fonte), as saídas resultantes da primeira camada oculta são por sua vez aplicadas à próxima camada oculta ou para a camada de saída da rede. A rede neural da Figura 12 é dita totalmente conectada, pois cada um dos nós de uma camada de rede está conectado a todos os nós da camada adjacente seguinte. Caso alguns dos elos de comunicação estiverem faltando, se diz que a rede é parcialmente conectada [36].

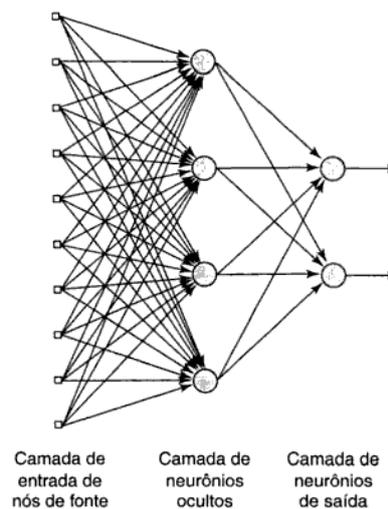


Figura 12 – Rede *feedforward* totalmente conectada com uma camada oculta e uma camada de saída. Fonte: [36].

A principal diferença entre uma rede neural *feedforward* e uma rede recorrente está no fato desta última ter pelo menos um laço de realimentação. Uma rede recorrente pode consistir de uma camada única onde cada neurônio esteja alimentando seu sinal de saída de volta para as entradas de todos os outros neurônios. A aplicação de realimentação permite que as redes recorrentes adquiram representações de estado, o que as torna dispositivos apropriados para aplicações tão diversas como previsão não-linear e modelagem, equalização de canais de comunicação, processamento de voz, entre outros [36].

Uma modelagem básica de sistema de localização *indoor* baseado em RNA pode empregar atributos do objeto e do ambiente (níveis de intensidade do sinal de rádio-frequência, por exemplo) como valores dos neurônios de entrada, e as coordenadas geográficas do objeto sendo representadas na camada de neurônios de saída. Maiores detalhes sobre modelagem de IPS baseado em RNA são apresentados na Seção 4.2.4.

2.2.2.2 Aprendizagem da Rede Neural

Uma rede neural aprende acerca do seu ambiente através de um processo interativo de ajustes aplicados a seus pesos sinápticos e níveis de bias. Idealmente, a rede se torna mais instruída sobre o seu ambiente após cada iteração do processo de aprendizagem [36].

O aprendizado no contexto das redes neurais é definido como um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. Os paradigmas de aprendizagem são [36], [73]:

- **Supervisionado:** o paradigma supervisionado também é conhecido por aprendizado com professor, isto porque se pode considerar o professor como tendo conhecimento sobre o ambiente, com este conhecimento sendo representado por um conjunto de exemplos de entrada-saída. A partir da entrada a rede realiza seu processamento e a saída obtida é comparada com a saída esperada. Caso não sejam iguais, um processo de ajuste de pesos é aplicado buscando-se um erro mínimo ou aceitável. Este processo de ajuste é conhecido como método de otimização ou algoritmo de treinamento.
- **Não-supervisionado:** como o nome implica, não há um elemento externo para supervisionar o processo de aprendizagem. Isto significa que não há exemplos rotulados da função a ser aprendida pela rede. Este paradigma possui duas divisões: aprendizado por reforço que consiste no mapeamento entrada-saída através da interação com o ambiente, e aprendizagem não-supervisionada ou auto organizada onde, a partir de métricas de qualidade do aprendizado ocorre a otimização dos parâmetros livres da rede.

O algoritmo de retro-propagação (*backpropagation*) é um algoritmo de aprendizado supervisionado baseado na correção de erros. O objetivo do algoritmo é encontrar na superfície de erro um mínimo global, ou seja, valores para os pesos sinápticos que

minimizam o erro da rede [36]. Uma das medidas de desempenho utilizadas em redes neurais *feedforward* é o erro médio quadrático (*Mean Square Error* - MSE), que representa a média quadrática do erro entre os valores previstos (p) pela rede e os valores reais (r) que estão sendo ensinados, aplicado a n amostras de treinamento. A equação (8) exhibe o cálculo do MSE [85].

$$mse = \frac{1}{n} \sum_{i=1}^n (e_i)^2 = \frac{1}{n} \sum_{i=1}^n (r_i - p_i)^2 \quad (8)$$

No algoritmo supervisionado, um sinal de erro se origina em um neurônio de saída da rede e se propaga para trás (camada por camada) através da rede. Sua computação por cada neurônio envolve uma função dependente do erro.

Para ilustrar como o treinamento funciona, pode ser considerado o método mais simples de otimização, denominado *gradient descent*. O algoritmo atualiza os pesos e bias da rede a partir da escolha do negativo do gradiente da função de erro como direção para cada passo de otimização [91]. Uma iteração do algoritmo pode ser escrita como

$$x_{k+1} = x_k - a_k g_k \quad (9)$$

onde x_k é um vetor dos pesos e bias atuais, g_k é o gradiente atual e a_k é a taxa de aprendizado. A taxa de aprendizado define o tamanho de cada passo de otimização.

A equação é processada por diversas iterações (épocas) até que exista uma convergência no desempenho da rede. Entretanto, de maneira geral, não se pode demonstrar que o algoritmo de retro-propagação convergiu, sendo que não existem critérios bem definidos para encerrar seu processamento. Existem alguns critérios razoáveis, que podem ser [36], [85]:

- *goal*: quando, percorrendo a superfície de erro, for encontrado um mínimo;
- *min_grad*: a magnitude do gradiente atingir um mínimo. O gradiente se torna um valor muito pequeno à medida que o treinamento alcança um desempenho mínimo;
- *epochs*: número máximo de iterações (épocas) atingir um limite;
- *max_fail*: número de iterações sucessivas em que o desempenho das amostras de validação não diminuam.

Em [85] pode ser encontrada uma lista com doze métodos de otimização para treinamento supervisionado. Devido à eficiência e disponibilidade, o algoritmo de retro-

propagação denominado Levenberg-Marquardt [86] é empregado neste trabalho. Este algoritmo aparenta ser o mais rápido método de treinamento para redes neurais de tamanho moderado (até centenas de pesos). Ele também está implementado de forma eficiente na ferramenta Matlab [86], pois a solução de aproximação é uma função embutida no software.

2.2.3 *Support Vector Machines (SVM)*

As máquinas de vetores de suporte (*Support Vector Machines - SVM*) constituem uma técnica de aprendizado de máquina solidamente fundamentada pela teoria de aprendizado estatístico, desenvolvida por Vapnik [90]. Essa teoria estabelece uma série de princípios que devem ser seguidos na obtenção de classificadores com boa generalização, definida como a sua capacidade de prever corretamente a classe de novos dados do mesmo domínio em que o aprendizado ocorreu [53].

As aplicações de SVM em problemas de classificação de padrões e também a sua extensão para o tratamento de problemas de regressão (*Support Vector Regression – SVR*) têm se mostrado altamente competitivas frente a outros modelos de aprendizado de máquina [3]. Embora o SVM possua vantagens, algumas limitações também podem ser listadas, tais como a dificuldade na escolha de parâmetros e a baixa velocidade do treinamento, que é menor em relação a outras técnicas, como redes neurais.

Inicialmente, é apresentado o algoritmo SVM com margens rígidas, que trabalha apenas com dados linearmente separáveis, o que faz com que sua aplicação não se estenda a muitas situações do mundo real. Na etapa de treinamento fornecem-se exemplos de entrada e saída $\{x_i, y_i\}$, para $i=1, \dots, n$, cujas coordenadas são conhecidas. O objetivo principal desta fase é ajustar os parâmetros livres do sistema e então, determinar uma função que relacione os respectivos valores de x_i e y_i . Para isso, penalizam-se os erros de estimativa e valorizam-se acertos, de forma que o sistema retenha informações relevantes e adquira conhecimento. Na fase de teste, utiliza-se um conjunto de dados sobre o qual o método será efetivamente aplicado. Como são conhecidos apenas os valores de entrada dessa amostra, o intuito é utilizar a função determinada no treinamento para gerar as saídas desejadas e encontrar os resultados. A partir de um conjunto de vetores de treinamento pertencentes a duas classes linearmente separáveis, a tarefa do SVM é determinar uma função que classifique corretamente os dados de testes. Em

outras palavras, significa encontrar um hiperplano¹ separador com boa capacidade de generalização [8].

Em um conjunto de treinamento S que contém pontos de duas classes, uma SVM separa as classes através de um hiperplano determinado por certos pontos de S , denominados de vetores suporte. Existem várias formas de separar a amostra de dados, sem erros (Figura 13). Entretanto, há somente uma maneira na qual se maximiza a margem de separação, conforme ilustrado na Figura 13d. Define-se margem de classificação como a distância entre o hiperplano separador e o vetor mais próximo de cada classe (Figura 14). O hiperplano que maximiza a distância entre estas duas margens é conhecido como hiperplano ótimo [8], [50], [51].

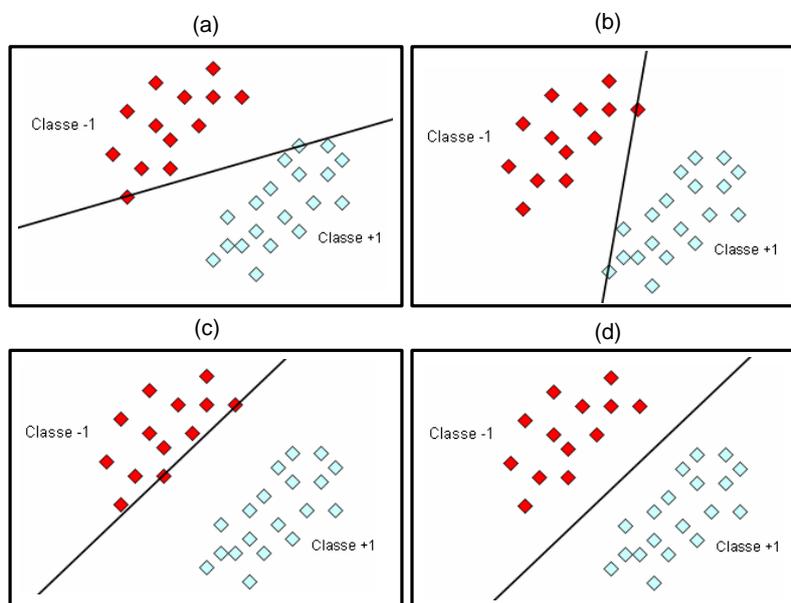


Figura 13 – Possibilidades de separação através de hiperplanos distintos. Fonte: [2].

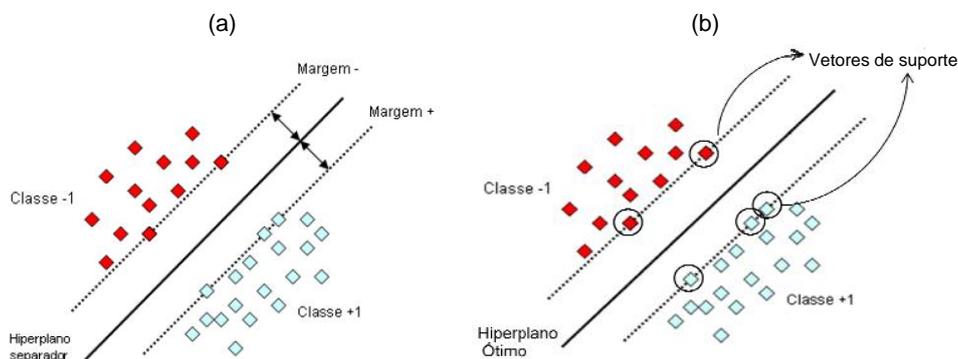


Figura 14 – Separação ótima de dois conjuntos linearmente separáveis (a) e vetores de suporte (b). Fonte: [2], [8].

¹ É a generalização do plano em diferentes números de dimensões. Em um espaço tridimensional um hiperplano é um plano habitual. Em um espaço bidimensional, um hiperplano é uma reta. Em um espaço monodimensional, um hiperplano é um ponto.

Em casos que o conjunto de treinamento não é linearmente separável, os quais caracterizam a maioria dos exemplos reais, o problema poderá separar vetores incorretamente, desviando seu valor na função de decisão. Embora nestas situações não seja possível construir um hiperplano separador sem erros de classificação, é possível encontrar aquele que minimiza a probabilidade de erro junto às amostras de treinamento. Para tratar problemas desta classe, introduz-se variáveis de folga $\xi_i \geq 0$ associadas a cada vetor de treinamento, que meçam os desvios encontrados em relação às condições ideais de separação (margens flexíveis). Logo, se um vetor está separado corretamente tem-se que $\xi_i = 0$ [8].

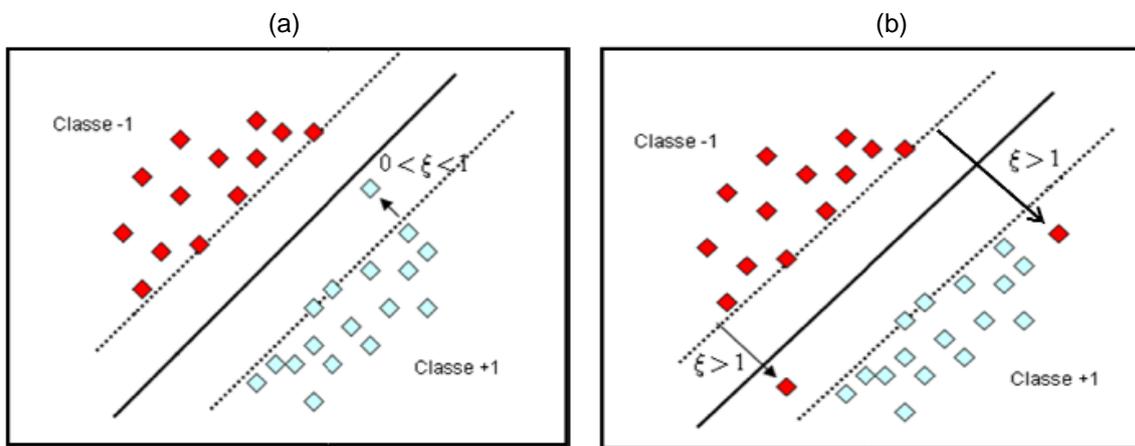


Figura 15 – Violação das margens de classificação no lado correto (a) e no lado incorreto da superfície de decisão (b). Fonte: [8].

Constata-se que esta situação pode ocorrer de duas maneiras distintas (Figura 15). A primeira possibilidade consiste no ponto x_i estar localizado dentro da região de separação e no lado correto da classificação. Neste caso considera-se que a variável de folga vale $0 < \xi_i \leq 1$. Na segunda possibilidade, o ponto x_i localiza-se no lado incorreto da superfície de decisão, podendo estar dentro ou fora da região de separação. Para este caso, a variável de folga correspondente terá valor $\xi_i > 1$ [8].

Um dos principais parâmetros na configuração de um modelo SVM diz respeito a constante de regularização (C). Em resumo, a constante de regularização estabelece o equilíbrio entre a complexidade do modelo e o erro de treinamento, determinando o quão sensível o modelo será à presença de pontos mal classificados. Valores grandes de C atribuem maior peso ao número de erros (permitindo poucos erros) e menor peso à margem do hiperplano (gerando uma margem pequena). Nessa condição, poucos erros são permitidos e gera-se uma menor margem de classificação. Ao se utilizarem valores pequenos de C , a situação inversa ocorre [13].

Em ambos os modelos (margens rígidas e flexíveis) foi considerado que a separação dos dados no espaço de entrada é realizada por funções lineares. No entanto, essa consideração pode ter baixo desempenho em problemas cujos pontos de treinamento não são linearmente separáveis. Uma das principais características do algoritmo SVM é a possibilidade de se projetar os dados de entrada para um espaço de dimensão maior, onde a separação linear torna-se possível. A Figura 16 mostra um exemplo dessa transformação, onde os dados não podem ser separados linearmente no espaço de entrada, mas podem ser no espaço característico [8].

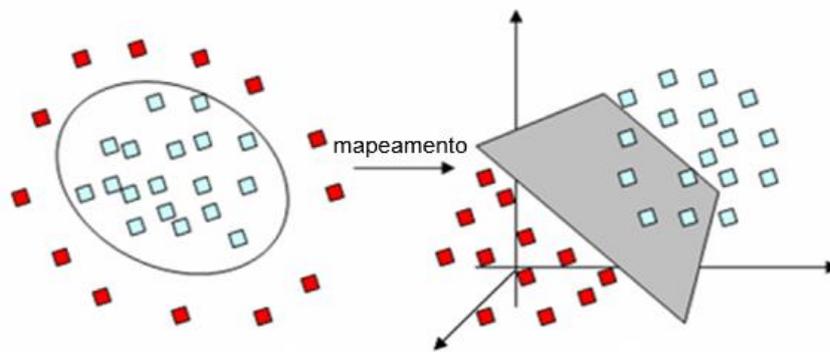


Figura 16 – Mapeamento não-linear do espaço de entrada para o espaço de características. Fonte: [10].

Na prática, a modificação necessária para implementar o referido mapeamento é mínima. Entretanto, o processamento necessário para esta transformação pode se tornar computacionalmente inviável quando aplicado diretamente ao espaço de características de alta dimensionalidade. Com o objetivo de se evitar o custo computacional, o mapeamento pode ser realizado pelas chamadas funções *kernel*, que dependem somente de variáveis do espaço de entradas. As funções *kernel* mais usadas são: linear, polinomial, função de base radial (RBF) e sigmoideal [13], [50].

2.2.3.1 Support Vector Regression (SVR)

Uma das maiores vantagens do SVM é a sua flexibilidade. Utilizando os conceitos de maximização de margem, dualidade e funções *kernel*, pode-se adaptar o problema de classificação binária (apenas duas classes), para resolver outros tipos de problema, como a regressão. Como o problema de localização de objetos depende de rótulos de classe discretos, ou seja, as coordenadas estimadas são valores contínuos, os vetores de suporte aplicados à regressão (SVR) são empregados neste trabalho [3].

De maneira geral, no SVR ocorre uma reversão do propósito dos modelos de classificação, ou seja, ao invés de forçar as amostras a se posicionarem o mais distante

possível do hiperplano ótimo, no espaço de características, o objetivo agora é posicionar as amostras o mais próximo possível do hiperplano ótimo, não importando o lado do hiperplano que a amostra se encontre. Apesar dos objetivos serem opostos, a ideia é a mesma: maximizar a capacidade de generalização [51].

Considera-se um conjunto de treinamento $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)\} \subseteq (X \times Y)^k$ no qual $x_i \subseteq \mathbb{R}^n$ é o vetor do espaço de entrada X . Porém, como se trata de um problema de regressão, as saídas desejadas podem assumir qualquer valor real, isto é, $y_i \subseteq \mathbb{R}$ para todo $i=1, \dots, k$. O objetivo do SVR é encontrar uma função $f(x)$ que apresente no máximo um desvio ε em relação aos valores alvo y_i . Em outras palavras, procura-se uma função com uma margem de erros caracterizada pelo intervalo $[y_i - \varepsilon, y_i + \varepsilon]$. Desta forma, desvios são permitidos desde que não ultrapassem a margem especificada [8].

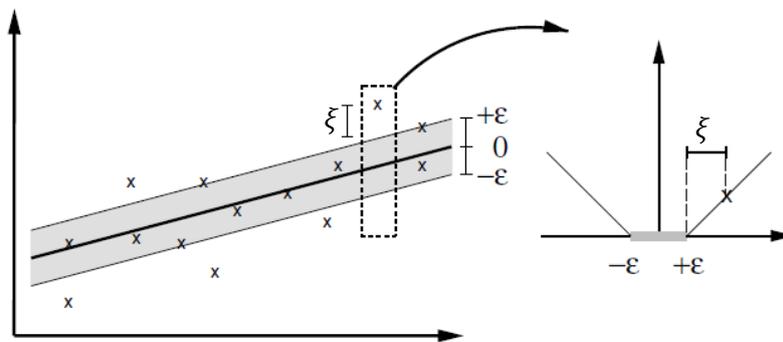


Figura 17 – Função de perda ε -insensitive. Fonte: [81].

Na regressão, há uma modificação da formulação original da função objetivo para que o erro, medido pela distância do valor estimado em relação ao valor real, é igual a zero para valores pequenos desta distância, e de valor crescente quando a distância ao valor real é maior do que um determinado limiar [3]. Uma função de perda introduz variáveis de folga não negativas ξ_i, ξ_i^* , cujo objetivo é penalizar dados que se situem fora da margem $|f(x_i) - y_i| \leq \varepsilon$. Esta função de perda, denominada ε -insensitive, é apresentada graficamente na Figura 17, onde somente os pontos localizados fora da região sombreada contribuem para o valor custo da função.

A Figura 18 descreve os passos do SVR. A entrada x (para a qual uma previsão deve ser feita) e os vetores de suporte $x_i, i=1, \dots, n$ são mapeados de forma não-linear pela função Φ para um espaço característico, onde as funções de *kernel* $K(x; x_i)$ são calculadas. Finalmente, os resultados são combinados através dos pesos e somados com o termo constante b para fornecer a previsão de x [19].

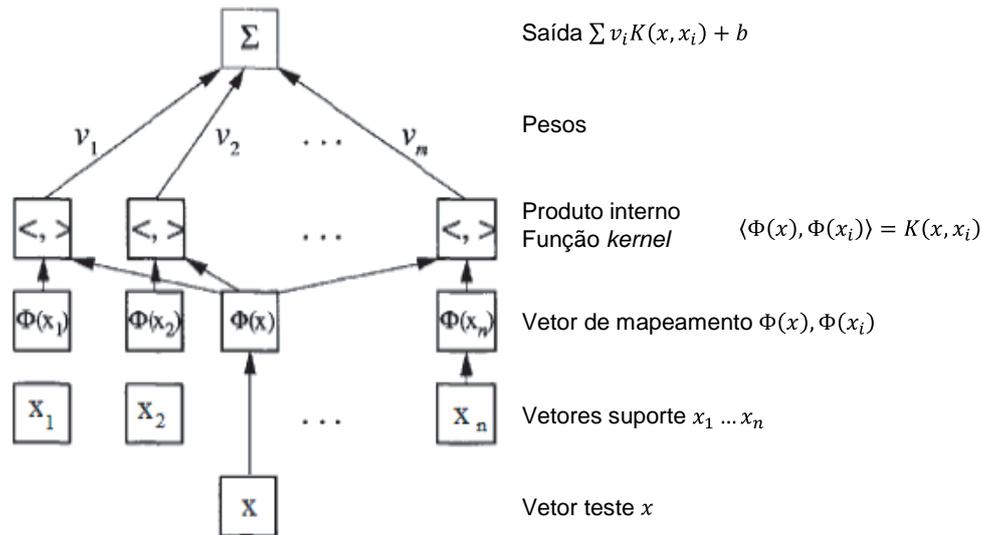


Figura 18 – Arquitetura de uma SVR. Fonte: [19], [81].

A Figura 19 ilustra resultados de aplicações do SVR, onde foram utilizados os mesmos dados de treinamento, a mesma constante de regularização $C=10$ e a mesma margem $\epsilon=20$, variando-se apenas as funções *kernel*. Dentre os resultados, o *kernel* polinomial não homogêneo (Figura 19b) obteve o melhor desempenho [8].

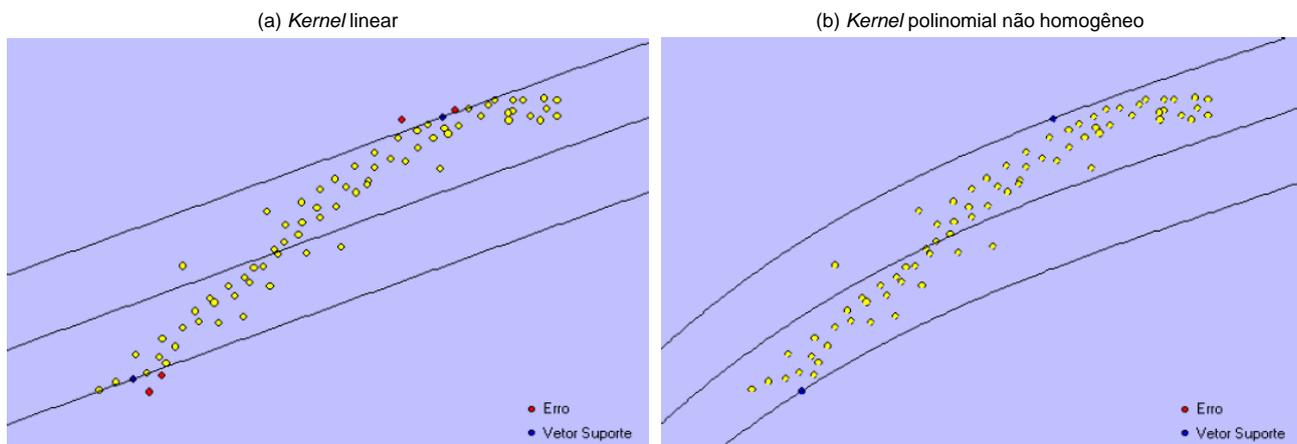


Figura 19 – Resultado ilustrado do SVR com duas funções *kernel* distintas. Fonte: [8].

2.2.4 K-Means

O agrupamento k-means é um método de particionamento que separa os dados em k *clusters* mutuamente exclusivos. O k-means atua sobre elementos observados e cria uma camada única de *clusters*. O k-means trata cada elemento observado como um objeto que possui uma posição no espaço. Ele encontra a partição cujos objetos dentro de cada *cluster* estão mais próximos um do outro, e mais longe de objetos presentes em outros *clusters* [87].

Dado um conjunto de elementos observados (x_1, x_2, \dots, x_n) , onde cada elemento é um vetor real de dimensão d , o agrupamento k-means busca particionar os n elementos em $k(\leq n)$ conjuntos $J = \{J_1, J_2, \dots, J_k\}$.

Cada *cluster* na partição é definido pelos seus membros (elementos observados) e pelo seu centroide, ou centro. O centroide de cada *cluster* é um ponto no qual a soma das distâncias a partir de todos os membros do *cluster* é mínima. O k-means usa um algoritmo iterativo, cujo objetivo é reduzir as distâncias entre cada objeto e o centroide do seu *cluster*, sendo aplicado a todos os *clusters*. Este algoritmo move os objetos entre os *clusters* até que a soma das distâncias não possa mais ser reduzida [87]. A Figura 20 exibe os *clusters* resultantes após três iterações. O algoritmo tem como objetivo minimizar uma função objetivo

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (10)$$

onde $\|x_i^{(j)} - c_j\|^2$ é a medida de distância escolhida entre o ponto $x_i^{(j)}$ e o centro do *cluster* c_j , k é o número de *clusters* e n é a quantidade de elementos observados [56].

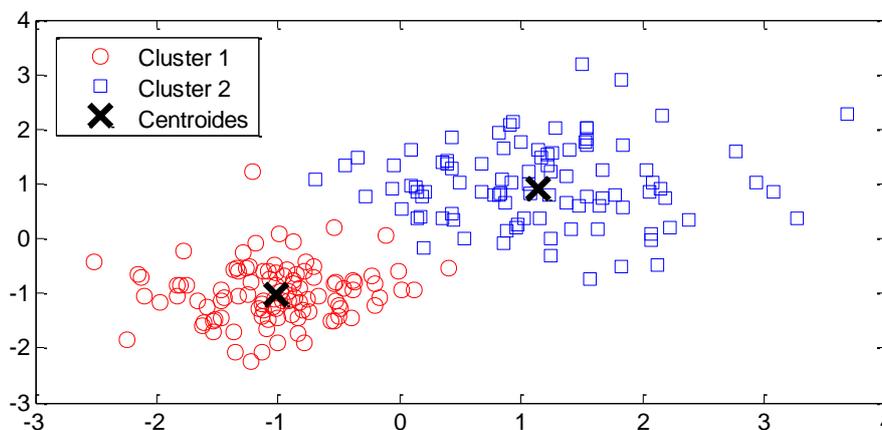


Figura 20 – Resultado de um agrupamento k-means. Os pontos com formas geométricas iguais estão no mesmo *cluster*. O centro dos *clusters* é indicado com uma cruz. Fonte: o autor.

3 TRABALHOS CORRELATOS

Em uma pesquisa bibliográfica sobre sistemas de localização *indoor* podem ser encontradas publicações envolvendo as mais diversas técnicas e tecnologias. Visando relacionar as publicações encontradas que seguem o foco deste trabalho de doutoramento, o escopo dos trabalhos analisados foi reduzido somente a IPSs RFID que empreguem modelos de aprendizado de máquina, ou que sejam híbridos e também utilizem técnicas de análise visual. A próxima seção relaciona de forma sucinta os IPSs RFID baseados em modelos de aprendizado de máquina. Na seção posterior, os sistemas híbridos que se assemelham às tecnologias e técnicas deste trabalho são apresentados.

3.1 IPSs baseados em RFID e Aprendizado de Máquina

Em Landmarc [66], etiquetas RFID de referência são distribuídas no ambiente e os valores de RSSI são medidos pelos equipamentos leitores. Ao mesmo tempo, etiquetas em posições desconhecidas são lidas e seus valores de RSSI são utilizados em um algoritmo k-NN para encontrar as etiquetas de referência mais próximas, estimando a posição das etiquetas desconhecidas (alvo). As etiquetas de referência obrigatoriamente devem estar presentes no cenário durante a localização das etiquetas desconhecidas. Na configuração padrão, quatro leitores RFID, 16 etiquetas de referência e oito etiquetas alvo são distribuídas no ambiente, conforme ilustrado na Figura 21. São empregadas somente etiquetas RFID ativas (longo alcance), operando na frequência de 308 MHz. O erro médio da localização foi próximo de 1 m com 50% de probabilidade de sucesso. O erro máximo observado foi igual a 2 m. Também foi constatado que utilizando um número maior de leitores e antenas, a precisão e o *overhead* de processamento do sistema aumentam.

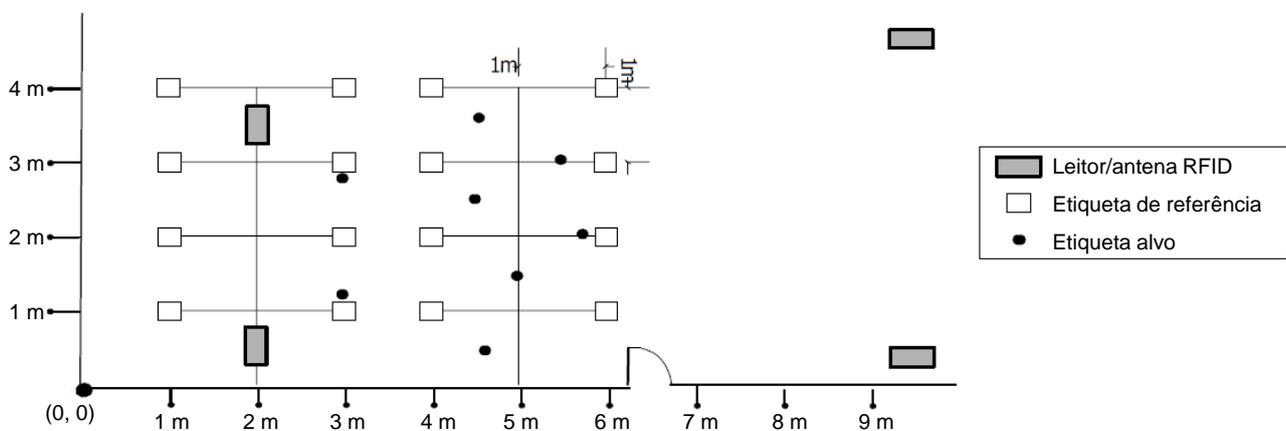


Figura 21 - Ambiente de experimentos do Landmarc: dimensões do cenário e organização dos componentes. Fonte: [66].

Em [12], a técnica Landmarc é comparada a um modelo de localização baseado em RNA (Figura 22a). Durante a fase de treinamento, para cada antena do leitor, o valor de RSSI das etiquetas de referência alimentam a entrada da rede. Na camada de saída, as coordenadas (x, y) e o ângulo de orientação (θ) das etiquetas é fornecido. No ambiente de experimentos (Figura 22b), as etiquetas de referência são separadas por uma distância de 50 cm. Os resultados demonstram que a precisão é 7 cm melhor que o sistema Landmarc em cenários sem obstáculos. Em cenários mais complexos (posicionamento de etiquetas mais denso, obstáculos, etc.), o algoritmo k-NN do Landmarc superou a rede neural em 20-25 cm.

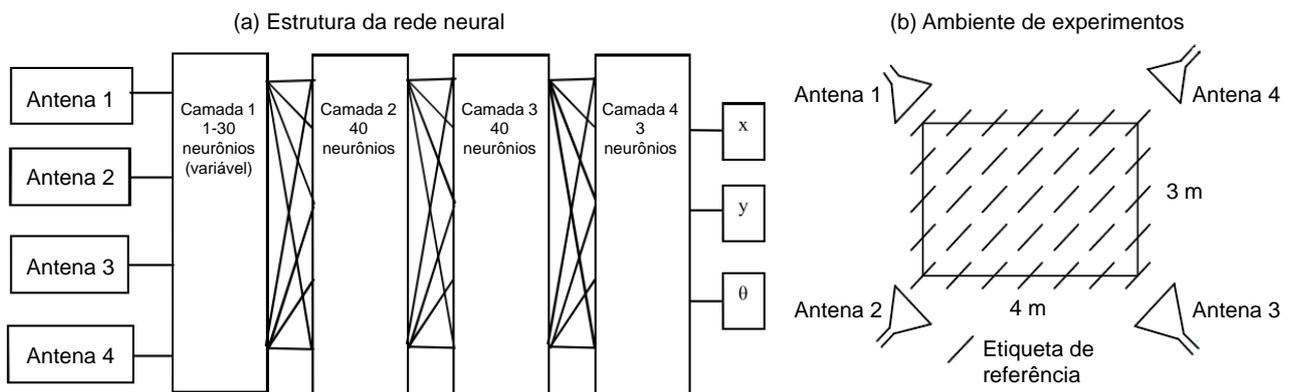


Figura 22 – Rede neural proposta (a) e configuração de ambiente dos experimentos (b).

Fonte: [12].

Em [74], um modelo de propagação que considera as perdas no caminho é usado para gerar o *fingerprint* RSSI no ambiente interno. O modelo empregado considera a atenuação do sinal causada por obstáculos e paredes, além das áreas de sombreamento¹. O RSSI de cada antena do leitor e as coordenadas das etiquetas são fornecidas como entradas e saídas de uma rede neural, respectivamente. Uma ferramenta de simulação é proposta baseada em etiquetas RFID ativas e considerando antenas isotrópicas² (inviáveis em ambientes reais). Os resultados utilizando três antenas demonstraram uma precisão de até 2 m com 80% de probabilidade de sucesso. Utilizando nove antenas, o percentual de sucesso foi de 98% na mesma precisão.

Sala *et al* [75] ampliaram o trabalho anterior para suportar agrupamentos de antenas. É proposto um método de agrupamento automático baseado em k-NN, sendo

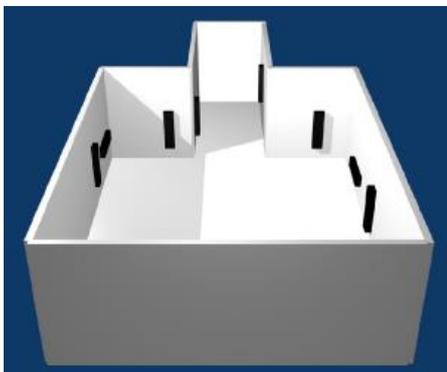
¹ O sombreamento é caracterizado por variações aleatórias na potência do sinal recebido devido às obstruções no percurso de propagação do sinal. Estas variações também são causadas por alterações nas superfícies refletoras e objetos dispersivos.

² Uma antena isotrópica irradia os sinais uniformemente em todas as direções. Este tipo de antena existe somente na teoria (não existe antena ideal), e sua finalidade é servir como padrão de referência na medição de outras antenas.

uma rede neural treinada para cada grupo. A rede neural é alimentada pelos valores de RSSI de cada antena do grupo e define as coordenadas (x, y) como saída. A precisão do sistema foi de 0,78 m com 90% de sucesso utilizando nove antenas na configuração com sete agrupamentos. Da mesma forma que o trabalho anterior, todos os experimentos foram realizados através de um simulador baseado em modelos de propagação, não sendo executados ensaios em ambientes reais.

Wille *et al* [96] apresentam um modelo de localização baseado em *Support Vector Regression* (SVR) aplicado a sistemas de cirurgia assistida por computador (navegação cirúrgica). Diferente dos IPSs RFID tradicionais, o deslocamento da fase¹ é utilizado como indicador para treinar e executar o modelo SVR. Os experimentos foram realizados dentro de um recipiente plástico (50 cm x 90 cm x 20 cm), projetado para emular o corpo e a cabeça humana. Oito antenas conectadas a um leitor RFID operando na frequência de 868 MHz foram posicionadas nas paredes internas do recipiente, conforme ilustrado na Figura 23a. Em um dos testes de validação (Figura 23b), um equipamento CNC (Comando Numérico Computadorizado) moveu uma etiqueta RFID passiva na trajetória que compreende o formato de um cubo (3 cm x 3 cm x 3 cm). A cada 1 mm, foram coletadas amostras do deslocamento da fase da etiqueta, medido pelas antenas presentes. A partir das amostras, um subconjunto de dados foi utilizado para treinar o modelo SVR e o conjunto completo foi empregado para verificar se as estimativas do modelo estavam corretas. Dois subconjuntos de treinamento com resolução de 5 mm e 10 mm foram avaliados. Os resultados demonstraram uma precisão de 1,3 mm e 6,8 mm nas resoluções de 5 mm e 10 mm, respectivamente.

(a) Leiaute do recipiente e antenas RFID



(b) Distribuição das amostras no formato de cubo dentro do recipiente

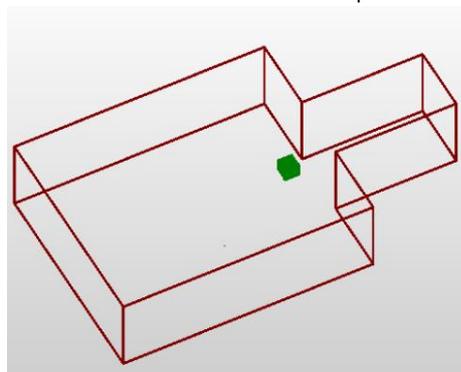


Figura 23 – Desenho do recipiente plástico e posicionamento das antenas RFID (a) e o cubo (verde) representando as amostras coletadas nos experimentos (b). Fonte: [96].

¹ Deslocamento da fase é a diferença, expressa em ângulo ou tempo, entre duas ondas eletromagnéticas que tenham a mesma frequência e em referência ao mesmo ponto no tempo.

Em [48], a abordagem Landmarc é combinada a uma rede neural com retro-propagação. Inicialmente, o Landmarc utiliza os valores de RSSI medidos para estimar as coordenadas da etiqueta alvo. Dado que a relação entre o RSSI e a distância é dinâmica, a rede neural ajusta estas coordenadas visando aumentar a precisão da localização. As entradas da RNA proposta são as coordenadas estimadas pelo Landmarc, enquanto as saídas são as coordenadas (x, y) ajustadas. Na fase *offline*, a rede neural aprende a relação não-linear entre as coordenadas calculadas pelo Landmarc e as coordenadas reais das etiquetas alvo. Na fase *online*, o processo de localização é executado em três fases: (i) os valores de RSSI da etiqueta alvo são agrupados através da média simples, resultando em um único valor de RSSI; (ii) o Landmarc estima a posição da etiqueta alvo; (iii) a posição estimada pelo Landmarc é utilizada como entrada para a rede neural treinada, fornecendo como saída a posição otimizada do alvo. Os resultados demonstraram um erro de 56 cm no cenário em que a distância entre as etiquetas de referência era de 30 cm. Em comparação ao Landmarc, o método proposto aprimorou a precisão em 31%.

3.2 Sistemas Híbridos de Localização Indoor

Nesta seção são abordados sete trabalhos correlatos que apresentam sistemas híbridos de localização *indoor* baseados na tecnologia RFID e técnicas de análise visual. As próximas seções descrevem estes trabalhos, considerando principalmente os detalhes que se destacam na implementação de cada um. A Seção 3.2.8 apresenta uma discussão sobre os trabalhos analisados e a correlação destes com o trabalho de doutoramento proposto.

3.2.1 Navegação de Robôs Móveis Baseada em RFID e Sensor Visual

Em [11] é apresentado um IPS aplicado à navegação de um robô, fazendo com que ele possa se locomover autonomicamente até um determinado destino. Nesta proposta, o robô móvel deve ter a capacidade de reconhecer a região em que está e também obter informações de orientação e a sua posição exata. As técnicas de localização empregadas nesta proposta basearam-se nas abordagens de proximidade, análise de padrões e angulação.

No método apresentado, o ambiente interno é dividido em regiões definidas previamente no sistema. Assim, a tecnologia RFID é empregada visando fornecer a região em que o robô está localizado no ambiente, denominada “localização global”. A implementação da tecnologia RFID se deu através de um leitor acoplado ao robô e de etiquetas ativas afixadas nas paredes de cada região (Figura 24). A ideia central da

proposta é a associação de diferentes pesos a cada uma das etiquetas detectadas pelo leitor, sendo que, quanto mais distante uma etiqueta está de outra região, maior será o respectivo peso. Desta forma, após uma varredura executada pelo sistema RFID, o algoritmo realiza a soma dos pesos obtidos em cada região. A região atual do robô é aquela que obtiver o maior peso.

Na proposta, recursos de visão computacional são empregados visando realizar um refinamento da localização global fornecida pelo modelo anterior. Uma *webcam* é acoplada ao robô, sendo utilizada inicialmente no processo de *setup*, chamado de construção do VisualMap. Neste processo, o robô é posicionado no centro de cada região e faz um giro de 360 graus. Para cada imagem capturada pela *webcam*, uma base de dados armazena a imagem, ângulo e a região correspondente.

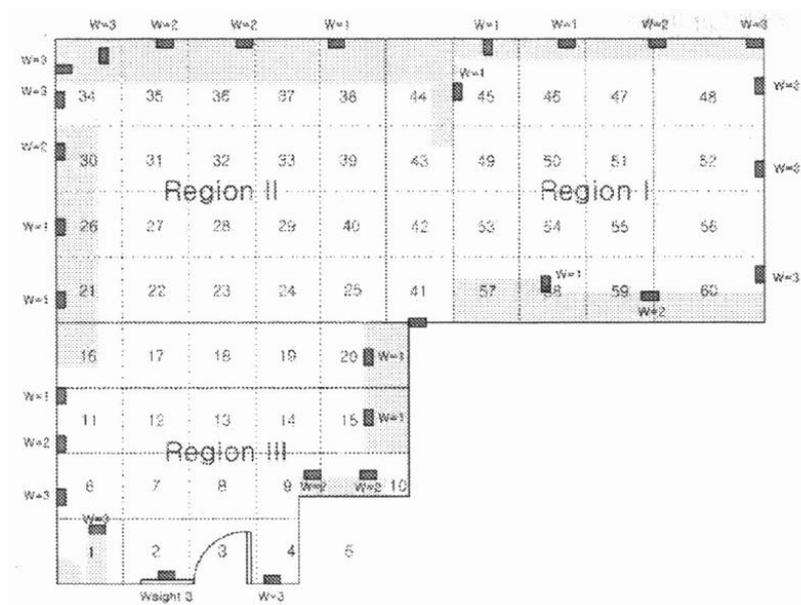


Figura 24 – Ambiente interno dividido em regiões e os pesos (W) das etiquetas afixadas.

Fonte: [11].

O processo de refinamento da localização global depende de uma imagem alvo do cenário, informado previamente pelo usuário do sistema. Esta imagem alvo pode ser uma foto da parede ou objeto em que o robô deve ser posicionar em frente. Assim, com o robô posicionado dentro da região adequada e tendo a base de dados sido coletada, uma técnica de análise de padrões é utilizada visando encontrar a imagem alvo dentro daquela região. O algoritmo *Scale-Invariant Feature Transform* (SIFT) é empregado a fim de realizar o reconhecimento da imagem alvo. Este algoritmo fornece estabilidade em ambientes com variação na iluminação e também possui bom funcionamento na presença de alterações de escala e rotação da imagem.

Por fim, imagens adjacentes à imagem alvo são reconhecidas e uma variação da técnica de angulação é aplicada ao sistema. A partir da captura das imagens adjacentes e das informações armazenadas na base de dados, funções trigonométricas são aplicadas visando obter a distância entre o robô e o marcador alvo.

Os resultados apresentados no trabalho baseiam-se na comparação entre a posição real do robô e a estimativa de localização fornecida pelo sistema proposto. Esta comparação é feita a partir das células da grade pré-definida no ambiente, ilustradas na Figura 24. Assim, é demonstrada uma precisão de 23 cm com probabilidade de sucesso de 91,5%.

3.2.2 RF Vision

Em [18] é apresentado um IPS, denominado RF Vision, que combina as tecnologias RFID, óptica e laser. O sistema de localização proposto permite que um robô se aproxime e capture (agarrando com um braço mecânico) um determinado objeto presente no ambiente. Para isto, um leitor RFID (com duas antenas), uma câmera e um *laser scanner* são acoplados ao robô. Uma técnica de localização baseada em análise de padrões é proposta a partir da tecnologia RFID. A localização visual foi desenvolvida a partir da técnica de histograma de cores das imagens. As informações da tecnologia laser são consideradas a fim de contemplar a escalabilidade tridimensional do ambiente. Por fim, um método de fusão baseado em Inferência Bayesiana é modelado e aplicado às respostas dos sensores.

No sistema proposto, inicialmente o robô varre o ambiente a fim de enumerar os objetos que possuem etiquetas RFID passivas afixadas. Baseado nesta resposta, o robô fornece uma interface para o usuário selecionar qual objeto deve ser capturado. Para a etiqueta do objeto selecionado pelo usuário, o sistema gera uma imagem contendo as informações de RSSI. Esta imagem é gerada da seguinte forma:

- i. É selecionada a antena RFID que está melhor direcionada para a etiqueta alvo, ou seja, a antena que obteve o RSSI mais alto para aquela etiqueta.
- ii. Inicia-se a leitura da etiqueta, armazenando-se o RSSI em uma base de dados.
- iii. Durante a leitura e registro do RSSI, a antena é movida mecanicamente nos sentidos horizontal (para esquerda e para direita) e vertical (para cima e para baixo). A posição atual da antena a cada movimento também é armazenado na base de dados.

- iv. Os resultados da base de dados são mapeados para uma imagem. Conforme ilustrado na Figura 25, um RSSI elevado é representado por cores mais quentes e um baixo RSSI é mapeado para cores mais frias. A fim de suavizar a imagem, um filtro gaussiano é aplicado utilizando um desvio de 45 *pixels*.

A partir da imagem RSSI, a intensidade de cada *pixel* é normalizada para a faixa de valores [0,0; 0,1], criando assim uma das matrizes de probabilidades que será utilizada pelo método de fusão de sensores.

Em relação à localização visual, uma câmera com resolução de 640x480 captura imagens do cenário e gera a matriz de probabilidades através de uma técnica de histograma de cores. Nesta técnica, regiões da imagem em que as cores mais se assemelham às cores do objeto são mapeadas para a matriz de probabilidade com um valor alto (próximo de 1), enquanto as demais regiões recebem valores próximos de zero. O sistema conhece a cor predominante do objeto, pois uma fotografia deste objeto é armazenada previamente no sistema. O sistema também conhece a cor predominante do cenário em geral, pois capturou imagens do ambiente enquanto o robô se movimentava.

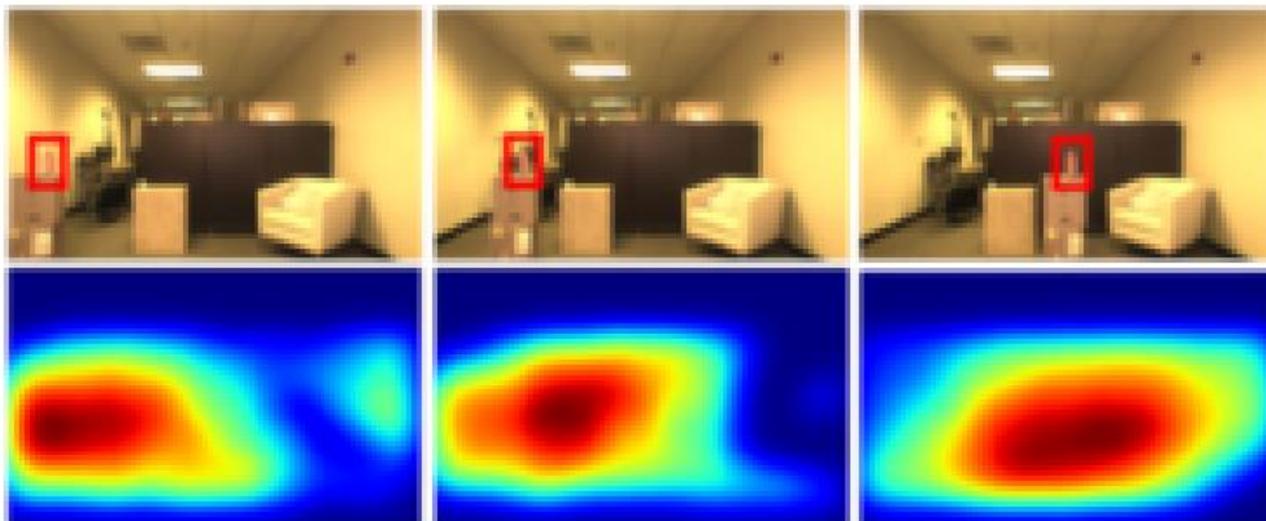


Figura 25 – Na parte superior, fotografias do ambiente destacando o objeto selecionado pelo usuário. Na parte inferior, imagens geradas a partir do nível de intensidade do sinal recebido pela antena RFID. Fonte: [18].

A partir da projeção do equipamento laser scanner, uma máscara binária é aplicada para cada ponto do cenário a fim de gerar a terceira matriz de probabilidades. Nesta matriz, caso o laser obtenha alcance suficiente no ponto (x, y) , a matriz recebe o valor 1, caso contrário recebe o valor zero. Além de fornecer mais uma matriz de probabilidades

para o método de fusão, o laser scanner também possui a função de estimar a distância em profundidade entre o robô e o objeto alvo.

Por fim, as matrizes de probabilidades são combinadas utilizando um modelo baseado em Inferência Bayesiana. Para cada ponto do cenário mapeado na matriz, o resultado do modelo estima a probabilidade do elemento representar a posição do objeto alvo. Assim, o elemento da matriz que possui a probabilidade máxima é determinado como o local onde o objeto se encontra. A partir da localização 2D, o equipamento laser scanner estima a distância 3D até o objeto, resultando na localização final (x, y, z) .

O sistema proposto obteve uma estimativa de localização 3D com precisão de zero cm em 17 das 18 tentativas (94,4%). Nos experimentos que não utilizaram localização RFID, a probabilidade diminuiu para 83,3%, demonstrando uma melhora de 11,1% na agregação desta tecnologia ao IPS.

3.2.3 Sistema de Identificação e Localização Baseado em RFID e Análise Visual

O sistema proposto em [93] tem como objetivo a localização *indoor* de pessoas. Para isto, técnicas de localização visual e RFID são empregadas. Além da localização, através da tecnologia RFID, a identificação das pessoas também é realizada. Neste sistema, etiquetas RFID do tipo ativa são afixadas no chão do ambiente, servindo como coordenadas de referência, enquanto um leitor RFID é carregado por cada pessoa que se deseja realizar a localização e identificação.

O IPS proposto exige uma fase inicial de configuração (*setup*), onde é criada uma tabela contendo um nível qualitativo de sinal, baseado no RSSI de cada etiqueta e na faixa de distância estimada entre a etiqueta e o leitor. A partir desta tabela, uma leitura futura de RSSI pode estimar a distância em que a etiqueta está do leitor.

No cenário de execução da localização RFID, as etiquetas são posicionadas em um intervalo de 2 metros de distância. Uma técnica denominada “*Sensing Overlap Area*” é proposta. Nesta técnica, o RSSI obtido pelo leitor a partir de cada etiqueta de referência é relacionado à tabela da fase de configuração, podendo assim obter a região estimada do leitor.

Na Figura 26, os números com fundo vermelho representam as etiquetas. O círculo roxo ao redor das etiquetas demonstra que, um leitor posicionado naquela região, está dentro do nível mais alto de intensidade de sinal (conforme a tabela da fase de *setup*). O bloco azul demonstra que o nível de sinal daquela região não alcançou o nível máximo

em nenhuma etiqueta, mas chegou ao segundo nível de intensidade para as etiquetas 2 e 5. O bloco amarelo indica que o nível de sinal da etiqueta 5 alcançou o segundo nível e as etiquetas 6 e 8 alcançaram o terceiro nível. Por fim, o bloco verde demonstra a posição de um leitor que obteve um nível de sinal fraco (terceiro nível) a partir das etiquetas 4, 5, 7 e 8.

No sistema de localização visual, é empregado o método de subtração de *background* a fim de estimar a posição de cada pessoa presente no ambiente. Assim, uma fotografia do ambiente é capturada sem nenhuma pessoa estar presente e, durante o processo de localização, uma nova imagem capturada é comparada a primeira fotografia. A diferença existente entre as imagens deve fornecer o bloco de *pixels* em que as pessoas estão posicionadas. Métodos morfológicos como erosão e dilatação também são empregados visando eliminar o ruído e oferecer maior precisão.

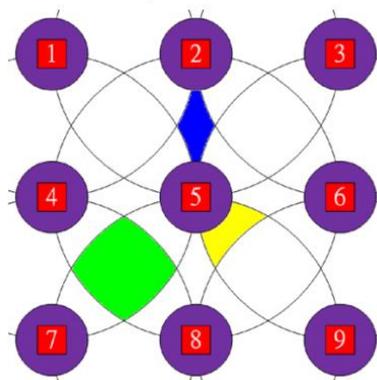


Figura 26 – Leiaute do ambiente na técnica empregada na localização RFID. Fonte: [93].

O método de fusão empregado utiliza a localização estimada pelo sistema RFID em conjunto às coordenadas fornecidas pela localização visual. Para cada leitor localizado pelo sistema RFID, o método de fusão busca as coordenadas mais próximas fornecida pelo sistema visual. Assim, a localização final da pessoa identificada pelo sistema RFID é dada pelas coordenadas fornecidas pela localização visual.

A Figura 27 ilustra um diagrama do cenário e das possíveis respostas para cada parte do sistema de localização. Em (a), os pontos azuis representam os leitores RFID identificados e localizados. Em (b), cada ponto vermelho representa uma pessoa presente no ambiente, obtido pela técnica de subtração de *background*. Finalmente, em (c), o método de fusão realiza a ligação entre as coordenadas das fases anteriores, fornecendo a localização final de cada pessoa.

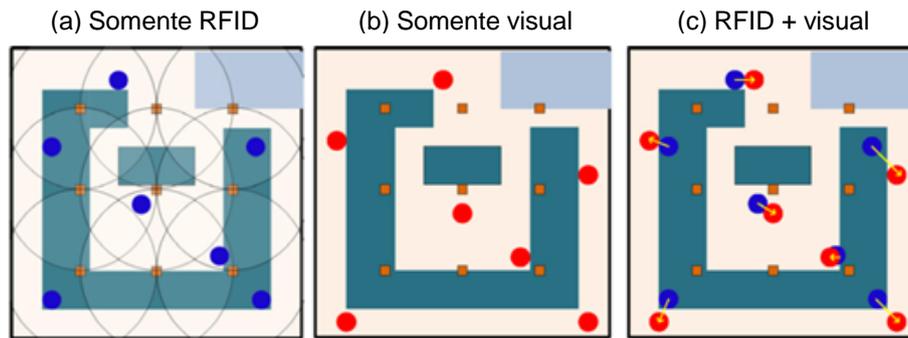


Figura 27 – Diagrama esquemático da localização RFID (a), visual (b) e o método de fusão proposto (c). Fonte: [93].

Os experimentos realizados, quando aplicados a múltiplas pessoas simultaneamente, demonstraram precisão de 1 metro de distância, tendo uma probabilidade de sucesso de 83% quando utilizado somente a localização RFID e 89% quando empregada somente a localização visual. A partir do método de fusão proposto, o sistema de localização foi capaz de identificar e localizar pessoas obtendo uma precisão de 1 metro com 95% de probabilidade de sucesso.

3.2.4 Otimização de IPS RFID através do Padrão de Movimentação de Objetos

Em [7], o método de fusão está baseado na extração da região de interesse (ROI). No IPS proposto, o leitor RFID é acoplado ao objeto alvo e, através de etiquetas de referência, obtém uma região reduzida no qual a localização visual deve atuar. No sistema visual, um algoritmo denominado *block matching* é aplicado, com o intuito de obter uma localização mais precisa do objeto.

Neste trabalho, o sistema de localização RFID necessita de etiquetas de referência (passivas) espalhadas no chão do ambiente, em padrão triangular. Previamente, em cada etiqueta de referência é armazenada a sua respectiva posição (x, y) em que está localizada no ambiente. O ROI obtido é uma área retangular, baseado nas leituras das etiquetas de referência que estavam ao alcance do leitor no momento da sua ativação.

A Figura 28 ilustra um exemplo de extração do ROI. As etiquetas de referência são representadas pelos pontos pretos e estão distribuídas em um intervalo de 10 cm. A área de cobertura do leitor é representada pela circunferência cinza. A partir desta cobertura, é determinada a dimensão da área retangular, cujas coordenadas serão informadas ao algoritmo de localização visual. Nos experimentos realizados, o leitor foi configurado para emitir um sinal fraco, alcançando no máximo três etiquetas para cada lado.

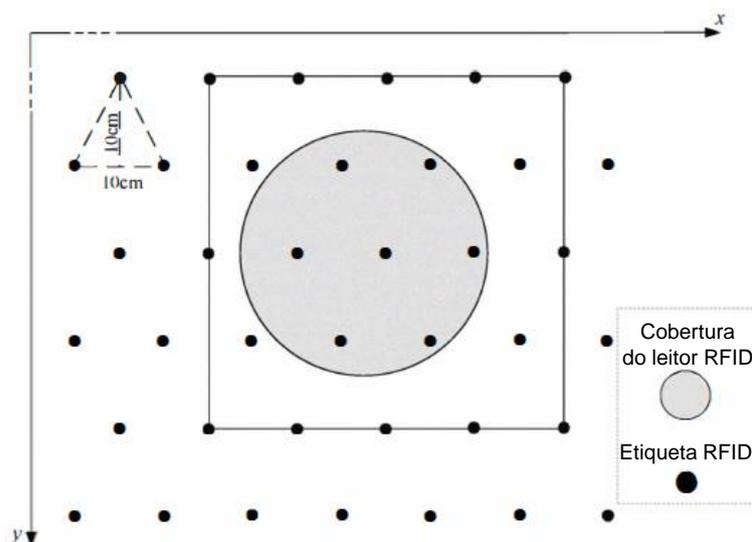


Figura 28 – Padrão triangular das etiquetas de referência, cobertura do leitor RFID e extração da região de interesse. Fonte: [7].

No que concerne à localização visual, uma câmera é afixada ao teto da sala e captura as imagens de acordo com o ROI estimado pela localização RFID. O trabalho proposto utiliza o algoritmo *block matching* para detectar o objeto alvo no cenário. O *block matching* é um algoritmo que realiza a comparação de blocos da imagem entre o *frame* atual (i) e o *frame* anterior ($i-1$). Assim, é analisada a movimentação do objeto e a diferença existente entre cada *frame*. Devido a dificuldades enfrentadas em relação à luminosidade do ambiente, um algoritmo morfológico também é empregado. No processo de morfologia, determinados *pixels* da imagem são alterados para seguir o padrão de movimentação do objeto. Além disso, somente componentes cromáticos são utilizados.

Foram realizados experimentos com o objeto presente em 48 posições distintas. O sistema proposto obteve um erro médio de 0,72 cm com uma variação de 0,16 cm.

3.2.5 Localização Óptica Utilizando WISP

O trabalho [76] difere de todos os demais, pois utiliza um hardware personalizado no processo de localização. Neste sistema, uma etiqueta RFID passiva, acoplada a um WISP (*Wireless Identification and Sensing Platform*), é afixada ao objeto alvo. O WISP ativa um LED a cada leitura RFID efetivada. Por sua vez, o LED, quando aceso, é reconhecido pelo sistema óptico de localização.

WISP é um sensor programável, agregado a uma plataforma computacional, projetado para explorar os recursos sensoriais de aplicações RFID. O WISP utiliza um

microcontrolador de 16-bits de baixa energia para emular o protocolo EPC Gen2¹ e executar tarefas computacionais e sensoriais. O WISP opera exclusivamente com a energia proveniente das ondas de radiofrequência, ou seja, é livre de baterias. Maiores detalhes sobre o projeto de WISPs pode ser encontrado em [77].

O funcionamento do sistema inicia com a consulta, pelo leitor RFID, do EPC da etiqueta afixada ao objeto alvo. Uma vez detectada a etiqueta, técnicas de localização por proximidade (Seção 2.1.1.2) são utilizadas para obter a possível região onde se encontra o objeto. Nesta região de interesse, duas imagens são capturadas por uma câmera industrial de alta resolução (NI 1764). A primeira captura fornece uma imagem com o LED do WISP aceso, enquanto a segunda captura provê a imagem com o LED apagado. Através de técnicas de visão computacional, um mapeamento diferencial entre as duas imagens é gerado e o local onde o LED se encontra é determinado, conseqüentemente estimando a posição do objeto alvo. A Figura 29 ilustra a arquitetura do sistema com o fluxo de funcionamento.

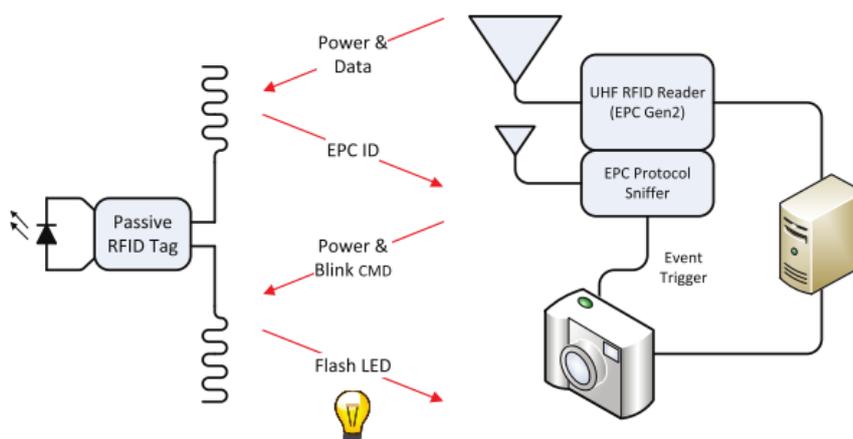


Figura 29 – Arquitetura do sistema de localização baseado em WISPs. Fonte: [76].

Um dos pontos chave deste trabalho é a sincronização entre a ativação do LED e a captura das imagens pela câmera. Neste processo, inicialmente o leitor RFID emite um comando de consulta para a etiqueta alvo. Se a etiqueta possui energia suficiente para ativar o LED, esta responde ao leitor. Na sequência, o leitor emite um comando de escrita. Ao receber o comando de escrita, a etiqueta aguarda um tempo pré-configurado e pisca o LED. De acordo com os componentes da arquitetura, o “EPC Protocol Sniffer” é um WISP adicional que também escuta o canal de comunicação, aguardando por comandos de

¹ EPC Gen2 é a denominação de um dos padrões regulados pela EPCGlobal que define requisitos físicos e lógicos para a operação do sistema RFID. Este padrão é equivalente a norma 18000-6C padronizada pela ISO.

escrita. Quando um comando de escrita é recebido, ele também aguarda o mesmo tempo pré-configurado e então aciona a captura da câmera, obtendo a imagem com o LED aceso. Após um curto período de tempo é realizada a segunda captura, que obtém a imagem com o LED apagado. Assim, o disparo da captura da imagem está sempre sincronizado com a ativação do LED pelo WISP afixado no objeto alvo.

Nos resultados apresentados, o sistema foi capaz de localizar objetos no raio de 1,5 a 2 metros de distância. A precisão do sistema se destaca entre os demais trabalhos, sendo possível localizar objetos com erros entre 1 e 2 cm.

3.2.6 Localização de Etiquetas RFID Auxiliada por Câmeras

O trabalho [67] visa obter a localização 2D de objetos em carrinhos que transportam encomendas na empresa alemã de correios. O sistema emprega recursos ópticos e a tecnologia RFID aliados à técnica de localização de análise de padrões. A partir dos resultados destes modelos, é aplicado um método de fusão de sensores baseado no modelo probabilístico do filtro de Kalman.

No cenário proposto, uma etiqueta passiva RFID é afixada em cada encomenda (caixa de cor amarela) que está dentro do carrinho. O sistema requer uma fase de *setup*, onde os valores de RSSI são medidos em diferentes distâncias, a fim de encontrar uma relação entre a intensidade do sinal e a distância. Utilizando um leitor e quatro antenas RFID acopladas no teto da sala, a relação entre o RSSI e as distâncias é armazenada para ser utilizada no método de fusão.

Relacionado ao emprego de visão computacional no método de localização, um algoritmo baseado em histograma de cores é utilizado. Para isto, câmeras digitais comuns são acopladas ao teto da sala, sendo a captura das imagens disparada quando as encomendas são detectadas pelo sistema RFID. Para o histograma de cores funcionar adequadamente, é necessária uma coleta prévia de amostra de imagens das caixas. A partir das amostras, operações morfológicas e de *thresholding* são aplicadas com o intuito de resultar no reconhecimento do objeto na imagem.

O método de fusão de sensores proposto baseia-se em uma variante do filtro de Kalman, denominado *Constrained Unscented Kalman Filter* (CUKF). O método, baseado em restrições, emprega o resultado da localização visual através da seguinte restrição: o objeto alvo está no cenário neste instante? Já o resultado da localização RFID é aplicado ao CUKF através dos valores RSSI medidos na fase de *setup* e a cada medição que ocorre durante a execução do sistema.

Os resultados deste trabalho foram demonstrados através do RMSE (*Root Mean Square Error*) da diferença entre a posição real do objeto e a estimativa de localização fornecida pelo modelo. Em um cenário estacionário, o RMSE obtido foi igual a 26 cm, enquanto no cenário dinâmico, com o carrinho em movimento, o RMSE foi 36 cm.

3.2.7 Localização Probabilística via Fusão de RFID e Visão Computacional

Em [30], o subsistema RFID proposto é baseado em uma abordagem de proximidade empregando um modelo bayesiano. A ideia é usar diversas antenas $A_1 \dots A_k$ que possam cobrir uma região de interesse específica. O modelo probabilístico incorpora os parâmetros de posição e padrão de irradiação das antenas, que necessitam ser aprendidos em uma fase prévia de calibração. Um padrão de irradiação de três antenas, medido a partir do RSSI, é ilustrado na Figura 30a. No caso, as antenas estavam afixadas no teto da sala e as etiquetas de referência estavam organizadas em forma de grade, a 30 cm de distância uma da outra. A Figura 30b ilustra a aproximação Gaussiana efetivamente utilizada no modelo. Assim, durante a fase de estimativas, o subsistema RFID fornece uma probabilidade para a posição de cada etiqueta baseado nos valores RSSI medidos. A Figura 30c exibe um exemplo do modelo aplicado a três antenas (A_1 , A_2 , A_3) e uma etiqueta (T_1). Como a etiqueta está mais próxima da antena A_1 , o valor do RSSI para esta antena é dominante em relação às demais.

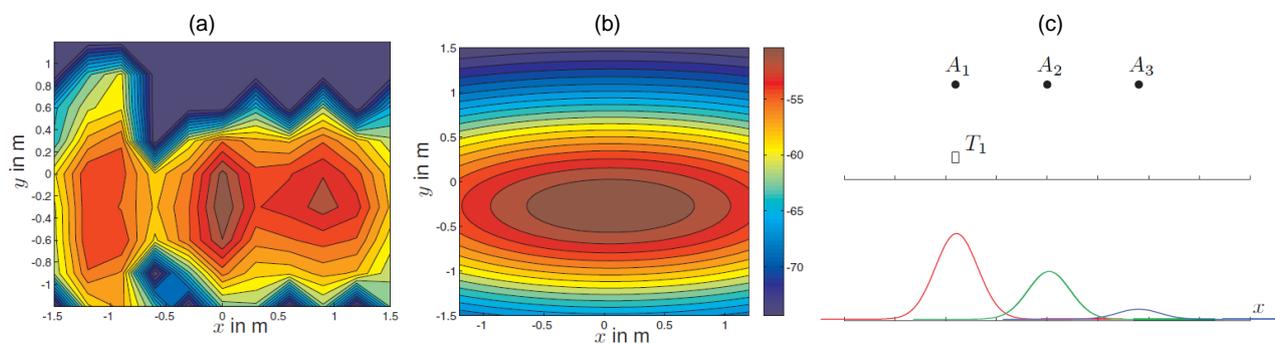


Figura 30 – Padrão de irradiação (a) e respectiva aproximação Gaussiana (b) do nível de intensidade de sinal medido. Fonte: [30].

Em relação ao subsistema de visão computacional, o trabalho implementa uma sequência de mecanismos para detecção e rastreamento de *blobs*¹. Assim, para cada quadro capturado por uma câmera monocular, os seguintes passos são realizados: (i) segmentação do *foreground* e do *background* através de uma técnica já conhecida de subtração de *background*, aplicado a objetos em movimento; (ii) detecção de sombras

¹ Em visão computacional, *blob* é uma região de uma imagem digital cujas propriedades são constantes ou variam dentro de uma faixa pré-definida de valores. Mais tecnicamente, um *blob* é um grupo de *pixels* que o computador identifica como um objeto.

para suprimir os falsos positivos do passo anterior; (iii) detecção de *blobs* no segmento de *foreground*, implementado por um algoritmo já estabelecido denominado rastreamento “Mean Shift”. Cada *blob* detectado é rastreado por vários quadros de imagem consecutivos. Consequentemente, as dimensões, a direção e a velocidade do *blob* também podem ser obtidas.

O sistema de fusão proposto visa estabelecer uma associação entre cada etiqueta e *blob* detectado. Além disso, o sistema precisa considerar a possibilidade de existirem etiquetas estacionárias que não se relacionam a nenhum *blob* presente no cenário. Uma matriz de probabilidades é criada relacionando cada etiqueta e o *blob* (ou o *background*, caso seja uma etiqueta estacionária). O valor da probabilidade é definido através da distância espacial entre a etiqueta e o *blob*, ou seja, quanto mais próximos, maior a probabilidade. No caso do *background*, é detectado que uma etiqueta não possui movimento através da análise de suas posições em relação ao tempo, sendo assim uma probabilidade maior é atribuída a este caso. Finalmente, uma etiqueta T_i é associada ao seu respectivo *blob* B_j encontrando o valor máximo em cada linha da matriz. Caso uma etiqueta seja associada ao *background* (estacionária), esta é ignorada, pois somente etiquetas associadas a *blobs* são consideradas na localização.

Os experimentos deste trabalho empregam métricas diferentes daquelas encontradas em outros trabalhos. A primeira métrica (P_d) diz respeito ao percentual de etiquetas detectadas corretamente em relação ao número de etiquetas que se movem no cenário. A segunda métrica (P_{fa}) indica a taxa de falsos positivos observados em relação ao número de etiquetas estacionárias na região de interesse. Em um ambiente com duas etiquetas estacionárias e uma etiqueta em movimento os resultados foram 96% (P_d) e 0% (P_{fa}). Em outro teste, quatro etiquetas estacionárias estavam presentes e duas pessoas caminhavam no cenário, sendo que uma delas carregava duas etiquetas. Os valores de P_d e P_{fa} para este teste foram 84,38% e 0,02%, respectivamente.

3.2.8 Discussão

Os trabalhos explorados abordam diferentes técnicas de localização, cada um com um uso distinto de sistemas ópticos e sistemas RFID. Fica visível que o desenvolvimento de IPSs está em constante evolução. À medida que os anos passam, métodos novos são desenvolvidos com o objetivo de melhorar ainda mais o desempenho deste tipo de sistema. Alguns trabalhos também visam otimizar o custo associado à implantação e manutenção do sistema, fazendo uso de componentes mais simples e se integrando à própria infraestrutura já existente no ambiente.

É importante destacar o uso de técnicas de localização distintas na maioria dos trabalhos (Tabela 3). Isto demonstra que não existe um método definitivo e que a pesquisa e o desenvolvimento na área são relevantes. As técnicas de localização empregadas pela tecnologia RFID baseiam-se em conceitos conhecidos, explorados na Seção 2.1.1. Na análise visual são empregados algoritmos de visão computacional com características funcionais distintas. Por fim, o método de fusão explorado por cada um dos trabalhos vislumbra o melhor modo de combinar as tecnologias.

Tabela 3 – Visão geral das técnicas de localização utilizadas nos trabalhos correlatos.

Trabalho	RFID	Visual	Método de fusão
Chae e Na [11]	Proximidade	SIFT; Angulação	ROI
Deyle <i>et al</i> [18]	Análise de padrões	Histograma e <i>thresholding</i>	Inferência Bayesiana
Wang e Cheng [93]	Análise de padrões (RSSI)	Subtração de <i>background</i>	Próprio
Babic <i>et al</i> [7]	Proximidade	<i>Block matching</i>	ROI
Sample <i>et al</i> [76]	Proximidade	Mapeamento diferencial	ROI
Nick <i>et al</i> [67]	Análise de padrões (RSSI)	Histograma e <i>thresholding</i>	Filtro de Kalman (variante: UCKF)
Göller <i>et al</i> [30]	Análise de padrões (RSSI)	Subtração de <i>background</i> e detecção de <i>blobs</i>	Próprio
(este trabalho)	Análise de padrões (RSSI)	Detecção de bordas	Próprio (MROI e k-means)

Além das técnicas de localização empregadas, outras características presentes nos trabalhos analisados estão diretamente relacionadas ao custo e desempenho associado a um sistema IPS. Dentre estas características, pode-se destacar: (i) tipo de etiqueta RFID (ativa ou passiva); (ii) necessidade de desenvolvimento de um novo hardware; (iii) escalabilidade 2D ou 3D; (iv) necessidade de fase de *setup*; (v) objetos alvo em movimento ou estáticos; (vi) precisão e (vii) probabilidade de sucesso. A Tabela 4 lista esta série de características para cada trabalho correlato, fornecendo uma visão geral das pesquisas mais recentes em relação ao tema abordado no trabalho de doutoramento.

Tabela 4 – Características dos sistemas propostos nos trabalhos correlatos analisados.

Trabalho	Etiqueta	Novo hardware	2D/3D	Setup	Dinâmico vs. estático	Precisão	Sucesso
Chae e Na [11]	Ativa	Não	2D	Sim	Estático ¹	23 cm	91,5%
Deyle <i>et al</i> [18]	Passiva	Sim ²	3D	Não ³	Estático	0 cm	94,4%
Wang e Cheng [93]	Ativa	Não	2D	Sim	Estático	1 m	95%
Babic <i>et al</i> [7]	Passiva	Não	2D	Sim	Dinâmico	0,72 cm	-
Sample <i>et al</i> [76]	Passiva (WISP)	Sim	3D	Não	Estático	2 cm	-
Nick <i>et al</i> [67]	Passiva	Não	2D	Sim	Ambos	26 cm (estático)	-
						36 cm (dinâmico)	-
Göller <i>et al</i> [30]	Passiva	Não	2D	Sim	Dinâmico	0 cm	96%
(este trabalho)	Passiva	Não	3D	Sim	Estático	9 cm (2D) 63 cm (3D)	-

¹ Apesar do robô se movimentar, ele pode estar parado no momento do processo de localização.

² É necessário um dispositivo mecânico para mover automaticamente as antenas RFID.

³ Uma imagem do objeto alvo deve estar previamente armazenada.

Restrições e deficiências puderam ser encontradas durante a análise dos trabalhos correlatos. É certo que, apesar de alguns trabalhos obterem uma boa precisão, o seu uso pode ser restrito ou possuir um custo elevado de implementação ou manutenção. As restrições e dificuldades detectadas foram:

- Chae e Na [11]: além do uso de etiquetas RFID ativas, fato que pode aumentar a manutenção e o custo do sistema, não é claro o tamanho que cada região do cenário deve possuir. É reconhecido que, quanto maior a região, maior deve ser o número de etiquetas utilizadas, aumentando assim o tempo dispendido na fase de *setup* do sistema, bem como o custo empregado na aquisição e manutenção de etiquetas ativas.
- Wang e Cheng [93]: a técnica de subtração do *background*, aplicada na localização visual, obteve bom desempenho na detecção de pessoas. Entretanto, não é claro se esta técnica também poderia ser empregada em objetos de tamanho reduzido.

- Babic *et al* [7]: a infraestrutura necessária para distribuir as etiquetas de referência exige uma alta densidade, ou seja, para uma melhor precisão e probabilidade de sucesso, mais etiquetas, em um intervalo menor de distância, devem ser instaladas no ambiente. Como restrição, o algoritmo de localização visual necessita que o objeto alvo esteja em movimento, pois uma análise estática não fornece diferenças entre as cenas analisadas pela técnica *block matching*.
- Nick *et al* [67]: o IPS demonstrado é bastante restrito aos objetos do cenário de uso empregado (caixas amarelas do correio alemão), não sendo validado em outro tipo de ambiente. No trabalho, é mencionado que a análise visual empregada pode não obter um bom desempenho em objetos de múltiplas cores. O trabalho não deixa claro se há uma associação entre a imagem da caixa (objeto) e a etiqueta RFID afixada, assim não é possível saber se o sistema localizaria múltiplos objetos simultaneamente. Também não é claro como o sistema funciona em cenários estáticos, pois o algoritmo Filtro de Kalman empregado tipicamente necessita de informações que se alteram em relação ao tempo.

Um padrão detectado nos trabalhos analisados, diz respeito ao escopo de uso de cada tecnologia. Foi verificado que a maioria dos trabalhos analisados utiliza a tecnologia RFID como sistema intermediário no processo de localização, sendo que o resultado final da localização do objeto alvo é fornecido pela localização visual. O motivo desta abordagem se deve principalmente à baixa precisão dos IPSs baseados em RFID, que não têm obtido uma boa resposta em distâncias no nível de poucos centímetros. Tipicamente, a causa deste baixo desempenho está relacionada a fenômenos interferentes e características intrínsecas à tecnologia (conforme explorado na Seção 2.1.3).

Um critério fundamental na adoção de IPSs e pouco considerado nos trabalhos correlatos diz respeito à escalabilidade geográfica, ou seja, a área mínima e máxima de cobertura do sistema. Nos trabalhos são relatados experimentos em determinadas distâncias, entretanto não definem se estas são as distâncias máximas de alcance do sistema proposto. Também não são realizados testes em distâncias superiores com o objetivo de definir os limites de atuação do sistema. Ainda, dentre os trabalhos estudados, somente dois deles fornecem uma localização *indoor* com escalabilidade 3D.

4 DESCRIÇÃO DO SISTEMA

O sistema de localização *indoor* proposto está baseado em um modelo híbrido, aplicando técnicas de análise visual em conjunto com a tecnologia RFID. A abordagem empregada aplica-se a um cenário onde etiquetas RFID e marcadores visuais estão afixados aos objetos que se deseja localizar. A etiqueta RFID e seu marcador visual correspondente identificam unicamente cada objeto do cenário. Neste trabalho, cada par de etiqueta RFID e marcador visual é denominado “marcador”. Assim, o IPS deve ser capaz de obter a posição de cada marcador presente no cenário. Em alinhamento às motivações e aos objetivos do trabalho, foram definidos os seguintes requisitos:

- i. Localizar objetos estáticos;
- ii. A informação de localização deve ser do tipo física e relativa;
- iii. Resultar a localização 2D (x, y) ou 3D (x, y, z) do objeto alvo;
- iv. Obter uma precisão de poucos centímetros e alta probabilidade de sucesso;
- v. Obter informações referentes aos objetos localizados (o código, por exemplo), sendo estes dados fornecidos pelos elementos afixados ao próprio objeto;
- vi. Possibilitar a execução com tempo de resposta de poucos segundos (até 3 segundos);
- vii. Localizar vários objetos simultaneamente.

As próximas seções detalham a pesquisa e o desenvolvimento da arquitetura e de cada parte do sistema.

4.1 Arquitetura do Sistema

Conforme discutido no Capítulo 2, os principais elementos de um IPS são os dispositivos sensoriais, um algoritmo ou técnica de localização e uma interface gráfica [70]. Neste trabalho, as seguintes tecnologias foram empregadas e associadas a cada camada: (i) estímulos de localização: RFID e câmera digital; (ii) técnica: modelos de aprendizado de máquina e algoritmos de visão computacional; (iii) interface gráfica: aplicação .NET. A Figura 31 exhibe os elementos associados a uma arquitetura em três camadas.

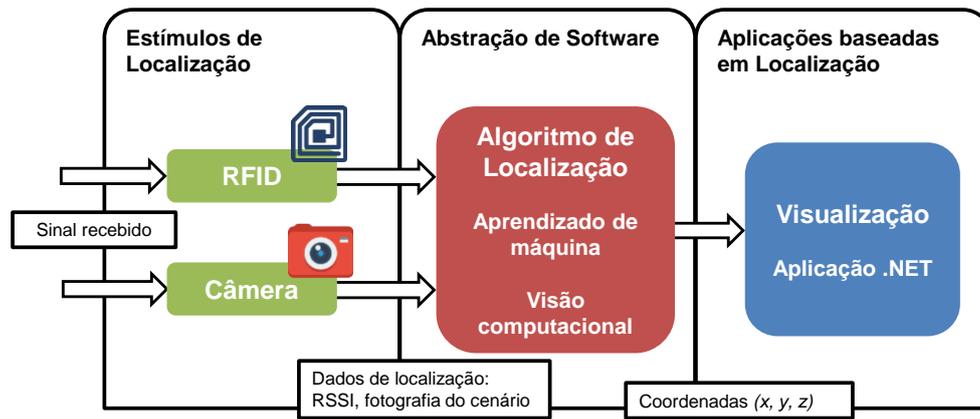


Figura 31 – Arquitetura em três camadas associada a um diagrama de blocos contendo os componentes fundamentais do IPS proposto. Fonte: o autor.

O sistema proposto está dividido em dois subsistemas. O primeiro emprega modelos de aprendizado de máquina para a estimativa da localização utilizando a tecnologia RFID, enquanto o segundo utiliza uma câmera e algoritmos de visão computacional para refinar as previsões do primeiro subsistema. A técnica de localização empregada neste trabalho baseia-se na análise de padrões (Seção 2.1.1.3), combinando as técnicas *fingerprinting* e análise visual. A técnica *fingerprinting* está dividida em duas fases: treinamento *offline* e estimativa *online*. A fase *offline* é executada uma única vez para o cenário escolhido. A fase *online* é executada quantas vezes for necessária, para cada objeto que se deseja localizar. A Figura 32 exhibe os dois subsistemas, as fases, os processos e os fluxos entre eles.

Na fase *offline*, os equipamentos RFID e a câmera são calibrados, as etiquetas de referência e os marcadores visuais são posicionados, e o modelo de aprendizado de máquina supervisionado é treinado. Dois modelos de aprendizado de máquina são propostos e comparados. Um modelo é baseado em Redes Neurais Artificiais e outro em *Support Vector Regression*. O subsistema RFID tem o papel de coletar os *fingerprints* RSSI (Seção 2.1.3) do ambiente para treinar os modelos. Nesta fase, o subsistema visual somente auxilia na captura de imagens do cenário a fim de configurar as posições dos marcadores de referência.

Na fase *online*, o subsistema RFID coleta *fingerprints* RSSI para serem usados como entrada dos modelos criados na fase *offline*. Uma abordagem k-means é proposta a fim de agrupar os resultados preliminares em *clusters*. Finalmente, o subsistema visual refina os resultados através da detecção visual de marcadores presentes no cenário. Ambos subsistemas e fases são apresentados em detalhes nas próximas seções.

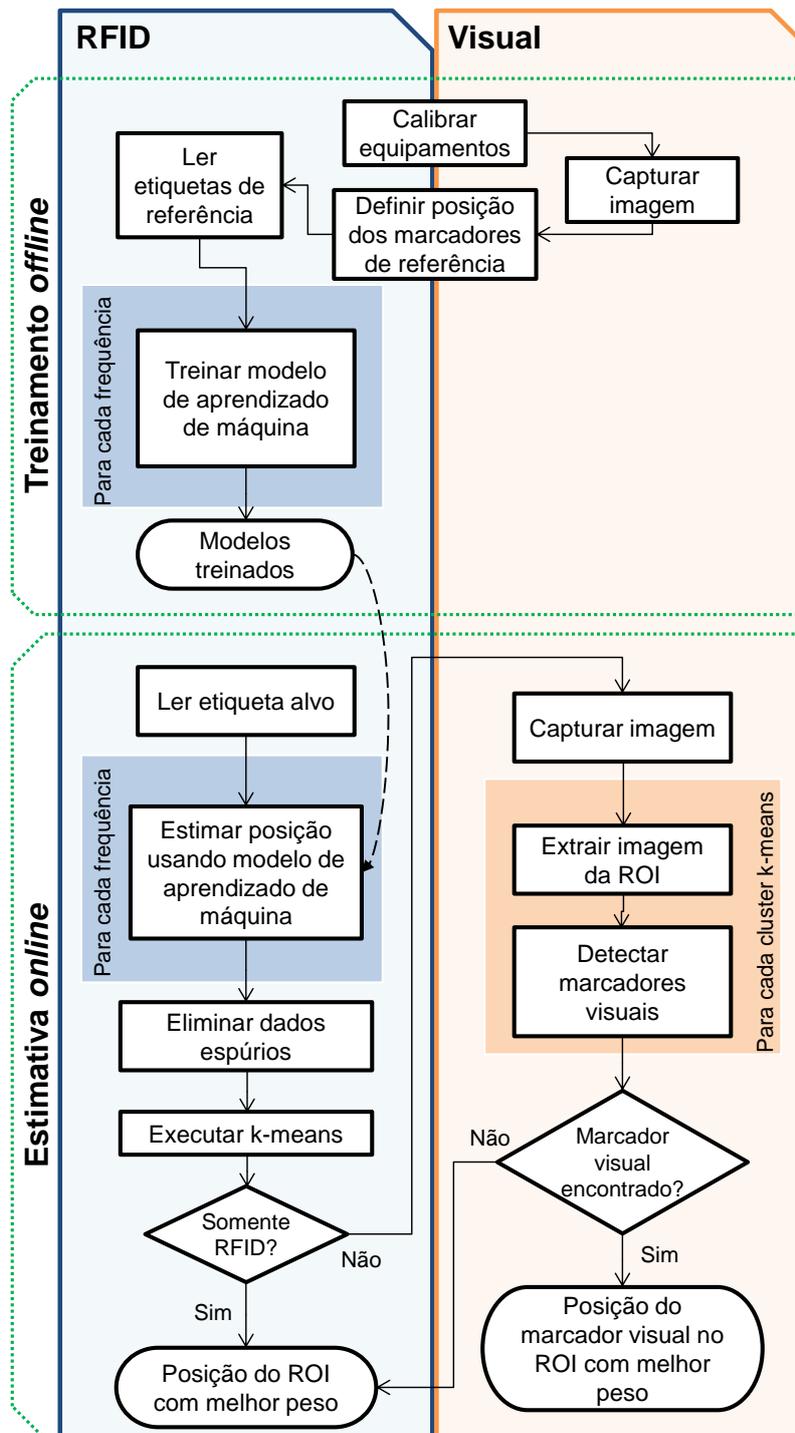


Figura 32 – Diagrama de blocos da arquitetura do sistema. As entradas são as leituras RFID e a imagem capturada. A saída é a posição do objeto alvo. Fonte: o autor.

O leitor RFID e a câmera são conectados a um computador executando o sistema. Uma aplicação .NET GUI (*Graphical User Interface*) foi desenvolvida para automatizar os experimentos e exibir graficamente os resultados para os usuários. Bibliotecas Matlab [14], [87], [88] foram integradas ao .NET C# para treinar e executar os modelos de aprendizado de máquina.

4.2 Fase Offline

Inicialmente, os equipamentos RFID e a câmera são ajustados para parâmetros e posições que devem permanecer fixos durante toda execução do sistema (*offline* e *online*). Algumas calibrações realizadas nesta fase são: potência e região de operação do leitor RFID, posição das antenas RFID, resolução e posição da câmera.

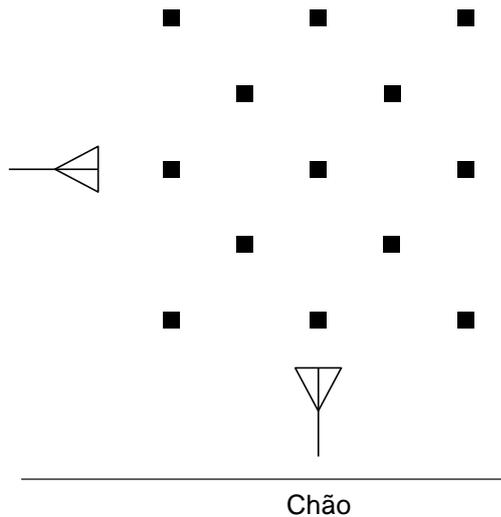
O subsistema RFID proposto baseia-se no valor do RSSI de cada etiqueta para inferir a sua localização no cenário. Este valor pode ser lido por uma ou mais antenas presentes no ambiente e que estejam conectadas a um leitor RFID compatível com o padrão LLRP, fornecendo assim o RSSI de cada etiqueta detectada. A posição das antenas do leitor é peça fundamental para o bom desempenho do sistema [43]. Foram realizados ensaios com diferentes posicionamentos a fim de identificar uma melhor configuração, sendo descritos na Seção 5.2.

Por se tratarem de modelos estatísticos, uma coleta de dados utilizando etiquetas de referência é necessária. Este passo requer que as etiquetas de referência sejam distribuídas no ambiente de maneira uniforme. Durante a execução dos experimentos, algumas formas de organização de etiquetas de referência foram avaliadas (Seção 5.3). A Figura 33a exibe um exemplo de organização dos componentes RFID. No caso, as etiquetas estão em formato de grade diagonal e as antenas posicionadas ao lado e abaixo do cenário alvo.

Nesta proposta, o ângulo de orientação das etiquetas também deve ser o mesmo entre as fases de treinamento e estimativa. Esta restrição se deve ao fato do nível de intensidade refletido pela etiqueta variar de acordo com a orientação da antena presente na etiqueta. Por exemplo, o RSSI de uma etiqueta na posição horizontal, que é associado a uma posição de referência, possivelmente não irá corresponder ao RSSI de uma etiqueta na orientação vertical, mesmo se posicionada no mesmo local [92], [99].

Após estes ajustes, a posição das etiquetas de referência deve ser armazenada na configuração do sistema. A posição (x, y) informada ao sistema está baseada nas coordenadas de uma grade virtual (Figura 33b) criada sobre uma imagem. Uma fotografia do cenário é capturada e uma grade é definida sobre ela. Desta forma, o sistema de localização pode resultar a posição de qualquer célula da grade dentro dos limites da imagem capturada. No caso da localização em escalabilidade 3D, a distância em profundidade (z) até as etiquetas deve ser medida manualmente (preferencialmente em centímetros) e registrada na configuração do sistema.

(a) Exemplo de organização dos componentes



(b) Grade virtual

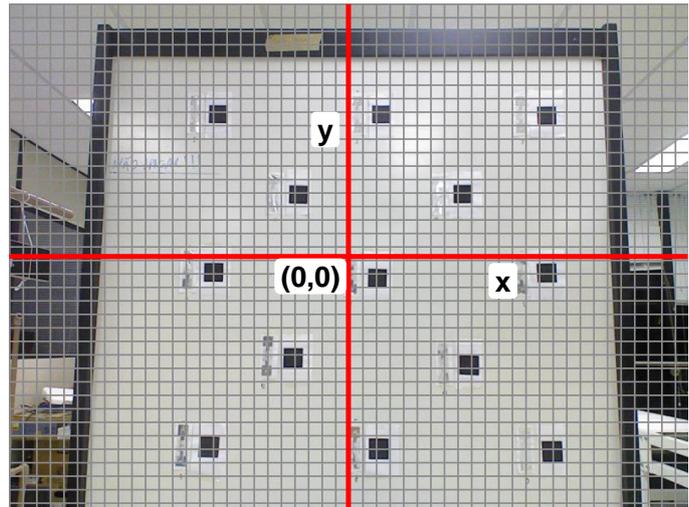


Figura 33 – Exemplo de posicionamento das etiquetas de referência e antenas (a) e grade virtual sobre uma fotografia capturada do cenário (b). Fonte: o autor.

Depois de realizadas as calibrações iniciais, as etiquetas de referência são lidas e os dados são coletados. O leitor RFID é ativado por um determinado período fixo de tempo e o sistema armazena o código da antena que efetuou a leitura, a frequência em MHz, o RSSI medido e as coordenadas das etiquetas de referência presentes no cenário. O Apêndice A fornece amostras de leituras coletadas pelo subsistema RFID.

Posteriormente, os dados coletados devem ser preparados para alimentar o treinamento dos modelos de aprendizado de máquina. As próximas seções descrevem as rotinas de ajuste e as modelagens propostas.

4.2.1 Multi-Frequência

Os valores de RSSI têm sido amplamente empregados em sistemas de localização. Entretanto, mesmo em ambientes estacionários, este valor pode sofrer variações de acordo com a frequência de operação do sistema RFID. Conforme o modelo de propagação apresentado na Seção 2.1.3.1, o comprimento de onda da frequência é uma das variáveis presentes na modelagem do canal de comunicação. Assim, a intensidade do sinal recebido pelo leitor pode ser diferente quando são utilizadas frequências distintas. Além disso, determinadas frequências podem sofrer interferências por intermodulação, quando existentes outros equipamentos de radiofrequência em funcionamento na mesma faixa de operação. Este fato pode resultar em uma atenuação dos níveis de intensidade recebidos [95].

Em conformidade a regulamentações federais como ANATEL (Brasil) e FCC (EUA), equipamentos RFID UHF não podem permanecer na mesma frequência por mais de 0,4 segundos dentro de um intervalo de 10 segundos [1], [24]. Para atender esta restrição, os leitores RFID varrem o espectro a cada 250 KHz, buscando canais disponíveis. Esta funcionalidade limita a possibilidade de execução do sistema RFID em uma frequência fixa.

Esta característica pode ser considerada uma restrição na utilização de equipamentos *off-the-shelf* em sistemas de localização, principalmente naqueles baseados em modelos probabilísticos. Por exemplo, os valores de RSSI aprendidos na frequência 915,25 MHz podem ser muito diferentes dos valores RSSI medidos durante estimativas na frequência 923,25 MHz.

Com o objetivo de superar estas limitações, neste trabalho os dados coletados em ambas as fases são particionados pela frequência de operação. Assim, na fase *offline* são criadas instâncias de modelos para cada frequência existente nas leituras coletadas (Figura 34). Consequentemente, na fase *online* cada instância de modelo utiliza as leituras da frequência respectiva para estimar a localização do objeto alvo.

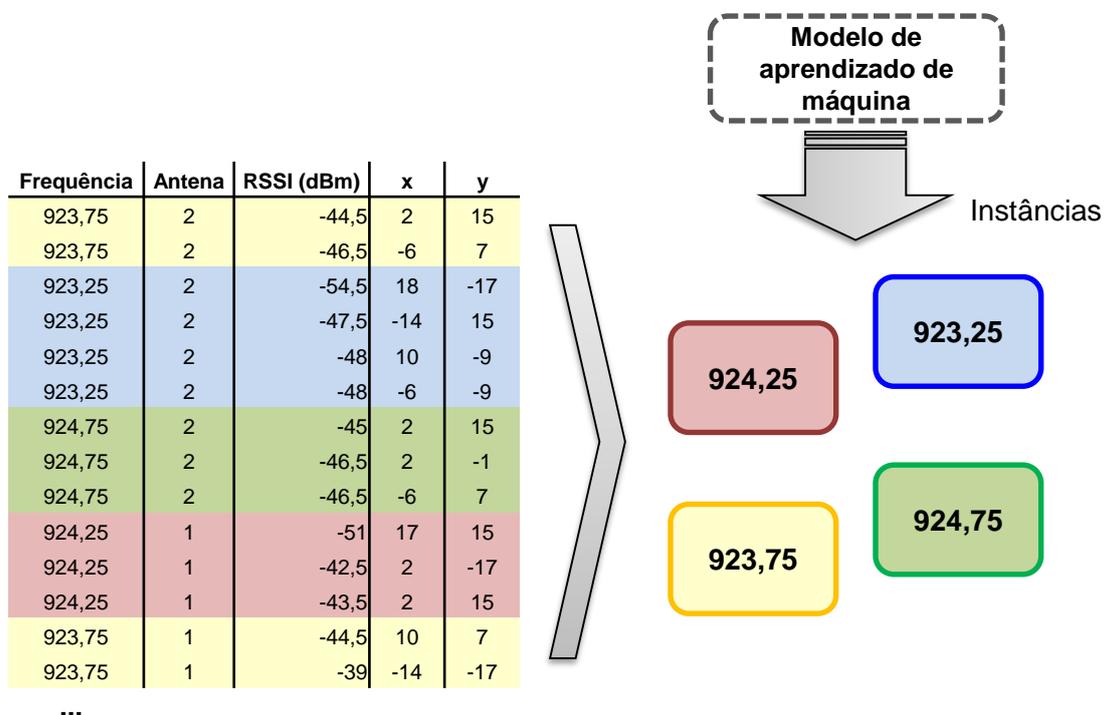


Figura 34 – Método de multi-frequência: criação de instâncias do modelo de aprendizado para cada faixa de frequência detectada na coleta de dados. Fonte: o autor.

Esta proposta visa separar os valores de RSSI de frequências distintas, evitando os problemas mencionados. Assim, é possível empregar com maior probabilidade de

sucesso modelos de localização baseados em análise de padrões usando equipamentos *off-the-shelf* e que atendem a regulamentações de radiofrequência. A Seção 5.5 apresenta experimentos e resultados que avaliam o método proposto.

4.2.2 Preparação dos Dados Coletados

A maioria dos leitores RFID coletam as leituras individualizadas por antena. Com o objetivo de apresentar as amostras para os modelos estatísticos, estes registros devem ser agrupados de tal forma que seja possível estimar a posição a partir da combinação dos valores de RSSI lidos por cada antena.

Assim, optou-se por aglutinar os registros de antenas distintas mais próximos na sequência de leituras e descartar os registros que ficaram sem par. Como resultado, para cada etiqueta posicionada no cenário, o número de amostras é igual ao número de registros da antena com menos leituras. A Figura 35 exibe um exemplo de amostras coletadas pelo leitor RFID e o agrupamento utilizado. O exemplo ilustra uma coleta realizada para o sistema na escalabilidade 2D (x, y), embora a sistemática do processo também se aplique à dimensão 3D (x, y, z). Devido à aplicação do método de multi-frequência (Seção 4.2.1), os registros de origem são sempre particionados por frequência (923,25 MHz, no exemplo).

Antena	RSSI (dBm)	Posição (x, y)
2	-61,5	(-7, 11)
1	-44,5	(-7, 11)
1	-44,5	(-7, 11)
2	-61	(-7, 11)
2	-60,5	(-7, 11)
1	-44,5	(-7, 11)
1	-44,5	(-7, 11)
1	-44,5	(-7, 11)
1	-44	(-7, 11)
...		
1	-49,5	(8, 3)
1	-50	(8, 3)
1	-50	(8, 3)
2	-55,5	(8, 3)
1	-49,5	(8, 3)
2	-56	(8, 3)
2	-55,5	(8, 3)
1	-49,5	(8, 3)
1	-50	(8, 3)
1	-50	(8, 3)
...		

Entrada: RSSI (dBm)		Saída: posição	
Antena 1	Antena 2	x	y
-44,5	-61,5	-7	11
-44,5	-61	-7	11
-44,5	-60,5	-7	11
...			
-49,5	-55,5	8	3
-50	-56	8	3
-50	-55,5	8	3
...			

■ Descartados

Figura 35 – Amostras coletadas individualmente pelo leitor RFID (esquerda) e agrupamento de dados (direita). Frequência: 923,25 MHz. Fonte: o autor.

Tendo a base de dados pronta, uma análise estatística deve verificar a correlação entre as variáveis independentes (explanatórias, entradas) e dependentes (resposta, saída). Na abordagem proposta, as variáveis independentes são os valores RSSI lidos por cada antena e as variáveis dependentes são as coordenadas (x, y) ou (x, y, z) das etiquetas de referência. Utilizando técnicas de aprendizado de máquina, modelos de regressão podem ser desenvolvidos com o intuito de estimar a posição de uma etiqueta alvo a partir dos valores de RSSI.

4.2.3 Modelagem RLM

Tendo uma coleta de dados da fase *offline* consolidada, foi utilizado o software IBM SPSS Statistics [41] para a execução da regressão linear múltipla. Na modelagem do problema, os valores de RSSI das antenas do sistema RFID foram configurados como variáveis independentes (explanatórias). Como a técnica de regressão somente permite uma variável dependente por modelo, e o problema de localização requer mais de um valor estimado, duas regressões foram executadas, sendo uma para a variável resposta da coordenada x e outra para a coordenada y .

A regressão linear múltipla aplicada sobre as variáveis que representam as coordenadas x e y no sistema de localização obtiveram um R^2 de 0,63 e 0,18, respectivamente. Este teste de validação ocorreu sob as mesmas configurações de ambiente apresentadas nos experimentos da Seção 5.6. Tais resultados demonstram uma não-linearidade nos dados, não sendo o método de regressão linear passível de obter uma boa predição utilizando tais amostras. A partir do baixo desempenho nos valores de R^2 , outros métodos de inferência que possam tratar dados não-lineares foram pesquisados, sendo apresentados nas seções seguintes.

4.2.4 Modelagem RNA

Redes neurais são modelos eficientes para problemas de localização, uma vez que elas possibilitam inferir posições através da interpolação das coordenadas conhecidas. Uma das suas principais características é que nenhum conhecimento prévio sobre a geometria do ambiente (posição das paredes, móveis e outros obstáculos) é necessário. Além disso, os fatores interferentes, como a propagação multi-caminho dos sinais de radiofrequência e a existência de outros equipamentos eletrônicos, já estão presentes nas amostras de treinamento coletadas no local [65], [75].

Os dados coletados e preparados na fase *offline* alimentam o processo de treinamento da RNA. Conforme descrito na Seção 4.2.1, os dados são separados pelos

valores de frequência, sendo que nesta modelagem é criada uma rede neural para cada frequência. A configuração da rede neural deste trabalho está baseado no modelo apresentado em [12]. O RSSI de cada antena e as coordenadas de cada etiqueta de referência na grade virtual são apresentados como dados de entrada e saída da rede, respectivamente.

Os dados de treinamento são divididos em três subconjuntos: treinamento, validação e teste. Os dados são separados aleatoriamente na taxa de 70%, 15% e 15%, respectivamente. O conjunto de treinamento é usado para o cálculo e processamento do gradiente e atualização dos pesos e bias da rede neural. O conjunto de validação assegura que não aconteça sobreajuste (*overfitting*) no resultado final. O conjunto de teste é útil para indicar uma divisão ruim no conjunto de dados.

De maneira geral, em redes neurais, o paradigma mais comum de aprendizado é o supervisionado, associado a uma rede direta multicamadas. Uma rede *feedforward* com retro-propagação usando quatro camadas foi modelada. A rede consiste de n neurônios na camada de entrada, 24 neurônios e 12 neurônios em cada camada oculta, e k neurônios na camada de saída. O número de neurônios na camada de entrada deve ser igual ao número de antenas utilizadas. O número de neurônios k na camada de saída depende da dimensão espacial do problema ($k=2$ para 2D, $k=3$ para 3D). A Figura 36 apresenta a configuração da rede na escalabilidade 2D. O número de camadas ocultas foi estipulado desta forma pois acima e abaixo destes valores o desempenho não melhorou e o tempo de treinamento aumentou consideravelmente.

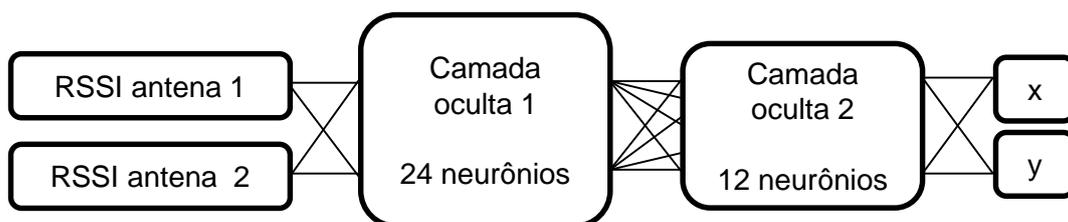


Figura 36 – Estrutura da rede neural artificial (escalabilidade 2D). Fonte: o autor.

Em uma validação inicial, com o auxílio da biblioteca de redes neurais do Matlab [88], uma rede neural foi treinada utilizando os algoritmos de treinamento Levenberg-Marquardt, Scaled Conjugate Gradient e BFGS Quasi-Newton [85]. Como resultado, o algoritmo Levenberg-Marquardt obteve o melhor desempenho, tanto no baixo erro de treinamento (MSE), quanto no menor tempo de processamento. Assim, este algoritmo de retro-propagação foi estabelecido para o trabalho. Os critérios de parada do algoritmo

(Seção 2.2.2.2) foram definidos em “ 1×10^{-5} ”, “1000” e “6” para os parâmetros *min_grad*, *epochs* e *max_fail*, respectivamente.

4.2.5 Modelagem SVR

Modelos SVR têm sido pouco empregados em problemas de localização, todavia esta técnica possui habilidade para tratar problemas não-lineares, além de ter excelente capacidade de generalização, especialmente para problemas com um conjunto reduzido de amostras [3].

Neste trabalho, a modelagem do problema de localização no SVR é definida de forma similar às redes neurais. O conjunto de treinamento do modelo conta com os valores de RSSI recebidos por cada antena presente no ambiente. A saída desejada é uma das coordenadas espaciais da posição da etiqueta RFID. Como a técnica SVR permite somente uma saída em cada formulação, o processo é realizado uma vez para cada coordenada. Por exemplo, em um problema de localização tridimensional são realizadas três formulações, uma para a coordenada *x*, outra para a *y* e a última para a distância em profundidade (*z*). Independente de qual coordenada se está treinando, os valores de RSSI de todas as antenas são considerados no vetor do espaço de entrada do modelo.

Esta modelagem também emprega o método de multi-frequência (Seção 4.2.1). Assim, as amostras coletadas estão separadas pelos valores de frequência e cada conjunto de dados alimenta o treinamento da sua respectiva frequência.

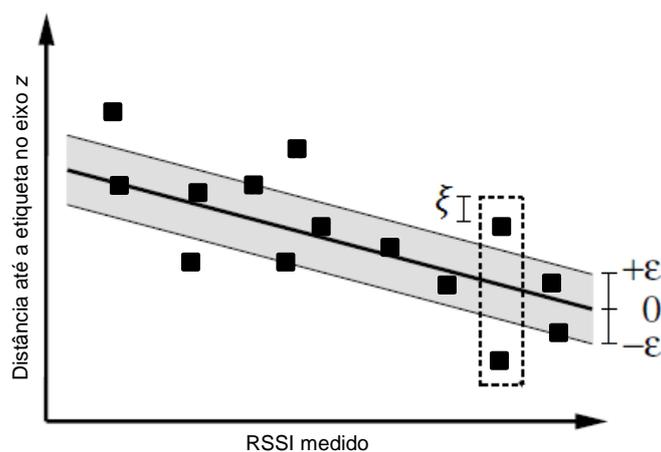


Figura 37 – Simplificação da modelagem SVR aplicada ao problema de localização.

Fonte: o autor.

A Figura 37 exibe uma simplificação da modelagem proposta, onde um determinado RSSI é medido e relacionado à distância em profundidade da etiqueta alvo.

No exemplo, o SVR procura encontrar uma função $f(x)$ que atinja no máximo um desvio ε em relação à posição z da etiqueta alvo. Na modelagem completa, podem existir diversos espaços de entrada x , um para cada antena presente no ambiente.

A modelagem SVR descrita foi implementada e executada com o auxílio de uma biblioteca SVR [14] disponibilizada na ferramenta Matlab [89]. Foi realizada uma validação cruzada dos coeficientes do modelo e, baseado nos resultados, foi configurada a constante de regularização $C=10$ e a margem $\varepsilon=0,00025$.

No que diz respeito à característica do SVM de projetar os dados em um espaço de dimensão maior, o mapeamento realizado neste procedimento foi definido pela função *kernel* denominada “*wavelet*”, publicada originalmente em [98]. Até o momento da implementação deste trabalho, a biblioteca SVR de [14] disponibilizava somente esta opção de *kernel*.

4.3 Fase *Online*

Esta fase executa os subsistemas RFID e visual com o objetivo de fornecer a localização final do objeto alvo. O subsistema RFID estima as possíveis posições do objeto alvo utilizando os modelos treinados, eliminando respostas espúrias e aplicando o método k-means. O objetivo do subsistema RFID nesta fase é a detecção de regiões de interesse que posteriormente serão utilizadas pelo subsistema de análise visual para indicar uma localização mais precisa do objeto alvo. A região de interesse (*Region of Interest* - ROI) pode ser definida como uma área de tamanho limitado e inferior às dimensões do cenário completo. Visto que outras técnicas de localização podem ser aplicadas a esta área reduzida, a tendência é que exista melhor desempenho no processo global de localização.

Durante a fase *online*, as etiquetas de referência não estão presentes no ambiente e uma etiqueta RFID desconhecida é lida durante um período fixo de tempo. Os dados coletados pelo leitor RFID são o código da antena, a frequência e o RSSI. As configurações do leitor RFID (por exemplo, região de operação e potência) devem ter os mesmos valores utilizados na fase *offline*. O modelo deve ser capaz de obter a estimativa de localização utilizando etiquetas e posições distintas às aquelas utilizadas durante o treinamento. Por exemplo, a etiqueta utilizada na fase *online* pode ser uma etiqueta nova que contenha um EPC diferente daquele usado na fase *offline*. O método de multi-frequência (Seção 4.2.1) e o procedimento de agrupamento dos dados (Seção 4.2.2) também são aplicados nesta fase.

Referente ao modelo RNA, uma vez que o modelo esteja treinado, os dados de uma etiqueta desconhecida são apresentados para a rede com o objetivo de estimar sua localização. Para cada frequência, a respectiva instância do modelo RNA é automaticamente escolhida e os dados de RSSI de cada antena alimentam a entrada da rede. A localização da etiqueta (coordenadas na grade virtual) é estimada e apresentada como saída da rede.

Na abordagem SVR, a fase *online* é bastante similar ao modelo RNA. O RSSI de uma etiqueta desconhecida é apresentado ao modelo SVR treinado e a localização da etiqueta é estimada. No SVR, além da frequência, cada coordenada estimada (x, y, z) tem sua própria instância de modelo. Por exemplo, em um sistema 2D, o RSSI de determinada amostra deve alimentar duas instâncias de modelo para obter as coordenadas de localização.

Após a predição ser realizada pela fase *online* dos modelos de aprendizado, as saídas são avaliadas para fornecer um conjunto de estimativas que possa ser combinado a outra tecnologia de localização, resultando assim na posição final do objeto alvo. As próximas seções descrevem os métodos utilizados para a definição deste conjunto de estimativas.

4.3.1 Eliminação de Dados Espúrios

É importante destacar que, em ambas as fases do sistema, o leitor RFID coleta dezenas de leituras para cada etiqueta. Atualmente, os equipamentos e protocolos existentes têm permitido uma coleta com dezenas de leituras em poucos segundos. Por exemplo, em nossos testes, durante 3 segundos lendo uma etiqueta, 46 leituras foram coletadas. Ainda, os valores de RSSI em cada leitura podem variar e o modelo de aprendizado de máquina pode estimar posições diferentes para a mesma etiqueta. Por exemplo, isto pode ocorrer caso o modelo de determinada frequência tiver um baixo desempenho no aprendizado. Assim, cada execução da fase *online* do sistema captura dezenas de leituras RFID, sendo que estratégias de agrupamento ou descarte de dados precisam ser aplicadas para resultar uma localização única (final) do objeto, e não um conjunto de posições.

Um refinamento inicial das estimativas é realizado através da remoção das amostras cuja as coordenadas estão fora dos limites da imagem do cenário. Portanto, uma amostra somente permanecerá na base de dados, caso:

$$g_{min} \times b < p < g_{max} \times b \quad (11)$$

onde p é o valor da coordenada (posição) na amostra avaliada, b é um percentual de tolerância a ser considerado além dos limites da imagem, e g_{min} e g_{max} são as coordenadas mínima e máxima do eixo na grade virtual, respectivamente.

Em um segundo passo, as estimativas que contém erros grosseiros de medição são removidas. Erros grosseiros são aqueles que comprovadamente foram causados por imprevistos, fontes involuntárias de erros, problemas em instrumentos de medição, operação, ou do operador que realiza o experimento, ou seja, todo fator externo, estranho ao experimento, que interfira diretamente no resultado da medição [63].

Com o intuito de eliminar estas estimativas, foi aplicado o método dos quartis. O método dos quartis consiste em analisar o quão distantes estão os extremos da base de dados (mínimo e máximo) do 1º e do 3º quartil¹ [64]. Uma amostra é considerada um dado espúrio, e deve ser removida da base de dados, se:

$$\begin{aligned} p &> Q_3 + 1,5 \times IQR && \text{ou;} \\ p &< Q_1 - 1,5 \times IQR \end{aligned} \tag{12}$$

onde p é o valor da coordenada (posição) que está sendo avaliada em cada eixo, IQR é a amplitude interquartil ($Q_3 - Q_1$) e Q_1 e Q_3 representam os quartis inferior e superior, respectivamente.

Ambos os métodos são aplicados de forma independente para cada coordenada (x , y , z) do conjunto de estimativas. Por exemplo, o valor da coordenada x de determinada amostra é avaliada perante o quartil inferior e superior de todos os valores de x do conjunto de dados.

4.3.2 K-Means

Após a eliminação de dados espúrios, o sistema ainda conta com um conjunto de estimativas fornecidas pelo modelo de aprendizado de máquina. Com o objetivo de agrupar as estimativas e também aprimorar o desempenho da localização baseada somente em RFID, optou-se por uma estratégia de particionamento das posições estimadas em regiões utilizando a técnica k-means (Seção 2.2.4). No caso, cada *cluster* extraído pela técnica k-means equivale a uma região de interesse, sendo o conjunto de *clusters* denominado “múltiplas regiões de interesse” (*Multiple Region of Interest – MROI*).

¹ Quartil é qualquer um dos três valores que divide o conjunto ordenado de dados em quatro partes iguais, e assim cada parte representa 1/4 da amostra ou população. Por exemplo, o quartil inferior (Q_1) é o valor aos 25% da amostra ordenada.

Assim, todas as regiões de interesse são exploradas posteriormente pela técnica de análise visual, abordada na Seção 4.4.

Baseado nos conceitos e definições da técnica k-means, é necessário definir quais são os elementos observados e a métrica de distância a serem utilizados. Neste trabalho, as coordenadas estimadas das etiquetas na grade virtual são os valores observados e a distância euclidiana quadrática é utilizada como métrica de distância.

Como as posições estimadas por instâncias de modelos de aprendizado de determinadas frequências podem diferir de outras frequências, o valor de k da técnica k-means foi definido como $k = f - 1$, onde f é o número de frequências distintas detectadas durante as leituras do sistema RFID. Desta forma, é mais provável que as estimativas observadas em frequências com ruído (interferências, obstáculos, etc.) sejam agrupadas nos seus respectivos *clusters*.

Também foi definido que o peso de cada *cluster* é representado pela quantidade de membros que ele possui. Em um *cluster* com peso maior existem mais membros do que um *cluster* com peso menor. Ou seja, um *cluster* com peso maior agrega mais posições próximas umas das outras do que um *cluster* com peso menor. Na maioria dos casos, um *cluster* com poucos membros apresenta posições espúrias ou posições em que o aprendizado de máquina não obteve uma resposta razoável. O *cluster* com o maior peso é aquele que teve a maior quantidade de posições similares estimadas pelo modelo de aprendizado de máquina.

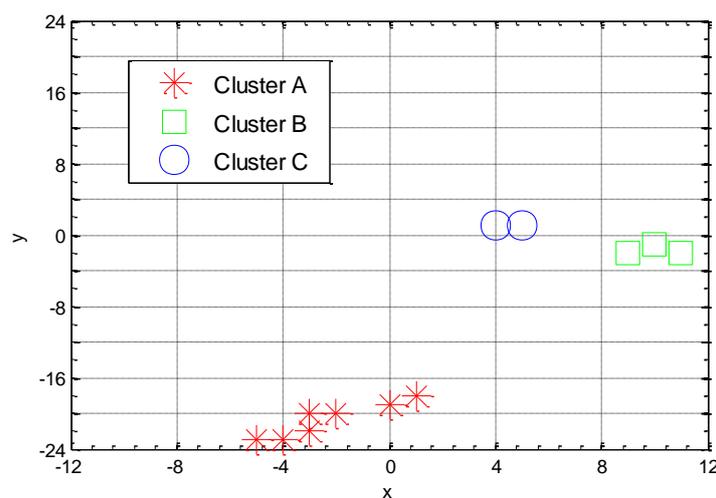


Figura 38 – Técnica k-means aplicada às posições estimadas sobre as coordenadas (x, y) da grade virtual. Fonte: o autor.

A Figura 38 exibe os *clusters* extraídos de um conjunto de posições estimadas para uma etiqueta específica. Amostras de leituras de quatro frequências de operação entre 923 e 924 MHz foram utilizadas. Alguns valores observados estão sobrepostos, sendo que os *clusters* A, B e C possuem 11, 7 e 3 membros, respectivamente. No caso, o *cluster* A é o que possui melhor peso.

A lista de MROI é criada sendo composta pelo centroide e peso de cada *cluster*. Esta lista é fornecida para a análise visual. A Tabela 5 apresenta uma lista de MROI baseada nos valores de exemplo do gráfico da Figura 38. No caso, os ROIs 1, 2 e 3 têm como origem os *clusters* A, B e C, respectivamente.

Tabela 5 – Exemplo de lista de MROI fornecida pela fase k-means para o subsistema visual.

#	Posição centroide	Peso
1	(-2,6; -20,9)	11
2	(9,8; -1,4)	7
3	(4,3; 1,0)	3

Apesar dos exemplos estarem focados somente nas coordenadas (x, y) , a técnica k-means também é aplicada sob escalabilidade 3D (x, y, z) . Entretanto, como no cenário 3D a coordenada z representa a distância no espaço do ambiente real (distância em cm), as coordenadas x e y são convertidas para esta mesma escala utilizando uma lista de divisores pré-definidos.

O sistema proposto também pode ser ajustado para utilizar somente o subsistema RFID. Esta opção é útil quando uma câmera não está disponível ou para comparações com o sistema híbrido. Para isto, é necessária alguma técnica que forneça a localização final ao usuário. Testes iniciais mostraram que a média simples das posições estimadas não traziam resultados satisfatórios. Assim, caso o sistema esteja em modo “somente RFID”, as coordenadas da posição centroide do ROI com melhor peso é considerada a localização final do objeto alvo. No exemplo da Tabela 5, a resposta seria $(-2,6; -20,9)$.

4.4 Integração com Análise Visual

Com o objetivo de refinar os resultados do subsistema RFID, este trabalho propõe o uso de análise visual a partir da lista de MROI resultante da fase *online* do subsistema RFID. Nesta abordagem, algoritmos de visão computacional processam pequenas regiões

da fotografia original do cenário a fim de detectar os marcadores visuais, possibilitando uma localização mais rápida e precisa.

O algoritmo da Figura 39 descreve a sequência de operações do subsistema visual. A lista de MROI fornecida pelo subsistema RFID contém as posições centroides e os pesos dos *clusters* k-means (Seção 4.3.2). Para cada região da lista de MROI, uma imagem pequena do cenário é recortada, criando uma subimagem. O centro da subimagem está baseada nas coordenadas do centroide da região. A subimagem possui dimensões entre 15 e 30% da imagem completa do cenário, sendo que este valor deve ser configurado pelo usuário (dimensão do ROI).

Algorithm 1: ObterLocalizacaoMarcadorVisual(*img*, *C*)

Input: *img*, fotografia capturada do cenário
Input: conjunto *C*, lista de MROI contendo posições (C^p) e pesos (C^w)
Output: Localização final do marcador

```

1 melhorPeso  $\leftarrow$  0, M  $\leftarrow$   $\emptyset$ , i  $\leftarrow$  0;
2 foreach regiao in C do
3   subImagem  $\leftarrow$  RecortarImagem(img, regiaop);
4   S  $\leftarrow$  DetectarPosicaoMarcador(subImagem);
5   foreach posicaoMarcador in S do
6      $M_i^p \leftarrow$  ObterPosicaoImagem(posicaoMarcador);
7      $M_i^w \leftarrow$  regiaow;
8     i  $\leftarrow$  i + 1;
9     if regiaow > melhorPeso then
10    |   melhorPeso  $\leftarrow$  regiaow;
11 if M  $\neq$   $\emptyset$  then // Marcador visual encontrado
12 |   B  $\leftarrow$  {indices( $M^w$ ) |  $M^w =$  melhorPeso};
13 |   F  $\leftarrow$  { $M_i^p$  | i  $\in$  B};
14 |   return ( $\frac{1}{n} \sum_{i=1}^{i=n} F_i^x, \frac{1}{n} \sum_{i=1}^{i=n} F_i^y$ );
15 else // Nenhum marcador encontrado (somente RFID)
16 |   return  $C_i^p$ , onde i = indice(max( $C^w$ ));

```

Figura 39 – Algoritmo contendo a lógica de funcionamento do subsistema visual. Fonte: o autor.

Para detectar a posição do marcador visual na subimagem (Linha 4), dois tipos de marcadores visuais podem ser escolhidos: codificados ou não-codificados. No uso de marcadores codificados, deve existir uma ligação entre o EPC da etiqueta RFID e a informação contida no marcador visual respectivo. Assim, o marcador que possui o código correspondente à etiqueta RFID lida é detectado e sua posição identificada. No caso de marcadores não-codificados, um algoritmo de reconhecimento de padrões ou formas deve

ser utilizado. Neste caso, as posições de todas as formas gráficas detectadas são retornadas. Os detalhes do processo de detecção dos marcadores visuais codificados e não-codificados são apresentados nas Seções 4.4.1 e 4.4.2, respectivamente.

Como o marcador visual é detectado na subimagem, é necessário conhecer a posição do marcador na imagem completa do cenário (Linha 6). Para isto, a equação (13) deve ser aplicada:

$$(x, y) = \left(x_m + x_s - \frac{w}{2}, y_m + y_s - \frac{h}{2} \right) \quad (13)$$

onde (x_s, y_s) são as coordenadas do centro da subimagem na imagem original, (x_m, y_m) são as posições do marcador visual na subimagem, w e h são a largura e a altura da subimagem, respectivamente.

Finalmente, a localização do objeto alvo é definida pela posição do marcador visual detectado no ROI com maior peso (Linhas 12-13). Caso mais de um marcador visual seja detectado nesta mesma região, o centroide simples das coordenadas destes marcadores é calculado. Caso nenhum marcador visual seja detectado, é informada somente a localização do subsistema RFID para o usuário.

4.4.1 Marcador Visual Codificado

Em uma primeira proposta de trabalho, o modelo de localização visual proposto baseou-se na detecção e leitura de marcadores visuais, codificados em imagens capturadas do ambiente. Esta abordagem de localização visual fornece distinção entre cada objeto do cenário, pois utiliza um identificador único para cada marcador [57].

Como primeira tentativa, foram realizados experimentos utilizando marcadores visuais coloridos, ou seja, cada marcador recebe uma cor distinta e esta cor é detectada e localizada no cenário. Entretanto, como já discutido na Seção 2.1.2.1, as condições de luminosidade geralmente são impróprias para realizar a identificação deste tipo de marcador. Tal fato foi experienciado nos testes realizados, existindo uma alta incidência de erros de detecção. Outra desvantagem desta abordagem é a existência, no cenário, de objetos de cor semelhante ao marcador alvo, causando uma falsa detecção do marcador visual desejado. Sendo assim, tal abordagem foi abandonada, partindo-se para alternativas que pudessem ser inseridas em ambientes mais próximos da realidade, ou seja, ambientes com variância na iluminação e contendo diversos objetos de cores semelhantes.

Assim, optou-se pela utilização de código de barras como modelo de marcador visual. Apesar de ser uma tecnologia existente há vários anos, pesquisas abordando a detecção, localização e leitura de código de barras em imagens digitais continuam sendo conduzidas, sendo um desafio na área de visão computacional [23], [97].

A escolha pelo código de barras se deu em função de ser um marcador visual codificado, ou seja, é um marcador que possui um código associado à imagem. Também optou-se por este modelo pela facilidade de implementação e bibliotecas de software disponíveis. O código de barras é um marcador visual bitonal, trazendo melhor desempenho ao algoritmo de detecção. Além disso, acredita-se que, através da utilização do código de barras, o uso e a validação dos demais componentes do IPS proposto possam ser conduzidos de forma satisfatória, não sendo necessária a implementação de um marcador visual de maior complexidade.

Dentre os códigos de barras mais populares e de larga utilização destacam-se o Code39 e o Code128 (Figura 40). Uma das desvantagens do Code39 em relação ao Code128 é a sua baixa densidade de dados, ou seja, o Code39 necessita de mais espaço para codificar e representar as informações contidas no seu código de barras. Levando em consideração que um dos requisitos do sistema proposto é a localização de objetos de pequeno porte, um marcador visual que ocupe menor espaço físico contempla melhor esta restrição. Tendo em vista as características de cada um dos modelos de código de barras e, principalmente, a alta densidade de dados do Code128, este trabalho optou pela utilização deste modelo na localização do marcador visual.



Figura 40 – Exemplo de códigos de barras Code128 e Code39. Fonte: [84].

A partir da captura realizada pela câmera, o sistema aplica sobre a imagem um algoritmo de detecção, leitura e localização do código de barras. Cada uma destas fases pode ser descrita como:

- i. Detecção: fase do algoritmo que tenta encontrar um possível código de barras no conteúdo da imagem.
- ii. Leitura: fase que analisa cada barra e verifica qual o código (identificador) que está armazenado no conjunto de barras.

- iii. **Localização:** fase que, a partir da detecção, deve retornar a posição (x, y) do código de barras no interior da imagem. Somente códigos de barras em que as fases de detecção e leitura obtiveram sucesso retornam sua informação de localização.

As fases de detecção e leitura foram implementadas a partir da biblioteca *BarcodeImaging* [21], que possui seu código-fonte aberto e é licenciada sob a licença *The Code Project Open License (CPOOL)* [15]. A fase de localização necessitou de alterações no código-fonte da biblioteca, podendo assim serem fornecidas as coordenadas (x, y) de cada código de barras detectado.

A biblioteca *BarcodeImaging* utiliza um modelo de algoritmo baseado em histogramas e no método *threshold*, a fim de detectar e ler o código de barras na imagem. Cada iteração do algoritmo é aplicada a uma porção de linhas da matriz de *pixels* (chamado no algoritmo de “banda”). A quantidade de linhas de cada banda é configurada pelo usuário ou desenvolvedor. O algoritmo sempre executa uma varredura completa da imagem, tanto na orientação horizontal quanto vertical.

Com o intuito de satisfazer os requisitos do sistema de localização proposto neste trabalho, o código-fonte da biblioteca *BarcodeImaging* foi alterado para fornecer as coordenadas de cada código de barras detectado e lido. Para isto, foram adicionadas estruturas de dados onde, a cada código de barras cujo número natural fosse obtido com sucesso, a posição da banda corrente (x) e a posição central horizontal do código de barras (y) fossem armazenadas e fornecidas como saída do algoritmo.

4.4.2 Marcador Visual Não-Codificado

Em uma segunda proposta, optou-se por um marcador visual não-codificado, ou seja, ele não possui informações que o identifiquem unicamente no cenário. Esta característica facilita a adoção de novos marcadores visuais e algoritmos de visão computacional, pois é facilmente replicado e basta o algoritmo reconhecer um único tipo de forma ou padrão. Como desvantagem, este marcador visual não possui capacidade de representar um código ou característica única que possa referenciar outro elemento do sistema.

Nesta abordagem, optou-se pela utilização de um marcador visual simples representado por um quadrado preto sobre um fundo branco, cujo lado mede 4,5 cm (20,25 cm²). Este tipo de marcador tem como vantagem a facilidade de criação, baixo

custo e rápida detecção (menos de 1 s). Pelo fato do marcador ser preto-e-branco, os problemas com iluminação do ambiente tendem a ser menores.

O marcador visual proposto também pode servir como indicativo da probabilidade de sucesso do IPS desenvolvido no que diz respeito à análise das múltiplas regiões de interesse fornecidas pelas estimativas RFID. Assim, os resultados obtidos com este marcador também podem auxiliar no desenvolvimento de novos marcadores visuais que contemplem cenários ainda mais complexos.



Figura 41 – Etapas do subsistema visual alinhadas aos passos fundamentais de um sistema de visão computacional. Fonte: o autor.

O diagrama da Figura 41 ilustra o fluxo das etapas do subsistema visual. Após a imagem ser capturada por uma câmera digital, a fase de pré-processamento converte a imagem para tons de cinza e aplica um filtro gaussiano com o objetivo de suavizar e reduzir o ruído na imagem.

Na implementação da fase de segmentação, um detector de bordas Canny (Seção 2.1.2.1.1) é empregado. O valor superior e inferior do *threshold* é definido em 180 e 120, respectivamente. A partir da imagem resultante do filtro de Canny, os contornos são detectados e as forma geométricas que possuem quatro ângulos entre 80 e 100 graus são selecionadas. Se a área da forma detectada for superior a 40 *pixels*, ela é reconhecida como um marcador visual presente no ambiente.

Para implementar este subsistema foi utilizada a biblioteca EmguCV [20], que é um *wrapper*¹ do OpenCV² [69] para plataforma .NET [59]. A Figura 42 exibe o recorte de uma região de interesse e a detecção de um marcador visual não-codificado.

¹ Função wrapper é uma sub-rotina em uma biblioteca de software cuja principal finalidade é chamar uma segunda sub-rotina ou uma chamada de sistema, com pouco ou nenhum processamento adicional.

² OpenCV (*Open Source Computer Vision*) é uma biblioteca de programação, de código aberto, desenvolvida inicialmente pela Intel Corporation. O OpenCV implementa uma variedade de ferramentas de interpretação de imagens, abrangendo operações simples e complexas, tais como a análise de movimentos e reconhecimento de padrões [55].

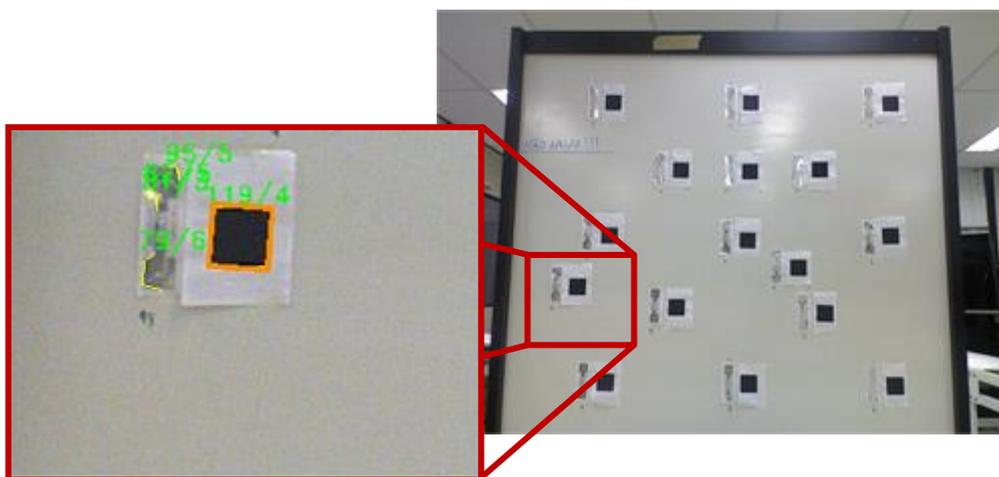


Figura 42 – Subimagem recortada da fotografia original do cenário e a detecção do marcador visual. Fonte: o autor.

5 EXPERIMENTOS E RESULTADOS

Este capítulo descreve os experimentos que foram realizados com o objetivo de avaliar as configurações e o desempenho do IPS proposto. O sistema contém uma série de variáveis que necessitam ser analisadas passo-a-passo a fim de encontrar uma melhor configuração dos componentes. Estas variáveis podem ser representadas como propriedades da configuração primária do sistema. A configuração primária que será avaliada inicialmente contempla os seguintes itens:

- i. Tipo de marcador visual: indica se o subsistema de análise visual utiliza marcadores visuais codificados ou não-codificados.
- ii. Posicionamento das antenas: refere-se à posição física e forma de organização das antenas conectadas ao leitor RFID.
- iii. Forma de organização das etiquetas de referência: posicionamento e quantidade de etiquetas de referência empregados na fase *offline* do sistema.
- iv. Tempo de leitura das etiquetas de referência: período em que o leitor RFID permanece ativado realizando a leitura das etiquetas na fase *offline*.
- v. Tempo de leitura das etiquetas na fase *online*: período de ativação do leitor RFID, para cada etiqueta alvo, na fase *online*.

As Seções 5.1 até 5.4 descrevem os experimentos e resultados que chegaram a configuração primária mais apropriada. Posteriormente, as Seções 5.5, 5.6 e 5.7 utilizam esta configuração para validar os métodos desenvolvidos e demonstrar o uso do sistema em cenários distintos. Além disso, os testes iniciais do sistema empregaram somente o modelo RNA como aprendizado de máquina. O modelo SVR é avaliado posteriormente, quando as configurações primárias já estão consolidadas. Os componentes e configurações constantes na Tabela 6 permaneceram fixas durante a realização de todos os ensaios.

No subsistema visual, uma câmera digital (estilo *webcam*) foi utilizada. As características da câmera são descritas na Tabela 7. A câmera possui baixa resolução, fornecendo imagens com tamanho reduzido e possibilitando um processamento mais rápido pelos algoritmos de visão computacional. Além disso, câmeras com estas características geralmente possuem custo mais baixo e são facilmente encontradas no mercado ou até na infraestrutura já existente das organizações.

Tabela 6 – Componentes e respectivas propriedades que permaneceram fixos em todas as seções de experimentos.

Componente	Propriedade	Valor
Leitor	Modelo	Impinj Speedway Revolution R420
	Frequências	923,25; 923,75; 924,25; e 924,75 MHz
	Potência	32,5 dBm
Antena	Modelo	Impinj Threshold IPJ-A0311-USA
	Faixa de frequência	902-928 MHz
	Ganho	5 dBi ¹
Etiqueta	Modelo	UPM RafSec DogBone 3000825
	Faixa de frequência	860-960 MHz
	Objeto em que está afixada	Quadro branco de material plástico
	Orientação	Vertical

Tabela 7 – Características da câmera digital utilizada no subsistema visual.

Propriedade	Valor
Fabricante	Microsoft
Sensor	Aptina MT9M112
Resolução	1,3 MP (1280x960) - RGB 32-bits
Sensor	1/4"
Formato de imagem	<i>raw</i>

Durante a fase *offline*, a interface gráfica do sistema (Figura 43) necessita ser configurada para registrar a posição conhecida dos marcadores. Para isto, um ícone contendo o EPC² da etiqueta deve ser posicionado sobre a sua respectiva posição no cenário. Além desta configuração, também deve ser informada a resolução da grade virtual. Em todos os ensaios, a resolução permaneceu configurada em 37 linhas e 50 colunas. Cabe destacar que o tamanho da célula no cenário real pode variar de acordo com a distância entre a câmera e o alvo. Por exemplo, em uma imagem que o objeto alvo está a 1 metro da câmera, cada célula (quadrada) da grade equivale à aproximadamente 3,5 cm de lado no cenário real (equação 14, variável r).

Todos os experimentos foram executados em laboratório (LSE – Laboratório de Sistemas Embarcados, PUCRS). O leitor RFID estava conectado a um computador através de rede Ethernet. Tanto na fase *offline* quanto na fase *online*, os marcadores

¹ dBi: dB isotrópico, medida utilizada para expressar o ganho de uma antena comparado a uma antena isotrópica (que distribui energia uniformemente em todas as direções).

² A fim de simplificar a leitura dos valores EPC, somente os últimos dígitos do código são apresentados.

(etiqueta e visual) estavam afixados em um quadro branco com área de 1,7 m². A câmera estava posicionada a 1,4 m do solo, de forma perpendicular ao centro do quadro.

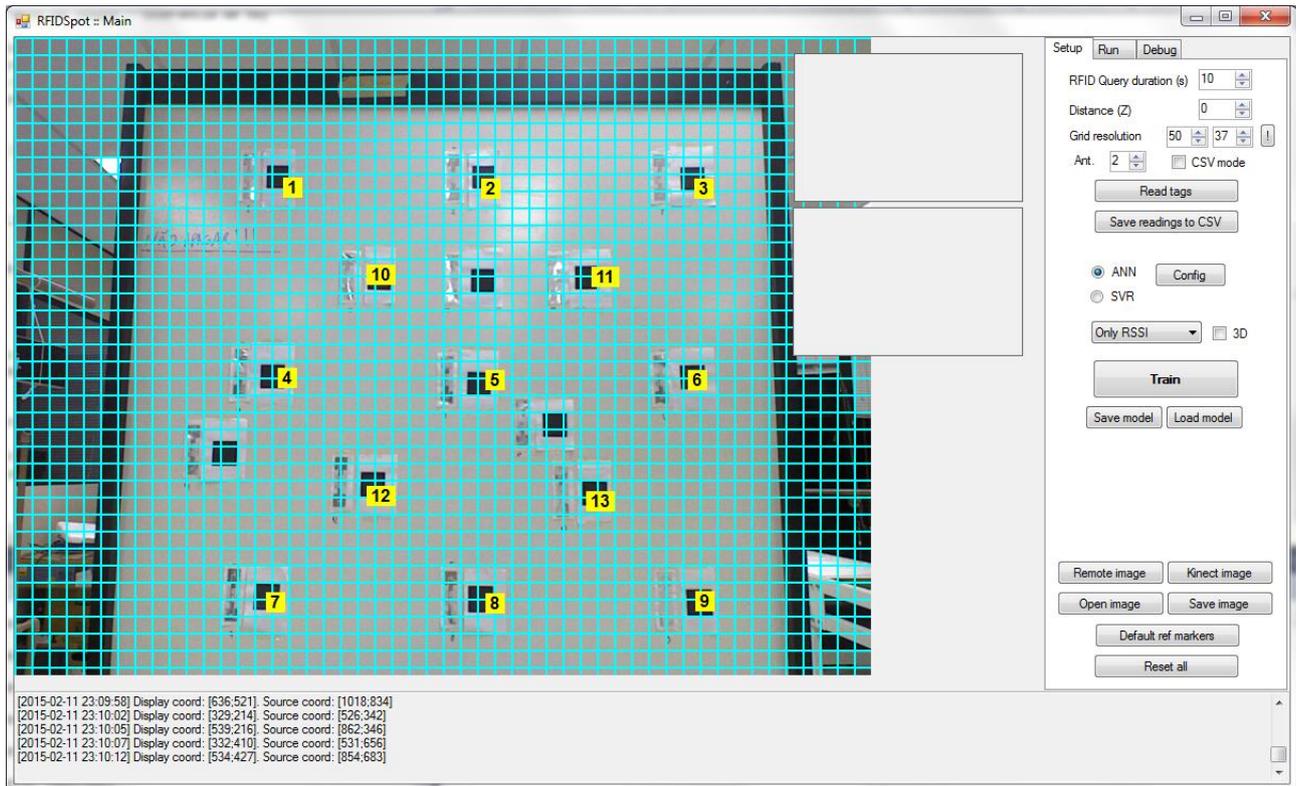


Figura 43 – Captura de tela do processo de configuração das posições conhecidas e resolução da grade virtual. Os ícones amarelos representam os marcadores conhecidos e seus respectivos EPCs. Fonte: o autor.

Após a execução da fase *online*, o sistema estima a localização dos marcadores alvo na grade virtual (coordenadas das células da grade). A fim de fazer a leitura dos resultados utilizando uma unidade de medida real e de fácil assimilação, a distância entre a célula estimada e a célula que corresponde a posição real do marcador precisa ser convertida para centímetros. Esta conversão foi realizada através da multiplicação

$$d_{cm} = r * d_g \quad (14)$$

onde r é o tamanho real da célula em centímetros, d_g é a quantidade de células entre o marcador estimado e real e d_{cm} é a distância estimada em centímetros. O valor de r é medido manualmente e pode variar de acordo com cada teste, sendo informado no decorrer das próximas seções.

Em relação à avaliação dos resultados, foram utilizadas as métricas definidas na literatura [35], [37], [52]. Tipicamente, o erro de localização (precisão) é utilizado como métrica de desempenho e comparação entre IPSs, sendo obtido através da distância

Euclidiana entre a localização estimada e a posição real. Em um espaço com duas dimensões, esta distância pode ser calculada através da equação (15):

$$e = \sqrt{(\hat{x} - x)^2 + (\hat{y} - y)^2} \quad (15)$$

onde (\hat{x}, \hat{y}) e (x, y) representam as coordenadas estimadas e as coordenadas reais de cada marcador, respectivamente. Todos os resultados de erro dos experimentos estão na unidade de centímetros, conforme conversão apresentada na equação (14).

Após este cálculo, o erro precisa ser ajustado de acordo com uma tolerância baseada no tamanho do marcador visual empregado. Isto é necessário pois o marcador visual possui área de 20,25 cm², sendo que muitas estimativas possíveis podem ocorrer em pontos dentro desta área. Nestes casos, optou-se por assumir que o erro é zero, pois a estimativa ficou dentro dos limites do marcador visual alvo. Assim, foi definido um ajuste no erro $e = \begin{cases} 0, & e \leq b \\ e, & e > b \end{cases}$, onde b é a tolerância em centímetros. No caso do marcador “não-codificado”, $b = 4,5$.

Com o objetivo de sintetizar os resultados de várias repetições e combinações de experimentos, a precisão da localização é demonstrada através da raiz do erro médio quadrático (*Root Mean Square Error* – RMSE). O RMSE é calculado pela diferença apresentada na equação (16) tal que

$$RMSE = \sqrt{\frac{\sum_{t=1}^k (\hat{x}_t - x_t)^2 + (\hat{y}_t - y_t)^2}{k}} \quad (16)$$

onde (\hat{x}_t, \hat{y}_t) são as coordenadas estimadas em cada localização realizada pelo sistema, (x_t, y_t) representam as coordenadas reais do marcador e k é o número de estimativas. A conversão de coordenada virtual para centímetros e o cálculo de fator de ajuste também são aplicados nesta métrica.

5.1 Tipo de Marcador Visual

Na primeira fase de experimentos, somente o subsistema de localização visual foi executado utilizando um cenário simplificado contendo um único marcador visual. Como marcador codificado, foi utilizado um código de barras no formato Code128 com dimensões de 7,6 cm de largura por 6 cm de altura. O papel no qual o código de barras está impresso foi posicionado em frente à câmera utilizada, em linha reta. No caso do marcador visual não-codificado, foi impresso um quadrado preto com lado de 4,5 cm em

uma folha de papel branca. Ambos tipos de marcador foram afixados em um quadro branco.

O objetivo do experimento foi verificar a distância mínima e máxima de detecção para cada tipo de marcador, sendo os resultados medidos através da precisão (erro de localização) e probabilidade de sucesso.

Foram executados testes em distâncias entre 20 e 220 cm de profundidade, com intervalo de 20 cm entre cada experimento. Para cada distância, o teste foi repetido cinco vezes e a média simples foi utilizada como resultado. O marcador visual estava posicionado em linha reta em relação à câmera. A resolução de imagem da câmera foi configurada em seu valor máximo (1280x960).

5.1.1 Resultados

Nesta fase, em todos os experimentos que houve detecção, a precisão foi ótima, sem erros na distância. Entretanto, em distâncias superiores a 60 cm, o sistema não foi capaz de detectar o marcador visual codificado nas imagens capturadas, obtendo probabilidade zero de sucesso na estimativa. Os resultados da probabilidade de sucesso para cada tipo de marcador e a respectiva distância em profundidade são apresentados no gráfico da Figura 44.

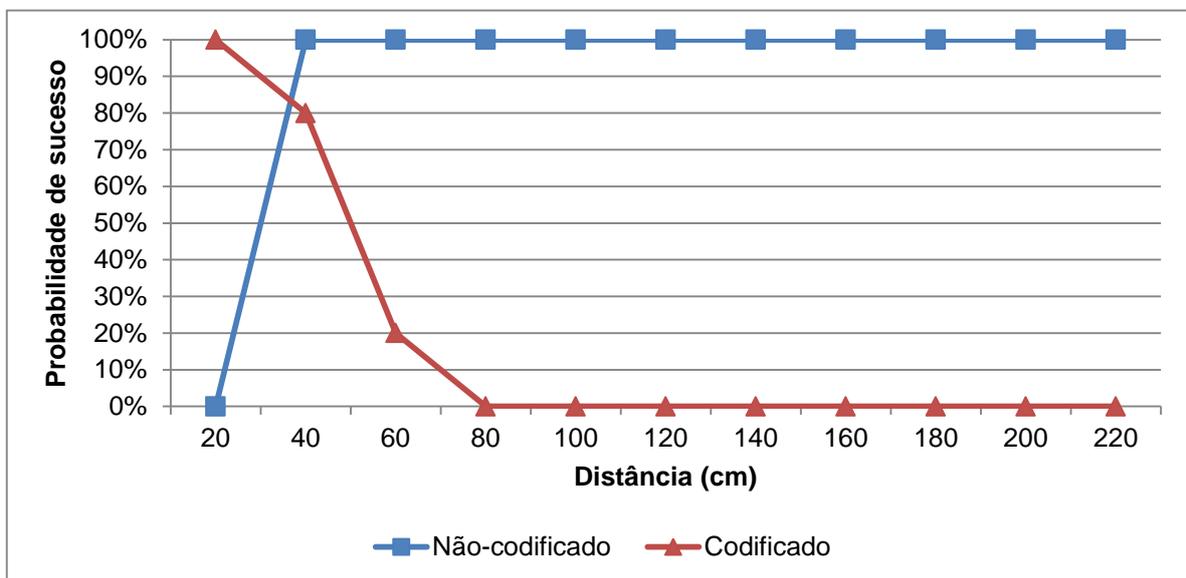


Figura 44 - Probabilidade de sucesso para cada tipo de marcador visual entre 20 e 220 centímetros. Fonte: o autor.

Na avaliação da localização visual, a precisão sem erros demonstra uma boa corretude nos algoritmos empregados, pois em todos os casos que houve sucesso na localização do marcador, a estimativa correspondia à posição real do marcador. No

entanto, em diversos casos do marcador codificado, houve insucesso na detecção do código de barras. A precisão do sistema foi diminuindo à medida que a distância aumentava, chegando ao limite máximo de 60 cm, com somente 20% de sucesso.

A Figura 45 ilustra, com um zoom de 1.200%, o código de barras capturado pela câmera a 1 metro de distância. O efeito de serrilhamento¹ (*aliasing*) pode estar sendo causado por dois motivos: a resolução do sensor é baixa em densidade (isto é, em *pixels* por polegada) ou a imagem está sendo ampliada pelo próprio processamento da câmera, ou ambos. Outra limitação identificada está presente nas lentes da câmera, que possivelmente está causando aberração cromática² entre o contraste do branco da folha e a cor preta das barras.



Figura 45 – Imagem do código de barras capturado. Zoom: 1.200%. Distância: 1 m. Fonte: o autor.

O nível de iluminação incidente sobre o marcador codificado durante os experimentos é um possível fator de instabilidade da qualidade das imagens capturadas. Apesar do laboratório permanecer com todas as luzes acesas durante o processo, algum fenômeno interferente pode ter reduzido a qualidade e a bitonalidade do código de barras, fazendo com que o algoritmo de localização visual não pudesse realizar a correta diferenciação de cor entre cada feixe de barras do marcador visual codificado.

Em relação ao marcador visual não-codificado, este modelo somente não foi detectado na distância de 20 cm. Nesta distância, a câmera capturou uma imagem em que o marcador visual ficou borrado (desfocado). Como consequência, o algoritmo que identifica os contornos das bordas detectadas na imagem resultaram em uma forma geométrica de 6 lados, ao invés de 4. Como os requisitos do IPS são mais prioritários no que diz respeito ao alcance em distâncias mais longas, optou-se por definir o tipo de marcador não-codificado na configuração primária do sistema.

¹ Serrilhamento (*aliasing*) é o efeito em forma de serra que se cria ao desenhar uma reta em diagonal. Uma vez que a divisão mínima em um recurso digital é discreta (*pixels*), surge o aparecimento dos "dentes" da serra ao longo da reta desenhada.

² Aberração cromática é a dispersão produzida por lentes que possuem diferentes índices de refração para diversos comprimentos de onda luminosa. Este fenômeno pode gerar interferências nas imagens obtidas, principalmente em arestas e bordas, onde os componentes da luz parecem se misturar em mais de um *pixel* da câmera, fazendo com que *pixels* de borda fiquem com uma cor não esperada.

5.2 Posicionamento das Antenas

O posicionamento das antenas é uma variável fundamental para o bom desempenho dos sistemas de localização, pois o valor de RSSI refletido por uma etiqueta pode sofrer grandes variações dependendo do ângulo e distância em que a respectiva antena esteja posicionada [43]. Este é o primeiro conjunto de ensaios que contempla todo o processo de localização do IPS proposto. Ou seja, todas as fases do sistema são executadas com o objetivo de localizar marcadores presentes no cenário. Este teste tem como principal objetivo verificar qual o melhor posicionamento para as antenas conectadas ao leitor RFID. A partir do resultado desta configuração, outros ensaios poderão utilizar somente o posicionamento de antenas mais adequado, não necessitando repetir os experimentos para cada forma de posicionamento.

Três formatos de posicionamento foram avaliados (Figura 46). A primeira estratégia, denominada 1AF (1 antena frontal), possui uma antena RFID posicionada em frente ao quadro. Esta posição pode trazer benefícios de organização aos componentes do sistema, pois neste formato a câmera digital (que está sempre posicionada em frente ao cenário), pode ser acoplada junto à antena. O formato 2AF (2 antenas frontais) emprega o mesmo estilo do formato anterior, entretanto são utilizadas duas antenas com o objetivo primário de possivelmente diferenciar os valores de RSSI coletados, principalmente nas etiquetas que estão em posições distintas verticalmente (eixo y). A última abordagem, denominada EIXO, também utiliza duas antenas, porém neste formato cada antena está ao lado e abaixo do cenário alvo. Ou seja, a antena lateral realiza as medições RSSI em relação ao eixo x (horizontal) e a segunda antena (inferior) mede o RSSI em relação ao eixo y (vertical).

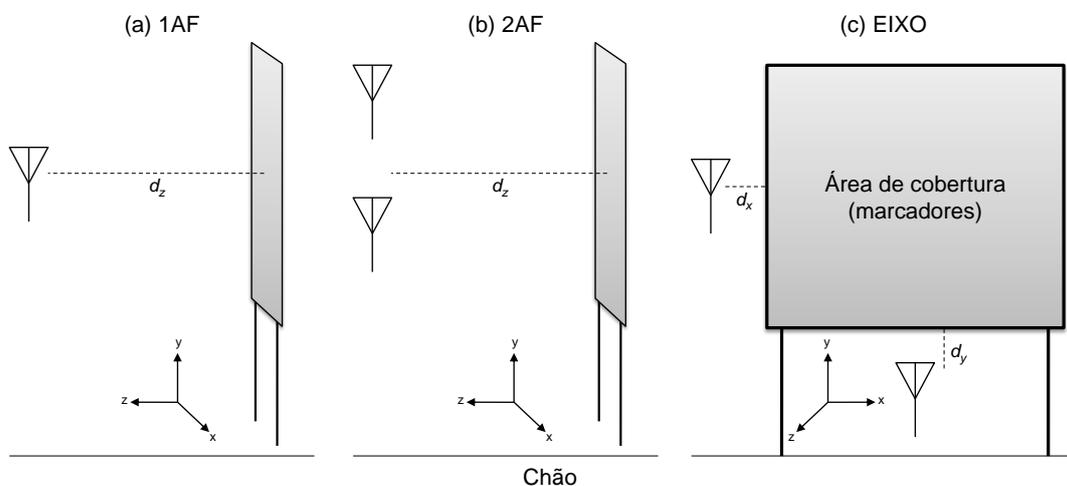


Figura 46 - Antenas e suas posições em relação ao quadro onde os marcadores são afixados. Fonte: o autor.

Em todos os formatos, a câmera do subsistema visual foi posicionada em frente ao cenário, de maneira centralizada, a 1 metro de distância do quadro. Esta era a distância mínima que possibilitava uma fotografia que capturasse todos os marcadores presentes. Nos formatos que utilizam antenas frontais, a distância (d_z) entre a antena e o quadro acompanhou a câmera, ou seja, foi fixada em 1 m. No formato EIXO, as distâncias até o quadro foram definidas em 20 cm (d_x) e 30 cm (d_y) para a antena lateral e inferior, respectivamente. Estas distâncias permaneceram fixas durante todas as fases do sistema (*offline* e *online*). A Figura 47 exibe o ambiente de testes e todos os componentes do sistema, sendo utilizado o formato de posicionamento EIXO.

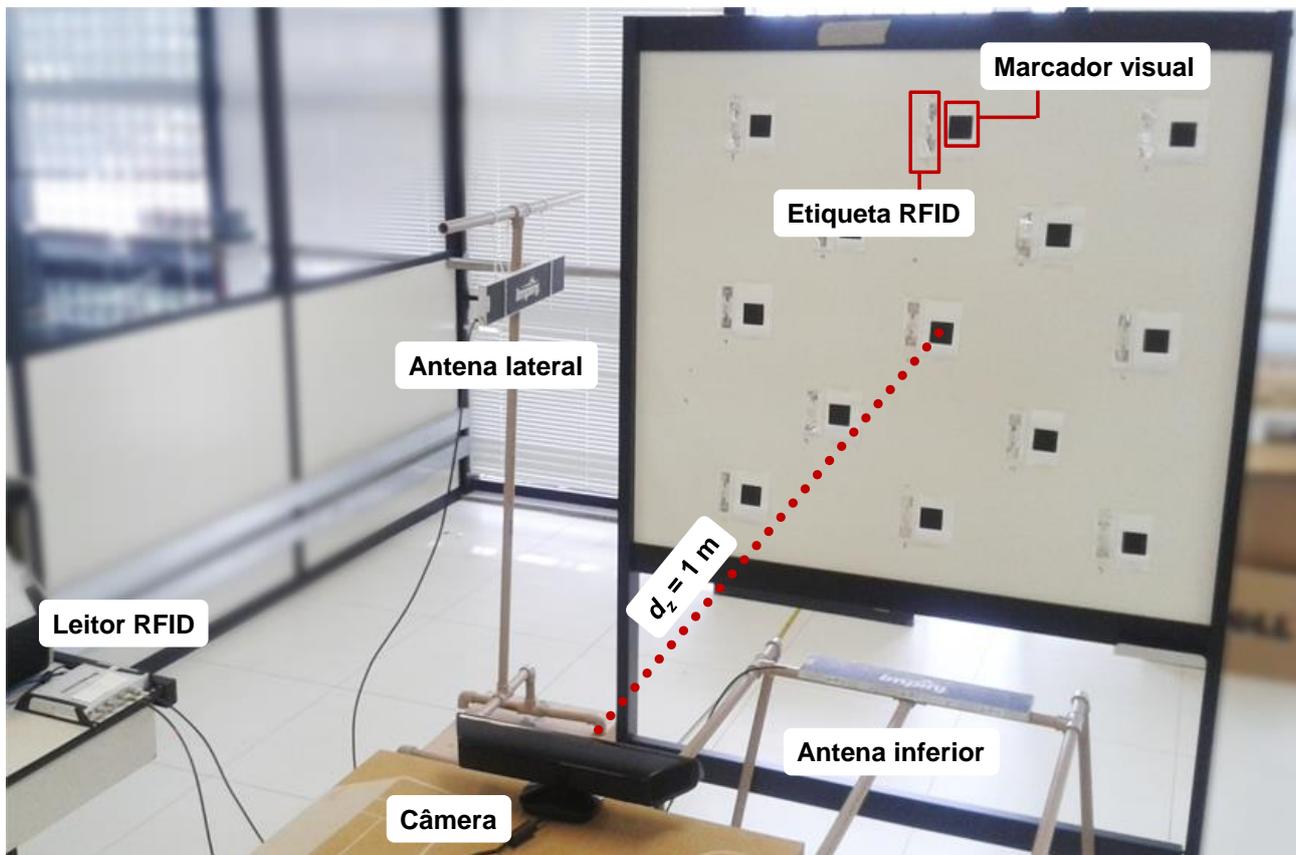


Figura 47 – Ambiente de testes e todos os componentes do sistema. Formato de posicionamento das antenas: EIXO. Fonte: o autor.

A fim de restringir o escopo de avaliação para tão somente o posicionamento das antenas, esta fase de testes manteve fixa as propriedades de algumas configurações primárias do sistema, conforme listado na Tabela 8. Por demonstrar um desempenho superior nos testes da Seção 5.1, o tipo de marcador visual empregado foi o não-codificado. As etiquetas de referência foram organizadas visando contemplar uma maior cobertura do cenário. Os tempos de leitura do leitor RFID foram definidos em valores que pudessem coletar um número mínimo de amostras, para cada frequência de operação.

Tabela 8 – Propriedades das configurações primárias do sistema empregados nos testes iniciais.

Propriedade	Valor
A. Tipo de marcador visual	não-codificado
B. Posicionamento das antenas	(variável)
C. Organização das etiquetas de referência	13 etiquetas em forma de grade diagonal
D. Tempo de leitura das etiquetas de referência	10 segundos
E. Tempo de leitura das etiquetas na fase <i>online</i>	3 segundos

Durante a fase *offline*, o leitor RFID coletou um total de 1604 amostras, sendo 533 para a estratégia 1AF, 525 para o formato 2AF e 546 registros no último formato (EIXO). Todas as etiquetas de referência foram detectadas por todas as antenas nas diferentes estratégias de posicionamento. A partir destas coletas, o modelo RNA foi treinado independentemente para cada uma das estratégias de posicionamento.

Na fase *online*, o objetivo dos experimentos era localizar seis marcadores distribuídos no cenário. Três marcadores estavam em posições já utilizadas na fase *offline* e outros três estavam em posições desconhecidas. A Figura 48 ilustra os marcadores que devem ser localizados (triângulos vermelhos) e seus respectivos identificadores (EPC). No caso, os marcadores 1, 5 e 9 estão em posições já utilizadas na fase *offline*, enquanto os marcadores 14, 15 e 16 estão em locais nunca antes treinados. Durante a fase *online*, as etiquetas de referência não estão presentes no cenário.

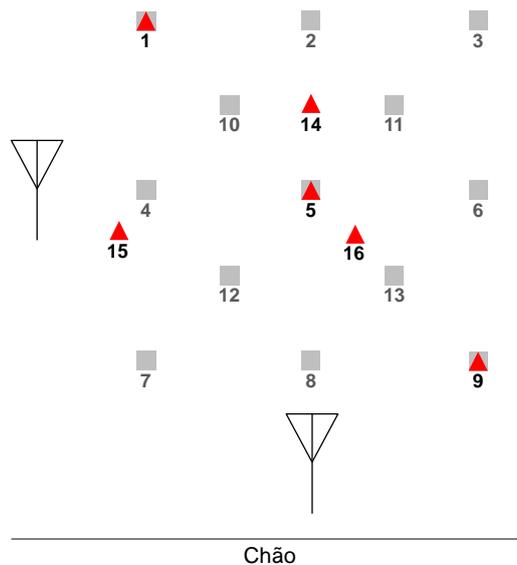


Figura 48 – Ilustração do cenário contemplando as etiquetas de referência (quadrados cinzas) e os marcadores alvo da fase *online* (triângulos vermelhos). Fonte: o autor.

Ainda dentro das configurações da fase *online*, o tamanho mínimo para detecção do marcador visual foi definido em 40 *pixels*. A dimensão do ROI (subimagem) foi configurada em 30%. Com este percentual, a região da imagem extraída possui dimensão aproximada de 35 x 27 cm (l x a) no cenário real.

A fase *online* do sistema foi avaliada nos modos de execução “somente RFID” e híbrido. Nos ensaios em modo híbrido, para cada teste, apenas um marcador visual estava presente no cenário. A localização de cada marcador é executada de maneira sequencial, ou seja, um marcador alvo é localizado por vez. Para cada execução, uma média de 58 amostras foram coletadas pelo leitor RFID. Visando minimizar erros estatísticos, cada experimento foi repetido três vezes.

5.2.1 Desempenho da Fase *Offline*

Como resultado da fase *offline*, os valores de MSE (Seção 2.2.2.2, equação 8) do treinamento das redes neurais modeladas são apresentados no gráfico da Figura 49. Como está sendo empregado o método de multi-frequência (Seção 4.2.1), uma rede foi criada para cada frequência detectada no conjunto de amostras. As médias gerais do MSE, considerando todas as frequências, foram de 50,2, 3,0 e 0,0 para os formatos 1AF, 2AF e EIXO, respectivamente. A unidade de medida utilizada no MSE é a mesma dos valores de saída do conjunto de treinamento (não é usada normalização), ou seja, o valor refere-se à distância nas coordenadas da grade virtual.

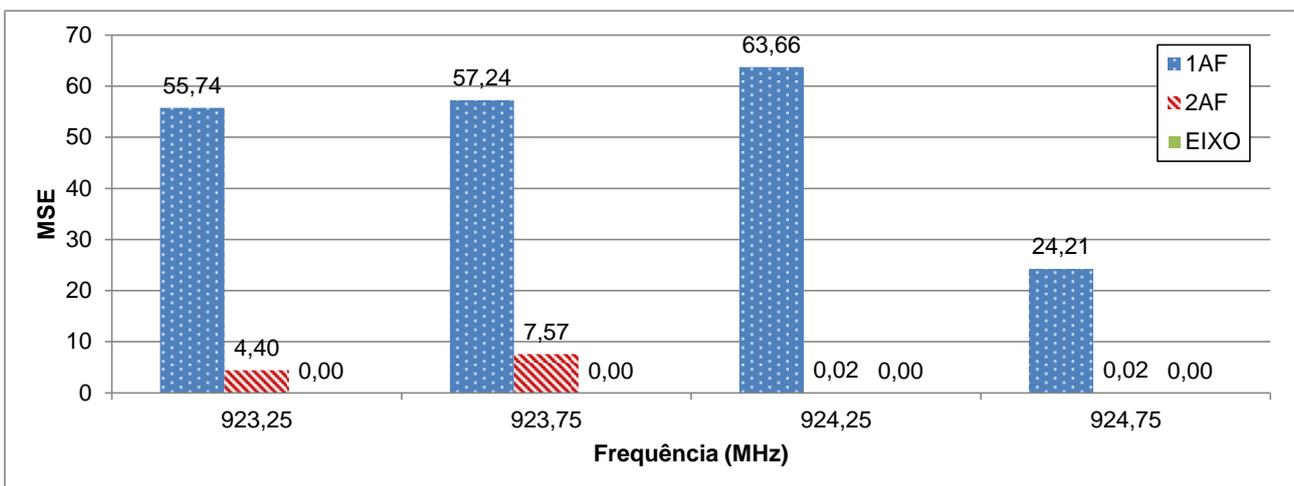


Figura 49 – MSE do treinamento para todas as instâncias de modelo RNA na avaliação do posicionamento das antenas. Fonte: o autor.

Conforme os resultados, existe um desempenho superior na estratégia de posicionamento EIXO, sendo que em todas as frequências o MSE foi zero. Ou seja, na

melhor época do algoritmo de retro-propagação da rede, para todas as entradas possíveis dos dados de treinamento, foi possível estimar com ótima precisão a saída.

Os gráficos da Figura 50 exibem o desempenho em cada época da rede, contemplando os conjuntos de dados de treinamento e validação. Nesta análise, é possível verificar que o algoritmo de retro-propagação ajusta os pesos e bias da rede conforme cada época é analisada. Nos gráficos, o círculo verde indica a época em que as amostras de validação atingiram o menor MSE, sendo esta época escolhida para fornecer os parâmetros utilizados pela rede durante a fase *online*.

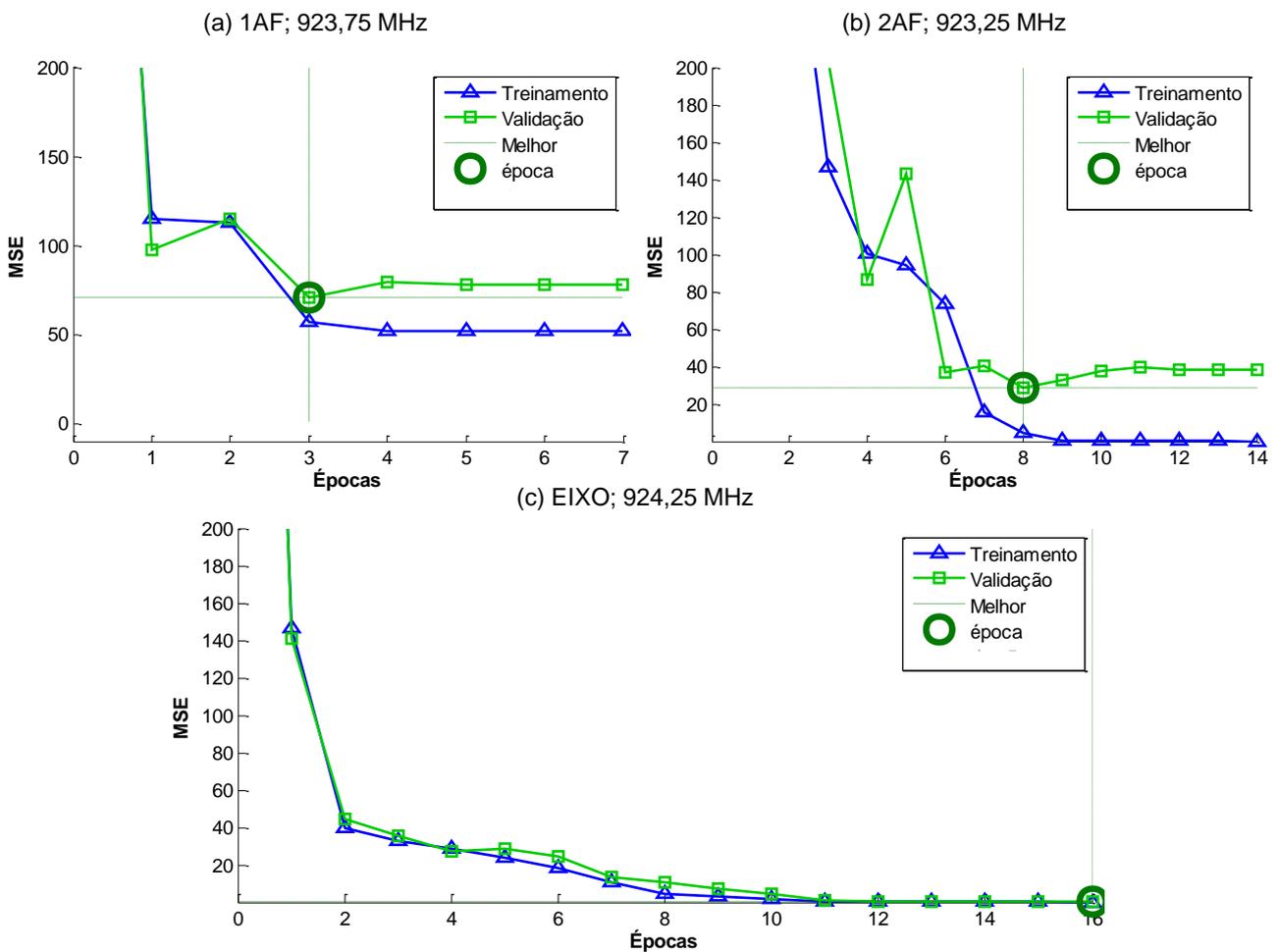


Figura 50 – Desempenho do treinamento das redes neurais (MSE em coordenadas da grade virtual) em relação às épocas do algoritmo de retro-propagação. Fonte: o autor.

No treinamento da rede neural para a frequência 923,75 MHz do formato 1AF (Figura 50a), a melhor época das amostras de validação obteve um MSE de 70,5 (época 3). Após esta época, o treinamento continuou por mais 4 épocas, sendo que o desempenho das amostras de treino melhorou. Entretanto, o desempenho da validação piorou, sendo consequência de um possível sobreajuste nos pesos da rede. Este mesmo comportamento ocorre no treinamento do formato 2AF (Figura 50b), onde o MSE da

validação em sua melhor época (8) foi igual a 28,5 e o MSE do treinamento nesta mesma época foi 4,4. Finalmente, a Figura 50c demonstra o treinamento para o formato EIXO em 924,25 MHz. De acordo com o gráfico, há uma similaridade entre o desempenho das amostras de validação e treino. O MSE atinge o valor zero em ambas as séries, sendo que não ocorre sobreajuste. Após 16 épocas o treinamento é encerrado em função do valor mínimo do gradiente ter sido alcançado (critério de parada *min_grad*).

5.2.2 Resultados da Fase *Online*

A Tabela 9 sintetiza os resultados dos experimentos realizados nesta fase. Os valores foram obtidos através do RMSE (equação 16) das estimativas dos seis marcadores alvo em três repetições executadas. A linha RMSE da tabela sintetiza o erro obtido de forma geral, ou seja, considerando ambos os modos de execução. O formato EIXO obteve desempenho 49% e 59% superior em relação às estratégias que utilizam uma e duas antenas frontais, respectivamente. A abordagem híbrida proposta neste trabalho superou a precisão alcançada no modo “somente RFID”, obtendo desempenho superior na maioria dos formatos.

Tabela 9 – Desempenho da localização (RMSE em cm) para cada formato de posicionamento.

	1AF	2AF	EIXO
Somente RFID	34,4	40,8	20,1
Híbrido	31,0	41,5	12,3
RMSE	32,7	41,1	16,7

A fim de detalhar os resultados percorrendo o erro da estimativa para todas as medições, um gráfico CDF¹ é exibido na Figura 51. Neste gráfico estão acumuladas as medições dos seis marcadores alvo, nas três repetições executadas, tanto do modo “somente RFID” quanto do modo híbrido. No gráfico, o formato 1AF possui poucas estimativas (38%) com precisão ótima (zero) e não possui erros superiores a 60 cm. Por este motivo, o formato 2AF obteve um desempenho geral (RMSE) inferior ao formato 1AF. O formato 2AF chegou a atingir um erro superior a 100 cm em algumas estimativas. O formato EIXO possui mais de 60% das estimativas com precisão zero e o maior erro não ultrapassou 50 cm.

¹ Um gráfico CDF (*Cumulative Distribution Function*) empírico é um instrumento utilizado para avaliar a distribuição dos dados, estimar percentis, e comparar distribuições amostrais distintas. Um gráfico CDF pode exibir cada valor único *versus* o percentual de valores na amostra que é menor ou igual a ele, conectando os pontos com uma linha em forma de degraus [60].

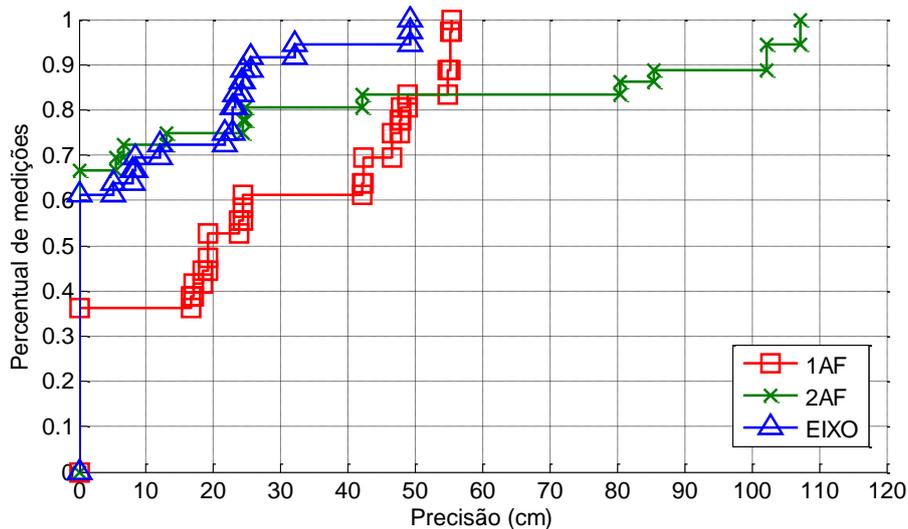


Figura 51 – Erro de localização acumulado (CDF) para os três posicionamentos avaliados.

Fonte: o autor.

Os resultados ruins dos formatos com antenas frontais ocorreram principalmente porque os valores de RSSI de etiquetas que estavam em posições distintas eram muito similares, sendo este também o motivo do baixo desempenho durante a fase de treinamento. Como consequência, não foi possível que o modelo de aprendizado de máquina realizasse uma interpolação razoável do RSSI durante a predição.

Os gráficos da Figura 52 ilustram a média simples dos valores de RSSI coletados durante a fase *offline* para cada estratégia de posicionamento abordado. A fim de preencher toda a superfície do gráfico, uma interpolação do tipo *thin plate spline*¹ foi aplicada. A Tabela 10 exibe a média simples do RSSI coletado para as etiquetas alvo durante a fase *online*.

É possível constatar similaridade nos valores de RSSI entre etiquetas distantes no treinamento do formato 1AF (Figura 52a). Por exemplo, as etiquetas de referência 5, 8, 10, 11 e 13 obtiveram um RSSI de -42,5, -42,5, -43, -42 e -43 dBm, respectivamente. Isso impossibilitou o modelo de predição a fornecer uma localização precisa que correspondesse a qualquer um destes valores de intensidade. No formato 2AF, este comportamento é menos comum durante o treinamento. Entretanto, na interpolação dos valores da fase *online*, as amostras de RSSI das etiquetas alvo não cruzam as medições realizadas pelas etiquetas de referência. Como exemplo, os valores médios de RSSI da etiqueta 15 foram de -56,1 e -46,0 dBm na antena superior e inferior, respectivamente

¹ *Thin-Plate Spline* (TPS) é uma técnica de interpolação de dados baseado em *splines* e possui alto nível de suavização. *Spline* é uma técnica de aproximação que consiste em se dividir o intervalo de interesse em vários subintervalos e interpolar, da forma mais suave possível, nestes subintervalos com polinômios de grau pequeno [29].

(Tabela 10). O valor medido pela antena superior, quando observado no gráfico de treinamento (Figura 52b), indica uma localização mais provável no lado direito do cenário, quando na verdade a etiqueta 15 está posicionada no extremo esquerdo do quadro. Já o valor medido pela antena inferior indica uma localização aproximada na parte de baixo do quadro, enquanto na realidade a etiqueta alvo está posicionada mais próxima ao centro do cenário.

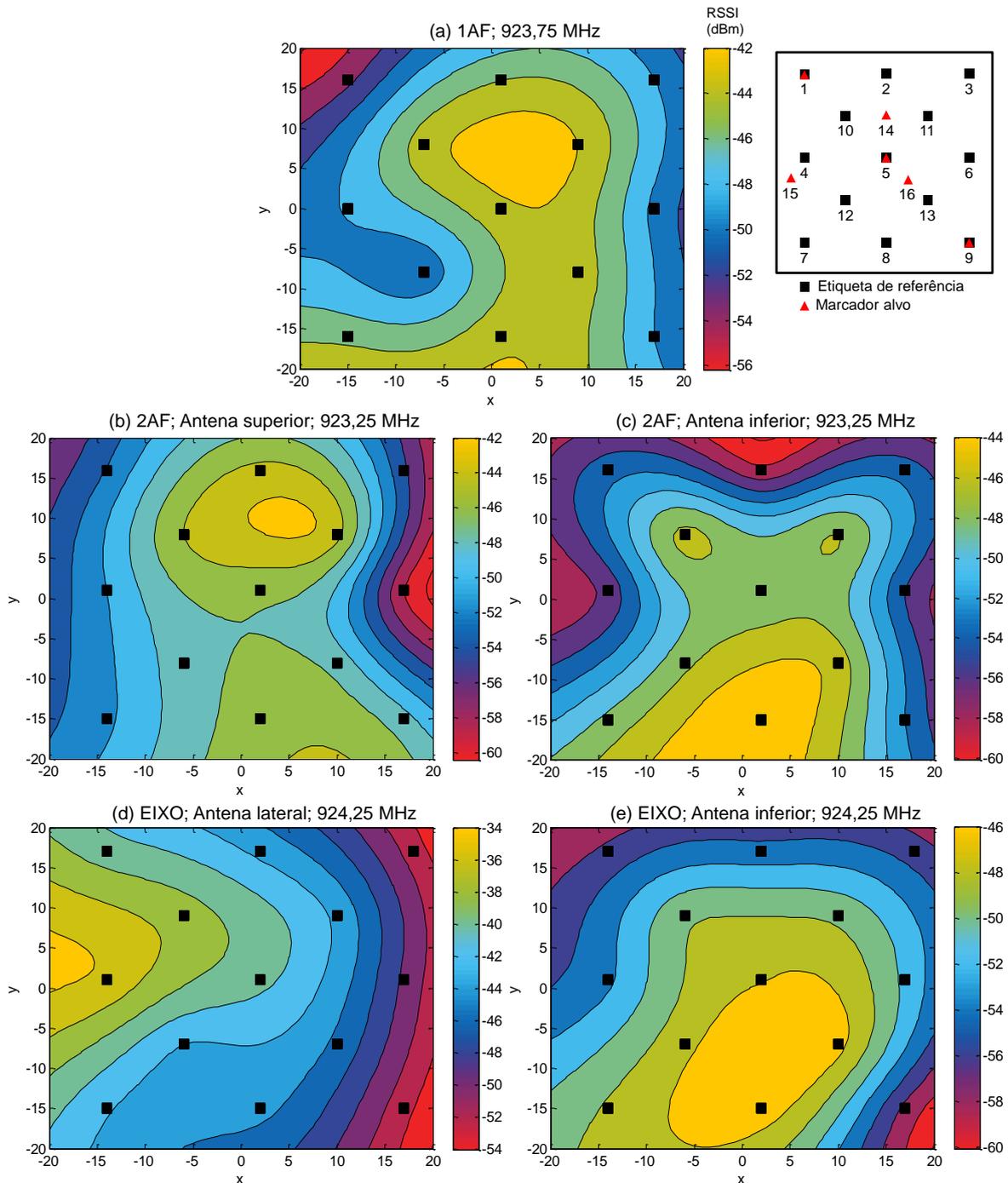


Figura 52 – Níveis de intensidade das etiquetas de referência para cada antena e formato de posicionamento. Os eixos dos gráficos representam as coordenadas na grade virtual.

Fonte: o autor.

Tabela 10 – Média do RSSI das etiquetas alvo em determinadas frequências para cada antena e estratégia de posicionamento.

EPC	1AF; 923,25 MHz	2AF; 923,75 MHz		EIXO; 924,25 MHz	
	Frontal	Superior	Inferior	Lateral	Inferior
1	-52,6	-51,8	-53,0	-39,0	-55,6
5	-43,0	-45,5	-47,7	-39,9	-46,2
9	-48,0	-47,3	-52,0	-51,5	-56,8
14	-43,5	-44,1	-50,8	-41,0	-53,8
15	-45,5	-56,1	-46,0	-36,0	-50,0
16	-44,0	-46,0	-49,3	-48,4	-49,7

Todos os valores estão em dBm.

Os gráficos do formato EIXO demonstram que o RSSI decrementa à medida que a distância em relação à antena aumenta, resultando em maior possibilidade de interpolação dos valores e, conseqüentemente, uma melhor precisão do sistema. Entretanto, apesar do *fingerprint* RSSI mais padronizado e um MSE mais alto na fase de treinamento, isto não significou um resultado ótimo na localização dos marcadores. Por exemplo, analisando os gráficos e os valores médios de RSSI da etiqueta 16, é possível identificar um comportamento semelhante ao descrito no formato 2AF, embora no formato EIXO o erro da distância seja menos acentuado. No caso, o valor medido pela antena inferior não apresenta falhas, aproximando-se da coordenada vertical real. Entretanto, quando observado o gráfico de treinamento da antena lateral, o RSSI de -48,4 dBm se aproxima mais da extrema direita do quadro, sendo que na realidade a etiqueta 16 está mais próxima do centro.

As estratégias de posicionamento com antenas frontais possuem uma manutenibilidade mais prática se comparado ao formato EIXO, isto porque a calibração do sistema RFID é de visada direta às etiquetas e também existe um alinhamento das antenas em relação à câmera. Caso um hardware único que contemplasse todos os componentes sensoriais do sistema pudesse ser vislumbrado, os formatos 1AF e 2AF seriam os mais adequados. Entretanto, a baixa precisão apresentada nos resultados demonstram ser um grande empecilho na adoção do sistema nestes formatos. Por sua vez, o formato EIXO fornece um melhor desempenho na precisão do sistema, embora possua uma calibração mais laborosa. Em um caso de uso que contemple um cenário que atinja uma distância em profundidade (d_z) maior, é certo que mais antenas necessitem ser instaladas, ou que as antenas sejam deslocadas manualmente. Este caso de uso é abordado em mais detalhes na Seção 5.6.

5.3 Organização das Etiquetas de Referência

O objetivo desta fase de testes foi avaliar o desempenho do sistema em relação à forma de organização e quantidade de etiquetas de referência. Neste teste, as antenas foram posicionadas no formato EIXO, cuja configuração obteve melhor resultado nos testes da Seção 5.2. A configuração primária do sistema também permaneceu fixa conforme a Tabela 8 (variáveis A, C, D, E).

Três formatos distintos de organização de etiquetas foram avaliados. A Figura 53a ilustra a primeira configuração (5ER), que contém cinco etiquetas de referência. A segunda configuração (9ER) utiliza nove etiquetas em forma de grade simples (Figura 53b). O último formato (13ER) emprega 13 etiquetas no formato de grade diagonal, conforme exibido na Figura 53c. Os experimentos e resultados do formato 13ER são os mesmos demonstrados na Seção 5.2 para o formato EIXO.

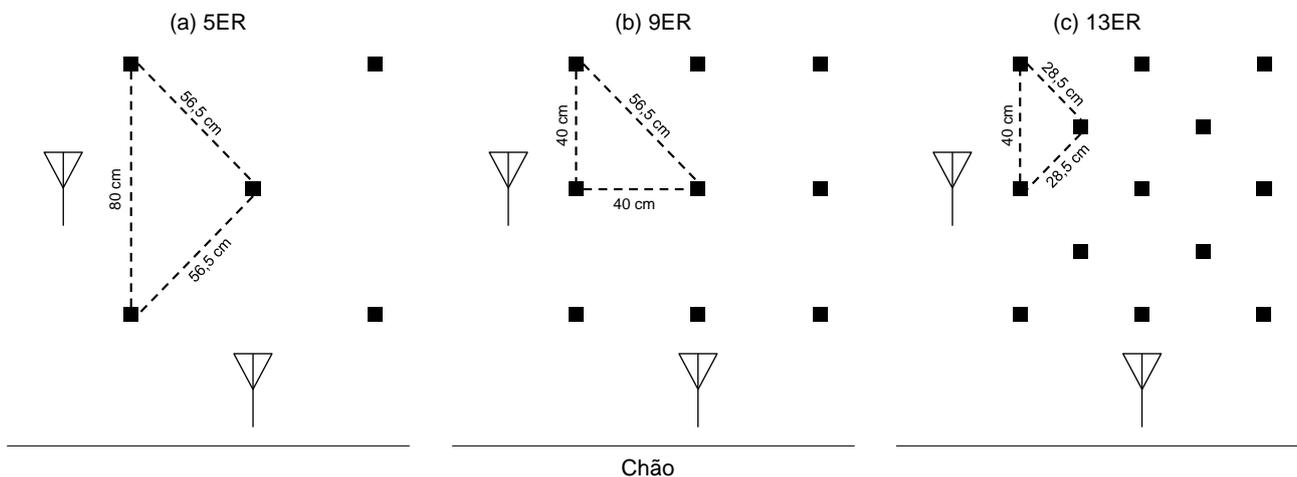


Figura 53 – Formas de organização das etiquetas de referência contendo 5, 9 e 13 elementos, respectivamente. Fonte: o autor.

Na fase *offline*, foi coletado um total de 1120 amostras, sendo 205 para a estratégia 5ER, 369 para o formato 9ER e 546 registros no formato 13ER. Todas as etiquetas de referência foram detectadas em todos os testes. A partir destas coletas, um modelo RNA foi treinado para cada forma de organização.

Na fase *online* foi utilizado o mesmo método de avaliação e as mesmas configurações (tempo de leitura, dimensão do ROI, etc.) descritas nos ensaios da Seção 5.2. Para as três formas de organização avaliadas, o objetivo foi localizar seis marcadores alvo, posicionados conforme ilustrado na Figura 48. Cada experimento foi repetido três vezes.

5.3.1 Resultados

O treinamento do modelo RNA resultou em um MSE de 0,012, 0,006 e 0,000 (média geral de todas as frequências) para os formatos 5ER, 9ER e 13ER, respectivamente. Assim, independente do formato e quantidade de etiquetas de referência utilizadas, cada rede neural respectiva conseguiu ajustar os pesos a fim de estimar os pontos amostrados no treinamento com ótima precisão.

A Tabela 11 exibe o RMSE do erro de localização em cada estratégia abordada. A configuração que emprega 13 etiquetas de referência obteve uma precisão 30,7% superior ao formato que emprega nove etiquetas. O formato 9ER superou em 38,6% a estratégia que utiliza somente cinco etiquetas de referência. Em todos os formatos, o sistema híbrido prevaleceu sobre a execução apenas do subsistema RFID. Na configuração 5ER, o sistema híbrido incrementou o resultado em somente 1 cm. Neste caso, a predição do subsistema RFID foi tão distante da posição real que nenhum dos ROIs estimados contemplou uma área em que o subsistema visual pudesse atuar.

Tabela 11 – Desempenho da localização (RMSE em cm) para cada estratégia de organização das etiquetas de referência.

	5ER	9ER	13ER
Somente RFID	39,8	29,5	20,1
Híbrido	38,9	17,0	12,3
RMSE	39,3	24,1	16,7

O gráfico CDF da Figura 54 exibe uma análise detalhada da precisão para todas as medições. Os formatos 5ER e 9ER possuem precisão ótima em metade dos experimentos, sendo que este percentual representa os testes onde foram utilizados marcadores alvo em posições já utilizadas na fase *offline*. Em todas as outras medições, a estimativa das posições de marcadores em posições desconhecidas não obteve uma interpolação razoável para as coordenadas dos marcadores, chegando a erros superiores a 80 cm no formato 5ER.

Os gráficos da Figura 55 auxiliam na visualização do comportamento da intensidade do sinal refletido pelas etiquetas de referência em cada uma das configurações. É notável que uma maior distribuição de etiquetas aumenta a possibilidade de cobertura do cenário para ambas as direções (horizontal e vertical).

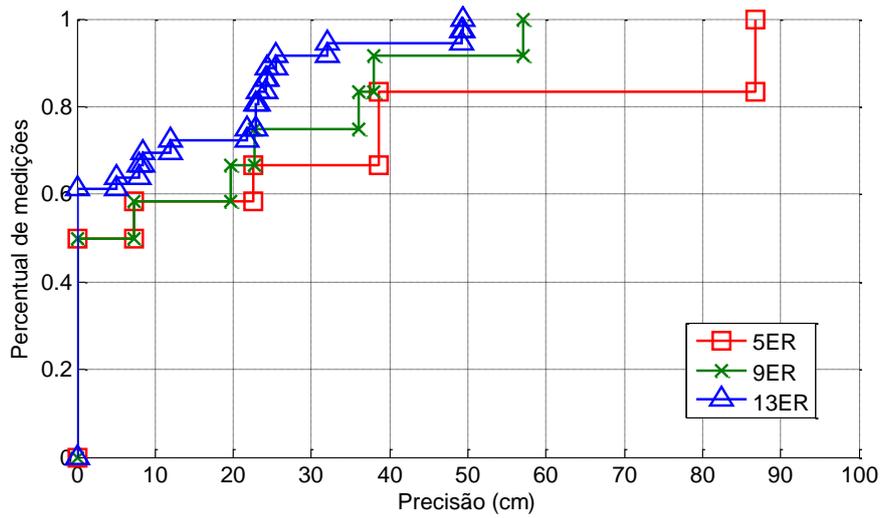


Figura 54 - Erro de localização acumulado (CDF) das três formas de organização de etiquetas. Fonte: o autor.

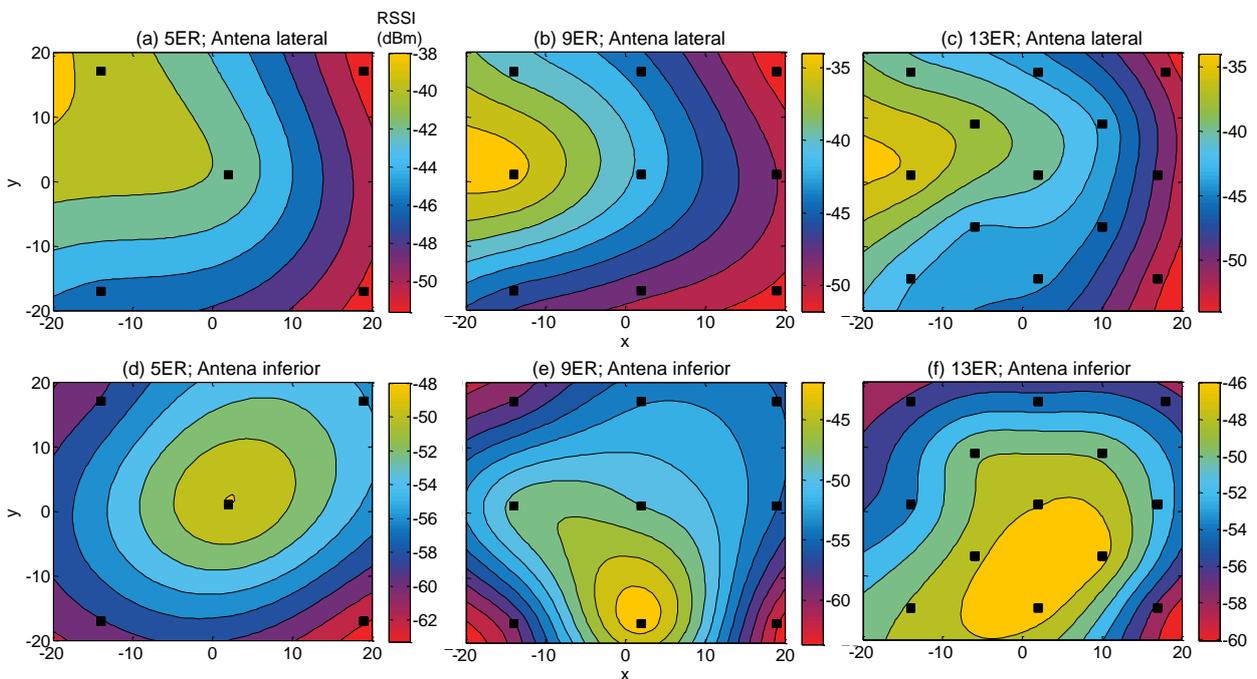


Figura 55 – *Fingerprint* RSSI das etiquetas de referência para cada antena e estratégia de organização. Frequência: 924,25 MHz. Fonte: o autor.

Em relação à baixa precisão do formato 5ER, é possível visualizar que o RSSI resultante da interpolação simples na posição (0, -20) do gráfico (d) não se aproxima do valor medido nos gráficos (e) e (f). O gráfico (d) mostra que o sistema não possui informações suficientes sobre a parte inferior e superior do centro do cenário a fim de estimar etiquetas próximas desta região. Entre a configuração 9ER e 13ER, apenas quatro etiquetas de referência foram adicionadas, sendo que a diferença do RMSE no sistema híbrido foi de somente 5 cm. As medições de RSSI nas posições extras do formato 13ER auxiliaram principalmente na localização mais apurada das etiquetas alvo

14 e 16, pois estas se encontram exatamente em uma interpolação das etiquetas adicionadas.

Apesar do formato 5ER possuir vantagens, tais como um custo menor dos componentes na implantação e manutenibilidade facilitada do sistema, esta configuração demonstra resultados que ficam muito abaixo do esperado para um IPS. O formato 9ER obteve resultados intermediários, sendo que o formato de disposição das etiquetas em grade simples pode ser um facilitador para alguns tipos de ambiente. Finalmente, o formato 13ER ampliou a cobertura dos sinais medidos, conseqüentemente tendo um resultado melhor. Este formato pode possuir uma manutenção um pouco mais elevada e um tempo de configuração maior, entretanto traz benefícios no que diz respeito à precisão do IPS.

5.4 Tempo de Leitura das Etiquetas

Esta fase de testes tem como objetivo avaliar o desempenho do sistema em relação ao tempo de ativação do leitor RFID para realizar a leitura das etiquetas, tanto na fase *offline* quanto na fase *online*. Em outras palavras, estes experimentos verificam se o tempo de coleta das amostras de RSSI das fases de treinamento e estimativa pode ser uma variável que afete a precisão do sistema. Assim, é possível observar se uma quantidade distinta de amostras, coletadas em durações diferenciadas, pode acrescentar variações relevantes e que afetem o modelo probabilístico. Da mesma forma, na fase *online*, os testes validam se um maior número de registros, apresentados ao modelo estatístico, ao agrupamento k-means e ao subsistema visual, pode afetar o desempenho geral do IPS.

Para a fase de treinamento, foram avaliados períodos de coleta de 10 (T10), 30 (T30), 60 (T60) e 120 (T120) segundos. Na avaliação da fase de estimativas (*online*), a coleta ocorreu por 3 (E3) e 30 (E30) segundos. Os experimentos e resultados dos testes T10 e E3 são os mesmos realizados na Seção 5.2 no formato EIXO. Isto implica que as configurações do sistema e o método de avaliação aplicados aos demais experimentos também foram os mesmos utilizados naquela fase.

5.4.1 Resultados

A Tabela 12 sumariza os resultados dos experimentos desta fase. A quantidade de amostras coletadas entre os treinamentos avaliados tem um crescimento significativo à medida que o tempo de leitura aumenta. Nos testes E3 e E30 foi considerada a média de amostras dos experimentos (seis marcadores alvo; três repetições de experimentos).

Tabela 12 – Quantidade de amostras, desempenho do treinamento (MSE) e da localização (RMSE em cm) para a avaliação do tempo de leitura das etiquetas nas fases *offline* e *online*.

	T10	T30	T60	T120	E3	E30
Quantidade de amostras ¹	546	1670	3304	6620	47	424
MSE do treinamento ²	0,0	0,4	2,2	5,0	-	-
RMSE da estimativa de localização (cm)	12,3	18,7	16,1	20,3	12,3	12,0

¹ E3 e E30: média simples entre todos os experimentos.

² Média geral do MSE dos modelos de todas as frequências.

Unidade de medida: coordenadas da grade virtual.

Os valores de MSE apresentados na tabela são uma média do MSE de todas as frequências. O gráfico da Figura 56 exibe o MSE de cada modelo e teste individualmente. No resultados apresentados, o desempenho do treinamento no modelo probabilístico sofre uma degradação à medida que o espaço amostral cresce. Nos testes T60 e T120, principalmente, o processamento da rede neural não conseguiu ajustar com êxito os pesos a fim de cobrir todos os valores de entrada (RSSI) amostrados, relacionando-os às respectivas saídas (coordenadas alvo). Devido ao MSE mais alto no teste T120, também foi realizada uma análise das amostras coletadas nas frequências 923,25 e 924,75 MHz (pior caso), a fim de verificar se existia uma possível variação na intensidade do sinal. Entretanto, o desvio padrão da média dos valores de RSSI se manteve entre 0,00 e 0,28 dBm para ambas as frequências, respectivamente.

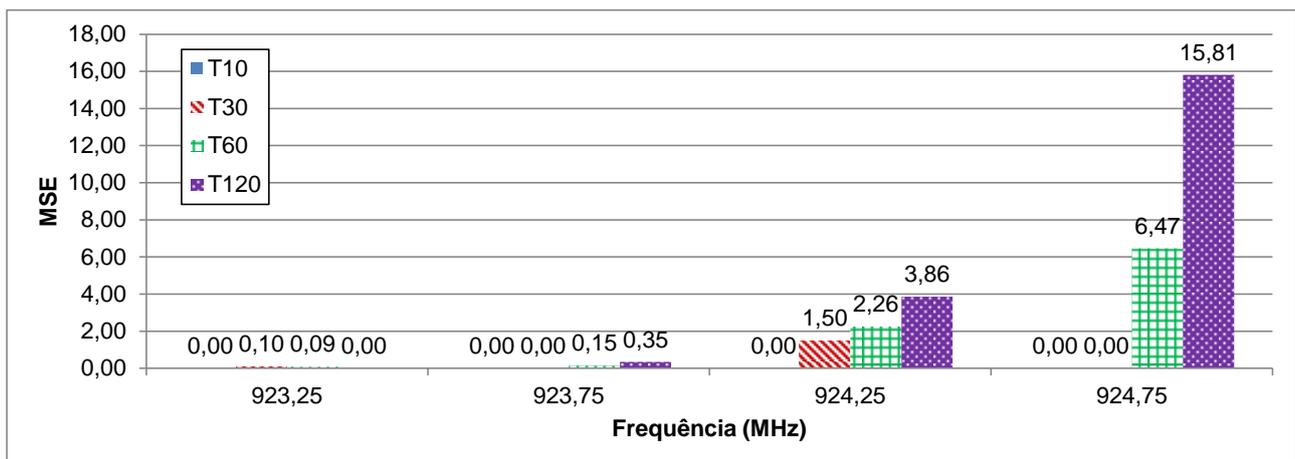


Figura 56 - MSE do treinamento para todas as instâncias de modelo RNA nos experimentos de tempo de leitura das etiquetas. Fonte: o autor.

Em relação à precisão do IPS frente aos diferentes tempos de coleta da fase *offline*, a Tabela 12 exibe o RMSE das estimativas de localização na execução do sistema

híbrido. Os erros apresentados possuem uma diferença de 8 cm entre o melhor (T10) e o pior (T120) teste. Nestes resultados, o erro não cresce exatamente à medida que o tempo de coleta aumenta. Essa variação também não segue o mesmo padrão apresentado no MSE de treinamento. No caso, o RMSE do teste T60 supera em 2 cm o teste T30.

Apesar do MSE de treinamento possuir resultados inferiores para algumas frequências, a abordagem híbrida proposta foi capaz de encontrar a posição dos marcadores alvo de maneira significativa. Por exemplo, o MSE dos testes T30 e T60 em 924,75 MHz foi de 0,00 e 6,47, respectivamente. Entretanto, devido às operações executadas pelos métodos multi-frequência e k-means, determinadas predições do modelo estatístico para esta frequência foram ignoradas, possivelmente por estarem em coordenadas distantes das demais estimativas.

No que diz respeito aos testes E3 e E30, apesar do número de amostras da coleta de 30 segundos ser nove vezes maior, não houve melhora significativa no erro de localização, reduzindo-o em somente 0,3 cm. As amostras coletadas para cada marcador alvo de ambos os testes resultaram em um RSSI praticamente constante, com variações pequenas (até 1,5 dBm) durante todo o período de leitura. Portanto, como os valores de entrada foram praticamente os mesmos, e ambas configurações empregaram o mesmo modelo de treinamento, a precisão praticamente não mudou entre um teste e outro. Ainda, no caso da fase *online*, o ônus de aguardar por mais tempo para receber uma resposta pode ser um empecilho na aceitação deste tipo de sistema. Desta forma, optou-se por definir a configuração primária do tempo de leitura na fase *online* em 3 segundos.

5.5 Multi-Frequência

Nesta seção é avaliado o método de multi-frequência proposto na Seção 4.2.1. Para isto, o sistema foi configurado de duas formas distintas:

- Modelo único (MU): o método de multi-frequência não é utilizado, sendo que todos os dados coletados pelo leitor alimentam uma única instância de modelo, ignorando a frequência de operação;
- Multi-frequência (MF): os dados são particionados por frequência, criando e alimentando uma instância de modelo para cada partição.

Os ensaios descritos nesta seção utilizam as configurações primárias definidas de acordo com os resultados das seções anteriores. A Tabela 13 exhibe as propriedades das configurações primárias empregadas neste e nos próximos experimentos do sistema.

Tabela 13 – Propriedades das configurações primárias definidas após série de experimentos.

Propriedade	Valor
A. Tipo de marcador visual	não-codificado
B. Posicionamento das antenas	EIXO
C. Organização das etiquetas de referência	13ER
D. Tempo de leitura das etiquetas de referência	10 segundos
E. Tempo de leitura das etiquetas na fase <i>online</i>	3 segundos

Os experimentos do teste MF reutilizaram os resultados do teste EIXO da Seção 5.2. Durante a fase *offline*, o teste MU empregou o mesmo conjunto de dados coletado na fase de treinamento do teste MF. Assim, um total de 546 amostras, não particionadas, alimentaram um único modelo RNA. O MSE de treinamento do teste MU foi de 0,17, enquanto o MSE médio entre todas as frequências do teste MF foi de 0,00.

O sistema foi executado nos modos “somente RFID” e híbrido, sendo cada experimento repetido três vezes. As configurações de cenário (distância entre a câmera e quadro, por exemplo) e o método de avaliação aplicados ao teste MU seguiram o mesmo padrão do teste MF, tendo como objetivo a localização de seis marcadores alvo, distribuídos conforme a Figura 48. As amostras de RSSI utilizadas durante a fase *online* do teste MU também foram as mesmas do teste MF.

5.5.1 Resultados

A Tabela 14 exhibe os resultados desta fase de testes. O método de multi-frequência obteve uma precisão superior à configuração MU, reduzindo o erro em até 15,1 cm. De maneira geral, a precisão do sistema foi otimizada em 39% a partir da aplicação do método desenvolvido.

Tabela 14 – Desempenho da localização (RMSE em cm) na avaliação do método de multi-frequência desenvolvido.

	MU	MF
Somente RFID	27,4	20,1
Híbrido	27,4	12,3
RMSE	27,4	16,7

A fim de analisar os resultados, a Figura 57 exhibe capturas de tela do sistema no modo “somente RFID”. Nas imagens, os quadrados amarelos representam as posições

auxiliares (resultantes somente do modelo RNA) e a elipse verde indica a posição final do marcador alvo (15). Cabe ressaltar que podem existir sobreposições de símbolos amarelos, ou seja, diversos símbolos na mesma posição. Ainda, apesar das telas estarem exibindo os marcadores visuais, eles não são utilizados pelo sistema no modo em que foi executado (“somente RFID”), servindo neste caso somente de referência para a leitura dos resultados.

Em relação ao teste MU (Figura 57a), as estimativas das posições auxiliares são próximas umas das outras, sendo que nenhuma se aproxima da posição real do marcador alvo. A causa deste comportamento é o baixo desempenho da interpolação fornecida pelo modelo único (RNA). Por exemplo, o par de RSSIs das antenas lateral e inferior (-36 e -50,5 dBm), quando apresentados como entrada ao modelo único, resultou nas coordenadas de saída (-11, 8). Este mesmo par de RSSIs, quando apresentado ao modelo da sua respectiva frequência no teste MF, resultou nas coordenadas (-14, -6), posição bastante próxima ao marcador real.

Como consequência do acúmulo de posições auxiliares próximas umas das outras no teste MU, nenhum dos ROIs explorados no modo híbrido contemplava um marcador visual válido, ou seja, em todos os casos a posição final estimada foi aquela fornecida pelo subsistema RFID. Assim, o RMSE do modo híbrido foi o mesmo do modo “somente RFID”.

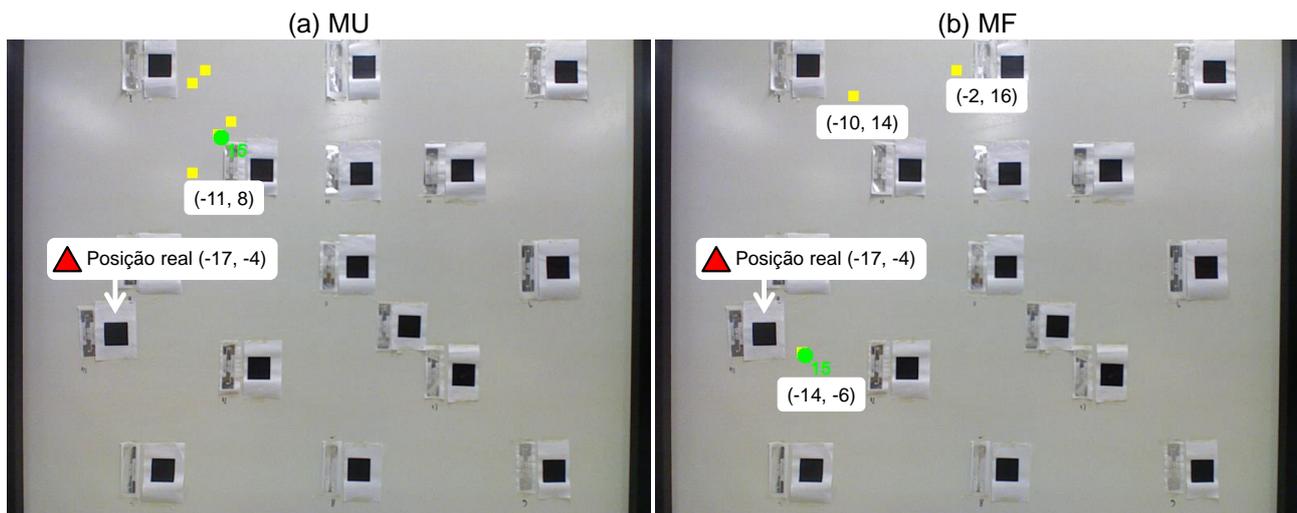


Figura 57 – Capturas de tela do sistema após tentativa de localização nos testes MU (a) e MF (b). Modo: “somente RFID”; modelo: RNA; marcador alvo: 15; precisão: 43 cm (a) e 8 cm (b). Fonte: o autor.

Conforme a Figura 57b, os modelos RNA do teste MF estimaram posições mais distantes entre si. Por exemplo, a posição auxiliar (-2, 16) foi uma das saídas fornecidas pela instância da rede neural na frequência 924,25 MHz, enquanto a posição (-14, 6) foi indicada pela instância da frequência 923,25 MHz. Assim, a posição final do marcador alvo foi estimada na posição centroide do agrupamento k-means que obteve o maior peso, ou seja, aquele que possuía o maior número de estimativas fornecidas pelos modelos RNA. Este comportamento também auxiliou na otimização dos resultados do subsistema visual, ampliando o escopo de busca por marcadores visuais.

5.6 Avaliação Geral do Sistema Híbrido em Escalabilidade 2D

O objetivo desta fase de experimentos foi avaliar o sistema de modo geral, abrangendo os dois modelos de aprendizado propostos (RNA e SVR), a cobertura do sistema na distância em profundidade, e a validação frente a cenários mais complexos. As configurações primárias do sistema utilizaram as mesmas propriedades constantes na Tabela 13.

O sistema foi avaliado em quatro cenários: C1, C2, C3 e C4. Em cada cenário, a distância d_z entre a câmera e o quadro era de 100, 140, 180 e 220 cm, respectivamente. Conforme descrito nos resultados dos testes sobre posicionamento das antenas (Seção 5.2.2), uma das desvantagens do formato EIXO reside no fato da cobertura do sistema no eixo z (profundidade) necessitar o deslocamento ou a adição de mais antenas nos eixos x e y. A principal razão dessa limitação está no fato do sistema RFID não detectar parte das etiquetas passivas presentes quando estas etiquetas não estão alinhadas às antenas. Como a infraestrutura dos ensaios contava com um número limitado de antenas, optou-se por deslocar as duas antenas presentes em cada cenário avaliado.

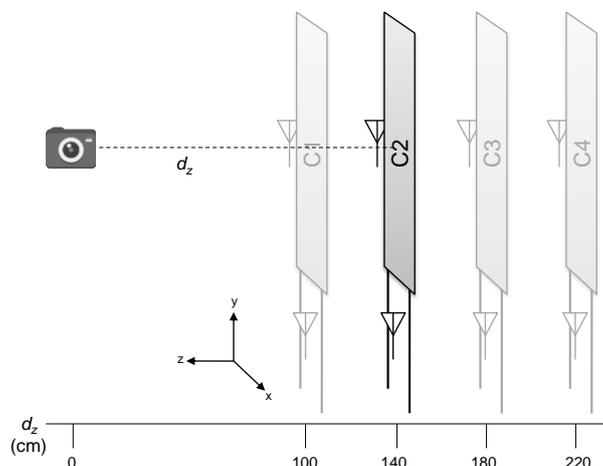


Figura 58 - Cenários avaliados e o deslocamento utilizado no treinamento e nas estimativas do sistema. Fonte: o autor.

Na ilustração de exemplo (Figura 58), o experimento corrente visa o cenário C2. Em seguida, para treinar e executar o IPS no cenário C3, o quadro com os marcadores e as antenas eram deslocados.

Na fase *offline*, foi coletada uma média de 536 amostras entre os quatro cenários avaliados. Em todos os cenários, houveram leituras em todas as etiquetas presentes. O conjunto de dados destes experimentos foi obtido a partir de uma nova rodada de ensaios, ou seja, não foram reaproveitados registros ou resultados de experimentos anteriores. Uma instância de modelo RNA e outra instância de modelo SVR foi criada para cada cenário (4) e frequência (4), totalizando um total de 32 instâncias. Os modelos SVR utilizaram os mesmos dados de treinamento das redes neurais.

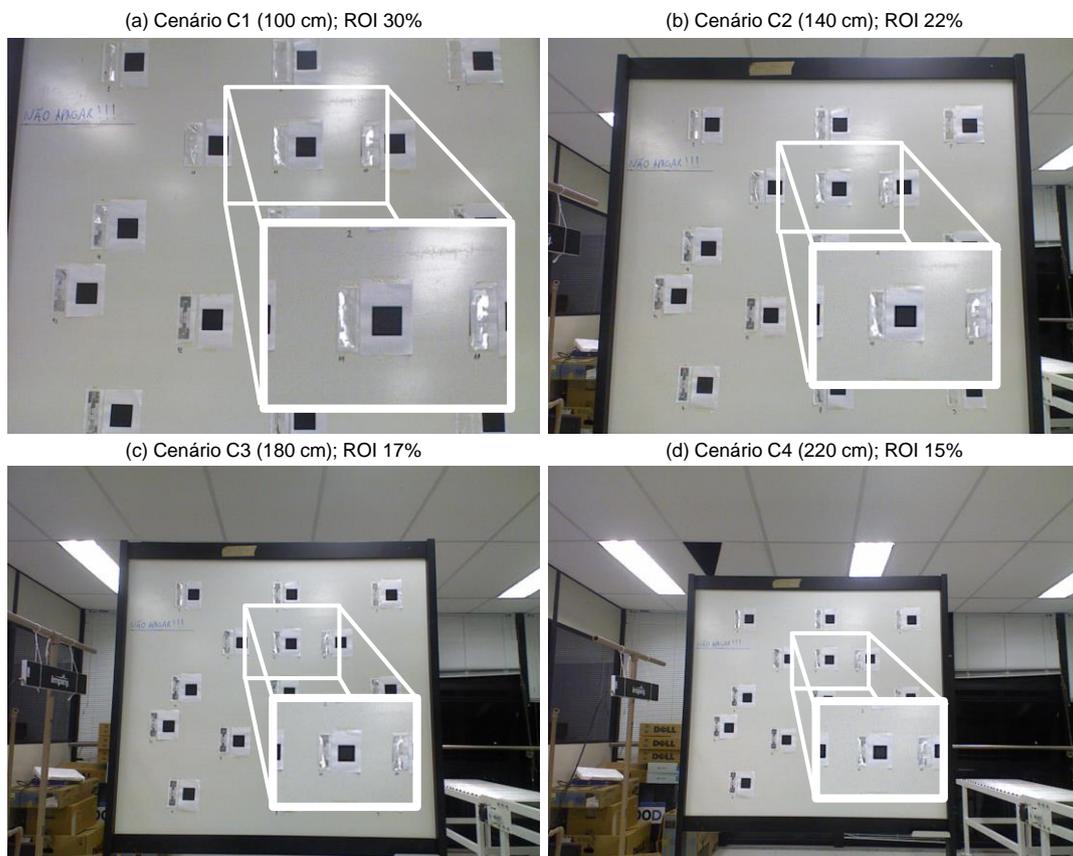


Figura 59 - Capturas de tela dos quatro cenários e amostras das dimensões do ROI.

Fonte: o autor.

Na fase *online*, o objetivo dos experimentos foi localizar seis marcadores alvo distribuídos no ambiente (Figura 48). Três testes de validação foram realizados. O primeiro era “somente RFID”, ou seja, o subsistema visual não foi utilizado. No teste “denso”, além dos seis marcadores alvo, outros dez marcadores visuais estavam afixados no quadro branco. Estes marcadores visuais adicionais não eram alvo de localização, e estavam presentes somente para simular um ambiente mais complexo. Neste teste, a

câmera capturava a imagem com todos os 16 marcadores visuais simultaneamente (Figura 62). No último teste, denominado “limpo”, somente um marcador visual estava presente no ambiente em cada execução do sistema híbrido. Na fase *online*, a execução do modelo SVR também utilizou as mesmas amostras informadas ao modelo RNA.

A dimensão do ROI (subimagem) foi configurada em 30, 22, 17 e 15% para os cenários C1, C2, C3 e C4, respectivamente. Com estes percentuais, a região da imagem extraída permaneceu em todos os cenários com uma dimensão aproximada de 35 x 27 cm (l x a) no ambiente real (Figura 59).

5.6.1 Resultados

Como resultado da fase *offline*, os valores de MSE das redes neurais modeladas para cada cenário e frequência são apresentados no gráfico da Figura 62. A média do MSE em cada cenário foi de 0,01, 0,01, 0,61 e 0,35 para os cenários C1, C2, C3 e C4, respectivamente. O treinamento dos modelos RNA foi concluído em menos de 30 segundos, enquanto o treinamento dos modelos SVR demorou aproximadamente 5 minutos. Infelizmente, a ferramenta e biblioteca utilizadas no modelo SVR não permitiam uma análise prática da fase de treinamento, sendo que somente os resultados da fase *online* deste modelo são apresentados.

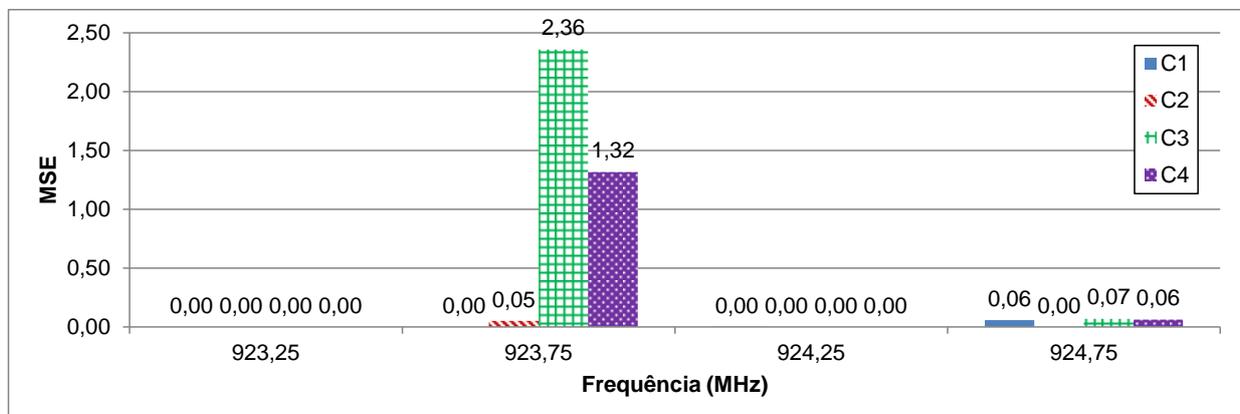


Figura 60 - MSE do treinamento para as instâncias RNA na avaliação geral do sistema híbrido. Fonte: o autor.

Conforme os resultados, a maior parte dos erros das amostras de treinamento foi zero, entretanto existe um desempenho inferior nos cenários C3 e C4 em algumas frequências. Por exemplo, no caso da frequência 923,75 MHz do cenário C3, as amostras de validação utilizadas somente obtiveram um MSE baixo até a época três do algoritmo de retro-propagação (Figura 61). Após esta época, o treinamento continuou por mais seis épocas, mas o erro da validação somente cresceu, caracterizando um sobreajuste no

treinamento. Assim, esta rede foi definida a partir dos pesos constantes na época três do algoritmo.

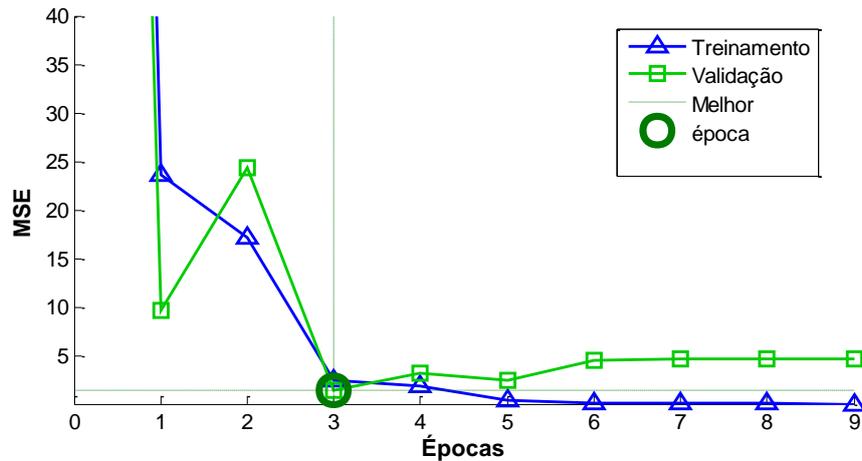


Figura 61 - Desempenho do treinamento da rede neural em relação às épocas do algoritmo de retro-propagação. Cenário: C3; frequência: 923,75 MHz. Fonte: o autor.

A partir dos modelos treinados para cada cenário, os testes de validação “somente RFID”, “denso” e “limpo” foram executados na fase *online*. A Figura 62 apresenta uma captura de tela do sistema exibindo as posições inferidas pelo subsistema RFID (quadrados amarelos) e a posição final do alvo (círculo verde). A interface gráfica auxilia os usuários a identificarem facilmente a localização do marcador. O tempo de processamento para localização de cada marcador é menos de 1 segundo, excluindo o tempo da coleta RFID.

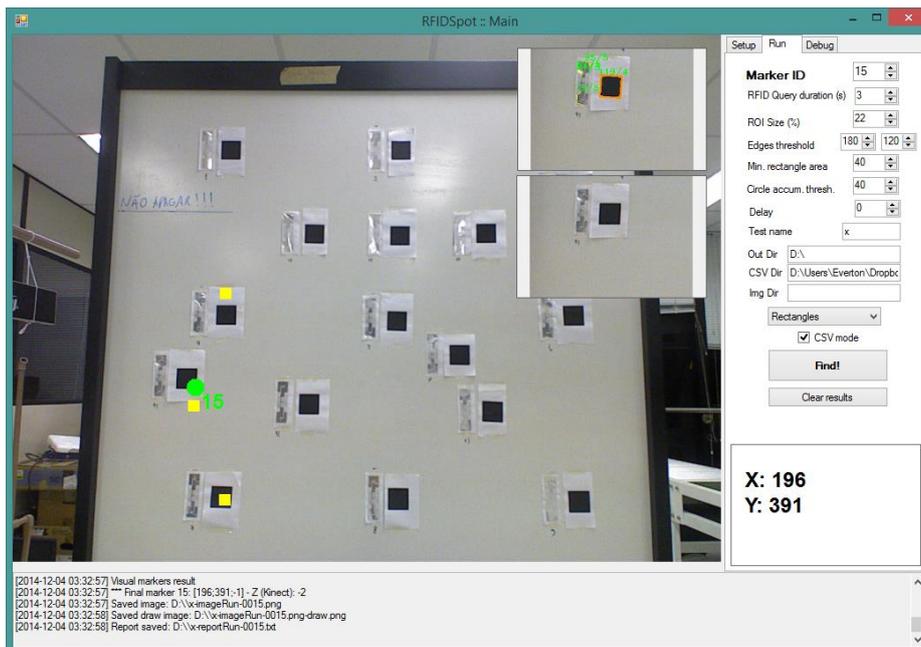


Figura 62 – Captura de tela do sistema em execução. Cenário: C2; teste: denso; modelo: RNA; marcador alvo: 15; precisão: 0 cm. Fonte: o autor.

A Tabela 15 resume o desempenho do sistema frente aos cenários, modelos e testes avaliados. Na comparação entre as abordagens RNA e SVR, o modelo RNA obteve desempenho 30% superior ao modelo SVR. Os resultados do modelo RNA na abordagem híbrida exibem um erro entre 8 e 24 cm. O modelo RNA foi inferior ao SVR apenas no modo “somente RFID” do cenário C3. Em uma comparação visando o teste “limpo”, o modelo RNA demonstrou uma precisão 1,7 vezes melhor em relação ao SVR.

Tabela 15 – Desempenho da localização (RMSE em cm) para cada cenário, teste de validação e abordagem de aprendizado de máquina.

Cenário	RNA			SVR		
	Somente RFID	Híbrido		Somente RFID	Híbrido	
		Denso	Limpo		Denso	Limpo
C1 (100 cm)	12,1	11,8	8,9	17,6	28,2	16,5
C2 (140 cm)	17,3	13,6	9,1	18,2	14,1	19,7
C3 (180 cm)	33,3	23,7	22,9	30,2	26,8	28,4
C4 (220 cm)	12,1	13,0	10,0	30,9	29,3	26,5
RMSE	20,6	16,2	14,0	25,0	25,4	23,3

Analisando os resultados do modelo RNA, a abordagem híbrida desenvolvida obteve desempenho superior ao subsistema RFID na maior parte dos casos. O modo híbrido resultou em uma precisão 21% e 32% superior ao subsistema RFID nos testes “denso” e “limpo”, respectivamente. Isto demonstrou a eficiência da melhoria trazida pela integração do subsistema visual, mesmo utilizando marcadores visuais simples em uma infraestrutura *off-the-shelf*.

O único resultado em que o modo “somente RFID” obteve uma precisão melhor foi no teste “denso” do cenário C4 (aproximadamente 1 cm). Neste caso, todas as estimativas do subsistema RFID resultaram em posições muito próximas a outro marcador visual presente no cenário (Figura 63). No modo híbrido, estas posições auxiliares forneceram regiões de interesse na área do marcador errado, sendo detectado como possível marcador alvo. Como o modo “somente RFID” não detecta marcadores visuais, e suas estimativas estavam mais próximas do marcador alvo correto, o erro de localização nesse modo foi menor.

Os resultados dos gráficos CDF (Figura 64) demonstram que o erro de localização é zero na maioria dos experimentos. No modelo RNA, 80% dos testes “limpos” exibem uma precisão ótima (0 cm) e os erros de ambos testes híbridos não excederam 40 cm.

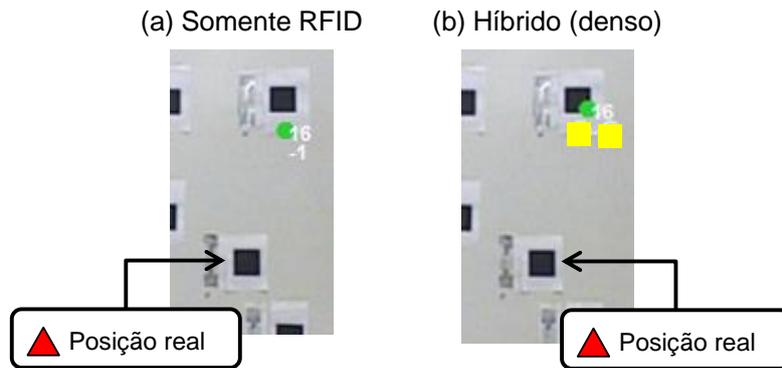


Figura 63 – Detalhe das estimativas de localização do subsistema RFID (a) e do teste “denso” (b). Cenário: C4; modelo: RNA; marcador: 16; precisão: 24 cm (a) e 29 cm (b).

Fonte: o autor.

O sistema híbrido não detectou o marcador visual em 20% dos testes limpos. Nestes casos, a localização do subsistema RFID foi utilizada como posição final. Isto ocorreu porque as imagens dos ROIs fornecidos ao subsistema visual de fato não possuíam nenhum marcador visual, ou seja, foram estimados em posições distantes ao alvo. Em todos os casos que o ROI possuía um marcador visual, o algoritmo de visão computacional o detectou com sucesso.

No teste “denso”, o sistema detectou o marcador visual errado em 25% dos experimentos, a maioria deles nos cenários mais distantes da câmera (C3 e C4). Os resultados do modelo SVR também foram melhores no teste “limpo”, enquanto os testes “somente RFID” e “denso” obtiveram resultados similares. Diferente do modelo RNA, o teste “denso” demonstrou o pior desempenho, ou seja, o modelo SVR inferiu mais posições próximas de marcadores visuais errados.

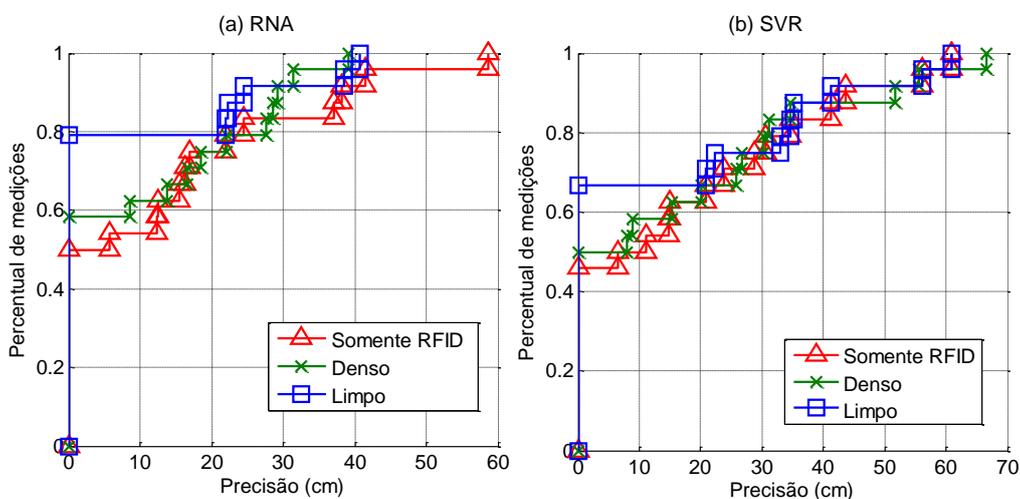


Figura 64 - Erro de localização acumulado (CDF) para ambos os modelos nos três testes de validação. Escalabilidade: 2D. Fonte: o autor.

O gráfico CDF da Figura 65 exibe os resultados de todos os experimentos individualmente por marcador alvo. Entre os marcadores alvo que foram posicionados em locais já utilizados na fase de treinamento (1, 5 e 9), somente o marcador 5 não obteve 100% de probabilidade de sucesso. O marcador 15 alcançou uma precisão inferior a 20 cm em mais de 80% das medições. A interpolação da posição deste marcador era um desafio, pois estava fora dos limites da grade de marcadores usada no treinamento, além de existir outro marcador visual próximo a ele no teste “denso”.

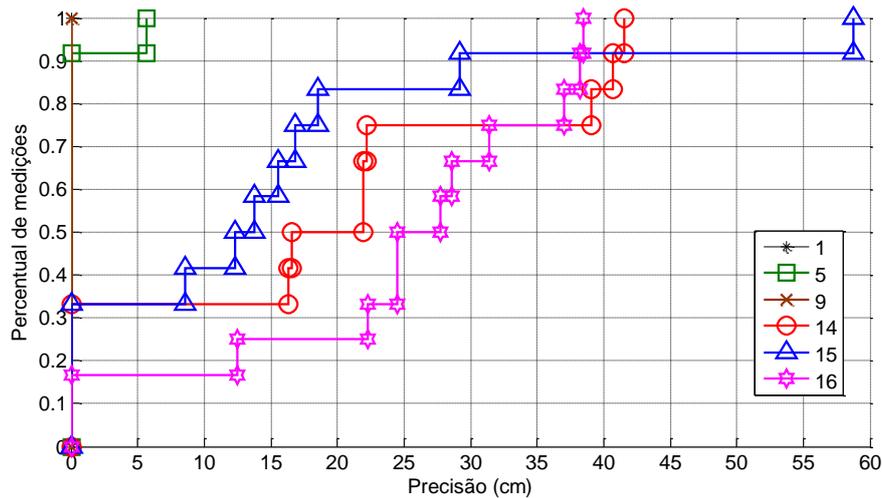


Figura 65 – Erro acumulado (CDF) na localização individual de cada marcador alvo.

Medições contemplam todos os cenários, modelos e testes de validação. Fonte: o autor.

Em contrapartida, o marcador 15 também obteve os piores erros, atingindo quase 60 cm. Estes erros (menos de 10%) ocorreram nos cenários C3 e C4, principalmente no modelo SVR e no modo “somente RFID”. Analisando o processo de localização, os agrupamentos k-means do marcador 15 no cenário C3 possuíam os seguintes centroides e pesos: (i) posição (-1, 5), peso 11; (ii) posição (-11, -5), peso 5; (iii) posição (2, -7), peso 6. A maior parte das estimativas auxiliares que estavam no agrupamento de maior peso tiveram como origem as leituras na frequência 923,75 MHz. O modelo desta frequência foi aquele que obteve o MSE mais alto durante a fase *offline* (Figura 60). Analisando as amostras coletadas nesta frequência, o RSSI médio da etiqueta 15 apresentou os valores -35 e -51,5 dBm para as antenas lateral e inferior, respectivamente. O *fingerprint* dos RSSIs no treinamento (Figura 66) demonstra que o nível de intensidade refletido pela etiqueta 15 condiz com a região do marcador alvo, que seria a extrema esquerda do centro do quadro. Assim, o baixo desempenho somente se explica em virtude dos pesos e parâmetros dos modelos de aprendizado de máquina não terem sido ajustados por completo, ou seja, durante o treinamento nem todas as variações de posicionamento das etiquetas puderam ser explicados pela relação entre o RSSI e as coordenadas.

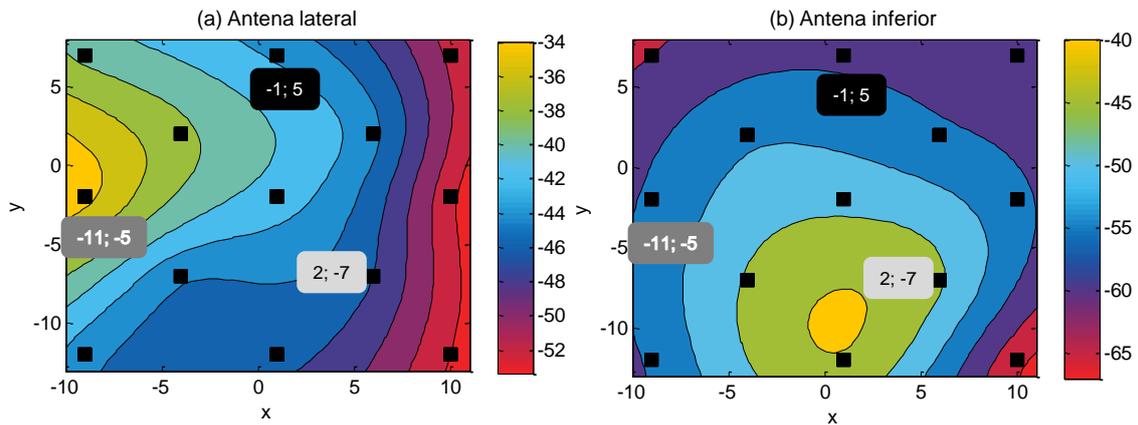


Figura 66 – *Fingerprint* RSSI das etiquetas de referência em ambas as antenas. A posição centroide de cada *cluster* k-means é exibida nas caixas de texto (etiqueta 15). *Clusters* com maior peso possuem cores escuras. Freq.: 923,75 MHz; cenário: C3. Fonte: o autor.

Em uma comparação dos resultados com os trabalhos correlatos que empregam RFID e aprendizado de máquina (Seção 3.1), o resultado da avaliação geral do sistema híbrido (cenário “limpo”) foi superior a todos aqueles trabalhos. O sistema híbrido proposto obteve um desempenho 85% superior ao trabalho [11], considerando que este trabalho possuía um cenário com quatro antenas RFID e etiquetas de referência a cada 50 cm. Em relação ao trabalho [48], a precisão foi aprimorada em 42 cm. Em [48] haviam cinco antenas presentes e as etiquetas de referências eram distribuídas a cada 30 cm. Referente ao único trabalho correlato listado que emprega modelos SVR [96], os resultados aqui apresentados não são passíveis de comparação devido às dimensões daquele ambiente serem extremamente pequenas.

No que diz respeito aos sistemas correlatos que combinam tecnologias RFID e visual (Seção 3.2), a precisão obtida no sistema apresentado foi superior em 9 cm, 86 cm e 12 cm aos trabalhos [11], [93] e [67] (estático), respectivamente. Em comparação a outros dois sistemas híbridos aplicados a ambientes estacionários [18], [76], o desempenho foi inferior em aproximadamente 13 cm. Em ambos os casos, o sistema dependia de um novo hardware ainda não disponível comercialmente.

5.7 Avaliação do Sistema Híbrido em Escalabilidade Tridimensional

Esta fase de experimentos avalia de modo geral o sistema híbrido sob a escalabilidade 3D, ou seja, nestes experimentos além de indicar a posição do marcador alvo na imagem plana, o sistema também fornece a distância em profundidade entre a câmera e o alvo. Esta fase também contempla os dois modelos de aprendizado propostos

(RNA e SVR) e valida o sistema frente a cenários com marcadores distribuídos de forma mais densa.

Nestes experimentos, as configurações primárias do sistema utilizaram as mesmas propriedades constantes na Tabela 13. Entretanto, o formato de posicionamento de antenas EIXO recebeu mais uma antena. A terceira antena foi posicionada junto à câmera, contemplando a distância no eixo z (Figura 67).

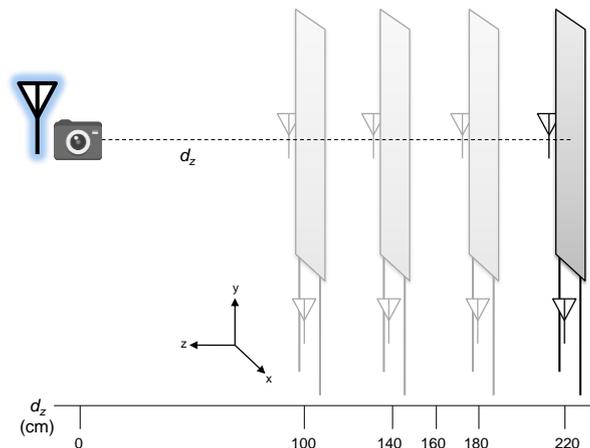


Figura 67 - Cenário avaliado sob escalabilidade tridimensional. Antena em destaque foi adicionada para contemplar o eixo z . Fonte: o autor.

Na fase *offline*, foram coletadas amostras nas distâncias d_z de 100, 140, 180 e 220 cm, totalizando 4400 registros. Em todas as distâncias, todas as etiquetas presentes foram lidas. As amostras de todas as distâncias foram combinadas em um único conjunto de dados que alimentou o treinamento dos modelos de aprendizado de máquina. Da mesma forma que ocorreu na Seção 5.6, as antenas dos eixos x e y tiveram que ser deslocadas em cada distância d_z avaliada. Na ilustração de exemplo (Figura 67), o treinamento visa a distância de 220 cm.

Na fase *online*, o objetivo dos experimentos era fornecer as coordenadas (x, y, z) de seis marcadores alvo distribuídos no ambiente (Figura 48). Foram avaliadas as distâncias d_z de 160 cm e 220 cm. Esta fase de experimentos também contemplou os três testes de validação realizados da Seção 5.6: “somente RFID”, “denso” e “limpo”. Cada execução foi repetida três vezes. Na fase *online*, a execução do modelo SVR utilizou as mesmas amostras que serviram de entrada para o modelo RNA.

A dimensão do ROI (subimagem) foi configurada em 20% e 15% para as distâncias d_z de 160 cm e 220 cm, respectivamente. Neste caso, considera-se uma limitação do sistema a entrada manual do tamanho do ROI, pois o usuário desconhece a distância d_z

durante a execução do sistema, ou seja, ele não saberia qual o percentual de recorte da imagem que corresponde a uma determinada distância no ambiente real.

5.7.1 Resultados

No resultado da fase *offline*, os valores de MSE foram 28,0, 21,1, 4,8 e 40,6 para os modelos RNA das frequências 923,25, 923,75, 924,25 e 924,75 MHz, respectivamente. A unidade de medida utilizada no MSE é a mesma dos valores de saída do conjunto de treinamento. No caso das coordenadas *x* e *y* refere-se à distância nas coordenadas da grade virtual, enquanto a coordenada *z* compreende a distância em cm no espaço real.

Os erros no treinamento demonstram que a adição de mais uma dimensão no espaço de saída dificultou o ajuste dos pesos e bias da rede, de tal forma que não foi possível explicar satisfatoriamente os posicionamentos tridimensionais através da relação entre os valores de RSSI e as coordenadas. A inserção de mais uma dimensão no espaço de saída em modelos de aprendizado de máquina é um problema conhecido na literatura, cunhado por Richard Bellman como a “maldição da dimensionalidade” (do inglês, “*curse of dimensionality*”). A medida que o número de respostas ou dimensões aumenta, a quantidade de informações também cresce. Basicamente, quando a dimensionalidade aumenta, o volume do espaço aumenta tão rapidamente que os dados disponíveis se tornam escassos. O fato de um sistema possuir mais respostas é atrativo, entretanto é possível que um modelo de predição tenha um desempenho pior do que aquele com menos saídas. O número de pontos de dados necessários para preencher um espaço pode aumentar exponencialmente com a dimensionalidade [72].

Em relação à fase *online*, a Tabela 16 resume o desempenho do sistema no cenário 3D. Diferente dos resultados na escalabilidade 2D, o modelo SVR obteve desempenho 5,3% superior ao modelo RNA. De maneira geral, os resultados dos testes de validação do sistema híbrido e do subsistema RFID foram muito similares. No modelo RNA, a abordagem híbrida aumentou o erro em 1 cm em relação ao subsistema RFID.

Tabela 16 - Desempenho da localização (RMSE em cm) sob escalabilidade tridimensional.

Distância (z)	RNA			SVR		
	Somente RFID	Híbrido		Somente RFID	Híbrido	
		Denso	Limpo		Denso	Limpo
160 cm	79,0	77,6	83,3	83,5	81,3	86,5
220 cm	80,7	83,7	78,2	62,7	65,3	74,6
RMSE	79,8	80,7	80,8	73,9	73,7	80,8

Analisando o modelo SVR, o sistema híbrido (teste “limpo”) obteve uma precisão até 12 cm maior do que o subsistema RFID. Nestes casos, dentre os ROIs analisados pelo subsistema visual, aquele que tinha o marcador visual era um ROI que resultou em coordenadas x e y precisas, mas uma distância em profundidade (z) com um erro alto na estimativa.

A Figura 68 exibe um exemplo desta situação em um dos experimentos do teste “limpo” em 220 cm. No caso, o subsistema visual detectou o marcador alvo “5” com precisão exata no plano bidimensional (Figura 68a). A Figura 68b ilustra as estimativas do subsistema RFID e a aplicação da técnica k-means, resultando as posições associadas a diversos ROIs. Conforme o algoritmo proposto (Figura 39, Seção 4.4), um marcador visual pode ser detectado em qualquer uma das regiões de interesse. Nesse exemplo, o marcador visual do alvo “5” foi detectado no ROI “C”.

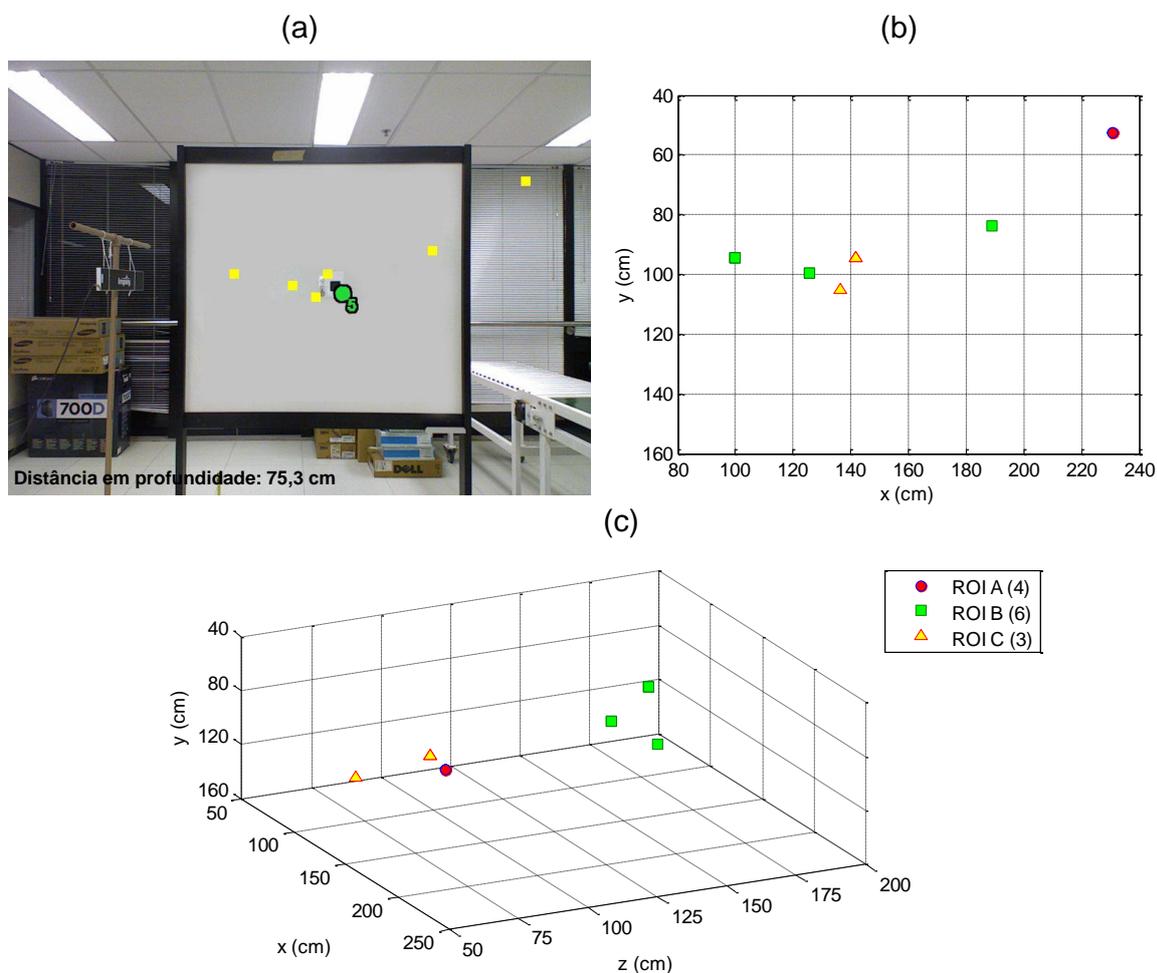


Figura 68 - Exemplo de experimento em escalabilidade 3D. Teste “limpo”; modelo: SVR; distância z real: 220 cm; marcador alvo: 5; precisão: 144,7 cm. Tela do sistema exibindo o resultado (a) e estimativas RFID e regiões de interesse na visão 2D e 3D em (b) e (c), respectivamente. Fonte: o autor.

Entretanto, como é observado na Figura 68c, a distância no eixo z deste ROI está afastada da distância de 220 cm, motivando os erros de localização apresentados no sistema híbrido. Por fim, é possível observar que o resultado que emprega somente o sistema RFID foi ligeiramente melhor, pois ele utiliza o centroide da região com maior peso como resposta, ou seja, o ROI “B” foi selecionado, sendo que esta região está mais próxima da distância em profundidade real.

Os gráficos CDF da Figura 69 demonstram que os erros nos três testes de validação são muito próximos. Isto ocorre principalmente porque muitos experimentos utilizaram somente o resultado do subsistema RFID, ou seja, não foi possível aprimorar as estimativas fornecidas por este subsistema através da análise visual.

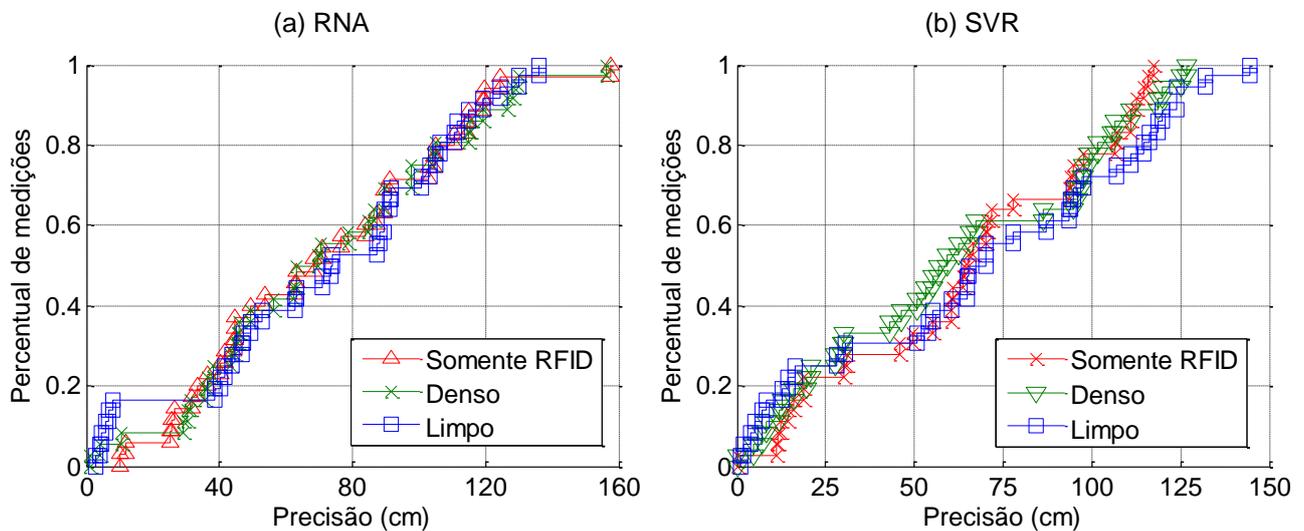


Figura 69 - Erro de localização acumulado (CDF) para ambos os modelos de aprendizado nos três testes de validação. Escalabilidade: 3D. Fonte: o autor.

O gráfico CDF da Figura 70 exibe uma comparação do RMSE para cada eixo. Ou seja, são apresentados os erros das estimativas nos eixos x (horizontal), y (vertical) e z (profundidade). Este gráfico contempla os resultados de ambos os modelos de aprendizado. É possível verificar certa similaridade na distribuição dos valores no gráfico até a precisão de 120 cm, sendo que, após esta distância, as medições contemplam somente o eixo z. O eixo x é aquele que possui a menor taxa de erros, sendo que mais de 90% das medições possuem precisão inferior a 60 cm. Os erros do eixo z após 120 cm ocorreram principalmente em virtude do erro de treinamento (MSE) no modelo RNA. Como exemplo, em um dos experimentos que visava a profundidade de 220 cm, somente a rede neural da frequência 924,25 MHz (melhor MSE no treinamento) resultou em uma distância z próxima a 2 metros, sendo que todas as demais redes estimaram distâncias entre 60 cm e 140 cm.

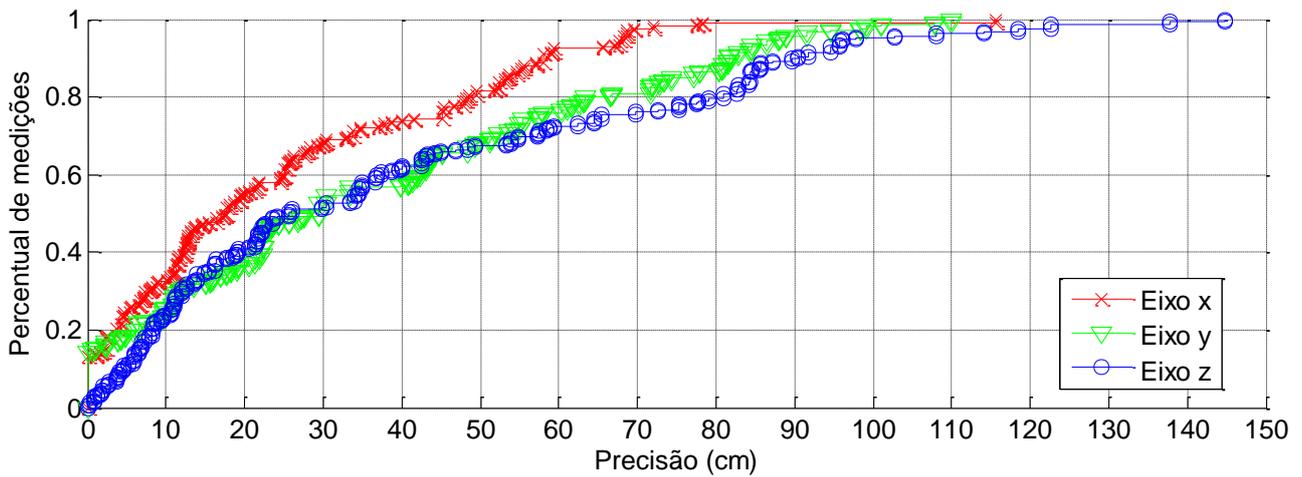


Figura 70 – Erro acumulado (CDF) para cada eixo no espaço de localização. Medições contemplam todos os cenários, modelos de aprendizado e testes de validação. Fonte: o autor.

Em comparação aos dois sistemas correlatos que contemplam ambientes tridimensionais [18], [76], o desempenho do IPS proposto foi pior em ambos os casos, aumentando o erro em aproximadamente 80 cm. Entretanto, ambos os trabalhos são propostos como um novo hardware e não poderiam ser implantados sob uma infraestrutura de componentes já existente. Além disso, o trabalho [18] estima a distância em profundidade utilizando uma tecnologia complementar (laser).

6 CONCLUSÃO E TRABALHOS FUTUROS

Foi apresentado um sistema híbrido de localização de objetos, aplicado a ambientes internos e estacionários, utilizando equipamentos *off-the-shelf*. A partir de etiquetas RFID afixadas em objetos, o sistema os localiza com o auxílio de marcadores visuais. São propostos modelos baseados em aprendizado de máquina, capazes de aprender os *fingerprints* RSSI e inferir a localização dos marcadores alvo. Um método de multi-frequência é proposto para superar restrições existentes no *hardware* RFID e em regulamentações de radiofrequência. Uma técnica k-means e algoritmos de visão computacional são aplicados a fim de refinar a localização.

Foram conduzidos experimentos em laboratório a fim de avaliar o desempenho da localização e comparar os modelos desenvolvidos. Os resultados em espaço bidimensional demonstraram uma precisão de 8,9 cm no melhor caso e a abordagem híbrida superou em até 32% a localização que emprega somente uma tecnologia (RFID). Estes resultados demonstraram que o sistema pode ser aplicado a uma localização de objetos com tamanho reduzido, embora a abordagem proposta ainda possua algumas limitações nos cenários onde muitos itens estejam próximos uns dos outros.

As próximas seções apresentam as contribuições do trabalho desenvolvido, as publicações obtidas durante o período de doutoramento e as possibilidades de trabalhos futuros.

6.1 Contribuições do Trabalho

Dentre as contribuições do trabalho, podem ser citadas como as mais importantes:

Modelos de aprendizado de máquina baseados em RNA e SVR: a primeira contribuição diz respeito aos modelos de predição da localização através dos valores de RSSI coletados por leitores RFID. Foram apresentados e comparados dois modelos de aprendizado de máquina, empregando as abordagens redes neurais artificiais e *support vector regression*. Ambas abordagens diferem-se da maior parte dos trabalhos correlatos por contemplarem a escalabilidade tridimensional. Além disso, o modelo RNA emprega o algoritmo de retro-propagação Levenberg-Marquardt, sendo validado em cenários distintos àqueles avaliados em outras pesquisas. O modelo SVR também diferencia-se por empregar um *kernel wavelet* e ser aplicado a ambientes de maiores proporções quando comparado às pesquisas existentes.

Método de multi-frequência: é proposto um método de particionamento das amostras de RSSI por frequência de operação. O conjunto de dados coletados pelo leitor RFID é particionado por frequência, criando e alimentando uma instância de modelo de aprendizado de máquina para cada partição. Este método tem como objetivo aprimorar a precisão das estimativas e, ao mesmo tempo, permitir o uso de equipamentos *off-the-shelf* no IPS. Este método diferencia-se dos trabalhos existentes, que geralmente empregam a média simples dos valores de RSSI sob todas as frequências, ou utilizam um *firmware* modificado para emissão de uma frequência única, o que atualmente não contempla regulamentações de radiofrequência.

Método para fusão das estimativas RFID e visual: uma importante contribuição do trabalho diz respeito ao método de combinação das tecnologias RFID e visual. É apresentado um modelo k-means aplicado às estimativas do subsistema RFID. Tal modelo fornece múltiplas regiões de interesse (MROI) para o subsistema visual, que por sua vez limita sua atuação sobre estas áreas reduzidas. Este método difere-se dos métodos de fusão empregados em sistemas híbridos de localização existentes. O método demonstrou uma boa precisão, principalmente em cenários que os marcadores visuais são posicionados de maneira mais esparsa no ambiente.

Experimentos e resultados da configuração primária do IPS: a fim de validar o IPS proposto, foram realizadas diversas sessões de experimentos com as mais variadas combinações de configuração. Os três experimentos que contemplam a definição da configuração primária do sistema possuem grande valor a projetistas e pesquisadores de sistemas de localização. O primeiro experimento testou o posicionamento das antenas RFID no cenário, avaliando a quantidade e a disposição dos elementos. O segundo teste teve como objetivo avaliar a quantidade e a forma de organização das etiquetas de referência. O último experimento verificou o melhor tempo de leitura para a coleta das amostras de RSSI das etiquetas, tanto na fase de treinamento quanto na fase de execução do sistema. A documentação detalhada da configuração dos ensaios e os resultados obtidos formam uma importante contribuição para o desenvolvimento de novos IPSs, principalmente aqueles baseados na tecnologia RFID.

6.2 Publicações

Os seguintes artigos científicos e um depósito de pedido de patente, relacionados a esta tese, foram publicados ou submetidos para publicação:

- Sapiras, Leonardo A. ; Berz, Everton L. ; Azambuja, Marcelo C. ; Hessel, Fabiano P. “Software para projeto de ambientes RFID aplicado a modelos de predição”. In: Escola Regional de Redes de Computadores (ERRC), 2011, São Leopoldo, pp. 47-50.
- Berz, Everton L. ; Hessel, Fabiano P. ; De Azambuja, Marcelo C. ; Ody, Julio C. “Prediction of RFID systems coverage applied to smart cards scenario”. In: IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2012, pp. 1478-1484.
- Azambuja, M. C. ; Hessel, F. P. ; Berz, Everton L. ; Porfírio, L. B. ; Valerio, P. R. ; Baladei, S. P. ; Jung, C. F. “Deterministic propagation model for RFID using site-specific and FDTD”. International Journal of Electronics (Online), 2014, pp. 1-14.
- Berz, Everton L.; Hessel, Fabiano P. “Sistema e Método de Localização de Coordenadas Espaciais de Objetos e Uso do Referido Sistema para Localização de Objetos Dispostos em Ambientes Internos”. Depósito de Pedido de Patente, INPI N° BR 10 2014 010722-3, 2014.
- Tesch, Deivid A.; Berz, Everton L.; Hessel, Fabiano P. “RFID Indoor Localization Based on Doppler Effect”. In: 16th International Symposium on Quality Electronic Design (ISQED), 2015, pp. 556-560.
- Berz, Everton L.; Tesch, Deivid A.; Hessel, Fabiano P. “RFID indoor localization based on support vector regression and k-means”. In: 24th IEEE International Symposium on Industrial Electronics (ISIE), 2015. Aceito para publicação.
- Berz, Everton L.; Tesch, Deivid A.; Hessel, Fabiano P. “A Hybrid RFID and CV System for Item-Level Localization of Stationary Objects”. Submetido para: Mobile Networks and Applications (Springer) - Special Issue: Intelligent Mobile Applications.

6.3 Trabalhos Futuros

Como trabalhos futuros, foram identificadas as seguintes possibilidades:

- Os experimentos realizados neste trabalho contemplaram uma área que pode ser aplicada a diversos casos de uso reais. Entretanto, é desejável avaliar o IPS sob dimensões geográficas maiores e que necessitem uma infraestrutura ampliada de comunicação. Assim, recomenda-se que a

escalabilidade geográfica do sistema seja avaliada através de experimentos em ambientes maiores.

- A fim de satisfazer o item anterior, pode ser necessário estender o sistema para suportar múltiplos leitores RFID e múltiplas câmeras digitais. Como a área de cobertura dos sistemas RFID passivos é bastante limitada, a principal maneira de expandir a cobertura seria aumentando o número de leitores e antenas RFID. Atualmente, o sistema já está preparado para receber amostras coletadas por múltiplas antenas, mas somente conectadas a um único leitor RFID. A combinação e a análise de fotografias capturadas por múltiplas câmeras presentes no cenário é outro desafio existente.
- O método de fusão desenvolvido neste trabalho não necessita permanecer limitado às tecnologias RFID e visual. Um trabalho de pesquisa que integre outras tecnologias ao método proposto é uma importante contribuição no desenvolvimento de sistemas híbridos de localização *indoor*. Os dispositivos infravermelhos e os sensores ultrassônicos são possibilidades de tecnologias que podem ser combinadas no método de fusão.
- Acredita-se que a proposta de subsistema visual pode ser ampliada para marcadores visuais que implementem métodos mais robustos de reconhecimento de padrões e rastreamento de objetos. Com o auxílio dos métodos e resultados descritos neste trabalho, pesquisas acerca de algoritmos de visão computacional mais avançados podem ser realizadas a fim de contemplar cenários mais complexos.
- Conforme citado na Seção 5.7, a configuração manual da dimensão do ROI pode ser um limitador do sistema quando se deseja contemplar a escalabilidade tridimensional. Sendo assim, o sistema deixa uma lacuna na definição deste parâmetro, possibilitando a pesquisa de técnicas que possam estimar o tamanho do ROI baseado em informações dos sensores visuais ou RFID.
- Sugere-se a realização de experimentos em casos de uso reais, executando o sistema em um centro de distribuição que necessite garantir de maneira automática a localização dos itens, ou seja, verificar se produtos, materiais ou equipamentos estão de fato nos locais em que deveriam estar. Outro tipo de caso de uso real que pode ser integrado e avaliado diz respeito ao uso de robôs que coletam objetos automaticamente.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Agência Brasileira de Telecomunicações (ANATEL). “Resolução 506 - Artigo 40”. Capturado em: <http://legislacao.anatel.gov.br/resolucoes/2008/104-resolucao-506>, Novembro 2014.
- [2] V. T. Ales. “O Algoritmo Sequential Minimal Optimisation para Resolução do Problema de Support Vector Machine: uma Técnica para Reconhecimento de Padrões”, Dissertação de Mestrado, Universidade Federal do Paraná, 2008, 150.
- [3] J. C. L. Alves. “Máquina de vetores de suporte aplicada a dados de espectroscopia NIR de combustíveis e lubrificantes para o desenvolvimento de modelos de regressão e classificação”, Tese de Doutorado, Instituto de Química, Universidade Estadual de Campinas, 2012, 272p.
- [4] Atlas RFID Solutions Inc. “Atlas RFID solutions - Active vs Passive”. Capturado em: <http://www.atlasrfid.com/Technology/ActivevsPassive.aspx>, Fevereiro 2015.
- [5] M. C. de Azambuja. “Modelos e técnicas para simulação de sistemas UHF de identificação por rádio frequência (RFID)”, Tese de Doutorado, Programa de Pós-Graduação em Ciência da Computação, PUCRS, 2011, 137p.
- [6] M. C. de Azambuja, F. P. Hessel, E. L. Berz, L. B. Porfírio, P. R. Valério, S. D. P. Baladei, C. F. Jung. “Deterministic propagation model for RFID using site-specific and FDTD”. *International Journal of Electronics*, vol. 102, 2014, pp. 932–945.
- [7] Z. Babic, M. Ljubojevic, V. Risojevic. “Indoor RFID localization improved by motion segmentation”. In: 7th International Symposium on Image and Signal Processing and Analysis (ISPA), 2011, pp. 271–276.
- [8] M. Beltrami. “Precificação de opções sobre ações por modelos de Support Vector Regression”, Dissertação de Mestrado, Universidade Federal do Paraná, 2009, 125p.
- [9] M. Bouet, A. L. dos Santos. “RFID tags: Positioning principles and localization techniques”. In: 1st IFIP Wireless Days, 2008, pp. 1–5.
- [10] B. P. R. de Carvalho. “O estado da arte em métodos para reconhecimento de padrões: Support Vector Machine”. In: Sucesu 2005 - Congresso Nacional de Tecnologia da Informação e Comunicação, 2005, pp. 1–7.
- [11] H. Chae, S. Na. “Mobile Robot Navigation Using RFID and Vision Sensor”. In: International Symposium on Information Technology Convergence (ISITC), 2007, pp. 52–57.
- [12] A. Chattopadhyay, A. R. Harish. “Analysis of low range Indoor Location Tracking techniques using Passive UHF RFID tags”. In: IEEE Radio and Wireless Symposium, 2008, pp. 351–354.

- [13] A. da C. F. Chaves. “Extração de Regras Fuzzy para Máquinas de Vetor Suporte (SVM) para Classificação em Múltiplas Classes”, Tese de Doutorado, Pontifícia Universidade Católica do Rio de Janeiro, 2006, 225p.
- [14] R. Clark. “A MATLAB implementation of Support Vector Regression (SVR)”. Capturado em: <http://www.mathworks.com/matlabcentral/fileexchange/43429-support-vector-regression>, Outubro 2014.
- [15] CodeProject. “The Code Project Open License (CPOL) 1.02”. Capturado em: <http://www.codeproject.com/info/cpol10.aspx>, Fevereiro 2015.
- [16] A. Conci. “Canny: detecção de borda”. Capturado em: <http://www2.ic.uff.br/~aconci/canny.pdf>, Fevereiro 2015.
- [17] G. Dash. “How RF Anechoic Chambers Work”, Relatório Técnico, 2005, 9p.
- [18] T. Deyle, H. Nguyen, M. Reynolds, C. C. Kemp. “RF vision: RFID receive signal strength indicator (RSSI) images for sensor fusion and mobile manipulation”. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009, pp. 5553–5560.
- [19] M. S. Dias. “O uso de Máquina de Suporte Vetorial para Regressão (SVR) na Estimação da Estrutura a Termo da Taxa de Juros do Brasil”, Dissertação de Mestrado, Pontifícia Universidade Católica do Rio de Janeiro, 2007, 79p.
- [20] EmguCV. “EmguCV”. Capturado em: <http://www.emgu.com>, Novembro 2014.
- [21] B. Engelbrecht. “Reading Barcodes from an Image - III”, *CodeProject*. Capturado em: <http://www.codeproject.com/Articles/42852/Reading-Barcodes-from-an-Image-III>.
- [22] EPCglobal Inc. “Low Level Reader Protocol (LLRP)”, Relatório Técnico, 2010, 198p.
- [23] L. Fang, C. Xie. “1-D Barcode Localization in Complex Background”. In: International Conference on Computational Intelligence and Software Engineering, 2010, pp. 1–3.
- [24] Federal Communications Commission (FCC). “Electronic Code of Federal Regulations - Section 15.247”. Capturado em: <http://www.ecfr.gov/>, Fevereiro 2015.
- [25] M. Fiala. “Designing highly reliable fiducial markers”. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32-7, jul. 2010, pp. 1317–1324.
- [26] K. Finkenzerler. “RFID Handbook: Fundamentals and Applications in Contact Less Smart Cards and Identification”. New York: Wiley & Sons, 2003, 2^a ed., 446p.
- [27] A. Gautham. “Practical evaluation and analysis of passive UHF RFID tags”, Dissertação de Mestrado, The University of Texas at Arlington, 2008, 117p.
- [28] B. Glover, H. Bhatt. “RFID Essentials”. Sebastopol: O’reilly, 2006, 1^a ed., 288p.
- [29] T. Godim. “Interpolação Spline”. Capturado em: <http://www2.ic.uff.br/~aconci/splineatual.html>, Fevereiro 2015.

- [30] M. Goller, C. Feichtenhofer, A. Pinz. “Fusing RFID and computer vision for probabilistic tag localization”. In: IEEE International Conference on RFID (IEEE RFID), 2014, pp. 89–96.
- [31] B. Gong. “CSC5280 Project 1: Canny Edge Detection”. Capturado em: <http://www-scf.usc.edu/~boqinggo/index.html>, Fevereiro 2015.
- [32] R. W. R. C. Gonzalez. “Processamento de Imagens Digitais”. São Paulo: Edgard Blucher, 2000, 509p.
- [33] M. Grewal, A. Andrews. “Kalman filtering: theory and practice using MATLAB”. Hoboken, New Jersey: John Wiley & Sons Inc., 2008, 3^a ed., 589p.
- [34] GS1. “EPCglobal - GS1”. Capturado em: <http://www.gs1.org/epcglobal>, Fevereiro 2015.
- [35] Y. Gu, A. Lo, I. Niemegeers. “A survey of indoor positioning systems for wireless personal networks”. *IEEE Communications Surveys & Tutorials*, vol. 11-1, 2009, pp. 13–32.
- [36] S. Haykin. “Redes Neurais: Princípios e prática”. Porto Alegre: Bookman, 2001, 2^a ed., 900p.
- [37] J. Hightower, G. Borriello. “Location systems for ubiquitous computing”. *Computer*, vol. 34-8, 2001, pp. 57–66.
- [38] J. Hightower, G. Borriello. “Location Sensing Techniques”, Relatório Técnico, 2001, 8p.
- [39] J. Hightower, B. Brumitt, G. Borriello. “The location stack: a layered model for location in ubiquitous computing”. In: Proceedings Fourth IEEE Workshop on Mobile Computing Systems and Applications, 2002, pp. 22–28.
- [40] S. Holm. “Hybrid ultrasound-RFID indoor positioning: Combining the best of both worlds”. In: IEEE International Conference on RFID, 2009, pp. 155–162.
- [41] International Business Machines (IBM) Corporation. “IBM SPSS Statistics”. Capturado em: <http://www-01.ibm.com/software/analytics/spss/>, Janeiro 2015.
- [42] IPIN. “International Conference on Indoor Positioning and Indoor Navigation (IPIN)”. Capturado em: <http://ipin-conference.org/>, Maio 2013.
- [43] B. Jachimczyk, D. Dziak, W. J. Kulesza. “RFID - Hybrid Scene Analysis-Neural Network system for 3D Indoor Positioning - Optimal System Arrangement Approach”. In: IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, 2014, pp. 191–196.
- [44] B. Jahne, H. Haubecker. “Computer vision and applications: a guide for students and practitioners”. Orlando, FL, USA: Academic Press Inc., 2000, 679p.

- [45] J. Karlsson, K. Ren, H. Li. "Tracking and Identification of Animals for a Digital Zoo". In: IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing, 2010, pp. 510–515.
- [46] G. Kim, E. M. Petriu. "Fiducial marker indoor localization with Artificial Neural Network". In: IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2010, pp. 961–966.
- [47] Z. L. Kovacs. "Redes Neurais Artificiais: Fundamentos e Aplicações". Livraria da Física, 2006, 4^a ed., 174p.
- [48] H.-Y. Kung, S. Chaisit, N. T. M. Phuong. "Optimization of an RFID location identification scheme based on the neural network". *International Journal of Communication Systems*, vol. 28-4, 2013, pp. 625–644.
- [49] B. Larson, R.; Farber. "Estatística aplicada". São Paulo: Pearson Prentice Hall, 2010, 4^a ed., 637p.
- [50] A. R. G. Lima. "Máquinas de Vetores Suporte na Classificação de Impressões Digitais", Dissertação de Mestrado, Universidade Federal do Ceará, 2002, 81p.
- [51] C. A. D. M. Lima. "Comitê de Máquinas: Uma Abordagem Unificada Empregando Máquinas de Vetores-Suporte", Tese de Doutorado, Universidade Estadual de Campinas, 2004, 378p.
- [52] H. Liu, H. Darabi, P. Banerjee, J. Liu. "Survey of Wireless Indoor Positioning Techniques and Systems". *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, vol. 37-6, nov. 2007, pp. 1067–1080.
- [53] A. C. Lorena, A. C. P. L. F. de Carvalho. "Uma Introdução às Support Vector Machines". *RITA - Revista de Informática Teórica e Aplicada*, vol. 14, 2007, pp. 43–67.
- [54] A. Malik. "RTLS for Dummies". Wiley Publishing Inc., 2009, 336p.
- [55] M. Marengoni, D. Stringhini. "Tutorial: Introdução à Visão Computacional usando OpenCV". *RITA - Revista de Informática Teórica e Aplicada*, vol. 16-1, 2009, pp. 125–160.
- [56] M. Matteucci. "A Tutorial on Clustering Algorithms". Capturado em: http://home.deib.polimi.it/matteucc/Clustering/tutorial_html/kmeans.html, Fevereiro 2015.
- [57] R. Mautz, S. Tilch. "Survey of optical indoor positioning systems". *2011 International Conference on Indoor Positioning and Indoor Navigation*, set. 2011, pp. 1–7.
- [58] C. A. Mello. "Segmentação - Detecção de Bordas". Capturado em: http://www.cin.ufpe.br/~cabm/visao/Aula05_EdgeDetection.pdf, Fevereiro 2015.
- [59] Microsoft Corporation. ".NET Framework". Capturado em: <https://msdn.microsoft.com/pt-br/vstudio/aa496123.aspx>, Fevereiro 2015.

- [60] Minitab Inc. “What is an empirical CDF plot?”. Capturado em: <http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/graphs/graphs-of-distributions/empirical-cdf-plots/empirical-cdf-plot/>, Fevereiro 2015.
- [61] M. C. Monard, J. A. Baranauskas. “Conceitos sobre Aprendizado de Máquina”. Editora Manole, 2003, 1ª ed., pp. 89–114.
- [62] K. P. Murphy. “Machine Learning: A Probabilistic Perspective”. The MIT Press, 2012, 1ª ed., 1104p.
- [63] M. Narciso. “Regressão Linear Múltipla - Eliminação de Dados Espúrios”. Capturado em: <http://condicao inicial.com/2010/09/regressao-linear-multipla-eliminacao-de-dados-espurios.html>, Fevereiro 2015.
- [64] W. Navidi. “Probabilidade e Estatística para Ciências Exatas”. McGraw-Hill, 2012, 1ª ed., 616p.
- [65] W. W. Y. Ng, Y.-S. Qiao, L. Lin, H.-L. Ding, P. P. K. Chan, D. S. Yeung. “Intelligent book positioning for library using RFID and book spine matching”. In: International Conference on Machine Learning and Cybernetics, 2011, pp. 465–470.
- [66] L. M. Ni, Y. Liu, Y. C. Lau, A. P. Patil. “LANDMARC: Indoor Location Sensing Using Active RFID”. *Wireless networks*, vol. 10-6, 2004, pp. 701–710.
- [67] T. Nick, S. Cordes, J. Gotze, W. John. “Camera-assisted localization of passive RFID labels”. In: International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2012, pp. 1–8.
- [68] K. Al Nuaimi, H. Kamel. “A survey of indoor positioning systems and algorithms”. In: International Conference on Innovations in Information Technology, 2011, pp. 185–190.
- [69] OpenCV. “OpenCV - Open Source Computer Vision Library”. Capturado em: <http://opencv.org>, Fevereiro 2015.
- [70] K. Pahlavan, J. Makela. “Indoor geolocation science and technology”. *IEEE Communications Magazine*, vol. 40-2, 2002, pp. 112–118.
- [71] PLANS. “IEEE/ION Position Location and Navigation Symposium (PLANS)”. Capturado em: <http://www.plansconference.org/>, Maio 2013.
- [72] K. L. Priddy, P. E. Keller. “Artificial Neural Networks: An Introduction (SPIE Tutorial Texts in Optical Engineering, Vol. TT68)”. SPIE Publications, 2005, 165p.
- [73] H. N. S. Rosa. “Redes Neurais na Previsão de Séries Temporais”. In: I Workcomp-Sul, 2004, pp. 1–12.
- [74] A. S. M. Sala, R. G. Quirós, E. E. López. “Using neural networks and Active RFID for indoor location services”. In: RFID System - European Workshop on Smart Objects: Systems, Technologies and Applications, 2010, pp. 1–9.

- [75] A. S. M. Sala, R. G. Quirós, E. E. López. “Active RFID Reader clustering and neural networks for indoor positioning”. In: The Third International EURASIP Workshop on RFID Technology, 2010, pp. 62–65.
- [76] A. P. Sample, C. Macomber, L. Jiang, J. R. Smith. “Optical localization of passive UHF RFID tags with integrated LEDs”. In: IEEE International Conference on RFID (RFID), 2012, pp. 116–123.
- [77] A. P. Sample, D. J. Yeager, P. S. Powledge, A. V. Mamishev, J. R. Smith. “Design of an RFID-Based Battery-Free Programmable Sensing Platform”. *IEEE Transactions on Instrumentation and Measurement*, vol. 57-11, nov. 2008, pp. 2608–2615.
- [78] T. Sanpechuda, L. Kovavisaruch. “A review of RFID localization: Applications and techniques”. In: 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2008, pp. 769–772.
- [79] L. G. Shapiro, G. Stockman. “Computer Vision”. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001, 1^a ed., 608p.
- [80] R. de A. Silva. “Um Sistema de Auxílio à Localização de Etiquetas RFID em Ambientes Internos”, Universidade Federal de Pernambuco, 2008, 111p.
- [81] A. J. Smola, B. Schölkopf. “A Tutorial on Support Vector Regression”, Relatório Técnico, 1998, 73p.
- [82] L. J. Spiegel, M. R.; Stephens. “Estatística”. Porto Alegre: Bookman, 2009, 4^a ed., 597p.
- [83] A. Stathakis, E. M. Petriu. “Robust pseudo-random fiducial marker for indoor localization”. In: IEEE International Symposium on Robotic and Sensors Environments (ROSE), 2011, pp. 19–24.
- [84] TEC-IT Datenverarbeitung GmbH. “Free Online Barcode Generator”. Capturado em: <http://barcode.tec-it.com/>, Fevereiro 2015.
- [85] The MathWorks Inc. “Train and Apply Multilayer Neural Networks”. Capturado em: <http://www.mathworks.com/help/nnet/ug/train-and-apply-multilayer-neural-networks.html>, Novembro 2014.
- [86] The MathWorks Inc. “Levenberg-Marquardt backpropagation - MATLAB trainlm”. Capturado em: <http://www.mathworks.com/help/nnet/ref/trainlm.html>, Novembro 2014.
- [87] The MathWorks Inc. “k-Means Clustering”. Capturado em: <http://www.mathworks.com/help/stats/k-means-clustering.html>, Novembro 2014.
- [88] The MathWorks Inc. “MATLAB Neural Network Toolbox”. Capturado em: <http://www.mathworks.com/products/neural-network/>, Novembro 2014.
- [89] The MathWorks Inc. “MATLAB - The Language of Technical Computing”. Capturado em: <http://www.mathworks.com/products/matlab/>, Fevereiro 2015.

- [90] V. N. Vapnik. "The Nature of Statistical Learning Theory". New York, NY, USA: Springer-Verlag New York, Inc., 1995, 1^a ed., 188p.
- [91] R. Vicente. "Redes Neurais para Inferência Estatística", Relatório Técnico, 2002, 75p.
- [92] S. Wagner, M. Handte, M. Zuniga, P. J. Marron. "On optimal tag placement for indoor localization". In: IEEE International Conference on Pervasive Computing and Communications, 2012, vol. March, pp. 162–170.
- [93] C.-S. Wang, L.-C. Cheng. "RFID & vision based indoor positioning and identification system". In: IEEE 3rd International Conference on Communication Software and Networks, 2011, pp. 506–510.
- [94] G. Welch, G. Bishop. "An Introduction to the Kalman Filter", Relatório Técnico, 2001, 81p.
- [95] A. Whitney, J. Fessler, J. Parker, N. Jacobs. "Received signal strength indication signature for passive UHF tags". In: IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2014, vol. 0503, pp. 1183–1187.
- [96] A. Wille, M. Broll, S. Winter. "Phase difference based RFID navigation for medical applications". In: IEEE International Conference on RFID, 2011, pp. 98–105.
- [97] H. Yang, A. C. Kot, X. Jiang. "Binarization of low-quality barcode images captured by mobile phones using local window of adaptive location and size.". *IEEE Transactions on Image Processing*, vol. 21-1, jan. 2012, pp. 418–425.
- [98] L. Zhang, W. Zhou, L. Jiao. "Wavelet Support Vector Machine". *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, vol. 34-1, fev. 2004, pp. 34–39.
- [99] L. Zhenzhong, M. Nezh, X. George, G. Liu, Y. Ono, B. Dayan. "Effects of orientation and obstacles on the RFID performance". In: 4th Annual Caneus Fly by Wireless Workshop, 2011, pp. 1–4.

APÊNDICE A – EXEMPLO DE AMOSTRAS COLETADAS PELO LEITOR RFID

Escalabilidade 2D e duas antenas

Antena;Frequência;RSSI;x;y
 2;923,75;-56,5;17;15
 2;923,75;-45;10;7
 2;923,75;-52,5;-14;-1
 2;923,75;-44,5;2;15
 2;923,75;-46,5;-6;7
 2;923,25;-54,5;18;-17
 2;923,25;-47,5;-14;15
 2;923,25;-48;10;-9
 2;923,25;-48;-6;-9
 2;923,25;-55,5;-14;-17
 2;923,25;-46;2;-1
 2;923,25;-50,5;17;-1
 2;923,25;-52,5;2;-17
 2;923,25;-52,5;2;-17
 2;923,25;-50,5;17;-1
 2;923,25;-55;-14;-17
 2;923,25;-45;10;7
 2;923,25;-48;10;-9
 2;924,75;-45;2;15
 2;924,75;-46,5;2;-1
 2;924,75;-46,5;-6;7
 2;924,75;-48,5;-6;-9
 2;924,75;-56,5;17;15
 2;924,75;-55,5;18;-17
 2;924,75;-52;-14;-1
 2;924,75;-48;-14;15
 1;924,75;-51;18;-17
 1;924,75;-39;-14;-17
 1;924,75;-34,5;-14;-1
 1;924,75;-44,5;10;7
 1;924,25;-51;17;15
 1;924,25;-42,5;2;-17
 1;924,25;-43,5;2;15
 1;924,25;-41,5;-14;15
 1;924,25;-42,5;10;-9
 1;924,25;-42;-6;7
 1;924,25;-48,5;17;-1
 1;924,25;-39,5;2;-1
 1;924,25;-37;-6;-9
 1;924,25;-48;17;-1
 1;923,75;-44,5;10;7
 1;923,75;-39;-14;-17
 1;923,75;-39,5;2;-1
 1;923,75;-41,5;-14;15
 1;923,75;-37;-6;-9
 1;923,75;-42;-6;7
 1;923,75;-34;-14;-1
 1;923,75;-51;18;-17
 1;923,75;-43,5;2;15
 ...

Escalabilidade 3D e três antenas

Antena;Frequência;RSSI;x;y;z
 1;923,75;-43,5;2;-16;100
 1;923,75;-43,5;2;-16;100
 2;923,75;-57,5;18;16;100
 2;923,75;-61;-15;0;100
 2;923,75;-48,5;10;8;100
 2;923,75;-45,5;-6;8;100
 2;923,75;-53;-15;16;100
 2;923,25;-56;-15;-16;100
 2;923,25;-56;18;0;100
 2;923,25;-45;2;16;100
 2;923,25;-52;2;0;100
 2;923,25;-49;-6;-8;100
 2;923,25;-48,5;10;-8;100
 3;924,25;-47;10;8;100
 3;924,25;-50,5;-15;-16;100
 3;924,25;-48;18;0;100
 3;924,75;-48,5;-15;0;100
 3;924,75;-42;2;0;100
 3;924,75;-44;2;16;100
 3;924,75;-44;2;-16;100
 3;924,75;-48;18;16;100
 1;923,25;-43,5;2;-16;100
 1;923,25;-50,5;18;16;100
 1;923,25;-41,5;-15;16;100
 1;923,25;-48,5;18;-16;100
 1;923,25;-42,5;10;8;100
 2;924,75;-48,5;7;5;180
 2;924,75;-52,5;7;-4;180
 2;924,75;-44;2;10;180
 2;924,75;-43,5;-2;5;180
 2;923,75;-52;2;-9;180
 2;923,75;-57,5;12;-9;180
 2;923,75;-48;-2;-4;180
 2;923,75;-46;2;0;180
 2;923,75;-49;-7;10;180
 2;923,75;-55,5;-7;0;180
 3;924,25;-43,5;-2;-4;180
 3;924,25;-43,5;7;-4;180
 3;924,25;-54,5;2;0;180
 3;924,25;-53,5;2;-9;180
 1;923,75;-51;12;0;180
 1;923,75;-35,5;-2;-4;180
 1;923,75;-41,5;-7;10;180
 1;923,75;-41,5;7;5;180
 1;923,75;-41;-2;5;180
 1;923,75;-44;2;-9;180
 1;923,75;-44,5;2;10;180
 1;923,25;-39,5;-6;8;100
 1;923,25;-38;2;0;100
 ...

ANEXO A – VISÃO COMPUTACIONAL: CONVOLUÇÃO E FILTRO GAUSSIANO

Fonte bibliográfica: [55]

Domínio de Espaço

O termo domínio espacial se refere à imagem em si, e métodos que atuam no domínio espacial estão baseados na manipulação direta dos *pixels* da imagem. Os processos no domínio espacial são caracterizados pela equação (17):

$$g(x, y) = T(f(x, y)) \quad (17)$$

onde $f(x, y)$ é a imagem original, $T(\cdot)$ é uma transformação na imagem e $g(x, y)$ é a imagem transformada. T é uma operação definida sobre uma vizinhança de influência do *pixel* que está localizado na posição (x, y) . A ideia de vizinhança de influência considera os *pixels* ao redor da posição (x, y) . Esta vizinhança é definida por uma região quadrada (ou retangular) e de tamanho (lado) ímpar. A Figura 71 mostra alguns exemplos de vizinhança com tamanhos variados, estas regiões, que definem matrizes nas operações de transformação, também são chamadas de máscaras

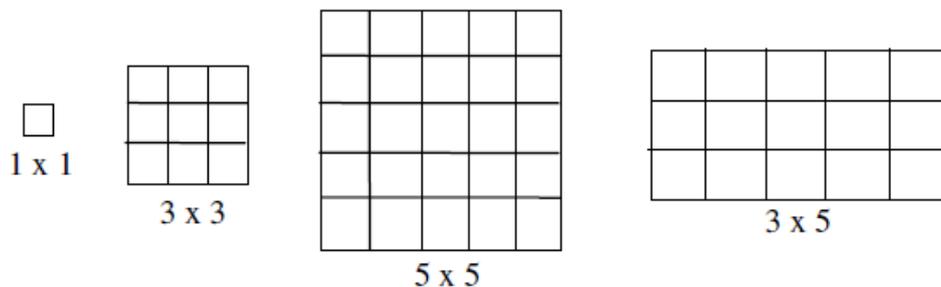


Figura 71 - Regiões de vizinhança (máscaras) de tamanhos diferentes, porém, todos com lado ímpar. Fonte: [55].

Filtragem no Domínio Espacial

Existem dois conceitos, de certa forma similares, que estão relacionados com a filtragem no domínio de espaço, são eles os de correlação e convolução. Como já mencionado, as transformações no domínio de espaço dependem de uma vizinhança de influência (máscara) do *pixel* que está sendo considerado. A ideia destas duas operações é a seguinte: cria-se uma máscara com dimensão d , onde cada posição da máscara

possui um determinado valor. Em seguida coloca-se a máscara com sua posição central sobre a imagem na posição (x, y) e, para cada posição da máscara executa-se o produto do valor da máscara pelo valor do *pixel*, faz-se a somatória destes valores obtidos na multiplicação e substitui-se o valor da posição (x, y) por este resultado. Esta operação é apresentada esquematicamente na Figura 72.

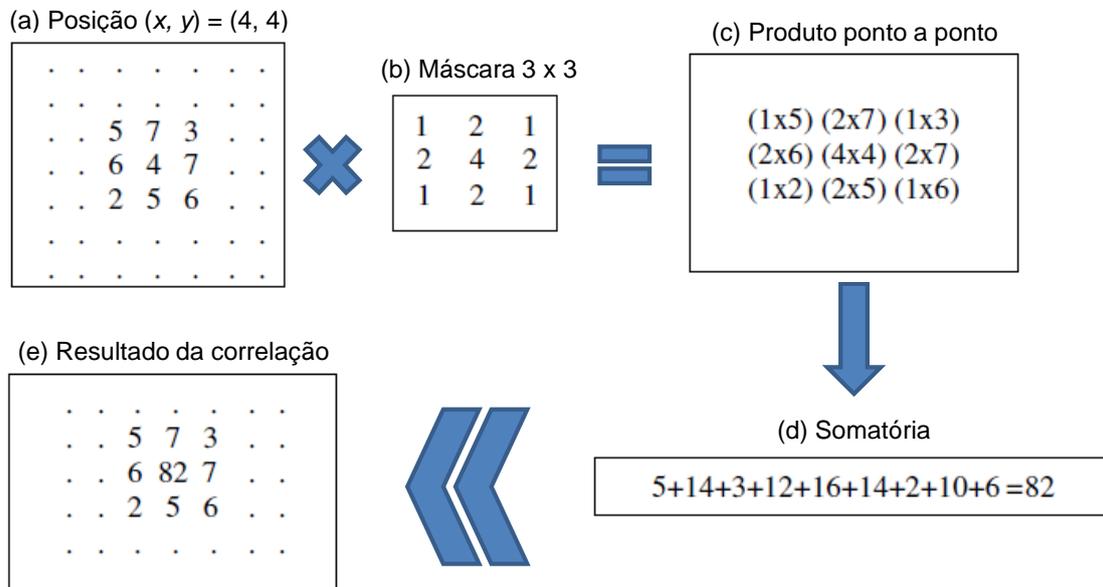


Figura 72 - Esquema da operação de correlação. Níveis de intensidade dos *pixels* na imagem original (a); máscara utilizada (b); produtos dos valores da máscara pelos valores dos *pixels* da imagem, ponto a ponto (c); somatória dos valores multiplicados (d); e, finalmente, o novo valor do *pixel* na imagem após a correlação (e). Fonte: [55].

A diferença entre a correlação e a convolução está na forma como a máscara é utilizada. Na correlação a operação é feita conforme indicado na Figura 72, já na operação de convolução a máscara é rotacionada em 180 graus. Note que, se a máscara for simétrica as operações de correlação e convolução são idênticas. As expressões que definem as operações de correlação e convolução são apresentadas nas equações (18) e (19), respectivamente.

$$g(x, y) = \sum_{i=-m/2}^{m/2} \sum_{j=-n/2}^{n/2} f(x+i, y+j) \times w(i, j) \quad (18)$$

$$g(x, y) = \sum_{i=-m/2}^{m/2} \sum_{j=-n/2}^{n/2} f(x-i, y-j) \times w(i, j) \quad (19)$$

onde w é a máscara utilizada no processo de correlação/convolução, e m e n são as dimensões da máscara. O tipo de filtro depende dos valores da máscara, no geral, as

máscaras são simétricas e, portanto, pode-se aplicar tanto a convolução como a correlação.

Filtros Gaussianos

Um filtro gaussiano tem os valores da máscara determinados a partir de uma função bidimensional gaussiana discreta, com média igual a zero e desvio padrão, conforme exibido na equação (20)

$$Gauss(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[\frac{-(x^2 + y^2)}{2\sigma^2}\right] \quad (20)$$

onde x e y são as posições na máscara e $Gauss(x, y)$ dá o valor a ser colocado na posição (x, y) da máscara. Os filtros gaussianos são utilizados para suavizar a imagem de forma ponderada e simétrica. Um exemplo de máscara gaussiana de tamanho 5 x 5 é apresentado na Figura 73.

$$Gauss(x,y) = 1/256 * \begin{array}{|c|c|c|c|c|} \hline 1 & 4 & 6 & 4 & 1 \\ \hline 4 & 16 & 24 & 16 & 4 \\ \hline 6 & 24 & 36 & 24 & 6 \\ \hline 4 & 16 & 24 & 16 & 4 \\ \hline 1 & 4 & 6 & 4 & 1 \\ \hline \end{array}$$

Figura 73 - Máscara gaussiana para um filtro do tipo passa-baixa. Fonte: [55].