

2 Fundamentos Teóricos

A presente pesquisa tem um caráter claramente multidisciplinar. Por isso foi introduzido neste documento o presente capítulo, cujo propósito é apresentar os fundamentos básicos teóricos e algumas técnicas das áreas relacionadas que são essenciais para a compreensão dos modelos e procedimentos apresentados nesse trabalho.

Leitores já iniciados nos temas aqui abordados poderão prescindir da leitura da seção correspondente ao capítulo sem prejuízo para a compreensão do restante deste trabalho.

Na primeira parte do capítulo é apresentada uma breve introdução sobre sensoriamento remoto e técnicas de processamento de imagens de sensores remotos, principalmente em alta resolução. Detalhes de correção geométrica e topográfica são apresentados.

Na segunda parte são apresentados algoritmos específicos de classificação supervisionada utilizados durante o trabalho. Também são descritos modelos de classificação baseados em lógica nebulosa e sistemas neuro-fuzzy e algoritmos genéticos utilizados para otimização de parâmetros.

O algoritmo de segmentação ou extração dos objetos a partir de imagens de sensores remotos é descrito na terceira parte do capítulo.

Na quarta e última parte apresenta-se uma breve descrição dos parâmetros de Haralick [37], utilizados como atributos de textura ao longo desse trabalho.

2.1 Sensoriamento remoto

Sensoriamento Remoto é a técnica ou arte de obter informações sobre objetos, áreas e fenômenos através da análise de informações obtidas sem contato direto com o objeto, área ou fenômeno de estudo [1]. Nesse contexto a análise de uma foto de uma determinada área geográfica pode ser denominada de sensoriamento remoto.

A principal ferramenta do sensoriamento remoto é o sensor de captação da imagem a ser analisada. Atualmente, existe uma grande variedade de sensores ativos e passivos para a captação de imagens [13]. Sensores ativos emitem energia em um determinado comprimento de onda e a captam de volta após ser refletida pelo material analisado. Através de medidas realizadas na emissão e captação é possível a estimação e criação de imagens remotas. Sensores passivos não emitem radiação, mas simplesmente captam a reflexão da radiação, principalmente a radiação eletromagnética solar, sobre os objetos. Esse trabalho concentrara-se em sensores passivos, sensíveis a radiações eletromagnéticas.

Esses sensores baseiam-se na propagação da energia eletromagnética e na sua interação com alvos naturais. Em poucas palavras a energia eletromagnética emitida pelo sol atinge alvos naturais e é refletida por esses, sendo então capturada por sensores remotos transportados por satélites, aeronaves, balões meteorológicos, etc. Essa reflexão eletromagnética sofre alguns fenômenos devidos à interação com a cobertura do solo ou à dispersão ou absorção pela atmosfera. As influências da atmosfera não são tratadas nesse trabalho e podem ser melhor estudadas em [1].

Ao atingir alguma região da superfície da Terra a radiação eletromagnética sofre interações de acordo com as características físico-químicas da cobertura do solo ou da composição do objeto; como resultado uma parte da energia eletromagnética é refletida e uma outra parte é absorvida pela superfície. Em termos gerais, tipos diferentes de cobertura de solo oferecem comportamentos diferentes, conforme o comprimento de onda por ele atingido pelo alvo, como o apresentado na Figura 1.

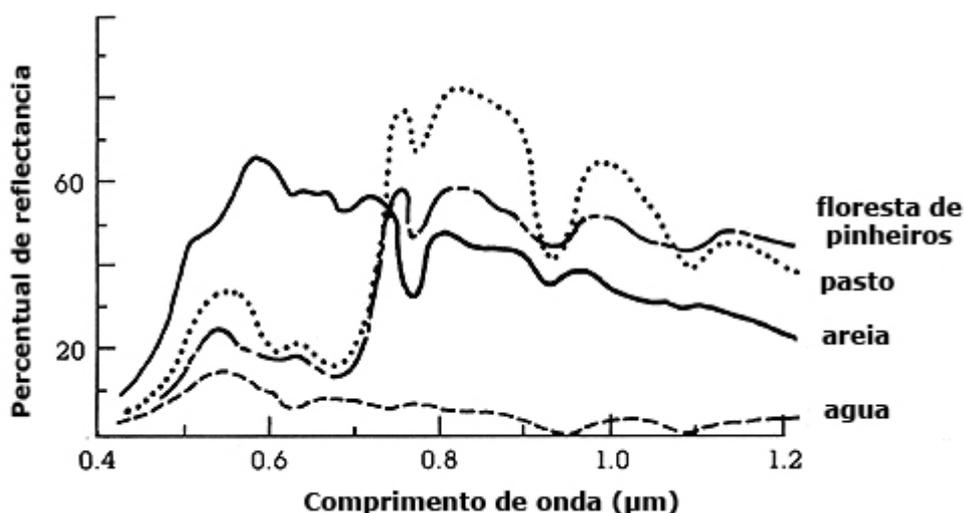


Figura 1 – Exemplo da influência da refletância e do comprimento de onda em diferentes tipos de cobertura de solo [14]

Os sensores remotos passivos que detectam energia eletromagnética refletida ou emitida pelo alvo permitem medir a proporção de energia refletida da superfície terrestre que os atinge, num determinado comprimento de onda.

2.1.1 Sensoriamento remoto em imagens de alta resolução

Os estudos em sensoriamento remoto em alta resolução espacial começaram a ganhar força quando câmeras a bordo de aviões foram utilizadas para a obtenção fotografias aéreas. Durante anos as tecnologias tanto de obtenção quanto de processamento e análise dessas imagens se desenvolveram intensamente, inicialmente nos setores militares e em seguida na administração pública e em setores privados [15]. Os mapas produzidos por fotos aéreas podem ser obtidos em diferentes escalas, dependendo basicamente do tipo de câmera utilizada e da altitude do sobrevôo.

O grande problema detectado com as imagens aéreas são os seus elevados custos de obtenção e a complexidade que cresce com a área a ser analisada. A análise de uma área relativamente grande com fotos aéreas é um processo caro e trabalhoso principalmente pelo complexo processo de obtenção que envolve um plano de vôo preciso e o controle necessário para garantir a homogeneidade na aquisição das imagens.

Com o avanço das tecnologias espaciais especialmente a partir de 1972, e o desenvolvimento de câmeras e sensores digitais, foi possível a aquisição de imagens a partir de satélites não militares. O programa LANDSAT surgiu ainda em 1972 com o objetivo de produzir imagens de toda a superfície do globo terrestre de maneira contínua e sistemática, e a custos razoáveis.

Desde o início da utilização não-militar de sensores remotos baseados em satélites, um grande número de diferentes sensores foram desenvolvidos. Estes sensores passam a forçar os desenvolvimentos científicos, e tecnológicos, tanto no que se refere a resolução espacial quando a resolução radiométrica. As imagens geradas estão sendo utilizadas amplamente, principalmente em tarefas de planejamento. A maior periodicidade na obtenção das imagens tornou possível uma análise de séries históricas em determinadas regiões que mostram claramente tendências de alteração de padrões de ocupação de solo. Com esses dados é possível um melhor planejamento a longo e médio

prazo de atividades que impactam direta ou indiretamente no padrão de cobertura do solo.

Os sensores remotos baseados em satélites tornaram possível uma utilização mais ampla e democrática da informação e do sensoriamento remoto como um todo. Porém, o tipo de informação extraída depende fundamentalmente do tipo de imagem obtida. Em outras palavras, uma imagem LANDSAT consegue representar objetos não menores do que 900 m^2 , assim como uma imagem de um satélite de metrologia como o AVHRR consegue representar informações a partir de 1 Km^2 .

Em meados de 1999 ocorreu um marco na área de sensoriamento remoto para aplicações não militares. O IKONOS, o primeiro satélite de alta resolução destinado exclusivamente para sensoriamento remoto e voltado a objetivos civis foi lançado. Com a resolução espacial de 1m no modo pancromático, as imagens IKONOS trouxeram para o mundo do sensoriamento remoto uma nova gama de possibilidades. Pela primeira vez era possível, de forma sistemática, em alguns casos à custos inferiores aos das fotos aéreas, a identificação e representação de objetos com até 1 m^2 .

O satélite IKONOS realiza a órbita em torno da Terra em aproximadamente 98 minutos em uma altitude de aproximadamente 680 km [16]. Lançado em uma órbita em sincronia com o sol o IKONOS consegue passar por uma dada longitude em praticamente o mesmo horário local (10h30min AM) diariamente. Com isso o IKONOS é capaz de produzir imagens de 1m de resolução de uma determinada área geográfica a cada 3 dias.

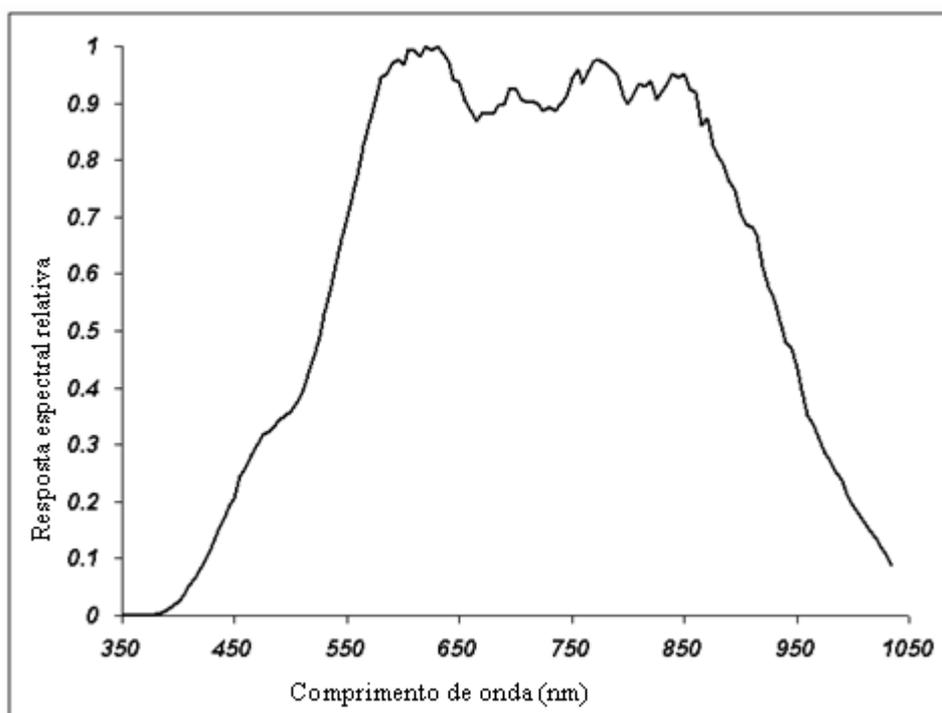


Figura 2 – resposta espectral relativa ao comprimento de onda na banda pancromática IKONOS [16]

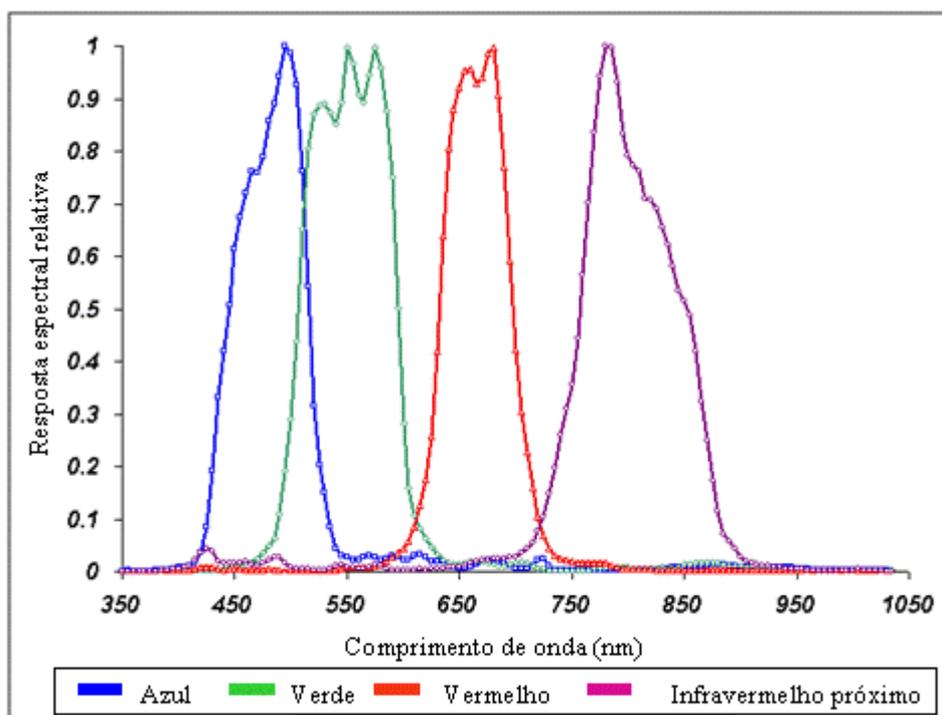


Figura 3 – resposta espectral relativa, conforme o comprimento de ondas para o sensor multiespectral IKONOS [16]

O satélite IKONOS produz dois tipos de imagens bem distintas em termos de resolução espectral e espacial. Com a resolução de apenas 1m por pixel a imagem

pancromática possui uma faixa de sensibilidade eletromagnética que cobre todo o espectro visível (Figura 2). A imagem multiespectral IKONOS possui uma resolução um pouco menor, 4 m por pixel, porém essa imagem possui quatro bandas de detecção oferecendo uma faixa dinâmica mais ampla em cada uma dessas bandas (Figura 3).

Todas as imagens IKONOS são geradas com uma resolução radiométrica de 11-bits e 8-bits. A resolução radiométrica de 11-bits permite uma faixa dinâmica de níveis muito maior, podendo chegar a 2048 níveis em cada banda. A resolução de 8-bits é fornecida por compatibilidade com softwares existentes. O processamento com 11bits fornece uma maior riqueza de detalhes já que são 1048 níveis de cinza e não 256 como no caso das imagens LANDSAT.

Assim como o IKONOS, outros satélites de alta resolução já foram e outros serão ainda lançados voltados a propósitos não militares. O QUIKCKBIRD, por exemplo, já operacional e trabalhando em uma resolução espacial de até 60cm e o SPOT 5 com uma resolução de até 2.5m já estão operacionais comercialmente.

O governo Brasileiro mostra o seu envolvimento com sensoriamento remoto em imagens de alta resolução através de um programa em conjunto com o Governo Chinês chamado CBERS (*China-Brazil Earth Resources Satellite*). Através desse programa satélites com câmeras de médias resoluções já se encontram operacionais. Esse mesmo programa promete colocar em órbita os satélites CBERS-3 e CBERS-4 em 2008 e 2010 respectivamente. Esses satélites devem operar com câmeras de média resolução capazes de produzir imagens com até 5m de resolução espacial na banda pancromática, e ainda com uma taxa de revista (período no qual o satélite passa no mesmo local) de aproximadamente cinco dias.

2.1.2 Pré-processamento

Para que várias imagens de sensores remotos sejam integradas entre si e com outros dados geográficos em um ambiente SIG (Sistema Geográfico de Informação), é necessário um processo prévio de correção.

O processo de georeferenciação é o responsável por converter as coordenadas das imagens de sensores remotos para uma base geográfica real. Esse processo associa à posição de cada pixel da imagem a coordenada geográfica do ponto imageado. Isso permite, por exemplo, a integração da informação contida na imagem com dados de SIG

disponíveis em todo o mundo. O georeferenciamento parte normalmente de pontos na imagem, cuja coordenada geográfica é bem conhecida, os chamados pontos de controle ou GCPs (“*ground control points*”). Os GCPs podem ser obtidos através de coleta de campo por GPS (*Global Position System* ou Sistema de Posicionamento Global) ou até mesmo a partir de algum dado já georeferenciado. Em outras palavras, os GCPs são um conjunto de pontos, facilmente identificáveis, em que se pode definir o posicionamento real ou geográfico de alguma maneira específica. A partir desses pontos procedimentos matemáticos são aplicados com o objetivo de realizar uma reamostragem na imagem como um todo segundo um sistema de coordenadas geográficas. Como resultado do georeferenciamento obtém-se uma nova imagem, denominada de “imagem retificada”, ou “georeferenciada”.

Um outro problema com imagens de satélite diz respeito à altimetria. Em regiões muito montanhosas ou em uma escala muito fina o tamanho de um pixel pode não estar compatível com a realidade apresentada em campo.

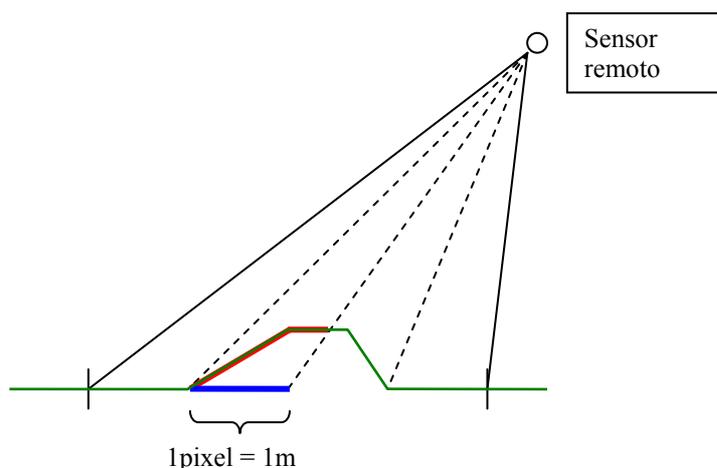


Figura 4 – Representação do problema da topografia em captura de imagens de sensores remotos

A Figura 4 ilustra o problema decorrente da elevação do terreno relacionado na captura de imagens de sensores remotos. Um pixel de 1m (área em azul) representa na realidade ou geograficamente falando uma área maior (em vermelho).

A solução para o problema da correção altimétrica está na georetificação em conjunto com a correção altimétrica com base no modelo digital de terreno, ou DTM (*Digital Terrain Model*). Esse processo é comumente chamado de “ortoretificação”. O DTM é uma representação digital da superfície com informações altimétricas, que pode

ser obtido a partir de isolinhas, ou seja, das curvas de nível da região estudada. Por ser digital possui uma resolução espacial e radiométrica assim como as imagens. A resolução, principalmente a espacial, do DTM utilizado afeta a qualidade da ortorectificação.

Além de alterar propriedades geométricas nas imagens capturadas, a topografia traz consigo outros efeitos, como a iluminação e a reflectância dos objetos. Técnicas de correção topográficas também são utilizadas para amenizar esses problemas. Essas técnicas levam em conta, além do DTM, vários outros fatores, tais como a posição do sol no ato da captura da imagem e a altitude do sensor.

Uma descrição detalhada dos procedimentos específicos, para a ortorretificação e para a correção topográfica, pode ser encontrada em [15, 17, 18 e 19].

2.2 Métodos computacionais

Esta seção aborda os fundamentos e as técnicas de processamento de imagem, inteligência computacional e reconhecimento de padrões necessários para a compreensão do restante do trabalho.

2.2.1 Classificação de padrões

O processo de classificação de padrões é o procedimento computacional capaz de atribuir a diferentes grupos ou classes alguns determinados padrões apresentados. Por exemplo, um procedimento que separe laranjas podres de laranjas sadias pode ser considerado um procedimento de classificação de padrões onde os padrões são os objetos analisados (laranjas). Os inúmeros fatores que tipicamente se usam para diferenciar uma laranja podre de uma sadia, por exemplo, o cheiro, a cor, a textura dentre outros são conhecidos como atributos dos padrões analisados.

Os procedimentos de classificação podem ser divididos em dois grandes conjuntos. Primeiro pode-se definir um grupo de procedimentos que não necessita de dados de treinamento, ou seja, não é necessária a apresentação das classes dos de uma parte do conjunto de padrões (conjunto de treinamento) de cada uma das classes para

que as mesmas sejam reconhecidas. Esses procedimentos normalmente trabalham dividindo o conjunto de padrões segundo os atributos apresentados ao sistema e são denominados “métodos de classificação não supervisionados”.

Um outro conjunto de métodos e algoritmos trabalha de maneira supervisionada, ou seja, existe uma fase de ajuste de parâmetros dos procedimentos denominada de fase de treinamento onde as classes de cada padrão são apresentadas ao modelo. Assim como o ser humano, que tem que aprender alguma tarefa, esses métodos utilizam-se da fase de treinamento para “aprender” a como diferenciar as classes apresentadas ao sistema com base no conjunto de atributos disponíveis. Esses métodos são mais robustos em termos de desempenho de classificação e em tarefas específicas de classificação de imagens sensores remotos tem mostrado melhor desempenho segundo a literatura exposta no Capítulo 2.

Nesse trabalho foi abordado um modelo de classificação supervisionada baseado em Redes Neurais Artificiais.

Em um segundo momento é abordado métodos supervisionados baseados em lógica nebulosa (ou lógica nebulosa) nas tarefas de classificações de padrões. Esse tipo de procedimento depende da elaboração de regras de inferência por um especialista. Essas regras interagem sobre os atributos dos padrões analisados inferindo a classe de cada padrão.

No final desse capítulo serão apresentadas resumidamente as teorias sobre sistemas *neuro-fuzzy* e algoritmos genéticos. Esses métodos serão utilizados mais a frente no modelo proposto, neste trabalho.

2.2.1.1 Redes neurais artificiais

O cérebro humano pode ser visto como um processador de informações altamente complexo, não-linear e paralelo [20]. Por causa dessa capacidade de aprendizado e generalização o funcionamento dos neurônios tem sido estudado e com o objetivo de reproduzi-los em ambientes computacionais.

O neurônio artificial é uma unidade de processamento de informação fundamental para a operação das chamadas “Redes Neurais Artificiais” [20].

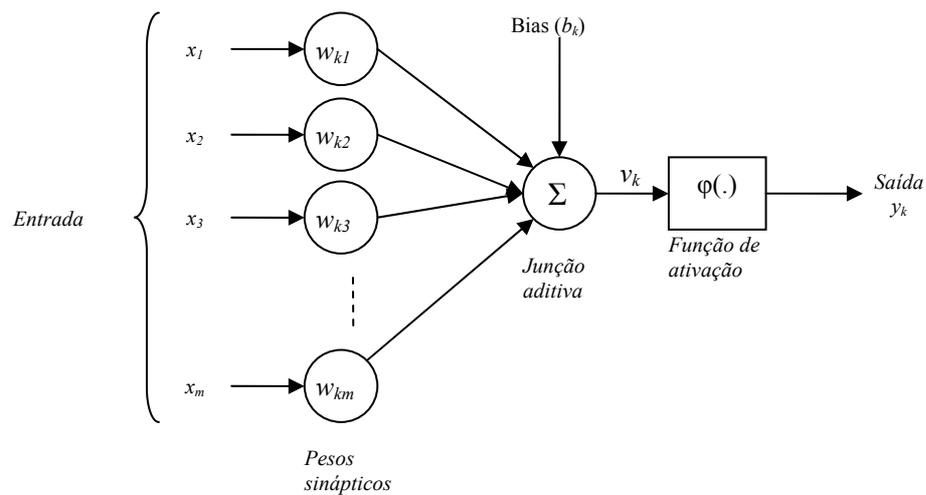


Figura 5 – Modelo de neurônio artificial não linear [20]

O modelo de um neurônio artificial é apresentado na Figura 5. O neurônio artificial consiste de m entradas ponderadas pelos pesos $w_{k1..m}$ e então somadas numa junção aditiva. Esses pesos são conhecidos como pesos sinápticos e são responsáveis pela capacidade de modelar e aprender de uma rede neural. Uma função de ativação é colocada em seguida ao somador, essa função é dita como restritiva segundo [20], o que em linhas gerais significa que ela limita o intervalo permissivo de saída do neurônio a que se referencia. O modelo prevê ainda um ajuste denominado *bias* que tem o efeito de aumentar ou diminuir o valor aplicado à função de ativação [20].

Em termos matemáticos pode-se definir a operação no neurônio k da Figura 5, que possui as entradas j ($j=1,2,\dots,m$) através das seguintes equações:

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

$$v_k = u_k + b_k \quad \text{Eq.(2)}$$

$$y_k = \varphi(v_k) \quad \text{Eq.(3)}$$

A função de ativação $\varphi(.)$ é responsável pela definição da saída de um neurônio em termos de uma entrada v . Essas funções podem ser representadas de inúmeras maneiras. Mais comumente, e nesse trabalho, utilizam-se funções sigmóides (Figura 6).

Basicamente o tipo de funções de ativação limita o comportamento no intervalo de saída. A função tangente hiperbólica sigmóide (*tansig*) por exemplo pode ter uma saída variando entre -1 e 1, enquanto a sigmóide logarítmica (*logsig*) tem uma saída sempre entre 0 e 1.

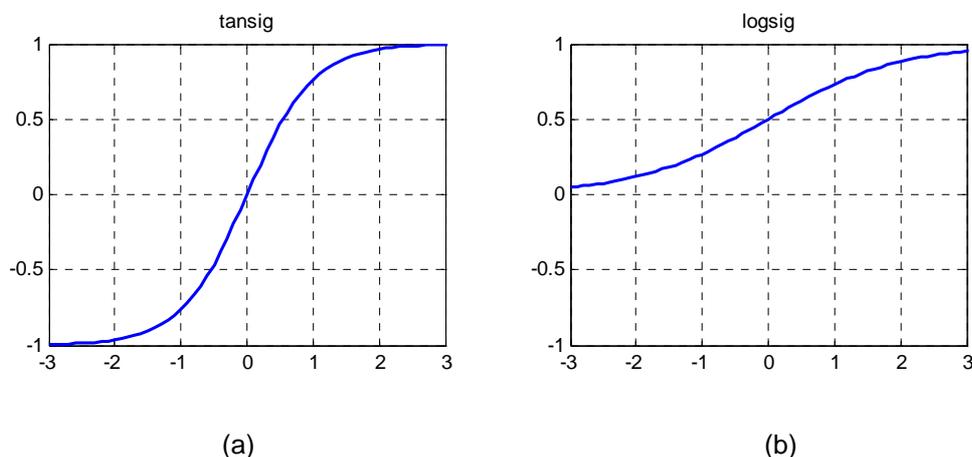


Figura 6 – Funções de ativação: (a) tangente hiperbólica sigmóide (*tansig*) e (b) sigmóide logarítmica (*logsig*) a direita.

Uma vez definida a característica de um neurônio pode-se definir a arquitetura de uma rede neural ligando neurônios entre si através dos pesos sinápticos. A configuração apresentada neste trabalho é a de uma rede neural de múltiplas camadas conhecida como “Perceptrons de múltiplas camadas” ou MLP (*multilayer perceptrons*). Essa configuração é ilustrada na Figura 7.

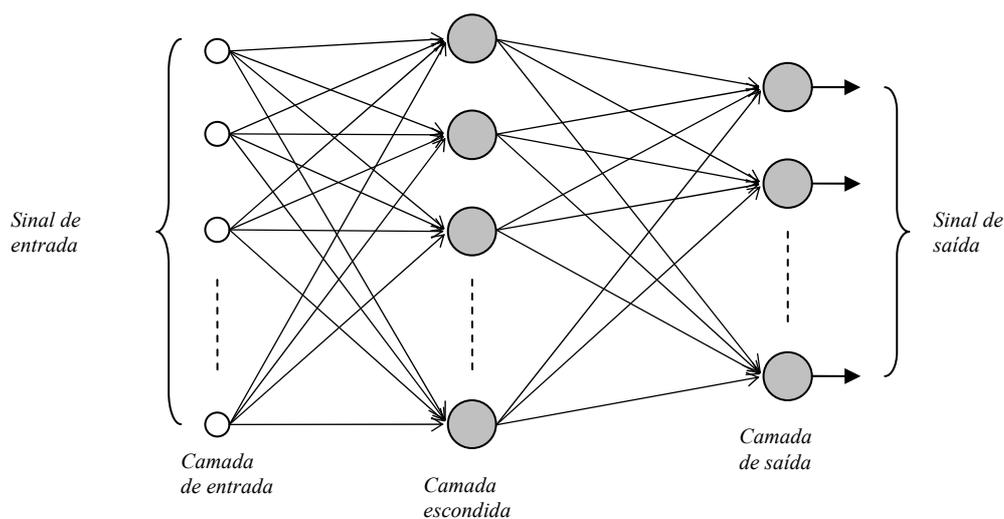


Figura 7 – Arquitetura de uma rede neural do tipo MLP com uma camada oculta

As redes MLPs consistem de um conjunto de unidades sensoriais, ou neurônios, arrumados em camadas, conforme o apresentado na Figura 7. Tem-se uma camada de entrada com nós não computacionais, uma ou mais camadas ocultas com nós computacionais e ainda uma camada de saída com nós também computacionais. Nesse trabalho exploram-se apenas redes neurais do tipo MLP com uma camada oculta. Os nós chamados computacionais realizam algum tipo de processamento, conforme o modelo do neurônio apresentado anteriormente. Os nós não computacionais são apenas entradas ou saídas de dados e não realizam qualquer processamento.

Cada arco que liga um neurônio a outro em uma rede MLP possui um peso sináptico associado denotado por w_{ji} . Esse peso pode ser entendido como um peso que associa à entrada do neurônio j a saída do neurônio i . Em uma rede neural do tipo MLP existe um grande número de pesos a serem definidos ou estimados. O algoritmo de retropropagação, ou BP (*backpropagation*), é um processo supervisionado de aprendizado baseado no gradiente descendente que estima os valores dos pesos sinápticos a partir de padrões de treinamento. Nesse algoritmo o erro de previsão se propaga da camada de saída para a camada oculta, procurando assim o melhor conjunto de valores para os pesos de toda a rede. Uma descrição detalhada do algoritmo de retropropagação é relativamente complexa e não foge ao escopo desse trabalho. Esse procedimento pode ser encontrado com mais detalhes de implementação em [20].

Um outro ponto interessante utilizado nesse trabalho sobre as redes neurais diz respeito à relação entre saída da rede neural e a probabilidade a posteriori. Segundo Haykin em [20] um classificador MLP aproxima de fato as probabilidades das classes a posteriori, desde que o tamanho do conjunto de treinamento seja suficientemente grande, garantindo que o processo de aprendizado por retropropagação não caia em um mínimo local.

2.2.1.2 Lógica nebulosa

A lógica nebulosa permite a expressão de forma clara e simples da incerteza e imprecisão inerentes às avaliações de um especialista, cuja experiência se deseja modelar num sistema de classificação baseado em conhecimento. Em outras palavras, a lógica nebulosa pode ser vista como uma maneira de expressar incerteza [21]. A lógica

nebulosa permite representar valores de pertinência (grau de verdade) intermediários entre os valores de verdadeiro e falso da lógica clássica (bivalente).

As variáveis de entrada em um sistema nebuloso são mapeadas em conjuntos nebulosos por meio de funções de pertinência. O mapeamento de uma entrada em um conjunto nebuloso gera um grau de pertinência dessa entrada ao conjunto nebuloso.

Por meio de regras SE-ENTÃO o sistema é capaz de expressar condições de pertinência. Tais regras são formuladas da seguinte forma: SE (ANTECEDENTES) ENTÃO (CONSEQUENTE). Nessa formulação os ANTECEDENTES expressam as condições a serem satisfeitas que resultam em um desdobramento representado pelo CONSEQUENTE da regra.

Através dos mecanismos da lógica nebulosa é possível a criação de sistemas de inferência nebulosa capazes de realizar tarefas complexas, como modelagem, classificação e previsão. Esse trabalho se concentra na elaboração de sistemas nebulosos voltados para a tarefa de classificação de padrões.

2.2.1.3 Sistema neuro-fuzzy

O grande problema dos sistemas de classificação nebulosos está na quantidade de parâmetros livres a serem estimados para um funcionamento adequado do modelo. Esses parâmetros livres se encontram na formação das funções de pertinência dos antecedentes e dos consequentes. Em outras palavras para se criar uma regra nebulosa do tipo SE ALGO FOR ALTO ENTÃO A SAÍDA É BAIXA precisa definir-se que tipo de função e quais os parâmetros das funções de pertinência ALTO e BAIXO nesse contexto.

Nesse contexto introduzem-se os sistemas adaptativos neuro-fuzzy. Unindo a flexibilidade dos sistemas fuzzy com a arquitetura de Redes Neurais. A ideia central por trás do algoritmo neuro-adaptativo de treinamento utilizado no ANFIS (*adaptive-network-based fuzzy inference system*) [22] é o aprendizado por meio de dados de treinamento por um algoritmo de retropropagação do erro [23] em conjunto com um método dos mínimos quadráticos. Esse método é muito semelhante ao utilizado para a aprendizagem em redes neurais artificiais.

Os sistemas ANFIS possuem uma arquitetura em rede similar a arquitetura das redes neurais artificiais, que mapeiam um conjunto de entradas em uma saída. Esse

mapeamento e feito através de funções de pertinência fuzzy. Os dados de entrada então são mapeados pelas funções de pertinência relacionadas a entrada e então por outras funções de pertinência associadas as saídas. Todas essas funções de pertinência possuem uma serie de parâmetros livres. A computação desses parâmetros e facilitada nos sistemas do tipo ANFIS pelo calculo do vetor de gradiente do erro relacionado a saída. Em outras palavras um sistema ANFIS computa os parâmetros das funções de pertinência nebulosa interativamente durante a fase de treinamento calculando a resposta obtida pelo sistema e a resposta esperada. E então pelo erro entre essas respostas o sistema automaticamente se adapta modificando os parâmetros das funções de pertinência fuzzy relacionadas com os antecedentes o os conseqüentes obtendo assim uma nova resposta para o sistema. Esse procedimento e executado com os dados de treinamento até que não se tenha mais diminuição significativa do erro relativo a saída esperada.

Para testar a capacidade de generalização do sistema fuzzy é utilizado um conjunto de dados de validação. Esses dados tem o mesmo formato que os dados de treinamento e os de teste. Geralmente procura-se um sistema de inferência fuzzy capaz de mapear uma dada entrada em uma saída de forma eficiente. O modelo dos sistemas ANFIS é fixo, ou seja, depois de treinado o sistema não modifica mais os seus parâmetros. Dessa forma existe uma tendência de *overfit* aos dados de treinamento o que pode causar um desempenho ruim durante os dados de teste. Nesse caso os dados de validação são usados e aplicados ao sistema de inferência fuzzy a cada passo do treinamento. Esse resultado e monitorado e tipicamente quando o erro de validação aumenta o sistema de inferência pode ser considerado em sua etapa de máxima generalização e o treinamento para.

Uma particularidade do sistema ANFIS e que se trata de um sistema fuzzy do tipo Takagi-Sugeno [24]. A principal conseqüência desse tipo de sistema e que somente é possível à representação de uma saída. Esse tipo de sistema foi desenvolvido inicialmente para previsão. Na pratica podemos ter em um sistema ANFIS inúmeras entradas, inúmeras regras porem apenas uma saída. Esse fato é altamente relevante nesse trabalho de classificação de padrões e define a arquitetura do classificador baseado em regras apresentado nesse trabalho.

2.2.1.4 Algoritmos Genéticos

Algoritmo genético é um método de resolução de problemas de otimização baseado nos conceitos do processo biológico de evolução das espécies conhecido como a seleção natural. A teoria da seleção natural das espécies foi desenvolvido por C. Darwin muito tempo antes do descobrimento de mecanismos genéticos [25]. A hipótese inicial de Darwin supõe que qualidades ou características de dois ou mais indivíduos se unem ou se fundem em seus filhos formando indivíduos que de alguma forma herdam características, sejam qualidades ou defeitos de seus parentes mais próximos.

Em algoritmos genéticos o vocabulário vem diretamente da genética natural. Indivíduos são soluções possíveis para um determinado problema. Esses indivíduos possuem um conjunto de características que pode ser denominado como o cromossomo do indivíduo. Cada característica individual de um cromossomo em um indivíduo é denominada gene [26], conforme a Figura 8.

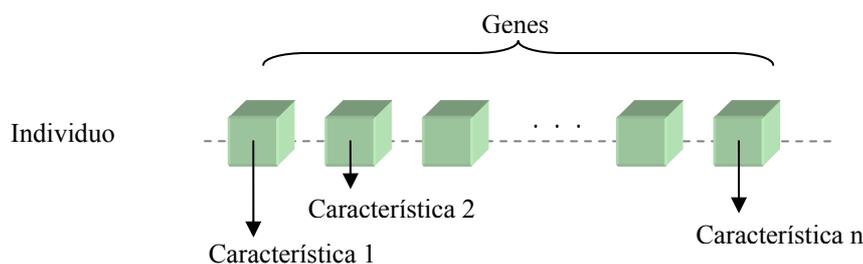


Figura 8 – Representação de um indivíduo e genes em algoritmos genéticos

Um conjunto de indivíduos em uma mesma geração é denominado de população e pode ser entendido como um conjunto de soluções potenciais para um determinado problema proposto.

Um processo evolucionário roda em uma população de cromossomos ou indivíduos correspondendo a uma procura de soluções sub-ótimas em um espaço de soluções potenciais [26]. Esse processo equilibra o interesse da busca de uma solução ótima e o interesse de exploração do espaço de soluções. Através das gerações a população evolui para uma solução ótima ou sub-ótima.

Os algoritmos genéticos utilizam regras para a criação de indivíduos em uma próxima geração a partir de indivíduos de uma geração corrente. Essas regras podem ser divididas em métodos de seleção e operadores genéticos.

O método de seleção indica a regra de escolha dos pais ou indivíduos que contribuirão para a criação da nova geração. Somente aos indivíduos selecionados são aplicados os operadores genéticos. Para que seja possível a avaliação dos indivíduos é necessário introduzir o conceito de função de avaliação. Essa função tem como entrada os genes de um indivíduo e indica a qualidade do indivíduo em uma população. Em outras palavras a função de avaliação é capaz de avaliar diferentes soluções para um problema e indicar qual é a melhor ou a menor, segundo algum critério utilizado.

Os operadores genéticos mais comuns são mutação e o *crossover*. Na mutação um gene ou uma característica de um determinado indivíduo é alterado, em geral aleatoriamente, com o objetivo de gerar um novo indivíduo em uma nova população. O *crossover* por sua vez, é a mistura de características de dois ou mais indivíduos de uma geração para produzir um indivíduo na próxima geração com características misturadas. Existem vários tipos de *crossover*, nesse trabalho utiliza-se somente *crossover* de um ponto e mutações de maneira aleatória. O *crossover* de um ponto realiza a mistura de maneira simples do material genético dos indivíduos. Na Figura 9 é exemplificado o *crossover* de um ponto. O indivíduo A em verde é “misturado” com o indivíduo B em amarelo. Um ponto de corte nos dois indivíduos é definido, em geral aleatoriamente. A partir desse ponto de corte são construídos dois indivíduos filhos que herdam o material genético dos dois pais relacionados. Dessa maneira os indivíduos gerados carregam as características ou genes, dos seus pais para a próxima geração.

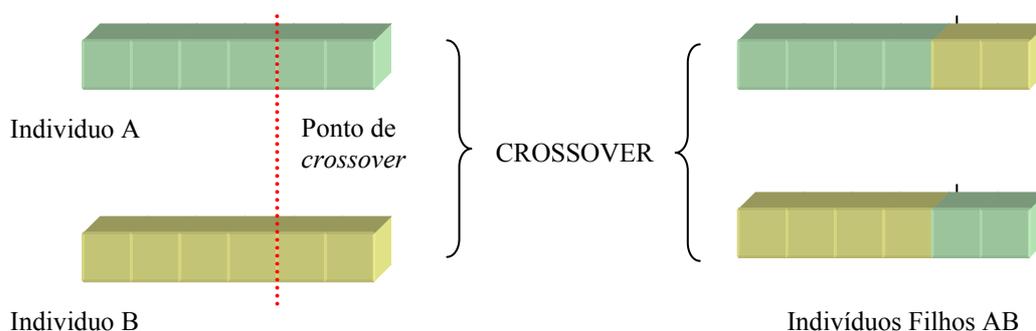


Figura 9 – demonstração do operador genético de *crossover* simples de um ponto.

2.3 Segmentação

2.3.1 Histórico

O processo de segmentação é um pré-requisito para a análise baseada em objetos.

A tarefa de segmentação divide a imagem em objetos ou segmentos. O nível dessa divisão depende do que se pretende expressar no procedimento de classificação. A tarefa de segmentação pode ser considerado “boa”, ou “aceitável”, quando os objetos de interesse do problema estiverem isolados, o que permitirá em tese discriminar adequadamente as classes da legenda [27].

As técnicas de segmentação procuram delimitar na imagem regiões com propriedades comuns. Podem-se dividir a maioria dos métodos de segmentação atuais em dois principais grupos de algoritmos: o primeiro são os métodos baseados em crescimento de regiões; e o segundo formado pelos métodos de agrupamento e divisões sucessivas de objetos [28].

Na abordagem de dividir e agrupar, conhecida como “*Split and Merge*”, os algoritmos subdividem ou agrupam segmentos sucessivamente. A decisão se um objeto deve ser subdividido, ou dois objetos adjacentes devem ser agrupados, baseia-se na avaliação de uma função de custo. As diferentes funções de custo definem as variantes dos algoritmos do tipo dividir e agrupar.

As técnicas do tipo crescimento de região ou “*region growing*” consideram como ponto de partida cada pixel como um objeto. Objetos adjacentes são agrupados formando objetos cada vez maiores. A cada passo do algoritmo a decisão de agrupar ou não objetos adjacentes leva em conta algum critério de homogeneidade do novo objeto que se formaria após o agrupamento. Os critérios de homogeneidade são específicos de cada algoritmo adotado.

Existem outros métodos mais particulares de segmentação. Esses métodos não se baseiam especificamente nem em crescimento de regiões e nem em agrupamento e divisão de objetos. Dentre eles destacam-se os métodos baseados em morfologia matemática, como o *watersheds*.

O método de segmentação alternativo baseado no divisor de águas (*watersheds*) é descrito sucintamente em [27]. Esse método utiliza a magnitude do gradiente da

imagem no seu critério de segmentação [29]. Ainda em [30] Beucher e Lantuejoul utilizam o algoritmo de *watersheds* com sucesso para a detecção de fraturas em imagens de micrografias de aço e detecção de bolhas em radiografia. Em 1991 [31] Beucher aplica o algoritmo de *watersheds* para a segmentação de imagens de sensores remotos e trata de um problema particular desta classe de algoritmos conhecido como *super segmentação*. Em [28] o algoritmo de *watersheds* é utilizado especificamente para segmentação de imagens de alta resolução.

Um método de segmentação não convencional, baseado nas características morfológicas dos componentes conectados foi proposto em [32]. Esse método parte dos resíduos de duas operações morfológicas sucessivas de abertura (*opening*) e fechamento (*closing*) [27]. O método foi testado em imagens de alta resolução tanto em áreas de campo quanto em áreas urbanas, mostrando um ótimo desempenho comparado com outros métodos de segmentação mais comuns. Uma melhor descrição das formas foi alcançada e o método intrinsecamente reduz o efeito da super segmentação devido a detalhes de implementação. Porém nesse método o custo computacional é bastante elevado, conforme mostram os experimentos apresentados no artigo, principalmente se houver nas imagens áreas muito uniformes em termos espectrais. Nesses casos os algoritmos de *watersheds* tendem a ter um desempenho melhor.

Vários algoritmos de segmentação são comparados em [33]. O estudo destacou dois softwares: o eCognition e o SPRING 4.0⁴. Ambos baseiam-se em crescimento de região, embora o SPRING ofereça também uma opção de segmentação por *watersheds*, um método que será abordado mais adiante nesta seção. Nesse trabalho também foram avaliados algoritmos de softwares como: Data Dissection Tools⁵, CAESAR 3.1⁶, InfoPACK 1.0⁷ e Image Segmentation (componente para o ERDAS Imagine)⁸.

2.3.2 Segmentação baseada em crescimento de regiões

Nesse trabalho utilizaremos um processo de segmentação baseado em crescimento de regiões desenvolvido por [34] e implementado comercialmente no

⁴ INPE - <http://www.dpi.inpe.br/spring>

⁵ INCITE, Stirling University – <http://www.incite.org.uk/projects>

⁶ N.A. Software Ltd. – <http://www.nasoftware.co.uk>

⁷ InfoSAR Ltd. - <http://www.infosar.co.uk>

software eCognition [35]. O procedimento pode ser descrito como um algoritmo de aglomeração de objetos. Inicialmente cada pixel é tratado como único objeto. Através de passos interativos esses objetos são agrupados com os seus vizinhos formando objetos cada vez maiores. Em cada passo é agrupado somente um dos vizinhos do objeto analisado, a escolha de qual vizinho será agrupado segue o critério de homogeneidade local. Em termos práticos o objeto é unido com o seu vizinho que oferece menor incremento possível na homogeneidade do objeto resultante. O processo de crescimento do objeto para quando o incremento de homogeneidade devido à união dos vizinhos é maior que uma determinada função de custo. Essa função é provida para o algoritmo como um parâmetro de escala.

O cálculo da homogeneidade local do objeto é realizado levando em conta n fatores espectrais e de forma do objeto. Em termos de forma dois parâmetros são analisados, a suavidade das bordas e a compactação do objeto gerado como um todo. Mais detalhes sobre esses cálculos podem ser obtido em [34]

Esse processo de segmentação permite uma variação no tamanho dos segmentos gerados a partir de um controle sobre o parâmetro de escala. Sendo assim a sua expansão para um procedimento em múltiplas escalas é simples e intuitiva, tornando o algoritmo muito produtivo para o processamento de imagens de alta resolução.

2.3.3 Segmentação em múltiplas escalas

A representação de objetos em múltiplas escalas é natural para o ser humano e crucial para todo o processo de reconhecimento, principalmente em imagens de alta resolução. Por exemplo, ao olhar um objeto como uma laranja muito de perto somente observa-se apenas a sua textura e não é possível a associar aquela superfície a uma casca de laranja. Ao se afastar um pouco do objeto analisado é possível definir que aquele objeto é de fato uma laranja.

A idéia da segmentação multi-escala é agrupar objetos de uma determinada escala em objetos maiores em uma escala maior. Essa aglomeração é realizada respeitando os mesmos critérios do procedimento de segmentação inicial porém mudando o parâmetro de escala na análise.

⁸ USDA Forest Service, Remote Sensing Application Center – <http://www2.erdas.com/SupportSite/>

Desta forma os objetos da imagem, ou segmentos, estão unidos por uma relação hierárquica. Nessa relação temos vários níveis de segmentações elaboradas com parâmetros de escalas diferentes. Em linhas gerais cada objeto tem um super-objeto em um nível superior que o contém; e ainda contém inúmeros objetos em um nível abaixo do seu, os chamados sub-objetos.

2.4 Parâmetros de textura

Além de informações espectrais, a textura é uma característica muito importante para a discriminação das classes na tarefa de classificação de cobertura de solo.

Os parâmetros de Haralick têm sido amplamente utilizados na classificação de imagens de sensores remotos (por exemplo [36, 45, 46 e 47]).

Os parâmetros de Haralick se baseiam na computação da matriz de co-ocorrência de níveis de cinza da imagem (GLCM – “*gray level cooccurrence matrix*”). A matriz GLCM representa a frequência de ocorrência de pares de valores de tom de cinza em uma determinada relação espacial. A partir da matriz GLCM uma série de parâmetros estatísticos pode ser extraída. Destes somente os seguintes parâmetros ou índices foram analisados nesse trabalho:

- Homogeneidade
- Contraste
- Dissimilaridade
- Entropia
- Segundo momento angular
- Média
- Desvio Padrão
- Correlação

A descrição detalhada dos cálculos dos parâmetros de Haralick pode ser encontrada em [37].