



UNESP – UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
CAMPUS DE PRESIDENTE PRUDENTE
FACULDADE DE CIÊNCIAS E TECNOLOGIA

**Classificação interativa de imagens orbitais utilizando técnica de
projeção multidimensional de dados**

Orientando: Tiago Oyan Aguiar

Orientador: Prof. Dr. Erivaldo Antônio da Silva

Co-orientador: Prof. Dr. Wallace Correa de Oliveira Casaca

PRESIDENTE PRUDENTE
OUTUBRO DE 2019

Tiago Oyan Aguiar

**Classificação interativa de imagens orbitais utilizando técnica de
projeção multidimensional de dados**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Cartográficas – PPGCC da Faculdade de Ciências e Tecnologia – UNESP.

Orientador: Prof. Erivaldo Antônio da Silva

Co-orientador: Prof. Dr. Wallace Correa de Oliveira Casaca

PRESIDENTE PRUDENTE

2019

AUTORIZO A DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

A282c	<p>Aguiar, Tiago Oyan</p> <p>Classificação interativa de imagens orbitais utilizando técnica de projeção multidimensional de dados / Tiago Oyan Aguiar. -- Presidente Prudente, 2019</p> <p>59 p. : il., fotos, mapas</p> <p>Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências e Tecnologia, Presidente Prudente</p> <p>Orientador: Erivaldo Antônio da Silva</p> <p>Coorientador: Wallace Correa de Oliveira Casaca</p> <p>1. Ciências Cartográficas. 2. Sensoriamento Remoto. 3. Local Affine Multidimensional Projection (LAMP). 4. Classificação de Imagens. 5. Processamento Digital de Imagens. I. Título.</p>
-------	---

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências e Tecnologia, Presidente Prudente. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

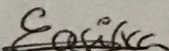
TÍTULO DA DISSERTAÇÃO: CLASSIFICAÇÃO INTERATIVA DE IMAGENS ORBITAIS UTILIZANDO
TÉCNICA DE PROJEÇÃO MULTIDIMENSIONAL DE DADOS

AUTOR: TIAGO OYAN AGUIAR

ORIENTADOR: ERIVALDO ANTONIO DA SILVA

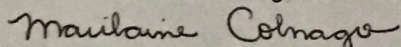
COORIENTADOR: WALLACE CORREA DE OLIVEIRA CASACA

Aprovado como parte das exigências para obtenção do Título de Mestre em CIÊNCIAS
CARTOGRÁFICAS, área: Aquisição, Análise e Representação de Informações Espaciais pela
Comissão Examinadora:



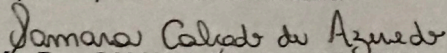
Prof. Dr. ERIVALDO ANTONIO DA SILVA

Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente



Profa. Dra. MARILAINE COLNAGO

Campus Experimental de Rosana



Profa. Dra. SAMARA CALÇADO DE AZEVEDO

UNIFEI / Universidade Federal de Itajubá

Presidente Prudente, 13 de setembro de 2019

Dedico,

A minha mãe Sonia pela base, afeto e amparo necessários para trilhar meu caminho. Aos meus avós Maria e Antônio, pelo amor e carinho incondicional. Ao professor Eivaldo por acreditar e apoiar meu desenvolvimento acadêmico, deste o segundo ano da minha jornada na UNESP. Dedico também, a minha esposa Fernanda, que de uma forma especial me deu forças, auxílio, amor e coragem para continuar e me desenvolver cada vez mais.

AGRADECIMENTOS

- Agradeço primeiramente a Deus, pela sabedoria e perseverança a mim concedidas e por sempre me guiar e proteger pelos caminhos da vida.

- Aos meus pais e avós, por todo apoio, incentivo, cuidado, paciência, amor e carinho ao longo da minha vida.

- A minha esposa, Fernanda, pelo amor, companheirismo e paciência, por enfrentarmos juntos cada dificuldade.

- A toda minha família, por todo carinho, companheirismo, incentivo e momentos de alegria.

- Aos professores Dr. Erivaldo Antônio da Silva e Dr. Wallace Correa de Oliveira Casaca, pelo aceite em me orientar, por todo conhecimento transmitido, pela paciência, amizade, compreensão e auxílio no meu desenvolvimento acadêmico.

- A UNESP, pela infraestrutura, por todo suporte e todos os funcionários, que me ajudaram diretamente ou indiretamente nas atividades desenvolvidas.

- A CAPES, pelo incentivo financeiro cedido durante o desenvolvimento do mestrado.

- Aos meus amigos do PPGCC (alunos, funcionários, professores), que me ajudaram ativamente em toda minha trajetória acadêmica do mestrado.

- Aos meus amigos da empresa Raízen Energia, pelo companheirismo e auxílio durante meu desenvolvimento acadêmico. Em especial ao meu gestor Vanderson, que me ajudou ativamente me apoiando e colaborando para o desenvolvimento desta pesquisa.

- A todos que não citei de forma direta, mas que me incentivaram e auxiliaram nesta jornada.

- O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“A vida é uma peça de teatro que não permite ensaios. Por isso, cante, chore, dance, ria e viva intensamente, antes que a cortina se feche e a peça termine sem aplausos.”

Charlie Chaplin

LISTA DE FIGURAS

Figura 1: Exemplo da aplicação do sensoriamento remoto na lavoura.	17
Figura 2: A reflectância da vegetação mostra alto gradiente entre as bandas 4 e 3.	18
Figura 3: Processamento pixel a pixel realizado nas operações aritméticas.	20
Figura 4: Projeção do conjunto de 1357 notícias, divididas em 9 classes.	25
Figura 5: Área teste correspondente ao comprimento de onda da banda do VERDE.	31
Figura 6: Área teste correspondente ao comprimento de onda da banda do VERMELHO.	31
Figura 7: Área teste correspondente ao comprimento de onda da banda do INFRAVERMELHO.	32
Figura 8: Área teste correspondente a composição colorida cor verdadeira.	33
Figura 9: Montagem da matriz base de processamento.	36
Figura 10: Sequência parcial de processamento utilizada no módulo de interação com o usuário.	39
Figura 11: Etapas de trabalho para o desenvolvimento da pesquisa.	40
Figura 12: Resultado da aplicação do índice RVI sobre a imagem teste.	41
Figura 13: Resultado da aplicação do índice NDVI sobre a imagem teste.	42
Figura 14: Resultado da aplicação do índice NDWI sobre a imagem teste.	42
Figura 15: Gráfico contendo os dados resultante da técnica de projeção LAMP (os eixos do gráfico se encontram na mesma escala da normalização).	44
Figura 16: Imagem resultante do processo de clusterização, gerado sem a intervenção do usuário (branco representa regiões com vegetação e preto representa regiões com solo exposto).	45
Figura 17: Processo de intervenção do usuário na clusterização.	46

Figura 18: Gráfico contendo o pixel selecionado (em vermelho), e os demais pontos da projeção (vegetação e não vegetação) - (os eixos do gráfico se encontram na mesma escala da normalização).....	47
Figura 19: Gráfico didático utilizado para compreender melhor o processo de geração do novo ponto de controle (baseado no pixel selecionado pelo usuário). ...	47
Figura 20: Imagem resultante do processo de clusterização, após a intervenção do usuário (branco representa regiões com vegetação e preto representa regiões com solo exposto).....	49
Figura 21: Produto cartográfico gerado com o processamento.....	50
Figura 22: Imagens resultantes do processamento, destacando as regiões que sofreram alterações com o processo de intervenção do usuário.	52

SUMÁRIO

RESUMO.....	11
ABSTRACT	12
1. INTRODUÇÃO	13
1.1. JUSTIFICATIVA.....	14
1.2. OBJETIVO GERAL	15
1.2.1. Objetivos específicos.....	15
2. FUNDAMENTAÇÃO TEÓRICA.....	16
2.1. SENSORIAMENTO REMOTO	16
2.2. COMPORTAMENTO ESPECTRAL DE ALVOS	17
2.2.1. Comportamento Espectral da Vegetação.....	18
2.2.2. Comportamento Espectral da Água.....	19
2.3. OPERAÇÕES ARITMÉTICAS DE IMAGENS	19
2.4. ÍNDICES ESPECTRAIS	21
2.4.1. Índice de Vegetação: RVI.....	21
2.4.2. Índice de Vegetação: NDVI	22
2.4.3. Índice de Corpos d'água: NDWI	23
2.5. DADOS DE ALTA DIMENSÃO	23
2.6. PROJEÇÃO MULTIDIMENSIONAL DE DADOS	24
2.6.1. Local Affine Multidimensional Projection (LAMP)	27
2.6.2. Clusterização K-MÉDIAS	28
2.7. CARTOMORPH	29
3. MATERIAIS E MÉTODOS	30
3.1. MATERIAIS UTILIZADOS.....	30
3.2. ÁREAS TESTE	30
3.3. METODOLOGIA	33
4. RESULTADOS E DISCUSSÃO	41
4.1. APLICAÇÃO DOS ÍNDICES ESPECTRAIS.....	41
4.2. PROCESSAMENTOS INICIAIS	43
4.3. PROJEÇÃO MULTIDIMENSIONAL LAMP	43

4.4.	CLUSTERIZAÇÃO	44
4.5.	INTERVENÇÃO DO USUÁRIO.....	45
4.6.	PRODUTO CARTOGRÁFICO	49
4.7.	DISCUSSÃO	51
5.	CONSIDERAÇÕES FINAIS	53
	REFERÊNCIAS BIBLIOGRÁFICAS	55

RESUMO

A extração de informação por meio do uso de imagens digitais orbitais é de grande importância para diversas áreas do conhecimento. Deste modo, essa pesquisa visa classificar imagens orbitais de alta resolução espacial, utilizando índices espectrais e a técnica de projeção multidimensional de dados Local Affine Multidimensional Projection (LAMP). Embora técnicas de projeções multidimensionais têm sido utilizadas com sucesso em grande conjunto de aplicações na área de visualização de dados, estas, por sua vez, não têm sido exploradas no contexto de cartografia. Observando esta lacuna, foi desenvolvido um método, no contexto da Cartografia e Geociências, a partir do uso e manipulação da técnica LAMP, em uma imagem orbital do satélite QuickBird, com a finalidade de apresentar à comunidade cartográfica e empresas interessadas, novas possibilidades de desenvolvimento de atualizações de bases cartográficas. É válido destacar os benefícios da técnica mencionada, principalmente para suprimir uma demanda cada vez mais recorrente no Sensoriamento Remoto, ligada ao aumento das resoluções espectrais e espaciais e consequente aumento do volume de dados para processamento, que, com o auxílio da técnica LAMP, podem ser manipulados e processados de forma adequada. Assim, mostra-se que a referida técnica, quando aliada a outras abordagens clássicas da área de Geociências, atinge um bom nível de classificação em imagens aéreas, definindo adequadamente as áreas de vegetação e solo exposto. Como produto da pesquisa, foi gerado um mapa de classificação da vegetação e um algoritmo utilizado no processamento dos dados. Vale ressaltar que o foco da pesquisa está voltado para a cultura da cana-de-açúcar, sendo essa, a feição de interesse neste estudo.

ABSTRACT

The extraction of information by using orbital digital images is of great importance for several areas of knowledge. Thus, this research aims to classify orbital images of high spatial resolution, using spectral indexes and the technique of multidimensional data projection site Affine Multidimensional Projection (LAMP). Although, multi-dimensional projection techniques have been successfully used in a large set of applications in the field of data visualization, however, have not been explored in the context of cartography. Observing this gap, a method was developed in the context of cartography and geosciences, from the use and manipulation of the LAMP technique, in an orbital image of the QuickBird satellite, to present to the cartographic community and companies new possibilities for the development of cartographic bases updates. It is worth highlighting the benefits of the aforementioned technique, mainly to suppress an increasingly recurrent demand in remote sensing, linked to the increase of spectral and spatial resolutions and consequent increase in the volume of data for processing, which, with the aid of the LAMP technique, can be manipulated and processed appropriately. Thus, it is shown that the aforementioned technique, when allied to other classical approaches in the area of geosciences, achieves a good level of classification in aerial images, appropriately defining the areas of vegetation and exposed soil. As a research product, a vegetation classification map was generated and an algorithm used in the data processing. It is noteworthy that the focus of the research is focused on the sugarcane crop, which is the feature of interest in this study.

1. INTRODUÇÃO

Os produtos cartográficos são de fundamental importância para o planejamento e o gerenciamento de projetos em diversas áreas e vários países enfrentam problemas para manter suas bases cartográficas atualizadas. Por isso diversas técnicas para atualizações de produtos cartográficos têm surgido para mitigar esses problemas, sendo o Sensoriamento Remoto uma ferramenta que têm contribuído de forma decisiva para minimização dessa desatualização.

O desenvolvimento de trabalhos que favoreçam as atualizações de produtos cartográficos é importante, principalmente visando a obtenção de bases cartográficas confiáveis e economicamente viáveis para aplicações diversas. Dentre as aplicações, têm-se o uso em planejamentos urbano (plano diretor) e rural (cadastro rural), estudos ambientais, classificação da cobertura do solo, dentre muitos outros que podem ser citados. Assim, integrando produtos de sensoriamento remoto e técnicas de processamento digital de imagens (PDI), é possível desenvolver ferramentas que possam suprir as atuais necessidades.

As técnicas de PDI existentes, foram sendo aprimoradas e utilizadas principalmente para desenvolver realce e restauração em imagens digitais, fato esse que têm atraído diversos pesquisadores, que contribuem com melhorias em diferentes processos, acarretando um crescimento expressivo, onde suas aplicações podem ser notadas nos mais variados ramos da atividade humana. Exemplos do uso do processamento e a interpretação automática de imagens captadas por satélites, são os trabalhos apresentados nas áreas de Geografia, Sensoriamento Remoto, Geoprocessamento e Meteorologia.

Neste contexto, esta pesquisa foi estruturada para estudar e avaliar o potencial de utilização da técnica de projeção multidimensional *Local Affine Multidimensional Projection* (LAMP), apresentada nos estudos de Joia et al. (2011) e Casaca et al. (2012), no campo da cartografia. Essa técnica supri uma demanda atual, relacionada ao aumento do volume de dados que são processados rotineiramente nas atividades de classificação em imagens orbitais. Portanto, a linha de pesquisa foi iniciada com base nos estudos desenvolvidos durante o período de realização da Iniciação Científica do acadêmico, que abordaram os índices espectrais e evoluída nessa pesquisa de mestrado; combinando o estudo dos

índices espectrais com a utilização da projeção multidimensional. Mais informações sobre os índices mencionados podem ser encontradas em trabalhos publicados pelo acadêmico em Aguiar et al. (2016), Aguiar et al. (2018) e Silva et al. (2018).

No caso dos índices, estes foram aplicados às imagens orbitais de alta resolução, com o intuito de realçar feições de interesse, antes que elas sejam submetidas ao processamento da projeção LAMP, que é o foco desta pesquisa. Os índices espectrais, segundo Schowengerdt (2006), tem como principal característica promover o realce de feições de interesse, como a vegetação e solo exposto, sendo muito utilizados para auxiliar na estimativa de biomassa e cobertura vegetal, por exemplo.

As imagens resultantes das aplicações dos índices espectrais, juntamente com as imagens originais (correspondentes ao comprimento de onda do verde, vermelho e infravermelho próximo), produzem uma série de dados (*features*), que estão localizados em um espaço de “n” dimensões. Esses dados são então mapeados para um espaço de recursos visuais, utilizando para isso, um método de projeção multidimensional iterativo. Conforme já mencionado, o método de projeção selecionado nesta pesquisa foi a técnica LAMP, que permite que o usuário interaja com a classificação, visando a obtenção de resultados mais precisos e customizados a fim de atender aos anseios do mesmo.

Uma diferença entre vários métodos de classificação extensivamente usados em estudos atuais e a metodologia proposta nesta pesquisa é a possibilidade de o usuário contribuir de forma mais proativa no processo de classificação, por meio de seus *insights* visuais e perspectivas de uso para o cumprimento de suas tarefas de interesse.

1.1. JUSTIFICATIVA

A pesquisa se justifica pela necessidade em buscar soluções consistentes e eficazes para a classificação de imagens digitais, visto que, são fontes de informação para uma série de aplicações. Nesse sentido, é estudado o uso de uma nova abordagem no campo da cartografia, que se fundamenta na aplicação da técnica de projeção multidimensional de dados *Local Affine Multidimensional Projection*, aliado a possibilidade de o usuário contribuir com este processo. Também se destaca a possibilidade de estudar uma alternativa que possa ser

eficiente no processo de classificação em imagens digitais orbitais, correlacionando com as necessidades do grupo Raízen Energia e a dinâmica de alteração da vegetação.

1.2. OBJETIVO GERAL

O estudo tem como principal objetivo estabelecer uma classificação em imagens orbitais, utilizando algoritmos clássicos de clusterização, aliado ao processo de interação com o usuário. Isso promove um aprimoramento da classificação a partir da técnica de projeção multidimensional de dados LAMP, procurando assim, analisar os benefícios do uso dessa técnica na cartografia.

1.2.1. Objetivos específicos

Os objetivos específicos do desenvolvimento desse estudo são:

- ❖ Aplicar e verificar a eficiência dos índices espectrais implementados no *software* CARTOMORPH, dando sequência a linha de pesquisa do acadêmico;
- ❖ Desenvolver e sistematizar um algoritmo que realize o processo iterativo de classificação, baseado em algoritmos clássicos de cluterização e na técnica de projeção multidimensional de dados LAMP;
- ❖ Disponibilizar, sob demanda, o algoritmo de processamento para os usuários, sempre considerando um desenvolvimento claro e conciso;
- ❖ Avaliar a classificação da imagem proposta, aplicada no estudo de áreas vegetativas;
- ❖ Obtenção de um produto cartográfico que possa ser útil para estudos diversos, principalmente no que diz respeito a áreas de vegetação do setor sucroenergético.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. SENSORIAMENTO REMOTO

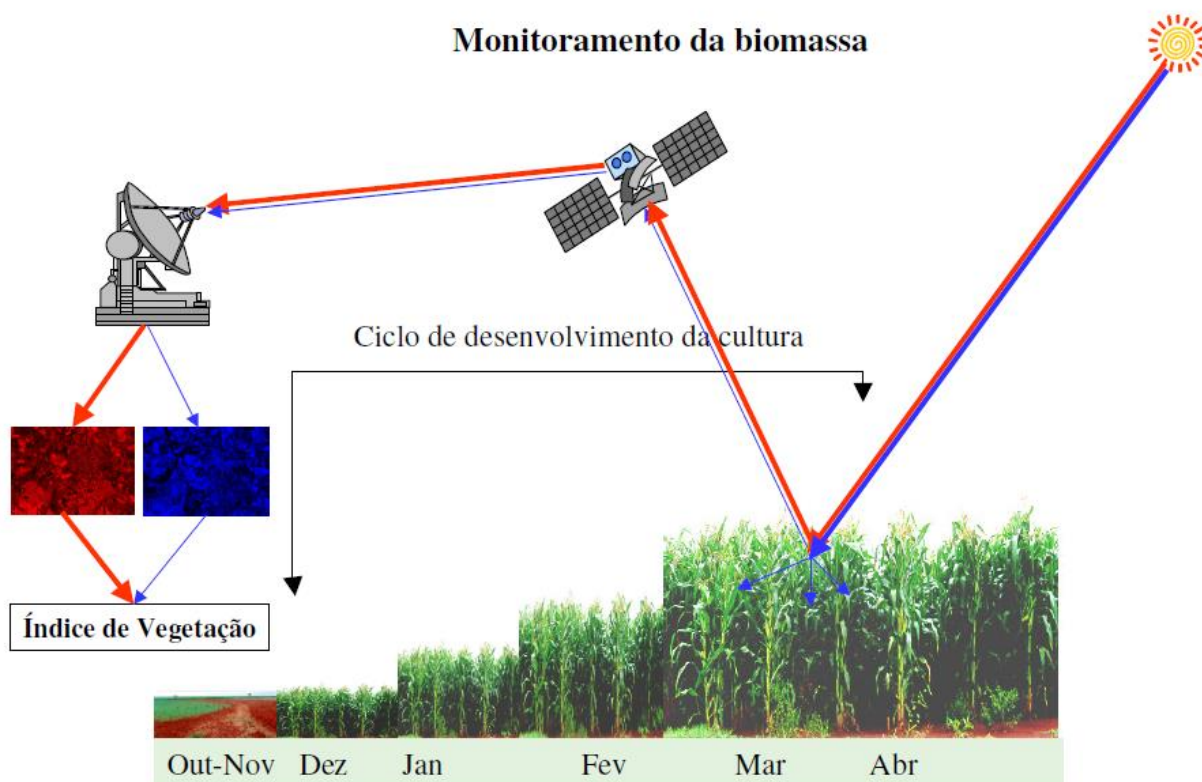
O sensoriamento remoto pode ser definido como uma ciência que obtém informações sobre um objeto, uma área ou um fenômeno, por meio da análise de dados, que são obtidos sem que haja contato físico direto com este objeto, esta área ou este fenômeno. O termo sensoriamento remoto foi criado para designar o desenvolvimento da tecnologia que através de instrumentos, foi capaz de obter imagens da superfície terrestre, onde a condição principal imposta é que o sensor deve estar a uma distância remota do objeto. Assim, a aquisição de dados inclui uma fonte de energia eletromagnética, um sensor que transforma a energia proveniente do alvo em sinal e um analisador que transforma esse sinal em informação (LILLESAND et al. 2004; SILVA, 1995; MENESES; ALMEIDA, 2012).

Algumas características muito importantes das imagens provenientes dos satélites são quanto: ao número e ao comprimento das bandas do espectro eletromagnético (resolução espectral), a menor área da superfície terrestre observada instantaneamente por cada sensor (resolução espacial), ao nível de quantização registrado pelo sistema sensor (resolução radiométrica) e ao intervalo entre duas passagens do satélite pelo mesmo ponto (resolução temporal) (CÂMARA et al. 2001). Com isso Meneses e Almeida (2012), afirmam que as resoluções espectrais das imagens obtidas pelos sensores atuais, ultrapassam as centenas de bandas e com a acurácia da resolução espacial, é possível utilizar as imagens de sensoriamento remoto em diversas áreas, como exemplo destacam-se, os levantamentos de recursos naturais, mapeamentos temáticos, monitoramento ambiental, detecção de desastres naturais, desmatamentos florestais, previsões de safras, cadastramentos territoriais multifinalitários, cartografia de precisão, defesa e vigilância; sendo essa, uma tecnologia empregada em escala mundial.

Diversas áreas aplicam os conceitos relacionados com sensoriamento remoto, como a agricultura, meio ambiente, geologia, recursos hídricos, estudo de solos e florestas. Na agricultura, pode-se utilizar o sensoriamento remoto quando se deseja realizar previsão de safras, mapeamento de culturas, definição de áreas de aptidão agrícola, zoneamento agro-ecológico e monitoramento de incêndios em lavouras e pastagens. Figueiredo (2005) mostra que na previsão de safras, por

exemplo, o sensoriamento remoto pode ser aplicado no dimensionamento das áreas de plantio e na estimativa do rendimento. A Figura 1 exemplifica o sensoriamento remoto sendo aplicado na estimativa de rendimento por meio do índice de vegetação.

Figura 1: Exemplo da aplicação do sensoriamento remoto na lavoura.



Fonte: Figueiredo (2005).

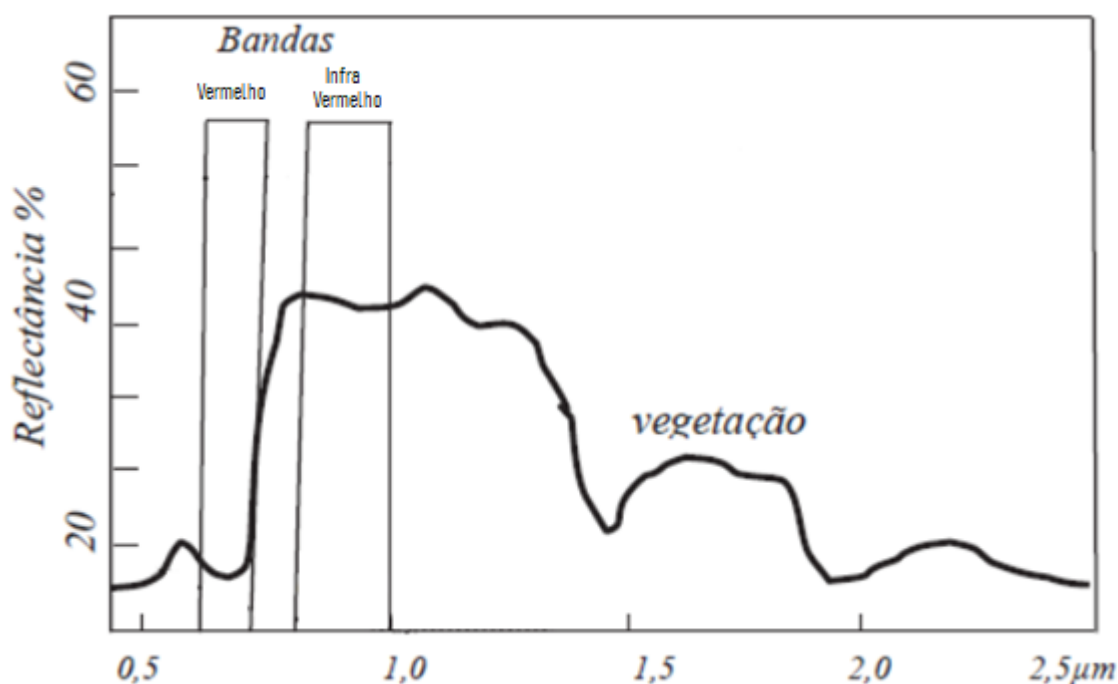
2.2. COMPORTAMENTO ESPECTRAL DE ALVOS

Os alvos apresentam comportamento espectral diferente entre si, dependendo principalmente das características físico-químicas dos materiais constituintes, os quais apresentam refletâncias próprias para cada banda do espectro eletromagnético. Assim, estudar o comportamento espectral de alguns alvos de interesse, torna-se imprescindível para a correta utilização dos índices espectrais (MATHER, 2004).

2.2.1. Comportamento Espectral da Vegetação

A vegetação apresenta um pico de absorção na região do visível, devido a presença da clorofila, que é encontrada nas folhas. Existem duas bandas de absorção distintas, estando uma centrada próximo ao comprimento de onda de 580 nm, relacionado a presença de carotenos e a outra próxima ao comprimento de onda de 750 nm, relacionado ao processo fotossintético. A Figura 2 apresenta um gráfico com a curva de reflectância da vegetação sadia, onde é possível verificar os comprimentos de onda correspondente às bandas de absorção da refletância, por conta da clorofila presente na vegetação (FIGUEIREDO, 2005; MENESES; ALMEIDA, 2012).

Figura 2: A reflectância da vegetação mostra alto gradiente entre as bandas 4 e 3.



Fonte: Adaptado de Meneses e Almeida (2012).

Entre as duas regiões de absorção mencionadas anteriormente, existe um pico de reflectância próximo aos 500 nm, que está relacionado a coloração verde observada na vegetação. Também se observa uma alta reflectância entre 700 nm a 1300 nm, relacionada à estrutura celular da vegetação. Entretanto, o comportamento espectral da vegetação pode variar com a interferência de diversos fatores externos, como exemplos destacam-se a cobertura vegetal, área foliar, estrutura da folha,

condições atmosféricas, biomassa, características do solo, estado fenológico, geometria de aquisição dos dados pelo satélite, dentre outros (FIGUEIREDO, 2005).

2.2.2. Comportamento Espectral da Água

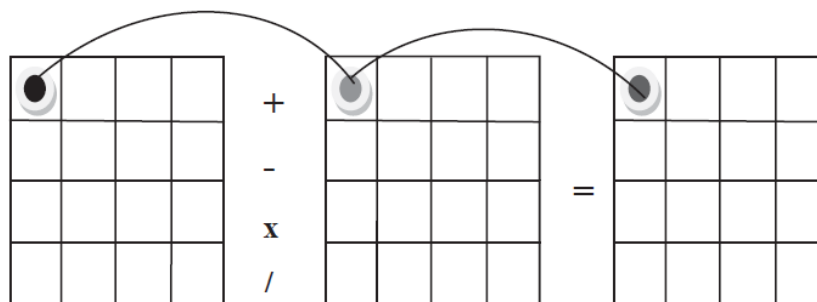
A água pode apresentar um comportamento espectral diferente, dependendo do estado físico em que se encontrar, no instante da aquisição das imagens. No estado líquido, os corpos d'água apresentam reflectância espectral maior entre 0,38 μ m e 0,70 μ m, sendo assim, eles irão absorver toda a radiação que esteja acima do comprimento de onda de 0,7 μ m (FIORIO, 2004).

No entanto, a concentração de materiais em suspensão e dissolvidos no corpo d'água podem alterar o comportamento espectral da água líquida. Assim, à medida que há um aumento nos sólidos em suspensão no corpo d'água, ocorre também um aumento na reflectância observada. (BARBOSA et al. 2003; FIORIO, 2004; KIRK, 1994). Novo (2008) também traz essa análise, mostrando que as medidas realizadas em campo e os experimentos realizados em laboratório, comprovam que o aumento da concentração de sólidos totais em suspensão na água, apresenta como principal efeito, o aumento do coeficiente de espalhamento da reflectância dela.

2.3. OPERAÇÕES ARITMÉTICAS DE IMAGENS

Segundo Meneses e Almeida (2012), a aritmética entre as bandas é uma das mais simples formulações algorítmicas de processamento de imagens. Em resumo, as operações aritméticas realizam um processamento nos dados obtidos pelos sensores, obtendo como resultado, uma nova imagem diferente das imagens iniciais. É importante destacar que o processo de transformação é feito pixel a pixel, por meio de uma lógica matemática pré-estabelecida, pelo algoritmo de processamento. A Figura 3 ilustra como é realizado este processamento e, a seguir, serão comentadas brevemente cada umas das operações existentes.

Figura 3: Processamento pixel a pixel realizado nas operações aritméticas.



Fonte: Meneses e Almeida (2012).

A operação aritmética de soma, é principalmente utilizada para realçar similaridades entre imagens de diferentes bandas ou datas, o que resulta no realce de alvos com respostas espectrais diferentes daqueles presentes nas imagens originais.

A operação de subtração é muitas vezes utilizada em um par de imagens, da mesma área, com objetivo de avaliar as variações ocorridas em diferentes épocas, ou seja, aplicada na detecção de mudanças temporais dos alvos presentes em uma determinada cena. Algumas aplicações desta operação, estão relacionadas a detecção de impactos ambientais e a avaliação de desmatamentos florestais (MATHER, 2004).

Já a operação aritmética de multiplicação, tende a realçar as feições morfológicas comuns às bandas que foram utilizadas, sendo útil para a geologia estrutural, a cartografia e a geomorfologia. É possível utilizar a multiplicação de imagens como recurso para obter uma imagem que agrega o destaque da informação de textura e do relevo, com a informação espectral FLORENZANO et al. (2001).

A divisão ou razão de bandas, pode ser entendida como a operação aritmética, que está baseada em parâmetros de aspectos físicos do sistema sensor e no comportamento da reflectância dos alvos. Assim pode-se obter pixels que apresentem valores de pequena magnitude e portanto, adota-se a multiplicação por um valor escalar (ganho), seguido da adição de uma constante (*offset*), para obter um deslocamento do histograma, para uma posição centralizada. A divisão entre bandas resulta em uma imagem que melhor reflete as verdadeiras características da superfície, sendo que, entre as principais vantagens, destacam-se a minimização das diferenças nas condições de iluminação de uma cena, a reflectância espectral de determinados alvos, a determinação de índices de biomassa e área foliar,

detecção de fitopatologias e fenologia, prospecção mineral, dentre muitos outros exemplos (IBGE, 2000 p. 92; MENESES; ALMEIDA, 2012).

2.4. ÍNDICES ESPECTRAIS

É possível aplicar diversas transformações nas imagens digitais, como exemplo, as operações aritméticas mencionadas anteriormente. Estas operações permitem gerar uma nova imagem, derivada de duas ou mais faixas, de uma imagem multiespectral ou multitemporal. O processo de divisão dos valores do pixel de uma imagem, pelos pixels homólogos de uma segunda imagem é conhecido como razão entre bandas, sendo provavelmente, a operação aritmética mais empregada em estudos de análises geológica, detecção da magnitude das diferenças entre as bandas espectrais, aplicações ecológicas e agrícolas. Um dos benefícios da razão espectral é a supressão do efeito de sombreamento topográfico, unido ao fato da razão, entre uma banda NIR (*Near InfraRed*) e uma banda vermelha, por exemplo, poder destacar a vegetação contida no local (MATHER, 2004; SCHOWENGERDT, 2006).

Mather (2004) destaca dois exemplos das aplicações de operações aritméticas em imagens digitais. O primeiro é a diferença numérica entre duas imagens observadas pelo mesmo sensor, porém em datas diferentes, que podem fornecer informações sobre eventuais mudanças ocorridas naquela área, como desmatamento, cultivo de culturas, dentre outros. O segundo é que a imagem derivada de uma operação aritmética, pode apresentar propriedades que a torna mais adequada, para um determinado fim, quando comparada com a imagem original; aqui se destacam os índices de vegetação como o RVI (*Ratio Vegetation Index*) e o NDVI (*Normalized Difference Vegetation Index*), que podem ser amplamente utilizados para a determinação de variáveis como a biomassa e índice de área foliar; e o índice de corpos d'água NDWI (*Normalized Difference Water Index*).

2.4.1. Índice de Vegetação: RVI

Segundo Ponzoni e Shimabukuro (2010), os índices de vegetação são algumas das transformações mais utilizadas para analisar o comportamento da

vegetação, sendo que, o índice mais conhecido da divisão de bandas, que visa o realce das variações de densidade da cobertura vegetal, é conhecido como RVI.

Este é universalmente o mais conhecido índice, sendo que, ele consiste basicamente na divisão direta da banda situada na região do comprimento de onda do infravermelho próximo (NIR), que é considerada a região que a vegetação mostra a mais intensa reflectância; dividido pela banda situada no comprimento de onda vermelho (RED), região onde a vegetação possui a mais alta absorção da luz solar visível, causada pela presença da clorofila, indicativo de uma vegetação sadia. Portanto, esse índice é expresso pela Equação 1:

$$RVI = \frac{\text{banda do INFRA VERMELHO}}{\text{banda do VERMELHO}} = \frac{NIR}{RED} \quad (1)$$

O resultado obtido com essa divisão é uma imagem monocromática, com a variação de tons de cinza indicando maior ou menor quantidade de biomassa ou o índice de área foliar, sendo que, quanto mais claro o tom de cinza, maior a densidade de vegetação; tons de cinzas médios indicam pouca vegetação e tons de cinza totalmente escuros indicam ausência de vegetação, o que pode ocorrer nas áreas de corpos de água, solos expostos, áreas cobertas por neve, dentre outras possibilidades. O RVI foi um dos primeiros índices utilizados em aplicações de sensoriamento remoto, porém, atualmente, existem vários outros índices utilizados para as mais diversas necessidades (MENESES; ALMEIDA, 2012; SCHOWENGERDT, 2006).

2.4.2. Índice de Vegetação: NDVI

Um índice similar ao RVI, porém com um grau de complexidade um pouco mais elevado, é chamado de NDVI. Este é um dos índices mais aplicados em diversos estudos ambientais, sendo realizada uma normalização da razão entre a diferença e a soma das bandas situadas no comprimento de onda do infravermelho próximo (NIR) e a banda situada no comprimento de onda do vermelho (RED), conforme apresentado na Equação 2:

$$NDVI = \frac{\text{banda do INFRAVERMELHO} - \text{banda do VERMELHO}}{\text{banda do INFRAVERMELHO} + \text{banda do VERMELHO}} = \frac{NIR - RED}{NIR + RED} \quad (2)$$

O fato de utilizar somas e diferenças entre bandas, pode fazer com que o NDVI seja mais apropriado para monitoramento de mudanças sazonais da vegetação, além de reduzir ruídos inerentes aos trabalhos com imagens multitemporais, principalmente no que se refere a diferença nas condições atmosféricas e de iluminação. Por isso, este índice é amplamente utilizado para realizar o monitoramento da vegetação em escalas continentais e globais (MATHER 2004; JENSEN, 2004; SCHOWENGERDT, 2006).

Quando comparado os índices RVI e NDVI, percebe-se que o segundo apresenta uma vantagem sobre o primeiro, isso porque o NDVI tende a ser linearmente mais proporcional à biomassa. Portanto, alguns exemplos de trabalhos que aplicam o NDVI são: medida de índice de área foliar, determinação da porcentagem de cobertura do solo e estimativas da radiação fotossinteticamente ativa (MENESES; ALMEIDA, 2012; SCHOWENGERDT, 2006).

2.4.3. Índice de Corpos d'água: NDWI

Segundo Gao (1996), baseado na curva de reflectância da água, pode-se criar um índice de diferença normalizada utilizando a imagem que corresponde ao comprimento de onda do verde (GREEN) e a correspondente ao comprimento de onda do infravermelho próximo (NIR). Portanto, o índice é expresso pela Equação 3:

$$NDWI = \frac{\text{banda do VERDE} - \text{banda do INFRAVERMELHO}}{\text{banda do VERDE} + \text{banda do INFRAVERMELHO}} = \frac{GREEN - NIR}{GREEN + NIR} \quad (3)$$

A partir da aplicação da razão entre bandas, formada pelo índice NDWI, é possível obter duas situações muito características. A primeira é que a imagem resultante da aplicação do NDWI apresenta tons de cinzas mais claros para os corpos d'água, solo exposto e rodovias e a segunda situação, é que para as áreas com vegetação, a imagem resultante apresenta tons de cinzas escuros. Assim o índice apresenta importantes aplicações na agricultura e na silvicultura (MCFEETERS, 1996; GAO, 1996).

2.5. DADOS DE ALTA DIMENSÃO

Segundo Pappu e Pardalos (2014), a precisão dos algoritmos de classificação convencionais tende a cair, quando aplicados em conjuntos de dados com

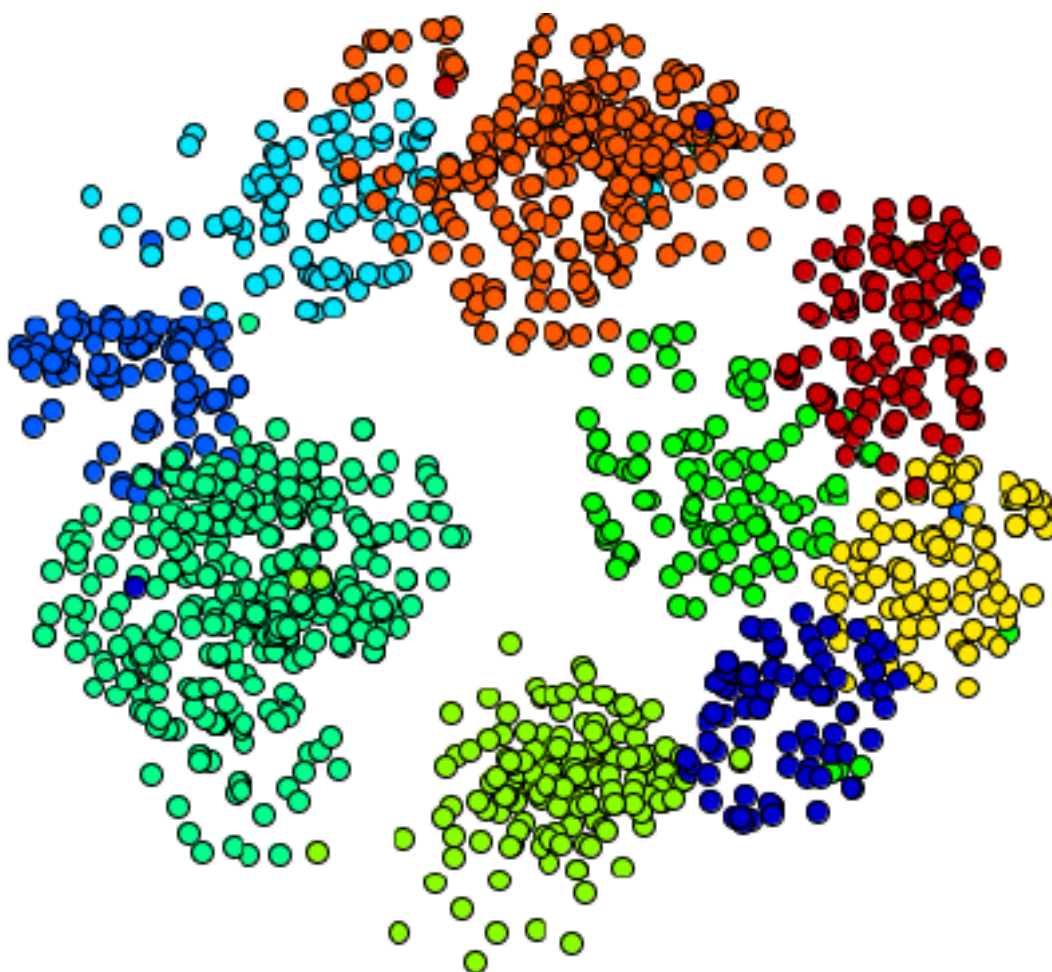
dimensões muito elevadas, ou seja, a medida que dimensão dos dados aumenta demasiadamente, a capacidade do algoritmo em convergir para um modelo de classificação satisfatório fica comprometida; o que é conhecido na literatura como "*the curse of dimensionality*".

Como visto, o emprego de um número muito elevado de dimensões pode trazer imprecisão nas classificações de dados, porém, quando se utiliza um número apropriado de dimensões e algoritmos clássicos de processamento de dados multidimensionais, tende-se a obter altas taxas de assertividade, conforme mencionam os estudos do estado-da-arte de classificação em espaços multidimensionais de dados, dos autores Berkhin (2002), Kogan et. al (2006) e Steinbach et. al (2004). Isso significa que, para um número razoável de dimensões, dependendo do tipo de aplicação/estudo, a utilização desse tipo de aparato computacional se torna uma ferramenta poderosa no agrupamento de dados semelhantes, além de promover a abertura ao usuário, para que o mesmo possa interagir diretamente com esses dados, de acordo com as suas necessidades.

2.6. PROJEÇÃO MULTIDIMENSIONAL DE DADOS

As técnicas de projeção multidimensional de dados consistem em projetar, em um espaço de dimensão igual a 1, 2 ou até 3D, amostras presentes em um espaço de dados de dimensão maior ou igual a três, em que não é possível visualizar, de forma direta, os padrões e relações presentes nesse espaço de dimensão superior (MAMANI, 2012). Assim, para que seja possível compreender melhor como as técnicas de projeção multidimensional de dados atuam, foi adaptada a Figura 4, de Nieto (2012), onde é possível verificar a projeção de um conjunto de dados com 1.357 notícias coletadas de diferentes agências, divididas em 9 classes. A partir do referido gráfico, é possível notar que a projeção em 2 dimensões desse conjunto de dados propicia ao usuário autonomia e experiência de interação em meio às 1.357 notícias categorizadas sob diferentes assuntos (*clusters*) de interesse. Dessa forma, há ainda a possibilidade da preservação da clusterização feita no espaço de origem (de alta dimensão), quando aplicada técnicas de projeção que preservem as relações da vizinhança existentes no espaço original.

Figura 4: Projeção do conjunto de 1357 notícias, divididas em 9 classes.



Fonte: Adaptado de Nieto (2012).

Existem diversas técnicas de projeção de dados que conseguem trabalhar, de forma satisfatória, com grandes volumes de dados multidimensionais, portanto, a escolha da técnica adequada, pode promover análises mais profundas e tarefas direcionadas em meio à um vultuoso conjunto de dados (MAMANI, 2012). Exemplos de técnicas que fazem uso desta abordagem podem ser consultadas nos trabalhos Mamani (2012), Nieto et al. (2014), Etemadpour et. al (2016), Nieto et al. (2016), por exemplo. As técnicas de projeções multidimensionais tendem a preservar as estruturas de uma vizinhança de um conjunto de dados, o que gera a observação de que, pontos que apresentam significativa relação semântica serão projetados mais próximos uns dos outros no espaço visual de dimensão reduzida (PAULOVICH et al. 2008).

Entende-se então que as técnicas de projeções multidimensionais, são ferramentas amplamente utilizadas na exploração de dados multidimensionais, pois as mesmas mapeiam os dados de um espaço de alta dimensão para outro de até 3

dimensões, de modo a preservar as relações de similaridades entre as instâncias desses dados, além de preservar as operações realizadas no espaço de alta dimensão como, por exemplo, a tarefa de classificação dos dados. As amostras que são projetadas próximas umas das outras, no plano visual de 2 ou 3 dimensões, apresentam alto grau de correlação, enquanto pontos mapeados distantes uns dos outros apresentam baixa correlação. Isso significa que, as técnicas de projeções procuram preservar a distância no espaço original, de 'n' dimensões, quando projetam os dados em um espaço visual de menor dimensão (SIEDLECKI et. al 1988, TEJADA et. al 2003).

Segundo Etemadpour et. al (2016), uma vantagem do uso das técnicas de projeção multidimensional, está relacionada à possibilidade de identificar padrões quando comparadas com outros métodos clássicos de análise de dados, o que pode fornecer uma maior confiabilidade dos resultados obtidos. Portanto, o emprego das técnicas de projeção multidimensional, é uma solução bastante efetiva quando se trata da interpretação e mineração de conjunto de dados de elevada dimensionalidade, sendo essa uma ferramenta relevante devido à sua capacidade de revelar estruturas que apresentam correlação em espaços multidimensionais, conforme reportado no estudo de Neves et. al (2015).

O objetivo de se utilizar técnicas de projeção multidimensional está relacionado com a criação de representações visuais, que sumarizem a informação original presente no espaço de alta dimensão dos dados, para que os usuários possam reconhecer, visualizar e explorar, padrões e estruturas que estejam presentes em um determinado conjunto de dados. Uma vez projetados em um espaço bidimensional, estes dados são convertidos em elementos geométricos (gráficos), passíveis de interação por parte do usuário, onde as posições geométricas desses elementos refletem tanto as relações de similaridade entre os dados como, ainda, relacionamentos *intra-clusters* que foram formados no espaço multidimensional de origem. Assim, os elementos que estão próximos uns dos outros na representação final, são interpretados como elementos de alto grau de similaridade, em relação a uma determinada medida e os elementos que estão distantes na representação final, indicam que os objetos não estão correlacionados (MACHADO, 2010).

2.6.1. Local Affine Multidimensional Projection (LAMP)

A LAMP é uma técnica de projeção multidimensional construída a partir da teoria do mapeamento ortogonal. Tal técnica foi selecionada para este trabalho por suas características de preservar com grande precisão as estruturas de vizinhança das amostras de dados, além de permitir alterações na projeção por meio da interação com o usuário no espaço visual de projeção, o que é uma característica importante no contexto deste trabalho (JOIA et al. 2011; TEJADA et al. 2003).

Para Joia et al. (2011), além de ser viável do ponto de vista do tempo de processamento computacional e da acurácia devido à preservação de vizinhança, a técnica LAMP é robusta no que diz respeito aos pontos de controles, apresentando baixa distorção mesmo quando utilizado um número reduzido destes pontos. No algoritmo apresentado pelo autor, os pontos de controle são escolhidos aleatoriamente em meio à nuvem de dados (amostras) e servem para direcionar a projeção do restante das amostras. A técnica conta também com a possibilidade da intervenção do usuário, onde ele pode a partir do conjunto pré-determinado de pontos de controle, adicionar novos pontos e/ou modificar pontos de baixa relevância.

De fato, essas particularidades favorecem seu uso, principalmente, quando aplicado a um conjunto de dados consideravelmente volumoso, em cenários supervisionados pelo usuário, onde o mesmo interage diretamente com os pontos de controle no espaço visual da projeção. Essa característica pode favorecer diversas tarefas afins, como por exemplo a classificação/clusterização de pixels em imagens digitais. Nesse sentido, através da interação com o usuário, é possível corrigir instâncias de dados que não puderam ser classificadas corretamente no espaço de alta dimensão.

O processamento presente na técnica LAMP utiliza uma transformação rígida, onde os dados são projetados durante o processo de mapeamento, evitando os efeitos de cisalhamento; assim sendo, consegue-se preservar o máximo possível as distâncias no espaço original, evitando propagação dos erros de posicionamento dos pontos, por meio da restrição de ortogonalidade imposta. Essa restrição garante que os erros contidos na etapa de posicionamento dos pontos, não sejam propagados durante a etapa da projeção dos dados, além do controle e minimização das

distorções após a projeção. Maiores detalhes sobre a projeção LAMP, podem ser obtidos em Joia et al. (2011).

Com o exposto, pode se perceber que a LAMP é um método de projeção multidimensional bastante interessante para lidar com dados de alta dimensão, em particular, na tarefa de classificação interativa de imagens de sensoriamento remoto, onde as classificações automáticas nem sempre trazem resultados satisfatórios.

2.6.2. Clusterização K-MÉDIAS

Os processos de clusterização são utilizados em diversas áreas da Ciência e da Engenharia, como exemplo a Psicologia, a Biologia, a Medicina, a Visão Computacional, as Comunicações e o Sensoriamento Remoto. E dentre os diversos métodos existentes, o *K-means* ou K-médias se destaca, principalmente pela sua popularidade de uso devido a sua facilidade de implementação e sua ordem de complexidade quanto ao número de padrões (KRISHNA; MURTY, 1999; JAIN et. al. 1999).

O método de clusterização K-médias é comumente utilizado para particionar automaticamente um conjunto de dados, em "k" grupos, sendo "k" definido a priori. Os algoritmos deste método são geralmente utilizados para o desenvolvimento de processos em um ambiente não supervisionado, com um conjunto de instâncias de dados que devem ser agrupadas de acordo com suas similaridades, onde, a forma de realizar esse agrupamento está relacionada com os padrões dentro de um grupo (*cluster*) que são mais semelhantes entre si (*intra-clusters*), do que os padrões existentes nos demais grupos (*inter-clusters*). Portanto o agrupamento é realizado a partir de alguma medida da similaridade entre o conjunto de padrões, o que faz com que esse método possa ser considerado eficaz, gerando bons resultados de clusterização para muitas aplicações práticas (MACQUEEN, 1967; WAGSTAFF et al. 2001; KRISHNA; MURTY, 1999; ALSABTI et al. 1997).

A clusterização gerada pelo k-médias é iniciada tomando centros de *clusters* (centroides), seguindo de um processo iterativo, onde uma instância de dados é processada e atribuída ao centro do *cluster* mais próximo. Na sequência, é gerada uma atualização deste centro, onde é calculado uma nova média que represente as instâncias dos dados que o constituem, ou seja, depois que um ponto é adicionado a um grupo, a média desse grupo é ajustada para levar em consideração o novo

ponto. O algoritmo converge quando não há mais mudança na atribuição de instâncias para os *clusters* (MACQUEEN, 1967; FONTANA; NALDI, 2009).

2.7. CARTOMORPH

Segundo Cardim (2015), o CARTOMORPH é um *software* desenvolvido para aplicações na área da Cartografia, Sensoriamento Remoto, Fotogrametria e áreas afins, construído com o propósito de processar imagens digitais orbitais e aéreas. Este *software* segue em aprimoramento, no entanto, já possui funções que podem ser utilizadas. Exemplos são os operadores relacionados à morfologia matemática e os operadores relacionados aos índices espectrais, utilizado em sensoriamento remoto.

Sua concepção foi estruturada possibilitando a utilização das funções de processamento já implementadas, além de permitir que alterações sejam feitas sempre que necessário, uma vez que ele têm o código aberto (*software open source*). Além do mencionado, o CARTOMORPH é de domínio público (gratuito), isto significa que não é necessário pagar para adquirir uma licença de uso, o que é uma vantagem perante a outros *softwares* de processamento similar.

Diversos algoritmos de processamento foram implementados em uma biblioteca de funções do *software* CARTOMORPH, a fim de possibilitar que os usuários possam utilizar essas funções em seu modo de programação, permitindo o uso de outras técnicas, algoritmos e funções da linguagem C++. Além disso, é possível utilizar uma interface gráfica, a qual possibilita aos usuários que não tenham conhecimento específico com a linguagem de programação, realizar todas as operações contidas na biblioteca do *software*. Assim a interface gráfica possibilita a utilização das funções de forma mais rápida, prática e simples.

3. MATERIAIS E MÉTODOS

3.1. MATERIAIS UTILIZADOS

- ❖ Imagens digitais das áreas de interesse (mais detalhes no item 3.2);

- ❖ Software CARTOMORPH;

- ❖ Software MATLAB R2013a;

- ❖ Software de Sistema de Informações Geográficas Q-Gis 3.4.

3.2. ÁREAS TESTE

A imagem teste utilizada na pesquisa, compreende a região de Maracaí, no estado de São Paulo e foi obtida pelo satélite QuickBird. Essa imagem foi previamente corrigida dos efeitos de espalhamento atmosférico, com o método *Dark Object Subtraction* (DOS), pela empresa fornecedora; sendo utilizada apenas ela, por compreender uma região específica de interesse dos apoiadores e estar disponível no banco de imagens do grupo de estudo.

A imagem possui um conjunto de 4 bandas multiespectrais, com resolução radiométrica de 8 bits (reamostrada dos 11bits original), com um *Ground Sample Distance* (GSD) de 2,4 metros e foi adquirida em 28 de abril de 2009. As Figura 5, Figura 6 e Figura 7 apresentam a imagem teste nas bandas correspondente aos comprimentos de onda do verde, vermelho e infravermelho próximo, respectivamente. É importante ressaltar que as Figura 5 e Figura 6 apresentadas no texto, estão com o seu valor de brilho acrescido em 20%, para que pudessem ser observadas com melhor qualidade visual.

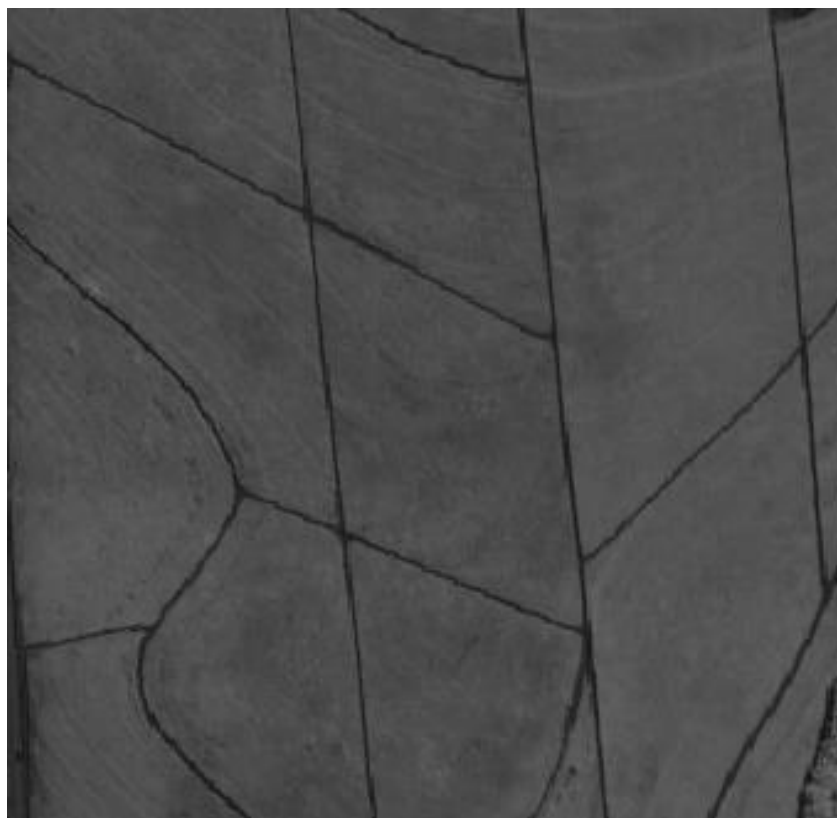
Figura 5: Área teste correspondente ao comprimento de onda da banda do VERDE.



Figura 6: Área teste correspondente ao comprimento de onda da banda do VERMELHO.



Figura 7: Área teste correspondente ao comprimento de onda da banda do INFRAVERMELHO.



Para ser possível compreender melhor as regiões da imagem teste, que correspondem as áreas de vegetação e solo exposto, foi gerada uma composição colorida, para ilustrá-las. Essa composição colorida, cor verdadeira, pode ser observada na Figura 8.

Figura 8: Área teste correspondente a composição colorida cor verdadeira.



3.3. METODOLOGIA

A pesquisa está baseada numa estrutura simples e objetiva, assim, os itens a seguir, foram fundamentais no desenvolvimento metodológico.

- ❖ Definição da área de estudo;
- ❖ Aplicação dos índices espectrais previamente selecionados;
- ❖ Desenvolvimento e implementações necessários para o processo de classificação, utilizando a técnica de projeção LAMP e o algoritmo de clusterização K-médias;
- ❖ Processamento das imagens testes;
- ❖ Geração dos produtos: mapa e algoritmo.

O trabalho foi iniciado com a aplicação dos três índices espectrais previamente selecionados, sendo usado o RVI e NDVI para promover o realce dos locais com vegetação e o NDWI para o realce das áreas de solo exposto. A partir destes índices foi possível realçar as feições de interesse, tendo como principal objetivo auxiliar no processo de classificação da área teste. O *software* CARTOMORPH foi utilizado para a aplicação dos índices mencionados, os quais foram implementados e serviram para a detecção das feições de interesse, conforme pode ser observado em AGUIAR et al. (2016), AGUIAR et al. (2018) e SILVA et al. (2018).

A partir das imagens resultantes da aplicação dos índices espectrais e das imagens correspondentes ao comprimento de onda do verde, vermelho e infravermelho próximo (totalizando 6 imagens/camadas), foi iniciada a próxima etapa, correspondente a classificação foco da pesquisa, obtendo como produto, um mapa que destaque as regiões cobertas por vegetação, das demais áreas.

Para a sequência do trabalho, foi utilizado o *software* MATLAB, onde foram desenvolvidos diversos procedimentos, buscando estruturar o processamento em funções, criadas para que fosse possível realizar os testes e manipulações das lógicas de programação, com maior facilidade e mitigando as chances de erros conceituais. Por isso, as principais funções serão comentadas na sequência da metodologia.

A estrutura desenvolvida no *software* MATLAB foi iniciada com as funções de parâmetros iniciais, onde foi possível definir a quantidade de *clusters* resultantes e a função de recorte das imagens, onde foi possível definir a região da imagem que será processada. Nesta pesquisa foram utilizadas 2 classes, uma correspondente as áreas com vegetação e outra correspondente as áreas sem vegetação. Essa forma de classificação foi projetada e baseada nas necessidades do grupo Raízen Energia.

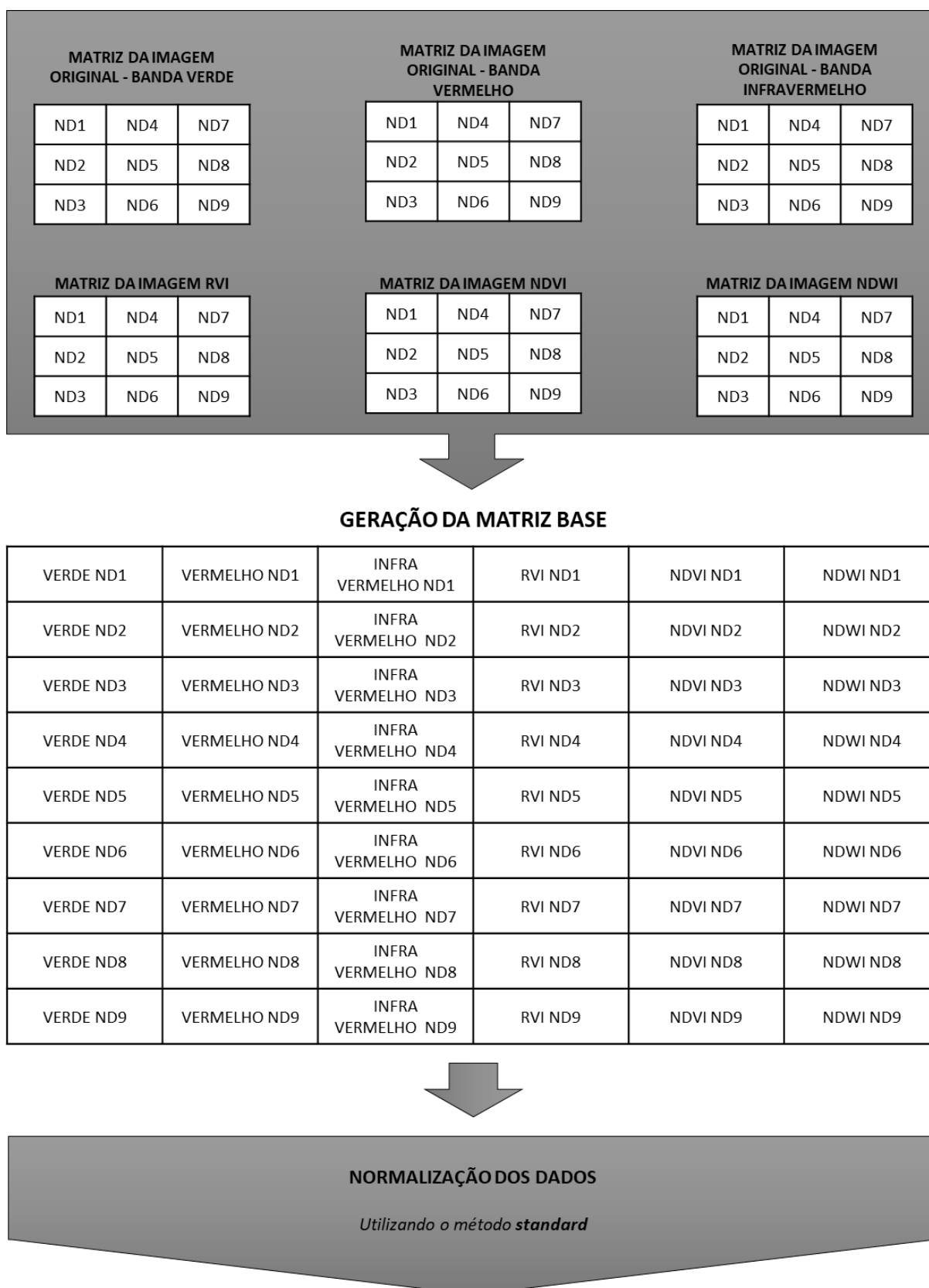
A partir do recorte gerado e da definição de classes, os pixels (tons de cinza) foram levados ao espaço R^n , ou seja, as imagens foram mapeadas para o espaço de “n” dimensão, cujas coordenadas estão relacionadas com o nível digital (ND) dos pixels. Assim foi possível gerar a matriz base de dados, observando o sistema de coordenadas coluna e linha, juntamente com o valor de ND do pixel correspondente, para cada imagem, preservando ao máximo as estruturas da vizinhança dos pixels.

O passo seguinte consistiu na normalização dos valores de ND da matriz base, para ser possível aplicar o processamento com o algoritmo da técnica de

projeção multidimensional LAMP. Este processo de normalização padroniza os dados que estão em diferentes escalas, para que possam ser interpretados em uma mesma magnitude, o que é requisito geral para muitos algoritmos de processamento de imagens digitais. Para isso, no MATLAB, foi utilizada a técnica de normalização de dados *standard*. Dessa forma os pixels foram organizados em colunas, em uma nova matriz, sendo essa a base tomada para aplicar a técnica de projeção multidimensional.

A Figura 9 apresenta um esquema ilustrativo, contendo o processo de montagem da matriz base e a normalização dos dados realizada.

Figura 9: Montagem da matriz base de processamento.



Fonte: Autor

A técnica de projeção LAMP, foi aplicada então na matriz normalizada, sendo importante ressaltar que essa técnica foi selecionada por apresentar propriedades de preservação de vizinhança e da perspectiva da interação com o usuário, tal como sugerido por estudos de Joia et al. (2011), Tejada et al. (2003) e Paulovich et al. (2008). Esta etapa da pesquisa foi adaptada de Casaca et al. (2012).

Posteriormente a aplicação da projeção multidimensional com a técnica LAMP, foi gerada a clusterização dos dados. Como a projeção utilizada apresenta a característica de preservar a distância, então a clusterização em baixa dimensão é similar à clusterização em alta dimensão, portanto, extrapolar a clusterização no espaço de 2 dimensões não altera de forma significativa os *clusters* que seriam gerados no espaço original dos dados. Para a aplicação da clusterização, foi necessário entrar com alguns argumentos iniciais, como a matriz numérica de dados (onde as linhas correspondem as observações e a as colunas correspondem as variáveis), e o número (inteiro e positivo) de *clusters* que se deseja obter, ou seja, o número de classes resultante do processo de agrupamento. Como saída, a função apresenta um vetor coluna contendo os índices dos *clusters*, ou seja, a associação de cada pixel da imagem com relação às classes previamente especificadas. Conforme mencionado anteriormente, a clusterização foi desenvolvida utilizando 2 classes, uma representando a vegetação e outra representando o que foi classificado como não vegetação.

A clusterização dos dados foi realizada pelo algoritmo *K-médias*, o qual agrupa pixels semelhantes, baseados em critérios de similaridades, de uma forma não supervisionada. Segundo a documentação contida na central de apoio do *software* MATLAB, o *K-médias* realiza o agrupamento dos dados a partir de uma matriz de dados ou observações. Os parâmetros utilizados nessa etapa da pesquisa correspondem ao *default* da própria função *K-médias* no MATLAB, sendo utilizada a métrica da Distância Euclidiana Quadrada onde cada centroide corresponde à média dos pontos em um determinado *cluster*, o número máximo de 100 iterações e o método para escolha do aglomerado inicial *plus*.

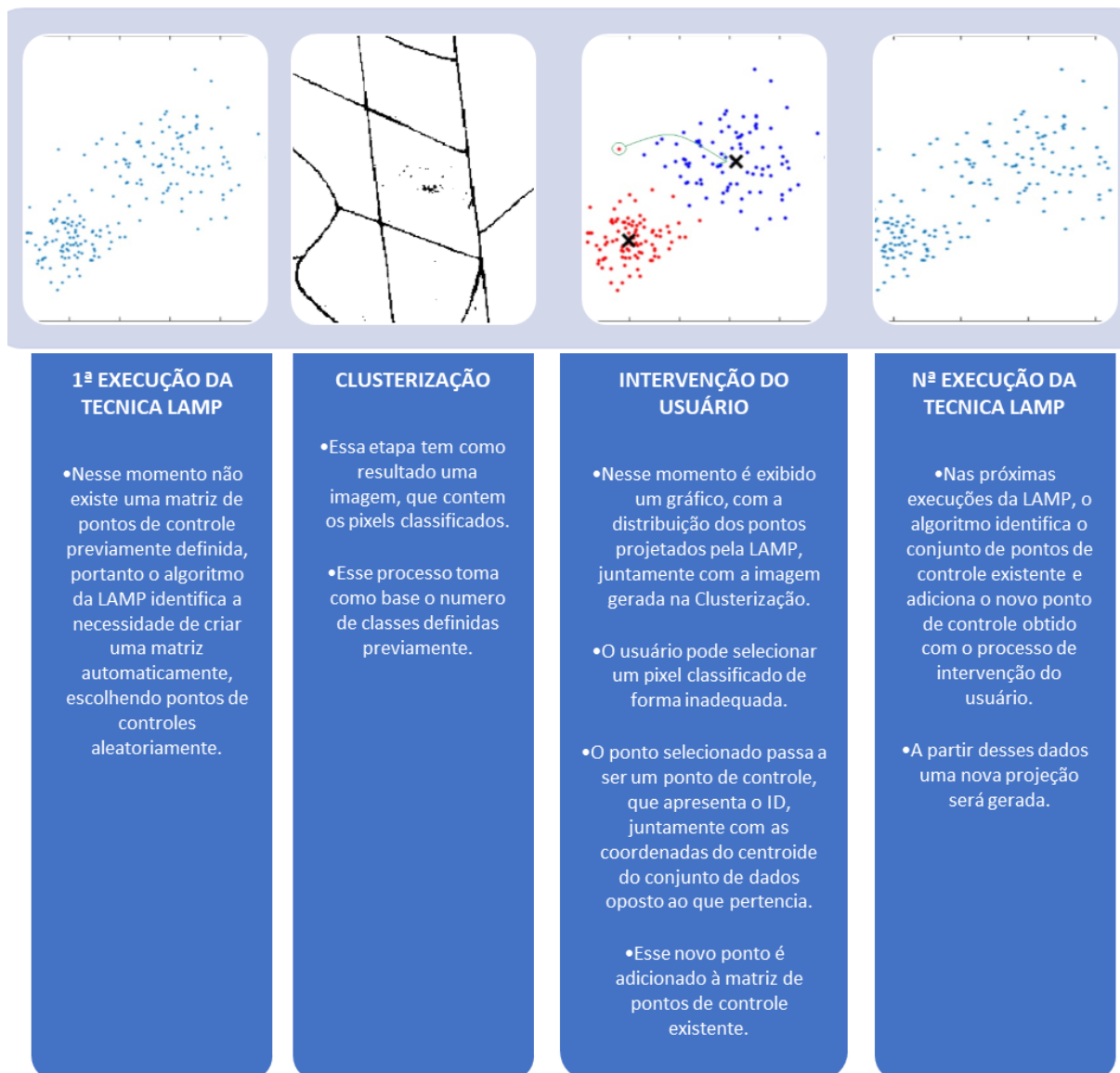
O passo seguinte do *pipeline* de processamento dos dados, foi pensado e projetado para ser o principal diferencial do trabalho, pois este está relacionado diretamente com prover ao usuário a possibilidade de interagir com o processo de classificação, procurando melhorar o resultado obtido. Desta forma, a estrutura de operação da técnica de processamento LAMP foi projetada utilizando-se os pontos

de controle, onde estes podem ser obtidos de duas formas distintas. Na primeira forma de entrada, que foi a utilizada no início do processamento, a técnica LAMP identifica que não existe nenhum conjunto de pontos de controle e a partir disso, esses pontos são gerados, automaticamente e aleatoriamente, pelo algoritmo. A partir destes pontos de controle o processo de classificação é realizado, conforme apresentado nos parágrafos anteriores.

Após a primeira classificação oriunda do algoritmo *K-médias*, no espaço de alta dimensão, o usuário tem a possibilidade de interagir com o resultado da projeção, mediante à manipulação dos pontos de controle gerados inicialmente pela LAMP. Desta maneira, é possível adicionar novos pontos de controle a partir da projeção já realizada, com o intuito de reclassificar possíveis *clusters* indexados de forma equivocada pela classificação inicial, essa é entendida então como a segunda forma de obtenção dos pontos de controle.

Essa possibilidade de intervenção pode ser repetida quantas vezes forem necessárias, a fim de que o usuário possa obter uma classificação mais próxima do real possível, sendo que, a cada novo ponto de controle adicionado ao conjunto de dados, gera-se uma influência em toda sua vizinhança, o que contribui na reclassificação de toda a vizinhança do ponto adicionado. O esquema apresentado na Figura 10 auxilia na compreensão dos passos descritos nos cinco parágrafos anteriores.

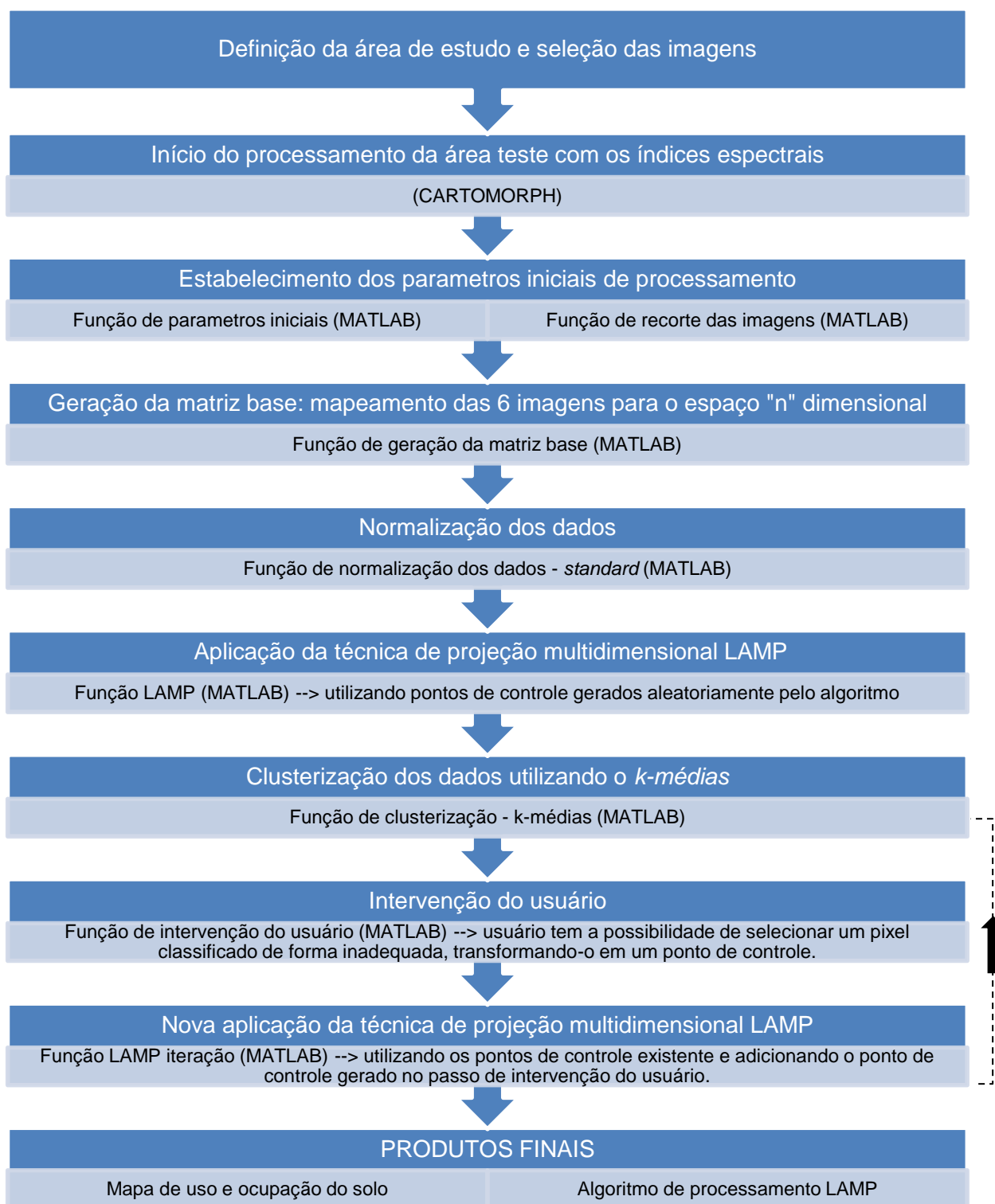
Figura 10: Sequência parcial de processamento utilizada no módulo de interação com o usuário.



Fonte: Autor

O esquema ilustrativo contido na Figura 11 apresenta uma síntese da metodologia desenvolvida em todo o trabalho e tem o intuito de facilitar o entendimento do processo como um todo.

Figura 11: Etapas de trabalho para o desenvolvimento da pesquisa.



Fonte: Autor

4. RESULTADOS E DISCUSSÃO

A partir do que foi exposto na metodologia, foi possível gerar os resultados da pesquisa, juntamente com as análises e considerações a respeito de todo o trabalho.

4.1. APLICAÇÃO DOS ÍNDICES ESPECTRAIS

Conforme comentado, a aplicação dos índices espectrais foi realizada por meio do *software* CARTOMORPH. As imagens resultantes da aplicação dos índices espectrais RVI, NDVI e NDWI, podem ser observadas nas Figura 12, Figura 13 e Figura 14, respectivamente.

Figura 12: Resultado da aplicação do índice RVI sobre a imagem teste.

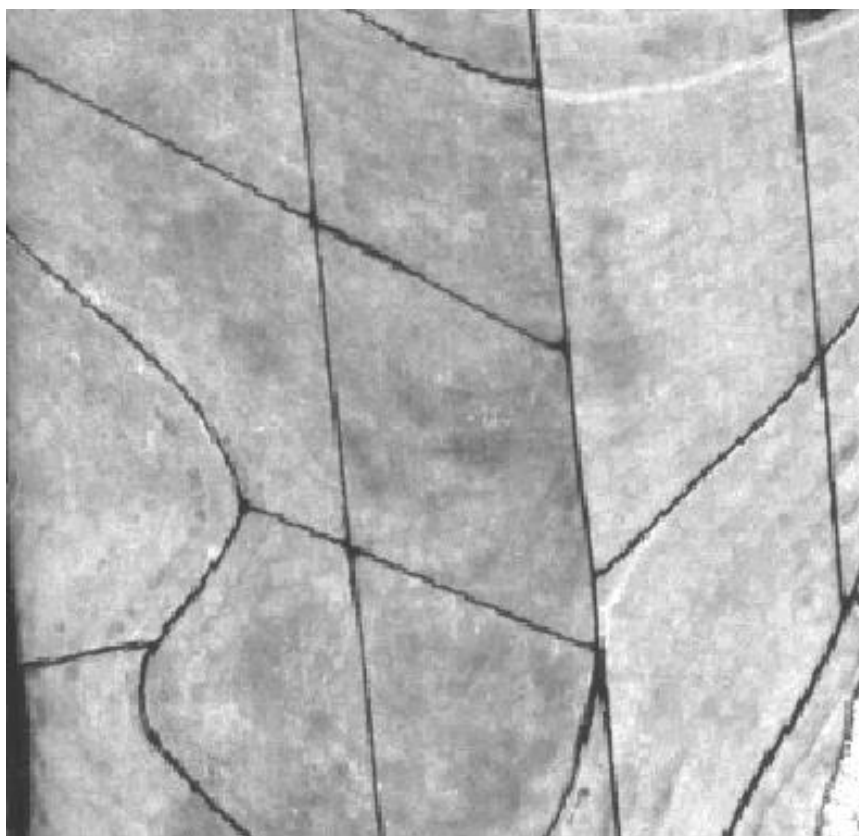
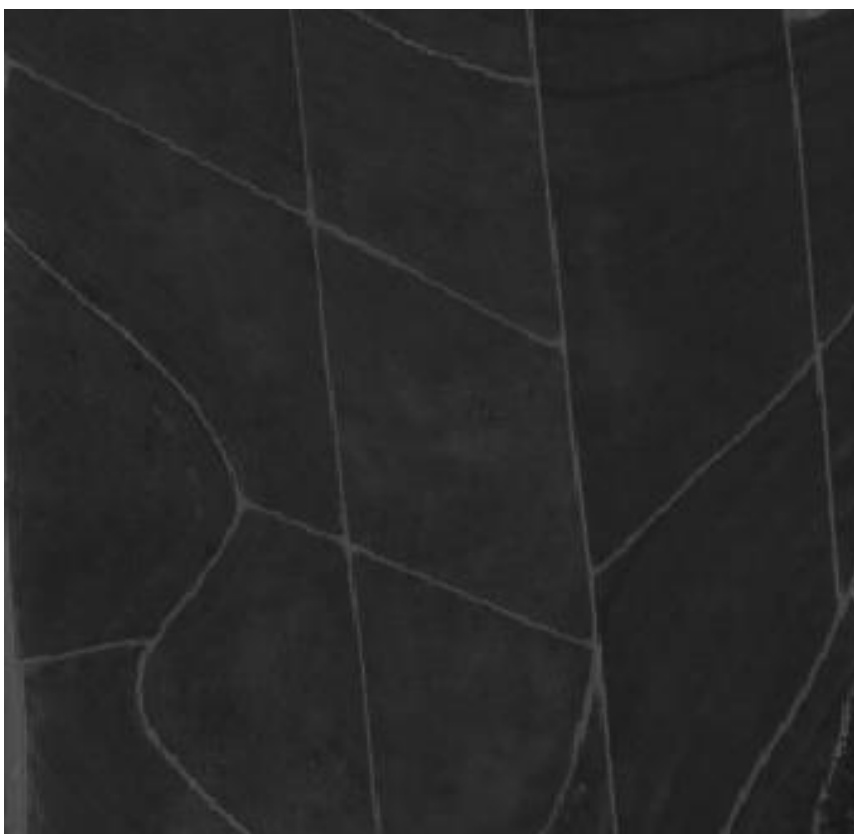


Figura 13: Resultado da aplicação do índice NDVI sobre a imagem teste.



Figura 14: Resultado da aplicação do índice NDWI sobre a imagem teste.



4.2. PROCESSAMENTOS INICIAIS

De posse das imagens resultantes da aplicação dos índices espectrais, foi iniciado o processamento com o *software* MATLAB, onde as funções foram estruturadas.

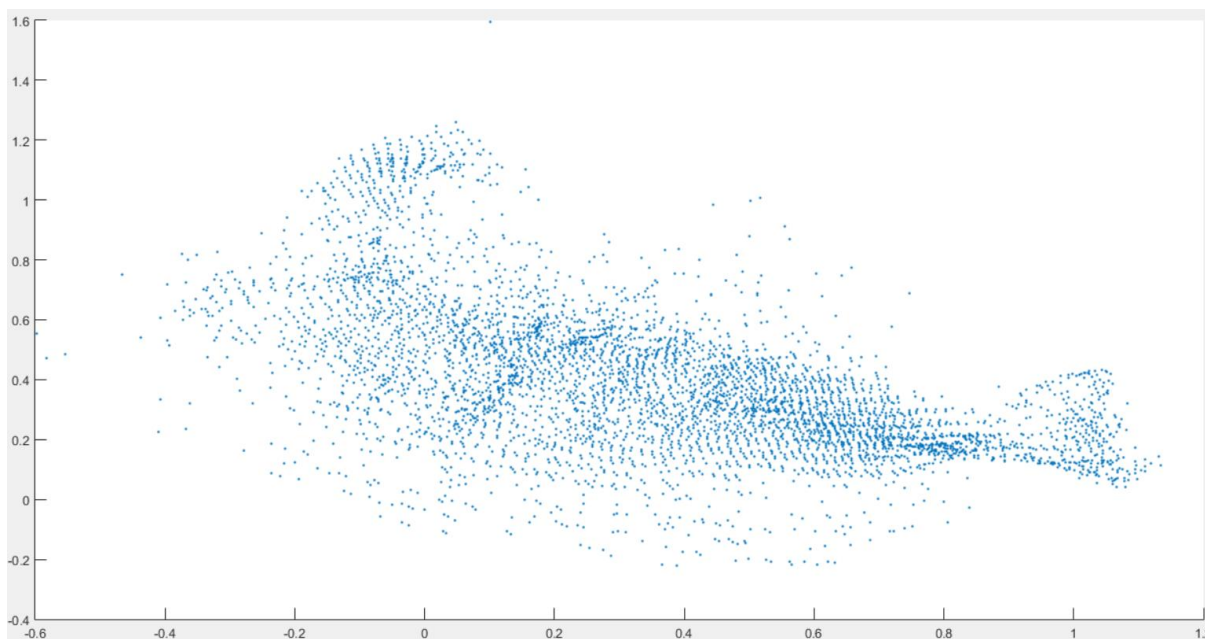
As duas primeiras funções correspondem aos parâmetros iniciais utilizados para desenvolver o processamento e a preparação das imagens, para posterior processamento. A primeira função desenvolvida possibilita que o usuário defina um *bounding box* para o processamento, delimitando uma região de interesse a ser processada, além de definir o número de classes que se deseja obter com o final do processo. Nesse estudo foram utilizadas 2 classes (vegetação e não vegetação), conforme explicado na metodologia. A segunda função, faz a leitura das camadas (seis imagens), confrontando-as com *bounding box* definido, para que seja considerada apenas a região de interesse.

A próxima etapa, utiliza as camadas, levando-as ao espaço “n” dimensional, para a geração da matriz base de processamento. Com a matriz base gerada, foi feito o processo de normalização dos dados, sendo que, conforme comentado na metodologia, foi utilizada a normalização de dados *standard*, para gerar a matriz base normalizada.

4.3. PROJEÇÃO MULTIDIMENSIONAL LAMP

A partir da matriz de dados normalizada, foi possível aplicar a técnica de projeção multidimensional (LAMP). A função gerada, busca o algoritmo da projeção multidimensional LAMP, e como ainda não existem pontos de controle previamente definidos, gera-os automaticamente, armazenando-os em variáveis padrão, definidas como “ctrlpts” e “y_ctrlpts”. O gráfico resultante dessa técnica de projeção pode ser observado na Figura 15.

Figura 15: Gráfico contendo os dados resultante da técnica de projeção LAMP (os eixos do gráfico se encontram na mesma escala da normalização).



Fonte: Autor

4.4. CLUSTERIZAÇÃO

Após a projeção ser concluída, foi feita a clusterização dos dados obtidos, utilizando o método *K-médias*, que agrupa os pixels semelhantes, de uma forma não supervisionada. Nessa clusterização foi utilizada uma quantidade de classes previamente definida, no início do processamento, por meio da variável “*nclusters*” da função de parâmetros iniciais.

Vale lembrar que nesse estudo, foram utilizadas 2 classes para gerar a clusterização, sendo que uma delas está relacionada com as áreas que contém vegetação e a outra está relacionada com as áreas que não contém vegetação. Esse critério de duas classes foi estabelecido observando uma demanda do grupo Raízen Energia, que busca classificar e identificar de forma automatizada, fazendas de cana-de-açúcar, separando principalmente áreas com vegetação (cana-de-açúcar), das áreas sem vegetação (carreadores, tanques de vinhaça, corpos d’água, rodovias, dentre outros). A Figura 16 apresenta a imagem resultante do processo de clusterização, gerado sem a intervenção do usuário.

Figura 16: Imagem resultante do processo de clusterização, gerado sem a intervenção do usuário (branco representa regiões com vegetação e preto representa regiões com solo exposto).

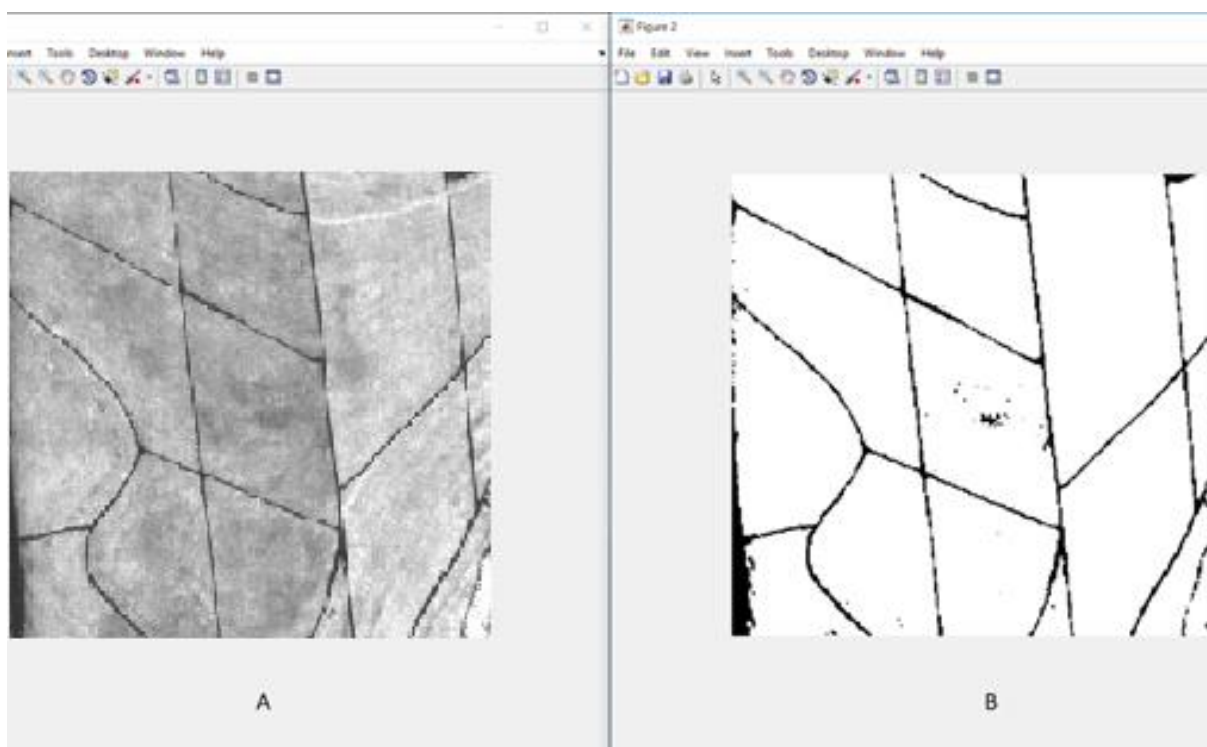


4.5. INTERVENÇÃO DO USUÁRIO

A partir dos dados obtidos com os processos de projeção multidimensional e clusterização, foi possível permitir ao usuário o poder de ajustar áreas classificadas de forma incorreta pelo algoritmo. Assim o usuário pode contribuir com sua percepção visual da cena, e por meio da seleção de pixels classificados de forma inadequada, influenciar no resultado obtido. Essa forma de identificação é feita utilizando uma imagem referência (Figura 17 – A), correspondente ao índice espectral RVI (selecionada por apresentar bom realce da vegetação) e a imagem resultante do processo de clusterização mais recente (Figura 17 – B). Assim o

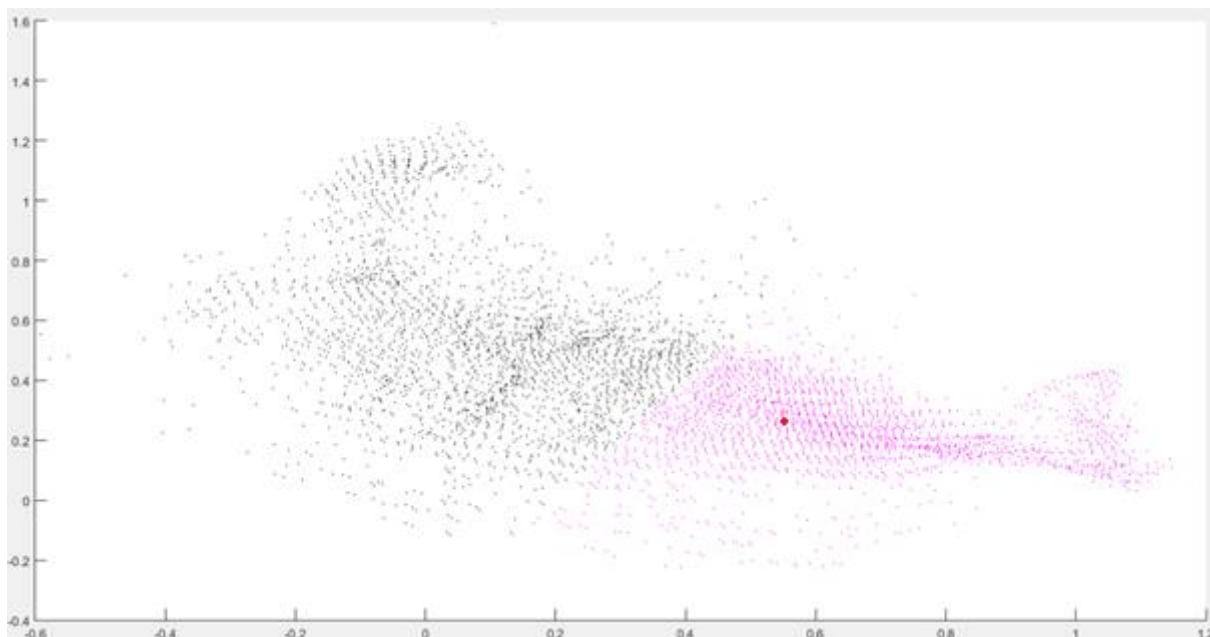
usuário poderá utilizar a imagem referência para “auditar” o processo de clusterização e a partir da identificação visual de um pixel que ele julgar estar classificado equivocadamente, selecioná-lo na imagem correspondente a clusterização (Figura 17 – B).

Figura 17: Processo de intervenção do usuário na clusterização.



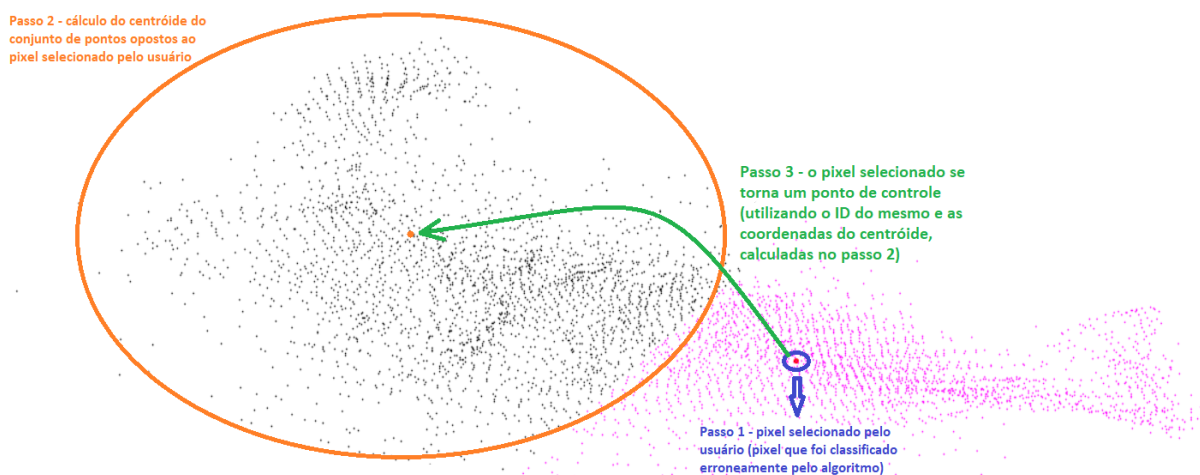
Com a seleção do pixel realizada, o algoritmo irá apresentar ao usuário, um gráfico que contém o pixel selecionado (definido pela cor vermelha na Figura 18), juntamente com os demais pontos da projeção. Os pontos da projeção que são mostrados, estão divididos entre os dois *clusters* definidos anteriormente, ou seja, um conjunto de pontos representa os pixels classificados como vegetação e o outro conjunto de ponto representa os pixels classificados como não vegetação. Essa relação de correspondência colorida, é possível de ser gerada a partir da identificação feita entre os dados da projeção e os dados clusterizados, portanto, ambos os dados apresentam um “ID” de controle/identificação, e a partir desse “ID” é possível fazer essa correlação.

Figura 18: Gráfico contendo o pixel selecionado (em vermelho), e os demais pontos da projeção (vegetação e não vegetação) - (os eixos do gráfico se encontram na mesma escala da normalização).



Com a identificação e amostragem do ponto, o algoritmo converte automaticamente o pixel selecionado, em um ponto de controle, atribuindo as coordenadas do centróide da região oposta que ele pertencia anteriormente. Para que seja possível compreender melhor esse processo, foi gerado um gráfico didático (apresentado na Figura 19), adaptado do gráfico obtido do processamento aqui descrito.

Figura 19: Gráfico didático utilizado para compreender melhor o processo de geração do novo ponto de controle (baseado no pixel selecionado pelo usuário).



Fonte: Autor

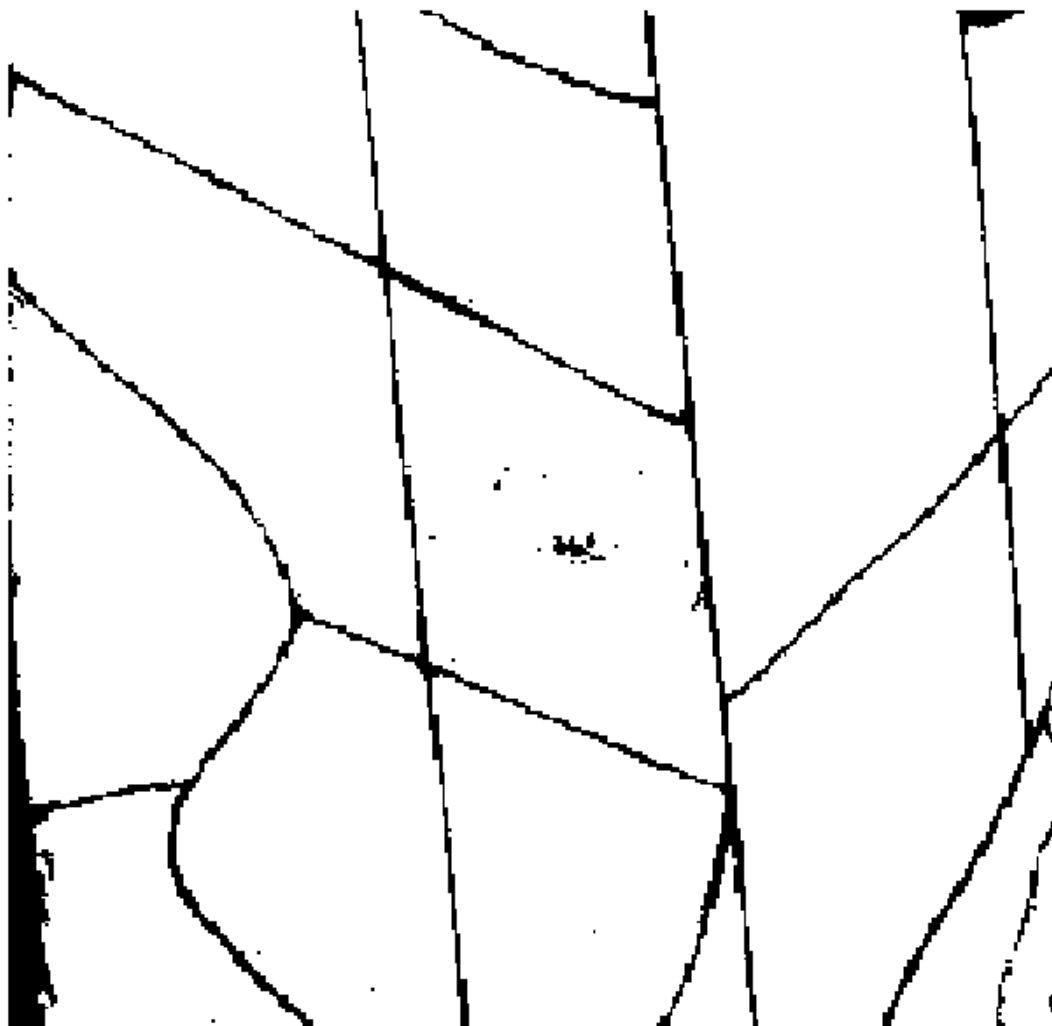
No gráfico apresentado na Figura 19, foi desenvolvido um esquema baseado em três passos, para compreender melhor como a intervenção do usuário irá agir no processamento. No primeiro passo, tem-se o pixel selecionado pelo usuário, ou seja, aquele que o usuário julgou estar classificado de forma inadequada. O segundo passo é responsável por fazer uma média das coordenadas dos eixos X e Y, do conjunto de dados oposto ao que pixel selecionado pertence, ou seja, calcula-se o centroide do conjunto de dados oposto. No terceiro passo o pixel selecionado se torna um ponto de controle, com seu "ID" e as coordenadas calculadas no segundo passo.

É importante observar que os pixels selecionados para serem corrigidos, afetam sua vizinhança, por meio da projeção multidimensional LAMP. Isso acontece pelo fato dos pixels selecionados se tornarem pontos de controle, que são adicionados aos pontos de controle existentes. Portanto a cada seleção de um novo pixel, ou seja, a cada processo de intervenção do usuário, é realizado um novo processo da projeção LAMP e o resultado obtido, é exibido ao usuário novamente por meio da clusterização, para que, se o usuário identificar outro pixel classificado de forma inadequada, ele possa corrigi-lo.

Para ser possível utilizar os pontos de controle gerados anteriormente, juntamente com o novo ponto de controle gerado pelo usuário, foi desenvolvida a função LAMP iteração, que utiliza os pontos de controle criado na primeira execução da LAMP, adicionando os novos pontos de controle que foram gerados a cada processo de intervenção executado pelo usuário. Portanto a função LAMP iteração não gera aleatoriamente os pontos de controle, como acontece na função LAMP, e sim utiliza os pontos existentes, adicionando os novos pontos de controle gerados pelos repetidos processos de intervenção do usuário.

O resultado obtido com a intervenção do usuário pode ser observado na Figura 20 e a respectiva análise, pode ser observada na seção 4.7 DISCUSSÃO, na página 51, desta dissertação.

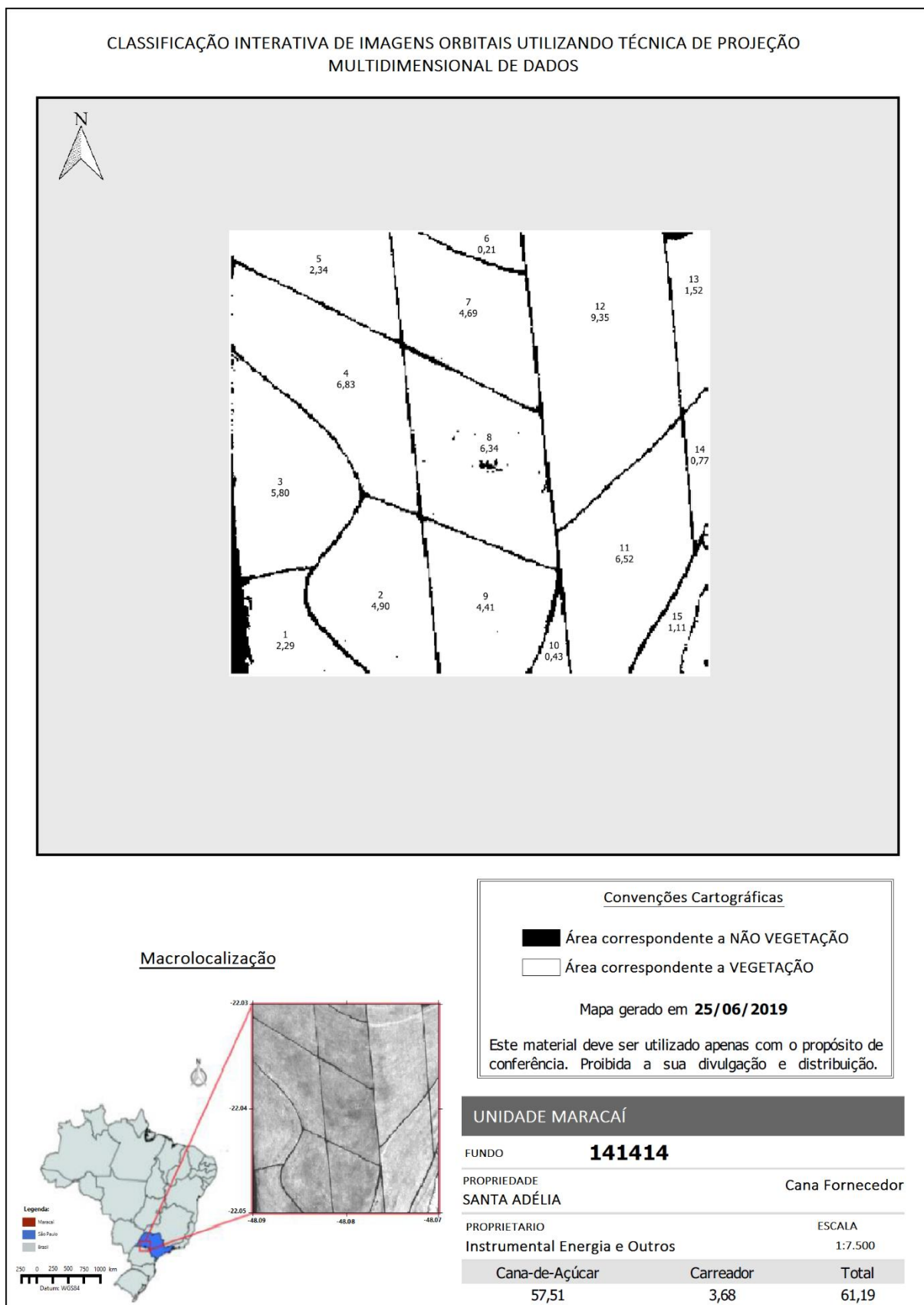
Figura 20: Imagem resultante do processo de clusterização, após a intervenção do usuário (branco representa regiões com vegetação e preto representa regiões com solo exposto).



4.6. PRODUTO CARTOGRÁFICO

Com a finalização do processamento, foi possível gerar o mapa que contém a classificação da área analisada, com as informações necessárias para seu uso em diversos estudos e uma macrolocalização da área estudada. Para o desenvolvimento deste mapa, foi utilizado o *software* Qgis, em sua versão 3.4. A Figura 21 apresenta o referido mapa, sendo importante ressaltar que as informações de “Fundo”, “Propriedade”, “Proprietário” e a grade de coordenadas, são dados fictícios, criados/modificados para proteger a identidade do proprietário do local. As áreas dos talhões evidenciadas no mapa, estão na unidade hectares.

Figura 21: Produto cartográfico gerado com o processamento.



Fonte: Autor

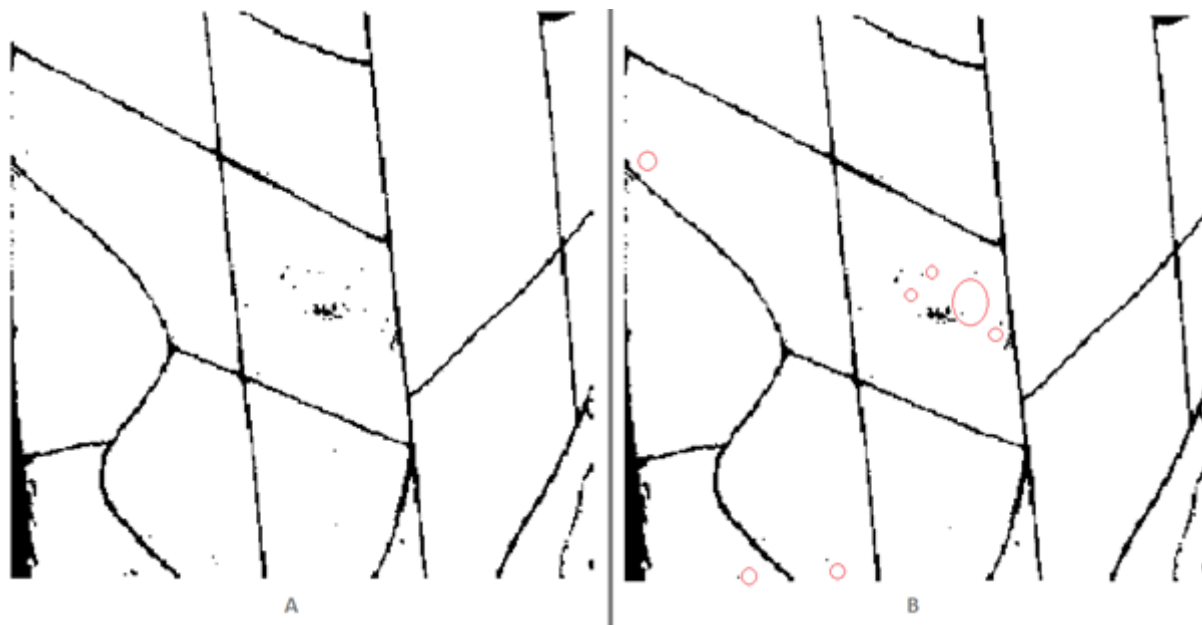
4.7. DISCUSSÃO

Observando inicialmente os resultados gerados pelos índices espectrais, foi possível verificar que, para os índices de vegetação, nos locais onde há ocorrência da vegetação, encontram-se os pixels que apresentam tons mais claros e nos locais com ausência de vegetação, como os solos expostos de carreadores, construções e corpos d'água, encontram-se os pixels que apresentam tons mais escuros, o que, de acordo com Schowengerdt (2007) e também com Meneses e Almeida (2012), é esperado. Para o índice de corpos d'água, observa-se que nos locais onde há ocorrência da vegetação, encontra-se os pixels que apresentam tons mais escuros e nos locais onde há ocorrência de solos expostos encontram-se os pixels que apresentam tons mais claros, o que, de acordo com Mcfeeters (1996), também é esperado.

As observações relatadas no parágrafo anterior são importantes para o processo da classificação desenvolvido neste estudo, pois as informações de ND dos pixels (tons de cinza) são utilizadas para promover a correta classificação dos dados. Assim, quando os pixels são levados ao espaço "n" dimensional, aplicado a técnica de projeção multidimensional LAMP e clusterizados, os índices espectrais tendem a contribuir no processo, devido a diferenciação de feições específicas como vegetação, corpos d'água, solo exposto e construções diversas.

Analisando a intervenção do usuário, foi possível verificar, por meio dos resultados obtidos, que houve melhorias no processo de classificação. Para ser possível observar esse ganho, colocou-se o resultado obtido antes da intervenção do usuário (Figura 22 - A) e o resultado obtido após a intervenção do usuário (Figura 22 - B), lado a lado, procurando destacar as regiões que sofreram alterações com o processo de intervenção.

Figura 22: Imagens resultantes do processamento, destacando as regiões que sofreram alterações com o processo de intervenção do usuário.



Na Figura 22 observam-se, destacados por círculos vermelhos, os locais onde ocorreram alterações com o processo de intervenção do usuário, sendo importante observar que, no geral, a intervenção do usuário contribuiu positivamente na classificação, ou seja, pixels que haviam sido classificados de forma incorreta, passaram a ser classificados de forma correta, de acordo com a verdade terrestre, que foi possível verificar pela composição colorida gerada na página 33, desta dissertação. Vale ressaltar também que foram feitos 4 processos de intervenção do usuário e os pixels selecionados nesses processos influenciaram pixels vizinhos, o que contribuiu para obtenção de um bom resultado.

O algoritmo de processamento, que será disponibilizado aos usuários sob demanda, foi construído de uma forma simples e objetiva. Assim os usuários poderão utilizá-lo, realizando adaptações para suas eventuais necessidades, de uma forma fácil e intuitiva. Portanto a divisão das funções na estrutura do algoritmo, contribuiu para o seu melhor entendimento e manutenção.

O mapa gerado contém as convenções cartográficas necessárias, podendo ser utilizado para o desenvolvimento de diversas atividades, como exemplo destaca-se a localização e estimativa para produção da cana-de-açúcar da fazenda.

5. CONSIDERAÇÕES FINAIS

Com o propósito de estudar alternativas para os processos de classificação em imagens digitais, esta pesquisa de mestrado foi desenvolvida, com foco na aplicação de uma técnica de projeção multidimensional de dados, já consolidada na área da visualização da informação, porém pouco explorada na área da cartografia. A técnica em questão é a *Local Affine Multidimensional Projection*, mais conhecida como *LAMP*.

Com a finalização da pesquisa, foi possível verificar que os valores de ND dos pixels, são as informações básicas para todo o desenvolvimento do processamento. Portanto, a utilização das imagens multiespectrais (correspondente as bandas proveniente do sensor do satélite) aliada as informações provenientes dos índices espectrais, contribuíram positivamente com os resultados obtidos; uma vez que os índices espectrais promovem realce de feições específicas, que no caso desta pesquisa estavam centradas na vegetação e solo exposto. Assim, a correta escolha dos índices espectrais utilizados, foi importante para alcançar um bom resultado de processamento, sendo altamente aconselhável conhecer os alvos de interesse, para escolher os melhores índices espectrais que devem ser utilizados.

O processo de intervenção do usuário contribuiu de forma positiva para a obtenção do resultado final, porém, vale ressaltar que este procedimento pode gerar resultados inversos ao almejado, caso o usuário não tenha total domínio ou conhecimento sobre quais pixels estão classificados de forma incorreta. Portanto é importante mencionar que o usuário deve estar familiarizado com a área e saber interpretar corretamente a imagem gerada, a fim de não contribuir de forma inadequada para o processamento, selecionando pixels que estejam classificados corretamente.

Conforme dito nas considerações iniciais do trabalho, os produtos da pesquisa foram o algoritmo de processamento e o mapa. Portanto foi possível concluir que os objetivos estabelecidos foram alcançados, sendo disponibilizado, sob demanda, o algoritmo de processamento, desenvolvido no ambiente do MATLAB e a geração do produto cartográfico (mapa). Assim observa-se que a técnica de processamento utilizada (LAMP), pode colaborar para trabalhos de

atualizações de bases cartográficas e outros, que estejam relacionados com a classificação de imagens de satélite.

Algumas considerações podem ser feitas sobre possibilidades futuras de aprimoramento deste estudo. Uma possível contribuição, seria a implementação do processamento em uma linguagem de programação que consiga trabalhar de forma mais rápida com o grande volume de dados, contribuindo principalmente para a atenuação do tempo de processamento, visto que, neste caso, o processamento completo demorou cerca de 3 a 4 horas para ser realizado. Uma possível alternativa seria a linguagem de programação C++. Além disso, estudar outras abordagens de classificação como a proposta por Negri et al. (2018), ou ainda outras técnicas de projeção multidimensional de dados, pode ser algo importante no sentido de contribuir para o desenvolvimento de processos mais eficazes e rápidos. Portanto, uma sugestão seria a reavaliação do processo com a técnica de projeção *Least Square Projection* (LSP).

REFERÊNCIAS BIBLIOGRÁFICAS

AGUIAR, T. O.; AZEVEDO, S. C.; PEDROSA, M. M.; CARDIM, G. P.; SILVA, E. A. **Multispectral Image Processing System Developed**. In CARTOMORPH Software - NDVI Module. *Advances in Remote Sensing*, v. 07, p. 91-100, 2018.

AGUIAR, T. O.; AZEVEDO, S. C.; SILVA, E. A. **Aplicação do índice espectral de corpos d'água - water index (WI) implementado no software CARTOMORPH**. In: Simposio Internacional SELPER 2016, 2016, Puerto Iguazu. *Actas do Simposio Internacional SELPER 2016*. Puerto Iguazu: SELPER, 2016. p. 1295-1306.

ALSABTI, K.; RANKA, S.; SINGH, V. **An efficient k-means clustering algorithm**. *Electrical Engineering and Computer Science*. 1997.

BARBOSA, C. C. F.; NOVO, E. M. L. M.; CARVALHO, J. C.; PEREIRA FILHO, W.; MANTOVANI, J.E. **Caracterização spectral das massas d'água amazônicas**. *Anais do Simpósio Brasileiro de Sensoriamento Remoto - 2003*. São José dos Campos: INPE, p. 1099-1106. 2003

BERKHIN, P. **A Survey of Clustering Data Mining Techniques**. 2002.

CAMARA, G.; DAVIS, C.; MONTEIRO, A. M. V. **Introdução à Ciência da Geoinformação**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2001.

CARDIM, G. P. **Desenvolvimento do Sistema Computacional CARTOMORPH para Processamento de Imagens de Sensoriamento Remoto**. Tese de Mestrado em Ciências Cartográficas—Presidente Prudente: Universidade Estadual Paulista “Júlio de Mesquita Filho” - UNESP, 2015.

CASACA, W. NIETO, E. G. FERREIRA, C. O. L. TAVARES, G. PAGLIOSA, P. PAULOVICH, F. NONATO, L. G. PAIVA, A. **Colorization by Multidimensional Projection**. 25th SIBGRAPI Conference on Graphics, Patterns and Images. Ouro Preto, Brazil. 2012.

ETEMADPOUR, R.; LINSEN, L.; PAIVA, J. G.; CRICK, C.; FORBES, A. G. **Choosing Visualization Techniques for Multidimensional Data Projection Tasks: A Guideline with Examples**. *Computer Vision, Imaging and Computer Graphics Theory and Applications*. Anais. Springer International Publishing, 2016.

FIGUEIREDO, D. **Conceitos Básicos de Sensoriamento Remoto**, 2005.

FIORIO, P. R. **Tópicos Básicos de Comportamento Espectral de Alvos**. (Material Didático). 2004.

FLORENZANO, T. G.; JÚNIOR, S. S. T.; LORENA, R. B.; MELO, D. H. C. T. B. **Multiplicação e Adição de Imagens Landsat no Realce de Feições da Paisagem.** Anais do X Simpósio Brasileiro De Sensoriamento Remoto, p. 1257–1263, 2001.

FONTANA, A.; NALDI, M. C. **Estudo de Comparação de Métodos para Estimação de Números de Grupos em Problemas de Agrupamento de Dados.** Universidade de São Paulo. 2009.

GAO, B. **NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space.** Remote Sensing of Environment, v. 58, n. 3, p. 257–266, 1 dez. 1996.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (ED.). **Introdução ao Processamento Digital de Imagens.** Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 2000.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. **Data clustering: a review.** p. 264–323. 1999.

JENSEN, J. R. **Introductory digital image processing: a remote sensing perspective.** Ed 2. Prentice Hall. 2004

JOIA, P.; COIMBRA, D.; CUMINATO, J.; PAULOVICH, F.; NONATO, L. G. **Local affine multidimensional projection.** IEEE Transactions on Visualization and Computer Graphics, v. 17, n. 12, p. 2563–2571, dez. 2011.

KIRK, J. T. O. **Light & photosynthesis in aquatic ecosystems.** Cambridge University Press, 2ed. 509p. 1994.

KOGAN, J.; NICHOLAS, C.; TEBoulLE, M. **Grouping Multidimensional Data: Recent Advances in Clustering.** 1. ed. [s.l.] Springer-Verlag Berlin Heidelberg, 2006.

KRISHNA, K.; MURTY, M. N. **Genetic K-Means Algorithm.** IEEE Transactions on Systems, Man, and Cybernetics. Cybernetics, Vol. 29, No. 3. June 1999.

LILLESAND, T.; KIEFER, R. W.; CHIPMAN, J. **Remote Sensing and Image Interpretation.** 5. ed. New Jersey: John Wiley & Sons, 2004.

MACHADO, B. B. **Exploração visual do espaço de características: uma abordagem para análise de imagens via projeção de dados multidimensionais.** text—[s.l.] Universidade de São Paulo, 13 dez. 2010.

MACQUEEN, J. B. **Some methods for classification and analysis of multivariate observations.** Proceedings of the Fifth Symposium on Math, Statistics and Probability. Berkeley, CA: University of California Press. p. 281-297. 1967.

MAMANI, G. M. H. **Empregando técnicas de projeção multidimensional para transformação interativa de espaços de características.** text—[s.l.] Universidade de São Paulo, 24 out. 2012.

MARQUES, F. O.; VIEIRA, N. H. **Processamento Digital de Imagens.** Rio de Janeiro: Brasport, 1999. ISBN 8574520098.

MATHER, P. M. **Computer Processing of Remotely Sensed Images: An Introduction.** 3. ed. University of Nottingham: John Wiley & Sons Ltd, 2004.

MathWorks - Documentation. Disponível em: <https://www.mathworks.com/help/stats/kmeans.html> - Acessado em 01 de maio de 2019.

MCFEETERS, S. K. **The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features.** International Journal of Remote Sensing, v. 17, n. 7, p. 1425–1432, 1 maio 1996.

MENESES, P. R.; ALMEIDA, T. DE. **Introdução ao Processamento de Imagens de Sensoriamento Remoto.** Brasília: Universidade de Brasília, 2012.

NEGRI, R. G.; SILVA, E.; CASACA, W. **Inducing Contextual Classifications with Kernel Functions Into Support Vector Machines.** IEEE Geoscience and Remote Sensing Letters, v. 15 (6), 962-966, 2018.

NEVES, T. T. DE A. T.; FADEL, S. G.; COIMBRA, D. B.; PAULOVICH, F. V. **Análise visual utilizando projeções multidimensionais.** Revista de Informática Teórica e Aplicada - RITA, v. 22, n. 2, p. 258–288, 2015.

NIETO, E. G.; CASACA, W.; MOTTA, D.; HARTMANN, I.; TAUBIN, G.; NONATO, L. **Dealing with Multiple Requirements in Geometric Arrangements.** IEEE Transactions on Visualization and Computer Graphics, v. 22 (3), 1223-1235, 2016.

NIETO, E. G.; ROMAN, F. S.; PAGLIOSA, P.; CASACA, W.; HELOU, E. S.; DE OLIVEIRA, M. C. F.; NONATO, L. **Similarity Preserving Snippet-Based Visualization of Web Search Results.** IEEE Transactions on Visualization and Computer Graphics, v. 20 (3), 457-470, 2014.

NIETO, E. M. G. **Projeção multidimensional aplicada a visualização de resultados de busca textual.** text—[s.l.] Universidade de São Paulo, 30 ago, 2012.

NOVO, E. M. L. DE M. **Sensoriamento Remoto: Princípios e Aplicações.** 3. ed. São Paulo: Edgard Blucher Ltda, 2008.

PAPPU, V.; PARDALOS, P. M. **High-Dimensional Data Classification.** In: Clusters, Orders, and Trees: Methods and Applications. Springer Optimization and Its Applications. [s.l.] Springer, New York, NY, 2014. p. 119–150.

PAULOVICH, F. V.; NONATO, L. G.; MINGHIM, R.; LEVKOWITZ, H. **Least Square Projection: A Fast High-Precision Multidimensional Projection Technique and Its Application to Document Mapping**. IEEE Transactions on Visualization and Computer Graphics, v. 14, n. 3, p. 564–575, jun. 2008.

PONZONI, F. J.; SHIMABUKURO, Y. E. **Sensoriamento Remoto no estudo da Vegetação**. Ed. São Jose dos Campos: Parentese. 2010

SCHOWENGERDT, R. A. **Remote Sensing: Models and Methods for Image Processing**. 3. ed. University of Arizona: Academic Press, 2006.

SIEDLECKI, W.; SIEDLECKA, K.; SKLANSKY, J. **An overview of mapping techniques for exploratory pattern analysis**. Pattern Recognition, v. 21, n. 5, p. 411–429, 1 jan. 1988.

SILVA, E. A. DA. **Extração de feições cartográficas de imagens multiespectrais fundidas**. Tese de Doutorado em Engenharia - São Paulo: Escola Politécnica da Universidade de São Paulo, 1995.

SILVA, E. A.; AGUIAR, T. O.; AZEVEDO, S. C.; CARDIM, G. P.; CASACA, W. C. O. **Aplicação do Índice Espectral de Vegetação (Ratio Vegetation Index) implementado no Software CARTOMORPH**. In: XVIII Simposio Latinoamericano de Percepción Remoto - SELPER, Cuba, 2018.

STEINBACH, M.; ERTÖZ, L.; KUMAR, V. **The Challenges of Clustering High Dimensional Data**. In: WILLE, L. T. (Ed.). New Directions in Statistical Physics: Econophysics, Bioinformatics, and Pattern Recognition. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004. p. 273–309.

TEJADA, E.; MINGHIM, R.; NONATO, L. G. **On Improved Projection Techniques to Support Visual Exploration of Multi-Dimensional Data Sets**. Information Visualization, v. 2, n. 4, p. 218–231, 1 dez. 2003.

WAGSTAFF, K.; CARDIE, C.; ROGERS, S.; SCHROEDL, S. **Constrained K-means Clustering with Background Knowledge**. Proceedings of the Eighteenth International Conference on Machine Learning. p. 577-584. 2001.