

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA**

Daniel Besen de Aguiar

**CLASSIFICAÇÃO DE SUPERFÍCIES UTILIZANDO POLINOMIAL
MAHALANOBIS E MÉTRICAS DE TEXTURA**

Florianópolis(SC)

2012

Daniel Besen de Aguiar

**CLASSIFICAÇÃO DE SUPERFÍCIES UTILIZANDO POLINOMIAL
MAHALANOBIS E MÉTRICAS DE TEXTURA**

Dissertação submetida ao Programa de Pós-Graduação em Ciências da Computação para a obtenção do Grau de Mestre em Ciências da Computação.

Orientador: Prof. Dr. rer. nat. Eros Comunello

Coorientador: Prof. Dr. rer.nat. Aldo von Wangenheim

Florianópolis(SC)

2012

Catálogo na fonte elaborada pela biblioteca da
Universidade Federal de Santa Catarina

A ficha catalográfica é confeccionada pela Biblioteca Central.

Tamanho: 7cm x 12 cm

Fonte: Times New Roman 9,5

Maiores informações em:

<http://www.bu.ufsc.br/design/Catalogacao.html>

Daniel Besen de Aguiar

**CLASSIFICAÇÃO DE SUPERFÍCIES UTILIZANDO POLINOMIAL
MAHALANOBIS E MÉTRICAS DE TEXTURA**

Esta Dissertação foi julgada aprovada para a obtenção do Título de “Mestre em Ciências da Computação”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciências da Computação.

Florianópolis (SC), 12 de Dezembro de 2012.

Prof. Dr. Ronaldo dos Santos Mello
Coordenador

Banca examinadora:

Prof. Dr. rer. nat. Eros Comunello
Orientador

Prof.^aLuciana Porcher Nedel

Prof.^aSílvia Modesto Nassar

Dr. Sylvio Luiz Mantelli Neto

A minha família, amigos e especialmente a
Renata

AGRADECIMENTOS

Agradeço ao Departamento de Informática e Estatística por toda a estrutura e conhecimento fornecido. Agradeço ao Prof. Aldo por me aceitar no Lapix e fornecer meios para que pudesse conduzir meus estudos. Agradeço ao meu orientador e amigo Prof. Eros por tudo que me foi ensinado e por tudo o que vivi no Lapix. Agradeço aos amigos Mathias, Rateke, Marcelo, Coser, Max, Eduardo, Vito, Rodrigo, Juliano, Leonardo, Fernando(Belém), Lucas(72) e especialmente ao Sylvio, Antônio e Adiel por todas as conversas e discussões engrandecedoras. Agradeço a minha família por todo o apoio e por sempre em mim acreditar. Agradeço especialmente a minha amada Renata por todo o apoio e dedicação. Enfim, agradeço a todos aqueles que de alguma forma me auxiliaram na realização desta conquista.

O destino é uma desculpa tola para o fracasso.

Ambrose Bierce

RESUMO

A classificação temática em imagens aéreas ou de satélite é uma etapa de fundamental importância para a área de sensoriamento remoto. Ela tem como objetivo determinar a classe a que uma determinada superfície encontrada em uma imagem pertence. Várias técnicas de classificação de imagens foram adaptadas para o domínio de sensoriamento remoto. Mesmo com os recentes avanços conseguidos na área de classificação, essas técnicas ainda não conseguem resultados similares aos obtidos por um observador humano. Um dos prováveis motivos é a capacidade de um operador humano utilizar uma gama variada de informações além da cor, sendo que entre essas informações adicionais se encontra as características de textura que uma determinada classe apresenta. Com o intuito de expandir o repertório de técnicas de classificação utilizadas na área de sensoriamento remoto se apresenta neste trabalho a utilização do classificador polinomial Mahalanobis juntamente com métricas locais de textura. Os resultados finais foram analisados pelos teste de acurácia obtendo taxas de acerto variando entre 80,0% e 99,0%. Esses resultados demonstram a viabilidade do uso da técnica utilizada como classificador no domínio do sensoriamento remoto.

Palavras-chave: sensoriamento remoto, classificação supervisionada, textura, polinomial Mahalanobis

ABSTRACT

Image classification is a fundamental step in remote sensing applications. It has the objective to identify and quantify surface patterns for resource assessments. Several techniques are used for that purpose, and a common difficulty is dealing with discrimination of multidimensional non-linear nearby patterns. Several classical methods like fuzzy, neural networks (NN), support vector machine (SVN) etc. are being used for that purpose, but every one have restricted performance specially in the case of non-linear classification functions. A novel approach never used in that area with a great potential to improve performance is the non-linear Polynomial Mahalanobis (PM). In the present work it is presented a PM supervised method that allows marking of a specific pattern based on a combination of multidimensional pixel value and texture features. The method was tested on air photogrametric images to assess specific features or cases like forests, water and cultivated areas for the State of Paraná, Brazil. The result analysis on the selected cases showed that the method presented a very good performance on classification allowing a better resource assessment.

LISTA DE FIGURAS

Figura 1	Representação de uma onda eletromagnética	33
Figura 2	Classificação do espectro eletromagnético segundo o comprimento de onda	34
Figura 3	Exemplo de reflexão, absorção e transmissão de energia.	36
Figura 4	Mapa temático dos remanescentes da cobertura vegetal Brasileira do PROBIO/MMA.	39
Figura 5	Processo de criação de Mapas Temáticos. (SCHOWENGERDT, 2007).	40
Figura 6	Interface da ferramenta	56
Figura 7	Exemplo de seleção de amostras da classe Recursos Naturais.	58
Figura 8	Exemplo de seleção de amostras da classe Plantações	59
Figura 9	Exemplo de seleção de amostras da classe Recursos Hídricos.	60
Figura 10	Mapa de probabilidades para uma classe tipo 2 (plantação).	61

LISTA DE TABELAS

Tabela 1	Vetores unidimensionais utilizados	51
Tabela 2	Definição das superfícies classificadas	57
Tabela 3	Tamanho dos conjuntos de treinamento de cada imagem	59
Tabela 4	Acurácia dos resultados positivos. Comparativo entre o Polinomial Mahalanobis e o padrão ouro	65
Tabela 5	Resultado dos testes realizados com a imagem c62nee.	66
Tabela 6	Resultado dos testes realizados com a imagem c62nof.	67
Tabela 7	Resultado dos testes realizados com a imagem c62neb.	68
Tabela 8	Resultado dos testes realizados com a imagem d13nof.	69
Tabela 9	Resultado dos testes realizados com a imagem d13sof.	70
Tabela 10	Resultado dos testes realizados com a imagem d41soe.	71

LISTA DE ABREVIATURAS E SIGLAS

SR	Sensoriamento Remoto	23
PM	Polinomial Mahalanobis	24
IHC	Interação Humano Computador	25
SR	Sensoriamento Remoto	31
SIG	Sistema de Informação Geográfica	38
VANT	Veículo Aéreo Não-Tripulado	38
PDI	Processamento Digital de Imagens	42
VC	Visão Computacional	42
JPEG	Joint Photographic Experts Group	55
PNG	Portable Network Graphics	55
HSV	<i>Hue, Saturation e Value</i>	56
R	Red	57
G	Green	57
B	Blue	57

SUMÁRIO

1 INTRODUÇÃO	23
1.1 CONTEXTUALIZAÇÃO	23
1.2 DEFINIÇÃO DO PROBLEMA	24
1.3 JUSTIFICATIVA	24
1.4 MOTIVAÇÃO	24
1.5 HIPÓTESE	25
1.6 TRABALHOS CORRELATOS	25
1.6.1 Classificação supervisionada utilizando K-NN	25
1.6.2 Classificação supervisionada utilizando GMM	26
1.6.3 Classificação utilizando logica fuzzy	27
2 OBJETIVO GERAL	29
2.1 OBJETIVOS ESPECÍFICOS	29
3 FUNDAMENTAÇÃO TEÓRICA	31
3.1 SENSORIAMENTO REMOTO	31
3.2 ENERGIA ELETROMAGNÉTICA	33
3.3 COMPORTAMENTO ESPECTRAL DE OBJETOS NATURAIS ..	35
3.4 IMAGENS ORTORRETIFICADAS	37
3.5 MAPAS TEMÁTICOS	38
3.5.1 Extração de Características	40
3.5.2 Treinamento do Classificador	41
3.5.3 Classificação	41
3.6 PRINCÍPIOS FUNDAMENTAIS DE SIMILARIDADE	42
3.6.1 Espaço vetorial e distância	43
3.6.2 Distância entre coordenadas no espaço	44
3.6.3 Distância em distribuições multivariadas	45
3.7 POLINOMIAL MAHALANOBIS	45
3.8 TEXTURA	47
3.8.1 Matriz de Co-ocorrência	48
3.8.2 Mascaras de Laws	51
3.8.3 Filtro de Gabor	53
4 MATERIAIS E MÉTODO	55
4.1 BANCO DE IMAGENS	55
4.2 TREINAMENTO	55
4.2.1 Classificação	59
5 RESULTADOS E DISCUSSÕES	63
6 CONSIDERAÇÕES FINAIS	73
6.1 RECOMENDAÇÕES	73

Referências Bibliográficas 75

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

Nas últimas décadas, o sensoriamento remoto tornou-se uma das principais fontes de informação sobre superfícies terrestres. Imagens aéreas e de satélite são utilizadas nas mais variadas aplicações como em estudos climáticos, controle de desmatamentos e gerenciamento de plantações (KRAMER, 1994). O processo de caracterizar um pixel de imagem em uma superfície qualquer é denominado de classificação de imagem e os algoritmos utilizados para essa classificação são denominados de classificadores de imagem (GONZALEZ; WOODS, 2001). A extração de informações sobre superfícies terrestres a partir de imagens aéreas ou de satélite tornou-se prioridade para o trabalho com SR. Alguns dos classificadores utilizados nesses estudos são classificadores tradicionais como K-NN e métricas estatísticas. Com o avanço tecnológico, novos classificadores passaram a ser utilizados na área de SR, como redes neurais artificiais (RUSSELL; NORVIG, 2010; NEWELL; SIMON, 1972), redes bayesianas, lógica nebulosa ou *fuzzy* (MOUSTAKIDIS et al., 2012) e *support vector machines* (LI et al., 2012). Apesar dos avanços tecnológicos nas áreas de desenvolvimento de técnicas e de equipamentos (computadores e sensores) dos últimos anos, ainda existe um campo vasto de conhecimento a ser trabalhado, seja por aprimoramento das técnicas existentes ou inclusão de novos classificadores de imagem.

Uma das formas utilizadas para melhorar a qualidade da classificação é obter mais informações sobre a imagem analisada. Essas informações podem ser dados geo-referenciados, multi-espectral, ou seja, captar a imagem com frequências variadas e informações sobre textura. Existe uma grande quantidade de parâmetros que podem ser extraídos de imagens de SR, tais como características estatísticas de textura, características geométricas de textura e informação multi-espectral. A utilização de métricas de textura permite distinguir entre diferentes tipos de estruturas nas imagens de forma mais precisa do que utilizando somente informações de cor. A análise de textura tem desempenhado um papel importante no processamento e interpretação de imagens digitais, principalmente pelo fato de adicionar informações sobre as características da imagem (PACIFICI; EMERY, 2009).

1.2 DEFINIÇÃO DO PROBLEMA

A classificação de superfícies em imagens aéreas ou de satélite é um procedimento complexo e, apesar dos grandes avanços na área de classificação automatizada, esses sistemas fornecem resultados inferiores aos obtidos por um observador humano. A problemática atual gira em torno do desenvolvimento de sistemas automatizados que obtenham resultados mais próximos dos obtidos por pessoas. Não existe um consenso no meio acadêmico quanto a melhor técnica para este fim. Observa-se como promissora a utilização de técnicas tradicionais de classificação na área de SR. O permite a representação do espaço criado durante a fase de treinamento ajustando-se de forma precisa ao conjunto de dados original. É uma extensão da distância de Mahalanobis, sendo esta uma métrica de similaridade multivariada com ampla utilização pela comunidade científica. O polinomial Mahalanobis possui como características ser altamente escalável e adaptável, permitindo o incremento de sua complexidade ao inserir novas características dos objetos analisados.

1.3 JUSTIFICATIVA

O processamento digital de imagens é um sub-conjunto de tarefas ligadas à visão computacional. É notório que sistemas de visão computacional ainda não possuem um conjunto de técnicas definitivas para solução global de problemas gerais. É válido salientar também que a classificação de superfícies em imagens aéreas ou de sensoriamento remoto utiliza um encadeamento de procedimentos até que se encontre um resultado final satisfatório. O refinamento do resultado de uma das etapas do encadeamento resulta em uma melhora no resultado final.

A classificação de uma tarefa é crucial dentro desse encadeamento e o aperfeiçoamento no resultado dessa classificação proporciona uma melhora no resultado final. Qualquer refinamento de resultado encontrado no processo de classificação é considerado positivo no sentido de aperfeiçoar os resultados obtidos e conquistados na área de sensoriamento remoto. Esse trabalho tem como objetivo à partir de atributos de cor e textura classificar imagens aéreas de SR utilizando PM.

1.4 MOTIVAÇÃO

O algoritmo de classificação utilizando polinomial Mahalanobis como métrica de similaridade apresenta excelentes resultados quando utilizado para

classificação de objetos utilizando somente informação de cor. O mesmo algoritmo é utilizado em aplicação de Interação Humano Computador IHC para a classificação de pele. Na área de robótica, também tem sido utilizado utilizado para identificar o caminho a ser seguido em imagens de vídeo utilizados para navegação autônoma. Apesar desses resultados, o algoritmo em questão ainda não foi utilizado na área de sensoriamento remoto.

O PM tem apresentado bons resultados com atributos multidimensionais de cor, e nesse trabalho é expandido para trabalhar com atributos de textura. Por ser a área classificação temática de imagens de SR uma área onde tanto os atributos de cor quando de textura são utilizados na classificação, optou-se por verificar a utilização do PM adicionando atributos de cor em imagens de SR.

1.5 HIPÓTESE

Espera-se que a classificação de um pixel em imagens de sensoriamento remoto utilizando atributos locais de textura, informações de cor à partir do algoritmo PM como métrica de similaridade apresente resultados satisfatórios quando comparados aos obtidos por um observador humano.

1.6 TRABALHOS CORRELATOS

1.6.1 Classificação supervisionada utilizando K-NN

Al-Kofahi et. al. (AL-KOFAHI et al., 2012) utilizaram fotografias aéreas de alta resolução (0,15m) e classificadores supervisionados orientados a objetos para mapear o uso do solo em áreas residenciais. O objetivo da pesquisa era de identificar diferentes tipo de vegetações, tais como: árvores, grama e arbustos. Os autores visitaram as todas as localidades estudadas, com o objetivo de criar um padrão ouro para avaliar o resultado da ferramenta analisada. Para a criação do padrão ouro todo tipo de superfície foi levada em consideração. Entre essas superfícies encontram-se: superfícies impermeáveis, árvores, arbustos, grama e piscinas.

O modulo de extração de características utilizado foi o *Feature Extraction Module (ENVI EX 4.7.1)*, que faz parte do software de processamento ENVI. Essa ferramenta realiza classificação supervisionada orientada a objetos. A ferramenta extrai objetos de imagens de alta resolução e classifica os mesmos com base nas suas características espacial, espectral e de textura. A ferramenta segmenta parte das imagens em grupos de pixels vizi-

nhos com características similares (principalmente cor) e gera uma estrutura de informação semântica e espacial dos segmentos da imagens que são utilizados para diferenciar objetos.

As informações espacial e textura foram utilizadas para classificar a cobertura do solo. A distribuição espacial dos segmentos dos objetos determina os atributos de textura, que são a distribuição espacial e a frequência dos valores dos pixels de uma imagem. Os autores utilizaram a média, variância, entropia e o *range* da textura para caracterizar a textura das imagens. Os valores de máximo, mínimo e desvio padrão não foram utilizados pois segundo os autores os mesmos reduzem a acurácia da classificação devido a similaridades na reflectância espectral dos diferentes tipos de vegetação e da alta correlação entre as bandas vermelha, verde e azul. O algoritmo utilizado foi K-NN *K-Nearest Neighbor*, utilizando a distância Euclidiana como medida de similaridade. O valor de *K* foi igual a 1, ou seja, a classe de um segmento depende da classe do do ponto do conjunto de treinamento mais próximo.

Os resultados obtidos pela ferramenta foram analisados através do teste de acurácia. Matrizes de confusão foram geradas e os resultados obtidos pelas mesmas foram refinados utilizando o coeficiente kappa e o teste Z. Os autores obtiveram uma taxa de acerto que variou de 89,0% até 97%. A técnica se mostrou eficaz ao classificar superfícies no solo.

1.6.2 Classificação supervisionada utilizando GMM

Quando se utiliza classificação supervisionada, os padrões espaciais em um conjunto de imagens são avaliados por um computador utilizando regras de decisão pré-definidas para determinar a identidade de cada pixel. Classificação supervisionada requer que dados previamente classificados por um especialista sejam utilizados de forma a permitir que pixels nas imagens sejam associados de forma correta a sua classe. Thapa et. al. (THAPA; MURAYAMA, 2008) identificaram amostras homogêneas nas imagens que foram utilizadas como amostras representativas para cada categoria de uso do solo para treinar o algoritmo para localizar pixels similares na imagem. Para cada tipo de uso do solo, entre 5 e 10 áreas de interesse foram preparadas como sendo as assinaturas das amostras de treinamento. As áreas de treinamento foram criadas de forma a discriminar as classes individuais. Os dados de referência no solo foram utilizados para preparar as assinaturas de treinamento. Após obter uma discriminação satisfatória entre as classes durante a avaliação da assinatura espectral, classificação supervisionada utilizando o classificador de máxima verossimilhança foi executado utilizando as quatro bandas da imagem. Esse classificador avalia quantitativamente a variância e a covariância

dos padrões de resposta da categoria espectral quando classificando um pixel desconhecido.

1.6.3 Classificação utilizando logica fuzzy

A classificação supervisionada utilizando fuzzy (THAPA; MURAYAMA, 2008) trabalha utilizando-se uma função de pertinência, onde o valor de um pixel é determinado se o mesmo encontra-se mais perto de uma classe do que de outra. Essa abordagem considera que cada pixel pode pertencer a várias classes diferentes sem fronteiras definidas. Assim, essa abordagem pode trabalhar com o problema de pixels misturados ou com o problema de representação de características heterogêneas. Para o teste dessa abordagem os autores prepararam entre 5 e 10 áreas de treinamento para cada classe de superfície no solo. Ao invés de delinear áreas de treinamento puramente homogêneas, uma combinação entre áreas puras e misturadas foram utilizadas. A mistura de vários tipos de características definiram os pesos das classes do conjunto de treinamento fuzzy. Um pixel classificado recebe um grau de filiação com respeito a cada classe de informação. Dois mapas foram criados, o mapa de classes de várias camadas e o mapa de distâncias. Uma convolução fuzzy foi então realizada para criar uma camada simples de classificação ao calcular o peso total inverso das distâncias de todas as classes em um kernel 3×3 . Essa operação atribui o pixel central do *kernel* a classe com a maior distância inversa total somada sobre todas as camadas de classificação do conjunto fuzzy. Classes com valores de distância muito baixos permanecem inalteradas enquanto que classes com valores de distância maiores podem mudar o seu valor caso haja um número suficientes de pixels na vizinhança com valores de classe e valores de distância pequenos.

2 OBJETIVO GERAL

O objetivo geral do trabalho é demonstrar a utilização de atributos locais de textura e cor no Polinomial Mahalanobis para a classificação de superfícies em imagens aéreas de SR.

2.1 OBJETIVOS ESPECÍFICOS

- Criar um banco de imagens de sensoriamento remoto;
- Desenvolver uma ferramenta para seleção de amostras das imagens;
- Definir uma métrica para comparar atributos de textura em imagens de SR.
- Aplicar atributos de textura local no polinomial Mahalanobis; e
- Avaliar os resultados encontrados.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 SENSORIAMENTO REMOTO

Sensoriamento Remoto é definido como a aquisição de características e propriedades de objetos encontrados na superfície da terra utilizando dados adquiridos a partir de aviões, radares ou satélites (SCHOWENGERDT, 2007). É assim uma tentativa de se extrair informações estando à distância do objeto analisado, sem a necessidade de ir até o local. Considerando que não há contato direto com os objetos de interesse, deve-se confiar na informação adquirida a partir de dados ópticos, acústicos ou de micro-ondas (SCHOWENGERDT, 2007). Pode-se afirmar que SR é uma outra forma de se olhar e interpretar o mundo.

Essa nova maneira de observar o mundo frequentemente fornece uma grande quantidade de dados que não poderiam ser adquiridos ou observados através de estudos ou observações locais, utilizando abordagens tradicionais de observação como por exemplo, estudo de superfícies. Estudos tradicionais podem se limitar a fornecer uma grande quantidade de dados de apenas algumas amostras tendo somente coletado dados de uma área pequena ou restrita. Sistemas de SR, especialmente os que utilizam informações de satélite, proporcionam uma visualização repetitiva e consistente da superfície do planeta imprescindíveis para observações de mudanças de curto e longo prazo (SCHOWENGERDT, 2007).

Os recursos oferecidos pelo sensoriamento remoto permitem a observação de florestas inteiras, regiões, continentes e até mesmo todo o planeta. Essa perspectiva permite que se combine dados coletados no solo de forma a se obter um resultado mais consistente. O SR permite observação em faixas espectrais diferentes à partir de sensores especiais. Esta característica permite que sejam observados outras informações como temperatura da superfície, quantidade de água, identificação de plantas e culturas.

Algumas das áreas de aplicação da tecnologia de sensoriamento remoto são listadas abaixo:

- Avaliação e monitoramento ambiental (crescimento urbano)
- Monitoramento e detecção de mudanças globais (aquecimento global, danos à camada de ozônio, danos florestais)
- Agricultura (avaliação de plantações, previsões de safra e erosão do solo)

- Exploração de recursos não-renováveis (petróleo, gás natural e mineração)
- Meteorologia (previsão do tempo, dinâmica atmosférica)
- Mapeamento (topografia, uso do solo, engenharia civil)
- Aplicações militares (estratégia política, reconhecimento)

Para atender as mais diferentes demandas, foram desenvolvidos vários sistemas de sensoriamento remoto, oferecendo uma grande variedade de parâmetros espaciais, temporais e espectrais. Análises meteorológicas necessitam frequentemente de dados repetitivos, de uma mesma área, com uma baixa variação temporal, ou seja, o tempo entre a coleta desses dados deve ser baixo. Aplicações de mapeamento necessitam de grandes resoluções, sem a necessidade de variação temporal. Aplicações militares necessitam de dados tanto com alta resolução quanto baixa variação temporal (SCHOWENGERDT, 2007).

Existe uma grande variedade de informações encontradas em uma imagem que, quando analisadas, fornecem informações valiosas sobre as superfícies e classes nela contida. Essa classificação é realizada por um analista experiente que encontra e identifica áreas ou características importantes. As principais características utilizadas por um especialista durante o processo de classificação são listadas abaixo:

- Forma: é a descrição geométrica do objeto. A forma de um objeto diz muito sobre a origem e função de um objeto analisado. Através da forma é possível determinar se um objeto em questão foi feito pelo homem ou não.
- Tamanho: ao se encontrar um objeto conhecido na cena é possível inferir sobre o tamanho de um objeto analisado e facilitar assim a sua classificação.
- Cor: computadores analisam cor melhor do que observadores humanos. É possível distinguir objetos simplesmente através da análise de cor.
- Textura: descreve a estrutura da variação de brilho em um objeto. Objetos com cores similares podem ser diferenciados ao comparar-se a sua textura.
- Padrão: são formas com geometria identificável.
- Localização: refere-se à localização geográfica da área analisada. Ao conhecer características do local é possível diminuir a possibilidade de classificação através de eliminação.

3.2 ENERGIA ELETROMAGNÉTICA

A principal fonte de informação utilizada por sistemas de sensoriamento remoto é a energia eletromagnética emitida ou refletida pelos objetos terrestres. A energia eletromagnética também pode ser chamada de radiação eletromagnética. O sol é a principal fonte de energia eletromagnética utilizada em SR. Entretanto, toda matéria que possua temperatura superior ao valor do zero absoluto (0 K ou -273°C) emite de forma contínua energia eletromagnética, embora essa energia seja de magnitude e composição espectral consideravelmente diferente. Assim, objetos encontrados na superfície terrestre são também fontes de energia eletromagnética (LILLESAND, 2006). Desta forma, todo corpo com temperatura superior ao zero absoluto é considerado uma fonte de energia eletromagnética.

A energia eletromagnética assume formas e denominações variadas. Ondas de rádio, calor, raios ultra-violeta e raios-X são algumas das mais conhecidas. Todas essas energias são inerentemente similares e são representadas de acordo com a teoria de ondas (LILLESAND, 2006). A energia eletromagnética é descrita como uma energia que viaja de forma harmônica e senoidal na velocidade da luz, conforme mostrado na Figura 1. O comprimento de onda λ representa a distância entre um pico de onda e o seu vizinho mais próximo. A frequência f informa a quantidade de picos de onda que passam por um ponto fixo no espaço por unidade de tempo.

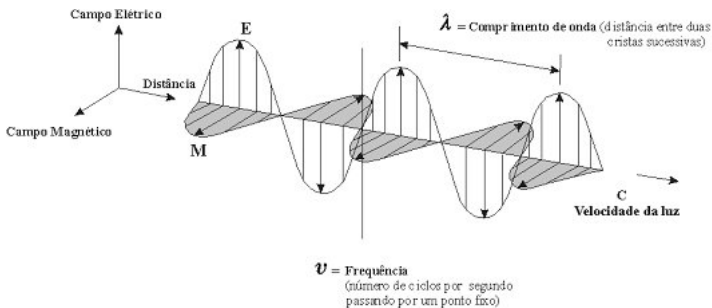


Figura 1 – Representação de uma onda eletromagnética

As ondas obedecem a seguinte equação geral 3.1:

$$c = \lambda * f \quad (3.1)$$

onde c é a velocidade da luz cujo valor é $3 \times 10^8 \text{ m/seg}$. Haja vista que

o valor c é constante, uma onda é caracterizada pelo seu comprimento de onda e sua frequência, sendo ambos inversamente proporcionais (LILLESAND, 2006). A energia eletromagnética pode ser ordenada de acordo com a sua frequência ou comprimento de onda, sendo essa ordenação chamada de espectro eletromagnético. O espectro eletromagnético apresenta subdivisões classificadas em função do tipo de processo físico que dá origem à energia eletromagnética, da interação entre radiação e objeto e transparência da atmosfera em relação à radiação atmosférica. A Figura 2 mostra as várias faixas do espectro eletromagnético classificadas de acordo com o comprimento de onda.

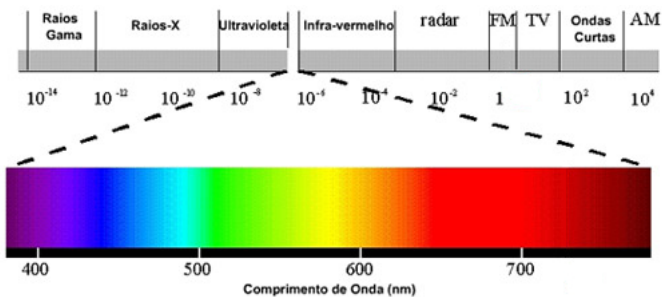


Figura 2 – Classificação do espectro eletromagnético segundo o comprimento de onda

O espectro magnético é dividido nas seguintes regiões:

- Radiação Gama: é o espectro com maior quantidade de energia. A radiação gama não possui um limite inferior quanto ao seu comprimento de onda, porém, seu limite superior encontra-se aproximadamente em $0,01 \mu m$. A radiação gama possui um alto poder de penetração. É utilizada em aplicações médicas e industriais.
- Raio-X: possui comprimento de onda um pouco menor do que a radiação gama, tornando arbitrária a diferenciação entre os dois. É comumente utilizado em aplicações médicas.
- Ultravioleta (UV): aproximadamente 10% de toda a radiação do sol pertence ao espectro ultravioleta e o mesmo encontra-se na faixa de $0,003m$ até aproximadamente $0,38m$. É nocivo às formas de vida terrestres, porém é bloqueado em grande parte pela camada de ozônio, o oxigênio e o hidrogênio presentes na atmosfera.

- Visível: é a faixa do espectro eletromagnético que pode ser detectada pelo sistema visual humano. Diferentes comprimentos de onda produzem sensações diferentes de cor. As cores estão associadas aos seguintes intervalos espectrais:

violeta: 0,38 a 0,45 μm

azul: 0,45 a 0,49 μm

verde: 0,49 a 0,58 μm

amarelo: 0,58 a 0,60 μm

laranja: 0,60 a 0,62 μm

vermelho: 0,62 a 0,70 μm

- Infravermelho (IV): encontra-se entre as frequências 0,7 e 1000 μm , e normalmente é dividida em em três sub-regiões:

IV próximo: 0,7 a 1,3 μm

IV médio: 1,3 a 6 μm

IV distante: 6 a 1000 μm

A teoria das partículas é outra forma de entender ou explicar como a energia eletromagnética reage com a matéria. Essa teoria sugere que a radiação eletromagnética é composta de várias unidades discretas, chamadas de fótons ou *quanta* (LILLESAND, 2006). A energia de um *quantum* é dada pela equação 3.2

$$Q = h.f \quad (3.2)$$

onde:

- Q = energia de um *quantum*, dada em *Joules* (J)
- h = constante de *Planck*, 6,626x10⁻³⁴F.seg
- f = frequência

3.3 COMPORTAMENTO ESPECTRAL DE OBJETOS NATURAIS

A energia eletromagnética que incide sobre uma superfície sofre influência do material que a compõe. Essa influência se dá através de reflexão, absorção ou transmissão de energia através do objeto. A capacidade de um

determinado tipo de material absorver, refletir ou transmitir a radiação eletromagnética recebe o nome de absorvância, reflectância e transmitância (BOWKER et al., 1985). A reflexão é a proporção entre o fluxo de energia electro-magnética incidente numa superfície e o fluxo que é refletido. A absorvância é a medida da quantidade de energia luminosa que é absorvida por um determinado objeto ou superfície. A transmitância é a fração da energia luminosa que atravessa um objeto. Um exemplo desses três comportamentos pode ser visto na Figura 3.

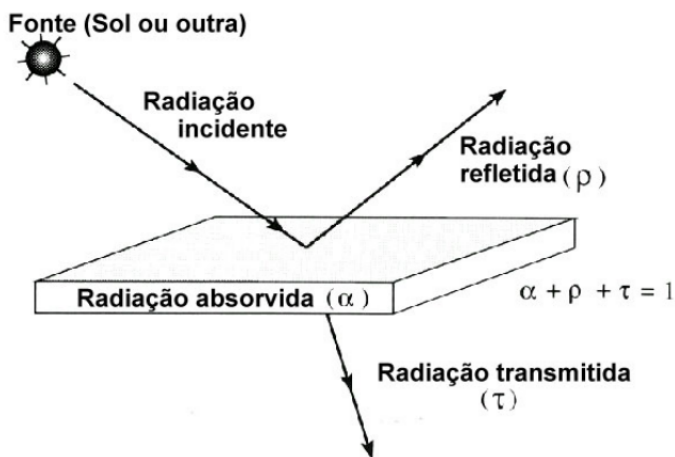


Figura 3 – Exemplo de reflexão, absorção e transmissão de energia.

O comportamento espectral de um objeto é a forma como ele reage quando exposto a um determinado espectro. Esses valores são também conhecidos como assinatura espectral. Superfícies com características físico-químicas e biológicas diferentes apresentam assinatura espectral diferente. Essa diferença permite que sistemas de sensoriamento remoto possam distinguir e reconhecer diferentes superfícies (BOWKER et al., 1985).

O conhecimento do comportamento espectral dos objetos terrestres é muito importante para a escolha da região do espectro sobre a qual pretende-se adquirir dados para determinada aplicação. As características básicas observadas no comportamento espectral destes objetos são:

- vegetação sadia: apresenta alta absorção da energia eletromagnética na região do espectro visível, que é capturada pela clorofila para a realização da fotossíntese. Dentro do espectro visível, a absorção é mais fraca na região que caracteriza a coloração da vegetação e a alta

reflectância no infravermelho próximo (até 1,3) é devida à estrutura celular, sendo que a partir deste comprimento de onda é o conteúdo de água na vegetação que modula as bandas de absorção presentes no comportamento espectral desta.

- o comportamento espectral dos solos é também dominado pelas bandas de absorção de seus constituintes. As combinações e arranjos dos materiais constituintes dos solos é o que define o seu comportamento espectral, sendo que os principais fatores são a constituição mineral, a matéria orgânica, a umidade e a granulometria (textura e estrutura) deste.
- A água pode se apresentar na natureza em três estados físicos, os quais apresentam comportamento espectral totalmente distintos. O comportamento espectral da água no estado líquido apresenta baixa reflectância (menor do que 10%) na faixa compreendida entre 0,38 e 0,7 μ m. O comportamento espectral de corpos de água é modulado principalmente pelos processos de absorção e espalhamento produzidos por materiais dissolvidos e em suspensão neles pois é verificado que a presença de matéria orgânica dissolvida em corpos de água desloca o máximo de reflectância espectral para o verde-amarelo, enquanto que a presença de matéria inorgânica em suspensão resulta num deslocamento em direção ao vermelho.
- O comportamento espectral de nuvens apresenta elevada reflectância (em torno de 70%) em todo o espectro óptico com destacadas bandas de absorção em 1 μ m, 1,3 μ m e 2 μ m.

3.4 IMAGENS ORTORRETIFICADAS

A fotografia aérea é uma das principais fontes de imagens utilizadas na produção de orto-imagens ou imagens ortorretificadas. Imagens ortorretificadas são imagens cujas distorções geométricas ocorridas durante o processo de aquisição foram corrigidas geometricamente (MATHER, 2005). Essa correção permite que objetos de interesse sejam mostrados em sua posição real no solo e com escala constante em toda a extensão da imagem. Essas imagens são criadas após a remoção das variações na escala e localização causadas pela inclinação da aeronave e o relevo do terreno. Essas imagens são utilizadas em várias aplicações de SR, tais como cartografia, monitoramento ambiental e planejamento de obras de grande porte e de cidades. A correção pode ser justificada por diversas circunstâncias, como, por exemplo,

transformar uma imagem para coincidir com um mapa, localizar pontos de interesse no mapa e na imagem, sobrepor sequências de imagens temporais da mesma área quando obtidas por diferentes sensores e sobrepor imagens e mapas em um SIG.

Cada sistema de aquisição de imagens produz distorções geométricas únicas nas imagens capturadas. Essas distorções podem variar consideravelmente devido a inúmeros fatores, sendo a plataforma utilizada para capturar a imagem (avião, VANT ou satélite), a órbita e o tipo de sensor (visível, infra-vermelho ou micro-ondas e resolução baixa ou alta) os fatores de maior influência. A forma mais comum de se capturar imagens aéreas é utilizando câmeras e lentes especiais e posicionando a câmera de forma perpendicular ao solo. As câmeras são montadas em giro-estabilizadores com o objetivo de diminuir os efeitos da inclinação da aeronave e as imagens são capturadas em um voo realizado em linha reta. Durante o voo, as imagens capturadas são sobrepostas a uma taxa que varia entre 60% e 80% e a área que se encontra entre duas imagens subsequentes permite a visualização estéreo das áreas comuns entre dois quadros sucessivos (SMITH, 2012).

Uma forma de retificação utilizada é o método de *anchor points* (MAYR; HEIPKE, 1998). Nesse método um conjunto de pontos de controle no solo, batizados pelo autor de *anchor points* são coletados em locais que possibilitem a sua identificação tanto na imagem quanto em um mapa correspondente. Após a coleta de uma quantidade suficiente de pontos de controle as coordenadas das imagens são modeladas como funções de um mapa de coordenadas utilizando o método dos mínimos quadrados para encaixar polinômios de baixa ordem aos dados. Uma das vantagens desse método é o baixo custo computacional.

3.5 MAPAS TEMÁTICOS

Uma mapa temático mostra a distribuição espacial de padrões identificáveis da superfície da terra e tem como objetivo fornecer informações sobre uma determinada área. Vários tipos de informações podem ser disponibilizadas tais como, por exemplo: categorização de tipos de solo, vegetação e profundidade de rios e lagos. É necessário para a construção de mapas temáticos utilizando imagens de SR que as categorias que se deseja classificar possam ser diferenciadas através dos dados coletados da imagem. Um exemplo de mapa temático é mostrado na Figura 4

Uma forma de criação de mapas temáticos a partir de imagens é com a utilização de algoritmos de classificação de imagens. Tradicionalmente, a classificação temática de uma imagem envolve alguns passos, conforme

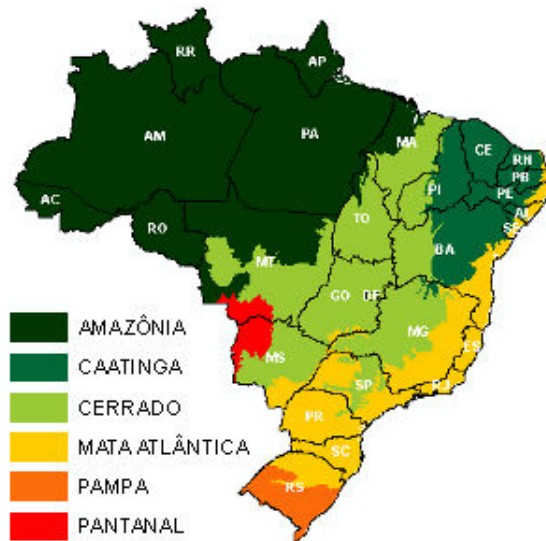


Figura 4 – Mapa temático dos remanescentes da cobertura vegetal Brasileira do PROBIO/MMA.

mostrado na Figura 5 e listados abaixo:

- Extração de Características: transformação de uma imagem multi-espectral para uma imagem comum. A imagem multi-espectral pode ser usada diretamente, sendo esse passo opcional.
- Treinamento: seleção de amostras da imagem que farão parte do conjunto de treinamento do classificador. Esse passo pode ser supervisionado por um especialista ou não-supervisionado, com a ajuda de algoritmos computacionais.
- Classificação: uso de métricas de similaridade para comparar os pixels das imagens com o conjunto de treinamento e assim determinar a qual classe pertence cada pixel.

O resultado é a caracterização das diferentes superfícies ou estruturas presentes no conjunto de dados inicial.

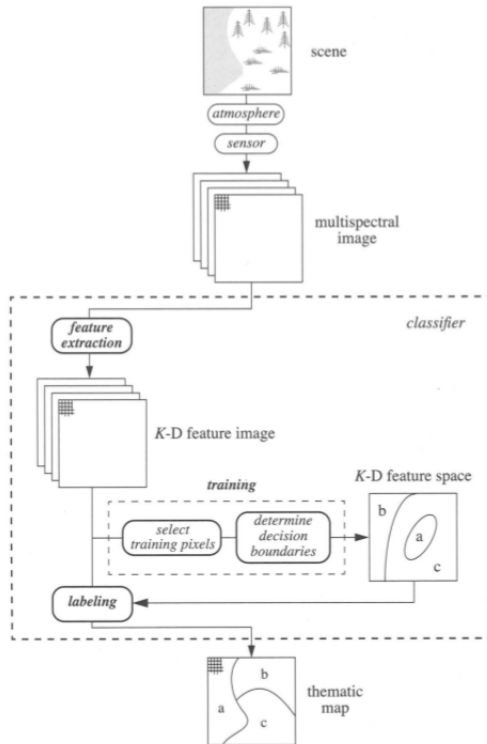


Figura 5 – Processo de criação de Mapas Temáticos. (SCHOWENGERDT, 2007)

3.5.1 Extração de Características

Dados de imagens multi-espectrais podem ser usados diretamente na classificação, porém contêm certas influências externas, tais como espalhamento atmosférico e relevo topográfico. Os dados são frequentemente correlacionados entre as bandas espectrais, resultando em uma análise ineficiente. Além disso, características derivadas da imagem, tais como medidas de estrutura espacial proporcionam informações mais úteis para a classificação. Assim, é prudente considerar várias transformações pré-classificação para extrair a maior quantidade de informações da imagem original. Algumas características tendem a suprimir variabilidades indesejadas em assinaturas

em sensoriamento remoto, por exemplo, relações espectrais podem suprimir variações topográficas. A utilização dessas características permitem uma melhor classificação entre classes espectrais distintas.

3.5.2 Treinamento do Classificador

Para que seja possível classificar as várias superfícies de uma imagem em categorias de interesse é necessário uma fase de treinamento para que seja possível para o algoritmo distinguir entre as várias classes. Conjuntos de treinamento, exemplos ou protótipos são utilizados como amostra do conjunto de treinamento. Após o treinamento, as regras ou características que representam ou definam cada classe são usadas para classificar os pixels da imagem como pertencente a uma classe (classificação *hard*) ou a mais de uma classe (classificação *soft*) dentre as classes definidas pelo conjunto de treinamento.

Os algoritmos de treinamento podem ser supervisionados, onde as amostras do conjunto de treinamento são classificadas previamente. Podem ser não-supervisionados, onde as amostras do conjunto de treinamento não são previamente classificadas. Quando se trabalha com treinamento supervisionado, a escolha das amostras de treinamento tem impacto direto nos resultados gerados pelos algoritmos de classificação. É importante que as classes de treinamento sejam constituídas de amostras homogêneas e ao mesmo tempo que incluam um grau de variabilidade para cada classe. Vale salientar que normalmente se utilizam mais de uma área de treinamento por classe durante a definição do conjunto de treinamento. Caso se verifique na imagem uma variabilidade entre classes considerável, a seleção de dados para a criação do conjunto de treinamento pode ser trabalhosa.

3.5.3 Classificação

O objetivo da fase de classificação é determinar quais áreas na imagem possuem características físicas similares. Essa classificação é realizada agrupando-se dados com características similares, chamados de assinaturas de classe. A questão principal é: quão bem a assinatura da classe de um dado na imagem corresponde às características físicas que distinguem uma característica da outra. Um exemplo do tipo de problema semântico que surge é em mapeamento do uso do solo em áreas urbanas tais como área residencial e área industrial. Essas áreas são tipicamente compostas de vários tipos de cobertura do solo tais como: vegetação, pavimento e de tipos diferentes de coberturas de teto, cada uma com diferentes assinaturas espectrais. Dentro

de uma área particular, várias classes espectrais são encontradas, resultando em uma assinatura espectral heterogênea cujas características dependem da proporção de cada componente da cobertura do solo. Essas características variam de pixel para pixel. Assim, para o mapeamento do solo, deve-se buscar relações mais complexas entre as características físicas, a imagem multi-espectral pixel a pixel e o mapa das classes de interesse. Às vezes, dados de sensoriamento remoto são insuficientes para a tarefa, necessitando-se de informações auxiliares.

É evidente que a assinatura espectral de um material em uma superfície qualquer não pode ser caracterizada por um vetor espectral simples determinístico, mas pode ser caracterizada através de uma distribuição de vetores.

A noção de mapas temáticos, conforme tradicionalmente usada em geografia, geologia e outras ciências, presume que cada ponto no solo pode ser classificado como pertencente a uma e somente uma categoria. Essa característica discreta é conveniente por sua simplicidade, porém não é um retrato fiel de superfícies reais; de fato, é inconsistente quando utilizados dados numéricos de alta resolução.

A maioria dos algoritmos de classificação produz uma função de verossimilhança ou similaridade para a atribuição de uma classe para cada pixel. Uma classificação *hard* é produzida selecionando-se a classe com a maior verossimilhança como sendo a correta. Caso os valores de verossimilhança sejam retidos, entretanto, permitindo-se múltiplas classificações para cada pixel, uma classificação *soft* é obtida. As fronteiras do espaço de decisão para um classificador *soft* pode ser definido como *fuzzy*.

3.6 PRINCÍPIOS FUNDAMENTAIS DE SIMILARIDADE

Similaridade é uma métrica que informa o grau de semelhança entre duas ou mais entidades. Funções de similaridade são utilizadas nas mais diversas áreas, tais como classificação, redes neurais e estatística. Na área de PDI e VC a similaridade pode ser calculada utilizando pixels ou agrupamento dos mesmos. Em algoritmos de classificação, dois pixels/regiões adjacentes Ω_i e Ω_j , são considerados similares quando uma função identifica que exista simetria entre eles. As funções de similaridade de uma forma geral calculam a proximidade entre dois vetores de ordem n . A distância euclidiana normalmente assume que cada componente do vetor possui peso ou importância igual e que esses mesmos componentes são independentes. Sabe-se, no entanto, que a distância euclidiana nem sempre pode ser satisfeita em aplicações reais, especialmente quando são utilizadas altas ordens dimensionais ($m > 3$)

(XIANG; NIE; ZHANG, 2008).

Outras funções são baseadas na geometria espacial produzida por alguma força de distorção com o objetivo de ponderar o resultado da similaridade, e outras ainda fazem uso de algum tipo de conhecimento prévio para a contração/repulsão da distância calculada entre as coordenadas no espaço. Uma função de similaridade é analogamente uma métrica de distância e essas métricas, para serem eficientes, devem ser capazes de identificar características importantes e ao mesmo tempo em que discriminam características irrelevantes ou de menor importância no contexto analisado.

3.6.1 Espaço vetorial e distância

Um espaço métrico M é um conjunto finito ou infinito onde a noção de distância entre os elementos ao longo das componentes vetoriais é definida. Um espaço métrico é dado por componentes m -dimensionais e a forma mais conhecida deste é a do espaço euclidiano tridimensional, onde a noção de tempo pode ser uma quarta dimensão, conforme descrito pela teoria geral da relatividade.

Em um espaço euclidiano tridimensional, o meio mais comum de se quantificar a distância entre dois vetores x e y é através da menor possível distância entre estes pontos. Uma distância mensurável em um espaço métrico é definida por uma métrica d , onde componentes m -dimensionais arbitrárias x e y apresentam:

1. $d(x,y) \geq 0$, propriedade de não-negatividade (uma distância sempre é um valor positivo).
2. $d(x,y) = 0$, se e apenas se $x = y$. Propriedade de identidade.
3. $d(x,y) = d(y,x)$, propriedade de simetria.
4. $d(x,y) \leq d(x,z) + d(z,y)$, onde z é outra coordenada arbitrária qualquer, satisfazendo a propriedade de desigualdade do triângulo.

As condições acima especificadas valem para todo x,y,z contidos no espaço métrico, e para qualquer dimensão m .

Tipicamente, todo espaço métrico é um espaço topológico de uma forma natural, e os teoremas sobre espaços topológicos podem ser empregados em espaços métricos. Isto pode ser demonstrado se considerando uma dada coordenada x em um espaço M , onde sobre x é aplicado um valor positivo de raio r , tal que: $B(x,r) = \{y \in M : d(x,y) \leq r\}$, onde B é um conjunto aberto que possibilita a geração de uma topologia em M , logo, um espaço

topológico. Variando d , é possível a obtenção de topologias não euclidianas e direcionadas à solução do problema em questão.

Espaços topológicos ou mapas topológicos são estruturas matemáticas que permitem uma definição formal de conceitos de convergência, continuidade e conectividade. Estes espaços consideram que subconjuntos podem estar inseridos no espaço, alterando a sua topologia de acordo com critérios de localização, agrupamento ou relevância de cada conjunto. As distâncias computadas neste tipo de espaço podem ser alteradas de acordo com a topologia, permitindo uma descrição mais precisa de similaridade.

3.6.2 Distância entre coordenadas no espaço

Em PDI e VC frequentemente é utilizada a álgebra linear para a representação de cores em espaços discretos. As funções de similaridade geralmente computam a similaridade entre duas coordenadas x e y , que são ou coordenadas reais ou coordenadas médias de populações multivariadas dispersas no espaço.

A norma vetorial é a forma mais simples de se quantificar a similaridade entre componentes vetoriais m -dimensionais. A norma vetorial verifica a distância entre pontos através de uma reta, e pode ser provada repetidamente pela aplicação do teorema de Pitágoras. Outras métricas variantes da distância euclidiana também podem ser aplicadas em um espaço euclidiano obtendo-se diferentes resultados. Em um espaço euclidiano \mathfrak{R}^m , a distância ou similaridade entre duas coordenadas $x(x_1, x_2, \dots, x_m)$ e $y(y_1, y_2, \dots, y_m)$ m -dimensionais podem ser computadas por:

1. $|x - y| = \sum_{i=1}^m |x_i - y_i|$, L_1 -norm ou distância de *Manhattan*, uma vez que utilizando somente valores absolutos a distância é quantificada univariadamente para cada coordenada em 4 conexos (não diagonais);
2. $\|x - y\| = \sum_{i=1}^m (|x_i - y_i|^2)^{1/2}$, L_1 -norm ou distância euclidianas propriamente dita, a menor distância entre 2 pontos em uma linha reta, conforme notação do teorema de Pitágoras;
3. $(\sum_{i=1}^m \frac{|x_i - y_i|^2}{\sigma_i^2})^{1/2}$, distância euclidiana escalada em relação a uma componente vetorial σ_i ;
4. $(\sum_{i=1}^m |x_i - y_i|^p)^{1/p}$, p -norm, similar à norma vetorial, onde p geralmente é determinado pela dimensão m ;
5. $\lim_{k \rightarrow \infty} (\sum_{i=1}^m |x_i - y_i|^k)^{1/k}$ denominada de L_∞ -norm e popularmente conhecida como distância no tabuleiro de xadrez, devido à similaridade

dade de número de movimentos que um rei deve fazer para percorrer as casas. Também pode ser analogamente representada pela distância de Manhattan utilizando 8 conexos (4 conexos + diagonais).

3.6.3 Distância em distribuições multivariadas

Em espaços métricos, subconjuntos vetoriais podem ser inseridos para alterar a topologia de um espaço métrico. Estes subconjuntos podem ser na forma de distribuições multivariadas que devem estar inseridas nas dimensões do espaço para a modelagem de uma topologia. As formas de modelagem destes mapas podem ser efetuadas de diversas maneiras, tanto por métricas globais como locais. Exemplo de uma métrica global é a distância de Mahalanobis (DUDA; HART; STORK, 2001). A distância de Mahalanobis é uma métrica de distância baseada na correlação estatística entre as componentes vetoriais. Por ser baseada na covariância de um dado conjunto de entrada, é invariante a escala e tamanho n do conjunto de entrada S . Dadas duas coordenadas vetoriais x e y , e um conjunto de entrada S correspondendo a uma distribuição multivariada no espaço, a distância de Mahalanobis é definida pela seguinte equação 3.4:

$$D_M(x) = [(x - \mu_x)^T \cdot \Sigma^{-1} \cdot (y - \mu_y)]^{1/2}, \quad (3.3)$$

onde μ é o vetor médio da distribuição S e Σ^{-1} é a inversa da matriz de covariância, a qual é obtida de S . Caso Σ seja uma matriz identidade, a distância de Mahalanobis reduz-se à distância euclidiana. Uma das vantagens da distância de Mahalanobis em relação às métricas euclidianas é a descrição de uma topologia elíptica em relação ao padrão de entrada S . Os pontos destacados em azul denotam o conjunto S correspondendo a um determinado padrão no espaço.

3.7 POLINOMIAL MAHALANOBIS

A distância polinomial de Mahalanobis, é a combinação polinomial dos vetores de entrada e a projeção destes vetores sob uma perspectiva dimensional superior. O conceito principal é a computação da distância de Mahalanobis em uma ordem dimensional maior, permitindo assim capturar as características não-lineares de maior complexidade das distribuições multivariadas dos padrões de entrada.

A distância de Mahalanobis é definida pela Equação 3.4. Essa equação define a primeira ordem do polinomial Mahalanobis. Considere $S : p_1, p_2, p_3, \dots, p_N$

sendo um conjunto de vetores m -dimensionais, onde $p_i \in \mathfrak{R}^m$ corresponde a um vetor de entrada do conjunto de treinamento, e N é a cardinalidade de S . O primeiro passo é computar a distância de Mahalanobis entre dois vetores de cores x e y , e então utilizar o conjunto de entrada S no cálculo da covariância (primeira ordem, ou a distância de Mahalanobis entre x e y). O segundo passo é calcular a segunda projeção polinomial (segunda ordem) através do mapeamento de todos os vetores m -dimensionais $pk = pk_1, pk_2, \dots, pk_m$ de S , para cada base $k = 1, \dots, N$, em todos os termos polinomiais de ordem q ou menor. Por exemplo, considerando um vetor bidimensional $pk = pk_1, pk_2$, o mapeamento deste vetor em sua segunda ordem polinomial é $(pk_1, pk_2, pk_1^2, pk_2^2, pk_1 pk_2)$. Uma forma simples de se obter a distância polinomial de Mahalanobis é pelo mapeamento direto de todos os dados de entrada de S em seus respectivos termos polinomiais em uma determinada q -ordem, e usá-los diretamente na Equação 3.4. Porém, o custo computacional para o cálculo de métricas de distância com polinômios de grande ordem é proibitivo. Uma maneira de se calcular a distância utilizando polinomiais de grande ordem foi proposta por (GRUDIC; MULLIGAN, 2006) e utiliza a seguinte equação modificada:

$$D_M(x, y) = dM_{\sigma^2}(x, y) + \sum_{l=1}^{L-1} dM_{\sigma^2}(g_l^i, g_l^j), \quad (3.4)$$

onde dM_{σ^2} é a distância de Mahalanobis computada com um pequeno escalar positivo σ^2 usado para anular limitações de inversão de matrizes se algum valor singular for zero, $L > 0$ é o número máximo de termos polinomiais de q -ordem ($q = 2^L$), e os argumentos g_l^i e g_l^j são as próximas projeções dos vetores x e y em seus termos polinomiais conforme segue:

1. A decomposição do valor singular (SVD) do padrão S é obtida e então x e y são projetados em k vetores dimensionais $(x_1, \dots, x_k), (y_1, \dots, y_k)$ em espaços polinomiais de segunda ordem:

$$\begin{aligned} g_l^i &= (x_1, \dots, x_k, x_1^2, \dots, x_k^2, x_1 x_2, x_1 x_3, \dots), \\ g_l^j &= (y_1, \dots, y_k, y_1^2, \dots, y_k^2, y_1 y_2, y_1 y_3, \dots); \end{aligned}$$

2. Da mesma forma, os pontos $pk = p_{k1}, p_{k2}, \dots, p_{km}$ em S , para $k = 1, \dots, N$ são projetados neste espaço polinomial g_1 , onde a matriz de covariância é calculada de $g_1^1, g_1^2, \dots, g_1^k$. Logo, os valores singulares denotados por w_1, w_2, \dots, w_m de g_1^k são obtidos, e $d_M \sigma^2(g_l^i, g_l^j)$ pode então ser calculada para a segunda ordem polinomial.

As próximas ordens polinomiais são da mesma forma construídas pela re-projeção destes termos g_l^i e g_l^j em uma nova segunda ordem polinomial. De

maneira projetiva reprojeta-se a segunda ordem polinomial na sua própria segunda ordem para se obter a quarta ordem. Da re-projeção da quarta ordem, se obtém a oitava ordem, e assim sucessivamente. Este mapeamento progressivo em espaços polinomiais de ordem superior continua até que a q-ordem seja igual a $2^{(L-1)}$.

3.8 TEXTURA

Não existe uma única definição de textura, e essa indefinição pode ser observada ao se constatar a grande quantidade de definições encontrada na literatura. Pode-se dizer que textura é uma repetição de padrões que caracterizam uma área ou objeto e esses padrões não são sempre uniformes, podendo variar quanto ao tamanho, forma, cor e orientação/inclinação em uma região (PEDRINI; SCHWARTZ, 2008). A textura, juntamente com as informações espectrais e contextuais são os três elementos fundamentais utilizados por observadores humanos para a interpretação de imagens coloridas (HARALICK; SHANMUGAM; DINSTEIN, 1973).

Diferentemente da característica espectral, que descreve a variação tonal média em várias bandas das porções visíveis e/ou infravermelha de um espectro eletromagnético, a textura contém ou sintetiza informações sobre a distribuição espacial das variações dos tons em uma determinada banda espectral. O conceito de tom baseia-se na variação do valor de um canal específico em um pixel de uma imagem (HARALICK; SHANMUGAM; DINSTEIN, 1973). A textura preocupa-se com a distribuição espacial/estatística dos tons de cada canal.

Embora seja bastante fácil para observadores humanos reconhecer e descrever texturas de forma empírica, o reconhecimento dela por sistemas computacionais se apresenta desafiadora. Dado que as propriedades de textura das imagens fornecem informações valiosas para a classificação e distinção de objetos ou superfícies, é de extrema importância o estudo e desenvolvimento de técnicas que permitam esse tipo de classificação. Rao et al (RAO; LOHSE, 1993) mostraram que observadores humanos usam essencialmente três características para interpretação de textura, sendo elas: repetição, direcionalidade e complexidade. Repetição e direcionalidade são representadas através da frequência e da orientação. Complexidade tem relação com a consistência da textura. Por exemplo, uma textura puramente senoidal teria baixa complexidade enquanto que uma textura sem uma forma ou distribuição bem definida seria considerada de alta complexidade.

Tonalidade e textura estão sempre presentes em uma imagem, embora uma propriedade possa sobrepujar outra em alguns casos (HARALICK;

SHANMUGAM; DINSTEIN, 1973). Em uma pequena área de uma imagem, se é possível verificar uma pequena variação na tonalidade, temos como característica dominante nessa área a tonalidade. Caso nesta mesma a variação de tonalidade seja grande então temos a textura como característica dominante. O tamanho destas pequenas áreas, o número relativo de características discretas e características distinguíveis são cruciais na determinação da característica dominante da imagem(HARALICK; SHANMUGAM; DINSTEIN, 1973). Conforme diminui-se a área analisada e a variação da tonalidade, aumenta-se a prevalência da mesma sobre as características de textura. Se a área analisada é mínima (no caso de imagens essa área mínima é um pixel) a única característica presente é a tonalidade. De forma contrária, conforme aumenta a variação da tonalidade em uma área da imagem, aumenta-se também a prevalência da textura sobre a tonalidade(HARALICK; SHANMUGAM; DINSTEIN, 1973).

3.8.1 Matriz de Co-ocorrência

Haralick et al (HARALICK; SHANMUGAM; DINSTEIN, 1973) assumiram que a informação do contexto da textura em uma imagem I está contida no relacionamento geral ou médio entre as tonalidades de cinza dessa imagem. Mais especificamente assume-se que a informação do contexto de textura é especificada adequadamente pela matriz de frequências relativas $P_{i,j}$ em que dois vizinhos separados por uma distância d e um ângulo α ocorrem na imagem, um com o tom de cinza i e o outro com o tom de cinza j . Essas matrizes representam o relacionamento angular e a distância entre dois pixels vizinhos.

A matriz do co-ocorrência informa o número de vezes em que valores de cinza x e y ocorrem, separados por uma distância d e um ângulo α . A matriz mapeia as ocorrências para todas as combinações de valores de cinza encontrados na imagem. Se uma imagem possui n valores de pixel, então a matriz de covariância será da forma $n \times n$. A matriz de co-ocorrência é calculada para os valores de α igual a 0° , 45° , 90° e 135° e para um valor de distância d fixo igual a 1. Temos assim um total de quatro matrizes de co-ocorrência para uma imagem. Matematicamente esse relacionamento é representado por:

$$1. P(i, j, d, 0) = \#(k, l), (m, n) \in N | k - m = 0, |l - n| = 0$$

$$2. P(i, j, d, 45) = \#(k, l), (m, n) \in N | k - m = d, l - n = -d$$

$$3. P(i, j, d, 90) = \#(k, l), (m, n) \in N \mid |k - m| = 0, l - n = 0$$

$$4. P(i, j, d, 135) = \#(k, l), (m, n) \in N \mid k - m = -d, l - n = d$$

onde # representa o número de elementos no conjunto, $k, l, m, n \in G$ e $I(k, l) = i$ e $I(m, n) = j$ e I representa uma imagem de tamanho $N \times N$ e $I(x, y : 0 \leq x \leq N - 1, 0 \leq y \leq N - 1)$ é a representação de uma imagem de tamanho $N \times N$.

Para melhor ilustrar a técnica vamos utilizar uma imagem de tamanho 4×4 com 4 níveis de cinza como exemplo. A imagem utilizada é representada segundo a matriz abaixo:

$$I(x, y) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 2 \\ 3 & 3 & 2 & 2 \end{pmatrix}$$

A partir da imagem acima, calcula-se o número de vezes em que uma transição entre dois pixels vizinhos, separados por uma distância d e ângulo α , ocorrem. Na imagem de exemplo acima existem quatro ocorrências de pixels com valor igual a 1, separados por uma distância $d = 1$ e com ângulo $\alpha = 0$. Matrizes de co-ocorrência são simétricas por natureza, isto é, $P(i, j; d; \alpha) = P(j, i; d; \alpha)$, sendo assim, para cada transação encontrada soma-se dois ao total de transações. Com os resultados para todos os pixels com os quatro ângulos previamente definidos e com distância fixa $d = 1$, temos as seguintes matrizes de covariância:

$$P(d = 1, \alpha = 0^\circ) =$$

	0	1	2	3
0	4	2	1	0
1	2	4	0	0
2	1	0	6	1
3	0	0	1	2

$$P(d = 1, \alpha = 90^\circ) =$$

	0	1	2	3
0	4	1	2	1
1	1	4	1	0
2	2	1	4	1
3	1	0	1	0

$$P(d = 1, \alpha = 45^\circ) =$$

	0	1	2	3
0	2	2	2	0
1	2	2	0	0
2	2	0	2	2
3	0	0	2	0

$$P(d = 1, \alpha = 135^\circ) =$$

	0	1	2	3
0	2	1	1	1
1	1	2	2	0
2	1	2	4	0
3	1	0	0	0

Após o cálculo das matrizes, normaliza-se os valores da mesma dividindo cada valor encontrado na matriz pela soma total de todos os valores. Haralick et al. propuseram 14 medidas estatísticas calculadas a partir da matriz de co-ocorrência; oito dessas medidas são listadas abaixo:

1. *Media* = $\sum_i \sum_j P(i, j) * i$
2. *Variância* = $\sum_i \sum_j P(i, j) * (i - \mu)^2$
3. *Homogeneidade* = $\sum_i \sum_j \frac{P(i, j)}{1 + |i - j|}$
4. *Contraste* = $\sum_i \sum_j P(i, j) * (i - j)^2$
5. *Entropia* = $\sum_i \sum_j - P(i, j) * \log_e P(i, j)$
6. *Segundomomentoangular* = $\sum_i \sum_j P(i, j)^2$
7. *Correlação* = $\sum_i \sum_j (i - \mu_x) * (j - \mu_y) * P(i, j)$
8. *Dissimilaridade* = $\sum_i \sum_j \frac{P(i, j)}{1 + |i - j|} * P(i, j)$

$$\text{onde } \mu = \frac{\sum P(i, j)}{n}, \mu_x = \frac{\sum P_j(i, j)}{n} \text{ e } \mu_y = \frac{\sum P_j(i, j)}{n}.$$

Cada medida de textura derivada da matriz de co-ocorrência é interpretada de forma única. Os indicadores média, homogeneidade e segundo momento angular são medidas de similaridade dos valores do pixel em um bloco da imagem (HARALICK; SHANMUGAM; DINSTEN, 1973). O contraste é uma medida de variação local presente na imagem. A correlação é

uma medida de previsibilidade dos valores do pixel nos domínios horizontal e vertical. A entropia pode ser descrita como uma medida de complexidade ou a medida de informação da imagem.

Embora seja amplamente utilizada em análise de texturas, a matriz de co-ocorrência apresenta alguns problemas, tais como a inexistência de um método para seleção da distância d entre os pixels, embora possa ser reduzido por meio do cálculo de matrizes para valores distintos de d , o que acarreta aumento do custo computacional (PEDRINI; SCHWARTZ, 2008). Outra desvantagem é o fato dessas matrizes não capturarem os aspectos da forma das primitivas contidas na textura. (PEDRINI; SCHWARTZ, 2008)

3.8.2 Máscaras de Laws

As Máscaras de Laws são um conjunto de máscaras de convolução usadas para o cálculo de indicadores de textura utilizando um conjunto de operadores de gradiente. Esses operadores de gradientes, também chamados de operadores lineares locais são baseados na combinação dos resultados das derivadas de primeira e segunda ordem de um bloco de imagem. Um conjunto de 25 máscaras de convolução é possível, utilizando combinações de um dos 5 vetores unidimensionais sugeridos por Laws. A Tabela 1 mostra os vetores unidimensionais. A variância ou o valor absoluto de uma convolução realizada em cada pixel de uma área da imagem é chamada de energia de textura.

<i>Level(L5)</i>	[1 4 6 4 1]
<i>Edge(E5)</i>	[-1 -2 0 2 1]
<i>Spot(S5)</i>	[-1 0 2 0 -1]
<i>Wave(W5)</i>	[-1 2 0 -2 1]
<i>Ripple(R5)</i>	[1 -4 -6 -4 -1]

Tabela 1 – Vetores unidimensionais utilizados

O vetor *Level(L5)* fornece uma média local ponderada e central. *Edge(E5)* é similar aos operadores de gradiente. *Spot(S5)* é baseado na segunda derivada e é similar à aplicação de uma Laplaciana sobre uma Gaussiana. *Wave(W5)* responde a pequenas mudanças na intensidade dos pixels em uma imagem e *Ripple(R5)* é usado para detectar ondulações em uma imagem. Essas máscaras são geradas através da multiplicação entre dois vetores unidimensionais quaisquer. Essa multiplicação é realizada entre um vetor vertical e um vetor horizontal, gerando assim as máscaras. As máscaras mais comuns encontram-se

abaixo.

$$\text{E5-E5} \quad \begin{pmatrix} 1 & 2 & 0 & -2 & -1 \\ 2 & 4 & 0 & -4 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -4 & 0 & 4 & 2 \\ -1 & -2 & 0 & 2 & 1 \end{pmatrix}$$

$$\text{R5-R5} \quad \begin{pmatrix} 1 & -4 & 6 & -4 & 1 \\ -4 & -16 & -24 & 16 & -4 \\ 6 & -24 & 36 & -24 & 6 \\ -4 & 16 & -24 & 16 & -4 \\ 1 & -4 & 6 & -4 & 1 \end{pmatrix}$$

$$\text{E5-S5} \quad \begin{pmatrix} -1 & 0 & 2 & 0 & -1 \\ -2 & 0 & 4 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -0 & 4 & 0 & -2 \\ -1 & 0 & 2 & 0 & 1 \end{pmatrix}$$

$$\text{L5.E5} \quad \begin{pmatrix} -1 & -2 & 0 & 2 & 1 \\ -4 & -8 & 0 & 8 & 4 \\ -6 & -12 & 0 & 12 & 6 \\ -4 & -8 & 0 & 8 & 4 \\ -1 & -2 & 0 & 2 & 1 \end{pmatrix}$$

$$\text{L5.S5} \quad \begin{pmatrix} -1 & 0 & 2 & 0 & -1 \\ -4 & 0 & 8 & 0 & -4 \\ -6 & 0 & 12 & 0 & -6 \\ -4 & 0 & 8 & 0 & -4 \\ -1 & 0 & 2 & 0 & -1 \end{pmatrix}$$

O resultado da convolução dessas máscaras e cada pixel da imagem é a de energia de textura do pixel. Essa energia é definida pela seguinte equação:

$$E_k(i, j) = I(i, j) * H_k(i, j) \quad (3.5)$$

onde $I(i, j)$ é o valor do pixel na posição i e j e $H(i, j)$ é o resultado da convolução de umas das máscaras escolhidas com a região da imagem cujo centro é na posição i e j . O valor da energia de textura geral E de uma imagem é dada pela seguinte equação:

$$E_k(i, j) = \frac{1}{P_k \cdot N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} E_k^2(i, j) \quad (3.6)$$

$$P_k = \sum_{i,j} H_k(i, j)^2 \quad (3.7)$$

Utilizando k máscaras de convolução diferentes obteremos K valores de energia de textura que caracterizam a textura da imagem.

3.8.3 Filtro de Gabor

Filtros de Gabor são detectores locais sensíveis à orientação. A sua forma de funcionamento consiste em realizar uma série de operações de convolução entre uma imagem e um conjunto de máscaras, chamadas de Gabores. Banco de Gabor é o nome dado ao resultado da convolução entre a imagem e diversas dessas máscaras. Esses Gabores são projetados com a finalidade de detectar certas frequências e certas orientações. O filtro de Gabor pode ser implementado para utilizar dois Gabores, um para detectar a parte real e outro para detectar a parte imaginária, porém é muito comum a sua utilização somente utilizando os Gabores para a parte real.

A análise de textura utilizando o filtro de Gabor é possível graças à capacidade que ele tem de modelar as frequências e orientações sensíveis ao sistema visual humano. O filtro de Gabor possui características que o tornam atraente para aplicações na área de visão computacional. Entre essas características incluem-se a sua simplicidade e a habilidade de simular o comportamento dos campos perceptivos do córtex visual humano, isolando frequências e orientações específicas (DAUGMAN, 1985).

O banco de filtros de Gabor é composto frequentemente por Gabores criados com 4 comprimentos de onda e frequências diferentes e com uma quantidade x de ângulos que normalmente variam entre 0° e 180° . A razão pela qual os ângulos cobrem somente metade do círculo deve-se ao fato de que o valor absoluto da resposta é o mesmo para ângulos simétricos, somente o valor da fase seria rotacionado em 180° . A resposta de cada filtragem é armazenada e é utilizada posteriormente como base para comparação. O filtro de Gabor é basicamente uma Gaussiana (com variâncias s_x e s_y ao longo dos eixos x e y respectivamente) modulada por uma senoide complexa e é definida pela fórmula 3.8.

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos 2\pi\left(\frac{x'}{\lambda} + \psi\right) \quad (3.8)$$

onde:

$x' = x \cos \theta + y \sin \theta$ e $y' = -x \sin \theta + y \cos \theta$ e λ é o valor do comprimento de onda, θ é o ângulo de inclinação das ondas paralelas do filtro, determina o desvio padrão da distribuição normal, e está relacionado com a largura da Gaussiana que modula o filtro, ψ determina o tamanho da janela do núcleo e determina a excentricidade do núcleo. A equação 3.8 gera uma função senoidal modelada por uma função Gaussiana e os valores x' e y' rocionam a equação de Gabor de acordo com o valor do ângulo θ .

4 MATERIAIS E MÉTODO

4.1 BANCO DE IMAGENS

De forma a facilitar a fase de testes do algoritmo, optou-se por utilizar fotos aéreas previamente ortoretificadas e com alta resolução. O Instituto de Águas do Paraná (Instituto de Aguas do Parana,) dispõe de um acervo com 141 imagens coloridas da Bacia do Alto Iguazu com essas características. As fotos foram escaneadas de um filme aéreo cuja resolução é de 28 micras utilizando um escâner de alta resolução. Para a geração destas fotos foram utilizados 327 modelos; entre esses encontram-se os modelos digitais do terreno definido pela base cartográfica em escala 1:10.000. Desse total de 141 imagens, 6 foram selecionadas de forma aleatória para o banco de imagens utilizado neste trabalho.

Apesar de possuírem uma alta resolução, as imagens encontram-se no formato JPEG e o mesmo apresenta ruído, principalmente em áreas de transição, como nas regiões de borda, devido ao algoritmo de compressão utilizado pelo padrão, gerando dados inconsistentes quando se opera em nível de pixel. De forma a evitar este tipo de problema optou-se por converter as imagens para o formato PNG sem compactação. As imagens no formato PNG apresentam uma transição mais suave nas bordas, diminuindo assim a quantidade de ruído nas imagens. Devido ao tamanho excessivo das imagens, fazendo com que houvesse falta de memória na máquina utilizada para executar os testes, optou-se por fazer uma redução no tamanho das mesmas. As imagens originais podem ser copiadas diretamente da página do Instituto das Águas do Paraná (Instituto de Aguas do Parana,).

4.2 TREINAMENTO

A fase de treinamento constitui-se como uma das mais importantes dentro do processo de classificação de imagens. Por se tratar de uma abordagem supervisionada, onde o usuário seleciona pontos pertencentes a uma classe qualquer, é de extrema importância que a ferramenta utilizada seja precisa e confiável. Dessa forma desenvolveu-se uma ferramenta de extração de características da imagem. O funcionamento da mesma se dá de forma bastante simples. O usuário seleciona pixels na imagem utilizando o *mouse* do computador e para cada pixel selecionado a ferramenta armazena as suas informações de cor R, G e B, assim como os dados de textura. A ferramenta

permite ao usuário selecionar os atributos de configuração do algoritmo. Esses atributos são listados abaixo:

1. número de classes: Essa variável determina quantas amostras o usuário deve coletar para realizar o treinamento;
2. ordem do polinômio: define a ordem do polinômio;
3. contraste: ;
4. espaço de cores: permite utilizar o espaço de cores RGB ou HSV;
5. tipos de textura: permite determinar se informações de texturas deveriam ser coletadas e também selecionar combinações de *kernels* específicos.

A Figura 6 mostra a interface do programa.



Figura 6 – Interface da ferramenta

À partir das superfícies identificadas nas imagens aéreas, tais como: floresta, pasto, rios e lagos, plantações e terrenos descampados, optou-se pela utilização de floresta, plantação, rios e lagos. A partir dessas superfícies criaram-se as classes que foram utilizadas para a classificação. A Tabela 2 mostra as classes escolhidas e a sua descrição.

As áreas pertencentes à classe floresta (classe 1) apresentam como características principais a predominância da cor verde e a presença também

Tabela 2 – Definição das superfícies classificadas

Número	Classes	Definição
1	Floresta	Vegetação natural e reflorestamento
2	Plantação	Terras irrigadas, área de plantação
3	Água	Lagos e rios

de textura, intercalando pixels de cor verde com pixels de cor preta, representando sombra ou falta de luminosidade nas copas das árvores. Um exemplo desse tipo de classe é mostrado na Figura 7. A classe plantação (classe 2) possui a cor bege como característica dominante e apresenta formato mais geométrico e textura variável. Devido à ausência de um especialista e também de material de apoio no banco de imagens que pudesse identificar de forma mais precisa qual o tipo de cultura presente na imagem (por exemplo: feijão, trigo, milho, arroz), considerou-se nesse trabalho áreas com essa característica como sendo da classe plantação. Um exemplo dessa classe é verificado na Figura 8. A classe água (classe 3) abriga todos os tipos de recursos hídricos encontrados nas imagens. Sempre que um rio, curso d'água ou qualquer outro recurso hídrico fosse encontrado na imagem analisada, amostras deles eram coletadas. Uma amostra dessa classe é verificada na Figura 9.

Para classificadores supervisionados, tanto o tamanho quanto a qualidade dos conjuntos de treinamento e de teste são determinantes para a qualidade do resultado final (THAPA; MURAYAMA, 2008). Dessa forma, para cada imagem analisada criou-se um conjunto de treinamento. Esse conjunto de treinamento era formado por n grupos de pixels, sendo n o número de classes presentes nessa imagem. Para cada classe pertencente a uma imagem, pixels eram selecionados utilizando a ferramenta descrita anteriormente. O número de uma amostra não poderia ser inferior a 50 pixels e não havia limite superior, contudo, nenhuma amostra possuía mais do que 2000 pixels. Testes previamente realizados com conjuntos de treinamento muito grandes (acima de 2000 pixels) não mostraram melhora no resultado final e aumentaram consideravelmente o tempo de treinamento. Para cada pixel, eram coletados os valores de R, G e B, assim como os seus valores de textura nos ângulos 0° , 45° , 90° , 135° para cada espaço de cor. O tamanho do conjunto de treinamento, representado aqui pelo número de pixels selecionados em cada imagem é mostrado na Tabela 3. Esses valores de textura foram calculados através da convolução de um *kernel* de tamanho 3×3 criado a partir da equação de Gabor. Optou-se pelo tamanho 3×3 por se tratar do menor *kernel* possível, permitindo assim o refinamento dos resultados ao aumentar as



Figura 7 – Exemplo de seleção de amostras da classe Recursos Naturais

dimensões do mesmo.

Após a criação do conjunto de treinamento partiu-se para a criação do espaço multidimensional. De forma a determinar quais parâmetros de configuração apresentam os melhores resultados, um conjunto de pré-testes foram realizados para a determinação da configuração dos parâmetros. Estes testes foram realizados de forma automática em uma imagem da base de dados escolhida aleatoriamente. As ordens do polinômio utilizadas foram: 1, 2, 3, 4, e 5 e os valores de contraste variaram entre 0,01 e 5. Após a realização do pré-teste os resultados fornecidos pelos mesmos foram analisados com o objetivo de determinar qual a melhor configuração para uma imagem. Devido a alta similaridade das imagens da base de dados, supôs-se que a configuração que apresentou os melhores resultados no pré-teste produziria também o melhor resultado nas imagens da base de dados. Sendo assim, os parâmetros que apresentaram o melhor resultado no pré-teste foram: ordem do polinômio igual a três para as classes Floresta e Água e ordem quatro para a classe Plantação. Os melhores valores de contraste c variaram entre 0.9 e 1.1, optando-se pelo valor fixo de 1.0 para a realização dos testes. Tendo os parâmetros e o conjunto de treinamento definidos criou-se então o espaço



Figura 8 – Exemplo de seleção de amostras da classe Plantações

Tabela 3 – Tamanho dos conjuntos de treinamento de cada imagem

Imagem	Amostras Floresta	Amostra Plantação	Amostra Água
c62neb	1023	1093	1031
c62nee	587	541	196
c62nof	1040	958	992
d13nof	974	928	não presente
d13sof	1010	953	1009
d41soe	968	-	não presente

multi-dimensional para a classificação das imagens.

4.2.1 Classificação

Seguida da fase de criação dos espaços multi-dimensionais passou-se à fase de classificação propriamente dita. Nesta, a distância $dist$ entre cada pixel de uma imagem específica e os espaços multi-dimensionais de cada classe presente na imagem era calculada. Esse valor de distância foi então transformado em um valor de probabilidade P utilizando a Equação 4.1:

$$P = \exp(-1 * c * dist) * 255 \quad (4.1)$$



Figura 9 – Exemplo de seleção de amostras da classe Recursos Hídricos

onde c é o valor de contraste, $dist$ é o valor da distância entre o pixel analisado e o espaço multi-dimensional que representa a classe sendo analisada e 255 é o valor máximo assumido por um canal de cor específico do pixel, considerando que todas as imagens possuem canais de cor com 8 bits de resolução. Quanto menor for o valor da distância, maior será o valor de probabilidade do pixel pertencer a classe analisada. Após percorrer toda a imagem tem-se como resultado final um mapa de probabilidades do espaço polinomial analisado. Esse mapa é representado na forma de uma imagem em tons de cinza. A Figura 10 mostra o mapa de probabilidades para uma classe do tipo 2 (plantação). Todos os mapas de probabilidades passaram por um processo de limiarização (GONZALEZ; WOODS, 2001), ou seja, cada pixel deste mapa teve o seu valor analisado, caso esse valor fosse menor do que 153 (valor do limiar utilizado), então esse pixel teria o seu valor atribuído para zero, caso contrário, o valor permanecia inalterado. O valor de 153 foi escolhido entre tantos outros testados. Esse valor representa uma probabilidade igual a 60%, sendo assim, todo pixel cujo valor fosse maior do que 153 teria uma probabilidade maior ou igual a 60% de pertencer a classe avaliada.

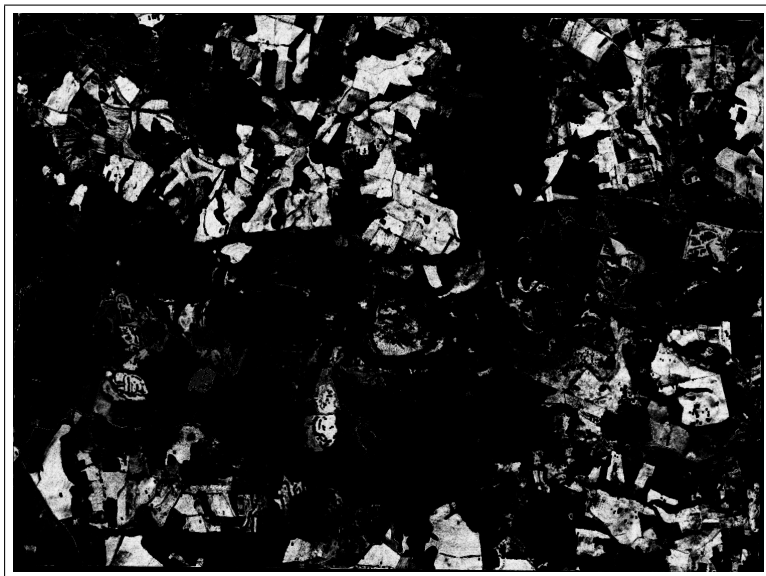


Figura 10 – Mapa de probabilidades para uma classe tipo 2 (plantação)

5 RESULTADOS E DISCUSSÕES

O resultado final da classificação é mostrado através de um mapa temático. Nesse mapa cada pixel assume uma cor correspondente a classe ao qual o mesmo pertence. É possível verificar que os mapas apresentam um máximo de quatro cores, sendo que a cor verde representa a classe Floresta, a cor vermelha representa a classe Plantação, a cor azul representa a classe Água e a cor preta representa alguma outra classe que não foi considerada nesse experimento. Dentre as 6 imagens analisadas, 4 delas apresentavam a classe Água, 5 apresentaram a classe plantação e todas tinham presente a classe Floresta. Apesar de a imagem d41soe apresentar a classe Plantação em sua superfície, a mesma foi desconsiderada devido a problemas no padrão ouro dessa classe na imagem.

A figura 5 apresenta os resultados encontrados para a imagem c62nee. Ao analisar o resultado para a classe Floresta é possível verificar que o classificador conseguiu diferenciar de forma consistente áreas verdes e descampados de áreas de mata, principalmente perto do leito do rio. Esse resultado mostra que as métricas de textura auxiliam na diferenciação entre superfícies de cores similares e textura diferente. O resultado para a classe Plantação se mostrou menos eficiente que os demais. É possível verificar que o classificador considerou que uma grande área branca no centro da imagem pertence a essa classe, apesar dessa área possuir características de cor e textura similares a classe definida, a mesma não pertence a classe Plantação, conforme mostra o padrão ouro (figura f) dessa imagem. Mesmo alterando os valores de limiarização não foi possível remover a mesma do resultado final. É possível verificar que o classificador conseguiu encontrar praticamente todo o rio, conforme mostrado na figura d da figura 5. O classificador apresentou uma taxa de acerto relativamente constante em praticamente toda a extensão do rio, porém, na margem com uma área de floresta mais densa e na parte do rio que apresenta uma ilha entre as margens houve uma diminuição dessa taxa,. A região do rio onde essa ilha se encontra é mais rasa do que o restante do leito do rio, fazendo assim com que a cor da água sofre influência dessa baixa profundidade. Já a região da margem do rio com intensidade de floresta maior sofre influência do reflexo da mesma, causando assim uma variação na cor da superfície da água, causando uma diferença entre essa região e o restante do rio.

A imagem c62nof apresenta resultados similares aos encontrados na imagem c62nee, conforme mostrado na figura 10. O classificador foi capaz de diferenciar entre florestas e copas de árvores de áreas de pasto e descampados conforme mostrado na imagem b. A classificação de áreas pertencentes a

classe Plantação apresentou poucos erros. O classificador foi capaz de identificar de forma correta as áreas de plantação mais relevante da imagem, porém, cometeu alguns erros quando compara-se o resultado com o padrão ouro. A classe Água apresentou um resultado bastante fiel quando analisa-se a resposta somente da área do rio. Algumas superfícies de aspecto azulado e de pouca granularidade da textura foram consideradas como pertencente a classe Água. Não se sabe a natureza dessas superfícies, porém, as mesmas não fazem parte da classe Água, logo, essa foi uma classificação errônea. Outro erro de classificação para a classe Água foi causado pelo aspecto esverdeado e textura sem variações da superfície da água. Uma região de pasto bem específica, próxima ao leito do rio apresentou características muito semelhante as da superfície da água, causando assim um erro de classificação para essa classe.

É possível verificar em todos os resultados que o classificador se comporta de forma semelhante quando analisando superfícies das classes Floresta e Água. O classificador foi capaz de detectar os rios e lagos encontrados no grupo de imagens. Houveram alguns erros principalmente nas margens dos rios e lagos. Essas áreas sofrem influência tanto da baixa profundidade quanto do reflexo de objetos encontrados próximos à margem. A classe Água, apesar de ser a mais homogênea entre as classes analisadas é a classe que mais sofre influência do meio ao seu redor, tornando a classificação de toda a sua extensão desafiadora. A classe Floresta não apresentou um resultado melhor devido a diferenças encontradas entre os resultados e o padrão ouro. Essa diferença se caracteriza principalmente por regiões de sombra extensas encontradas nas copas das árvores. Enquanto que o observador humano considera essas regiões como pertencente a classe Floresta, o classificador fornece uma resposta muito baixa para essas áreas, dependendo de sua extensão, aumentando assim a taxa de erros para essas áreas. O classificador para a classe Plantação cometeu alguns erros, principalmente ao confundir regiões de descampados com áreas de plantação, conforme mostrado de forma mais clara na figura 5. É possível verificar também que o classificador identificou de forma correta as áreas de plantação de tamanho mais significativo, cometendo erros somente em pequenas áreas ao longo das imagens.

A acurácia dos mapas temáticos derivados de processos de classificação de imagens é usualmente analisada em estudos de sensoriamento remoto (THAPA; MURAYAMA, 2008). Acurácia é um termo para comparação de resultados encontrados pelos métodos automáticos de classificação e dados de referência geográfica que, assume-se, sejam verdadeiros ou que pertençam a uma padrão ouro (THAPA; MURAYAMA, 2008; MACROBERTS; F., 2012; AL-KOFAHI et al., 2012; BEREZOWSKI et al., 2012). Geralmente esse padrão ouro é definido por especialistas e é um procedimento onde um especi-

alista desloca-se até o local das imagens para validar o resultado da classificação. Os padrões ouro utilizados nesse trabalho foi realizado por especialistas das áreas de processamento de imagens e visão computacional. Esses padrões foram feitos de forma manual, onde o especialista, com o auxílio de uma ferramenta de edição foi pintando nas imagens as áreas pertencentes as classes utilizadas nesse trabalho. Um padrão foi realizado para cada imagem selecionada. Os padrões ouro utilizados nesse trabalho não foram validados com resultados de campo. A acurácia é, então, o percentual de amostras classificadas corretamente quando comparadas com o padrão ouro. Uma matriz de confusão usualmente é empregada para a criação dos mapas temáticos. A matriz informa a correspondência entre o resultado obtido e as imagens do padrão ouro e o resultado obtido pode ser visto na figura 4.

Tabela 4 – Acurácia dos resultados positivos. Comparativo entre o Polinomial Mahalanobis e o padrão ouro

Imagem	Acurácia Floresta	Acurácia Plantação	Acurácia Água
c62neb	81.05%	93.54%	91.90%
c62nee	85,79%	92.25%	99.11%
c62nof	84.90%	92.67%	98.66%
d13nof	86.69%	97.18%	não presente
d13sof	87.40%	97.63%	98.92%
d41soe	87.49%	-	não presente

Os resultados mostram que o classificador foi capaz de identificar de forma eficiente pixels nas imagens que pertencem as suas respectivas classes. Conforme esperado, o algoritmo foi mais eficiente quando classificava objetos pertencentes a classe Água. Com exceção da imagem c62neb, onde o algoritmo errou na margem com floresta mais densa e onde o mesmo é mais raso. No restante das imagens que continham objetos da classe Água, o mesmo cometeu alguns erros perto da margem classificou erroneamente alguns pixels no solo. A área de floresta apresentou um resultado pior devido as características das áreas de floresta nas imagens do padrão ouro. Um observador humano tende a considerar as copas das árvores, juntamente com as sombras mais densas encontradas nessas áreas como pertencentes a mesma área. Sendo o algoritmo um classificador menos flexível, o mesmo fornece uma baixa probabilidade para essas áreas de sombra, fazendo assim com que a sua taxa de erros aumente.

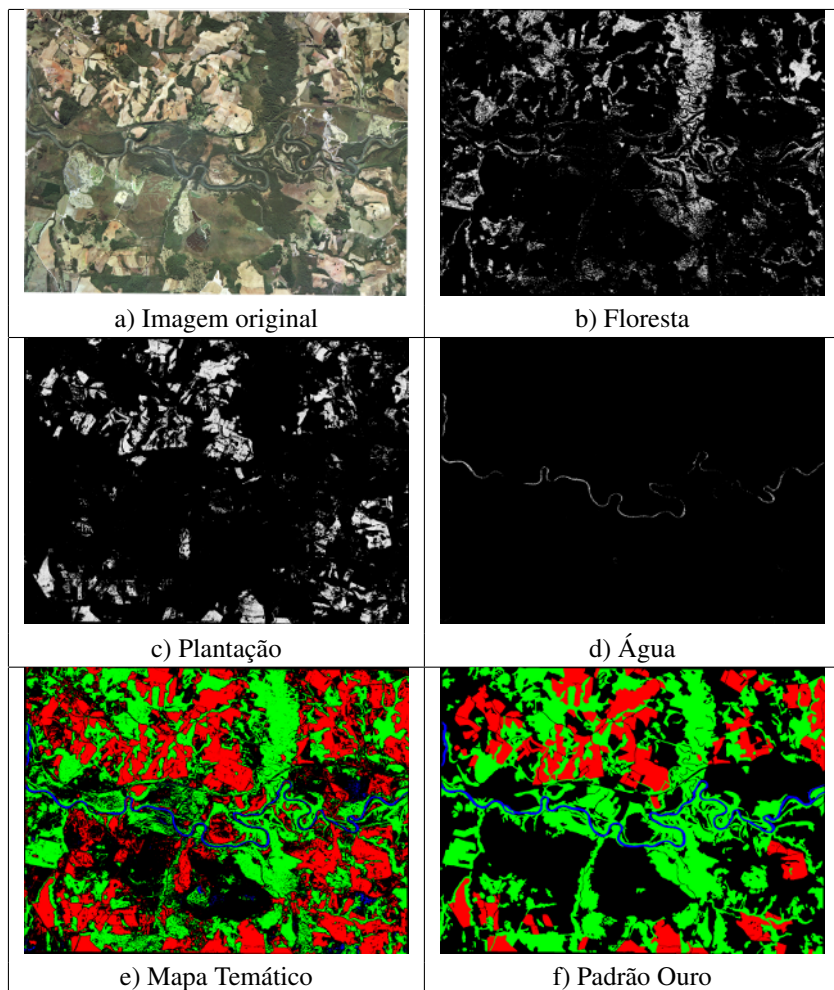


Figura 5 – Resultado dos testes realizados com a imagem c62nee.

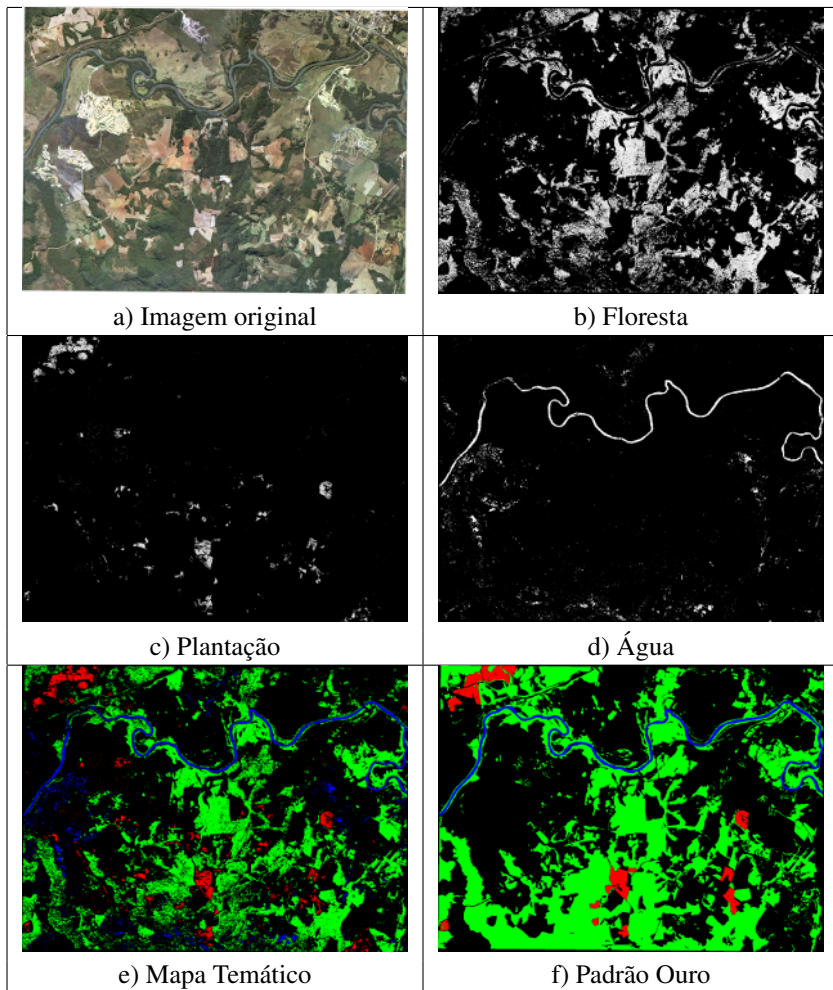


Figura 6 – Resultado dos testes realizados com a imagem c62nof.

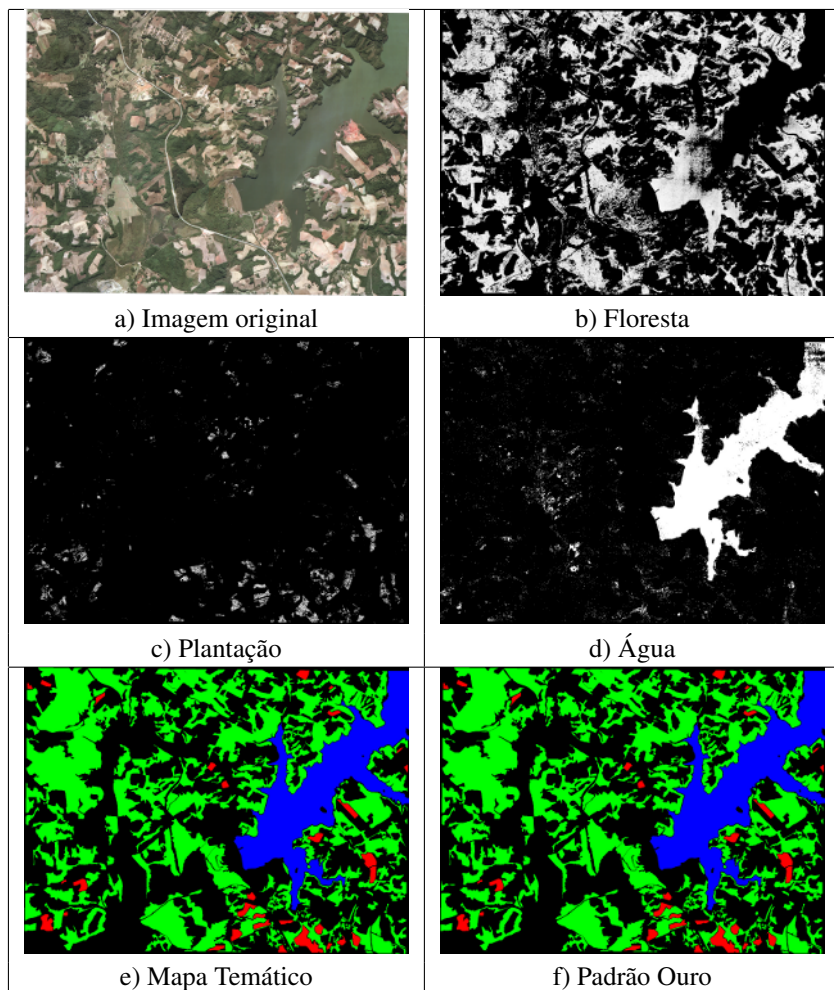


Figura 7 – Resultado dos testes realizados com a imagem c62neb.

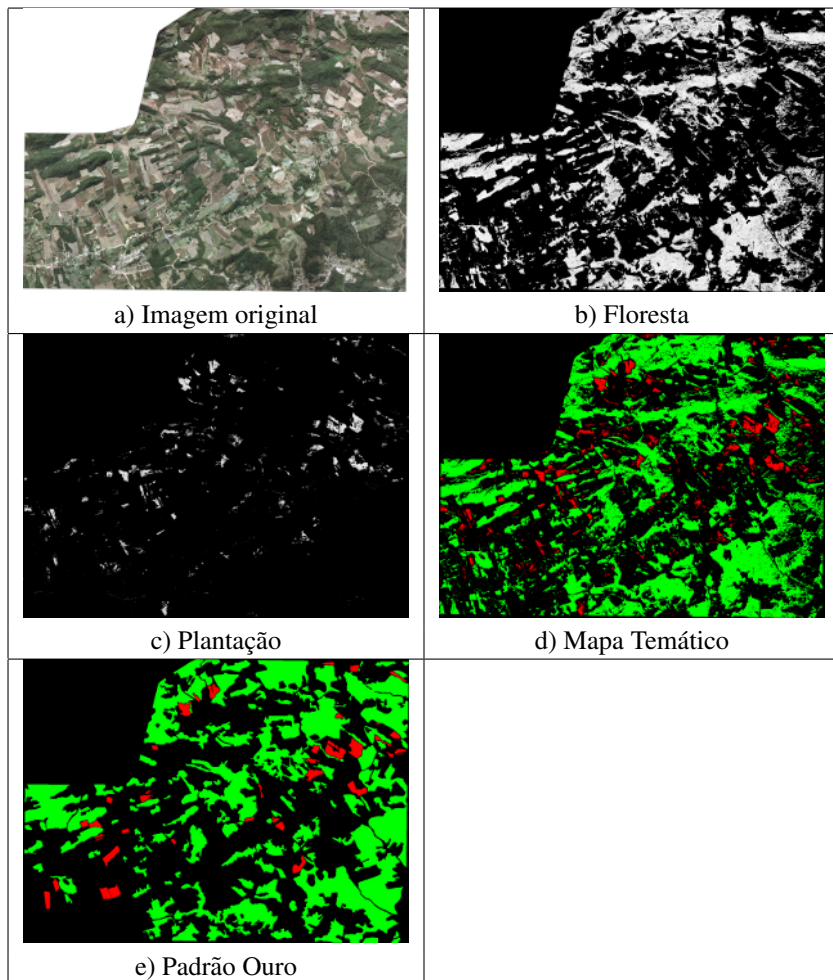


Figura 8 – Resultado dos testes realizados com a imagem d13nof.

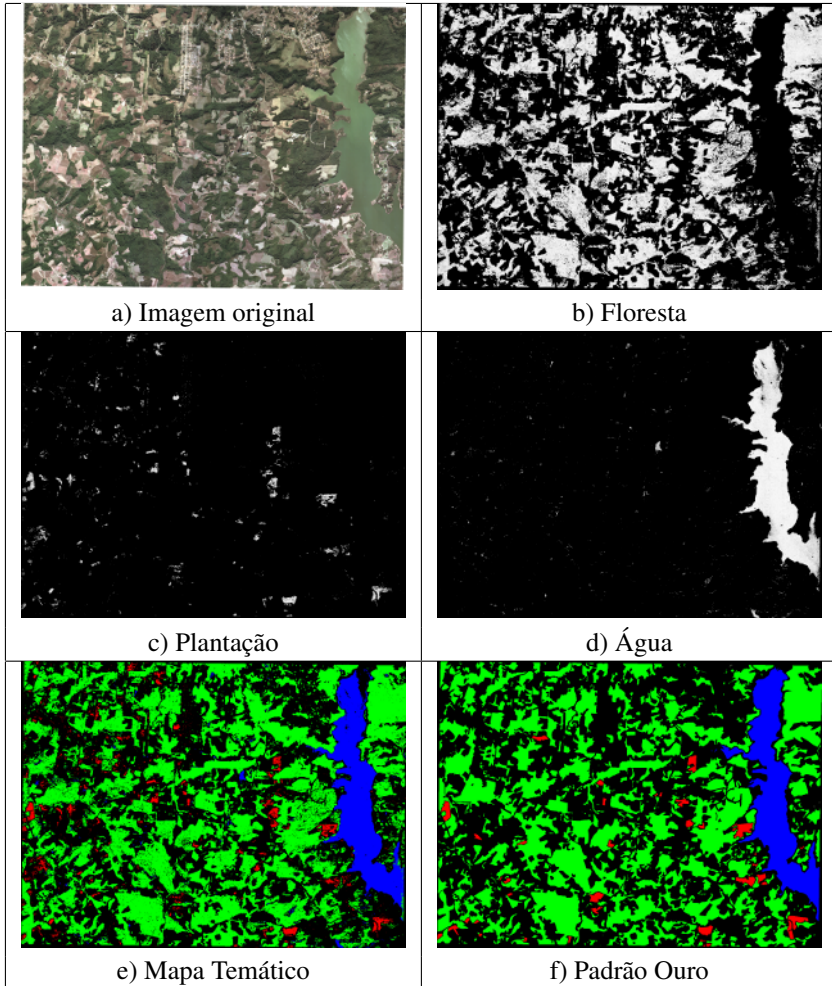


Figura 9 – Resultado dos testes realizados com a imagem d13sof.

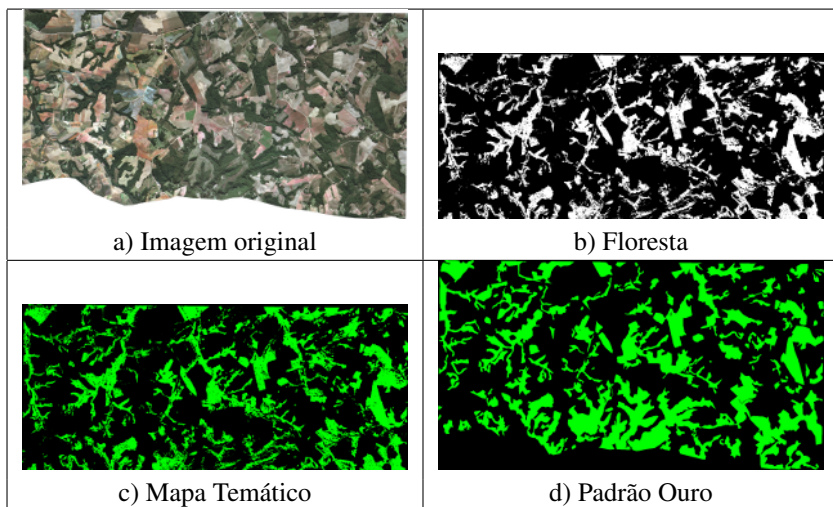


Figura 10 – Resultado dos testes realizados com a imagem d41soe.

6 CONSIDERAÇÕES FINAIS

A utilização de métricas locais de textura e cor no Polinomial Mahalanobis se mostrou promissora para a classificação de superfícies em imagens aéreas de SR. O classificador para a classe água apresentou o melhor resultado, provavelmente devido a sua maior homogeneidade superficial, tanto de cor quanto de textura. Alguns lagos apresentavam uma coloração semelhante às áreas de pasto e de mata, porém, o algoritmo conseguiu classificar de forma eficiente as instâncias dessas classes. O classificador para as classes Floresta e Plantação apresentaram menor precisão, porém, o resultado foi considerado satisfatório quando aplicado o teste de acurácia. O algoritmo conseguiu diferenciar regiões de pasto de regiões de floresta, mesmo tendo as duas regiões valores de cores bem semelhantes. Foi possível verificar que as informações de textura fizeram a diferença quando classificando-se classes com valores de cores similares.

Como desvantagem pode-se citar o tempo dispendido para a criação do descritor de classe. Esse tempo aumenta consideravelmente quando se aumenta a ordem do polinômio. O grande número de dimensões de cada pixel, quinze ao total, também impacta neste tempo. Os *kernels* utilizados se mostraram eficazes, possibilitando o refinamento de resultado quando comparado aos resultados encontrados somente utilizando informações de cor. Além disto a relação de tamanho dos *kernels* utilizados para o cálculo de textura influencia negativamente estes achados, devido ao seu tamanho fixo, em imagens com alta resolução não é possível obter informações de textura. Com esse tipo de imagem acredita-se que *kernels* de tamanhos maiores sejam necessários, porém essa informação é somente uma hipótese, sendo necessários mais testes para ser comprovada.

Por fim foi possível verificar o potencial de utilização da técnica desenvolvida e aplicada neste trabalho.

6.1 RECOMENDAÇÕES

De forma obter um padrão ouro mais preciso e assim obter uma análise dos resultados mais precisa, sugere que os padrões ouro sejam construídos por um especialista. A falta de um operador capaz de reconhecer as diferentes classes encontradas dentro da classe de plantações nas imagens tornou o experimento mais generalista, devido à incapacidade de reconhecer ou diferenciar essas classes previamente. Caso houvesse a participação de um operador com esse conhecimento específico, poder-se-ia determinar com mais clareza a real

eficiência do método e verificar se ele é capaz de classificar diferentes tipos de cultura, como, por exemplo, plantações de feijão, trigo, milho ou qualquer outra que fosse encontrada na base de imagens utilizada. Não sendo possível realizar um experimento com essas características seria imprudente afirmar que o método em questão é capaz de diferenciar e classificar corretamente essas culturas.

Caso estivessem disponíveis imagens de satélite de diferentes espectros para uma mesma área, seria possível utilizar essa informação adicional para possivelmente refinar os resultados encontrados ou até mesmo classificar diferentes tipos de árvores encontrada na zona de mata. Por exemplo, plantas de diferentes espécies sintetizam taxas diferentes de gás carbônico. Esta diferença pode ser observada através do espectro infravermelho. A inclusão de dados sobre o contexto, extraídas de levantamentos cartográficos, poderiam ser igualmente avaliadas. Sugere-se que testes utilizando esse tipo de informação sejam realizados.

Sugere-se que os seguinte tópicos sejam abordados no futuro:

1. utilização de imagens de infravermelho;
2. testes com tamanho de *kernel* variado;
3. comparação dos resultados com outras técnicas;

para fins de classificação em imagens de sensoriamento remoto.

REFERÊNCIAS BIBLIOGRÁFICAS

AL-KOFAHI, S. et al. Mapping land cover in urban residential landscapes using very high spatial resolution aerial photographs. *Urban Forestry and Urban Greening*, v. 11, n. 3, p. 291 – 301, 2012. ISSN 1618-8667. <<http://www.sciencedirect.com/science/article/pii/S161886671200060X>>.

BEREZOWSKI, T. et al. Impact of remotely sensed land-cover proportions on urban runoff prediction. *International Journal of Applied Earth Observation and Geoinformation*, v. 16, n. 0, p. 54 – 65, 2012. ISSN 0303-2434. <<http://www.sciencedirect.com/science/article/pii/S0303243411001723>>.

BOWKER, D. E. et al. *Spectral Reflectances of Natural Targets for Use in Remote Sensing Studies*. [S.l.]: NASA Reference Publication, 1985.

DAUGMAN, J. G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A, OSA*, v. 2, n. 7, p. 1160–1169, Jul 1985. <<http://josaa.osa.org/abstract.cfm?URI=josaa-2-7-1160>>.

DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification (2nd Edition)*. 2. ed. [S.l.]: Wiley-Interscience, 2001. Hardcover. ISBN 0471056693.

GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. 2nd. ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2001. ISBN 0201180758.

GRUDIC, G.; MULLIGAN, J. Outdoor path labeling using polynomial mahalanobis distance. In: *Proceedings of Robotics: Science and Systems*. Philadelphia, USA: [s.n.], 2006.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural Features for Image Classification. *Systems, Man and Cybernetics, IEEE Transactions on, IEEE, SMC-3*, n. 6, p. 610–621, nov. 1973. ISSN 0018-9472. <<http://dx.doi.org/10.1109/TSMC.1973.4309314>>.

Instituto de Aguas do Parana. <<http://www.aguasparana.pr.gov.br/>>. Acessado em 20 abril. 2012.

KRAMER, H. *Observation of the earth and its environment: survey of missions and sensors*. [S.l.]: Springer-Verlag, 1994.

LI, C.-H. et al. A spatial contextual support vector machine for remotely sensed image classification. *Geoscience and Remote Sensing, IEEE Transactions on*, n. 3, p. 784–799, march 2012. ISSN 0196-2892.

LILLESAND, T. M. *Remote Sensing and Image Interpretation*. [S.l.]: John Wiley & Sons, 2006. ISBN 0470088273.

MACROBERTS, R. E.; F., W. B. Statistical inference for remote sensing-based estimates of net deforestation. *Remote Sensing of Environment*, Elsevier, v. 124, p. 394–401, 2012.

MATHER, P. *Computer Processing of Remotely-Sensed Images: An Introduction*. [S.l.]: Wiley, 2005. ISBN 9780470021019.

MAYR, W.; HEIPKE, C. A contribution to digital orthophoto generation. *International Archives of Photogrammetry and Remote Sensing, Kyoto, Japan*, v. 27, p. IV430–IV439, 1998.

MOUSTAKIDIS, S. et al. Svm-based fuzzy decision trees for classification of high spatial resolution remote sensing images. *Geoscience and Remote Sensing, IEEE Transactions on*, v. 50, n. 1, p. 149–169, jan. 2012. ISSN 0196-2892.

NEWELL, A.; SIMON, H. A. *Human Problem Solving*. [S.l.]: Englewood Cliffs, 1972. 920 p.

PACIFICI, M. C. F.; EMERY, W. J. A neural network approach using multi-scale textural metrics from very high resolution panchromatic imagery for urban land-use classification. In: *Remote Sensing of Environment*. [S.l.: s.n.], 2009.

PEDRINI, H.; SCHWARTZ, W. *Análise de imagens digitais: princípios, algoritmos e aplicações*. [S.l.]: THOMSON PIONEIRA, 2008. ISBN 9788522105953.

RAO, A. R.; LOHSE, G. L. Identifying high level features of texture perception. *CVGIP: Graph. Models Image Process.*, Academic Press, Inc., Orlando, FL, USA, v. 55, n. 3, p. 218–233, maio 1993. ISSN 1049-9652. <<http://dx.doi.org/10.1006/cgip.1993.1016>>.

RUSSELL, S.; NORVIG, P. *Artificial Intelligence A modern Approach*. 3. ed. [S.l.]: Prentice Hall, 2010. 1132 p.

SCHOWENGERDT, R. A. *Remote Sensing: Models and Methods for Image Processing*. 3rd. ed. [S.l.]: Elsevier, 2007. 560 p. ISBN 978-0-12-369407-2.

SMITH, G. S. *Digital Orthophotography and GIS*. 2012.
<<http://proceedings.esri.com/library/userconf/proc95/to150/p124.html>>.
Acessado em 27 julho. 2012.

THAPA, R.; MURAYAMA, Y. Urban mapping, accuracy, & image classification: A comparison of multiple approaches in Tsukuba City, Japan. *Applied Geography*, out. 2008. ISSN 01436228.

XIANG, S.; NIE, F.; ZHANG, C. Learning a mahalanobis distance metric for data clustering and classification. *Pattern Recognition*, v. 41, n. 12, p. 3600 – 3612, 2008. ISSN 0031-3203.
<<http://www.sciencedirect.com/science/article/pii/S0031320308002057>>.