

UNIVERSIDADE FEDERAL DE ITAJUBÁ

**PROGRAMA DE PÓS-GRADUAÇÃO
EM ENGENHARIA ELÉTRICA**

**Metodologia e Desenvolvimento
de Redes Neuro-Especialistas
Aplicadas à Navegação Autônoma**

Ana Sophia Cavalcanti Alves

Itajubá, Novembro de 2010

UNIVERSIDADE FEDERAL DE ITAJUBÁ

**PROGRAMA DE PÓS-GRADUAÇÃO
EM ENGENHARIA ELÉTRICA**

Ana Sophia Cavalcanti Alves

**Metodologia e Desenvolvimento
de Redes Neuro-Especialistas
Aplicadas à Navegação Autônoma**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Elétrica como parte dos requisitos para obtenção do Título de Mestre em Ciências em Engenharia Elétrica

Área de Concentração:

Automação e Sistemas Elétricos Industriais

Orientador:

Prof. Dr. Leonardo de Mello Honório

Novembro de 2010

Itajubá - MG

Resumo

Uma rede neural artificial é responsável por mapear as entradas e saídas de um dado sistema. Isto é realizado através da captura e representação do conhecimento em sua rede interna de pesos e seus respectivos neurônios contendo funções de ativações genéricas. Desta forma, toda a informação contida no sistema deveria ser mapeada e armazenada. Entretanto, em sistemas reais de grande complexidade é impossível ter todo o conhecimento mapeado. Logo a parte do conhecimento não mapeada gera uma perda de informação a respeito do sistema desejado que se traduz nos erros produzidos pelo modelo. Para melhorar a modelagem do sistema e gerar uma captura de informação mais precisa, este trabalho apresenta um modelo de rede neuro-especialista onde, para um conjunto de neurônios, as funções genéricas de ativação foram substituídas por um modelo impreciso do sistema. Desta forma, com parte da informação sobre o sistema já presente, a rede deve ter que capturar apenas o restante. Isso torna o aprendizado muito mais rápido e eficaz. Para demonstração a metodologia foi utilizada em um sistema de navegação autônoma.

Abstract

An artificial neural network is responsible for mapping a set of inputs/outputs of a given system. This is accomplished by capturing and representing the systems knowledge into its internal network weights and their respective generic activation functions. Thus, all information contained in the system should be mapped and stored. However, in complex real systems it is impossible to have all the knowledge mapped. The unmapped part of this knowledge will generate information to be lost which results in errors produced by the model. To improve the modeling of the system this paper presents a model of specialist neuro-network where, for a set of neurons, the generic activation functions will be replaced by an inaccurate, but representative, model of the system. Thus as some of the information is already present on the system, the network should have to capture only the remainder. This makes learning much faster and efficient. To demonstrate the methodology was used in a system of autonomous navigation.

Sumário

1	Introdução	13
2	Veículo Autônomo Inteligente	16
2.1	Constituição	17
2.1.1	Ação e Locomoção	17
2.1.2	Percepção	18
2.1.3	Controle e Inteligência.....	19
2.1.4	Comunicação	20
2.2	Navegação.....	20
2.2.1	Localização	20
2.2.2	Detecção	21
2.2.3	Planejamento da Trajetória.....	21
3	Sistema De Posicionamento	22
3.1	Modelo Cinemático.....	22
3.1.1	Representação da Posição do Veículo	23
3.1.2	Modelo Desenvolvido.....	23
3.2	DGPS	26
3.2.1	Fundamentos da Navegação por Satélite.....	27
3.2.2	Sistemas de Coordenadas	27
3.2.3	Precisão.....	29
3.3	Encoder	30
3.3.1	Odometria	32
3.3.2	Precisão.....	32
4	Processamento De Dados	34

4.1	Trajectoria.....	34
4.2	Filtro de Kalman	36
4.2.1	Filtro α - β	37
4.2.2	Filtro α - β Aplicado na Estimativa da Posiçao	38
4.2.3	Filtro α - β Aplicado na Estimativa do Angulo	39
5	Redes Neurais Artificiais.....	41
5.1	Conceitos Preliminares	41
5.1.1	Neuronio Biologico	42
5.1.2	Modelo de um Neuronio Artificial	43
5.1.3	Funçao de Ativaçao	44
5.2	Arquitetura da Rede	45
5.3	Aprendizagem	46
5.3.1	Aprendizado Supervisionado.....	47
5.4	Treinamento	48
5.4.1	Dinamica de Treinamento	48
5.4.2	Criterios de Parada do Treinamento	49
6	Redes Neuro-Especialistas	50
6.1	Modelo de um Neuronio Especialista.....	51
6.2	Funçao de Ativaçao.....	52
6.3	Arquitetura da Rede	52
6.4	Treinamento	53
6.5	Exemplo	53
7	Projeto De Uma RNA.....	57
7.1	Definiçao do Problema	58
7.2	Coleta de Dados para Treinamento e Testes.....	58

7.3	Pré-Processamento	65
7.4	Projeto da RNA.....	71
7.4.1	Definição dos Parâmetros da Rede Neural	71
7.4.2	Arquitetura da Rede Neural	75
7.4.3	Arquitetura da Rede Neuro-Especialista	76
7.5	Treinamento da RNA.....	78
7.5.1	Treinamento da Rede Neural	79
7.5.2	Treinamento da Rede Neuro-Especialista	80
7.6	Teste da RNA.....	81
7.6.1	Teste da Rede Neural.....	81
7.6.2	Teste da Rede Neuro-Especialista	86
7.7	Validação da RNA	91
8	Conclusão	94
9	Bibliografia.....	95
	Apêndice A Algoritmo Genético	98

Lista de Figuras

Figura 2.1 - Cinemática <i>Ackermann</i>	18
Figura 3.1 - Representação da Posição do Veículo	23
Figura 3.2 - Modelo Cinemático.....	24
Figura 3.3 - Sistema de Coordenadas	28
Figura 3.4 - <i>Encoder</i> de Movimento Linear	31
Figura 3.5 - <i>Encoder</i> de Movimento Angular	31
Figura 4.1 - Trajetória Percorrida pelo Veículo (Posição)	35
Figura 4.2 - Trajetória Percorrida pelo Veículo (Ângulo).....	35
Figura 4.3 - Filtro α - β Aplicado na Estimativa de Posição	39
Figura 4.4 - Filtro α - β Aplicado na Estimativa de Ângulo.....	40
Figura 5.1 - Neurônio Biológico	43
Figura 5.2 - Neurônio Artificial.....	44
Figura 5.3 - Função Sigmóide	45
Figura 5.4 - Arquitetura da Rede Neural	46
Figura 5.5 - Aprendizado Supervisionado.....	48
Figura 6.1 - Mapeamento do Sistema.....	51
Figura 6.2 - Neurônio Especialista	51
Figura 6.3 - Arquitetura Rede Neuro-Especialista	52
Figura 6.4 - Neurônio Especialista (Exemplo)	54
Figura 6.5 - Arquitetura Rede Neuro-Especialista (Exemplo)	54
Figura 6.6 - Função Seno.....	55
Figura 6.7 - Resultado Rede Neural (Exemplo)	56
Figura 6.8 - Resultado Rede Neuro-Especialista (Exemplo).....	56
Figura 7.1 - Etapas do Projeto de uma RNA	57
Figura 7.2 - Percurso 1 - Posição do Veículo	59
Figura 7.3 - Percurso 1 - Ângulo do Veículo	59
Figura 7.4 - Percurso 2 - Posição do Veículo	60
Figura 7.5 - Percurso 2 - Ângulo do Veículo	60
Figura 7.6 - Percurso 3 - Posição do Veículo	61
Figura 7.7 - Percurso 3 - Ângulo do Veículo	61

Figura 7.8 - Percurso 4 - Posição do Veículo	62
Figura 7.9 - Percurso 4 - Ângulo do Veículo	62
Figura 7.10 - Percurso 5 - Posição do Veículo	63
Figura 7.11 - Percurso 5 - Ângulo do Veículo	63
Figura 7.12 - Percurso 6 - Posição do Veículo	64
Figura 7.13 - Percurso 6 - Ângulo do Veículo	64
Figura 7.14 - Percurso 1 - Posição do Veículo com Filtro	65
Figura 7.15 - Percurso 1 - Ângulo do Veículo com Filtro.....	66
Figura 7.16 - Percurso 2 - Posição do Veículo com Filtro	66
Figura 7.17 - Percurso 2 - Ângulo do Veículo com Filtro.....	67
Figura 7.18 - Percurso 3 - Posição do Veículo com Filtro	67
Figura 7.19 - Percurso 3 - Ângulo do Veículo com Filtro.....	68
Figura 7.20 - Percurso 4 - Posição do Veículo com Filtro	68
Figura 7.21 - Percurso 4 - Ângulo do Veículo com Filtro.....	69
Figura 7.22 - Percurso 5 - Posição do Veículo com Filtro	69
Figura 7.23 - Percurso 5 - Ângulo do Veículo com Filtro.....	70
Figura 7.24 - Percurso 6 - Posição do Veículo com Filtro	70
Figura 7.25 - Percurso 6 - Ângulo do Veículo com Filtro.....	71
Figura 7.26 - Definição do Número de Neurônios na Camada Escondida	73
Figura 7.27 - Excesso de Treinamento	74
Figura 7.28 - Arquitetura da Rede Neural	76
Figura 7.29 - Arquitetura da Rede Neuro-Especialista.....	77
Figura 7.30 - Percurso 1 - Treinamento (Ângulo).....	79
Figura 7.31 - Percurso 1 - Treinamento (Posição).....	79
Figura 7.32 - Percurso 1 - Treinamento (Ângulo).....	80
Figura 7.33 - Percurso 1 - Treinamento (Posição).....	80
Figura 7.34 - Percurso 2 - Teste (Ângulo).....	81
Figura 7.35 - Percurso 2 - Teste (Posição)	82
Figura 7.36 - Percurso 3 - Teste (Ângulo).....	82
Figura 7.37 - Percurso 3 - Teste (Posição)	83
Figura 7.38 - Percurso 4 - Teste (Ângulo).....	83
Figura 7.39 - Percurso 4 - Teste (Posição)	84
Figura 7.40 - Percurso 5 - Teste (Ângulo).....	84
Figura 7.41 - Percurso 5 - Teste (Posição)	85

Figura 7.42 - Percurso 6 - Teste (Ângulo).....	85
Figura 7.43 - Percurso 6 - Teste (Posição)	86
Figura 7.44 - Percurso 2 - Teste (Ângulo).....	86
Figura 7.45 - Percurso 2 - Teste (Posição)	87
Figura 7.46 - Percurso 3 - Teste (Ângulo).....	87
Figura 7.47 - Percurso 3 - Teste (Posição)	88
Figura 7.48 - Percurso 4 - Teste (Ângulo).....	88
Figura 7.49 - Percurso 4 - Teste (Posição)	89
Figura 7.50 - Percurso 5 - Teste (Ângulo).....	89
Figura 7.51 - Percurso 5 - Teste (Posição)	90
Figura 7.52 - Percurso 6 - Teste (Ângulo).....	90
Figura 7.53 - Percurso 6 - Teste (Posição)	91

Lista de Tabelas

Tabela 2.1 - Sensores.....	19
Tabela 6.1 - Comparação Numérica.....	55
Tabela 7.1 - Relação Entre Sensores e Dados Coletados	58
Tabela 7.3 - Comparação dos Resultados para a Posição (Erro Total)	92
Tabela 7.4 - Comparação dos Resultados para a Posição (Erro Médio)	92
Tabela 7.5 - Comparação dos Resultados para a Posição (Erro Máximo)	92
Tabela 7.6 - Comparação dos Resultados para o Ângulo (Erro Total).....	93
Tabela 7.7 - Comparação dos Resultados para o Ângulo (Erro Médio).....	93
Tabela 7.8 - Comparação dos Resultados para o Ângulo (Erro Máximo)	93

Lista de Abreviaturas e Siglas

VI	Veículo Inteligente
ABS	Sistema de Anti Bloqueio de Frenagem
IA	Inteligência Artificial
IEEE	Instituto de Engenheiros Eletricistas e Eletrônicos
GPS	Sistema de Posicionamento Global
DGPS	Sistema de Posicionamento Global Diferencial
TOA	Tempo de Chegada
ECEF	Centrado na Terra, Fixado na Terra
RNA	Rede Neural Artificial
AG	Algoritmo Genético

1 INTRODUÇÃO

Veículos Inteligentes (VI) não são mais tecnologias visionárias que poderiam estar presentes em situações reais em um futuro não muito distante. Desde aplicações mais simples de auxílio ao motorista, tais como o controle eletrônico de velocidade ou sistemas de frenagem antitravamento (ABS), até as mais complexas, tais como sistemas militares autônomos de alta tecnologia ou sistemas automáticos de estacionamento de veículos, estas tecnologias estão se tornando cada vez mais comuns na atualidade (1).

O termo “Veículo Inteligente” é de certa forma, muito genérico. Veículos podem ser terrestres, aéreos ou aquáticos, assim como a designação inteligente pode ser empregada com diferentes significados (2). Nesta proposta o termo VI será utilizado para designar veículos terrestres não-holonômicos (tipo carro) com capacidade de perceber o ambiente onde ele está, através de sensores, e tomar decisões estratégicas autônomas de navegação e de controle de direção.

No meio acadêmico e industrial há muita pesquisa sendo realizada nesta área visando tanto o desenvolvimento de hardware quanto de algoritmos baseados em técnicas de Inteligência Artificial (IA). Este crescente interesse pode ser observado através dos diversos workshops, conferências, publicações e desafios promovidos pelo IEEE, IEE e outras organizações relevantes.

A análise das publicações nesta área mostra que os principais campos de pesquisa em VI são:

1. **Instrumentação de Alta Precisão** – Atualmente, a instrumentação disponível para aplicações automotivas é de alta precisão. Os Sistemas de Posicionamento Global Diferencial (DGPS), por exemplo, atingem precisão sub-métrica (2), podendo ser empregados para detectar com precisão a posição do veículo. Esta precisão e a confiabilidade dos dados fornecidos pelos sensores aumentam quando se utiliza técnicas de registro simultâneo e fusão de sensores (3) (4), temas muito pesquisados e em desenvolvimento.

2. **Técnicas de Inteligência Artificial** – Poderosos algoritmos utilizando técnicas de IA são capazes de lidar com dados imprecisos e incompletos, permitindo a modelagem do sistema de uma maneira muito confiável (5) (6). Através da utilização destes algoritmos é possível receber dados de diferentes sensores, identificar claramente a situação atual, planejar determinadas decisões para executar uma tarefa, fazer re-planejamento caso uma situação inesperada ocorra, prever certas condições evitando situações desastrosas, entre outras.

3. **Visão Computacional** – Trata-se de umas das áreas mais pesquisadas em IA. A visão computacional integra a utilização de *hardware* de alto desempenho com *software* de tempo real. Pesquisas neste campo (7) (8) oferecem um largo espectro de características, tais como, visão *stereo*, movimento, aparência, identificação de bordas, entre outras.

Dentro deste cenário uma funcionalidade básica é o veículo conhecer sua posição absoluta no mundo. Para isso um DGPS associado a um Filtro de Kalman, ou sua variante convergida, o Filtro α - β é suficiente. Entretanto, existem várias situações onde o sinal de DGPS falha colocando em risco a integridade do veículo. Neste caso somente a cinemática do veículo não é capaz de apresentar resultados satisfatórios para sua localização. O motivo disso é que a cinemática é uma aproximação do problema real que é muito precisa para pequenas distâncias, mas com um erro acumulativo muito grande para que o veículo seja capaz de se orientar apenas por esta metodologia. Para melhorar o modelo matemático da cinemática do veículo este documento propõe o desenvolvimento de uma rede neural especializada.

Basicamente uma rede neural é um mapeador de entradas e saídas. Ou seja, em uma situação de treinamento supervisionado, um conjunto de dados representando diversas situações de entradas e saídas de um sistema de um mundo real são passadas para um algoritmo de treinamento que tenta capturar este conhecimento em forma dos pesos internos na rede neural.

Em uma situação ideal, se todo o conhecimento for capturado, quando novas entradas forem apresentadas para a rede, uma saída idêntica ao sistema modelado será apresentada. Entretanto, esta situação ideal quase nunca se apresenta em problemas de grande

complexidade, logo o conhecimento não preenchido ou mapeado pela rede neural retorna através dos erros fornecidos.

Uma forma de melhorar este mapeamento é apresentar, durante o treinamento da rede, parte do conhecimento que esta deve adquirir. Desta forma, com parte do conhecimento já mapeado, a rede deverá apenas aprender o restante. No caso da cinemática, todas as fórmulas que representam grande parte do conhecimento e comportamento do veículo são previamente conhecidas. Apesar de mostrarem um erro de longo prazo, preenchem quase todo o conhecimento sobre o sistema. Desta forma, incluir isso em uma rede deixa o treinamento mais eficiente, uma vez que nesta situação apenas o restante do conhecimento do processo é que deve ser capturado.

Para desenvolver esta idéia, foi projetada uma rede neuro-especialista, onde a função de ativação de parte dos neurônios são as equações cinemáticas do veículo. Com esta informação presente dentro da rede, os pesos restantes conseguem reduzir a perda de informação e, desta forma, mapear melhor o conhecimento real do comportamento do veículo.

Para demonstrar esta idéia, esta dissertação está dividida da seguinte forma: Nos Capítulos 2 e 3 será mostrado uma revisão de veículos autônomos junto com seu modelo matemático. Nos Capítulos 4, 5 e 6, serão apresentados o Filtro α - β , Redes Neurais Artificiais e Redes Neuro-Especialistas. Por fim, os Capítulos 7 e 8 mostram os resultados e as conclusões.

2 VEÍCULO AUTÔNOMO INTELIGENTE

Dirigir é uma tarefa exigente que requer muito mais do que um comportamento robótico repetitivo. Uma condução bem sucedida requer atenção, alerta e respostas instintivas de diferentes condições da estrada, dos obstáculos e de segurança. Todas essas questões exigem um fluxo constante de decisões.

Para melhorar as capacidades do condutor, a computação autônoma pode auxiliar significativamente na tomada de decisões. Os fabricantes têm integrado com sucesso câmbio automático, *cruise control* (sistema que mantém a velocidade do veículo) e estacionamento totalmente automático em carros comerciais. Veículos operados remotamente têm realizado com sucesso missões espaciais em ambientes hostis por mais de duas décadas. A indústria tem utilizado sistemas robóticos para a movimentação de materiais entre as localidades em linhas de produção durante meio século. Essas conquistas foram principalmente devido à medição precisa, à resposta controlada e à atuadores precisos e confiáveis com um recurso de acionamento manual para garantir a segurança (9).

Veículos autônomos inteligentes são aqueles que operam sem a intervenção humana. Esses veículos têm a capacidade de perceber o ambiente e tomar decisões para se adaptar a diferentes condições.

O modo como uma pessoa dirige um veículo se difere do modo como um sistema autônomo atua sobre o veículo. Essa diferença de comportamento ocorre principalmente devido ao modo que cada um percebe o ambiente. Um condutor humano opera em um mundo contínuo, em contrapartida, o sistema de condução da máquina não funciona em um mundo contínuo, pois sua percepção de mundo é discretizada (10).

Outra diferença entre os motoristas humanos e o sistema autônomo é o domínio do conhecimento de condução que cada um possui. A diferença está na capacidade da uma pessoa interpretar mudanças no estado do ambiente com base em um conhecimento prévio de como esse ambiente se comporta.

Para desenvolver adequadamente projetos na área de veículos autônomos é necessário que primeiramente ocorra uma análise a respeito dos aspectos relativos à constituição do

veículo, que serão discutidos na Seção 2.1. Além disso, o sistema deve ser capaz de gerar um plano de ação e executar este plano. Na Seção 2.2 serão discutidos os problemas advindos do processo de navegação (mover de forma segura de um lugar a outro do ambiente), uma das importantes tarefas que um carro autônomo deve executar.

2.1 Constituição

Os seguintes aspectos relativos à constituição de um veículo autônomo devem ser considerados (11):

- Locomoção - Como o veículo irá se deslocar no ambiente;
- Percepção - Como o veículo irá perceber o ambiente e monitorar parâmetros próprios;
- Controle e Inteligência - Como o veículo irá transformar suas percepções e conhecimentos prévios adquiridos em ações, ou seqüências de ações, a serem executadas;
- Comunicação - Como o veículo irá se comunicar com um operador humano, ou mesmo com outros veículos.

2.1.1 Ação e Locomoção

O sistema de locomoção é caracterizado pelo comportamento do veículo durante seu deslocamento no ambiente. Em um carro tradicional o motor é responsável por girar as rodas de um mesmo eixo movimentando o veículo e o volante é responsável por girar a barra de direção que irá direcioná-lo. Este tipo de veículo terá um comportamento seguindo a cinemática *Ackermann*, como mostra a Figura 2.1.

Um veículo autônomo deve ser equipado com atuadores, que serão responsáveis pela execução de suas ações e, principalmente, pelo seu deslocamento no ambiente. Os atuadores são capazes de mudar a marcha, a direção e frear o veículo.

O conhecimento do modelo cinemático (equações matemáticas que descrevem o movimento do veículo) é essencial para que se possa aplicar um mecanismo adequado de controle, ou mesmo, simular o funcionamento do veículo.

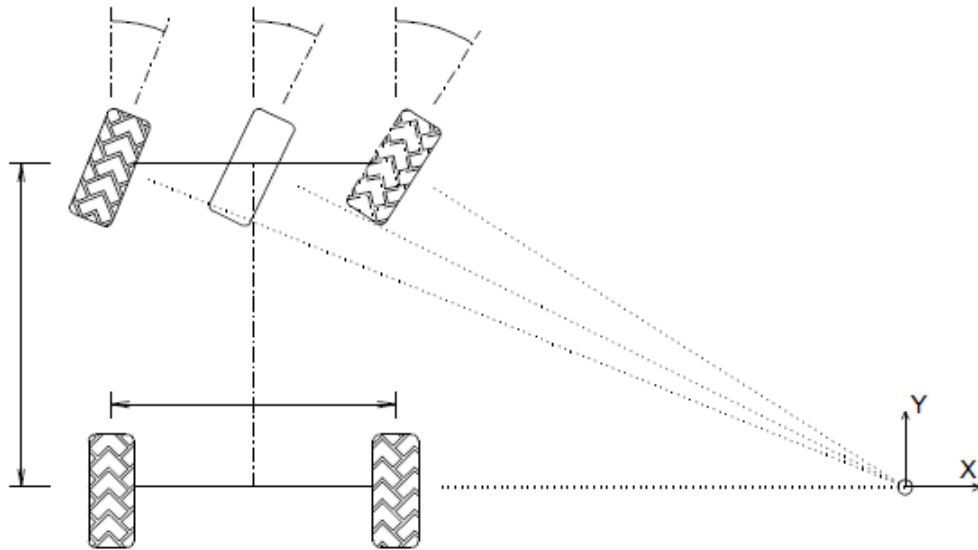


Figura 2.1 - Cinemática *Ackermann*

2.1.2 Percepção

A percepção é um dos componentes principais do veículo autônomo, pois é através dos sensores que o veículo monitora o ambiente e a si mesmo, garantindo assim, um maior nível de autonomia e robustez ao sistema.

O uso adequado dos sensores, bem como a correta interpretação dos dados fornecidos por estes, permitirá ao sistema perceber o ambiente, podendo assim, reagir a ele. Os sensores podem ser integrados (fusão sensorial) de forma a explorar a complementaridade entre eles, que possuem diferentes escalas de sensibilidade e confiabilidade. A Tabela 2.1 apresenta alguns exemplos de sensores utilizados em veículos autônomos e suas funções.

Sensor	Função
<i>Encoder</i> (de rodas)	Mede o número de rotações das rodas, permitindo assim, obter informações do espaço percorrido pelo veículo.
Sensor <i>Laser</i>	Detecta obstáculos no campo de atuação do sensor informando a posição em que cada um se encontra.
Sonar (ultra-som)	Determina a distância entre o sensor e um obstáculo através da reflexão sonora.

Bússola Eletrônica	Identifica a orientação do veículo em relação ao campo magnético da terra.
GPS	Identifica a posição absoluta do veículo no globo terrestre, baseando-se na rede de satélites GPS.
Visão Artificial	O uso de imagens permitem que sejam adquiridas descrições do ambiente e que se implementem técnicas de determinação de posicionamento, estimativa de deslocamento e detecção de obstáculos.
Unidade de Medida Inercial	Mede a velocidade, a orientação e forças gravitacionais, usando uma combinação de acelerômetros e giroscópios.

Tabela 2.1 - Sensores

2.1.3 Controle e Inteligência

O sistema de controle que inclui “comportamentos inteligentes” torna o veículo mais confiável, menos dependentes da intervenção de seres humanos e capaz de lidar com diversas situações podendo resolver e executar tarefas complexas.

O sistema de controle deve realizar tarefas como as descritas a seguir (11):

- Garantir a preservação da integridade física dos seres humanos, bem como não causar danos ao ambiente onde está inserido;
- Garantir a preservação da integridade física do veículo evitando colisões contra pessoas e objetos (obstáculos estáticos ou móveis);
- Garantir a manutenção de informações que permita uma melhor execução das tarefas, o que pode incluir: construção e atualização de um mapa do ambiente, determinação da localização do robô em relação a este mapa, definição de trajetórias que permitam que ele se desloque de um ponto a outro (considerando o mapa disponível);
- Integrar as informações provenientes de diferentes sensores, interpretando estas informações e levando em consideração problemas de ruído e erros associados às mesmas;

- Gerar os comandos, na seqüência correta, de modo a realizar passo-a-passo as tarefas que lhe são atribuídas;
- Definir soluções alternativas para situações imprevistas que se apresentem e que por ventura possam prejudicar a execução das ações que estavam planejadas;
- Ter a capacidade de se adaptar, aprender com as experiências passadas e corrigir seus erros.

2.1.4 Comunicação

O sistema de comunicação é usado de forma a estabelecer uma troca de informações homem-máquina ou entre veículos. A interface homem-máquina possibilita ao usuário apresentar os objetivos ao sistema autônomo e também permite uma supervisão em tempo real do sistema. A comunicação entre veículos é de suma importância em situações relacionadas ao controle de fluxo em comboios nas estradas.

2.2 Navegação

O sistema de navegação fornece os meios para que um veículo autônomo se mova de forma segura de um local a outro do ambiente. O problema de navegação de um veículo autônomo está dividido em três tarefas básicas (10):

- Localização - saber onde o veículo está;
- Detecção - ver o que está ao redor do veículo;
- Planejamento da trajetória - determinar como chegar a um destino.

2.2.1 Localização

O principal problema que dificulta a navegação é o problema da localização. Sem saber a posição do veículo em relação a sua representação do ambiente (mapa), pouquíssimos sistemas são capazes de controlar o veículo de forma adequada.

Utilizando as leituras atuais dos sensores, o veículo deve ser capaz de determinar e manter atualizada a sua posição e orientação em relação a este ambiente, mesmo que os sensores apresentem erro e/ou ruído.

A abordagem principal deste trabalho é desenvolver um sistema de posicionamento preciso que possibilite a navegação do veículo autônomo em ambientes *indoor* e *outdoor*. Esse sistema será descrito no próximo capítulo e aperfeiçoado (como proposta deste trabalho) nos capítulos seguintes.

2.2.2 Detecção

O sistema de navegação utiliza as informações sensoriais locais para controlar diretamente a movimentação do veículo. Estas informações possibilitam que o veículo opere em ambientes desconhecidos e/ou imprevisíveis. A cada instante de tempo o veículo utiliza seus sensores para localizar os obstáculos mais próximos, e então planeja a próxima parte de seu caminho. O sistema de detecção de um veículo autônomo deve ser capaz de distinguir entre regiões seguras e inseguras.

2.2.3 Planejamento da Trajetória

O planejamento de trajetória deve encontrar uma rota de um local do ambiente para outro, onde esta rota deve seguir uma trajetória livre de colisões durante o período de tempo em que o veículo se movimentar e o movimento planejado deve ser consistente com as restrições cinemáticas do veículo. Um algoritmo de planejamento de trajetória deve garantir um caminho até o destino, ou indicar se o destino é inacessível.

3 SISTEMA DE POSICIONAMENTO

Determinar a posição precisa de um veículo autônomo é uma tarefa extremamente desafiadora. Para poder navegar de forma eficiente, o veículo autônomo precisa ser capaz de determinar sua posição de forma rápida e precisa, contudo, não há nenhuma maneira direta de medir sua posição instantaneamente. Estimativas de posição relativamente precisas podem ser obtidas integrando as informações cinemáticas do veículo com os seus sensores.

A solução do problema de posicionamento é a base para que toda a navegação seja validada, sem uma posição com um nível mínimo de precisão as tarefas executadas pelo veículo autônomo acabam ficando limitadas.

O processo de compreensão dos movimentos de um veículo começa com a descrição da contribuição de cada roda para seu movimento. Cada uma das rodas contribui para o movimento e, ao mesmo tempo, impõe restrições. Na Seção 3.1, é apresentada uma descrição do movimento do veículo através de seu modelo cinemático.

Os sensores são utilizados para obter informações sobre o posicionamento local e global do veículo no ambiente. Dois tipos de sensores são utilizados: um DGPS (*Differential Global Positioning System*) para a obtenção de coordenadas por meio de satélites, descrito na Seção 3.2 e um *encoder*, descrito na Seção 3.3, responsável por monitorar e enviar informações sobre os movimentos das rodas.

3.1 Modelo Cinemático

Um modelo cinemático oferece uma descrição matemática do movimento de veículos, sem considerar as forças que afetam o movimento. As equações de movimento são baseadas puramente em relações geométricas que regem o sistema (12).

3.1.1 Representação da Posição do Veículo

Neste trabalho, o ambiente em que o veículo está inserido é considerado bidimensional, onde existe apenas movimentação no plano horizontal representado pelos eixos X e Y , como mostra a Figura 3.1. Para descrever o movimento do veículo é necessário um sistema de coordenadas globais onde (X, Y) representa a localização do centro de gravidade do veículo no ambiente, enquanto ψ descreve a orientação do veículo.

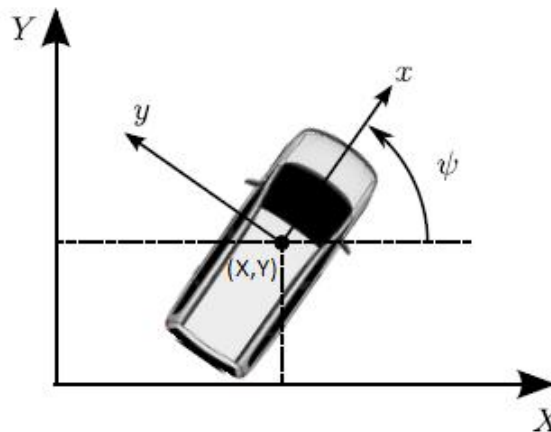


Figura 3.1 - Representação da Posição do Veículo

Para a orientação do veículo em relação a esse plano é necessário um sistema de coordenadas locais onde o eixo x representa a movimentação longitudinal do veículo, o eixo y representa sua movimentação lateral e a origem do sistema de coordenadas é o centro de gravidade do veículo.

3.1.2 Modelo Desenvolvido

Considere um modelo bicíclico do veículo, como mostrado na Figura 3.2. No modelo bicíclico, as duas rodas dianteiras, esquerda e direita, são representadas por uma única roda no ponto A . Do mesmo modo as rodas traseiras são representadas por uma roda traseira central no ponto B . O ângulo da direção para a roda dianteira é representada por δ . O centro de gravidade do veículo está no ponto C . A distância dos pontos A e B ao centro de gravidade do veículo são l_f e l_r , respectivamente e a distância entre eixos do veículo é $L = l_f + l_r$.

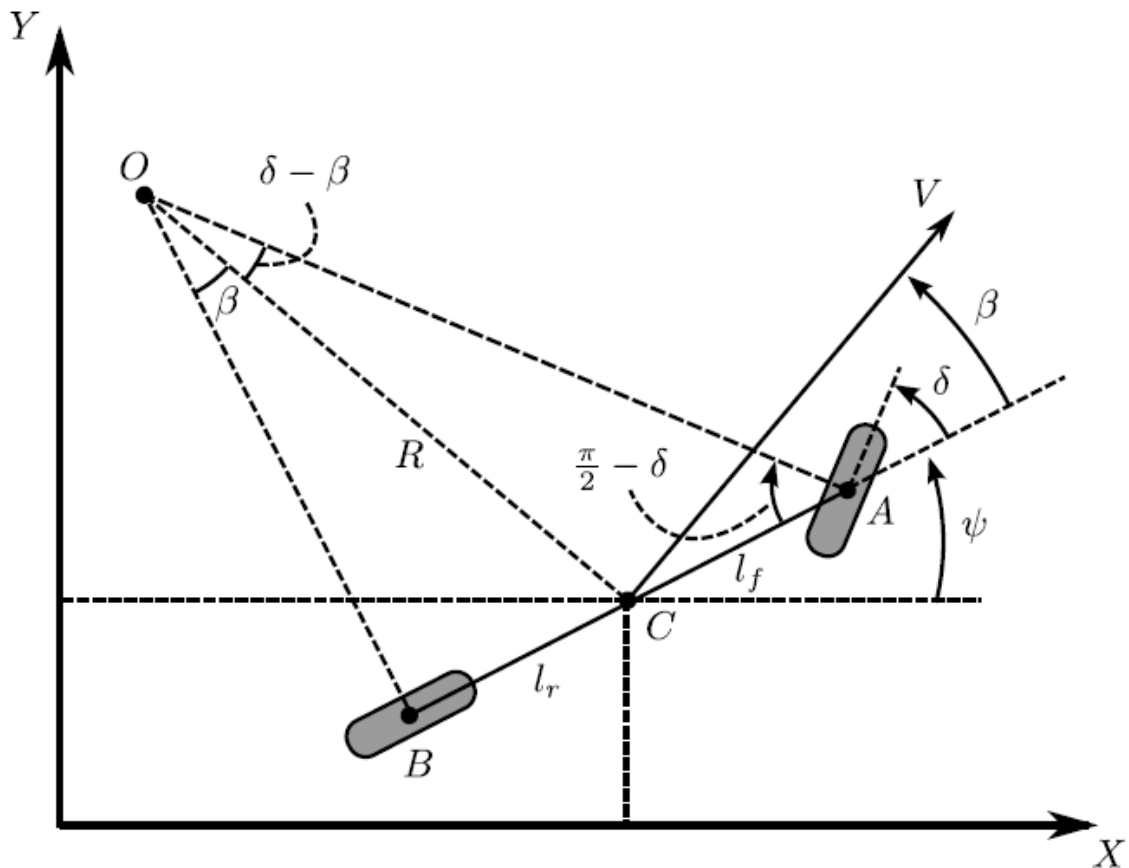


Figura 3.2 - Modelo Cinemático

Como dito anteriormente, (X, Y) representa a localização do centro de gravidade do veículo, enquanto ψ descreve a orientação do veículo. A velocidade no centro de gravidade do veículo é denotada por V e faz um ângulo de β com o eixo longitudinal do veículo. O ângulo β é chamado de ângulo de deslizamento do veículo.

A principal consideração feita no desenvolvimento do modelo cinemático é que os vetores velocidade nos pontos A e B estão na direção da orientação das rodas dianteiras e traseiras, respectivamente. Em outras palavras, o vetor velocidade na roda dianteira faz um ângulo δ com o eixo longitudinal do veículo. Isso é equivalente a assumir que os ângulos de escorregamento em ambas as rodas são zero. O ângulo de escorregamento numa roda é o ângulo entre a direção em que a roda aponta e a direção real em que a roda se desloca. Esta é uma suposição razoável para o movimento de baixa velocidade do veículo, por exemplo, para velocidades inferiores a 5 m/s. À baixas velocidades, a força lateral gerada sobre os pneus é pequena (12).

Para um pequeno período de tempo a velocidade pode ser considerada constante e o ângulo das rodas dianteiras do veículo se mantém fixos. Nessas condições o veículo realiza um movimento circular com um raio de curvatura constante, chamado de curva de estado estacionário (13). O ponto O é o centro instantâneo de curvatura do veículo e é definido pela interseção das linhas AO e BO , que são perpendiculares à orientação das duas rodas.

O raio da trajetória do veículo R é definido pelo comprimento da linha OC que liga o centro de gravidade C ao centro instantâneo de curvatura O . A velocidade no centro de gravidade é perpendicular a linha OC .

Aplicando a regra dos senos no triângulo OCA ,

$$\frac{\sin(\delta - \beta)}{l_f} = \frac{\sin(\frac{\pi}{2} - \delta)}{R} \quad (3.1)$$

Da Eq. 3.1,

$$\frac{\sin(\delta) \cos(\beta) - \sin(\beta) \cos(\delta)}{l_f} = \frac{\cos(\delta)}{R} \quad (3.2)$$

Multiplicando os dois lados da Eq. 3.2 por $\frac{l_f}{\cos(\delta)}$,

$$\tan(\delta) \cos(\beta) - \sin(\beta) = \frac{l_f}{R} \quad (3.3)$$

Aplicando a regra dos senos no triângulo OCB ,

$$\frac{\sin(\beta)}{l_r} = \frac{\sin(\frac{\pi}{2})}{R} \quad (3.4)$$

Logo,

$$\sin(\beta) = \frac{l_r}{R} \quad (3.5)$$

Somando as Eq. 3.3 e 3.5,

$$\tan(\delta) \cos(\beta) = \frac{l_f + l_r}{R} \quad (3.6)$$

Se assumirmos que o raio da trajetória do veículo muda lentamente, devido à baixa velocidade do veículo, então a taxa de mudança da orientação do veículo deve ser igual à velocidade angular do veículo (12).

Portanto,

$$\dot{\psi} = \frac{V}{R} \quad (3.7)$$

Usando a equação 3.7, a equação 3.6 pode ser reescrita por

$$\dot{\psi} = \frac{V}{l_f + l_r} \cos(\beta) \tan(\delta) \quad (3.8)$$

As equações gerais do movimento são, portanto, dadas por

$$\dot{X} = V \cos(\psi + \beta) \quad (3.9)$$

$$\dot{Y} = V \sin(\psi + \beta) \quad (3.10)$$

$$\dot{\psi} = \frac{V}{l_f + l_r} \cos(\beta) \tan(\delta) \quad (3.11)$$

O ângulo β pode ser obtido das equações 3.3 e 3.5,

$$\beta = \tan^{-1} \left(\frac{l_r}{l_f + l_r} \tan(\delta) \right) \quad (3.12)$$

As equações obtidas anteriormente representam o modelo cinemático do veículo e são utilizadas para descrever seu movimento neste trabalho.

3.2 DGPS

Localização é um dos problemas fundamentais na robótica móvel autônoma. Utilizando as leituras dos sensores, o veículo autônomo deve ser capaz de determinar e manter atualizada a sua posição e orientação em relação ao ambiente.

O DGPS é um dos sensores utilizados para obter uma informação precisa sobre a posição do veículo autônomo.

3.2.1 Fundamentos da Navegação por Satélite

O sistema de navegação global (GPS - *Global Positioning System*) foi criado pelo Departamento de Defesa dos Estados Unidos no programa de satélites NAVSTAR (*Navigation Satellite with Time and Ranging*). O sistema inclui 24 ou mais satélites ativos uniformemente dispersos em torno de seis órbitas circulares com quatro ou mais satélites cada. São necessários no mínimo quatro satélites para determinar a posição de um receptor em qualquer lugar na superfície da Terra (14).

O GPS utiliza o conceito de tempo de chegada (TOA - *Time of arrival*) para determinar a posição do receptor. Este conceito implica em medir o tempo que um sinal emitido por um emissor, numa posição conhecida, demora para alcançar o receptor (15).

Este intervalo de tempo, denominado tempo de propagação do sinal é então multiplicado pela velocidade do sinal, por exemplo, velocidade da luz, para obter a distância emissor-receptor. Ao medir o tempo de propagação do sinal de transmissão de múltiplos emissores (satélites), em locais conhecidos, o receptor pode determinar a sua posição.

Uma vez conhecidas as distâncias a cada um dos satélites, tem-se que calcular as coordenadas da posição.

3.2.2 Sistemas de Coordenadas

O sistema de coordenadas geodésicas, representada na Figura 3.3 fornece um modelo de forma elipsoidal da Terra. Neste modelo, seções transversais da terra paralelas ao plano equatorial são circulares (Paralelos). A seção equatorial da Terra tem 6.378,137 km de raio (eixo maior a), que é o raio médio equatorial da Terra. Seções transversais da Terra normal ao plano equatorial são elipsoidais (Meridianos). A seção polar da Terra tem 6.356,752 km de raio (eixo menor b) (15).

A longitude λ é a mais curta distância angular ao longo do equador a partir do primeiro meridiano até a linha formada entre o ponto C no centro da Terra e o ponto P de referência. Ela é medida na direção leste ou na direção oeste de 0° a 180° .

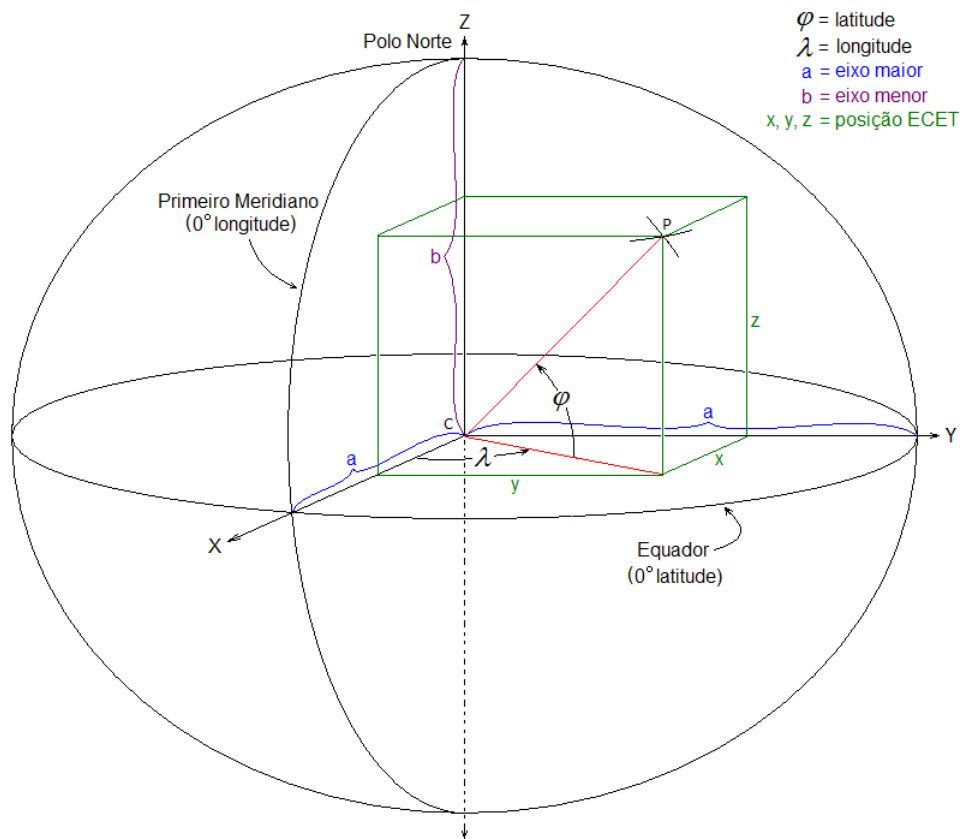


Figura 3.3 - Sistema de Coordenadas

A latitude φ é a distância angular norte ou sul a partir do equador até a linha formada entre o ponto C no centro da Terra e o ponto P de referência, medida ao longo do meridiano em graus de 0° (do equador) até 90° (aos pólos).

Para formular os modelos matemáticos do problema da navegação por satélite, é necessário escolher um sistema de coordenadas de referência. O sistema adotado é conhecido como ECEF (*Earth-centered Earth-fixed*), também representado na Figura 3.3. O sistema ECEF é um sistema de coordenadas cartesianas que tem origem no centro de massa da Terra e gira com ela.

As equações de transformação de coordenadas geodésicas em coordenadas cartesianas do sistema ECEF são fornecidas a seguir. Dados os parâmetros geodésicos λ , φ , e h , podemos calcular x , y , e z da seguinte forma (15):

$$x = \frac{a \cos \lambda}{\sqrt{1 + (1 - e^2) \tan^2 \varphi}} + h \cos \lambda \cos \varphi \quad (3.13)$$

$$y = \frac{a \sin \lambda}{\sqrt{1 + (1 - e^2) \tan^2 \phi}} + h \sin \lambda \cos \phi \quad (3.14)$$

$$z = \frac{a(1 - e^2) \sin \phi}{\sqrt{1 - e^2 \sin^2 \phi}} + h \sin \phi \quad (3.15)$$

onde e é definido como:

$$e = \sqrt{1 - \frac{b^2}{a^2}} .$$

Para transformar as coordenadas ECEF em coordenadas globais (descritas na Subseção 3.1.1), escolhe-se um ponto de referência no local. Se o ponto de referência escolhido for encontrado em $p = (x_p, y_p, z_p)$ e o receptor em $r = (x_r, y_r, z_r)$, o vetor que aponta do ponto de referência para o receptor será $\vec{r} - \vec{p}$.

3.2.3 Precisão

A precisão com que um receptor pode determinar sua posição ou velocidade, ou sincronizar a hora do sistema GPS, depende de uma complexa interação de vários fatores. Em geral, o desempenho da precisão do GPS depende do número e da geometria dos satélites usados e da qualidade da medição da distância receptor-satélite (*pseudorange*) (15).

Alguns fatores que podem influenciar na precisão do GPS (16):

- Atraso na Ionosfera e Troposfera - O sinal de satélite reduz a velocidade quando atravessa a atmosfera. O GPS usa um padrão de sistema embutido, que calcula parcialmente o tempo comum de demora, para corrigir este tipo de erro.
- Sinal *Multipath* - Isto ocorre quando o sinal de GPS é refletido em objetos, como edifícios altos ou superfícies com pedras grandes, antes de localizar o receptor. Isto aumenta o tempo de viagem do sinal, causando erros.
- Erros no receptor de relógio - O relógio embutido de um receptor não é tão preciso quanto o relógio atômico dos satélites de GPS. Então, pode haver erros de cronometragem muito leves.
- Erros orbitais - Também conhecido como erros de efemérides, são inexatidões do local informado do satélite.

- Número de satélites visíveis - Quanto maior o número de satélites que um receptor de GPS puder captar, tanto melhor será a precisão. Edifícios, terrenos, interferência eletrônica, ou às vezes até mesmo folhagem densa, podem bloquear notoriamente a recepção, causando erros de posição ou possivelmente nenhuma leitura de posição. Tipicamente, unidades de GPS não funcionarão em lugar fechado, subaquático ou subterrâneo.

A precisão de um receptor GPS é de cerca de 15 metros. Existem muitas aplicações, no entanto, em que os níveis de exigência da precisão vão além do que um receptor GPS pode oferecer. Para essas aplicações, existe o GPS Diferencial (DGPS).

GPS Diferencial é uma técnica para reduzir o erro na posição GPS usando dados adicionais a partir de um receptor GPS de referência em uma posição conhecida. A forma mais comum de DGPS envolve determinar os efeitos combinados das efemérides, ionosfera e erros do relógio em uma estação de referência e transmitir as correções da distância receptor-satélite (*pseudorange*) em tempo real para o receptor, que aplica as correções no processo para determinar a sua posição (14). A precisão de um receptor DGPS é de cerca de 1 metro.

3.3 Encoder

Os *encoders* são transdutores de movimento capazes de converter movimentos lineares ou angulares em informações elétricas (pulsos). Armazenando precisamente os movimentos das rodas utilizando *encoders*, o veículo pode determinar o quanto ele se locomoveu e o quanto ele rotacionou. O número de pulsos por rotação da roda determina a resolução do *encoder*. Conhecendo o raio da roda é possível calcular o movimento do veículo baseando-se nas informações obtidas do *encoder*.

O veículo autônomo em questão utiliza dois *encoders*: um *encoder* de movimento linear e um *encoder* de movimento angular.

Encoder de movimento linear:

O sensor *ABS* (*Anti-lock Braking System*) do veículo, mostrado na Figura 3.4, foi adaptado como *encoder*. Ele está localizado nas rodas traseiras do veículo e é utilizado para determinar seu espaço percorrido e apresenta uma resolução de 29 pulsos por rotação da roda.

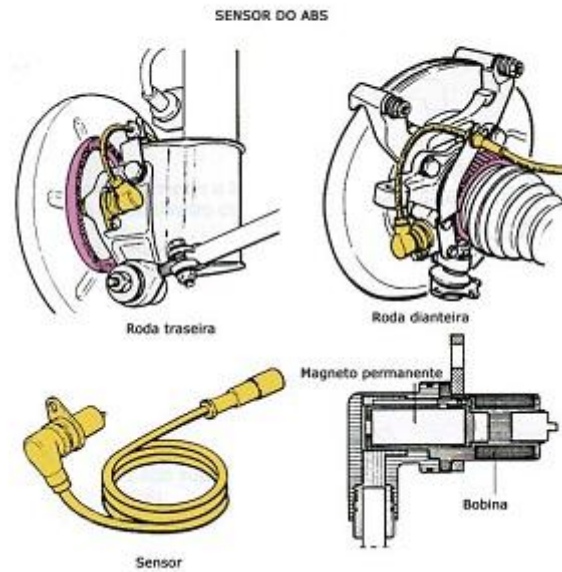


Figura 3.4 - *Encoder* de Movimento Linear

Encoder de movimento angular:

Localizado no volante do veículo, como mostra a Figura 3.5, este *encoder* é utilizado para determinar sua rotação que pode variar de -35° a 35° e apresenta uma resolução de 7142 pulsos por grau rotacionado.



Figura 3.5 - *Encoder* de Movimento Angular

A localização de robôs móveis utilizando *encoders* é conhecida como odometria.

3.3.1 Odometria

Odometria é o método de navegação mais utilizado para o posicionamento de um robô móvel. Sabe-se que a odometria fornece uma boa precisão de curto prazo, é barato e permite taxas de amostragem muito altas. No entanto, a ideia fundamental de odometria é a integração das informações do movimento incremental ao longo do tempo, o que leva inevitavelmente à acumulação de erros. Particularmente, o acúmulo de erros de orientação irá causar erros de posição de grande porte que aumentam proporcionalmente com a distância percorrida pelo robô. Apesar dessas limitações, a maioria dos pesquisadores concorda que a odometria é uma parte importante de um sistema de navegação de robôs e que as tarefas de navegação serão simplificadas se a precisão odométrica puder ser melhorada (17).

Odometria é baseada em equações simples, que são facilmente implementadas e que utilizam dados de *encoders* incrementais de baixo custo. No entanto, odometria é também baseado no pressuposto de que as revoluções da roda podem ser traduzidas em deslocamento linear em relação ao piso. Esta hipótese é apenas de validade limitada, uma vez que está sujeita a erros.

3.3.2 Precisão

Junto com o caso extremo de derrapagem total, existem várias outras razões mais sutis para imprecisões na tradução de leituras do *encoder* em um movimento linear. Todas estas fontes de erro se encaixam em uma das duas categorias: erros sistemáticos e erros não-sistemáticos (17).

As principais causas de erros sistemáticos são:

- Diferentes diâmetros das rodas;
- Diâmetro da roda difere da especificação;
- Desalinhamento das rodas;
- Resolução do *encoder* limitada.

As principais causas de erros não-sistemáticos são:

- Movimento sobre solos não uniformes (rugosos);
- Movimento sobre obstáculos inesperados no solo;

- Escorregamento das rodas:
 - Solo escorregadio;
 - Grandes acelerações do veículo;
 - Rotações rápidas.

A clara distinção entre erros sistemáticos e não-sistemáticos é de grande importância para a efetiva redução dos erros de odometria. Por exemplo, os erros sistemáticos são particularmente graves porque se acumulam constantemente. Na maioria das superfícies lisas os erros sistemáticos contribuem muito mais para os erros de odometria do que os erros não-sistemáticos, no entanto, em superfícies ásperas com irregularidades significativas, os erros não-sistemáticos são dominantes.

4 PROCESSAMENTO DE DADOS

Os sensores descritos no capítulo anterior são responsáveis por fornecer os dados necessários para traçar a trajetória do veículo autônomo. O traçado do veículo é obtido de duas maneiras diferentes: a primeira usando as informações do DGPS e a segunda usando as informações da odometria juntamente com o modelo cinemático. Os dados são armazenados numa frequência de 10 Hz para posterior processamento.

Na Seção 4.1 será discutido como a trajetória do veículo é traçada utilizando os dados dos sensores. Na Seção 4.2 será descrito um Filtro de Kalman para conseguir uma maior precisão para a trajetória do veículo.

4.1 Trajetória

Traçar a trajetória do veículo utilizando os dados dos sensores é uma maneira de analisar a precisão dos sensores e conseqüentemente a precisão do posicionamento do veículo. Os dois métodos utilizados neste trabalho para obter o traçado do veículo são descritos a seguir.

Os dados fornecidos pelo DGPS, latitude e longitude, são transformados em coordenadas globais e representam o posicionamento do veículo no ambiente. A antena do DGPS se encontra localizada no centro de massa do veículo e com a posição X e Y obtida a cada instante de tempo, é possível traçar a trajetória percorrida pelo veículo.

A odometria é responsável por fornecer as variáveis de entrada das equações cinemáticas. Com o valor do ângulo da roda obtida pelo *encoder* do volante e o espaço percorrido pelo veículo obtido pelo *encoder* das rodas traseiras é possível calcular a posição X e Y em coordenadas globais utilizando o modelo cinemático. As posições calculadas pela cinemática também são utilizadas para traçar a trajetória do veículo.

A Figura 4.1 e a Figura 4.2 mostram a trajetória percorrida pelo veículo obtida com informações do DGPS e utilizando a odometria e a cinemática.

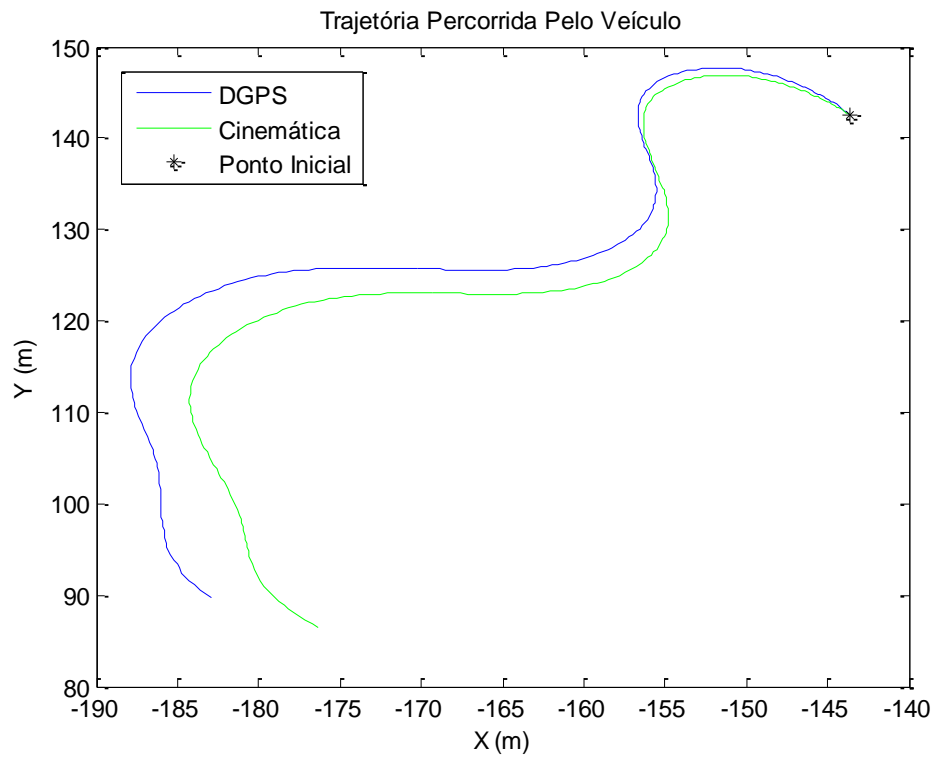


Figura 4.1 - Trajetória Percorrida pelo Veículo (Posição)

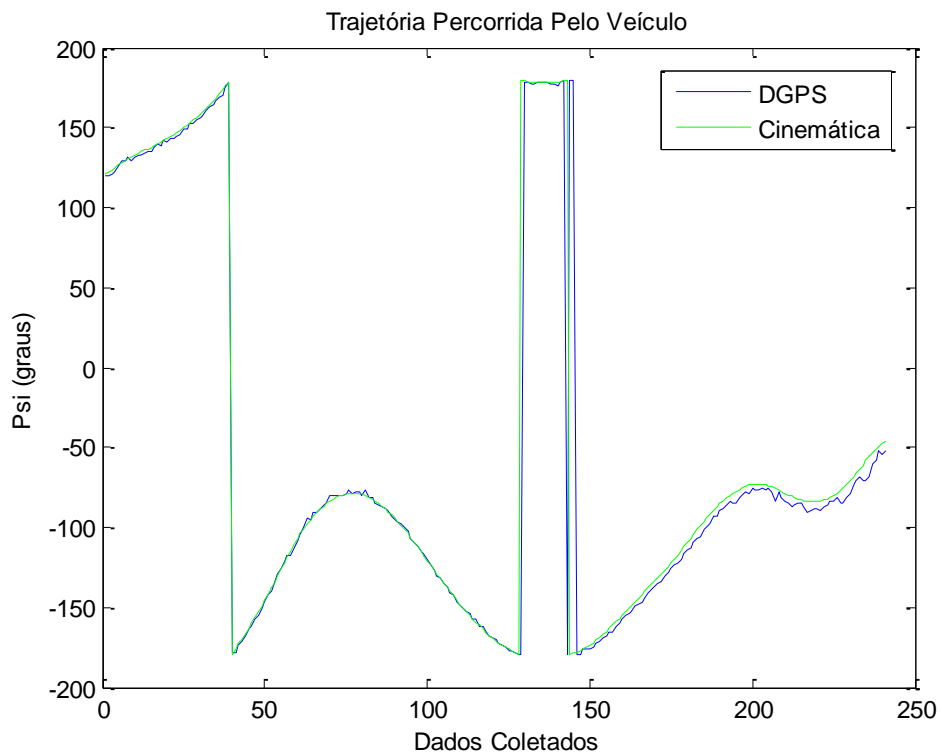


Figura 4.2 - Trajetória Percorrida pelo Veículo (Ângulo)

A Figura 4.1 representa o posicionamento em X e Y do veículo ao longo da trajetória e a Figura 4.2 representa a variação do ângulo ψ do veículo. A posição X e Y é dada em metros e o ângulo ψ é dado em graus e pode variar de -180° a 180° . A trajetória percorrida pelo veículo representada pelos dois gráficos descritos será apresentada dessa forma, ao longo deste trabalho, para outros percursos.

Como pode-se notar, as duas trajetórias apresentam divergências que se devem às imprecisões dos sensores e a simplificação do modelo cinemático. No caso da cinemática, o erro do ângulo do veículo é pequeno, mas causa um erro significativo no posicionamento do veículo já que uma informação depende da outra.

4.2 Filtro de Kalman

O filtro de Kalman é frequentemente usado para combinar dados obtidos de diferentes sensores em uma estimativa estatisticamente ótima. Se um sistema pode ser descrito através de uma modelo linear e as incertezas dos sensores e do sistema podem ser modelados como ruídos Gaussianos brancos, então o filtro de Kalman proverá uma estimativa estatisticamente ótima para os dados fundidos. Isto significa que, sob certas condições, o filtro de Kalman é capaz de encontrar a melhor estimativa baseada na correção de cada medida individual.

O filtro de Kalman é um estimador linear que dá uma estimativa ótima imparcial do atual estado do sistema com base em todas as medições anteriores (18).

O modelo de estado é dado por

$$X_k = AX_{k-1} + w_{k-1} \quad (4.1)$$

onde X_k é o vetor de estado do sistema no instante k , A é a matriz que define a dinâmica do sistema, e w_k é um ruído Gaussiano branco com covariância Q . A equação de medição é dada por

$$Z_k = HX_k + v_k \quad (4.2)$$

onde Z_k é a medida no instante k , H é uma matriz que define a saída de medição, e v_k é um ruído Gaussiano branco com covariância R . Os ruídos de processos v_k e w_k são considerados independentes.

Considerando que $X_{k|j}$ e $P_{k|j}$ denotam as estimativas do estado e da covariância em k dadas as medições no instante j , o filtro de Kalman é descrito com as seguintes equações

Atualização do Tempo (Predição)

$$X_{k|k-1} = AX_{k-1|k-1} \quad (4.3)$$

$$P_{k|k-1} = AP_{k-1|k-1}A^T + Q \quad (4.4)$$

Atualização da Medição (Correção)

$$X_{k|k} = X_{k|k-1} + K_k(Z_k - HX_{k|k-1}) \quad (4.5)$$

$$P_{k|k} = (I - K_kH)P_{k|k-1} \quad (4.6)$$

onde

$$K_k = P_{k|k-1}H^T [HP_{k|k-1}H^T + R]^{-1} \quad (4.7)$$

4.2.1 Filtro α - β

Um filtro α - β é uma versão de estado estacionário do filtro de Kalman. Sob condições de processos de ruído estacionário e uma taxa fixa de medição, o filtro α - β é equivalente ao filtro de Kalman correspondente ao estado estacionário. Devido às condições restritivas, o filtro α - β apresenta estimativas de qualidade inferior, no entanto, a simplicidade e eficiência computacional do filtro α - β justifica a sua utilização em diversas aplicações (19).

Um filtro α - β é um filtro de Kalman para um modelo de movimento com velocidade quase constante sob condições de estado estacionário, que são resultado de uma taxa de dados constante e processos de ruído estacionário. O modelo de estado para o filtro α - β é um vetor de estado simples de dois elementos dado por

$$X_k = [x_k \quad \dot{x}_k]^T \quad (4.8)$$

onde x_k é posição e \dot{x}_k é velocidade. O sistema dinâmico para o filtro α - β é modelado com

$$A = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \quad (4.9)$$

onde T é o tempo entre as medições. Substituindo 4.3 e 4.5 com

$$H = [1 \quad 0] \quad \text{e} \quad K_k = \left[\alpha \quad \frac{\beta}{T} \right]^T \quad (4.10)$$

o filtro α - β é descrito por:

Atualização do Tempo

$$x_{k|k-1} = x_{k-1|k-1} + T\dot{x}_{k-1|k-1} \quad (4.11)$$

ou

$$x_{k|k-1} = x_{k-1|k-1} + dx$$

$$\dot{x}_{k|k-1} = \dot{x}_{k-1|k-1} \quad (4.12)$$

Atualização da Medição

$$x_{k|k} = x_{k|k-1} + \alpha(x_k - x_{k|k-1}) \quad (4.13)$$

$$\dot{x}_{k|k} = \dot{x}_{k|k-1} + \frac{\beta}{T}(x_k - x_{k|k-1}) \quad (4.14)$$

onde

$$0 < \alpha < 1 \quad (4.15)$$

$$0 < \beta < 1 \quad (4.16)$$

4.2.2 Filtro α - β Aplicado na Estimativa da Posição

As Equações 4.11 e 4.13 são utilizadas para estimar e corrigir a posição do veículo, respectivamente. A estimativa do valor de x ($x_{k|k-1}$) é dado pela cinemática com base nos dados anteriores ($x_{k-1|k-1}$) e a correção do valor de x estimado ($x_{k|k}$) é calculada considerando o valor de x_k como o valor obtido pelo DGPS. O valor de α escolhido foi de 0.15. Esse procedimento é realizado também para y .

A Figura 4.2 mostra o filtro α - β aplicado na estimativa da posição do veículo.

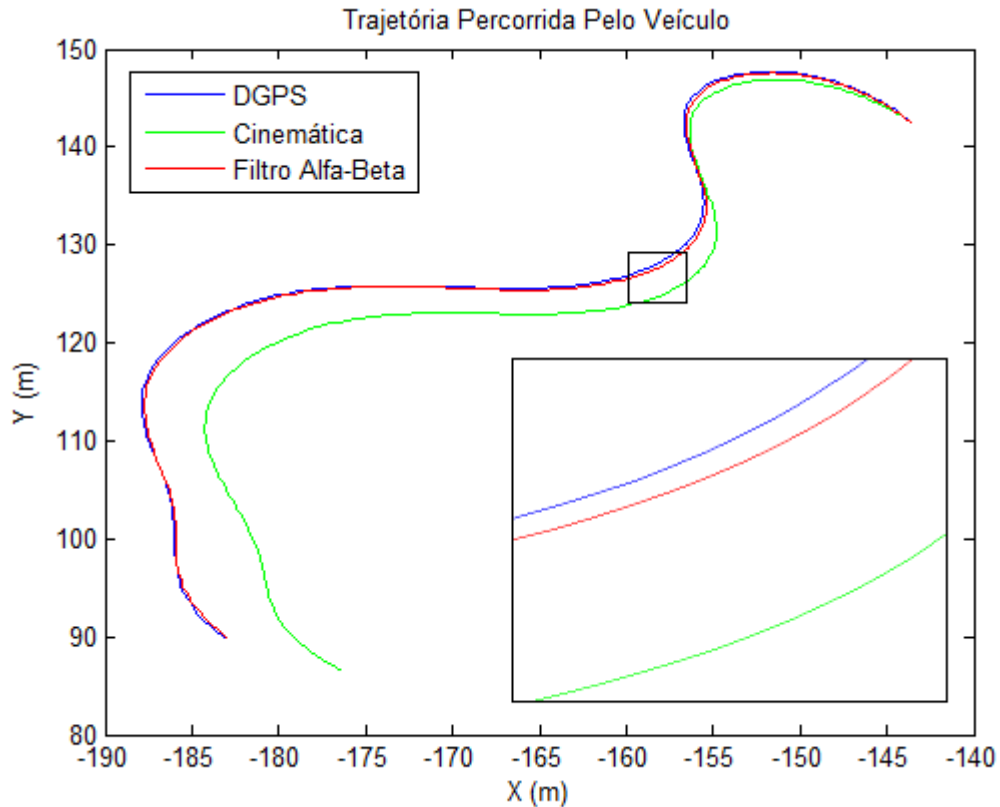


Figura 4.3 - Filtro α - β Aplicado na Estimativa de Posição

4.2.3 Filtro α - β Aplicado na Estimativa do Ângulo

Assim como para a posição, as Equações 4.11 e 4.13 são utilizadas para estimar e corrigir o ângulo do veículo, respectivamente. A estimativa do valor de ψ ($\psi_{k|k-1}$) é dado pela cinemática com base nos dados anteriores ($\psi_{k-1|k-1}$) e a correção do valor de ψ estimado ($\psi_{k|k}$) é calculada considerando o valor de ψ_k como o valor calculado utilizando os dados do DGPS, ou seja

$$\psi = \tan^{-1} \left(\frac{dy}{dx} \right)$$

O valor de α escolhido para este caso foi de 0.10.

A Figura 4.3 mostra o filtro α - β aplicado na estimativa do ângulo do veículo.

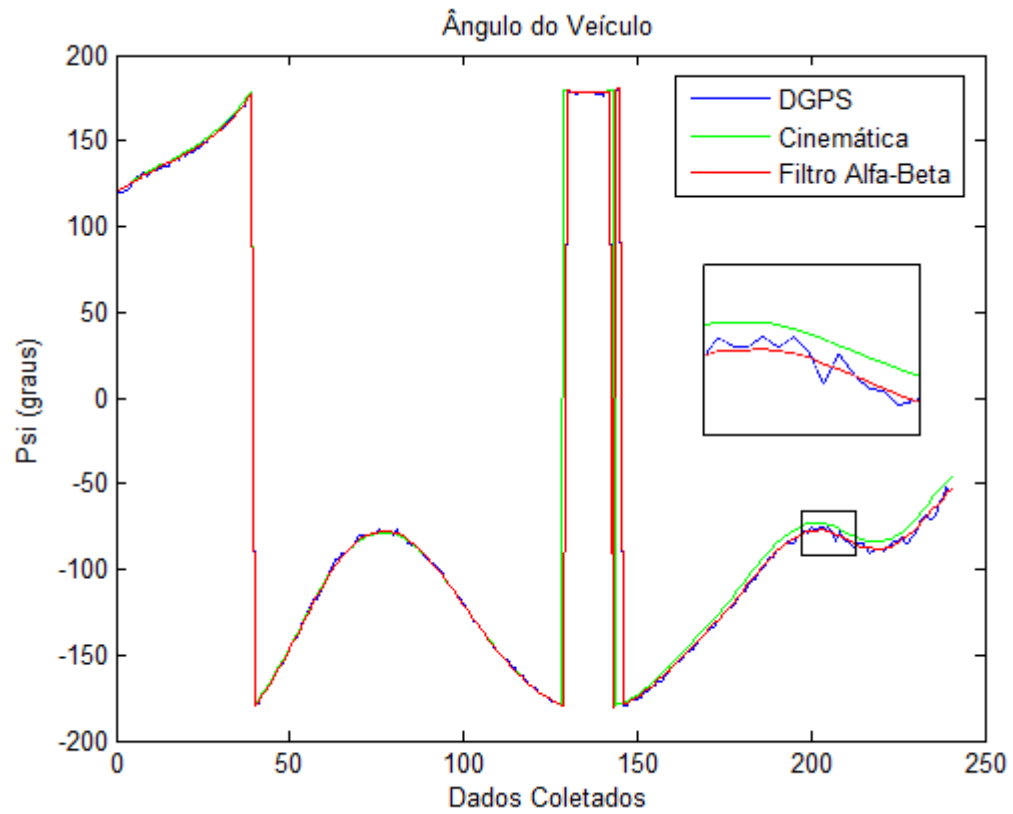


Figura 4.4 - Filtro α - β Aplicado na Estimativa de Ângulo

5 REDES NEURAIIS ARTIFICIAIS

As redes neurais artificiais são um método para solucionar problemas através da simulação do cérebro humano, inclusive em seu comportamento, ou seja, aprendendo. São técnicas computacionais que apresentam um modelo inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência.

As redes neurais são utilizadas nas seguintes aplicações reais:

- aproximação de funções;
- previsão de séries temporais;
- classificações;
- reconhecimento de padrões.

Neste capítulo serão abordadas as características adotadas na construção de uma rede neural aplicada como aproximador universal de funções. O teorema de aproximação universal afirma que uma rede neural multicamadas é capaz de realizar uma aproximação uniforme, dado um conjunto de entradas e saídas suficientemente significativas para representar uma função.

A Seção 5.1 apresenta uma breve introdução sobre redes neurais artificiais. Uma rede neural pode ser caracterizada por quatro aspectos principais: 1- a representação do neurônio artificial e sua função de ativação, 2- o padrão de conexões entre os neurônios (arquitetura), 3- o método de aprendizado e 4- o método de determinação dos pesos das conexões (algoritmo de treinamento). Esses quatro aspectos serão abordados nas Seções 5.1, 5.2, 5.3 e 5.4 respectivamente.

5.1 Conceitos Preliminares

O cérebro é um computador altamente complexo, não linear e paralelo. Ele tem a capacidade de organizar seus constituintes estruturais, conhecidos como neurônios, assim como executar cálculos determinados (por exemplo, reconhecimento de padrões, percepção e controle motor), muitas vezes mais rápido que o mais rápido computador existente hoje (20).

Uma rede neural artificial é um modelo que emula uma rede neural biológica. Os nós de uma rede neural artificial são baseados numa representação matemática simples do que pensamos ser a aparência real dos neurônios (21).

Redes neurais artificiais podem ser mais adequadamente caracterizadas como "modelos computacionais" com propriedades específicas, tais como a capacidade de se adaptar e aprender, para generalizar, agrupar ou organizar os dados, e que a operação é baseada em processamento paralelo (22).

É evidente que o poder de computação de uma rede neural deriva, primeiramente, de sua estrutura maciçamente paralela e distribuída, e também de sua capacidade de aprender e, portanto, generalizar. Generalização refere-se à rede neural produzir saídas razoáveis para as entradas que não encontrou durante o treinamento (aprendizagem). Estas duas capacidades de processamento de informação possibilitam às redes neurais a resolução de problemas complexos que outrora eram considerados intratáveis (20).

5.1.1 Neurônio Biológico

Os neurônios são divididos em três seções: o corpo da célula, os dendritos e o axônio, cada um com funções específicas, porém complementares. Os dendritos têm por função receber as informações, ou impulsos nervosos, oriundas de outros neurônios e conduzi-las até o corpo celular. No corpo celular a informação é processada e novos impulsos são gerados. Estes impulsos são transmitidos a outros neurônios, passando através do axônio até os dendritos dos neurônios seguintes. O ponto de contato entre a terminação axônica de um neurônio e o dendrito de outro é chamado de sinapse. São pelas sinapses que os neurônios se unem funcionalmente, formando redes neurais. As sinapses funcionam como válvulas, e são capazes de controlar a transmissão de impulsos, isto é, o fluxo da informação, entre os neurônios na rede neural. O efeito das sinapses é variável, e é esta variação que dá ao neurônio capacidade de adaptação. A Figura 5.1 ilustra, de forma simplificada, os componentes do neurônio (23).

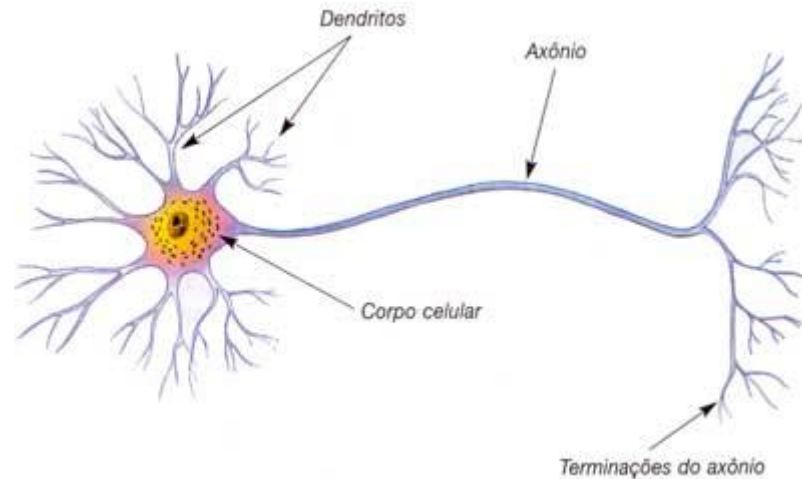


Figura 5.1 - Neurônio Biológico

Os sinais oriundos dos neurônios pré-sinápticos são passados para o corpo do neurônio pós-sináptico, onde são comparados com os outros sinais recebidos pelo mesmo. Se o percentual em um intervalo curto de tempo é suficientemente alto, a célula "dispara", produzindo um impulso que é transmitido para as células seguintes (neurônios pós-sinápticos). Este sistema simples é responsável pela maioria das funções realizadas pelo nosso cérebro. A capacidade de realizar funções complexas surge com a operação em paralelo de todos os 10^{11} nodos do nosso cérebro.

5.1.2 Modelo de um Neurônio Artificial

O modelo de neurônio proposto é uma simplificação do neurônio biológico. Sua descrição matemática resultou em um modelo com n terminais de entrada x_1, x_2, \dots, x_n , (que representam os dendritos) e apenas um terminal de saída y (representando o axônio). Para emular o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados w_1, w_2, \dots, w_n , cujos valores podem ser positivos ou negativos, dependendo de as sinapses correspondentes serem inibitórias ou excitatórias. O efeito de uma sinapse particular i no neurônio pós-sináptico é dado por $x_i w_i$. Os pesos determinam "em que grau" o neurônio deve considerar sinais de disparo que ocorrem naquela conexão. Uma descrição do modelo está ilustrada na Figura 5.2 (23).

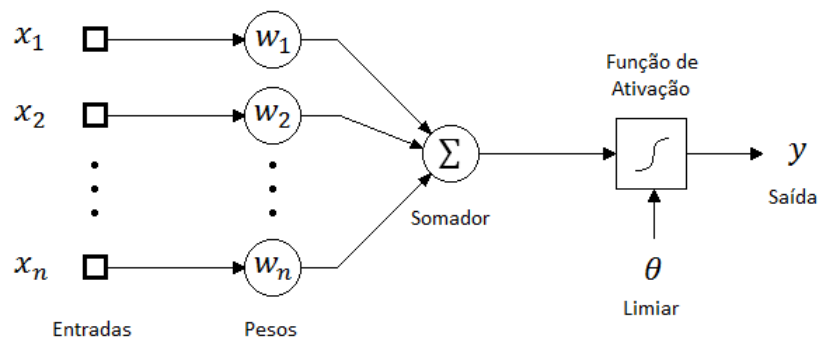


Figura 5.2 - Neurônio Artificial

Um neurônio biológico dispara quando a soma dos impulsos que ele recebe ultrapassa o seu limiar de excitação (*threshold*). O corpo do neurônio, por sua vez, é emulado por um mecanismo simples que faz a soma dos valores $x_i w_i$ recebidos pelo neurônio (soma ponderada) e decide se o neurônio deve ou não disparar comparando a soma obtida ao limiar do neurônio. A ativação do neurônio é obtida através da aplicação de uma "função de ativação", que ativa ou não a saída, dependendo do valor da soma ponderada das suas entradas.

5.1.3 Função de Ativação

O neurônio terá então sua saída ativa quando:

$$\sum_{i=1}^n x_i w_i \geq \theta$$

onde n é o número de entradas do neurônio, w_i é o peso associado à entrada x_i e θ é o limiar (*threshold*) do neurônio.

A função sigmóide, cujo gráfico é em forma de S, é a forma mais comum de função de ativação utilizada na construção de redes neurais artificiais. É definida como uma função estritamente crescente que exhibe um equilíbrio entre o comportamento linear e não linear (20). A função de ativação utilizada neste trabalho é definida por:

$$f(x) = \frac{2}{1 + e^{(-ax)}} - 1$$

onde a é o parâmetro de inclinação da função sigmóide. Ao variar o parâmetro a da função, obtemos sigmóides de diferentes inclinações, como ilustrado na Figura 5.3. A função sigmóide assume um intervalo contínuo de valores de -1 a 1.

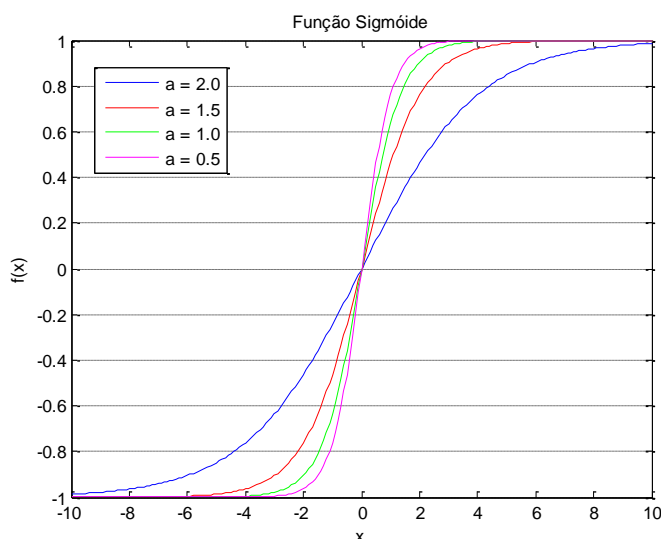


Figura 5.3 - Função Sigmóide

5.2 Arquitetura da Rede

A rede escolhida para este trabalho é classificada como uma rede neural multicamada, acíclica, completamente conectada.

Uma rede neural acíclica (*feedforward*) se distingue pela presença de uma ou mais camadas ocultas, cujos nós são correspondentemente chamados neurônios escondidos ou unidades escondidas. A função dos neurônios ocultos é intervir entre a entrada externa e a saída da rede, de alguma maneira útil. Ao adicionar uma ou mais camadas ocultas, a rede está habilitada para extrair estatísticas de ordem superior. Em outras palavras, a rede adquire uma perspectiva global, apesar de sua conectividade local, devido ao conjunto extra de conexões sinápticas e da dimensão extra de interações neurais. A capacidade dos neurônios escondidos para extrair estatísticas de ordem elevada é particularmente valiosa quando o tamanho da camada de entrada é grande (20).

A fonte de nós na camada de entrada da rede fornece os respectivos elementos do padrão de ativação (vetor de entrada), que constituem os sinais de entrada aplicados aos neurônios na segunda camada (ou seja, a primeira camada escondida). Os sinais de saída da

segunda camada são utilizados como entradas para a terceira camada, e assim por diante para o resto da rede. Normalmente os neurônios em cada camada da rede têm como entradas, apenas os sinais de saída da camada anterior. O conjunto de sinais de saída dos neurônios na camada de saída da rede constitui a resposta global da rede para o padrão de ativação fornecido pela fonte de nós na camada de entrada. O gráfico da Figura 5.4 ilustra o *layout* de uma rede neural acíclica (*feedforward*) multicamadas para o caso de uma única camada escondida. A rede tem dez nós de entrada, quatro neurônios na camada oculta, e dois neurônios de saída.

A rede neural da Figura 5.4 é dito ser totalmente conectada no sentido de que cada nó em cada camada da rede está ligado a cada outro nó na camada adjacente.

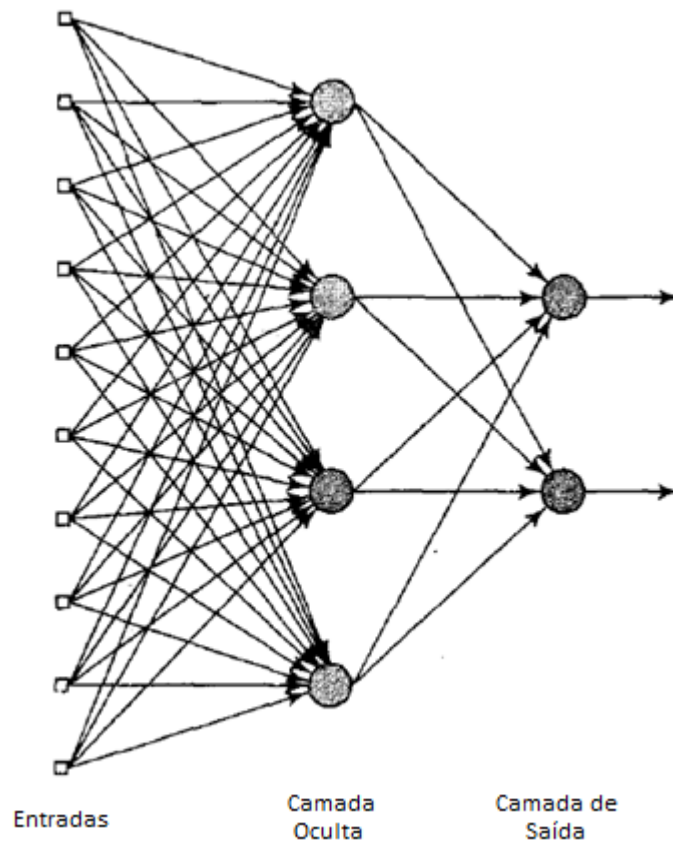


Figura 5.4 - Arquitetura da Rede Neural

5.3 Aprendizagem

A utilização de uma RNA (Rede Neural Artificial) na solução de uma tarefa passa inicialmente por uma fase de aprendizagem, quando a rede extrai informações relevantes de

padrões de informação apresentados para ela, criando assim uma representação própria para o problema. A etapa de aprendizagem consiste em um processo iterativo de ajuste de parâmetros da rede, os pesos das conexões entre as unidades de processamento, que guardam, ao final do processo, o conhecimento que a rede adquiriu do ambiente em que está operando (23).

A capacidade de aprender através de exemplos e de generalizar a informação aprendida é, sem dúvida, o atrativo principal da solução de problemas através de RNAs. A generalização, que está associada à capacidade de a rede aprender através de um conjunto reduzido de exemplos e posteriormente dar respostas coerentes para dados não-conhecidos, é uma demonstração de que a capacidade das RNAs vai muito além do que simplesmente mapear relações de entrada e saída. As RNAs são capazes de extrair informações não-apresentadas de forma explícita através dos exemplos.

5.3.1 Aprendizado Supervisionado

Aprendizado Supervisionado foi o método de aprendizado escolhido para este trabalho. É chamado de aprendizado supervisionado porque a entrada e saída desejadas para a rede são fornecidas por um supervisor (professor) externo. O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos. A Figura 5.5 ilustra o mecanismo de aprendizado supervisionado. O professor indica explicitamente um comportamento bom ou ruim para a rede, visando direcionar o processo de treinamento. A rede tem sua saída corrente (calculada) comparada com a saída desejada, recebendo informações do supervisor sobre o erro da resposta atual. A cada padrão de entrada submetido à rede compara-se a resposta desejada (que representa uma ação ótima para ser realizada pela rede) com a resposta calculada, ajustando-se os pesos das conexões para minimizar o erro. A minimização da diferença é incremental, já que pequenos ajustes são feitos nos pesos a cada etapa de treinamento, de tal forma que estes caminhem, se houver solução possível, para uma solução. A soma dos erros quadráticos de todas as saídas é normalmente utilizada como medida de desempenho da rede e também como função de custo a ser minimizada pelo algoritmo de treinamento (23).

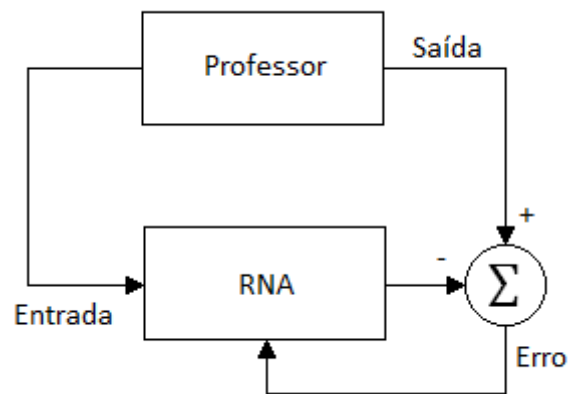


Figura 5.5 - Aprendizado Supervisionado

O problema passa a ser então como treinar estas redes.

5.4 Treinamento

Existem atualmente vários algoritmos para treinar redes neurais. A rede neural desenvolvida neste trabalho foi treinada usando algoritmo genético.

A aplicação de redes neurais acíclicas (*feedforward*) multicamadas em alguns problemas reais tem sido dificultada pela falta de um algoritmo de treinamento confiável que encontre um conjunto de pesos globais quase ideais em um tempo relativamente curto. Algoritmos genéticos são uma classe de procedimentos de otimização que são bons em exploração de um espaço amplo e complexo de uma forma inteligente para encontrar valores próximos ao ótimo global. Assim, eles são bem adaptados para o problema de treinamento de redes neurais acíclicas (*feedforward*) multicamadas (24). Uma breve explanação sobre Algoritmo Genético e a definição de seus parâmetros adotados para treinamento das RNAs, podem ser vistos no Apêndice A.

5.4.1 Dinâmica de Treinamento

A dinâmica de treinamento adotada neste trabalho é por ciclo (*epoch*). No treinamento por ciclo os pesos são atualizados após todos os exemplos de treinamento terem sido apresentados. Esta técnica é geralmente estável, mas pode ser lenta se o conjunto de treinamento for grande.

5.4.2 Critérios de Parada do Treinamento

Critério de parada é a determinação do momento em que o treinamento de uma rede neural deve ser encerrado. Uma boa determinação destes critérios é fundamental para um bom treinamento e conseqüentemente uma boa generalização.

O critério de parada adotado neste trabalho é por número de iterações ou número de ciclos, ou seja, o número de vezes em que o conjunto de treinamento é apresentado à rede. Um número excessivo de ciclos pode levar a rede à perda do poder de generalização (*overfitting*). Por outro lado, com um pequeno número de ciclos a rede pode não chegar ao seu melhor desempenho (*underfitting*).

6 REDES NEURO-ESPECIALISTAS

Basicamente, uma rede neural é um mapeador de entradas e saídas. Ou seja, em uma situação de treinamento supervisionado, um conjunto de dados representando diversas situações de entradas e saídas de um sistema de um mundo real são passadas para um algoritmo de treinamento que tenta capturar este conhecimento em forma dos pesos internos na rede neural.

Em uma situação ideal, se todo o conhecimento for capturado, quando novas entradas forem apresentadas para a rede, uma saída idêntica ao sistema modelado será apresentada. Entretanto esta situação ideal quase nunca se apresenta em problemas de grande complexidade, logo o conhecimento não preenchido ou mapeado pela rede neural retorna através dos erros fornecidos.

Uma forma de melhorar este mapeamento é apresentar, durante o treinamento da rede, parte do conhecimento que esta deve adquirir. Desta forma, com parte do conhecimento já mapeado, a rede deverá apenas aprender o restante. O modelo do sistema preenche quase todo o conhecimento necessário. Desta forma, incluir isso em uma rede deixa o treinamento mais eficiente, uma vez que nesta situação apenas o restante do conhecimento do processo é que deve ser capturado.

Para desenvolver esta idéia, foi projetada uma Rede Neuro-Especialista, onde a função de ativação de parte dos neurônios são as equações do modelo do sistema. Com esta informação presente dentro da rede, os pesos restantes conseguem reduzir a perda de informação e, desta forma, mapear melhor o conhecimento real.

A Figura 6.1 ilustra a diferença entre o mapeamento do sistema por uma Rede Neural pura e por uma Rede Neuro-Especialista.

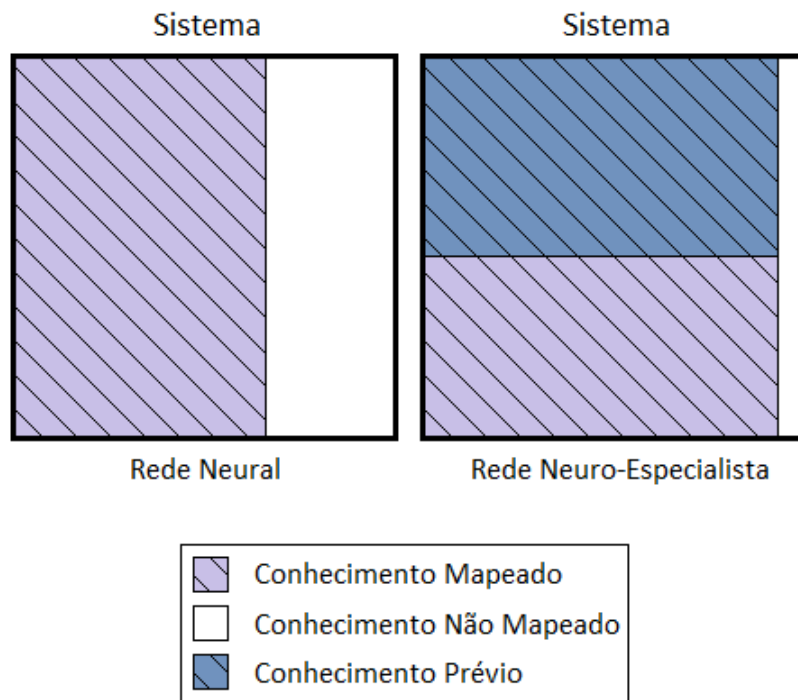


Figura 6.1 - Mapeamento do Sistema

6.1 Modelo de um Neurônio Especialista

O modelo de um Neurônio Especialista apresenta n terminais de entrada x_1, x_2, \dots, x_n e um terminal de saída y , onde n corresponde ao número de variáveis da função de ativação. Os terminais de entrada do neurônio têm pesos acoplados w_1, w_2, \dots, w_n , cujos valores podem ser positivos ou negativos. Uma descrição do modelo está ilustrada na Figura 6.2.

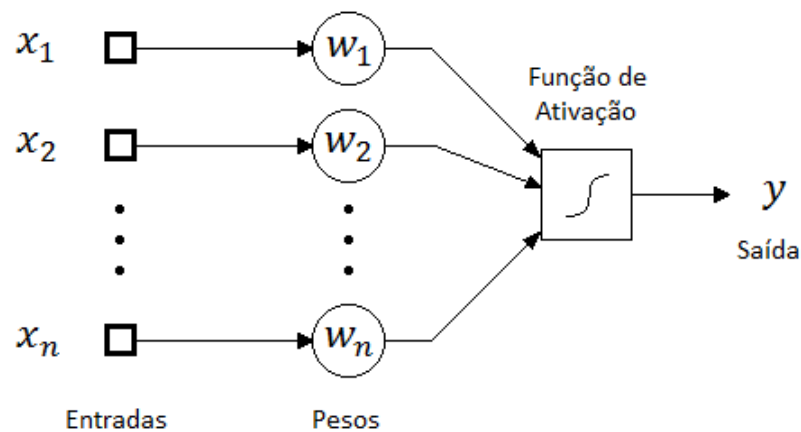


Figura 6.2 - Neurônio Especialista

6.2 Função de Ativação

A função de ativação de um Neurônio Especialista corresponde à equação do modelo do sistema e apresenta como variáveis, as entradas do neurônio multiplicadas por seus respectivos pesos, ou seja, cada entrada do neurônio corresponde a uma variável da função e não o somatório das entradas como em um neurônio artificial comum.

A função de ativação é representada por

$$f(x_1w_1, x_2w_2, \dots, x_nw_n)$$

6.3 Arquitetura da Rede

A Rede Neuro-Especialista, assim como uma Rede Neural pura, também é classificada como uma rede multicamada, acíclica, completamente conectada. A diferença para um Rede Neural pura é que a Rede Neuro-Especialista apresenta uma camada a mais denominada camada especialista que contém um ou mais neurônios especialistas.

A Figura 6.3 ilustra o *layout* de uma Rede Neuro-Especialista. A rede tem dois nós de entrada, um neurônio na camada especialista, cinco neurônios na camada oculta e um neurônio de saída.

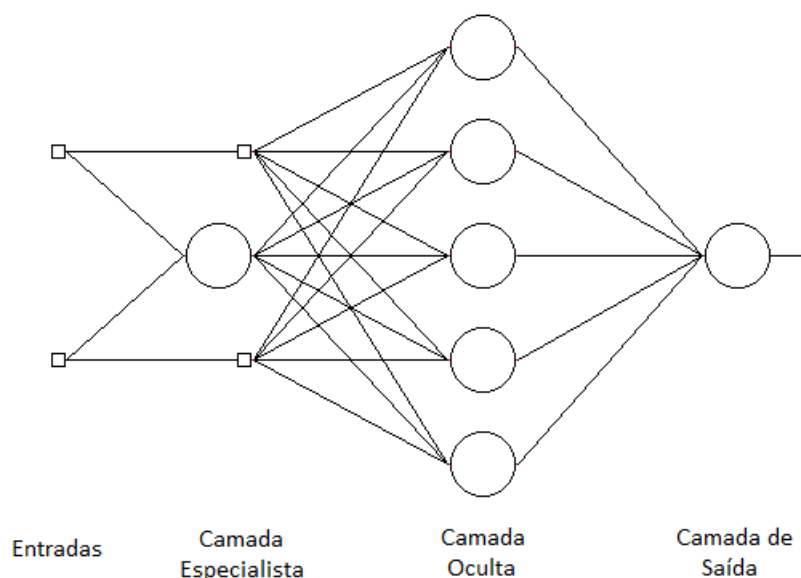


Figura 6.3 - Arquitetura Rede Neuro-Especialista

6.4 Treinamento

A princípio os pesos do neurônio especialista foram arbitrados como um, para que dessa forma a saída do neurônio possa ser fiel ao modelo do sistema. Em trabalhos futuros pretende-se estudar a influência do neurônio especialista na rede ao variar os pesos dessas ligações.

Considerando então que os pesos do neurônio especialista são fixos, o treinamento da Rede Neuro-Especialista é realizado da mesma forma que o treinamento da Rede Neural pura, ou seja, utiliza-se um supervisor com o objetivo de ajustar os pesos da rede para que se possa encontrar uma ligação entre os pares de entrada e saída fornecidos.

6.5 Exemplo

Para exemplificar a metodologia, propõe-se uma Rede Neuro-Especialista para aproximar a função $\sin(a + b)$, onde a função de ativação do neurônio especialista é a Função Seno definida pela Série de Taylor.

A Série de Taylor que define a Função Seno é dada por

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \quad (6.1)$$

Considerando apenas os sete primeiros termos da Série de Taylor, tem-se

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \frac{x^{13}}{13!} \quad (6.2)$$

Para a função $\sin(a + b)$ a eq. 6.2 fica

$$\sin(a + b) = (a + b) - \frac{(a + b)^3}{3!} + \frac{(a + b)^5}{5!} - \frac{(a + b)^7}{7!} + \frac{(a + b)^9}{9!} - \frac{(a + b)^{11}}{11!} + \frac{(a + b)^{13}}{13!} \quad (6.3)$$

Logo, a função de ativação do neurônio especialista é dada por

$$f(a, b) = (a + b) - \frac{(a + b)^3}{3!} + \frac{(a + b)^5}{5!} - \frac{(a + b)^7}{7!} + \frac{(a + b)^9}{9!} - \frac{(a + b)^{11}}{11!} + \frac{(a + b)^{13}}{13!} \quad (6.4)$$

A Figura 6.4 mostra o neurônio especialista para este exemplo.

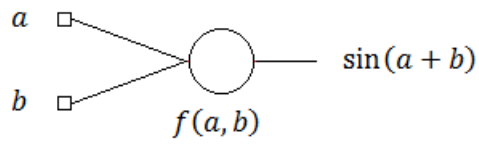


Figura 6.4 - Neurônio Especialista (Exemplo)

A Figura 6.5 mostra a Rede Neuro-Especialista para este exemplo.

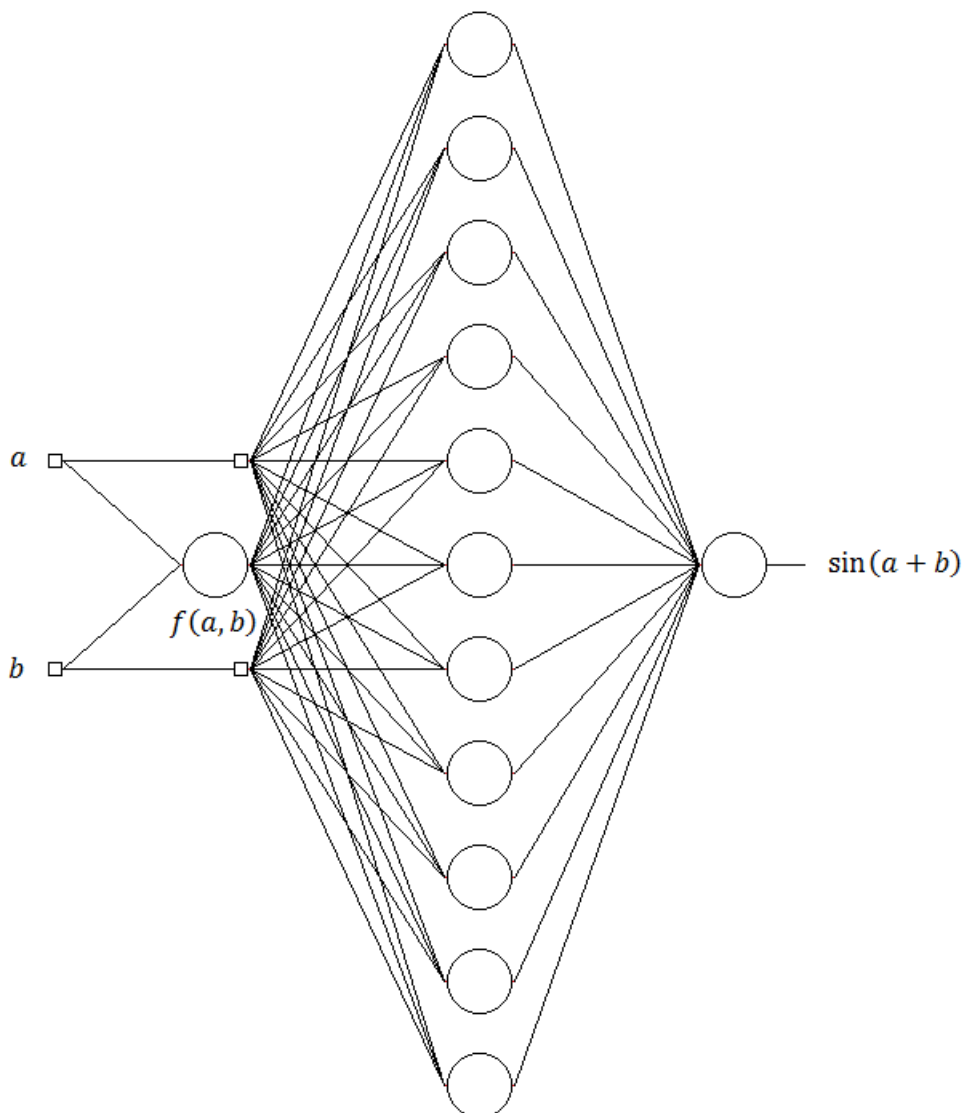


Figura 6.5 - Arquitetura Rede Neuro-Especialista (Exemplo)

A Figura 6.6 compara a saída desejada da rede neural (Função Seno) e a saída do neurônio especialista (Função Seno definida pela Série de Taylor) e a Tabela 6.1 apresenta alguns resultados numéricos para essa comparação. Pode-se notar que os valores do modelo do sistema e os valores desejados são muito próximos, o que contribui para o conhecimento prévio da rede.

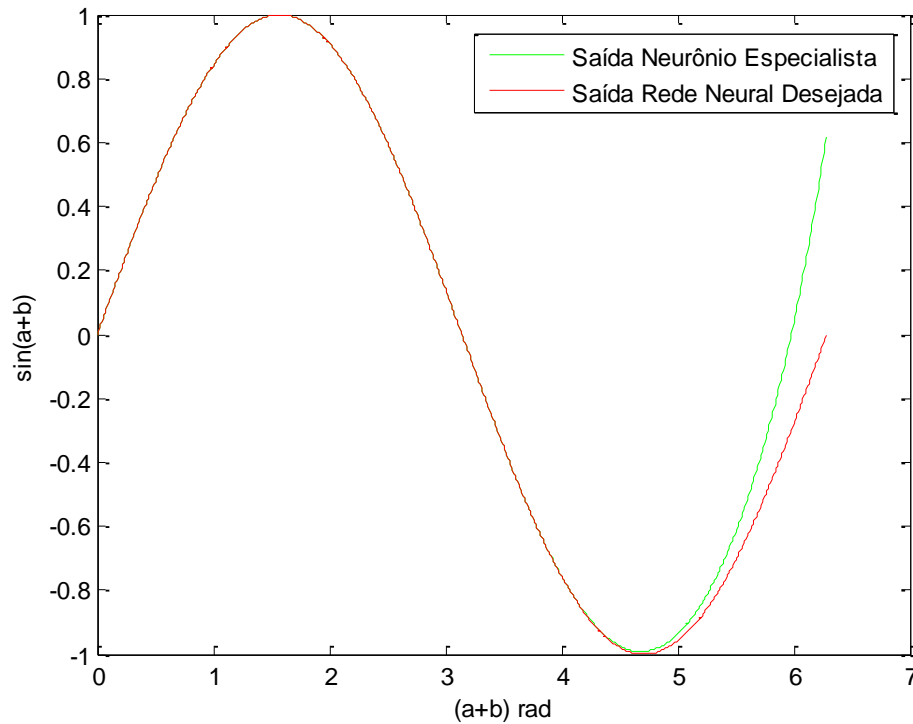


Figura 6.6 - Função Seno

Entradas (rad)		Saída Neurônio Especialista	Saída Desejada (Professor)
a	b	$\sin(a + b)$	$\sin(a + b)$
0.524	1.047	1.000	1.000
1.047	2.094	0.000	0.000
1.571	3.141	-0.991	-1.000
2.094	4.189	0.6244	0.000

Tabela 6.1 - Comparação Numérica

Para salientar a eficiência da Rede Neuro-Especialista as Figuras 6.7 e 6.8 mostram os resultados de uma Rede Neural pura e uma Rede Neuro-Especialista para esse exemplo,

definidas com os mesmos parâmetros, onde o gráfico em vermelho representa a saída desejada (Função Seno) e o gráfico em azul o resultado da rede.

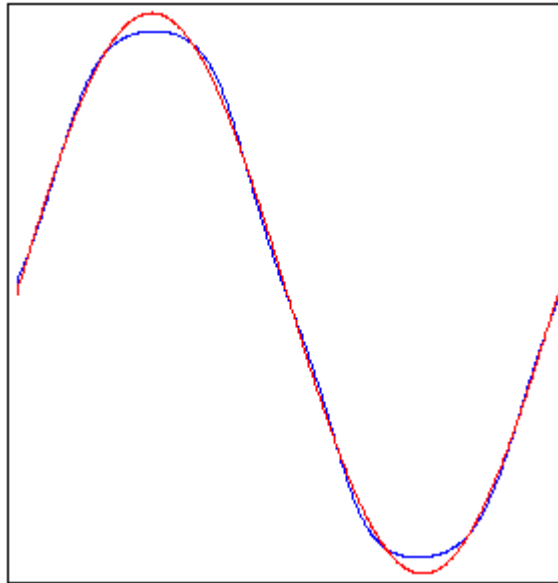


Figura 6.7 - Resultado Rede Neural (Exemplo)

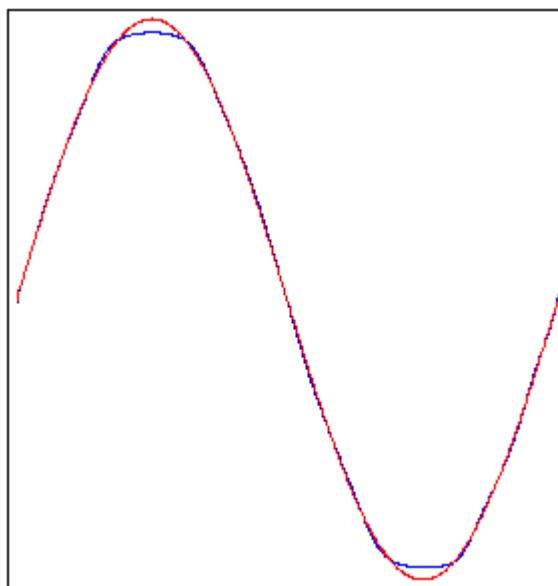


Figura 6.8 - Resultado Rede Neuro-Especialista (Exemplo)

Os resultados comprovam a transferência do conhecimento prévio do sistema à rede pelo neurônio especialista.

7 PROJETO DE UMA RNA

Para o desenvolvimento de um projeto baseado em redes neurais artificiais são necessárias diversas etapas como mostra a Figura 7.1, tais como: definição do problema, coleta dos dados de treinamento e de teste, pré-processamento dos dados, projeto da estrutura da rede, treinamento, teste e validação. Essas etapas são descritas nas Seções 7.1, 7.2, 7.3, 7.4, 7.5, 7.6 e 7.7 respectivamente.

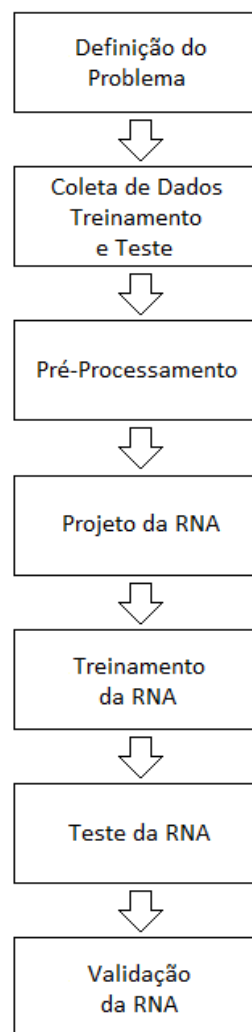


Figura 7.1 - Etapas do Projeto de uma RNA

7.1 Definição do Problema

No Capítulo 3 foi discutida a importância do posicionamento de um veículo autônomo para uma navegação eficiente, contudo, o modelo cinemático, descrito na Seção 3.1, utilizado para determinar a trajetória do veículo autônomo apresenta uma série de simplificações. Uma rede neural artificial aplicada como aproximador universal de funções será utilizada para “corrigir” as equações do modelo cinemático baseando-se em valores reais de entrada e saída.

A ideia de se “corrigir” o modelo cinemático parte do princípio de que traçando a trajetória do veículo autônomo sem precisar utilizar os dados do DGPS, contando apenas com as informações da odometria, será possível fazer navegações *indoor*, ou seja, em ambientes fechados onde não pega sinal GPS, além de baratear o custo do projeto, uma vez que um aparelho DGPS custa caro e é necessário pagar por sua informação diferencial.

7.2 Coleta de Dados para Treinamento e Testes

Como já discutido na seção anterior, o objetivo da RNA que será projetada é corrigir a trajetória do veículo autônomo calculada pelo modelo cinemático. Um conjunto de dados obtidos pelos sensores foi utilizado para traçar a trajetória percorrida pelo veículo autônomo, como já explicado na Seção 4.1. Esses dados são necessários para que a RNA possa ser treinada, testada e validada. A Tabela 7.1 mostra a relação entre os sensores e os dados obtidos.

Sensores	Dados Coletados
DGPS	Coordenadas Globais X e Y
<i>Encoder</i> Volante	Ângulo da Roda
<i>Encoder</i> Rodas Traseiras	Distância Percorrida

Tabela 7.1 - Relação Entre Sensores e Dados Coletados

A base de dados coletada refere-se aos dados de 6 trajetórias diferentes percorridas pelo veículo autônomo. Os gráficos que representam essas trajetórias, como exemplificado na Seção 4.1, podem ser vistos nas figuras a seguir. As posições X e Y do veículo durante as trajetórias são apresentadas nas Figuras 7.2, 7.4, 7.6, 7.8, 7.10 e 7.12 e a variação do ângulo do veículo nas Figuras 7.3, 7.5, 7.7, 7.9, 7.11 e 7.13.

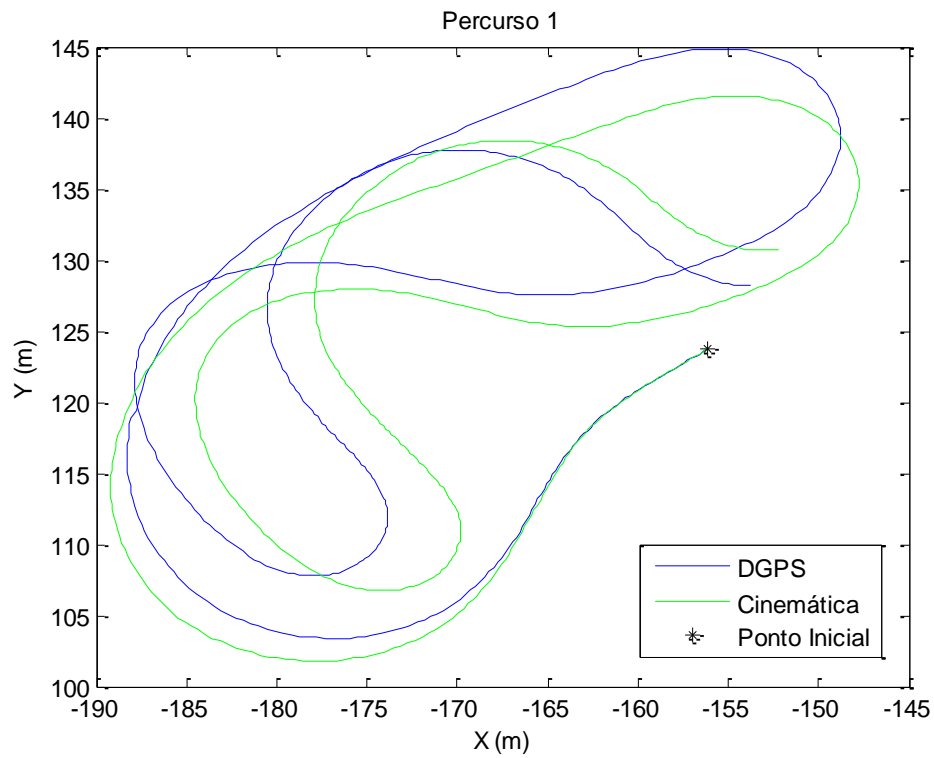


Figura 7.2 - Percurso 1 - Posição do Veículo

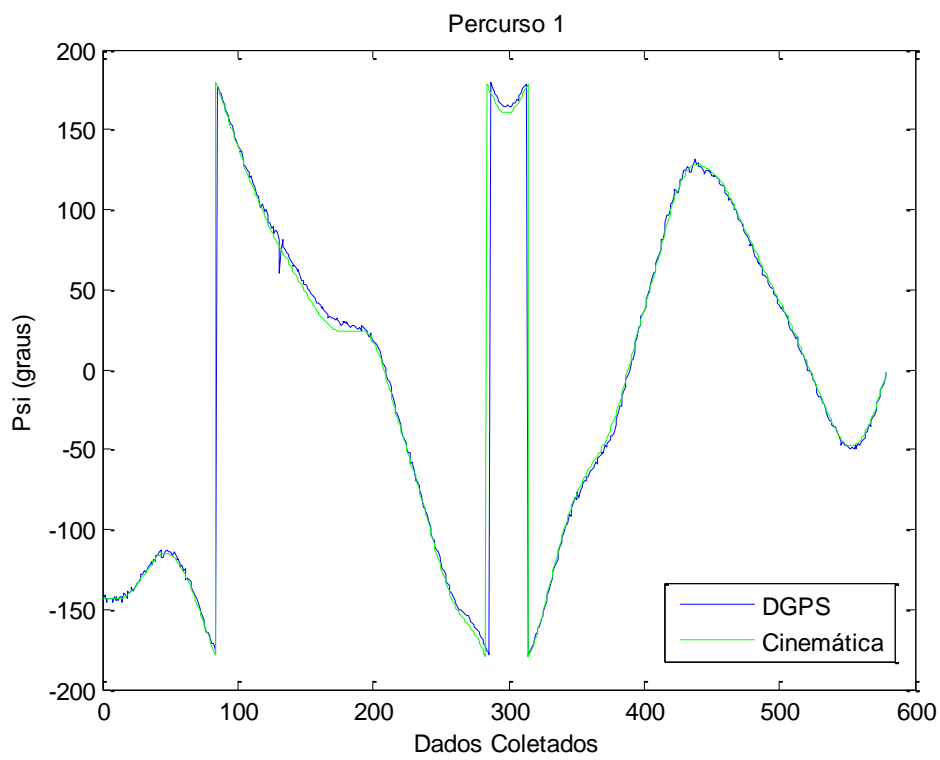


Figura 7.3 - Percurso 1 - Ângulo do Veículo

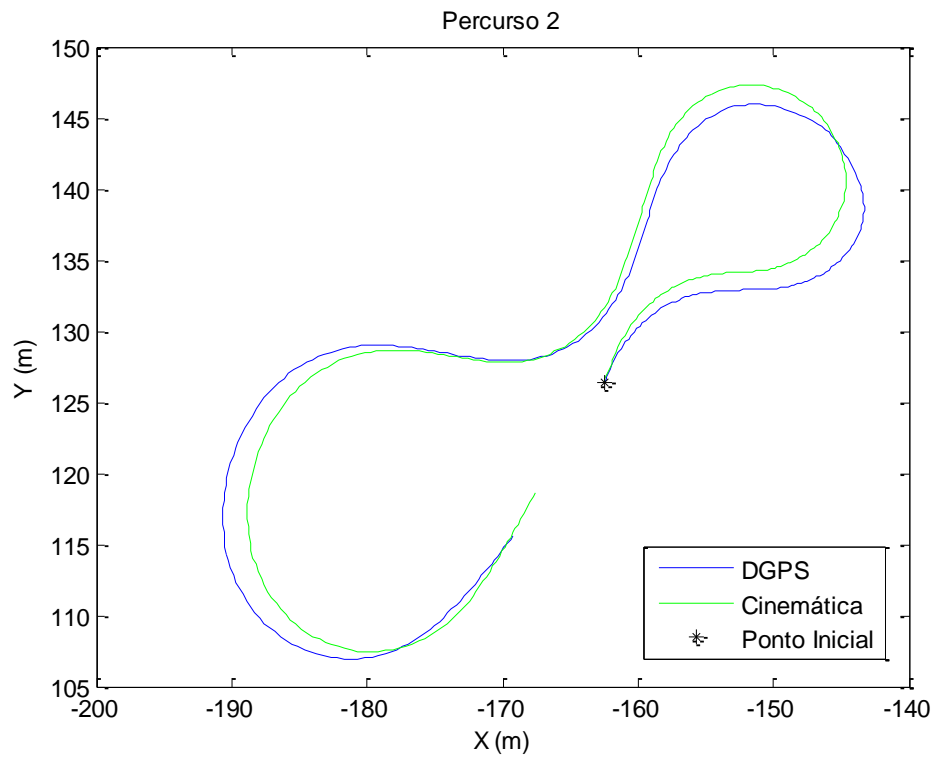


Figura 7.4 - Percurso 2 - Posição do Veículo

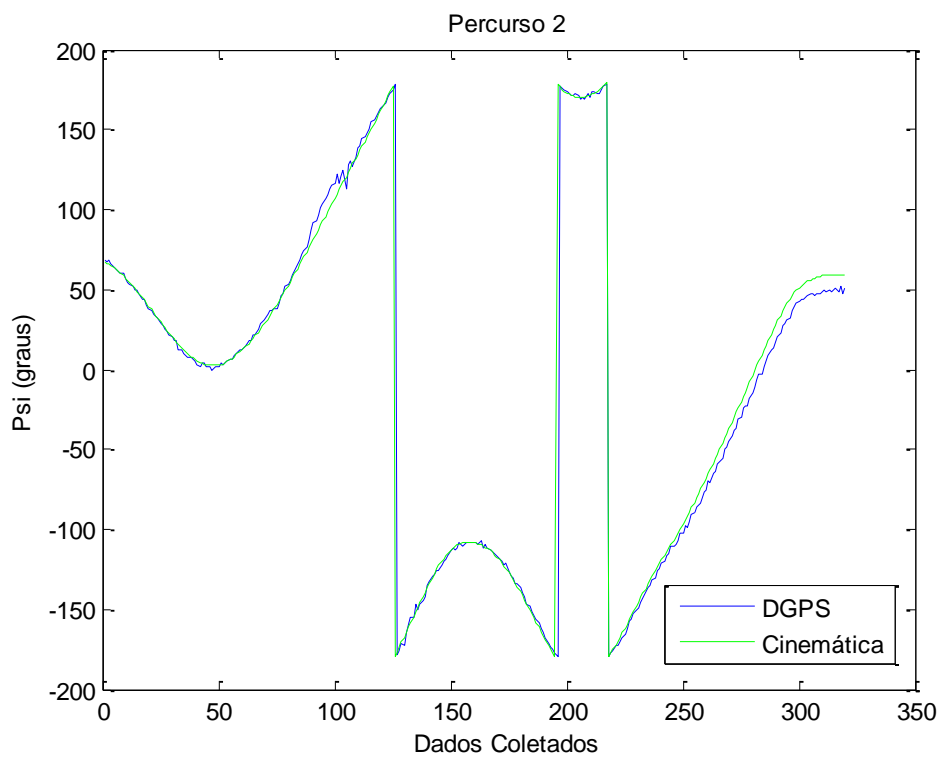


Figura 7.5 - Percurso 2 - Ângulo do Veículo

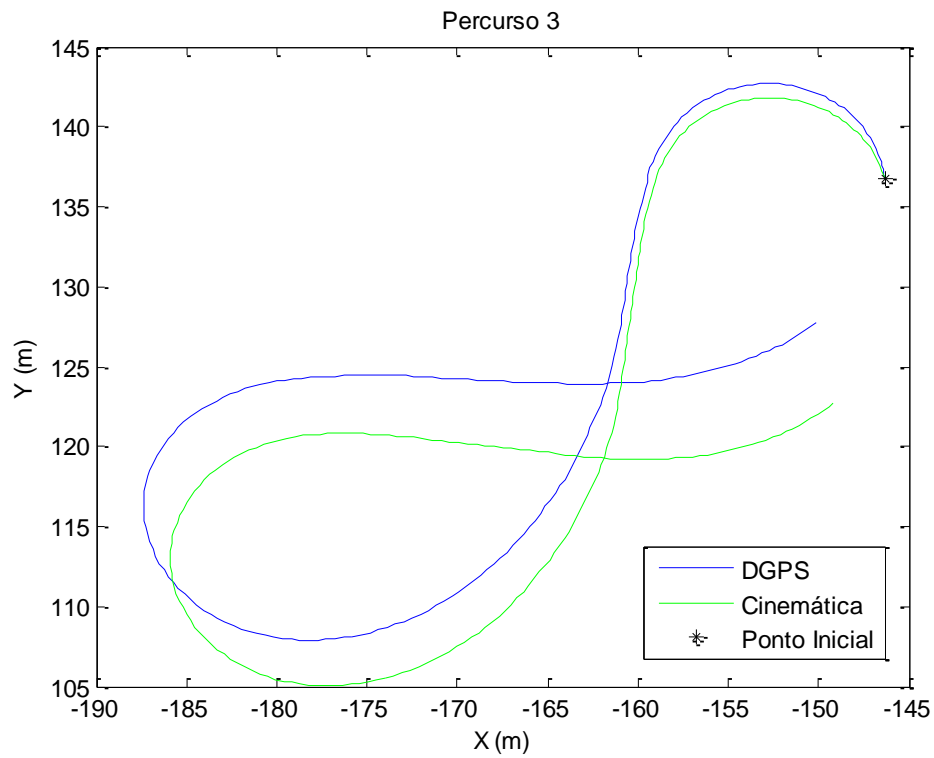


Figura 7.6 - Percurso 3 - Posição do Veículo

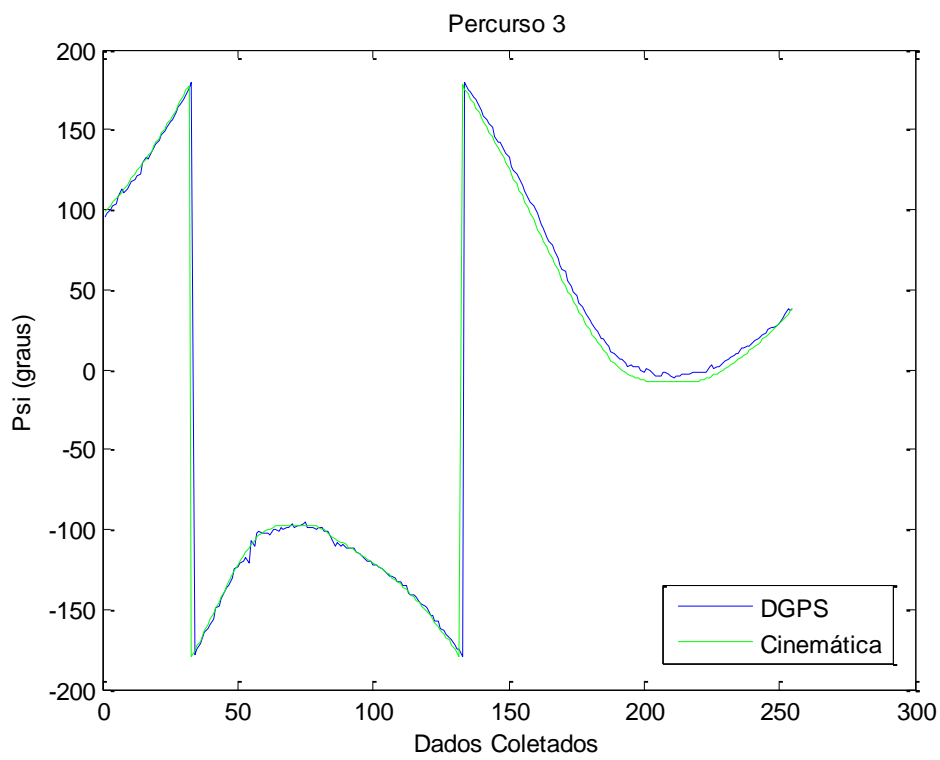


Figura 7.7 - Percurso 3 - Ângulo do Veículo

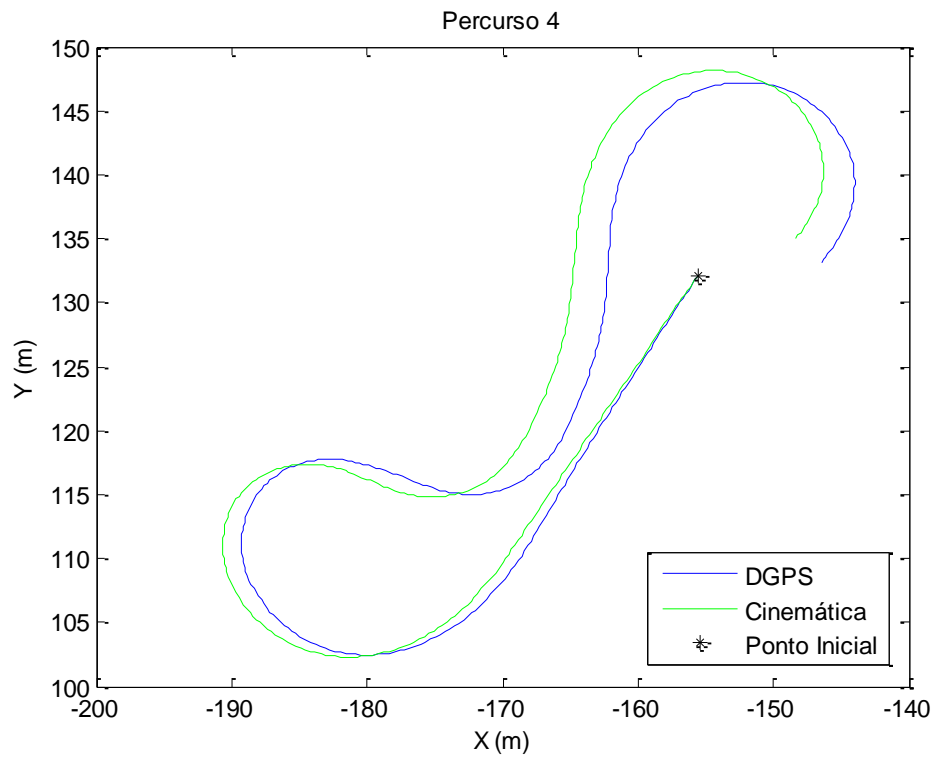


Figura 7.8 - Percurso 4 - Posição do Veículo

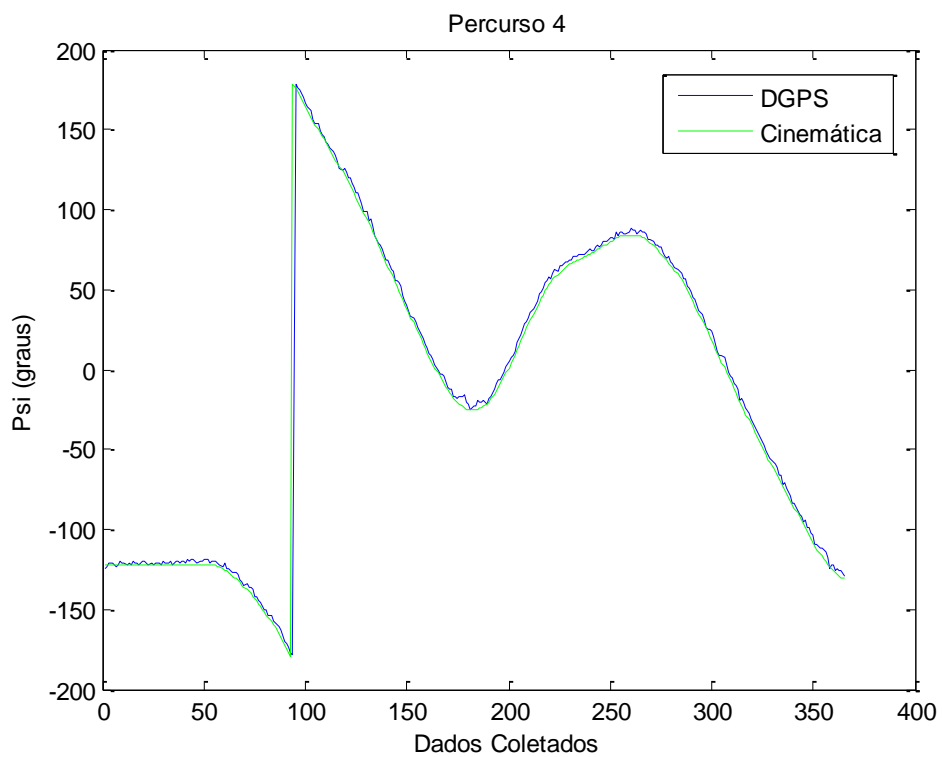


Figura 7.9 - Percurso 4 - Ângulo do Veículo

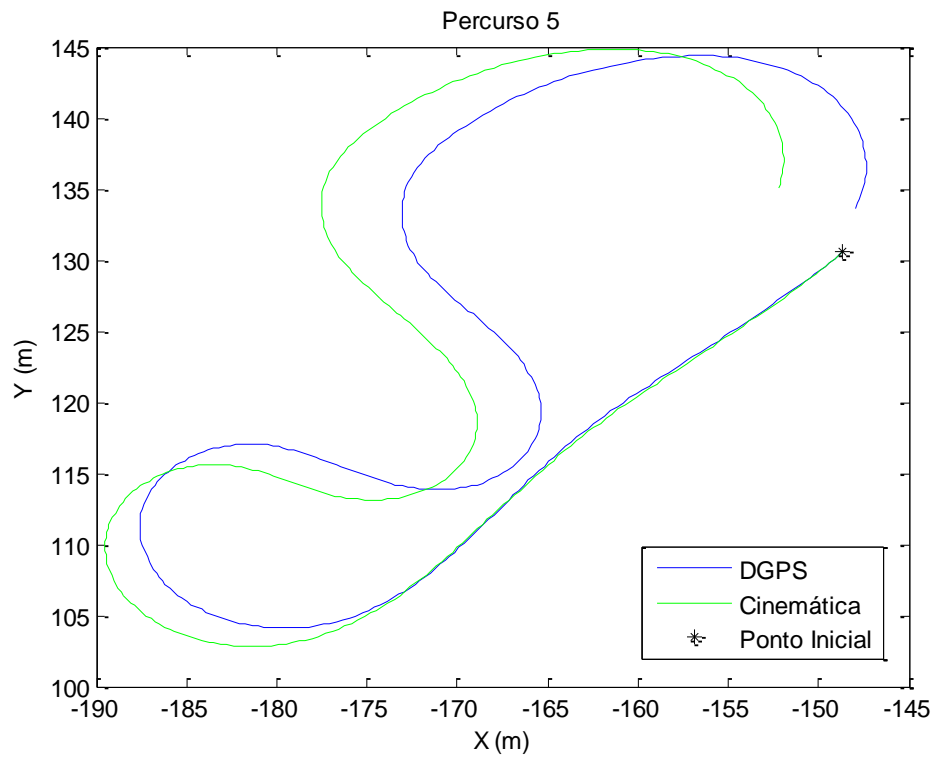


Figura 7.10 - Percurso 5 - Posição do Veículo

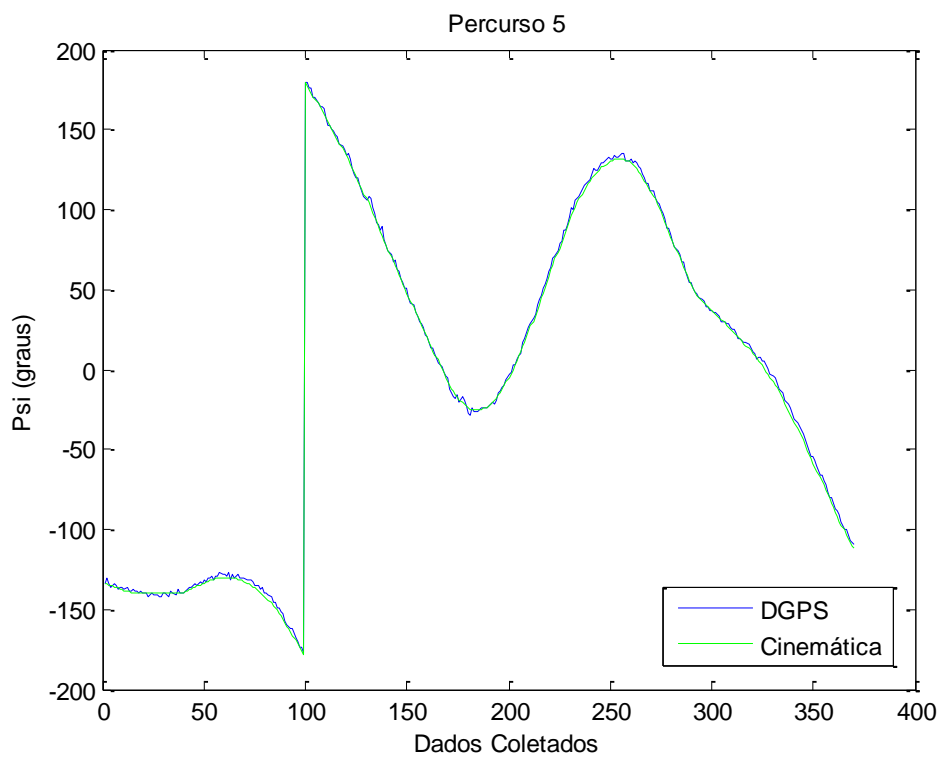


Figura 7.11 - Percurso 5 - Ângulo do Veículo

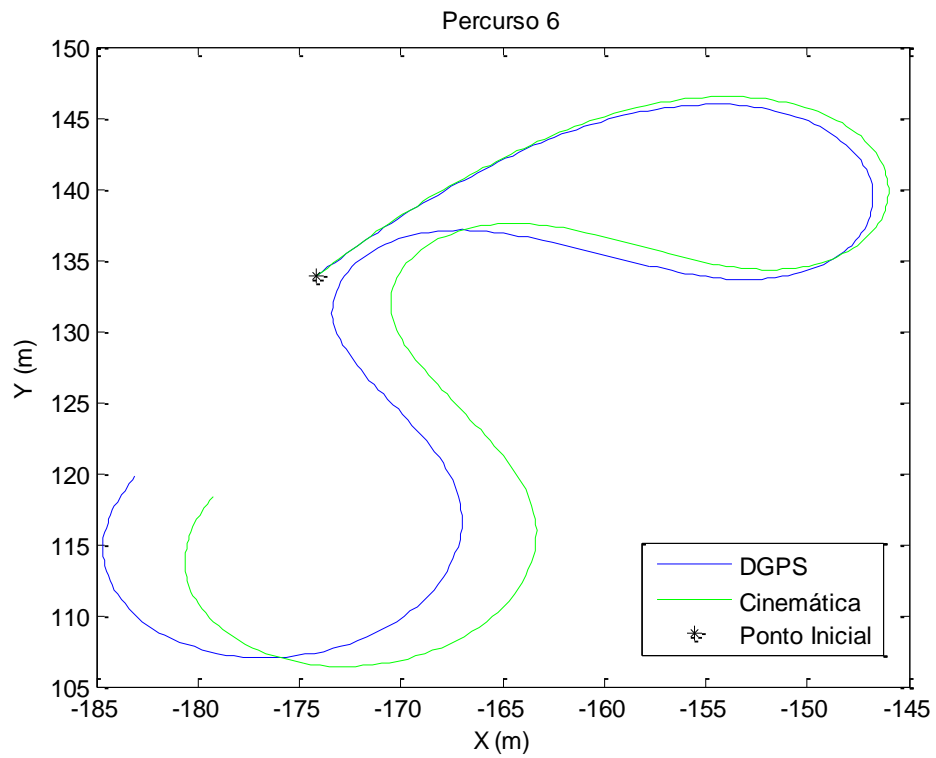


Figura 7.12 - Percurso 6 - Posição do Veículo

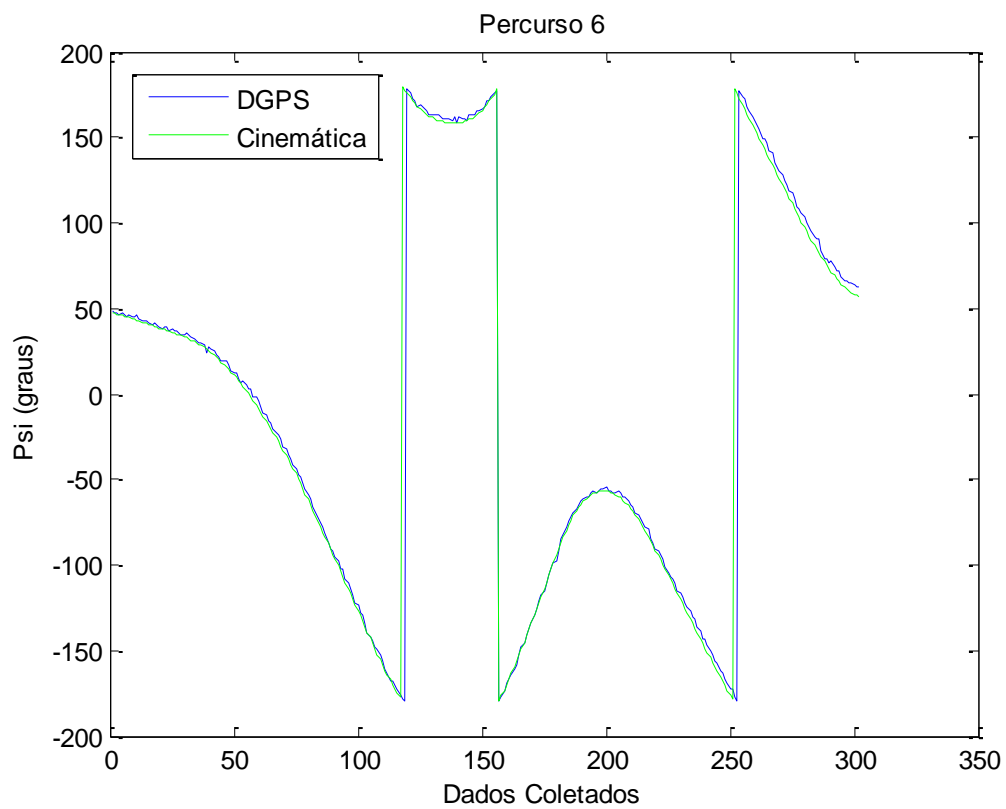


Figura 7.13 - Percurso 6 - Ângulo do Veículo

7.3 Pré-Processamento

Os dados coletados para cada trajetória passam pelo Filtro α - β descrito na Seção 4.2. Os valores calculados com o Filtro serão a solução que a rede neural terá que encontrar, ou seja, é a saída desejada fornecida pelo supervisor (professor) à rede. As Figuras 7.14, 7.15, 7.16, 7.17, 7.18, 7.19, 7.20, 7.21, 7.22, 7.23, 7.24 e 7.25, mostram o resultado do Filtro para cada percurso.

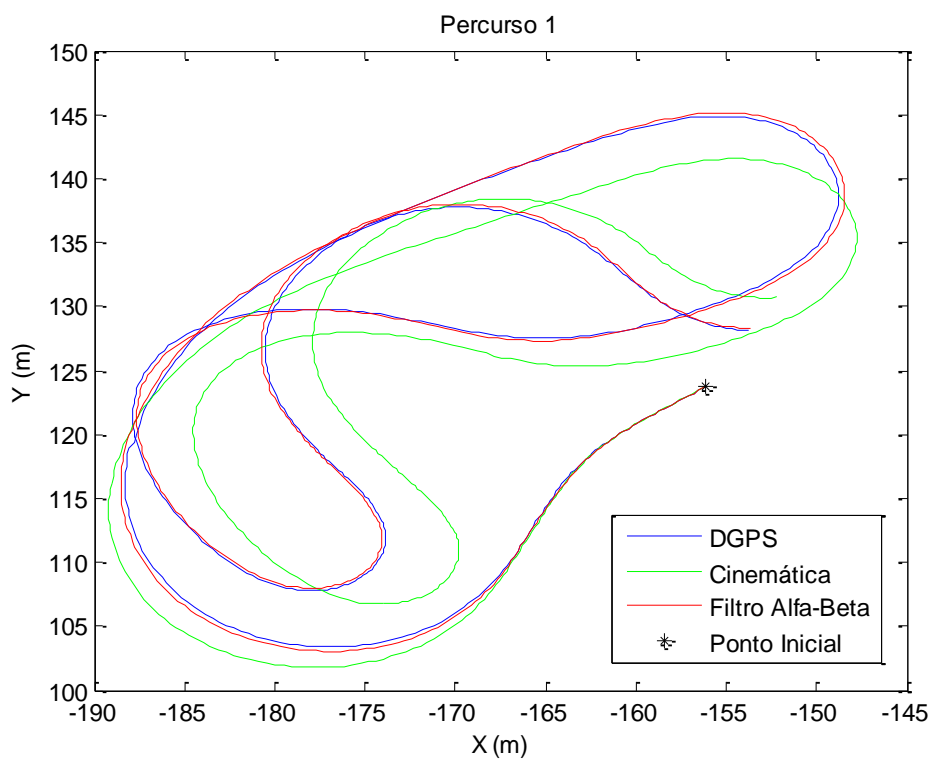


Figura 7.14 - Percurso 1 - Posição do Veículo com Filtro

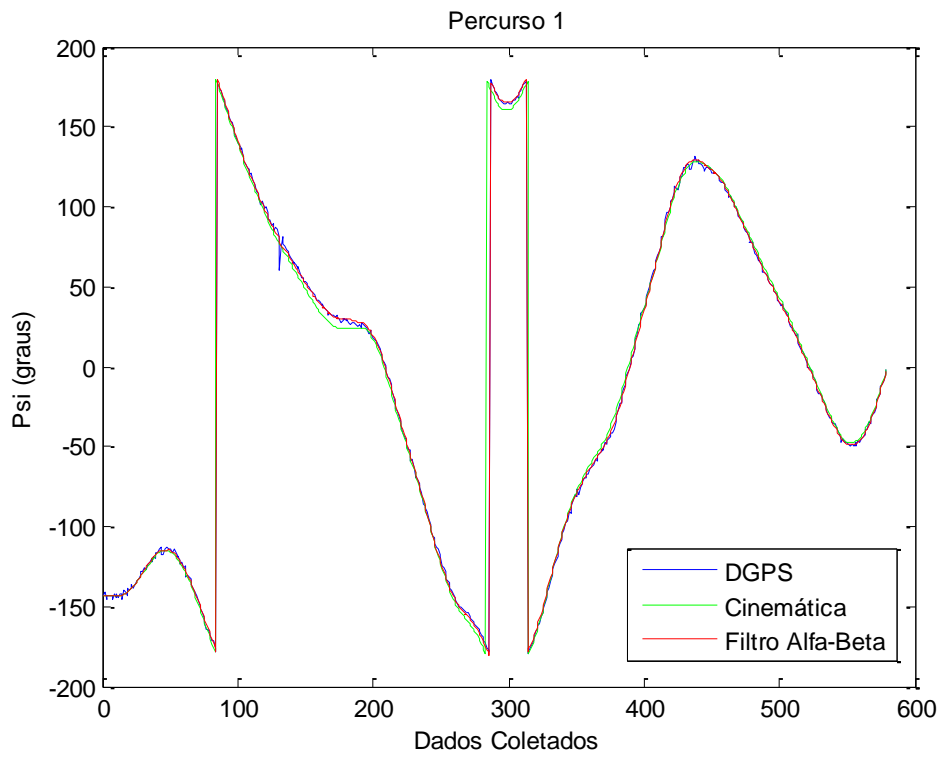


Figura 7.15 - Percurso 1 - Ângulo do Veículo com Filtro

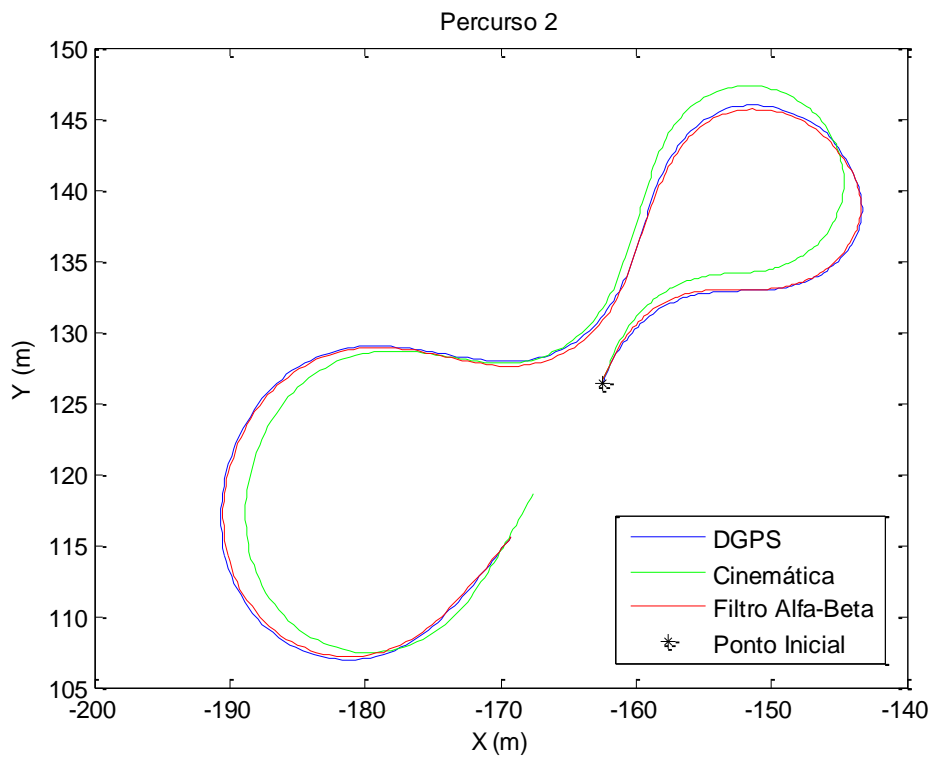


Figura 7.16 - Percurso 2 - Posição do Veículo com Filtro

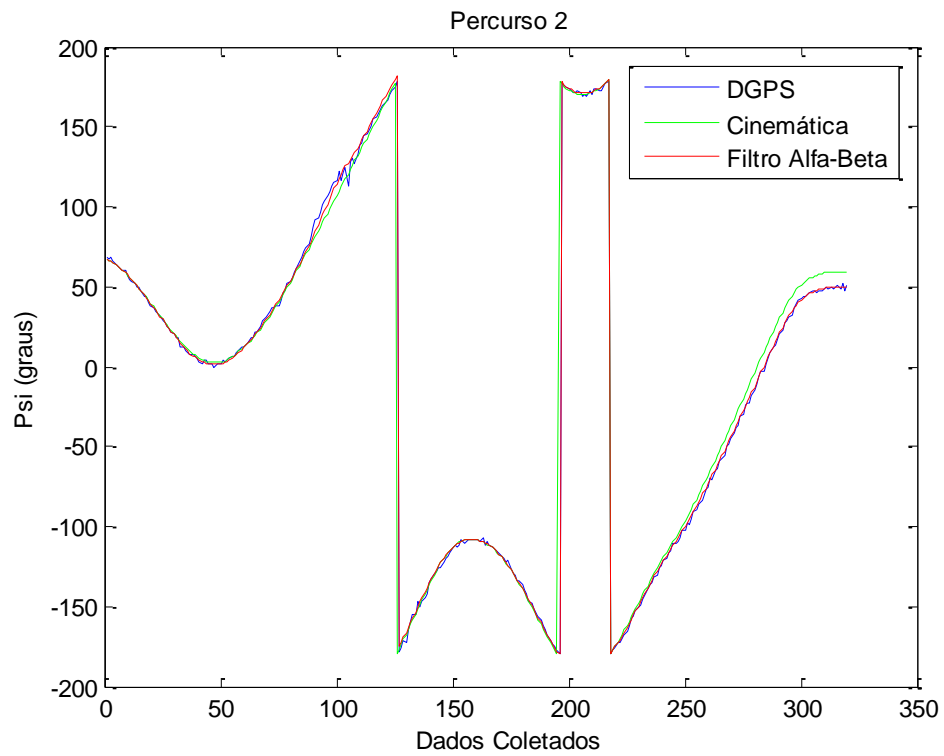


Figura 7.17 - Percurso 2 - Ângulo do Veículo com Filtro

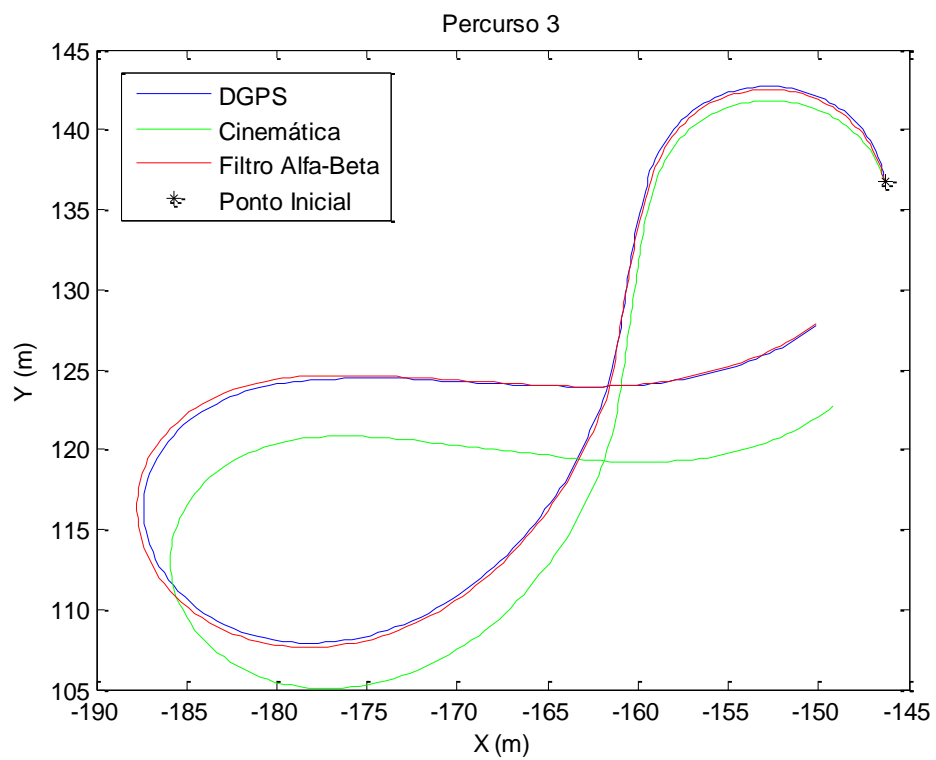


Figura 7.18 - Percurso 3 - Posição do Veículo com Filtro

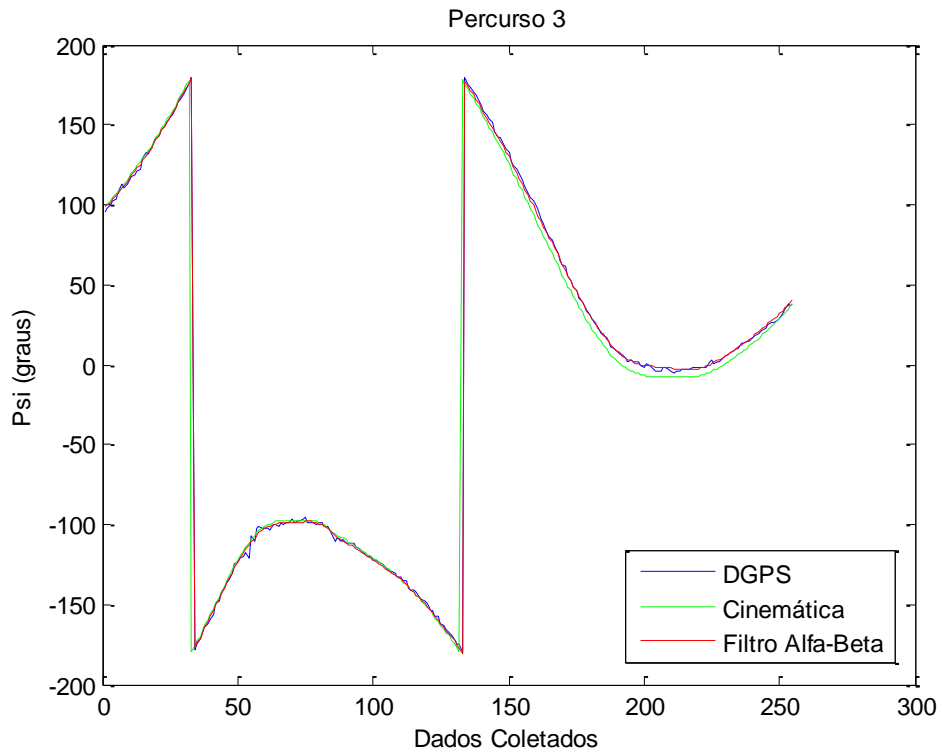


Figura 7.19 - Percurso 3 - Ângulo do Veículo com Filtro

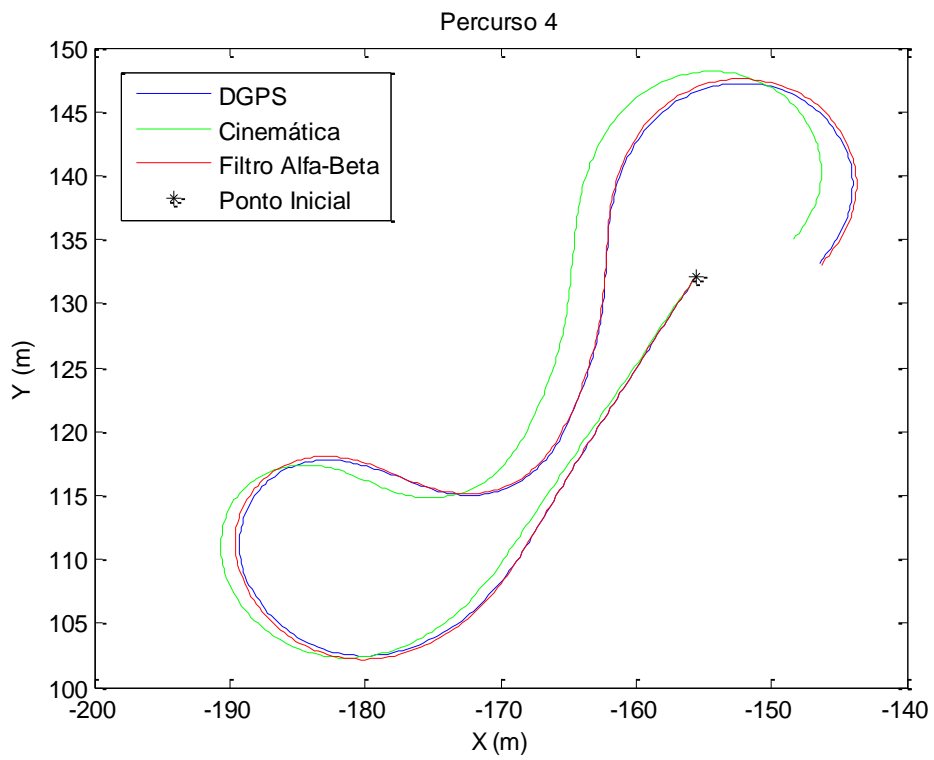


Figura 7.20 - Percurso 4 - Posição do Veículo com Filtro

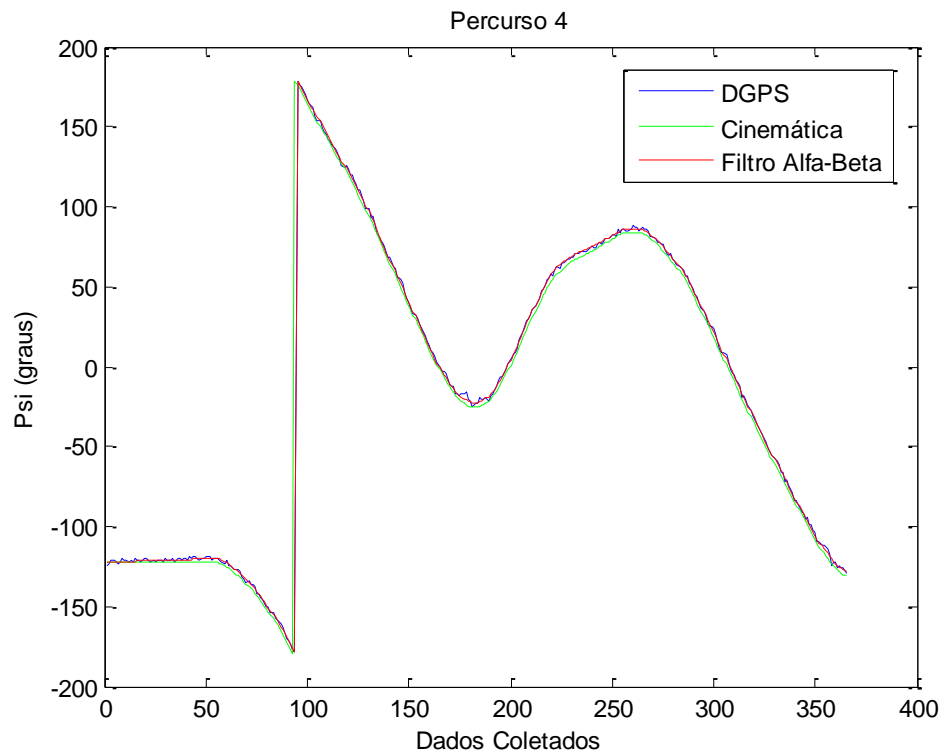


Figura 7.21 - Percurso 4 - Ângulo do Veículo com Filtro

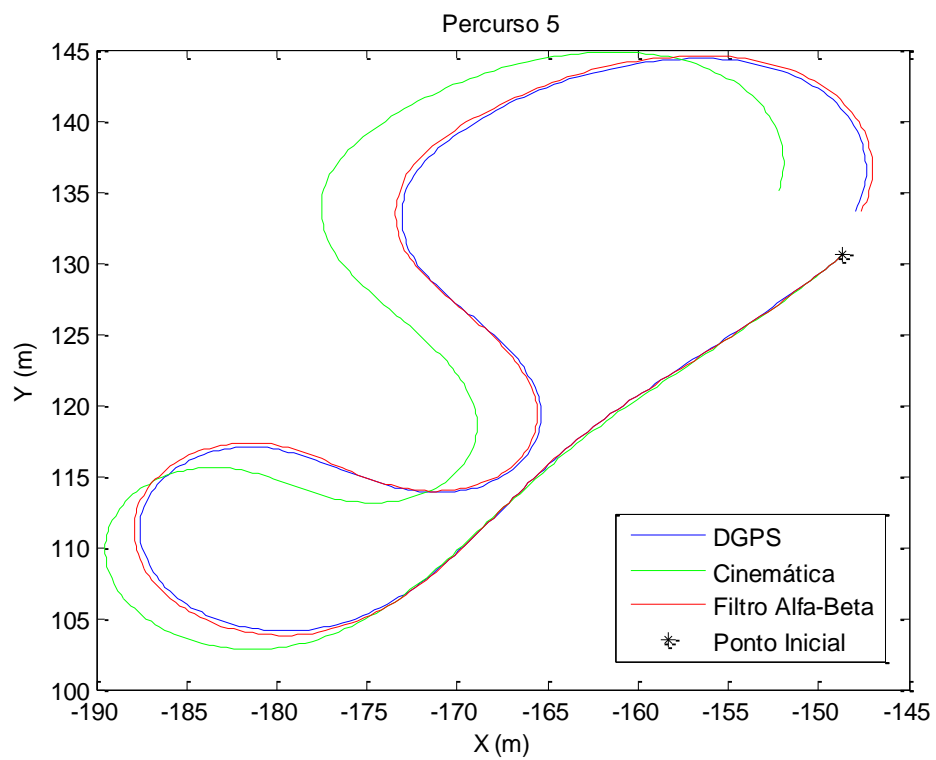


Figura 7.22 - Percurso 5 - Posição do Veículo com Filtro

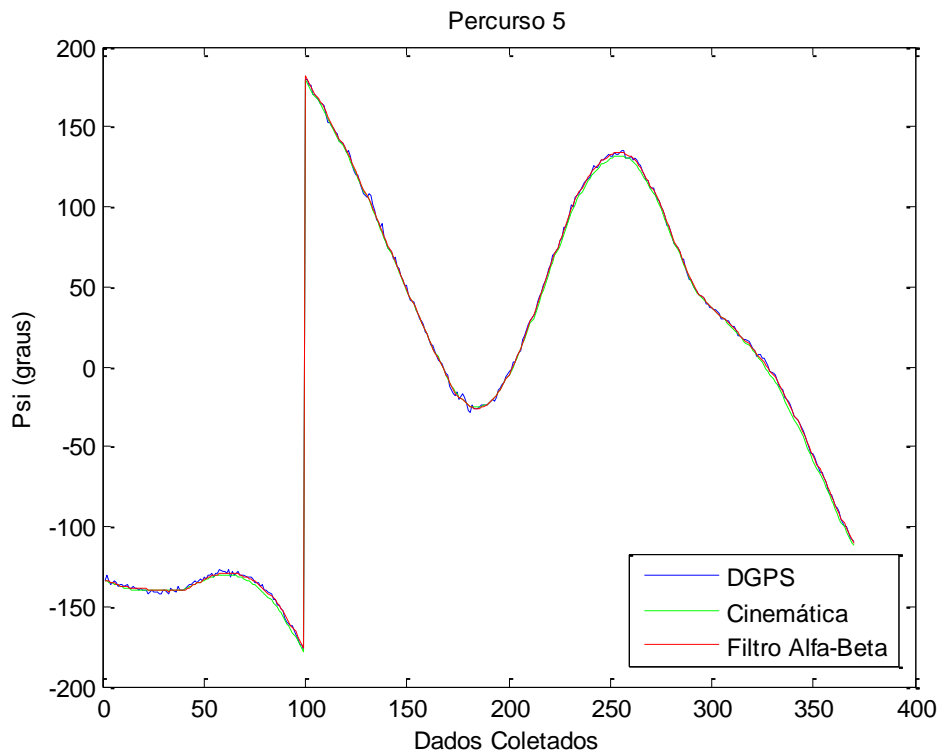


Figura 7.23 - Percurso 5 - Ângulo do Veículo com Filtro

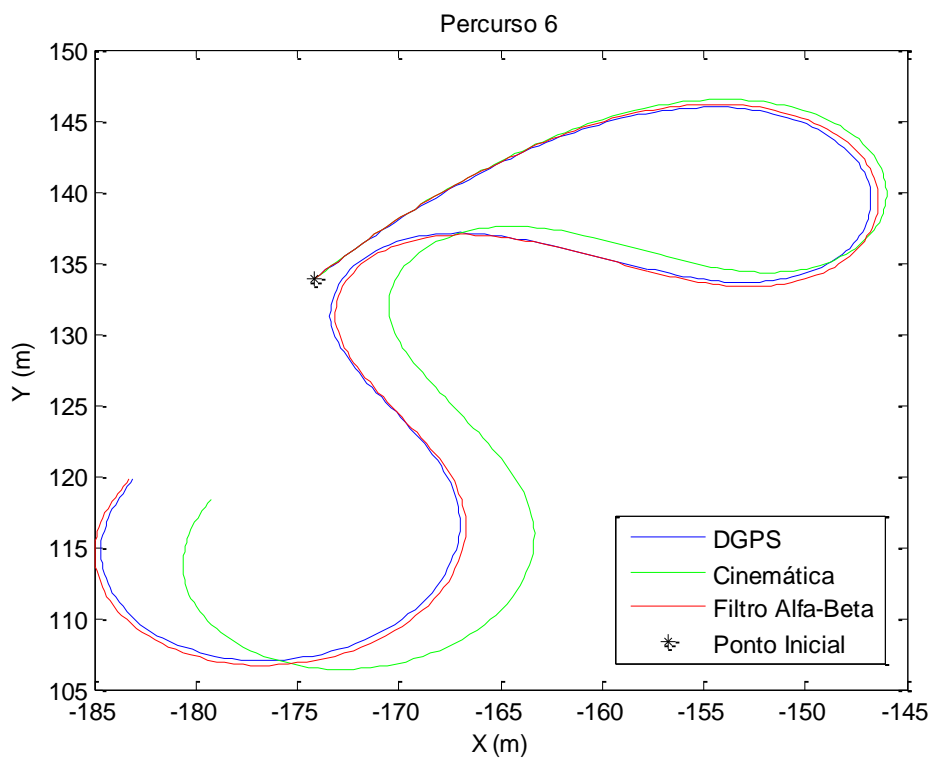


Figura 7.24 - Percurso 6 - Posição do Veículo com Filtro

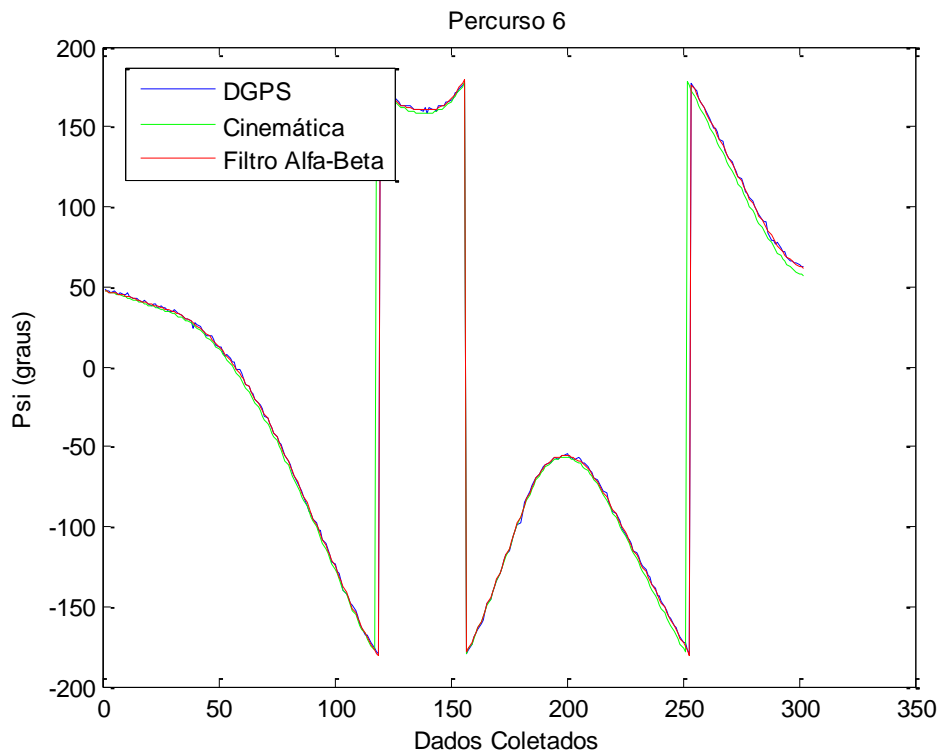


Figura 7.25 - Percurso 6 - Ângulo do Veículo com Filtro

7.4 Projeto da RNA

Os sistemas baseados em redes neurais artificiais dependem fortemente da topologia destas redes, assim como de seus parâmetros. Como resultado, a determinação da arquitetura da rede afeta muito o seu desempenho, ou seja, está relacionada com a velocidade e a exatidão do aprendizado e com a capacidade de generalização. É grande a dificuldade de se projetar redes neurais eficientes. Existem algumas técnicas, nas quais são envolvidos alguns conhecimentos empíricos, que serão abordadas nesta seção.

Neste trabalho duas redes neurais foram propostas, uma chamada de Rede Neural e a outra de Rede Neuro-Especialista. Essas redes serão descritas, comparadas e analisadas nas próximas seções.

7.4.1 Definição dos Parâmetros da Rede Neural

É quase impossível especificar uma arquitetura eficiente, dadas as especificações de um problema. Isso pode ser feito através da experimentação, ou seja, empiricamente.

7.4.1.1 Número de Camadas Escondidas

A rede neural contém uma ou mais camadas de neurônios ocultos que não são parte da entrada ou saída da rede. Esses neurônios escondidos ativam a rede para aprender tarefas complexas, extraindo características cada vez mais significativas dos padrões de entrada.

Uma arquitetura do tipo multicamadas pode ser vista como uma ferramenta prática geral para fazer um mapeamento não-linear de entrada-saída. Especificamente, seja m o número de entradas da rede e n o número de saídas. A relação entrada-saída da rede define um mapeamento de um espaço euclidiano de entrada m -dimensional para um espaço euclidiano de saída n -dimensional, que é infinitamente continuamente diferenciável (20).

Teorema da Aproximação Universal: Seja $f(\cdot)$ uma função contínua não-constante, limitada e monotonicamente crescente. Seja I_m um hipercubo unitário p -dimensional $[0,1]^m$. O espaço das funções contínuas em I_m é denominado $C(I_m)$. Então, dada qualquer função $g \in C(I_m)$ e $\varepsilon > 0$, existe um inteiro M e conjuntos de constantes reais α_1 e w_{ij} , onde $i = 1, \dots, M$ e $j = 1, \dots, m$, tais que pode-se definir

$$F(x_1, x_2, \dots, x_m) = \sum_{i=1}^M \alpha_i f \left(\sum_{j=1}^m w_{ij} x_j - \theta \right)$$

como uma aproximação da função $g(\cdot)$ tal que,

$$|F(x_1, x_2, \dots, x_m) - g(x_1, x_2, \dots, x_m)| < \varepsilon$$

para todo $\{x_1, \dots, x_m\} \in I_m$.

O teorema afirma que uma rede neural multicamadas com uma única camada intermediária é capaz de realizar uma aproximação uniforme, dado um conjunto de treinamento representado pelo conjunto de entradas x_1, \dots, x_m e uma saída desejada $f(x_1, \dots, x_m)$. No entanto, o teorema não diz que uma única camada escondida é ótima no sentido de tempo de aprendizagem, facilidade de implementação ou generalização.

7.4.1.2 Número de Neurônios na Camada Escondida

Decidir o número de neurônios nas camadas escondidas é uma parte muito importante de decidir a arquitetura geral da rede neural. Embora essas camadas não interajam diretamente

com o ambiente externo, elas têm uma enorme influência sobre o resultado final. Tanto o número de camadas ocultas e o número de neurônios em cada uma destas camadas devem ser cuidadosamente considerados.

Usar poucos neurônios nas camadas escondidas pode resultar em algo chamado *underfitting*. *Underfitting* ocorre quando a rede não converge durante o seu treinamento e gasta tempo em excesso tentando encontrar uma representação ótima. Usar muitos neurônios nas camadas escondidas pode resultar em *overfitting*. *Overfitting* ocorre quando a rede memoriza os dados de treinamento, ao invés de extrair as características gerais que permitirão a generalização. Devido a estas dificuldades é recomendado manter o número de neurônios escondidos baixo, mas não tão baixo quanto o estritamente necessário (25).

Existem muitos métodos para determinar o número de neurônios nas camadas escondidas. O principal deles envolve definir o número de neurônios em função da dimensão das camadas de entrada e saída da rede. Contudo, nesse trabalho o número de neurônios na camada escondida foi definido empiricamente, como mostra a Figura 7.26, por ter apresentado melhores resultados do que utilizar alguma regra para sua definição.

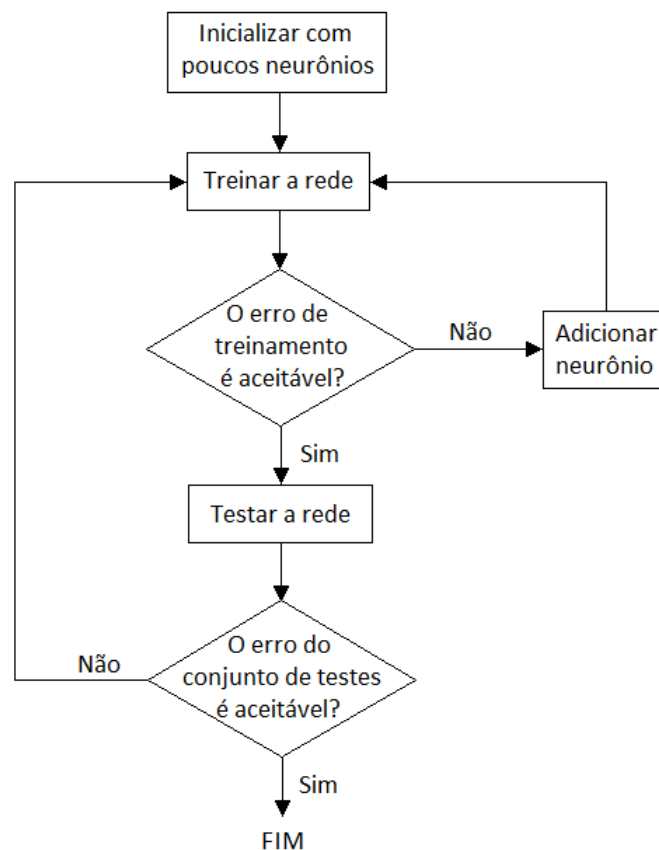


Figura 7.26 - Definição do Número de Neurônios na Camada Escondida

7.4.1.3 Número de Iterações

As redes neurais podem sofrer excesso de treinamento (*overtraining*). Existe um quantidade ideal de treinamento e passar desse ponto melhora o desempenho do conjunto de treinamento, mas degrada o desempenho da rede em geral. Isso pode ocorrer se a rede e/ou o conjunto de treinamento forem projetados de forma incorreta (26).

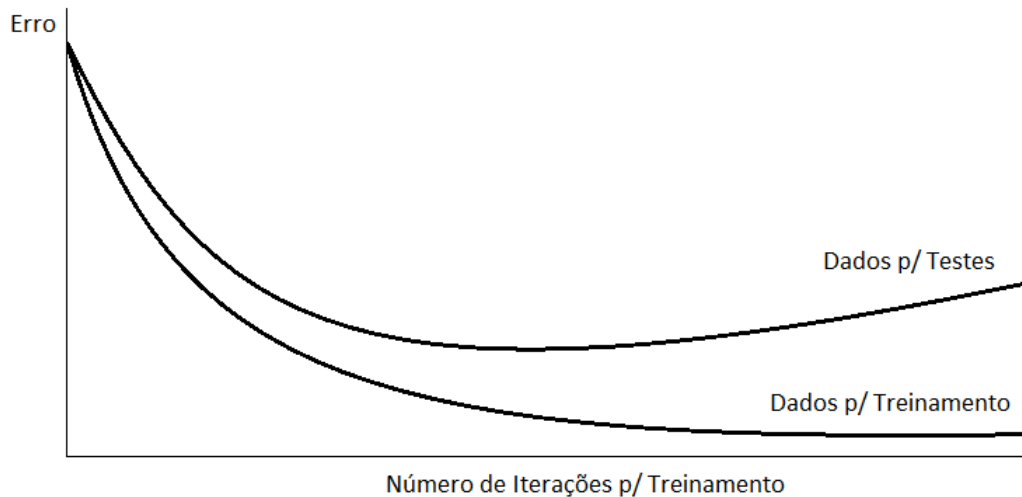


Figura 7.27 - Excesso de Treinamento

A Figura 7.27 mostra o erro de uma rede neural para dois diferentes conjuntos de dados em relação ao número de iterações de treinamento realizadas. Um conjunto de dados é o conjunto de treinamento. Como esperado, o erro para esse conjunto decresce, aproximando-se de uma assíntota. O outro conjunto de dados, chamado de conjunto de validação, é extraída da mesma população como o conjunto de treinamento, mas não é utilizado para o treinamento. Novamente, o erro diminui no início do treinamento. Mas, dando continuidade ao treinamento, depois de um certo ponto o erro do conjunto de validação passa a aumentar.

A solução aparentemente óbvia, é de parar o treinamento quando o erro do conjunto de validação começa a crescer. Outra solução seria corrigir a rede reduzindo o número de neurônios ocultos, a ponto de a rede não ter a capacidade de aprender as idiossincrasias, ou aumentar o tamanho e a diversidade do conjunto de treinamento.

7.4.2 Arquitetura da Rede Neural

Com base na definição dos parâmetros da rede neural é possível estruturar sua arquitetura, faltando apenas a definição das entradas e saídas da rede. Como o objetivo da rede neural é aproximar uma função, suas entradas serão as variáveis da função e a saída o resultado da equação. As funções que serão aproximadas estão descritas na Seção 3.1 como as equações do movimento cinemático. Como dt é constante para uma amostragem com frequência de 10 Hz, as funções foram reescritas a seguir:

$$dx = dS \cos(\psi + \beta) \quad (7.1)$$

$$dy = dS \sin(\psi + \beta) \quad (7.2)$$

$$d\psi = \frac{dS \cos \beta}{l_f + l_r} \tan \delta \quad (7.3)$$

onde

$$\beta = \tan^{-1} \left(\frac{l_r}{l_f + l_r} \tan(\delta) \right)$$

dx e dy são as variações das coordenadas globais X e Y , respectivamente;

$d\psi$ é a variação do ângulo do veículo ψ ;

dS é o espaço percorrido pelo veículo;

δ é o ângulo da roda;

$l_f + l_r$ é o entre eixos do veículo.

Sendo assim, as arquiteturas das Redes Neurais estão representadas na Figura 7.28. A primeira rede neural representa a aproximação da função 7.3 e é composta por 2 entradas, 1 camada escondida com 11 neurônios e 1 saída. A segunda rede neural representa a aproximação das funções 7.1 e 7.2 e é composta por 3 entradas, 1 camada escondida com 13 neurônios e 2 saídas.

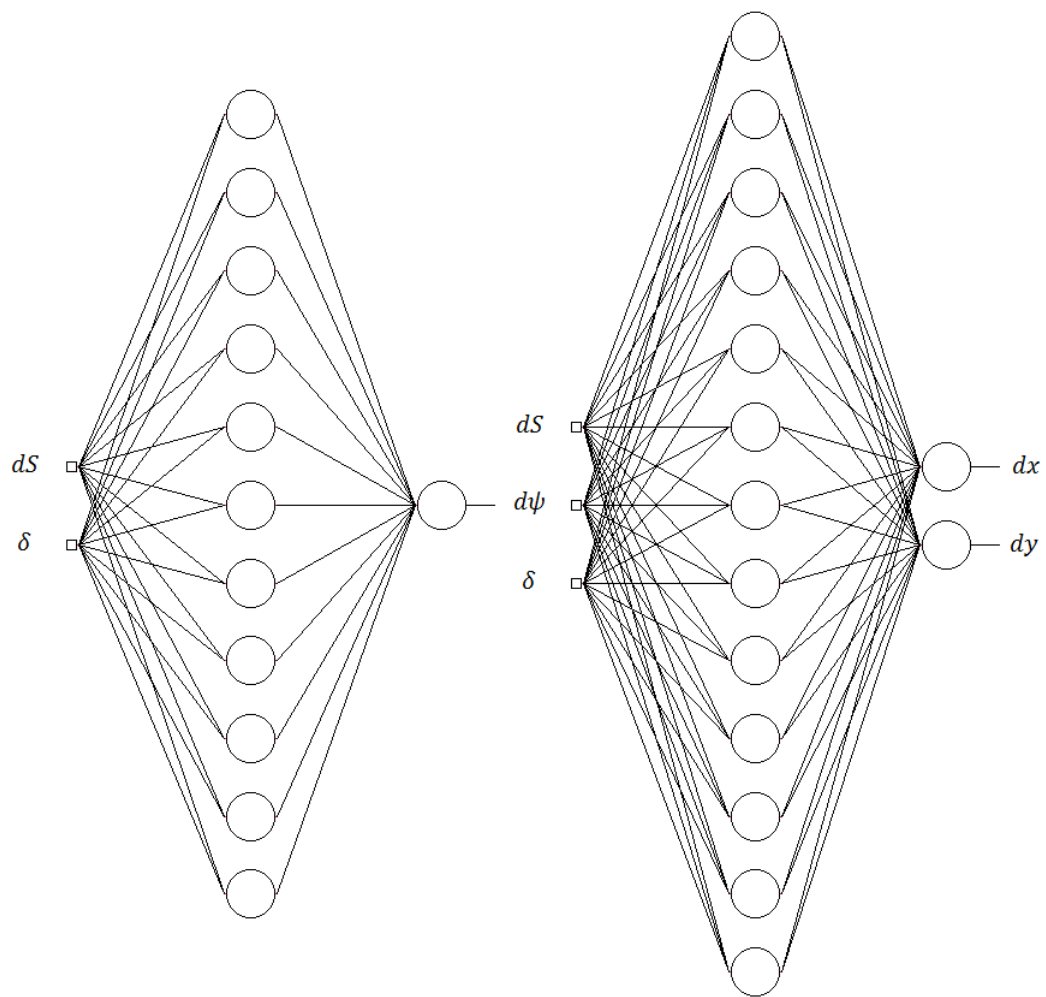


Figura 7.28 - Arquitetura da Rede Neural

7.4.3 Arquitetura da Rede Neuro-Especialista

A seção anterior descreve uma das redes neurais utilizadas neste trabalho. A outra rede que será utilizada se trata de uma Rede Neuro-Especialista. Na Rede Neuro-Especialista a função de ativação de parte dos neurônios são as equações cinemáticas do veículo. Com esta informação presente dentro da rede, os pesos restantes conseguem reduzir a perda de informação e, desta forma, mapear melhor o conhecimento real do comportamento do veículo.

Desta forma, com parte do conhecimento já mapeado, a rede deverá apenas aprender o restante. No caso da cinemática, todas as fórmulas que representam grande parte do conhecimento e comportamento do veículo são previamente conhecidas. Apesar de mostrarem um erro de longo prazo, preenchem quase todo o conhecimento sobre o sistema.

Incluir isso em uma rede deixa o treinamento mais eficiente, uma vez que nesta situação apenas o restante do conhecimento do processo é que deve ser capturado.

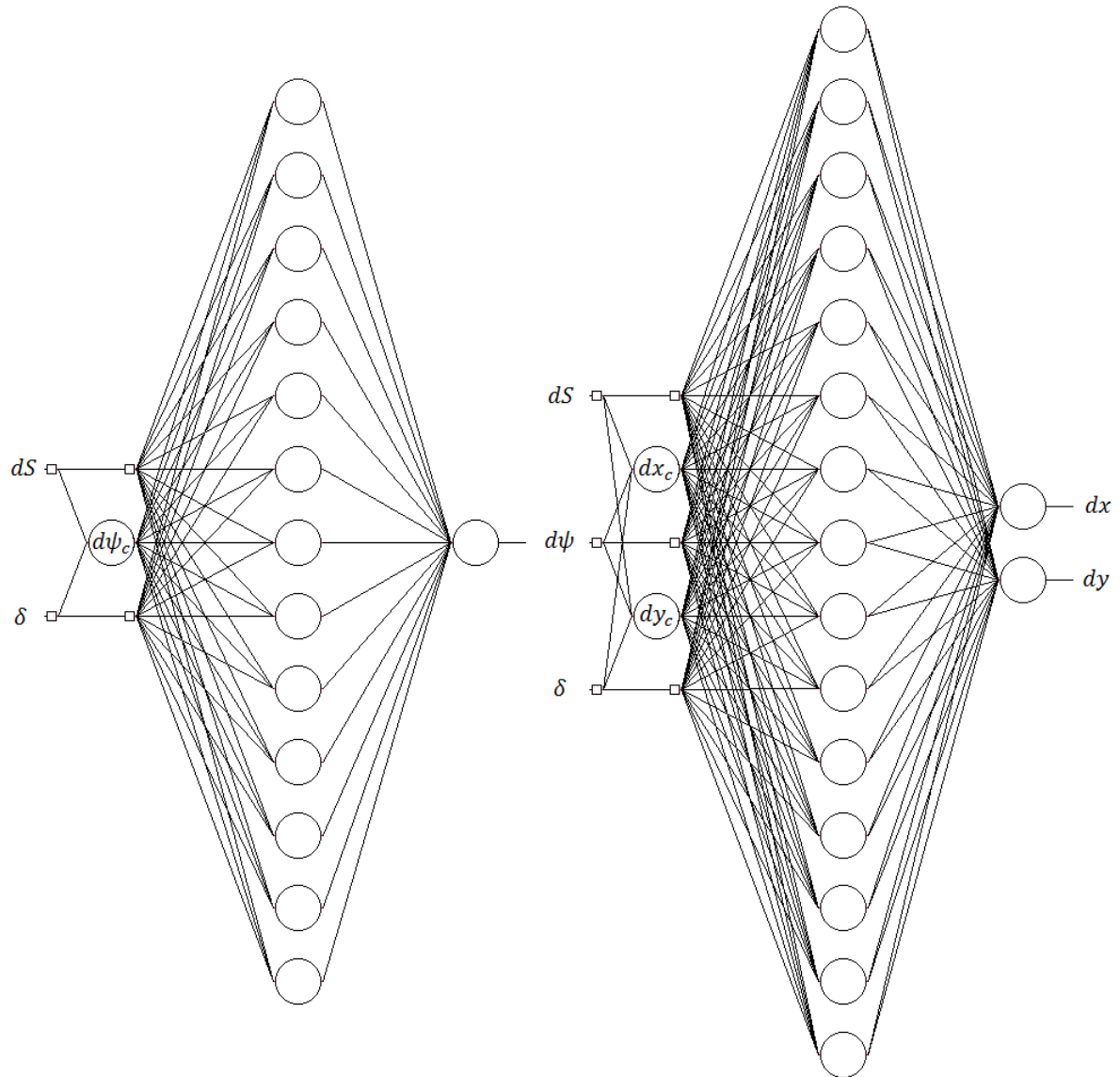


Figura 7.29 - Arquitetura da Rede Neuro-Especialista

Sendo assim, as arquiteturas das Redes Neuro-Especialistas estão representadas na Figura 7.29. A primeira rede neuro-especialista representa a aproximação da função 7.3 e é composta por 2 entradas, 1 camada com 1 neurônio-especialista, 1 camada escondida com 13 neurônios e 1 saída. A segunda rede neuro-especialista representa a aproximação das funções 7.1 e 7.2 e é composta por 3 entradas, 1 camada com 2 neurônios-especialistas, 1 camada escondida com 15 neurônios e 2 saídas.

7.5 Treinamento da RNA

Os pesos sinápticos designados às conexões entre os neurônios artificiais são os responsáveis pelo armazenamento do conhecimento nas RNAs. São eles os parâmetros que devem ser ajustados através do processo de "treinamento", para que a rede habilite-se a responder o mais corretamente possível a quaisquer outros estímulos que lhe forem apresentados em uma fase posterior ao treinamento, denominada de "teste". (20)

Primeiro, é apresentado para a rede um conjunto de dados de treinamento. Esse conjunto de dados contém exemplos de comportamentos observados. O treinamento da rede neural segue a seguinte metodologia: o conjunto de dados de treinamento sem a resposta é apresentado para a rede. A rede gera uma resposta que é comparada com a resposta dos dados de treinamento e os pesos são ajustados. Isso é feito repetidas vezes até que a rede consiga acertar a maioria das respostas, ou seja, quando a taxa de erro for pequena.

Para o treinamento das redes foram utilizados os dados do Percurso 1. Os resultados do treinamento da Rede Neural são apresentados nas Figuras 7.30 e 7.31 e os resultados do treinamento da Rede Neuro-Especialista são apresentados nas Figuras 7.32 e 7.33.

Para realizar o treinamento e os testes das redes neurais utilizou-se uma biblioteca para o software Visual Studio com programação em C#. A biblioteca é conhecida como Aforge.net e é livre. Algumas modificações na biblioteca foram realizadas para se chegar na ferramenta computacional mostrada nas figuras a seguir.

A interface da ferramenta computacional apresenta 3 grupos de dados. O Grupo denominado "Function" apresenta dois gráficos: o gráfico em vermelho representa a função a ser aproximada (resultado esperado) e o gráfico em azul representa a função obtida (resultado da rede neural). O Grupo "Settings" mostra a configuração dos parâmetros da Rede Neural e do Algoritmo Genético. E por fim, o Grupo "Current Iteration" apresenta o número de iterações realizadas e os erros da Rede Neural para tal iteração.

7.5.1 Treinamento da Rede Neural

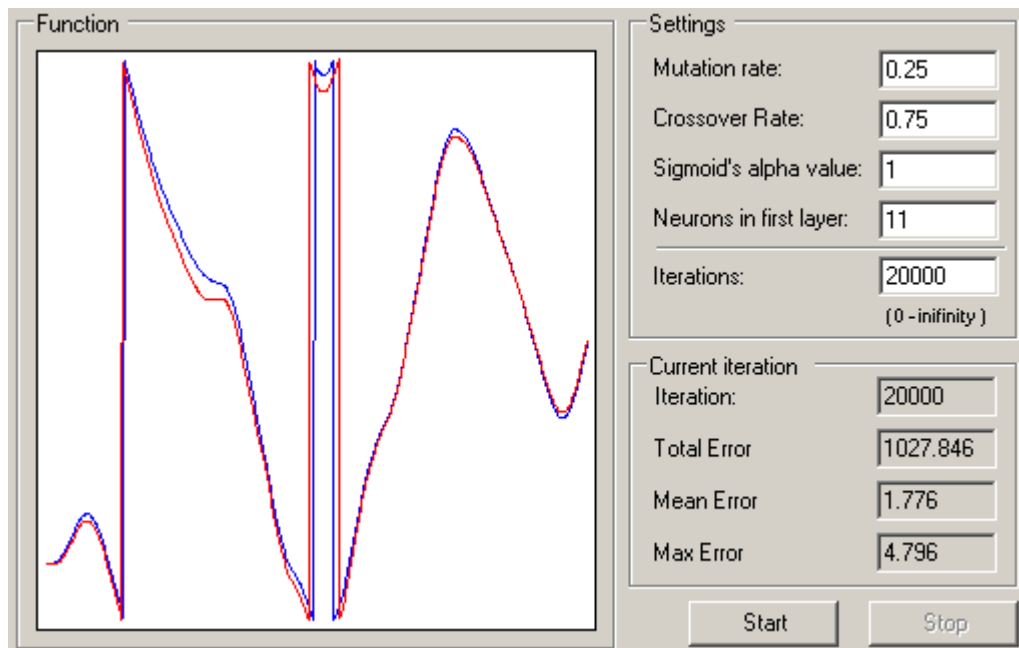


Figura 7.30 - Percurso 1 - Treinamento (Ângulo)

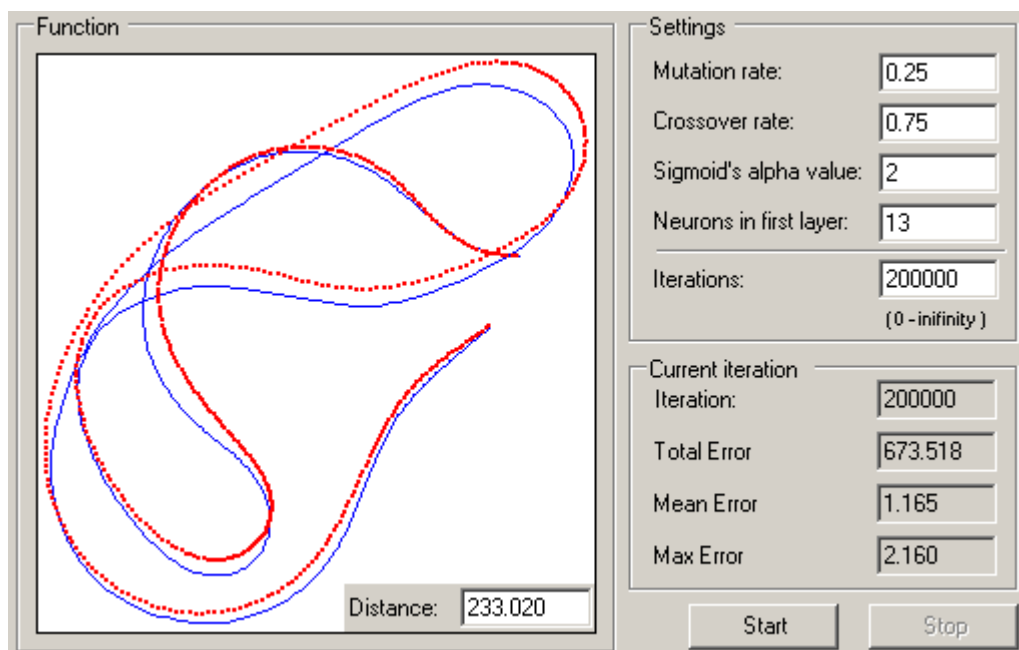


Figura 7.31 - Percurso 1 - Treinamento (Posição)

7.5.2 Treinamento da Rede Neuro-Especialista

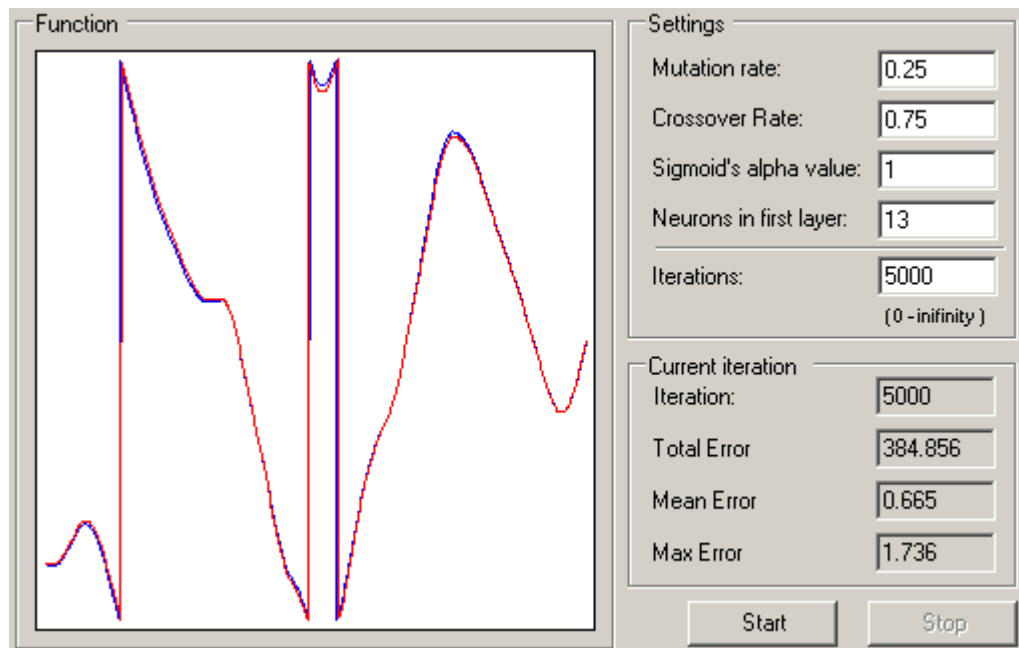


Figura 7.32 - Percurso 1 - Treinamento (Ângulo)



Figura 7.33 - Percurso 1 - Treinamento (Posição)

7.6 Teste da RNA

Depois da fase de aprendizagem, é apresentado à rede neural um conjunto de dados, com informações semelhantes, mas diferentes, dos dados apresentados durante a fase de aprendizagem. Os resultados obtidos serão utilizados para estimar a qualidade das arquiteturas neurais. Para o teste das redes foram utilizados os dados dos Percursos 2, 3, 4, 5 e 6 e os resultados do teste da Rede Neural são apresentados nas Figuras 7.34, 7.35, 7.36, 7.37, 7.38, 7.39, 7.40, 7.41, 7.42 e 7.43 e os resultados do teste da Rede Neuro-Especialista são apresentados nas Figuras 7.44, 7.45, 7.46, 7.47, 7.48, 7.49, 7.50, 7.51, 7.52 e 7.53. O gráfico em vermelho representa o resultado esperado e o gráfico em azul o resultado obtido pela rede neural.

7.6.1 Teste da Rede Neural

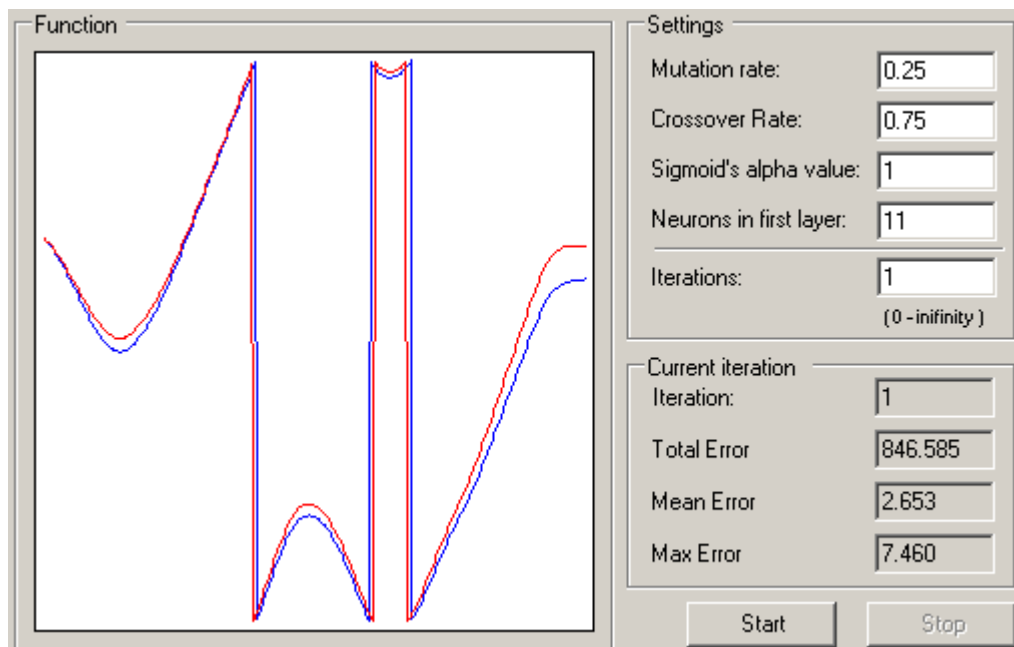


Figura 7.34 - Percurso 2 - Teste (Ângulo)



Figura 7.35 - Percurso 2 - Teste (Posição)

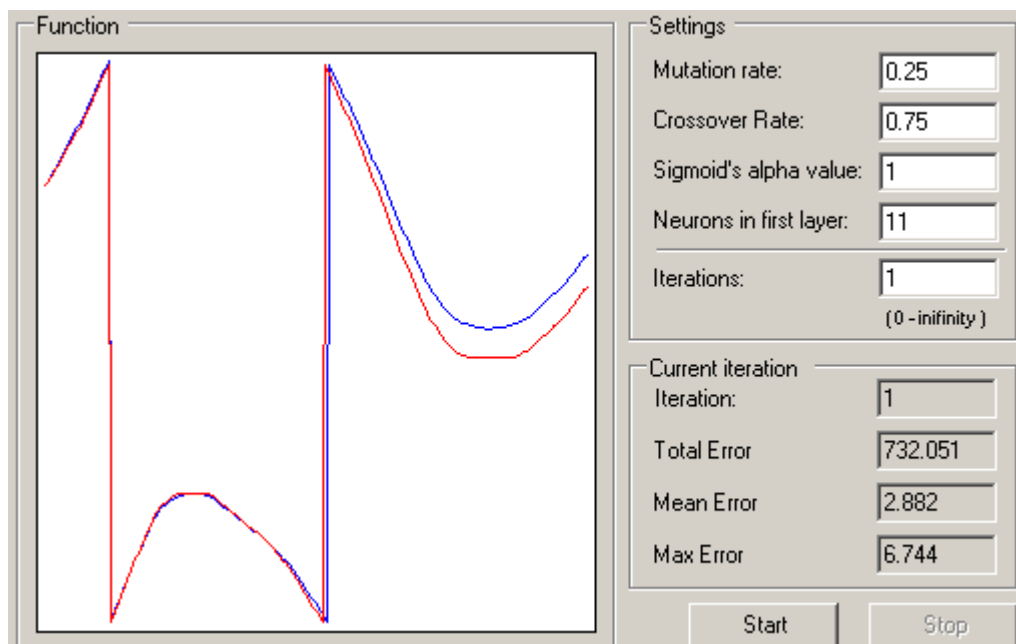


Figura 7.36 - Percurso 3 - Teste (Ângulo)



Figura 7.37 - Percurso 3 - Teste (Posição)

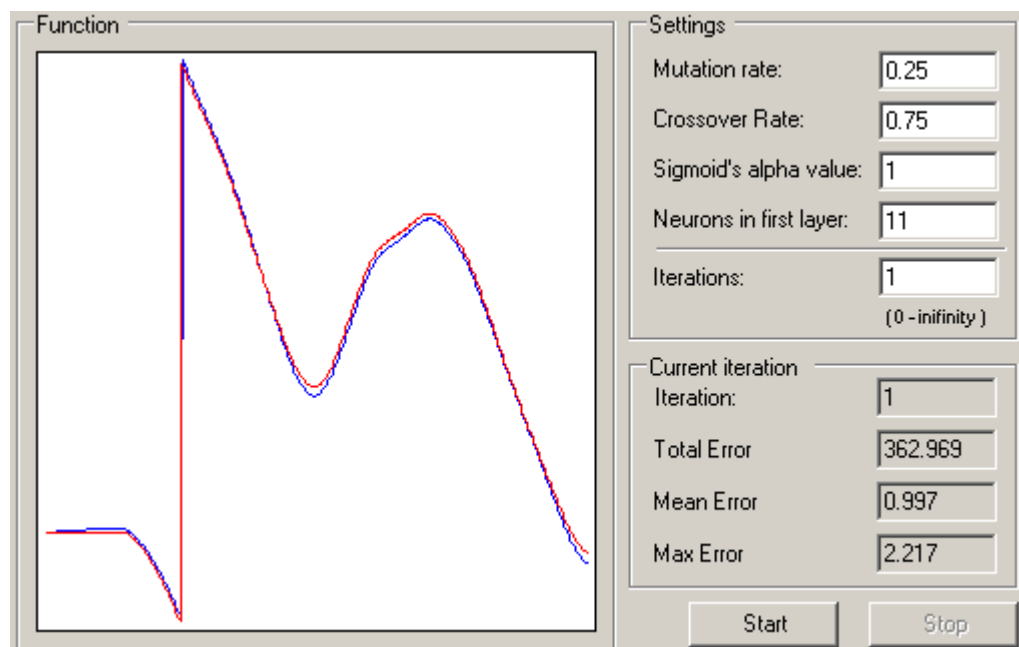


Figura 7.38 - Percurso 4 - Teste (Ângulo)



Figura 7.39 - Percurso 4 - Teste (Posição)

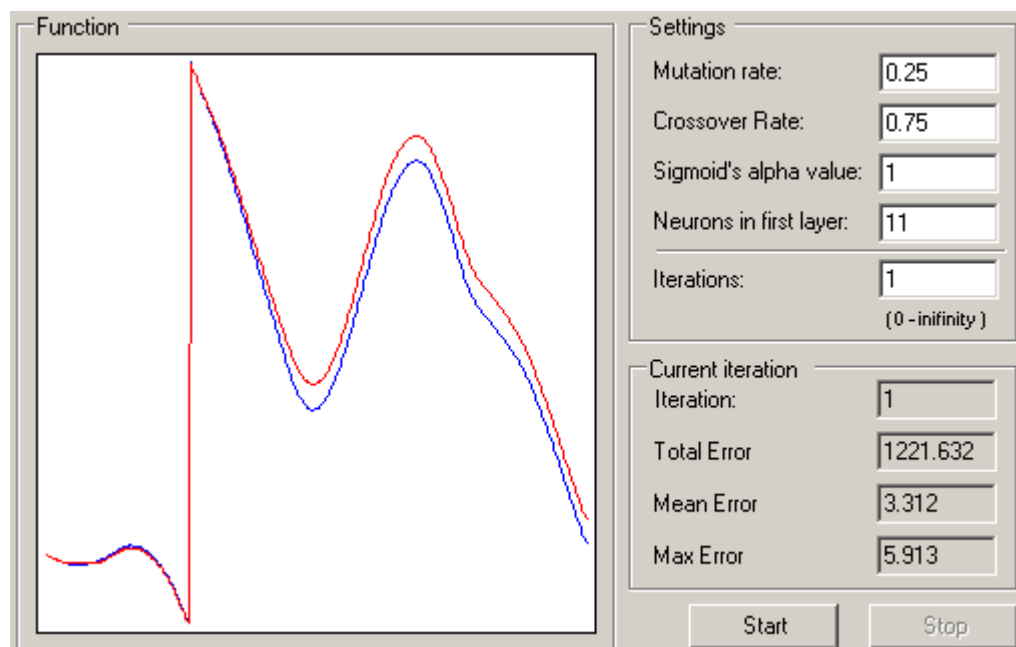


Figura 7.40 - Percurso 5 - Teste (Ângulo)

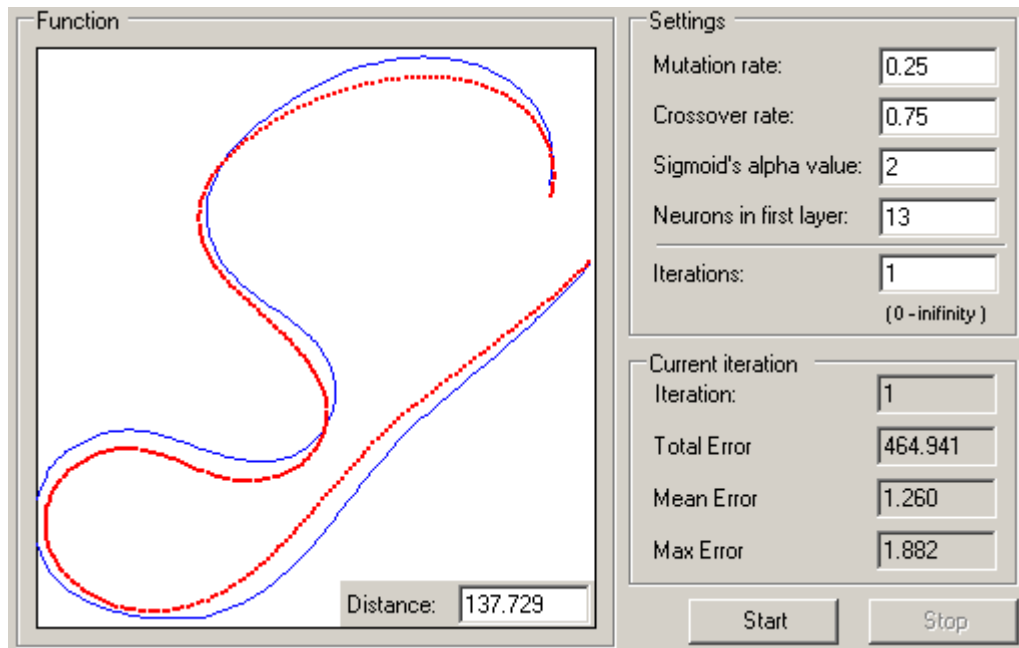


Figura 7.41 - Percurso 5 - Teste (Posição)

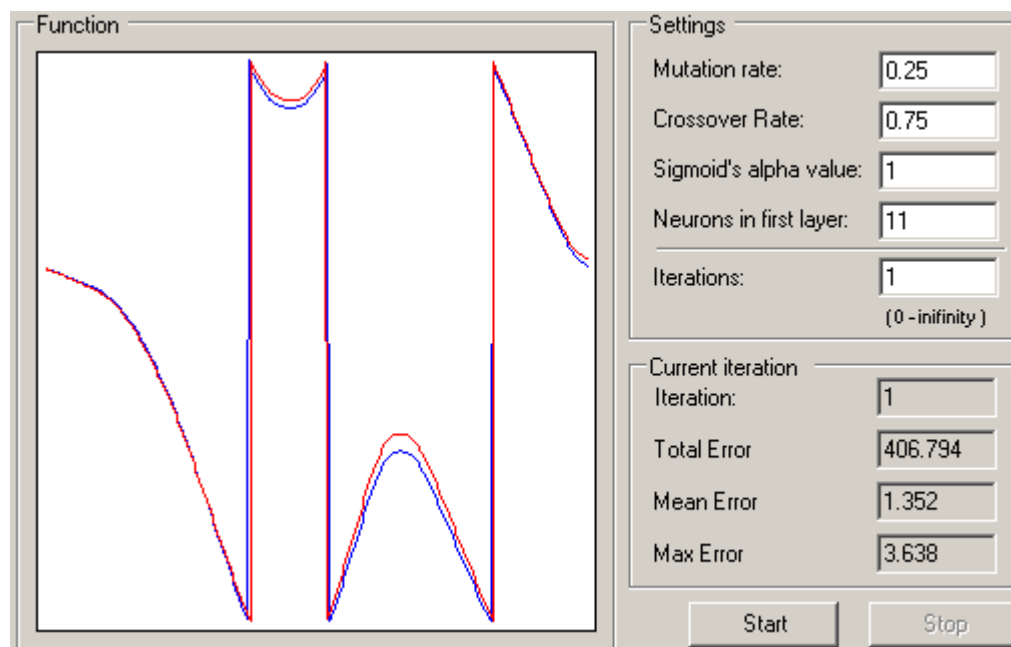


Figura 7.42 - Percurso 6 - Teste (Ângulo)

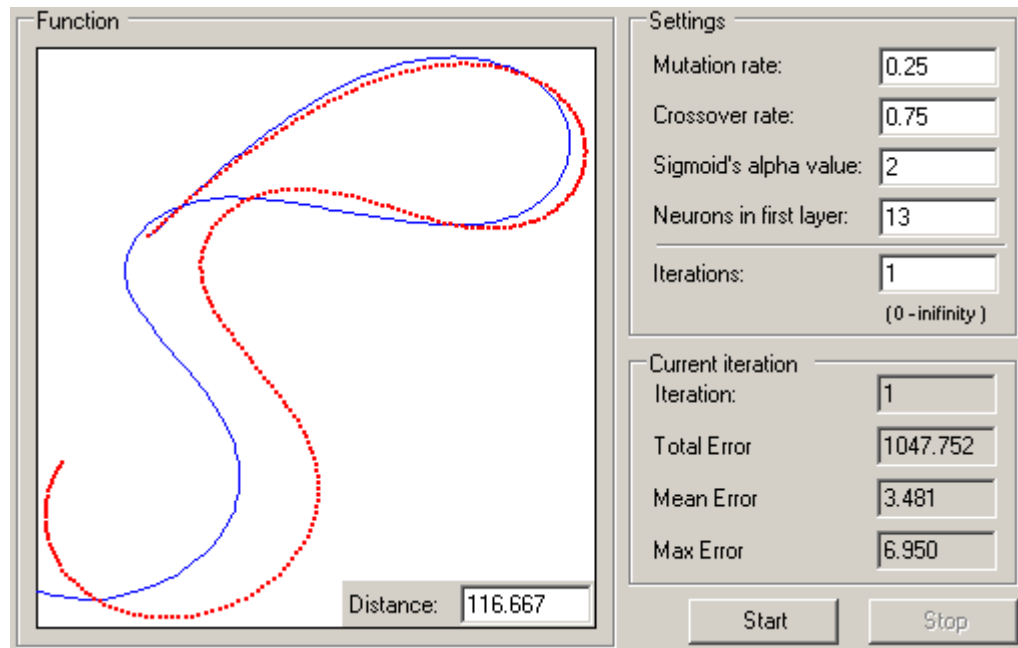


Figura 7.43 - Percurso 6 - Teste (Posição)

7.6.2 Teste da Rede Neuro-Especialista

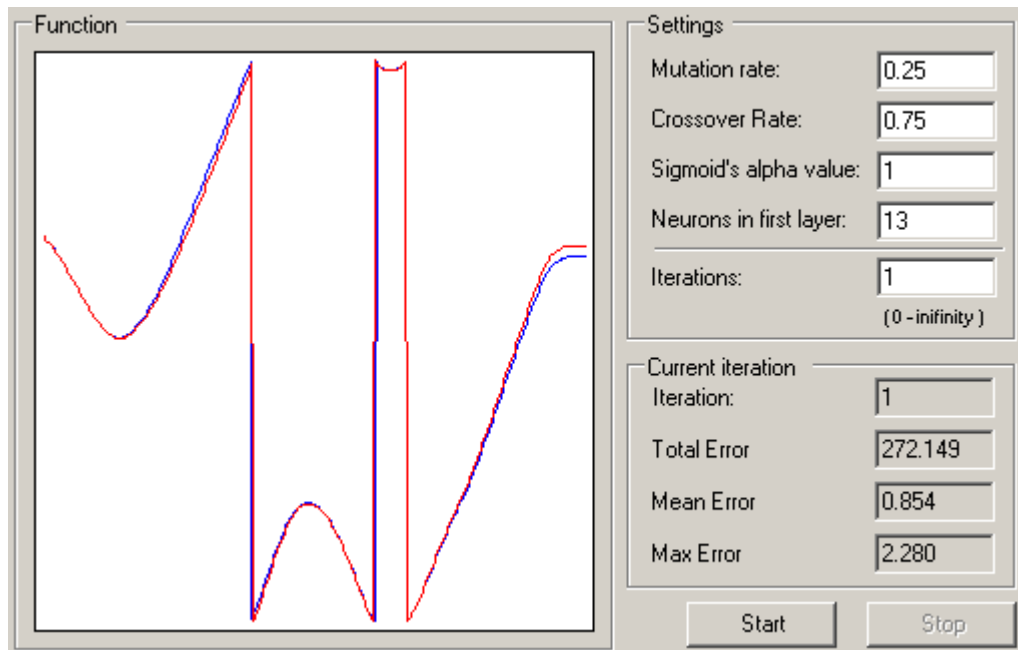


Figura 7.44 - Percurso 2 - Teste (Ângulo)



Figura 7.45 - Percurso 2 - Teste (Posição)

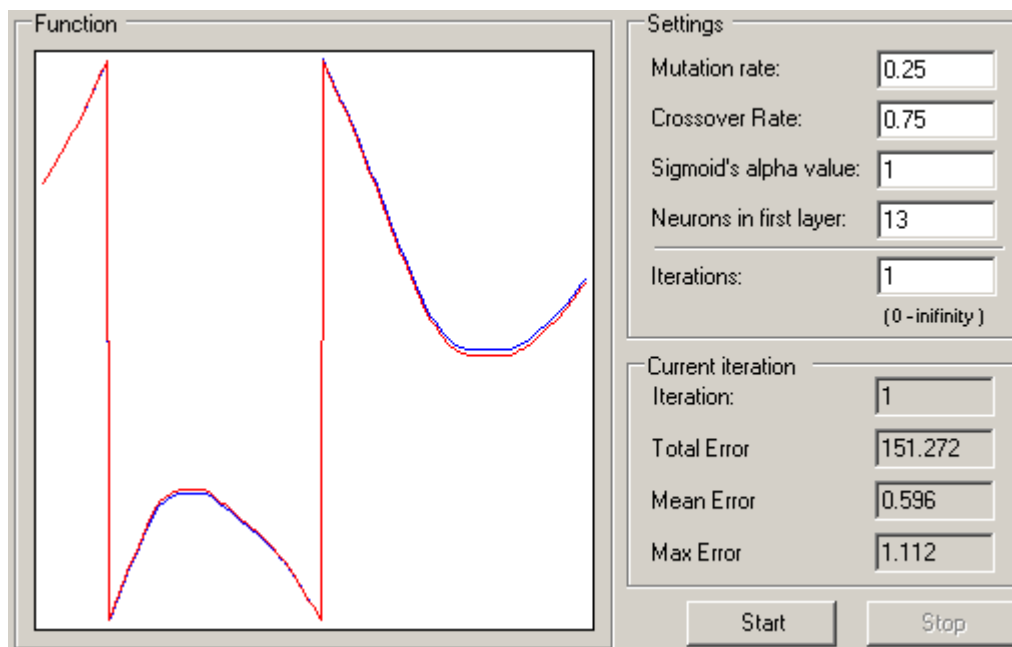


Figura 7.46 - Percurso 3 - Teste (Ângulo)

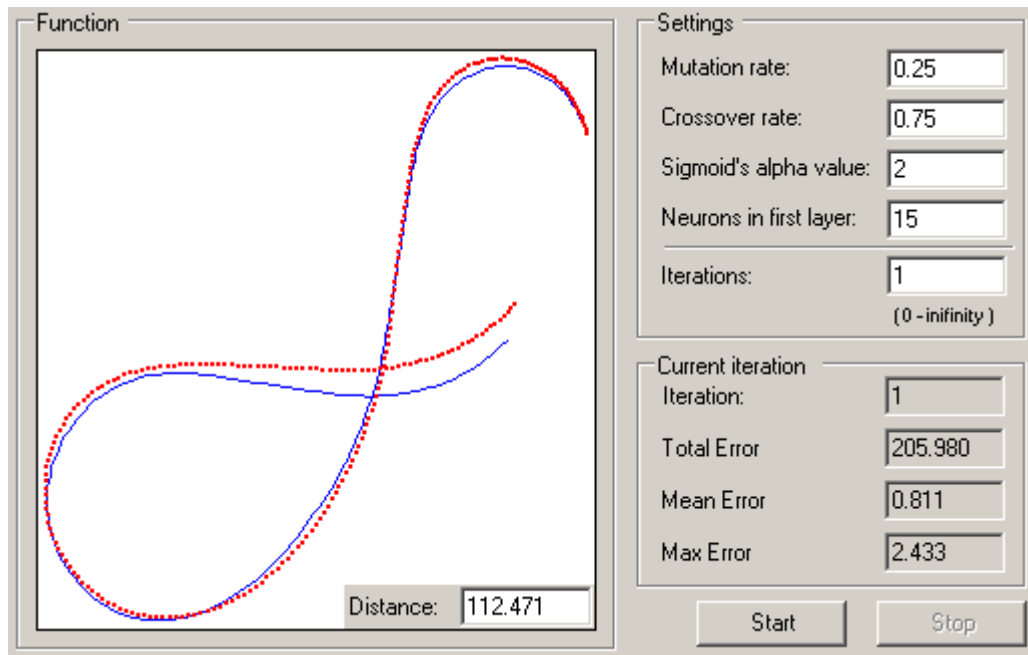


Figura 7.47 - Percurso 3 - Teste (Posição)

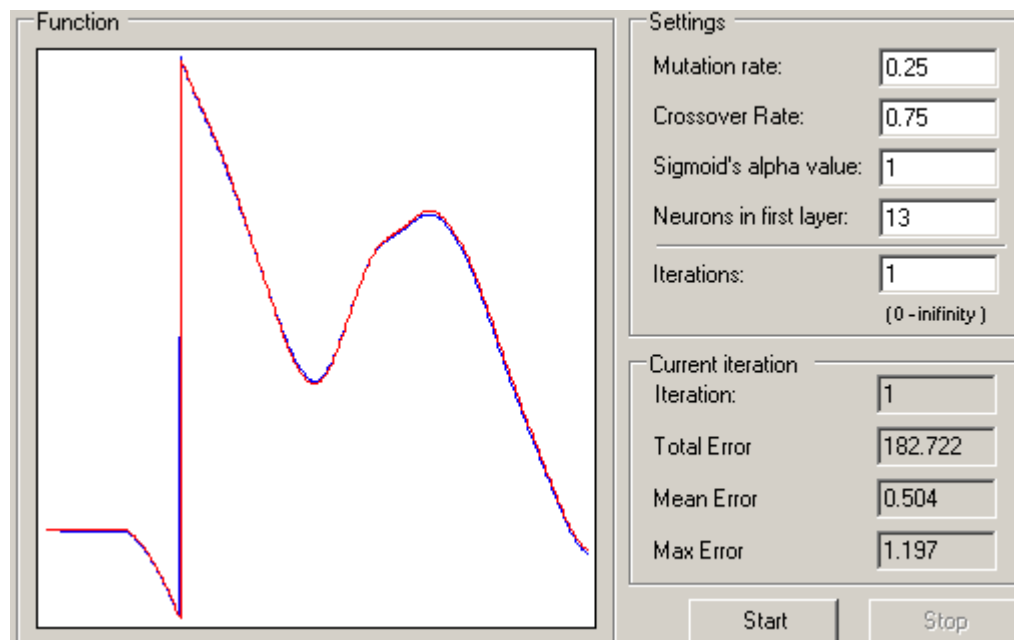


Figura 7.48 - Percurso 4 - Teste (Ângulo)



Figura 7.49 - Percurso 4 - Teste (Posição)

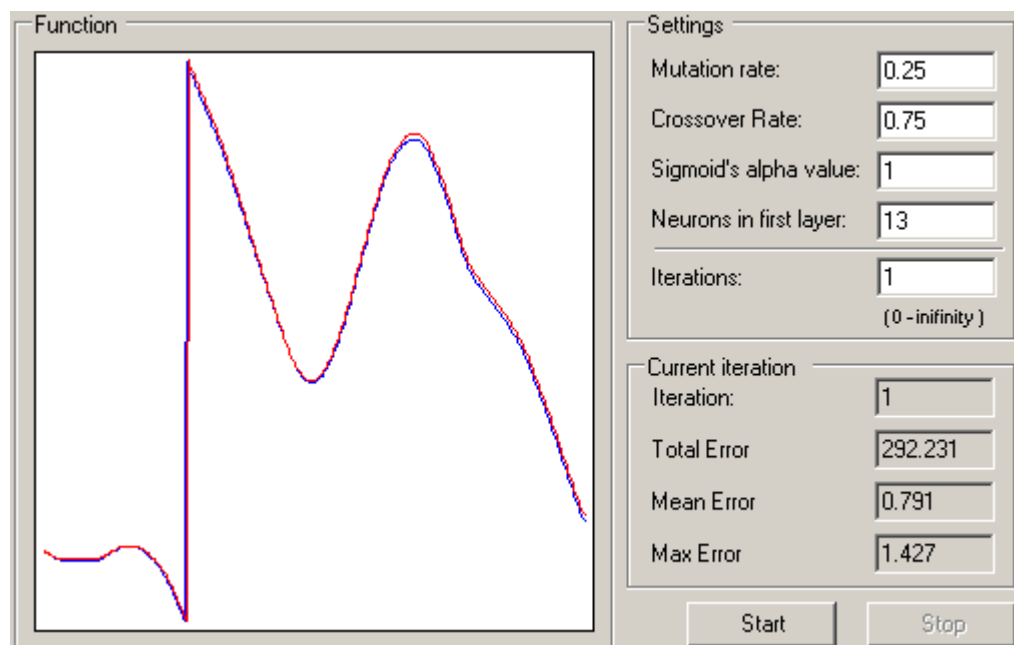


Figura 7.50 - Percurso 5 - Teste (Ângulo)

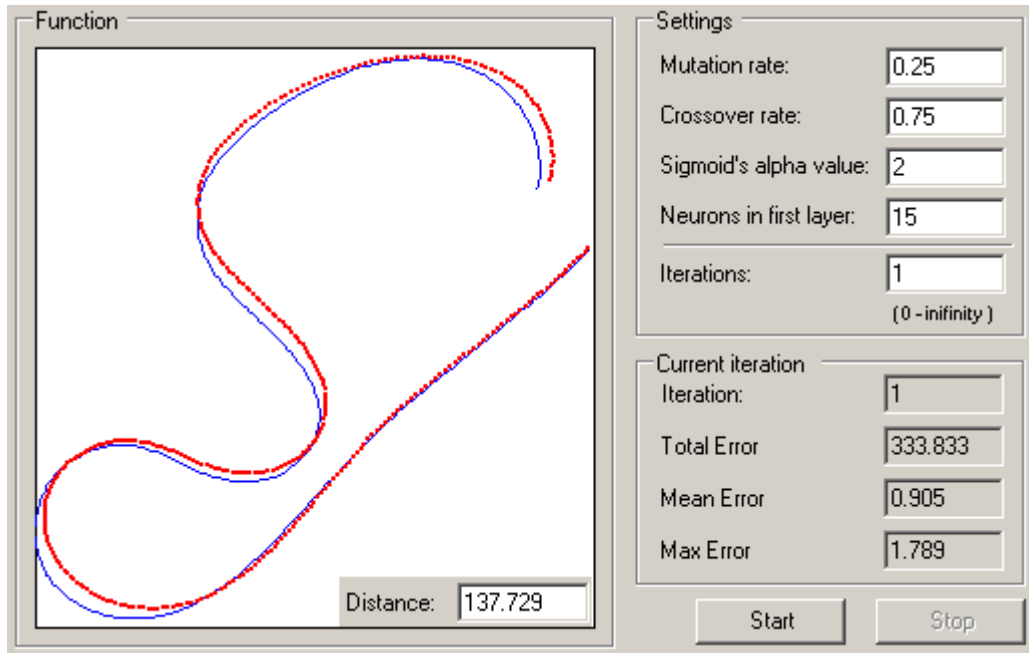


Figura 7.51 - Percurso 5 - Teste (Posição)

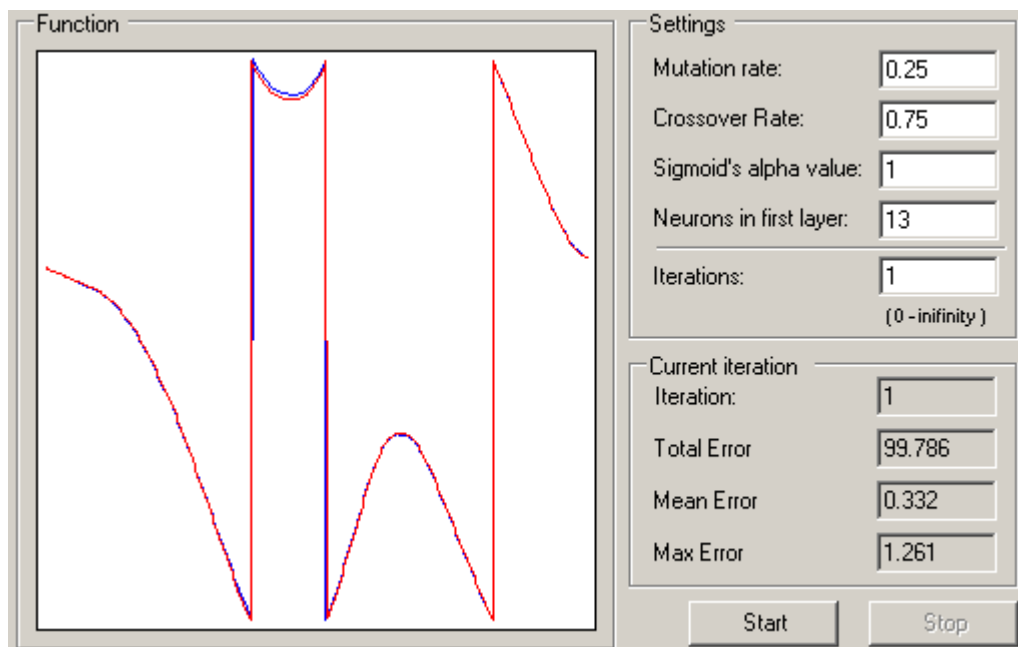


Figura 7.52 - Percurso 6 - Teste (Ângulo)

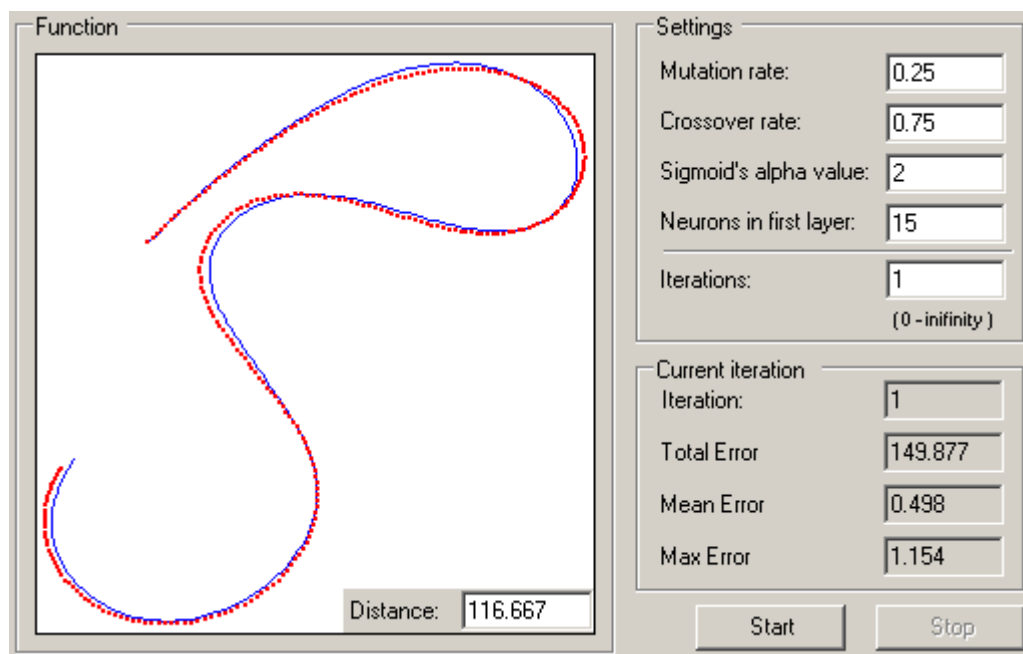


Figura 7.53 - Percurso 6 - Teste (Posição)

7.7 Validação da RNA

Para validar as redes neurais, três erros foram calculados (Erro Total, Erro Médio e Erro Máximo) para que os resultados das redes pudessem ser comparados com os resultados da cinemática. O Erro Total é o somatório da diferença entre o resultado pretendido e o resultado alcançado pela cinemática ou pela rede neural. O Erro Médio é o Erro Total dividido pelo número de dados coletados. E o Erro Máximo é a maior diferença entre o resultado pretendido e o resultado alcançado.

As Tabelas 7.3, 7.4 e 7.5 e as Tabelas 7.6, 7.7 e 7.8 mostram os resultados obtidos para a posição e o ângulo, respectivamente.

Analisando os resultados mostrados nas Tabelas pode-se validar a Rede Neuro-Especialista, pois esta apresenta para todos os erros calculados valores abaixo dos erros apresentados pela cinemática. Já a Rede Neural não apresentou bons resultados para a Posição, sendo o Percurso 5 o único percurso com resultados melhores que os da cinemática. E no caso da Rede Neural para o Ângulo apenas o Percurso 5 não apresentou um resultado melhor que os da cinemática. Com esses resultados, a Rede Neural para o Ângulo e a Rede Neural para a Posição devem ser reestruturadas.

	Erro Total (m)		
	Cinemática	Rede Neural	Rede Neuro-Especialista
Percurso1	1554.200	673.518	345.425
Percurso2	551.016	612.942	309.825
Percurso3	741.135	864.926	205.980
Percurso4	719.707	786.081	306.747
Percurso5	1109.600	464.941	333.833
Percurso6	732.027	1047.752	149.877

Tabela 7.2 - Comparação dos Resultados para a Posição (Erro Total)

	Erro Médio (m)		
	Cinemática	Rede Neural	Rede Neuro-Especialista
Percurso1	2.684	1.165	0.598
Percurso2	1.722	1.921	0.971
Percurso3	2.906	3.405	0.811
Percurso4	1.972	2.169	0.843
Percurso5	2.999	1.260	0.905
Percurso6	2.424	3.481	0.498

Tabela 7.3 - Comparação dos Resultados para a Posição (Erro Médio)

	Erro Máximo (m)		
	Cinemática	Rede Neural	Rede Neuro-Especialista
Percurso1	4.893	2.160	1.316
Percurso2	3.555	3.722	3.134
Percurso3	5.291	5.235	2.433
Percurso4	3.338	3.072	1.591
Percurso5	5.566	1.882	1.789
Percurso6	4.757	6.950	1.154

Tabela 7.4 - Comparação dos Resultados para a Posição (Erro Máximo)

	Erro Total (graus)		
	Cinemática	Rede Neural	Rede Neuro-Especialista
Percurso1	1253.746	1027.846	384.856
Percurso2	923.379	846.585	272.149
Percurso3	733.271	732.051	151.272
Percurso4	952.370	362.969	182.722
Percurso5	528.210	1221.632	292.231
Percurso6	644.807	406.794	99.786

Tabela 7.5 - Comparação dos Resultados para o Ângulo (Erro Total)

	Erro Médio (graus)		
	Cinemática	Rede Neural	Rede Neuro-Especialista
Percurso1	2.177	1.776	0.665
Percurso2	2.865	2.653	0.854
Percurso3	3.266	2.882	0.596
Percurso4	2.635	0.997	0.504
Percurso5	1.432	3.312	0.791
Percurso6	2.120	1.352	0.332

Tabela 7.6 - Comparação dos Resultados para o Ângulo (Erro Médio)

	Erro Máximo (graus)		
	Cinemática	Rede Neural	Rede Neuro-Especialista
Percurso1	5.959	4.796	1.736
Percurso2	9.798	7.460	2.280
Percurso3	6.761	6.744	1.112
Percurso4	4.297	2.217	1.197
Percurso5	3.438	5.913	1.427
Percurso6	5.157	3.638	1.261

Tabela 7.7 - Comparação dos Resultados para o Ângulo (Erro Máximo)

8 CONCLUSÃO

O trabalho apresentado mostrou a utilização de uma nova arquitetura de redes neurais, denominada Rede Neuro-Especialista. A ideia foi adicionar neurônios cuja função de ativação fosse a modelagem pré-existente do sistema. Desta forma, mesmo que impreciso, tal modelo aumenta a região de informação sobre o mundo, deixando para o resto da rede apenas ajustes de sintonia.

Para demonstrar esta proposta, um problema de navegação autônoma foi abordado onde a cinemática do veículo foi modelada na forma de neurônios especialistas. Feito isso, a rede foi treinada e os resultados foram comparados com a forma tradicional (apenas a cinemática), com uma rede neural artificial pura e com a proposta dessa dissertação. Os resultados mostram uma eficiência muito grande tanto em tempo de treinamento quando em qualidade de resposta.

O método para treinamento da rede foi genérico, logo como proposta de continuação deste trabalho será definido modelos específicos de treinamentos para este tipo de arquitetura, novas funções de avaliação e aplicações em outras áreas do conhecimento.

9 BIBLIOGRAFIA

- (1) BISHOP, R. Intelligent Vehicle Applications Worldwide. **IEEE Intelligent Systems and Their Applications**, v. 15, n. 1, p. 78-81, Jan-Feb 2000.
- (2) BISHOP, R. **Intelligent vehicle technology and trends**. [S.l.]: Artech House, 2005.
- (3) LI, W.; LEUNG, H. Simultaneous registration and fusion of Multiple Dissimilar Sensors for Cooperative Driving. **IEEE Transactions on Intelligent Transportation Systems**, v. 5, n. 2, p. 84-98, June 2004.
- (4) BENTO, L. et al. Sensor Fusion for Precise Autonomous Vehicle Navigation in Outdoor Semi-structured Enviroments. **8th International IEEE Conference On Intelligent Transportation Systems**, Vienna, September 2005. 245-250.
- (5) THRUN, S. et al. **Particle filters for mobile robot localization**. [S.l.]: Springer, 2001.
- (6) LI, Q.; ZHENG, N.; CHENG, H. Springrobot: A Prototype Autonomous Vehicle and Its Algorithms for Lane Detection. **IEEE Transactions on Intelligent Transportation Systems**, v. 5, n. 4, p. 300-308, December 2004.
- (7) MCCALL, J. C.; TRIVEDI, M. M. Video-Based Lane Estimation and Tracking for Driver Assistance: Survey, System, and Evaluation. **IEEE Transactions on Intelligent Transportation Systems**, v. 7, p. 20-37, 2006.
- (8) THORPE, C. et al. Vision and Navigation for the Carnegie-Mellon Navlab. **IEEE Trans. Patt. Analysis Mach. Intell**, v. 10, n. 3, p. 362-373, December 1988.
- (9) SEETHARAMAN, G.; LAKHOTIA, A.; BLASCH, E. P. Unmanned Vehicles Come of Age: The DARPA Grand Challenge. **IEEE Computer Society**, p. 26-29, December 2006.
- (10) MILLER, I.; GARCIA, E.; CAMPBELL, M. To Drive Is Human. **IEEE Computer Society**, p. 52-56, December 2006.

- (11) JUNG, C. R. et al. **Computação Embarcada: Projeto e Implementação de Veículos Autônomos Inteligentes. XXV Congresso da Sociedade Brasileira de Computação**, São Leopoldo, 22-29 Julho 2005. 1358-1406.
- (12) RAJAMANI, R. **Vehicle Dynamics and Control**. [S.l.]: Springer, 2006.
- (13) ABE, M. **Vehicle Handling Dynamics: Theory and Application**. [S.l.]: Elsevier, 2009.
- (14) GREWAL, M. S.; WEILD, L. R.; ANDREWS, A. P. **Global Positioning Systems, Inertial Navigation and Integration**. Second. ed. [S.l.]: Wiley, 2007.
- (15) KAPLAN, E. D.; HEGARTY, C. J. **Understanding GPS: Principles and Applications**. Second. ed. [S.l.]: Artech House, Inc, 2006.
- (16) **GPS Center**. Disponível em: <<http://www.gpscenter.com.br/index64.html>>. Acesso em: 29 Outubro 2010.
- (17) BORENSTEIN, J.; EVERETT, H. R.; FENG, L. **Where am I? Sensors and Methods for Mobile Robot Positioning**. [S.l.]: [s.n.], 1996.
- (18) WELCH, G.; BISHOP, G. **An Introduction to the Kalman Filter**. University of North Carolina. [S.l.], p. 16. 2006.
- (19) OGLE, T. L.; BLAIR, W. D. Fixed-Lag Alpha-Beta Filter for Target Trajectory Smoothing. **IEEE Transactions on Aerospace and Electronic Systems**, v. 40, n. 4, p. 1417 - 1421, October 2004.
- (20) HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. Second. ed. [S.l.]: Prentice Hall, 1999.
- (21) MEDSKER, L. R. **Hybrid Intelligent Systems**. [S.l.]: Kluwer Academic Publishers, 1995.
- (22) KROSE, B.; SMAGT, P. V. D. **An Introduction To Neural Networks**. Eighth. ed. [S.l.]: [s.n.], 1996.
- (23) BRAGA, A. P.; CARVALHO, A. F.; LUDERMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. [S.l.]: LTC, 2000.

- (24) MONTANA, D. J.; DAVIS, L. Training Feedforward Neural Networks Using Genetic Algorithms. **11th International Joint Conference on Artificial Intelligence**, Detroit, 1989. 762-767.
- (25) HEATON, J. **Introduction to Neural Networks for C#**. Second. ed. [S.l.]: Heaton Research, 2008.
- (26) MASTERS, T. **Practical Neural Network Recipes in C++**. [S.l.]: Academic Press, 1993.
- (27) CARVALHO, A. F. **Algoritmos Genéticos**. Disponível em:
<<http://www.icmc.usp.br/~andre/research/genetic/index.htm>>. Acesso em: 29 Outubro 2010.

APÊNDICE A ALGORITMO GENÉTICO

A.1 Conceito

Algoritmos Genéticos são algoritmos de otimização e busca baseados nos mecanismos de seleção natural e genética. Enquanto os métodos de otimização e busca convencionais trabalham geralmente de forma sequencial, avaliando a cada instante uma possível solução, os AGs trabalham com um conjunto de possíveis soluções simultaneamente.

Na terminologia de AGs (Algoritmos Genéticos), uma solução candidata é chamada de indivíduo ou cromossomo. Neste trabalho cada indivíduo é formado pelos pesos das ligações neurais. Ao conjunto de indivíduos simultaneamente avaliados é dado o nome de população. A cada indivíduo é associado um grau de adaptação ou aptidão, que mede a capacidade da solução, representada pelo indivíduo, para resolver um dado problema. AGs não garantem uma solução ótima, mas geralmente encontram soluções quase ótimas em um tempo aceitável. Esta técnica emprega uma estratégia de busca paralela e estruturada, embora com um forte componente aleatório, que é voltada em direção à busca de pontos de "maior aptidão", ou seja, pontos nos quais a função a ser minimizada (ou maximizada) tem valores relativamente baixos (ou altos) (23).

Apesar de aleatórios, eles não definem caminhos desordenados em direção a uma solução ótima, uma vez que utilizam informações históricas para definir novos pontos de busca, onde são esperados melhores desempenhos. Isto é realizado através de processos iterativos, em que cada iteração é chamada de geração. Durante cada geração, princípios de seleção e reprodução são aplicados a uma população de candidatos. A população pode variar de tamanho, dependendo da complexidade do problema e dos recursos computacionais disponíveis. Espera-se que ao final de um número de gerações o algoritmo venha a convergir para soluções ótimas ou quase ótimas.

O primeiro passo é a geração da população inicial. Em seguida deve ser determinado como os indivíduos serão codificados, permitindo a sua posterior manipulação pelos operadores genéticos. A este processo dá-se o nome codificação ou representação. Durante o

processo evolutivo, a população é avaliada: para cada indivíduo é dada uma nota, ou índice, refletindo sua habilidade de adaptação a um determinado ambiente.

Uma porcentagem dos mais aptos é mantida, enquanto os outros são descartados (darwinismo). Os membros mantidos pela seleção podem sofrer modificações em suas características fundamentais através de mecanismos como mutações e *crossovers*, gerando descendentes para a próxima geração. Este processo, chamado de reprodução, é repetido até que uma solução satisfatória seja encontrada.

A.1.1 Seleção

O processo de seleção determina quais indivíduos da população podem participar da fase de reprodução. Os indivíduos são selecionados de acordo com uma probabilidade dada pelos seus índices ou notas de aptidão. Assim, os indivíduos com maior adaptação relativa têm maiores chances de participar da reprodução. O princípio básico do funcionamento dos Algoritmos Genéticos é que um critério de seleção faz com que, após várias gerações, o conjunto inicial de indivíduos produza indivíduos mais aptos (23).

A.1.2 Reprodução

Os indivíduos escolhidos na fase de seleção participam da fase de reprodução, em que podem ser combinados ou modificados, produzindo os indivíduos da próxima geração. Estas combinações e modificações são realizadas por um conjunto de operadores chamados operadores genéticos. Os operadores genéticos são utilizados para transformar uma população através de sucessivas gerações, estendendo a busca até a obtenção de um resultado satisfatório (23).

Os principais operadores de reprodução são: *crossover* (ou cruzamento) e mutação. Eles são utilizados para assegurar que a nova geração apresente novos indivíduos (soluções), mantendo características desejáveis adquiridas em gerações anteriores.

A.1.2.1 *Crossover*

O *crossover* é o operador responsável pela recombinação de características genéticas dos pais durante a reprodução, permitindo que elas sejam herdadas pelas próximas gerações.

Ele é considerado o operador genético predominante, por isso é aplicado com uma probabilidade, chamada taxa de *crossover*, maior que a taxa de mutação (23).

A.1.2.2 Mutação

O operador de mutação é responsável pela introdução e manutenção da diversidade genética na população, alterando arbitrariamente um ou mais genes de um cromossomo escolhido aleatoriamente.

A mutação assegura que a probabilidade de se chegar a qualquer ponto do espaço de busca nunca será zero, além de contornar o problema de mínimos locais, por permitir a alteração da direção da busca. O operador de mutação é aplicado aos indivíduos com uma probabilidade dada pela taxa de mutação, que é geralmente pequena (23).

A.2 Definição dos Parâmetros do Algoritmo Genético

Determinar a melhor configuração para os parâmetros dos operadores genéticos é de suma importância para o desempenho do algoritmo genético, uma vez que este é fortemente dependente dos operadores do problema em questão.

Os parâmetros do algoritmo genético são os responsáveis pelo controle do processo evolucionário e são descritos a seguir (27):

População - refere-se ao número de indivíduos de cada população, ou seja, o número de pontos do espaço de busca. O tamanho da população afeta o desempenho global e a eficiência do AG. Com uma população pequena o desempenho pode cair, pois deste modo a população fornece uma pequena cobertura do espaço de busca do problema. Uma grande população geralmente fornece uma cobertura representativa do domínio do problema, além de prevenir convergências prematuras para soluções locais ao invés de globais. No entanto, para se trabalhar com grandes populações, são necessários maiores recursos computacionais, ou que o algoritmo trabalhe por um período de tempo muito maior.

Gerações - Total de ciclos de evolução de um algoritmo genético. Neste trabalho, o número de gerações corresponde ao número de iterações da rede neural.

Taxa de Crossover - Probabilidade de recombinação de material genético entre indivíduos. Quanto maior for esta taxa, mais rapidamente novas estruturas serão introduzidas na população. Mas se esta for muito alta, estruturas com boas aptidões poderão ser retiradas mais rapidamente e a maior parte da população será substituída. Com um valor baixo, o algoritmo pode tornar-se muito lento.

Taxa de Mutação - probabilidade de cada gene do cromossomo sofrer modificação. Uma baixa taxa de mutação previne que uma dada posição fique estagnada em um valor, além de possibilitar que se chegue a qualquer ponto do espaço de busca. Com uma taxa muito alta a busca se torna essencialmente aleatória.

Taxa de Seleção - Controla a porcentagem da população que será substituída durante a próxima geração. Com um valor alto, a maior parte da população será substituída, mas com valores muito altos pode ocorrer perda de estruturas de alta aptidão. Com um valor baixo, o algoritmo pode tornar-se muito lento.

A configuração dos parâmetros do algoritmo genético para o treinamento das RNAs desenvolvidas neste trabalho foi determinada empiricamente. Os valores adotados são mostrados na Tabela A.1.

Parâmetro Genético	Valor Escolhido
População	50
Taxa de Cruzamento	0.75
Taxa de Mutação	0.25
Taxa de Seleção	0.2

Tabela A.1 - Parâmetros do Algoritmo Genético