

DANILO PEREIRA BARBOSA

**AVALIAÇÃO DE DESCRITORES TEXTURAIS GEOESTATÍSTICOS E
DE HARALICK PARA O RECONHECIMENTO DE PLANTAS
DANINHAS**

Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Estatística Aplicada e Biometria,
para obtenção do título de *Magister
Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2009

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

B823a
2009

Barbosa, Danilo Pereira, 1984-

Avaliação de descritores texturais geoestatísticos e de Haralick para o reconhecimento de plantas daninhas / Danilo Pereira Barbosa. – Viçosa, MG, 2009. xii, 89f.: il. (algumas col.) ; 29cm.

Inclui apêndices.

Orientador: Nerilson Terra Santos.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 58-64.

1. Análise espacial (Estatística). 2. Sensoriamento remoto. 3. Geoestatística. 4. Processamento de imagens. 5. Plantas daninhas. I. Universidade Federal de Viçosa. II. Título.

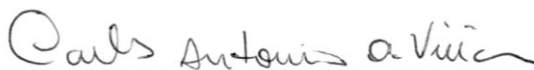
CDD 22.ed. 519.5

DANILO PEREIRA BARBOSA

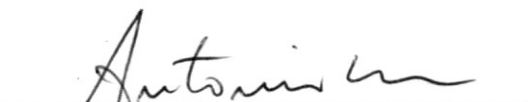
**AVALIAÇÃO DE DESCRITORES TEXTURAIS GEOESTATÍSTICOS E
DE HARALICK PARA O RECONHECIMENTO DE PLANTAS
DANINHAS**

Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Estatística Aplicada e Biometria,
para obtenção do título de *Magister
Scientiae*.

APROVADA: 17 de fevereiro de 2009.



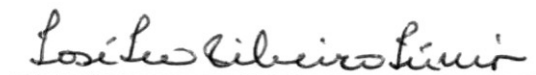
Carlos Antônio Oliveira Vieira



Antônio Policarpo Souza Carneiro



Sebastião Martins Filho



José Ivo Ribeiro Júnior



Nerilson Terra Santos
(Orientador)

À minha mãe Marly, e
À Deus
Com amor.

AGRADECIMENTOS

A Deus, por tudo.

A meus pais pela vida e aos meus amigos pelo apoio.

Ao professor orientador Nerilson Terra Santos e co-orientadores, Luiz Alexandre Peternelli e Francisco de Assis Carvalho Pinto pela orientação e amizade ao longo do curso.

À Universidade Federal de Viçosa, por intermédio do Programa de Pós Graduação de Estatística Aplicada e Biometria, pela oportunidade.

À Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), pela concessão da bolsa de estudos.

Aos demais colegas da pós-graduação, especialmente ao Antônio Carlos Fava de Barros pelo apoio nos apêndices e Melissa Pizaroglo de Carvalho, pelo apoio no andamento da dissertação.

Ao colega Altino Alves de Sousa Filho, pelas valiosas sugestões no decorrer do curso.

A todos que, de uma forma ou de outra, contribuíram para a realização deste trabalho.

BIOGRAFIA

DANILO PEREIRA BARBOSA, filho de Marly Pereira e Getúlio Oliveira Barbosa, nasceu em Rio Verde, Goiás, no dia 28 de agosto de 1984.

Em agosto de 2002 iniciou o curso de Ciências – Habilitação em Matemática na Universidade de Rio Verde – FESURV, Goiás, o qual concluiu em julho de 2006.

Em setembro de 2006, iniciou o curso de Mestrado em Estatística Aplicada e Biometria, na Universidade Federal de Viçosa, Minas Gerais, defendendo a tese em Fevereiro de 2009.

SUMÁRIO

	Páginas
RESUMO	vii
ABSTRACT	x
1. INTRODUÇÃO	13
1.1. Objetivo geral	15
1.2. Objetivos específicos	15
2. REVISÃO DE LITERATURA	16
2.1. Mapeamento das plantas daninhas	16
2.2. Reconhecimento de padrões	17
2.3. Análise textural	19
2.3.1. Tamanho de blocos	19
2.3.2. Descritores texturais	20
2.4. Teoria das variáveis regionalizadas	21
2.4.1. Aspectos gerais	21
2.4.2. Localização e continuidade	21
2.4.3. Função aleatória	22
2.4.4. Características das variáveis regionalizadas	23
2.5. Técnicas Geoestatísticas para classificação de imagens	27
2.6. Classificadores Estatísticos	28
2.6.1. Desempenho e avaliação de classificadores	31

2.6.2. Uso do método da validação cruzada	31
3. MATERIAL E MÉTODOS	33
3.1. Aquisição de imagens digitais	33
3.2. Pré-Processamento das imagens	35
3.3. Processamento	38
3.4. Classificação para as amostras de treinamento.....	45
3.4.1. Matriz de Erro	49
3.4.2. Índice Kappa.....	50
3.4.3. Avaliação dos descritores	52
3.4.4. Comparação dos descritores geoestatísticos com os de Haralick.....	52
3.4.5. Avaliação da combinação de descritores texturais	53
3.5. Validação do classificador	53
4. RESULTADOS E DISCUSSÕES.....	54
4.2. Definição das distâncias para os descritores texturais geoestatísticos	56
4.3.1. Variograma	57
4.3.2. Madograma	60
4.3.3. Variograma Cruzado.....	62
4.3.4. Pseudo variograma Cruzado	63
4.4. Descritores de Haralick.....	64
4.5. Comparação dos descritores geoestatísticos com os de Haralick.....	66
4.6. Avaliação da combinação de descritores texturais	67
4.7. Validação	68
5. CONCLUSÕES	69
6. REFERÊNCIAS BIBLIOGRÁFICAS	70
APÊNDICES	77

RESUMO

BARBOSA, Pereira Danilo, M.Sc., Universidade Federal de Viçosa, fevereiro de 2009. **Avaliação de descritores texturais geoestatísticos e de Haralick para o reconhecimento de plantas daninhas**. Orientador: Nerilson Terra Santos. Coorientadores: Francisco de Assis de Carvalho Pinto e Luiz Alexandre Peternelli.

A preocupação em minimizar a quantidade de produtos químicos utilizado em lavouras vem aumentando. O uso de sistemas de visão artificial tem demonstrado um grande potencial para o uso de taxas variadas de insumos, como por exemplo, a aplicação de herbicidas somente em locais onde é detectada a presença de planta daninha. O bom desempenho de um sistema desenvolvido para esta finalidade depende principalmente do uso de descritores que permitam diferenciar padrões de plantas daninhas do padrão da espécie cultivada. Sendo assim, objetivo geral do presente trabalho foi desenvolver e avaliar um descritor para o reconhecimento dos padrões planta de milho e planta daninha. Os objetivos específicos foram: a) identificar qual imagem, excesso de verde ou o índice de vegetação de verde normalizado, tende a proporcionar melhor classificação; b) comparar a classificação obtida por descritores geoestatísticos, com a obtida ao usar os descritores de Haralick. Com esta finalidade, foram adquiridas aos 29 dias após a emergência, período em que normalmente é feita a aplicação de herbicidas, nove imagens de milho (*Zea Mays L.*) e de três espécies de plantas daninhas avaliadas neste experimento: leiteira (*Euphorbia heterophylla L.*), capim-milhã (*Digitaria horizontalis Willd*) timbête (*Cenchrus echinatus L.*). Seis destas imagens foram utilizadas para a seleção do descritor que promove melhor desempenho na classificação. As três restantes foram utilizadas para a validação do descritor selecionado. Cada uma das seis imagens de treinamento foi recortada em 100 blocos de 68x68 pixels. Para cada um dos blocos foi obtido o valor dos descritores texturais geoestatísticos (variograma, o madograma, variograma cruzado e pseudo variograma cruzado) e os de Haralick (momento angular, média, variância, entropia, correlação, momento do produto, momento

inverso da diferença e medidas de correlação). Adicionalmente, descritores geoestatísticos e não-geoestatísticos foram obtidos considerando diferentes ângulos (0, 45, 90 e 135°) de relacionamento entre pixels. Descritores geoestatísticos foram, também, obtidos para diferentes distâncias (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) de pareamento entre pixels. Os descritores variograma e madograma foram calculados partir da imagem excesso de verde e GNDVI. Já os descritores variograma cruzado e pseudo variograma cruzado foram calculados com o uso do Greenness Method nos blocos usando as combinações das bandas RxG, GxB e IVxG. Os descritores de Haralick foram calculados a partir das imagens do excesso de verde e GNDVI. O desempenho dos descritores, assim propostos, foi avaliado usando análise discriminante. Os descritores selecionados foram aqueles que apresentaram maior valor para o índice kappa. Adicionalmente, novos descritores foram obtidos a partir de combinações dos descritores selecionados. Estas combinações, também, tiveram o seu desempenho avaliado usando a análise discriminante com o objetivo de identificar qual combinação proporciona melhor desempenho na classificação. Posteriormente, o poder de generalização da combinação selecionada foi avaliado usando as três imagens de cada espécie reservadas para a etapa de validação. As conclusões obtidas com relação aos objetivos propostos nesta pesquisa foram a) a imagem que tendeu a apresentar os melhores resultados do índice kappa foi a imagem excesso de verde; b) os descritores obtidos a partir da função madograma e os de Haralick foram os que forneceram os melhores resultados; c) o descritor geoestatístico madograma nas 10 distâncias e ângulo 0° apresentou melhores resultados de classificação quando usado sem combinação de outros descritores; d) os descritores geoestatísticos e os de Haralick, quando usados isoladamente não apresentaram resultados tão bons quanto combinados; e) o uso de descritores que consideram a continuidade dos valores de pixel, no reconhecimento de padrões pode ser uma ferramenta fundamental no processo de classificação.

ABSTRACT

BARBOSA, Pereira Danilo, M.Sc., Universidade Federal de Viçosa, February, 2009. **Evaluation of geostatistic textural descriptor and of Haralick for the recognition of harmful plants**. Adviser: Nerilson Terra Santos. Co-advisers: Francisco de Assis de Carvalho Pinto and Luiz Alexandre Peternelli.

The concern in minimizing the amount of chemical products used in farmings is increasing. The use of artificial vision systems has been demonstrating a great potential for use of varied taxes of inputs, as for instance, the application herbicides only in places where the presence of harmful plant is detected. The good acting of a system developed for this purpose depends mainly of the characteristic use that they allow to differentiate patterns of harmful plants of the pattern of the cultivated species. Like this, the objective of the present work was to develop and to evaluate a characteristic for the recognition of the patterns corn plant and harmful plant. The specific objectives were: the) to identify which image, green excess or the index of vegetation of normalized green, tends to provide better classification; b) to compare the classification obtained by characteristics geoestatistics, obtained when using the characteristics of Haralick. With this purpose, were acquired to the 29 days after the emergency, period in that it is usually made the application of herbicides, nine corn images (*Zea Mays* L.) and of each one of the species of appraised harmful plants in this experiment: *Euphorbia heterophylla* L., *Digitaria horizontalis* Willd, *Cenchrus*

echinatus L. Six of these images were used for the selection of the characteristic that promotes better acting in the classification. The remaining three were used for the validation of the selected characteristic. Each one of the six training images was cut out in 100 blocks of 68x68 pixels. For each one of the blocks was obtained the value of the characteristic textural geostatistic (variogram, the madograma, cross variogram and pseudo cross variogram) and the one of Haralick (angular moment, average, variance, entropy, correlation, moment of the product, inverse moment of the difference and correlation measures). Additionally, characteristic geostatísticos and no-geostatísticos they were obtained considering different angles (0, 45, 90 and 135°) of relationship among pixels. Characteristic geostatísticos were, also, obtained for different distances (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) of in pairs among pixels. The characteristic variograma and madograma were calculated to leave of the image green excess and GNDVI. Already the characteristic cross variogram and pseudo cross variogram were calculated with Greenness Method use in the blocks using the combinations of the bands RxG, GxB and IVxG. The characteristic of Haralick were calculated starting from the images of the green excess and GNDVI. The acting of the characteristic, proposed like this, it was evaluated using discriminate analysis. The selected characteristic were those that presented larger value for the index kappa. Additionally, new characteristic were obtained starting from combinations of the selected characteristic. These combinations, also, had appraised acting using the discriminant analysis with the objective of to identify which combination provides better classification. Later, the power of generalization of the selected combination was evaluated using the three images of each species reserved for the validation stage. The conclusions obtained regarding the objectives proposed in this research were a) the image that tended to present the best results of the index kappa was the image excess of green; b) the characteristic obtained starting from the function madograma and the one of Haralick were the ones that supplied the best results; c) the characteristic geostatístico madograma in the 10 distances and angle 0° presented better classification results when used without combination of other characteristic; d)

the characteristic geostatísticos and the one of Haralick, when used separately didn't present such good as combined results; e) the characteristic use that consider the continuity of the pixel values, in the recognition of patterns can be a fundamental tool in the classification process.

1. INTRODUÇÃO

Novos empresários rurais têm buscado estratégias de manejo integrado de plantas daninhas e pragas em sítio específico, objetivando aperfeiçoar a eficiência do uso de insumos aplicando-os a uma taxa variável, ao contrário da agricultura convencional, que os aplica a taxas uniformes. Desta forma recursos tecnológicos inovadores, como sensores com sistemas de posicionamento global têm sido utilizados. A adoção de tais tecnologias possibilita ao produtor ter conhecimento detalhado de sua propriedade, permitindo ao mesmo avaliar a necessidade de manejar de forma diferenciada zonas dentro de sua área.

No manejo convencional as aplicações de insumos e defensivos são baseadas nas necessidades médias da área. Caso exista variabilidade na quantidade necessária destes insumos, essas aplicações são excessivas em determinadas áreas do campo e insuficiente em outras, pois uma área raramente é uniforme. As aplicações excessivas dos insumos podem tornar-se fontes de contaminação das águas subterrâneas e superficiais (ZHANG e CHAISATTAPAGON, 1995).

Com o objetivo de reduzir a quantidade de defensivos aplicados e o impacto sobre o meio ambiente, pesquisadores como Willis e Stoller (1990), Chancellor e Goronea (1993) e Tian et al. (1999) estudaram a variabilidade da distribuição espaço-temporal das plantas daninhas. Tais estudos foram realizados pelo mapeamento de áreas infestadas ao longo dos anos.

O mapeamento de plantas daninhas pode ser realizado usando GPS. Porém, fatores como o tempo e o custo têm limitado o seu uso comercial. Em face destas limitações, sistemas de visão artificial (SVA), que se baseiam em técnicas de sensoriamento remoto têm sido avaliados para o controle de plantas daninhas (ORLANDO, 2003 e SENA JUNIOR, 2005).

As principais funções de um SVA são adquirir imagens e reconhecer os padrões existentes nestas. Para desenvolver estas funções um SVA básico é composto por uma câmera e um computador. Para que um SVA funcione em

tempo real, faz-se necessário que a câmera esteja conectada ao computador. Este deve estar equipado com uma placa de captura de vídeo e implementado com softwares capazes de processar as imagens. As etapas envolvidas neste processamento abrangem o pré-processamento e a segmentação das imagens. Numa etapa seguinte são extraídas características utilizadas para o reconhecimento dos padrões existentes nas imagens.

O uso de sistemas de visão capazes de reconhecer regiões infestadas por planta daninha tende a diminuir o custo com a aplicação de herbicidas. Hummel e Stoller (2002) compararam o custo da aplicação de herbicida no manejo convencional com a aplicação a taxa variada de acordo com a densidade de infestação de plantas daninhas. O custo da aplicação a taxa variada foi, em média, U\$7/área menor do que a aplicação convencional.

Contudo, a maioria dos SVA desenvolvidos até o presente, utiliza descritores texturais que não levam em consideração a autocorrelação existente entre os pixels que compõem uma imagem (descritores de Haralick).

Esta autocorrelação entre pixels pode ser considerada no processo de reconhecimento de plantas daninhas em uma imagem ao utilizar descritores de textura que se baseiam em funções geoestatísticas. Estas funções são utilizadas na análise de autocorrelação espacial e de estruturas da variabilidade espacial (RAMSTEIN e RAFFY, 1989).

A textura de uma imagem está relacionada com a variabilidade espacial dos valores dos pixels. Uma textura mais rugosa indica que pixels vizinhos não possuem valores semelhantes, ao passo que, uma textura mais fina indica que os pixels vizinhos sejam bem semelhantes. Algumas funções geoestatísticas, como por exemplo, a semivariância, tem por objetivo medir esta semelhança entre valores observados em pontos vizinhos e, portanto, poderiam ser utilizadas para descrever a textura de uma imagem.

Alguns trabalhos (CARR, 1996; LACAZE et al., 1994; MIRANDA e FONSECA, 1996) indicaram a existência de potencial no uso de funções geoestatísticas para descrever a textura e assim reconhecer os padrões existentes em uma imagem.

Estes resultados revelam um potencial promissor para o desenvolvimento de um SVA para o reconhecimento de plantas daninhas usando funções geoestatísticas como descritores texturais. Uma imagem adquirida de uma cultura infestada por plantas daninhas irá conter padrões que poderiam ser reconhecidos como cultura ou planta daninha. Estes dois padrões, em princípio, apresentam texturas diferentes em uma imagem. A hipótese deste trabalho, portanto, é que descritores texturais geoestatísticos são tão eficientes em reconhecer padrões quanto descritores não-geoestatísticos.

1.1. Objetivo geral

Avaliar um descritores geoestatísticos para o reconhecimento dos padrões na cultura do milho e ervas daninha.

1.2. Objetivos específicos

- a) Identificar qual imagem, excesso de verde ou o índice de vegetação de verde normalizado, tende a proporcionar melhor classificação;
- b) Comparar a classificação obtida por descritores geoestatísticos, com a obtida ao usar os descritores de Haralick;

2. REVISÃO DE LITERATURA

2.1. Mapeamento das plantas daninhas

Lutman e Perry (1999) dividiram as técnicas de mapeamento das plantas daninhas em dois grandes grupos. O primeiro grupo, composto por técnicas de detecção por amostragem, se baseia na divisão da área em grades (“grids”) e avaliação visual da presença de plantas daninhas. O segundo, composto por técnicas de detecção automatizada, se baseia no uso de unidades móveis de vôo de baixa altitude, satélites e outras formas de sensoriamento.

O mapeamento pelo método de amostragem em grade permite amostragens sistemáticas sobre a comunidade de plantas daninhas, gerando amostras que representem a subárea. Estas informações são convertidas em mapas de infestações de plantas daninhas por krigagem ou por outro método relacionado (Lutman e Perry, 1999). Porém ainda existem limitações sobre sua utilização na prática devido ao tempo necessário para a confecção de mapas de aplicações de herbicidas, principalmente em áreas extensas.

Em virtude da maioria das plantas daninhas infestarem as áreas agrícolas em “reboleiras”, para fins de tratamento, sua variabilidade no campo deve ser controlada com aplicações localizadas de herbicidas.

Segundo Stafford e Miller (1996) a aplicação localizada de herbicidas pode ser realizada de duas maneiras: (i) Baseada em mapas de infestações de plantas daninhas e (ii) Através de um sistema de detecção em tempo real (*Real time*), de modo que a pulverização de herbicidas é realizada logo após a detecção das plantas daninhas.

O segundo tipo de aplicação localizada de herbicida, *real time*, não necessita de mapeamento das plantas daninhas, pois a pulverização é baseada em sensores acoplados ao pulverizador que detectam a planta daninha acionando a pulverização (ANTUNIASSI e GADANHA JUNIOR, 2000).

Steward e Tian (1998) desenvolveram um sistema composto por câmeras de vídeo montadas sobre um trator para identificar, em tempo real, plantas

daninhas na cultura de soja. O algoritmo utilizado no sistema de identificação apresentou resultados com alto índice de acerto e baixo tempo computacional.

A identificação das plantas daninhas é realizada por meio da extração e reconhecimento automático dos padrões contidos em uma cena.

2.2. Reconhecimento de padrões

Atividades involuntárias de reconhecer e classificar objetos no cotidiano são realizadas por seres humanos. O termo objeto ou padrão, diz respeito aquilo que temos interesse em reconhecer. Recebemos as informações, que são captadas pelos nossos sentidos, e muitas vezes reconhecemos sem nenhum esforço aparente. Exemplo disso é o reconhecimento de vozes no telefone, identificação de pessoas que há tempos não vemos e discriminação entre tipos de plantas.

O padrão a ser reconhecido às vezes está associado a alguma classe que tem padrões semelhantes. O reconhecimento do padrão pode, portanto, ter o objetivo de identificar a qual classe o padrão pertence.

Com este objetivo, em meados de 1960, foram elaboradas teorias, e desenvolvidas técnicas para o desenvolvimento de sistemas capazes de realizar reconhecimento de padrões. Dentre os vários domínios de aplicação destes sistemas, Jain et al. (2000) destacam-se:

- Análise de seqüência de DNA;
- Classificação de documentos de internet;
- Inspeção visual para automação industrial;
- Sensoriamento remoto por imagens multiespectrais;
- Reconhecimento de voz.

A técnica utilizada para o reconhecimento ou classificação de um padrão é feita usando uma das duas metodologias apresentadas a seguir (CAMPOS, 2000):

- Classificação supervisionada - O padrão de entrada é identificado como um membro de uma das classes pré-estabelecidas durante a fase de treinamento do classificador;

- Classificação não-supervisionada - O padrão é associado a uma das classes que são definidas durante o processo de classificação.

As etapas de um algoritmo simplificado para o reconhecimento de padrões são apresentadas na Figura 1 (Adaptado de Webb, 2004):

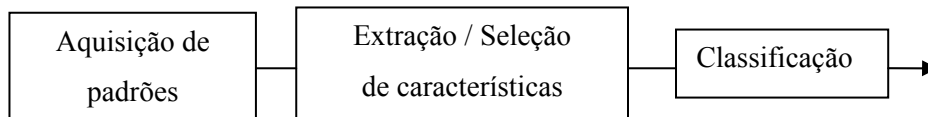


Figura 1 - Etapas de um algoritmo para a classificação de padrões

- Aquisição dos padrões – consiste em adquirir os padrões a serem classificados usando um sensor, exemplo, câmera digital;
- Extração de características – técnicas que extraem características dos padrões a serem classificados. No caso de imagens, estas características são obtidas a partir dos valores dos pixels que compõem uma imagem;
- Seleção de características – analisa o conjunto de características extraídas anteriormente e elimina as irrelevantes.
- Reconhecimento ou Classificação – com base nas características selecionadas, designa o padrão a uma das classes possíveis.

Sendo assim, o desafio em reconhecer padrões consiste na escolha de técnicas para realizar eficientemente as três etapas. Desta maneira, o reconhecimento pode ser definido como a categorização de padrões de entrada, em classes identificáveis, usando características significantes ou atributos extraídos dos padrões de entrada (TOU e GONZALES, 1974).

No caso de reconhecimento de plantas daninhas as pesquisas têm usado características que mensuram a forma e a textura da planta (TANG et al., 1999; FAN et al. 1998). No entanto, a identificação das plantas daninhas por características da forma, principalmente da folha, não tem se mostrado eficiente devido a problemas como sobreposição, orientação e movimentos provocados pelo vento (SENA JUNIOR, 2005). Por esse motivo, apenas características de

textura têm sido utilizadas no desenvolvimento de algoritmos para segmentação das plantas daninhas dos outros objetos na imagem (SENA JUNIOR, 2005).

2.3. Análise textural

A variação dos tons de cinza ou de cor, quando relacionadas a qualidade de uma imagem pode ser definida como textura (RAO, 1990). E, por conseguinte, a análise textural consiste em avaliar a correlação espacial entre estas variações, e podem ser implementadas no domínio espacial ou de frequência, sendo que o uso da segunda opção possui informação de todos os pixels da imagem.

Para Todd e Hans Du Buf, 1993, citado por Orlando (2003), o objetivo da análise textural em imagens é mapear diferenças na estrutura espacial dos níveis de cinza (ou cor), para posteriormente agrupá-los em áreas homogêneas.

Para Park e Chen (2001), a análise textural é robusta para mensurar a intensidade dos níveis de cinza ou da cor, pois se baseia em funções estatísticas que levam em consideração o domínio espacial da imagem.

2.3.1. Tamanho de blocos

Na análise textural para reconhecimento de padrões é comum extrair a informação da textura a partir de blocos recortados da imagem (ORLANDO, 2003). Blocos pequenos tendem a representar a textura da folha. Por outro lado, blocos maiores, dependendo da resolução da câmera, tendem a representar a textura do dossel (ORLANDO, 2003).

O tamanho do bloco utilizado nos diversos trabalhos relacionados ao reconhecimento de plantas daninhas apresenta grande variação. Shearer e Holmes (1990), Burks et al. (2000) e Park e Chen (2001) utilizaram blocos de 64 x 64 pixels. Por outro lado, Yang et al. (1998) utilizaram blocos de 100 x 100 pixels, enquanto que Yang et al. (1999), utilizou o tamanho de 80 x 80 pixels. Esta

grande variação é um indicativo de que a escolha do tamanho do bloco depende do fenômeno sendo estudado.

O tamanho do bloco também pode influenciar a eficiência do sistema de reconhecimento dos padrões. Tang et al. (1999) verificaram que o tempo de processamento e a classificação foram influenciados pelo tamanho dos blocos. Dentre os tamanhos de blocos 9 x 9, 13 x 13 e 17 x 17 pixels avaliados, o de maior dimensão apresentou os melhores resultados. Resultado semelhante foi obtido por Orlando (2003) que avaliou os tamanhos de blocos 17x17, 34x34 e 68x68 pixels.

2.3.2. Descritores texturais

A razão entre a soma do produto dos pixels em uma imagem pela probabilidade marginal da matriz de co-ocorrência é uma forma de medir a textura. Esta maneira utilizada é conhecida como média. Logo a média fornece a informação da textura, sendo denominada como descritor textural.

Segundo Haralick (1973), da matriz de co-ocorrência podem ser gerados 16 descritores texturais: segundo momento angular, contraste, correlação, soma de quadrados, momento do produto, momento inverso da diferença, momento da diferença, soma das médias, soma das variâncias, soma das entropias, entropia, diferença da variância, diferença da entropia, medidas de correlação e máximo coeficiente de correlação. Diversos autores como Alves et al. (2002) e Tang et al. (1999) utilizaram os descritores de Haralick.

Orlando (2003) utilizou os descritores, momento angular, média, variância, entropia, correlação, momento do produto e momento inverso da diferença em um classificador neuronal. Os melhores resultados de exatidão global da classificação usando estes descritores variaram entre 81 e 86,5% para imagens monocromáticas, e, entre 78,8 e 86,2% para imagens excesso de verde dependendo da condição de iluminação.

De acordo com Matsuyama (1980) e Sun e Quin (1993), citado por Olmo e Hernández (2000), os descritores para classificação com base na textura da

imagem são divididos em três métodos: estrutural, espectral e estatístico. Os métodos estruturais consideram a textura como uma repetição de padrões primitivos básicos, os métodos espectrais são baseados na Transformada de Fourier, e os métodos estatísticos se baseiam em parâmetros estatísticos locais.

Uma alternativa aos descritores acima mencionados são os descritores geoestatísticos.

Estes descritores têm como fundamento a teoria de variáveis regionalizadas desenvolvida por Matheron (1971), e vêm sendo usados em diferentes áreas do conhecimento, quando as variáveis estudadas apresentam dependência espacial. A teoria de variáveis regionalizadas é também conhecida com o nome de geoestatística, devido sua aplicação ter sido direcionada, inicialmente, a problemas de geologia e mineração.

2.4. Teoria das variáveis regionalizadas

2.4.1. Aspectos gerais

Diz-se que toda variável distribuída no espaço e que apresenta uma estrutura espacial correlacionada, é regionalizada. Assim, se $z(x)$ é o valor de uma variável aleatória Z em um ponto x , $Z(x)$ é uma variável regionalizada (VR) (OLIVEIRA, 2003). No estudo da dependência espacial, alguns aspectos importantes das variáveis regionalizadas devem ser considerados, como a localização e continuidade.

2.4.2. Localização e continuidade

Uma das características qualitativas da VR, diretamente ligada à estrutura do fenômeno espacial conforme Guerra (1988), é a localização.

Outra característica importante é a continuidade espacial, pois quanto menor a distância entre dois pontos, maior é a tendência destes pontos

apresentarem valores semelhantes. A diferença entre estes pontos reflete o grau de dependência entre um valor e outro (GUERRA, 1988).

A continuidade espacial foi expressa por Valente (1989) como:

$$\lim_{x \rightarrow x_0} E[z(x) - z(x_0)] = 0 \quad (1)$$

Em muitas situações esta continuidade não se verifica, surgindo assim o efeito pepita (τ^2), que é definido como:

$$\lim_{x \rightarrow x_0} E[z(x) - z(x_0)]^2 = \tau^2 \neq 0 \quad (2)$$

2.4.3. Função aleatória

Suponha que em uma região D são selecionados alguns pontos de amostragem. Destes pontos são obtidas as suas coordenadas e as mensurações de atributos de interesse.

Desta forma, o processo estocástico envolvido neste processo de amostragem pode ser expresso por:

$$\{Z(x_i) : x_i \in D \subset R^d\} \quad (3)$$

Em que, Z é a variável aleatória que varia continuamente em D , x é a localização da variável, considerada fixa, D é a região em estudo e R^d é o espaço d -dimensional ($d= 1, 2, 3$ ou 4). Para os casos em que $d=1$ os dados estão na transeção da malha (distribuídos em uma linha transversal da malha), para $d=2$, em um plano, para $d=3$, em um volume e para $d=4$ inclui-se o tempo (OLIVEIRA, 2003).

Assim, o conjunto de valores $z(x_i)$ em todos os pontos x_i na região do domínio D , é considerada uma particular realização do conjunto de variáveis aleatórias $\{Z(x_i) : x_i \in D \subset R^d\}$ (SOARES, 2000).

Segundo Oliveira (2003), a função aleatória expressa aspectos aleatórios e estruturados de uma variável regionalizada de tal forma que: a) localmente, para um ponto x_i , $z(x_i)$ é uma variável aleatória e, b) $Z(x_i)$ é uma função aleatória no sentido de que cada par de pontos (x_i) e $(x_i + h)$ é, em geral, correlacionado, expressando a estrutura espacial da variável regionalizada $Z(x_i)$.

2.4.4. Características das variáveis regionalizadas

Oliveira (2003), a geoestatística modela os valores de um atributo, dentro de uma região, como uma função aleatória apresentada pela Equação 3. Para cada posição, $x \in D$ o valor de um atributo é modelado como uma variável aleatória $Z(x)$. Isto indica que em cada posição, a variável aleatória pode assumir distintos valores para o atributo, e cada uma destas variáveis está associada a uma distribuição (SAMPER e RAMIREZ, 1990; FELGUEIRAS, 1999).

Sua função de distribuição acumulada multivariada, composta por variáveis aleatórias, é definida pela Equação 4:

$$F_{Z(x_1), Z(x_2), \dots, Z(x_k)}(Z_1, Z_2, \dots, Z_k) = \text{Prob}[Z(x_1) \leq Z_1, Z(x_2) \leq Z_2, \dots, Z(x_k) \leq Z_k] \quad (4)$$

A interpretação probabilística de uma variável regionalizada como realização de uma função aleatória $Z(x)$ tem sentido somente se for possível inferir, pelo menos em parte, qual é a função de distribuição ou lei de probabilidade de $Z(x)$. Em geral, uma das limitações dos dados com dependência espacial está relacionada com a impossibilidade de repetir indefinidamente um experimento e realizar inferência a partir de uma única realização. Sendo assim, para que se possa fazer inferências somente com uma realização, são necessárias algumas hipóteses restritivas que permitam conhecer um fenômeno da natureza que tem continuidade espacial.

2.4.5. Hipóteses restritivas

a) Estacionaridade

Diz-se que a função é estritamente estacionária, se sua função de distribuição de probabilidades apresentada na Equação 2 é invariante em relação a qualquer translação do vetor h , ou então, se a função de distribuição do vetor aleatório $[Z(x_1), Z(x_2), \dots, Z(x_k)]$ é idêntica à do vetor $[Z(x_1 + h), Z(x_2 + h), \dots, Z(x_k + h)]$ para qualquer h .

b) Estacionaridade de primeira ordem

O momento de primeira ordem da função aleatória $Z(x)$ é a esperança matemática, definida como $E[Z(x)] = m(x)$ (OLIVEIRA, 2003). Dessa forma, uma função aleatória é estacionária de primeira ordem se admitimos que todas as variáveis aleatórias $Z(x_i)$, $i = 1, \dots, N$, têm a mesma média em todos os locais da região em estudo. Esta hipótese, definida por Trangmar et al. (1985), é expressa por:

$$E[Z(x_1)] = E[Z(x_2)] = \dots = E[Z(x_i)] = E[Z(x)] = m \quad (5)$$

Quando existe a variância de $Z(x)$, ela pode ser definida, como o momento de segunda ordem em relação à média.

c) Estacionaridade de segunda ordem

Para definir a estacionaridade de segunda ordem é necessário, antes, definir a covariância, a qual é o momento de segunda ordem da função aleatória $Z(x)$.

Seja duas variáveis aleatórias $Z(x_i)$ e $Z(x_j)$, nos pontos x_i e x_j , respectivamente. A covariância é definida por (OLIVEIRA, 2003):

$$C(Z(x_i), Z(x_j)) = E[[Z(x_i) - m(x_i)][Z(x_j) - m(x_j)]] \quad (6)$$

ou ainda, de forma equivalente,

$$C(Z(x_i), Z(x_j)) = E \{ [Z(x_i) - E(Z(x_i))][Z(x_j) - E(Z(x_j))] \} \quad (7)$$

A covariância, definida na Equação 6, deve ser positiva definida para assegurar a existência de todas as distribuições acumuladas.

Sendo assim, uma função aleatória $Z(x)$ é estacionária de segunda ordem se, além de atender a hipótese de estacionaridade de 1ª ordem, apresentar a característica de que, para cada par de valores $Z(x)$ e $Z(x+h)$, a covariância (Equação 6) existe e depende apenas da distância h (JOURNEL e HUIJBREGTS, 1978; VIEIRA et al., 1983; TRANGMAR et al., 1985; GONÇALVES, 1997). As equações relacionadas a esta hipótese como a esperança matemática, covariância e variância são:

$$E[Z(x)] = m, \quad \text{para todo } x \quad (8)$$

$$C(Z(x+h), Z(x)) = C(h) = E[Z(x+h)Z(x)] - m^2 \quad (9)$$

Ou

$$C(h) = E[Z(x) - m][Z(x+h) - m] \quad (10)$$

A hipótese de estacionaridade de segunda ordem pressupõe a existência de covariância e em conseqüência, uma variância *a priori* finita. A expressão para variância, neste caso, será:

$$\text{Var}[z(x)] = E[z(x) - m]^2 = C(0) \quad (11)$$

Na natureza existem numerosos fenômenos que apresentam dispersão ilimitada. Neste caso, tais fenômenos não podem ser descritos de forma correta se lhes forem atribuída uma variância finita (SAMPER e RAMIREZ, 1990).

Esta hipótese de variância finita é muito forte e raramente assumida. Esta hipótese usualmente é substituída por uma hipótese mais fraca chamada de hipótese intrínseca.

d) Hipótese intrínseca

Esta, não é definida através das distribuições das variáveis aleatórias pontuais e sim das diferenças entre elas, também chamada de estacionaridade dos incrementos ou estacionaridade fraca (RIBEIRO JUNIOR, 1995).

A hipótese intrínseca é definida pelas seguintes condições:

- O incremento entre dois pontos tem esperança nula

$$E[Z(x+h) - Z(x)] = 0 \quad (12)$$

- Para qualquer vetor h o incremento tem variância finita que independe da posição x .

$$Var[Z(x+h) - Z(x)] = E[Z(x+h) - Z(x)]^2 = 2\gamma(h) \quad (13)$$

sendo $2\gamma(h)$ a definição de variograma.

A hipótese intrínseca apresentada nas Equações 13 e 14 requer que para todos os vetores de h , a variância do acréscimo $Z(x+h) - Z(x)$, dada pela Equação 14, seja finita e independente da localização dentro da região.

Além de possuir as hipóteses referidas, é possível obter várias realizações da variável por meio da diferença entre os valores em pontos separados por uma distância h , ou seja, $[Z(x_i + h) - Z(x)]$ (DAVID, 1977; CRESSIE, 1993; RIBEIRO JUNIOR, 1995; GONÇALVES, 1997; KITANIDIS, 1997).

Dividindo a Equação 13 por dois, obtém-se a semivariância $\gamma(h)$, a qual depende do vetor de distâncias h . Espera-se que o valor de $\gamma(h)$ aumente à

medida que a distância h aumente, até um valor máximo no qual a semivariância tende a estabilizar o seu valor.

A função de semivariância é a metade da esperança do quadrado da variável aleatória $[Z(x) - Z(x+h)]$ (JOURNEL e HUIJBREGTS, 1978), ou seja:

$$\gamma(x, h) = \frac{1}{2} E[Z(x+h) - Z(x)]^2 \quad (14)$$

O variograma é, portanto, apresentado por $2\gamma(x_i, x_j)$, mas existem autores que usam indistintamente ambas as formas para referir-se à função $\gamma(x_i, x_j)$ (OLIVEIRA, 2003).

2.5. Técnicas Geoestatísticas para classificação de imagens

Segundo Atkinson e Lewis (1999), todas as técnicas geoestatísticas são caracterizadas por um modelo de função de covariância espacial ou, mais freqüentemente, o variograma.

Para variáveis contínuas, tal como a refletância de um dado comprimento de onda, de uma banda espectral, tem-se que, o variograma experimental é definido como a metade do valor médio do quadrado da diferença entre os valores de pixels separados por vetor h que é específico para uma direção e distância (ATKINSON e LEWIS, 1999).

O variograma experimental $\gamma(h)$ pode ser obtido por $\{x, x+h\}$ pares de observações $\{Z(x), Z(x+h)\}$ definidos por uma localização $x = 1, 2, \dots, P(h)$ separadas a uma distância fixada h .

Rerzfeld e Chris (1996) utilizaram o variograma direcional, como função de estrutura espacial na geoestatística para distinguir rugosidade geológica e morfológica, no fundo do oceano atlântico. A aplicação deste método conseguiu classificar, consistentemente, as propriedades morfológicas e geológicas, no fundo do oceano.

Berberoglu et al. (2000) classificaram imagens do Landsat, da região litoral do sudeste da Turquia. Para a classificação espectral e textural de diferentes coberturas vegetais, os autores utilizaram o variograma com diferentes distâncias de separação. Os resultados obtidos mostraram que a melhor acurácia foi obtida com 1 unidade de distância de separação entre os pixels, obtendo um total de 89% de exatidão na classificação dos diferentes tipos de cobertura da superfície em análise.

Outros exemplos de incorporação do variograma em classificação de imagens podem ser observados na literatura. Miranda et al. (1996) utilizaram um classificador de semivariograma textural, para discriminar vegetação a partir de imagens. Ramstein e Raffy (1989) utilizaram modelos de variograma. Carr (1996), Lark (1996), Miranda et al. (1998), Miranda et al., (1994) e Miranda et al. (1995) usaram parâmetros geoestatísticos deduzidos pela função de variograma, dentre outros.

O que também tem sido notado é um grande aumento no uso de funções derivadas do variograma (MIRANDA et al. 1992, 1996; RAMSTEIN e RAFFY, 1989).

Carr (1996) usou a semivariância e a informação espectral separadamente e em combinação, com classificação supervisionada. Este autor relata que a acurácia da classificação foi maior quando usou a semivariância em combinação com a informação espectral, do que quando usada sozinha.

2.6. Classificadores Estatísticos

Num algoritmo de classificação (Figura 1), após selecionar as características, ou seja, os descritores, é necessário introduzi-las em um classificador. Segundo Jain et al. (2000) as quatro abordagens mais conhecidas de reconhecimento de padrões são: casamento (*template matching*), estatística, sintática e redes neurais.

Nesta seção é apresentada, em detalhes, a abordagem estatística. Não serão descritas as outras abordagens, visto que não pertencem ao escopo deste

trabalho. Maiores detalhes podem ser vistos em Jain et al. (2000) e Theodoridis (1999).

A utilização estatística para reconhecimento de padrões tem sido amplamente discutida no meio científico. Jain et al. (2000), realizaram um rápido levantamento dentre os artigos publicados na revista *IEEE Transactions on Pattern Analysis and Machine Intelligence* e descobriram que, no período de 1979 a 2000, 350 artigos se enquadravam na área de reconhecimento de padrões, dos quais 300 relacionavam-se a abordagem estatística.

Um classificador estatístico assume que, cada padrão é uma observação obtida aleatoriamente e pode ser representado por um vetor x de características.

De acordo com Duda e Hart (1973) citado por Jayas et al. (2000), os classificadores estatísticos baseiam-se na regra do erro mínimo de Bayes:

$$x \in \Omega_k \quad \text{se} \quad P(w_k | x) > P(w_j | x) \quad \forall j \neq k \quad (15)$$

em que,

x = vetor de características de um objeto desconhecido; e

$P(w_j/x)$ = probabilidade a posteriori de que um objeto com um vetor de características x pertença à classe w_j .

A regra afirma que para minimizar a probabilidade média de erro, um objeto deveria ser classificado como pertencente à classe w_i que maximiza a probabilidade a posteriori $P(w_i/x)$. Uma formulação mais prática pode ser obtida com a aplicação do teorema de Bayes (JAYAS et al., 2000):

$$P(w_i | x) = \frac{P(w_i)p(x | w_i)}{p(x)} \quad (16)$$

obtendo-se:

$$x \in \Omega_k \quad \text{se} \quad P(w_k)p(x | w_k) > P(w_j)p(x | w_j) \quad \forall j \neq k \quad (17)$$

em que,

$P(w_i)$ = probabilidade a priori de que um objeto venha da classe w_i ;

$p(x)$ = função densidade probabilidade de x ; e

$p(x/w_i)$ = função densidade probabilidade de x dado a classe i .

Na maioria das aplicações a probabilidade a posteriori, ou seja, a função densidade probabilidade (fdp) condicional da classe é desconhecida, necessitando, portanto, ser estimada. A estimativa pode ser obtida usando métodos paramétricos ou não paramétricos (JAYAS et al., 2000). O método paramétrico assume que, a fdp condicional da classe, tem a forma de uma distribuição normal multivariada. Portanto, para estimar a fdp, é necessário estimar os parâmetros média (μ_i) e variância (Σ_i). Por outro lado, o método não-paramétrico calcula a probabilidade *a posteriori* $P(w_i/x)$ diretamente dos dados de treinamento, sem assumir nenhum tipo de distribuição para a fdp (JAYAS et al., 2000).

Em um processo de classificação usam-se exemplos para treinamento do classificador. Desta forma o desempenho do classificador depende tanto do número de amostras do treinamento como também dos padrões contidos nestas amostras. Ao mesmo tempo, o objetivo é classificar futuros padrões que não sejam os mesmos usados no treinamento. Por isto nem sempre o processo de classificação produzirá resultados satisfatórios. Campos (2000) cita alguns fatores que podem contribuir para não produzir tais resultados.

O sobre-ajuste (*overfitting*) acontece quando o classificador é mais complexo que o necessário fazendo com que o poder de generalização diminua. Isso ocorre devido ao fato de que o classificador se adapta a detalhes específicos da base de treinamento (THEODORIDIS, 1999).

Por outro lado, o sobre-treinamento (*overtraining*) ocorre quando o classificador é treinado usando um conjunto muito grande de exemplos de padrões (THEODORIDIS, 1999). Uma solução é utilizar, para treinamento, padrões que realmente representem todas as possíveis variações entre padrões de mesma classe.

Outro fator que pode prejudicar a classificação é a redução de dimensionalidade. O termo dimensionalidade refere-se à dimensão do espaço de

características. Duas principais razões para que a dimensão seja a menor possível são o custo de medição e a precisão do classificador. Há um fator analisado pelo teorema do “patinho feio” de Watanabe (1985), onde é possível fazer dois padrões arbitrários ficarem similares se esses forem codificados com um número suficientemente grande de características similares. Isso enfatiza a necessidade de uma escolha cuidadosa de características.

Entretanto no final do treinamento para que o classificador possa ser utilizado e aceito, é necessário avaliá-lo.

2.6.1. Desempenho e avaliação de classificadores

Devido a amostra ter, em geral, tamanho reduzido, cada conjunto de características é utilizado para treinar e também avaliar um classificador (CAMPOS et al., 2000). Um critério de avaliação do classificador se baseia na taxa de acerto no reconhecimento. Quanto maior a taxa de acerto, melhor é o conjunto de características. Segundo Kohn (1998), não se deve empregar o conjunto de treinamento (utilizado para projetar o classificador) para estimar a probabilidade de erro, pois o classificador estará ajustado especificamente para o conjunto de treinamento e a estimativa da probabilidade de erro será tendenciosa. Dois problemas ocorrem nesta abordagem. O primeiro é que quando a razão entre o tamanho do conjunto de amostras e o do conjunto de características é pequena, o erro de classificação por si só não pode ser confiavelmente estimado. Além disso, a escolha de um classificador é um problema por si só, e o subconjunto selecionado ao final claramente depende do classificador (JAIN et al. 2000).

2.6.2. Uso do método da validação cruzada

A validação cruzada é um dos métodos que podem ser utilizados para avaliar o desempenho de um classificador. De acordo com Figueira et al. (2006), em cada iteração, esta técnica é subdividida em dois conjuntos de amostras: o de

validação e o de treinamento. O subconjunto de validação é apenas um elemento. O subconjunto de treinamento é composto então por todas as amostras, com exceção da amostra usada no primeiro subconjunto. Desta maneira, o desempenho da classificação de cada elemento é estabelecido individualmente e, ao final, obtêm a performance global do classificador, com relação a todos os elementos do conjunto. Este método de validação cruzada é também conhecido como método da validação cruzada deixando um de fora, pois conforme apresentado anteriormente, a cada iteração uma amostra é deixada de fora para avaliar o desempenho do classificador com relação à mesma.

3. MATERIAL E MÉTODOS

3.1. Aquisição de imagens digitais

As imagens utilizadas nesta tese são parte das imagens utilizadas na elaboração da tese de Roberto Carlos Orlando apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Agrícola, para obtenção do título de “*Doctor Scientiae*”.

Para a execução do projeto de pesquisa da tese de Orlando (2003) foram adquiridas imagens em um experimento instalado segundo o esquema fatorial, durante o ano de 2001 nas dependências da Embrapa Milho e Sorgo, localizada em Sete Lagoas, Minas Gerais. A coleta das imagens foi realizada no período de oito dias, por duas câmeras diferentes resultando em três tipos de imagens, em quatro níveis de iluminação e três alturas das câmeras para quatro espécies de plantas daninhas e a planta de milho, com três repetições, totalizando 1620 imagens. O objetivo da tese de Orlando (2003) foi avaliar o efeito dos fatores tipo de câmera, tipos de imagem, nível de iluminação, altura da câmera, espécie de planta daninha na acurácia da classificação de plantas daninhas e milho usando classificadores do tipo neural e estatístico.

As espécies de plantas daninhas avaliadas no experimento foram: o picão preto (*Bidens pilosa* L.), leiteira (*Euphorbia heterophylla* L.), capim-milhã (*Digitaria horizontalis* Willd) e timbête (*Cenchrus echinatus* L.). A escolha das espécies de plantas daninhas foi feita com base no critério de maior incidência para as condições de manejo, clima e solo, presentes na Embrapa Milho e Sorgo.

No controle de iluminação artificial foram utilizadas lâmpadas alógenas com 600, 700, 800 e 900 watts de potência. As imagens adquiridas foram do tipo: colorida (bandas vermelha, verde e azul), falsa-cor (bandas infravermelho próximo, vermelho e verde) e monocromática. As alturas das câmeras foram: 2,04; 1,58 e 1,32m para imagens coloridas e falsa cor, e 2,04; 1,61 e 1,35m para imagens monocromáticas.

O crescimento da planta pode alterar algumas de suas características, como por exemplo, a cor e a textura. Sendo assim utilizou as imagens tomadas nos dias 10, 13 e 17 de julho, correspondentes a 22, 25 e 29 dias após a emergência, período em que normalmente é feita a aplicação de herbicida.

A obtenção das características texturais foi feita em blocos da imagem com dimensões de 17 x 17, 34 x 34 e 68 x 68 pixels, recortados em regiões de interesse das imagens coletadas.

Tendo como base as conclusões obtidas por Orlando (2003), que indicaram os níveis daqueles fatores que proporcionaram maior acurácia, foram utilizadas neste trabalho apenas as imagens coloridas (RGB) e falsa cor (bandas infravermelho próximo, vermelho e verde) adquiridas a uma altura de câmera correspondente a 2,04m e iluminação com 900 watts de potência das espécies leiteira (*Euphorbia heterophylla* L.), capim-milhã (*Digitaria horizontalis* Willd), timbête (*Cenchrus echinatus* L.) e milho (*Zea Mays*). No total, portanto, foram utilizadas 36 imagens RGB e 36 falsa cor, sendo nove imagens para cada espécie. Das nove imagens de cada espécie, seis foram usadas para a seleção dos descritores que apresentaram melhor desempenho no reconhecimento dos padrões de plantas. Esta etapa de seleção dos descritores, daqui por diante será referida como treinamento. As três imagens restantes de cada espécie foram destinadas para a validação do desempenho dos descritores selecionados.

Computadores do Laboratório de Projeto de Máquinas e Visão Artificial (PROVISAGRO) do Departamento de Engenharia Agrícola da Universidade Federal de Viçosa e computadores do Laboratório de Análises e Pesquisa em Estatística (LAPe) foram utilizados para desenvolver algoritmos (Apêndices A a L), implementados no software MATLAB® para realizar as etapas de pré-processamento, processamento e classificação dos padrões presentes em cada uma das imagens.

3.2. Pré-Processamento das imagens

Esta etapa teve por finalidade diminuir ou anular o efeito de fatores que pudessem dificultar o reconhecimento das plantas daninhas e de milho. Inicialmente foi necessário definir qual o tamanho do bloco em torno do pixel que se deseja reconhecer o padrão ao qual pertence.

3.2.1. Recorte dos blocos das imagens

Orlando (2003) avaliou a acurácia de classificação obtida a partir de blocos de dimensões 17x17, 34x34 e 68x68 pixels. O tamanho de 68x68 pixels foi o que apresentou melhor acurácia. Portanto, para este trabalho as imagens foram recortadas em blocos de 68 x 68 pixels. Os recortes dos blocos foram realizados por um algoritmo implementado no MATLAB® (Apêndice A). Para a execução deste algoritmo foi necessário, inicialmente, demarcar a região de interesse da imagem com três pontos. Estes três pontos correspondem aos cantos superiores direito e esquerdo, e ao canto inferior direito desta região. Desta maneira, o algoritmo cortou blocos de 68x68 pixels limitados a região escolhida na imagem, como mostra a Figura 2. Para a combinação dos fatores definidos (imagem RGB, altura da câmera de 2,04m e 900watts de potência), foram recortados 100 blocos de cada uma das 6 imagens destinadas para treinamento. Portanto no total, para cada espécie foram utilizadas 600 amostras para treinamento.

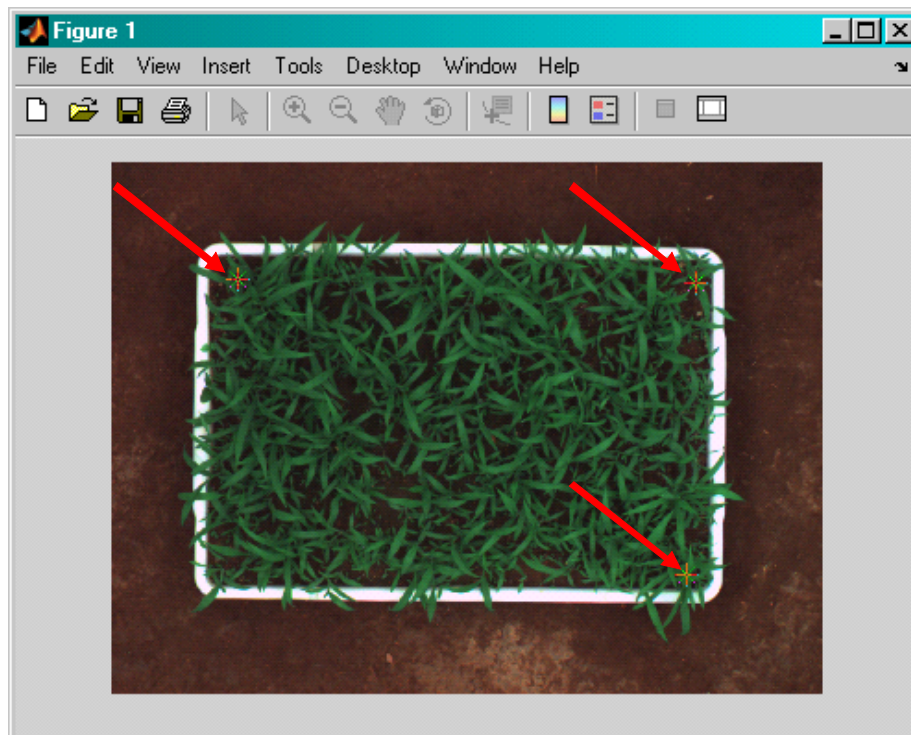


Figura 2 - Demarcação dos três pontos na imagem para recorte dos blocos.

3.2.2. Greenness Method

Este método pressupõe que os padrões verdes existentes em uma imagem são plantas (YANG et al., 2002). Esta pressuposição é válida em experimentos controlados, tal como o realizado neste trabalho em que os padrões existentes foram apenas planta, solo e bandeja. Espera-se, portanto, que em uma imagem RGB, os pixels para os padrões planta apresentem maior valor para a banda verde (G) do que o solo ou o vaso.

Nos blocos recortados das imagens de treinamento, cada pixel foi classificado quanto a ser de uma planta ou não. Para isto, nas imagens RGB os valores das bandas vermelha (R), verde (G) e azul (B) foram comparados. Aqueles pixels que apresentaram o valor da banda G maior do que os da banda R e B foram utilizados na análise. Para os demais pixels atribui-se o valor zero nas bandas utilizadas. Para as imagens falsa cor infravermelha (CIR), utilizou-se apenas as bandas infravermelho próximo (IV) e verde. Os valores dos pixels na

banda IV foram comparados. Aqueles que apresentaram valores maiores do que os da banda G foram utilizados na análise. Para os demais pixels atribui-se o valor zero nas bandas utilizadas.

Portanto, este método desconsidera os pixels de padrões indesejáveis da imagem, pois a imagem resultante preserva apenas os pixels dos padrões que se deseja reconhecer (neste caso as plantas daninhas e milho) e elimina as outras informações.

O algoritmo elaborado para a implementação deste método é apresentado nos Apêndices C, E, G e I.

3.2.3. Realce e Normalização

A imagem excesso de verde foi obtida pela Equação 18 a seguir, com a finalidade de realçar a informação verde. Já a normalização, divisão pela soma das 3 bandas, tem como objetivo corrigir esta variação na iluminação, os valores de pixels foram normalizados usando o índice de excesso de verde nas imagens RGB. Este índice foi obtido por

$$ExV = \frac{2 * G - R - B}{G + R + B} \quad (18)$$

em que

ExV = valor do pixel na imagem de excesso de verde normalizado;

G = valor do pixel na banda verde;

R = valor do pixel na banda vermelha;

B = valor do pixel na banda azul.

Para as imagens falsa cor, utilizou-se o índice de vegetação da diferença de verde normalizado.

$$GNDVI = \frac{IV - G}{IV + G} \quad (19)$$

onde:

GNDVI = índice de vegetação da diferença de verde normalizado e representa o valor do pixel na imagem falsa cor;

IV = valor do pixel na banda do infravermelho próximo.

O algoritmo para a obtenção destes índices na linguagem do MATLAB é apresentado nos Apêndices C e E.

3.3. Processamento

Esta etapa de processamento consistiu em obter descritores texturais para cada um dos 600 blocos recortados das imagens originais. O algoritmo para a obtenção destes descritores na linguagem do MATLAB é apresentado nos Apêndices B, D, F e H (descritores geoestatísticos) e J (descritores de Haralick). Os descritores obtidos foram os de Haralick e os geoestatísticos apresentados a seguir.

3.3.1. Descritores de Haralick

Para cada uma dos 600 blocos de imagem de excesso de verde e GNDVI foi gerada uma matriz de co-ocorrência normalizada na direção de um ângulo, digamos θ , do pixel seguinte, ou seja, supondo uma distância entre pixels de uma unidade ($d = 1$).

Shearer (1987), citado por Orlando (2003), expressa a criação de uma matriz de co-ocorrência para uma imagem $Im(x,y)$ com quatro níveis de valores numéricos do pixel, considerando o vizinho seguinte ($d = 1$), na direção horizontal ($\theta = 0$).

Nesse modelo, a imagem

$$\text{Im}(x, y) = \begin{bmatrix} 0 & 0 & 3 & 1 \\ 2 & \mathbf{1} & 0 & 2 \\ 3 & 2 & 0 & 3 \\ \mathbf{1} & \mathbf{2} & \mathbf{1} & 3 \end{bmatrix}$$

resultou a matriz de co-ocorrência

$$P_0^1(i, j) = \begin{bmatrix} 2 & 1 & 2 & 2 \\ 1 & 0 & 3 & 2 \\ 2 & 3 & 0 & 1 \\ 2 & 2 & 1 & 0 \end{bmatrix}.$$

Pode-se mencionar como exemplo, o valor $P_0^1(2, 3) = 3$ está representando os três pixels da imagem $\text{Im}(x, y)$ que tem o nível dois (valor numérico 1) seguido, na mesma linha ($\theta = 0$), do nível três (valor numérico 2) ou nível três seguido de dois. Estas ocorrências estão destacadas na $\text{Im}(x, y)$. Em outras palavras, quantas vezes o valor um ocorre ao lado do valor dois.

A matriz de co-ocorrência normalizada na direção horizontal $\theta = 0$ e distância do pixel, $d = 1$, é dada a seguir:

$$P(i, j) = \frac{p(i, j)}{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i, j)} \quad (20)$$

em que:

$P(i, j)$ = matriz de co-ocorrência normalizada;

$p(i, j)$ = elementos da matriz de co-ocorrência;

N_g = número total de níveis do atributo.

Para cada uma das 600 matrizes de co-ocorrência geradas, foram obtidos os descritores de Haralick: momento angular, média, variância, entropia, correlação, momento do produto, momento inverso da diferença e medidas de correlação. Estes descritores são apresentados a seguir.

Momento angular: Essa característica é uma medida da homogeneidade da imagem.

$$C_1 = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} [p(i,j)]^2, \quad (21)$$

em que:

C_1 = momento angular;

$p(i, j)$ = elementos da matriz de co-ocorrência;

N_g = número total de níveis do atributo.

Média: é uma medida da intensidade da imagem representada em uma matriz de co-ocorrência.

$$C_2 = \sum_{i=0}^{N_g-1} i \times p_x(i), \quad (22)$$

em que:

C_2 = média;

i = valor observado;

$p_x(i)$ = probabilidade marginal da matriz de co-ocorrência (soma da linha).

Variância: contraste da imagem (momento de ordem 2).

$$C_3 = \sum_{i=0}^{N_g-1} (i - C_2)^2 p_x(i), \quad (23)$$

em que:

C_3 = variância;

Entropia: é uma medida da aleatoriedade, mede a informação contida em p.

$$C_4 = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i,j) \ln [p(i,j)], \quad (24)$$

em que:

C_4 = entropia;

\ln = logaritmo neperiano.

Correlação: é uma medida de quão correlacionado está um pixel com seu vizinho na imagem, variando de -1 a 1.

$$C_5 = \frac{\left[\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} ij p(i, j) \right] - C_2^2}{C_3}, \quad (25)$$

em que:

C_5 = correlação;

i e j = coordenadas do pixel observado;

Momento do produto: quando os valores altos da matriz de co-ocorrência estiverem próximos da diagonal, têm-se um valor baixo para esse descritor.

$$C_6 = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i - C_2)(j - C_2) p(i, j) \quad (26)$$

Momento inverso da diferença: quando os valores altos da matriz de co-ocorrência estiverem próximos da diagonal, este descritor possuirá também valor alto.

$$C_7 = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{p(i, j)}{1 + (i - j)^2} \quad (27)$$

Medidas de correlação: Segundo Shearer (1987), é uma representação da taxa de entropia da matriz de co-ocorrência e foi determinada pela expressão:

$$C_8 = \frac{C_2 - HXY1}{HX} \quad (28)$$

em que:

$$HX = - \sum_{i=0}^{N_g-1} p_x(i) \ln[p_x(i)] \quad (29)$$

$$HXY1 = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i, j) \ln[p_x(i)p_x(j)] \quad (30)$$

em que:

C_8 = medidas de correlação;

$p_x(j)$ = probabilidade marginal da matriz de co-ocorrência (soma da coluna).

Os descritores de Haralick foram gerados para os ângulos θ de 0, 45, 90 e 135° entre pixels. Como existem 8 funções e 4 ângulos, foram avaliados no total, 32 descritores de Haralick. Os algoritmos implementados no MATLAB para a obtenção destes descritores em cada um dos blocos de imagens ExV e GNDVI, são apresentados no Apêndice J.

3.3.2. Descritores geoestatísticos

Descritores geoestatísticos foram avaliados com relação à capacidade de quantificar a continuidade dos valores de pixels que existe em cada um dos padrões a serem classificados. Pressupõe-se que os padrões a serem classificados apresentem diferentes graus de continuidade nos valores do pixel.

3.3.2.1. Variograma

Para obter o variograma foi necessário, para cada intervalo de separação, calcular o valor médio da semivariância entre pares de pontos que satisfizessem ao intervalo de separação. O gráfico das semivariâncias em função da distância de separação entre pares de pixels é referido como variograma. Segundo Silva et al. (2004) os parâmetros que definem um variograma são o patamar, o alcance e o efeito pepita e podem ser observados na Figura 3. O patamar é o valor de estabilização do valor da semivariância, o alcance é a distância correspondente ao patamar. O efeito pepita é o valor da semivariância considerando distância de separação nula.

Supondo que o valor de um pixel, digamos i , de uma imagem seja identificado por um número digital (dn), a função do semivariograma é definida como (OLMO e HERNÁNDEZ, 2000):

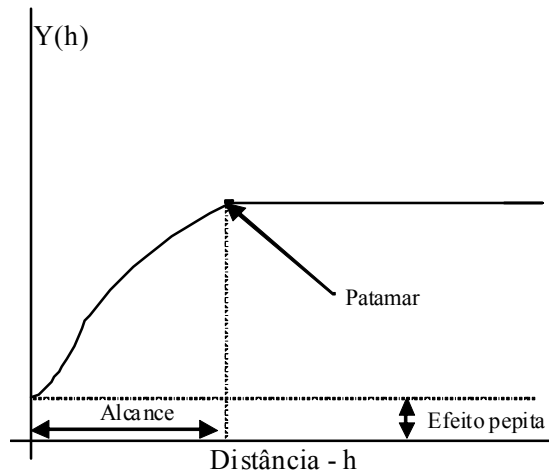


Figura 3 - Variograma

$$Y(h_\theta) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \{dn(x_i) - dn(x_i + h)\}^2 \quad (31)$$

Onde:

- $Y(h_\theta)$ - semivariância estimada para pares de pontos separados por h na direção θ ;
- $n(h)$ - é o número pares de pontos que satisfazem a distância;
- dn - são os valores dos pixels nas posições x_i e $x_i + h$;

3.3.2.2. Madograma

Semelhante ao variograma, mas em vez do quadrado da diferença, utiliza a diferença absoluta, dado por (DEUTSCH e JOURNAL, 1992).

$$y(h_\theta) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} |dn_k(x_i) - dn_k(x_i + h)| \quad (32)$$

3.3.2.3. Variograma cruzado

Quantifica a junção da variabilidade espacial (correlação cruzada) entre duas bandas. Ou seja, definido como a metade do produto médio dos h-pares relativos às faixas radiométricas j, k.

$$y_{jk}(h_0) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \{[dn_j(x_i) - dn_j(x_i + h)].[dn_k(x_i) - dn_k(x_i + h)]\} \quad (33)$$

- j e k – bandas do sensor

3.3.2.4. Pseudo variograma cruzado

O pseudo variograma cruzado representa a semivariância dos pares de pixels cruzados, em vez da covariância direta dos pares, como o variograma cruzado.

$$y_{jk}(h_0) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \{dn_j(x_i) - dn_k(x_i + h)\}^2 \quad (34)$$

Os variogramas apresentados representam diferentes medidas para quantificar a continuidade dos valores de pixel. Os dois primeiros são funções univariadas e os outros dois são funções cruzadas. A diferença absoluta (madograma) e o quadrado da diferença (variograma) realçam o contraste da textura. Por sua vez, as funções cruzadas (variograma cruzado e pseudo variograma cruzado) medem a variância da textura entre duas diferentes bandas radiométricas.

Estes descritores foram calculados usando varias distâncias de separação entre pixels e quatro diferentes ângulos para cada um dos 600 blocos, como mostra Tabela 1 a seguir.

Tabela 1 - Descritores geoestatísticos, ângulos e distâncias de separação entre pixels utilizados

Descritores	Ângulos	Distância
<i>Variograma</i>	0, 45, 90, 135	1, 2, 3, ... ,10
<i>Madograma</i>	0, 45, 90, 135	1, 2, 3, ... ,10
<i>Variograma cruzado</i>	0, 45, 90, 135	1, 2, 3, ... ,10
<i>Pseudo variograma cruzado</i>	0, 45, 90, 135	1, 2, 3, ... ,10

Dessa forma, de acordo com a Tabela 1 e o modo de utilização acima foram utilizados 4 descritores, cada um deles avaliados em 4 ângulos e 10 distâncias, totalizando 160 descritores geoestatísticos.

Os descritores variograma e madograma foram obtidos com base na imagem do excesso de verde e GNDVI utilizando apenas pixels de plantas, ou seja, que não apresentaram valor nulo após a aplicação do *greenness method*. Os algoritmos implementados no MATLAB para a obtenção destes descritores são apresentados nos Apêndices B, C, D e E.

Por outro lado, os descritores variograma cruzado e pseudo variograma cruzado foram obtidos com base nos valores das bandas RGB e falsa cor utilizando apenas pixels de plantas, ou seja, que não apresentaram valor nulo após a aplicação do *greenness method*. As combinações de bandas testadas foram (R,G), (B,G) e (IV,G). A combinação cruzada RxB não foi utilizada pois conteve pouca informação a respeito dos padrões de planta. Os algoritmos implementados no MATLAB para a obtenção destes descritores são apresentados nos Apêndices F, G, H e I.

3.4. Classificação para as amostras de treinamento

Para cada um dos 600 blocos de imagem de cada espécie aplicou-se os descritores de Haralick e geoestatísticos. A avaliação do desempenho destes descritores em reconhecer padrões foi avaliada pela exatidão de classificação proporcionada pelos mesmos ao usar a análise discriminante.

A análise discriminante consiste em obter funções que permitem classificar um padrão, digamos x com base em medidas de um número p de características do mesmo, em uma de varias populações π_j , ($j = 1, 2, \dots, g$), distintas.

Gonzales e Woods (2000) demonstram a obtenção das funções de decisão envolvendo duas classes de padrões ($g = 2$), que apresentam distribuição normal, com médias m_1 e m_2 e desvios padrão σ_1 e σ_2 , respectivamente. Os padrões são escalares, denotados por x . Para esse caso, as funções de decisão, $d_j(x)$, possuem a forma:

$$d_j(x) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left[-\frac{(x - m_j)^2}{2\sigma_j^2}\right] P(w_j) \quad \text{para } j = 1, 2 \quad (35)$$

em que,

σ = desvio padrão da classe;

m_j = vetor de médias das classes; e

$P(w_j)$ = frequência relativa da classe w_j .

Para o caso n dimensional, cada densidade é especificada por seu vetor de média m_j e matriz de covariância C_j . Devido à forma exponencial da densidade gaussiana, utiliza-se então o logaritmo natural da função de densidade de probabilidade (GONZALES e WOODS, 2000). Considerando que a densidade associada a π_j é normal multivariada, então a função discriminante é linear e dada por:

$$d_j(x) = \ln P(w_j) - \frac{1}{2} \ln |C_j| - \frac{1}{2} \left[(x - m_j)^T C_j^{-1} (x - m_j) \right] \quad (36)$$

em que,

x = observação (ou padrão) a ser classificada;

$|C_j|$ = determinante da matriz de covariância da classe j ; e

m_j = vetor de médias da classe j .

De acordo com Gonzales e Woods (2000), supondo que a probabilidade das classes sejam idênticas e que as matrizes de covariância sejam idênticas, chega-se à função de decisão

$$d_j(x) = -\frac{1}{2}[(x - m_j)^T |C|^{-1} (x - m_j)] \quad (37)$$

Desta forma a regra de decisão para a classificação do padrão x é a seguinte:

$$x \in \pi_j \Leftrightarrow d_j = \max(d_1(x), d_2(x), \dots, d_g(x)) \quad (38)$$

Porém na prática raramente se conhece os parâmetros m_j e C . Utilizam-se então as estimativas ao invés dos parâmetros, de modo que o processo de classificação continua essencialmente o mesmo.

Sendo assim, o algoritmo do classificador estatístico baseado na análise discriminante compreendeu os seguintes passos:

- a) leitura dos arquivos que contém os descritores texturais separados em colunas e identificados com suas respectivas classes. A identificação das classes se deve ao fato da classificação ser supervisionada.
- b) Cada descritor contém 600 observações. Cada observação se refere ao valor do descritor para cada um dos 100 blocos de cada uma das 6 imagens de treinamento;
- c) de acordo com o método da validação cruzada deixando um de fora, retira-se a primeira observação, restando 599 observações;
- d) estas 599 observações são separadas por classes;
- e) para cada classe são calculadas a média e a covariância. Em seguida é calculada a variância comum para todas as classes.

- f) para a observação que foi deixada de fora, calcula-se o valor da função de decisão para cada uma das 4 classes (milho, leiteira, capim milha e timbete) , utilizando a Equação 37.

$$d_j(x) = -\frac{1}{2} \left[(x - m_j)^T |C|^{-1} (x - m_j) \right]$$

em que

x = valor do descritor para a observação que foi deixada de fora;

$|C|$ = determinante da matriz de covariância comum; e

m_j = vetor de médias da classe j ;

- g) em seguida atribui-se a observação à classe que apresentar o maior valor

$$[d_j(x)], \text{ ou seja, } x \in \pi_j \Leftrightarrow d_j = \max(d_1(x), d_2(x), \dots, d_g(x))$$

- h) a classificação da observação é então avaliada se foi feita corretamente. O resultado da classificação é salvo.
- i) a primeira observação retorna para o vetor que continha as 599 observações, totalizando 600. Em seguida, retira-se a segunda observação e segue os passos anteriores de “a” até “h”. Desta maneira as 600 observações foram classificadas.
- j) Com base nos resultados das classificações das 600 observações é gerada uma matriz de erro de classificação para o descritor. Com base nesta matriz de erro é calculada a exatidão global, o índice kappa e a variância do índice kappa associado ao descritor.
- k) Os passos de “a” até “j” foram realizados para cada um dos 32 descritores de Haralick e dos 160 descritores geoestatísticos.

O código fonte para o classificador estatístico e a matriz de erro é apresentado no Apêndice K.

3.4.1. Matriz de Erro

A matriz de erro (Tabela 2) favorece a visualização dos resultados da classificação temática e expressa a relação entre os dois tipos de erros associados ao sistema de classificação: os de inclusão (comissão) e os de omissão (exclusão), (CARD, 1982).

Tabela 2 - Matriz de erro de classificação

	1	2	...	c	Total
1	x_{11}	x_{12}	...	x_{1c}	$x_{1.}$
2	x_{21}	x_{22}	...	x_{2c}	$x_{2.}$
...
c	x_{c1}	x_{c2}	...	x_{cc}	$x_{c.}$
Total	$x_{.1}$	$x_{.2}$...	$x_{.c}$	n

As colunas usualmente representam os dados de referência enquanto as linhas indicam os resultados da classificação. A matriz de erros é o ponto de partida para uma série de técnicas estatísticas descritivas e analíticas (CONGALTON, 1991). O autor sugere que a técnica descritiva mais simples talvez seja a exatidão global, que pode ser computada dividindo-se o total de amostras corretamente classificadas (a diagonal da matriz de erros) pelo número total de amostras na matriz. A exatidão global pode ser calculada utilizando-se a equação a seguir:

$$EG = \frac{\sum_{i=1}^c x_{ii}}{n} \quad (39)$$

em que:

EG = exatidão global da classificação;

x_{ii} = valor na linha i e coluna i (diagonal) da matriz de erros;

n = número total de amostras.

A exatidão global (Equação 39), calculada através da matriz de erro, (Tabela 2), mede a concordância entre o número de padrões de cada espécie de planta, classificados corretamente, e o número total de padrões. A quantidade de padrões classificados corretamente é apresentada na diagonal da matriz. Por outro lado, fora da diagonal é apresentada a quantidade de padrões classificados incorretamente.

Após feito a análise de exatidão do classificador pela matriz de erro, seguiu-se para o cálculo do índice kappa, que finalmente avalia o desempenho da classificação.

3.4.2. Índice Kappa

O coeficiente ou índice kappa leva em consideração toda a informação contida na matriz de erros, incluindo os erros de omissão e de comissão. Este índice mede o grau de concordância da classificação digital com a realidade.

$$k = \frac{\sum_{i=1}^c x_{ii} - \sum_{i=1}^c (x_{i+} x_{+i})}{n^2 - \sum_{i=1}^c (x_{i+} x_{+i})} \quad (40)$$

Onde:

k = estimador Kappa de concordância;

n = número de observações (pontos amostrais);

c = número de classes;

x_{ii} = observações na linha i e coluna i da matriz de erros,

x_{i+} = total marginal da linha i ;

x_{+i} = total marginal da coluna i .

Quando o coeficiente kappa é igual a zero o acerto obtido pelo classificador é igual a uma classificação aleatória. Valores positivos de kappa ocorrem quando o acerto é maior que uma classificação aleatória. Por outro lado, valores negativos ocorrem quando o acerto é pior do que uma classificação

aleatória. O maior valor de Kappa (1,00) ocorre quando existe uma classificação perfeita (Cohen, 1960, citado Rosenfield e Fitzpatrick-Lins, 1986).

Segundo Figueiredo e Vieira (2007), não existe uma fundamentação teórica para recomendar os níveis mínimos aceitáveis deste índice numa classificação. Entretanto, na Tabela 3 são apresentados níveis de desempenho da classificação para o valor de Kappa (LEÃO et al., 2007).

O código fonte para obtenção do Índice Kappa é apresentado no Apêndice L.

Tabela 3 - Índice Kappa e correspondente classificação do desempenho da classificação.

Índice Kappa	Desempenho
$k < 0$	Péssimo
$0 < k \leq 0,2$	Ruim
$0,2 < k \leq 0,4$	Razoável
$0,4 < k \leq 0,6$	Bom
$0,6 < k \leq 0,8$	Muito Bom
$0,8 < k \leq 1,0$	Excelente

Um teste estatístico para testar a eficiência dos classificadores foi utilizado por Vieira (2000). Onde um valor de Z foi calculado na matriz de erro. Hipótese Ho: kappa igual a zero.

$$Z = \frac{\hat{k}}{\sqrt{Var(\hat{k})}} \quad (41)$$

Em que *Var* é a variância para grandes amostras do índice kappa e Z é padronizado e normalmente distribuído. Sendo a variância determinada pela Equação 42 (BISHOP et al., 1975 citados por HUDSON e RAMM, 1987):

$$Var(\hat{k}) = \frac{1}{n} \left[\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2 - \theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4 - 4\theta_2^2)}{(1-\theta_2)^4} \right] \quad (42)$$

$$\theta_3 = \sum_{i=1}^c x_{ii} (x_{i+} + x_{+i}) / n^2 \quad (43)$$

$$\theta_4 = \sum_{i=1}^c \sum_{j=1}^c x_{ij} (x_{j+} + x_{+i}) / n^3 \quad (44)$$

Se o valor crítico $Z_{\alpha/2}$ for menor que o $Z_{\text{calculado}}$, a classificação é significativamente melhor que uma classificação ao acaso. Por outro lado para testar a significância de dois índices kappa independentes a um determinado nível de significância tem-se:

$$Z = \frac{(\hat{k}_1 - \hat{k}_2) - (k_1 - k_2)}{\sqrt{\hat{Var}(\hat{k}_1) + \hat{Var}(\hat{k}_2)}} \quad (45)$$

3.4.3. Avaliação dos descritores

Os desempenhos dos descritores geoestatísticos e os de Haralick foram avaliados usando o índice kappa (Equação 40), levando em consideração os efeitos do ângulo entre pixels e o tipo de imagem. Escolheu-se os descritores que apresentaram valores de kappa maiores que 0,2. Conforme a Tabela 3 um kappa com valor 0,2 indica um desempenho de classificação razoável. Na avaliação dos descritores geoestatísticos foi levado também em consideração o efeito da distância entre pixels.

3.4.4. Comparação dos descritores geoestatísticos com os de Haralick

Esta etapa teve por objetivo identificar quais descritores apresentaram melhor desempenho com base na comparação dos valores do índice kappa usando o teste Z. Para isto, foram escolhidos os descritores que apresentaram índice kappa maior que 0,4.

3.4.5. Avaliação da combinação de descritores texturais

Geralmente, um descritor não é capaz de extrair toda a informação que um padrão possui. É possível, portanto, que combinações de descritores texturais sejam capazes de juntos extraírem mais informações do que cada descritor isoladamente. Para avaliar esta possibilidade, diferentes combinações dos descritores texturais foram avaliadas com relação à eficiência da classificação. A comparação do desempenho das combinações foi realizada comparando os índices kappa pelo teste Z e Equação 45.

3.5. Validação do classificador

Esta etapa teve o objetivo de avaliar o poder de generalização de classificação ao utilizar os descritores selecionados na etapa anterior. Para isto, foram utilizadas imagens que não foram empregadas para a seleção dos descritores texturais e/ou combinações de descritores.

Três imagens de cada espécie foram utilizadas. De cada uma destas imagens, foram recortados 100 blocos. De cada um destes blocos foram obtidos os descritores texturais. O vetor de características, contendo um descritor e/ou combinações de descritores, serviu como entrada para o classificador. A eficiência da classificação foi então avaliada pelo índice kappa e pelo teste Z. O algoritmo usado na validação é o mesmo utilizado na classificação para as amostras de treinamento (item 3.4).

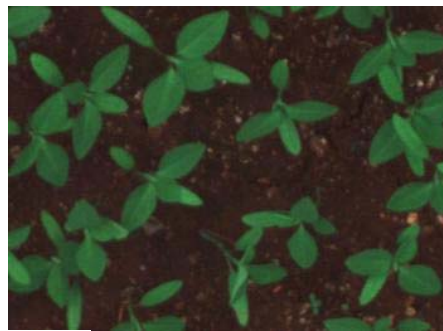
4. RESULTADOS E DISCUSSÕES

4.1. Tipo de imagem coletada

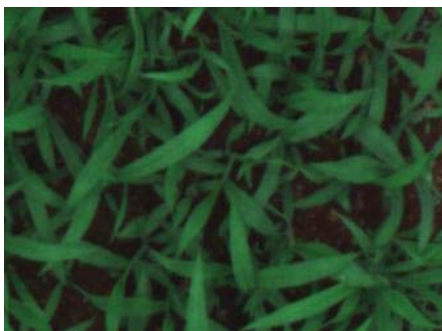
Na Figura 4 e Figura 5 são apresentadas imagens em RGB e falsa cor infravermelha, respectivamente, referentes às espécies avaliadas neste experimento. Estas imagens RGB foram adquiridas aos 29 dias após a emergência com a câmera a uma altura de 2,04m sob uma iluminação de 900 watts de potência.



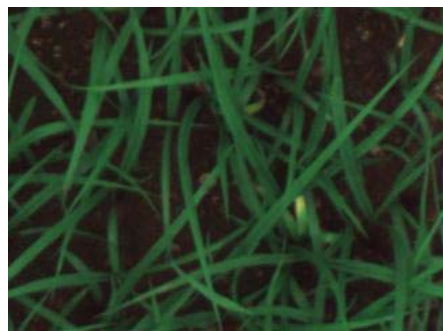
(a)



(b)



(c)



(d)

Figura 4 - Exemplo de imagens RGB adquiridas aos 29 dias após a emergência, nível de iluminação de 900 Watts e altura da câmera de 2,04 m:(a) milho, (b) leiteira, (c) capim milhã e (d) timbête.

Nota-se na Figura 4 e Figura 5 que cada classe apresenta diferentes tipos de dossel, o que solidifica a hipótese que a discriminação pode ser feita por análise textural (ORLANDO, 2003). Em contra partida, verifica-se que a resolução não foi suficiente pra registrar a textura da folha em uma análise visual.

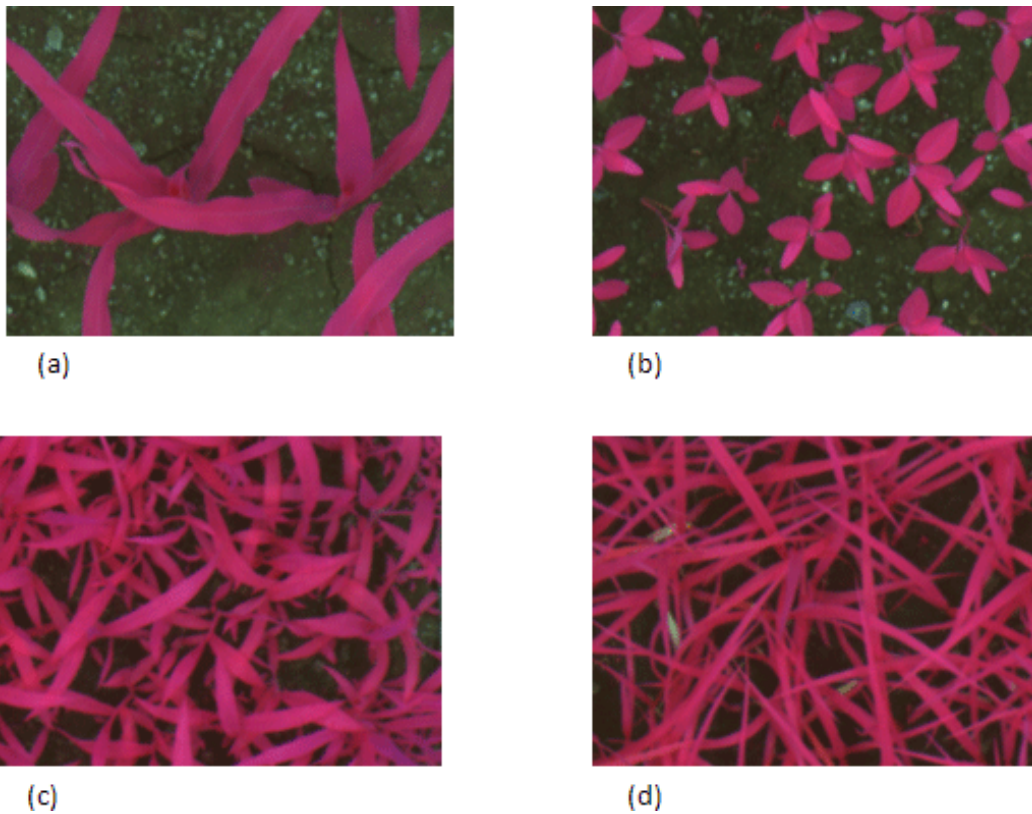


Figura 5 - Exemplo de imagens falsa cor infravermelha adquiridas aos 29 dias após a emergência, nível de iluminação de 900 Watts e altura da câmera de 2,04m: (a) milho, (b) leiteira, (c) capim milhã e (d) timbête.

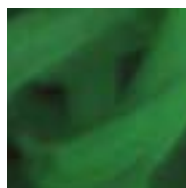


Figura 6 - Bloco de dimensão 68x68 pixels.

Na Figura 6 é apresentado um exemplo de um bloco recortado de uma imagem. Como pode ser visto o tamanho dos blocos (68 x 68 pixels) contém apenas parte da folha.

4.2. Definição das distâncias para os descritores texturais geoestatísticos

Para os cálculos dos descritores geoestatísticos, é necessário estabelecer distâncias específicas. Observando a Figura 7, neste bloco não foi possível obter informação textural usando a semivariância com ângulo de 135° e distância ≥ 20 pixels. Sendo assim, com a intenção de padronizar as distâncias, e sabendo então que o limite máximo não pode ultrapassar 20 pixels, adotou valores de 1 a 10 para todos os outros ângulos no cálculo da semivariância e os demais descritores como mostrado na Figura 8.

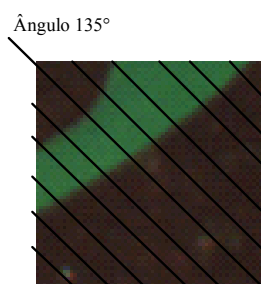


Figura 7 - Recorte de bloco 68x68 pixels da planta milho.

4.3. Descritores texturais geoestatísticos

Esta seção teve como objetivo principal avaliar se o uso da informação da dependência espacial, por meio dos descritores geoestatísticos fornece bons resultados na classificação. Além disso, teve o propósito de avaliar para cada descritor, se o ângulo de pareamento e a distância entre pixels têm efeito no desempenho da classificação. Adicionalmente avaliou-se qual tipo de imagem ou qual combinação de bandas tende a proporcionar melhores resultados na classificação.

4.3.1. Variograma

Na Tabela 4 são apresentados os valores do índice kappa obtidos para o descritor variograma considerando as distâncias de 1 a 10 e ângulos 0, 45, 90 e 135° nas imagens RGB e falsa cor infravermelha. Para corrigir a variação de iluminação, foram obtidas a partir destas imagens, respectivamente, a imagem do excesso de verde e (ExV) e a imagem do índice de vegetação da diferença de verde normalizado (GNDVI).

Observando a Tabela 4, nota-se que a imagem ExV tendeu a apresentar maiores valores de kappa comparando com a imagem GNDVI. Considerando o kappa para a imagem ExV, em cada um dos ângulos, observa-se que este apresentou uma tendência crescente até a distância 4. Após esta distância, o kappa apresentou tendência de queda. Este resultado era esperado, pois espera-se que à medida que a distância entre pixels aumenta, diminui a similaridade entre eles. Não sendo, portanto, conveniente usar informações de um pixel muito distante de um para o qual se deseja classificar.

Deste modo, ao usar um vetor de características obtido a partir de pixels com pouca variação, tem-se uma menor chance de confundimento entre as classes, o que acarreta em um índice kappa maior.

De forma contrária, uma grande distância no cálculo do descritor variograma gera um vetor de características oriundo de pixels não correlacionados. Sendo assim, estes pixels não correlacionados possuem então grande variação entre seus valores. A classificação de um vetor obtido desta forma pode causar em um confundimento entre as classes durante a classificação, resultando assim, em um menor kappa.

Ainda é possível observar uma tendência dos diferentes ângulos avaliados não exercerem influência na classificação. Deste modo temos então um efeito isotrópico (independe da direção). Assim pode dizer que o padrão da autocorrelação entre pixels tende a ser a mesma em todas as direções.

Tabela 4 - Índice kappa ao usar o descritor variograma considerando as distâncias de 1 a 10 e ângulos 0, 45, 90 e 135° para as imagens de excesso de verde (ExV) e índice de vegetação da diferença de verde normalizado (GNDVI)

Distâncias	Ângulos							
	0°	45°	90°	135°	0°	45°	90°	135°
	<i>ExV</i>				<i>GNDVI</i>			
1	0,434	0,377	0,389	0,399	0,305	0,246	0,227	0,253
2	0,428	0,392	0,369	0,408	0,336	0,261	0,220	0,259
3	0,437	0,415	0,402	0,428	0,291	0,286	0,234	0,273
4	0,449	0,400	0,414	0,401	0,295	0,291	0,257	0,285
5	0,431	0,364	0,404	0,182	0,304	0,291	0,251	0,278
6	0,409	0,361	0,378	0,383	0,308	0,285	0,267	0,267
7	0,376	0,359	0,360	0,375	0,311	0,278	0,277	0,257
8	0,356	0,347	0,363	0,364	0,318	0,271	0,276	0,245
9	0,388	0,340	0,370	0,355	0,302	0,262	0,267	0,239
10	0,384	0,312	0,377	0,341	0,291	0,245	0,263	0,233

Na Tabela 5 é apresentada a matriz de erro obtida ao usar o descritor variograma para a imagem ExV nas distâncias 1, 5 e 10 e ângulo 0° ao usar 600 amostras de treinamento. Nesta Tabela podemos observar que, para a distância 1, das 600 observações da classe milho, 555 foram classificadas corretamente, 39 foram classificadas para a classe leiteira, 4 para a classe capim milhã e 2 para a classe timbete.

A classe milho foi a que obteve maior número de acertos na classificação e, por conseguinte menor confundimento entre as classes. Isso pode ser observado também para a distância 5 (distância intermediária) e para a distância 10. Enfim, para todas as distâncias de 1 a 10, a classe milho apresentou maior porcentagem de acerto.

Por outro lado, para as classes leiteira e capim milhã, nas distâncias de 1, 5 e 10 houve confundimento entre elas. Já para a classe timbete, observa-se que na distância 1, 385 observações foram classificadas corretamente das 600. Aumentando as distâncias nota-se que a classe timbete, aumenta também o confundimento entre as classes, ou seja, reduz o número de observações classificadas corretamente.

Tabela 5 - Matriz de erro de classificação obtida ao usar o descritor variograma para a imagem ExV nas distâncias 1, 5 e 10 e ângulo 0° ao usar 600 amostras de treinamento

Distância 1					
<i>Classes</i>	<i>Milho</i>	<i>Leiteira</i>	<i>Capim Milhã</i>	<i>Timbete</i>	<i>Total</i>
<i>Milho</i>	555	72	64	5	696
<i>Leiteira</i>	39	302	238	95	674
<i>Capim Milhã</i>	4	119	139	115	377
<i>Timbete</i>	2	107	159	385	653
<i>Total</i>	600	600	600	600	2400

Distância 5					
<i>Classes</i>	<i>Milho</i>	<i>Leiteira</i>	<i>Capim Milhã</i>	<i>Timbete</i>	<i>Total</i>
<i>Milho</i>	555	68	26	1	650
<i>Leiteira</i>	40	357	211	128	736
<i>Capim Milhã</i>	4	125	138	145	412
<i>Timbete</i>	1	50	225	326	602
<i>Total</i>	600	600	600	600	2400

Distância 10					
<i>Classes</i>	<i>Milho</i>	<i>Leiteira</i>	<i>Capim Milhã</i>	<i>Timbete</i>	<i>Total</i>
<i>Milho</i>	544	99	19	7	669
<i>Leiteira</i>	42	306	191	236	775
<i>Capim Milhã</i>	8	123	310	226	667
<i>Timbete</i>	6	72	80	131	289
<i>Total</i>	600	600	600	600	2400

O confundimento entre as classes leiteira, capim milhã e timbete pode ser devido a estas classes apresentarem padrões de dependência espacial, ou seja, textura semelhantes. Com esta finalidade foi gerado o variograma empírico para cada uma das classes. Como observado anteriormente, os diferentes ângulos aparentaram não exercer influência na classificação. Sendo assim, para a geração do variograma empírico, o valor da semivariância, para cada uma das distâncias, foi obtido a partir da média das semivariâncias obtidas nos ângulos 0, 45, 90 e 135°. O variograma empírico resultante é apresentado na Figura 9. Pode-se verificar que as classes leiteira, capim milhã e timbete tendem a ter padrões de textura semelhantes, pois elas tendem a apresentar valores próximos para o patamar, efeito pepita e alcance. Contudo, observa-se que o padrão da textura para a classe milho tende a ser distinta das demais classes. Esta diferença nos

padrões pode ser uma das explicações para o erro de classificação desta classe ter sido menor do que nas outras classes.

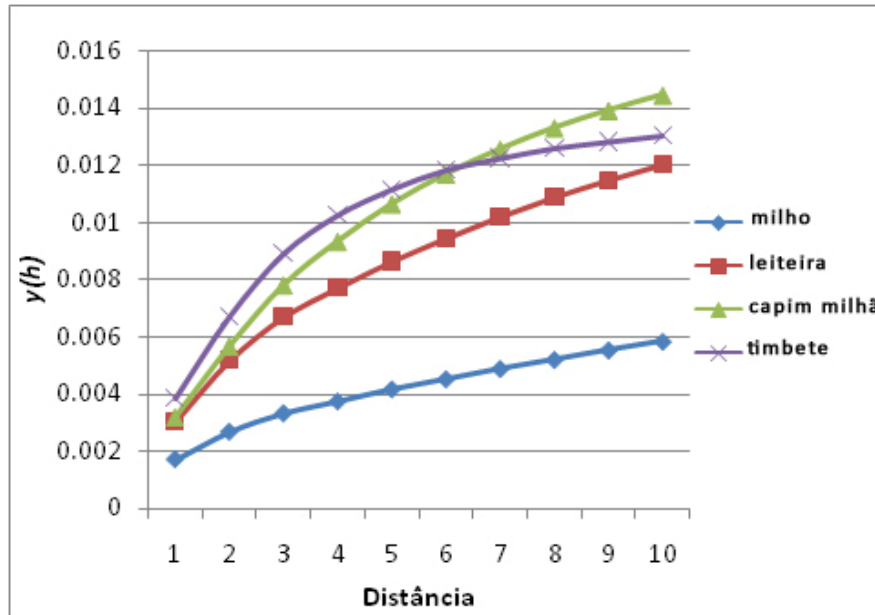


Figura 9 - Valores do variograma empírico obtidos em função da distância para as quatro espécies de plantas avaliadas.

4.3.2. Madograma

De acordo com os índices kappa, apresentados na Tabela 6, obtidos para o descritor madograma, a imagem ExV tendeu a apresentar maiores valores de kappa em relação a imagem GNDVI. Os valores do kappa em função dos ângulos, tipo de imagem e distância apresentaram comportamento semelhante ao apresentado pelo variograma.

Tabela 6 - Índice kappa obtido para o descritor madograma considerando as distâncias de 1 a 10 e ângulos 0, 45, 90 e 135° para as imagens de excesso de verde (ExV) e índice de vegetação da diferença de verde normalizado (GNDVI)

<i>Distâncias</i>	<i>Ângulos</i>							
	<i>0°</i>	<i>45°</i>	<i>90°</i>	<i>135°</i>	<i>0°</i>	<i>45°</i>	<i>90°</i>	<i>135°</i>
	<i>ExV</i>				<i>GNDVI</i>			
1	0,516	0,469	0,478	0,472	0,336	0,277	0,243	0,268
2	0,492	0,460	0,452	0,472	0,354	0,318	0,262	0,291
3	0,502	0,473	0,481	0,483	0,357	0,342	0,258	0,318
4	0,514	0,441	0,480	0,462	0,366	0,359	0,278	0,338
5	0,503	0,421	0,475	0,436	0,379	0,364	0,297	0,350
6	0,487	0,394	0,447	0,399	0,380	0,356	0,306	0,327
7	0,451	0,271	0,430	0,398	0,364	0,332	0,327	0,313
8	0,423	0,364	0,401	0,378	0,362	0,321	0,330	0,291
9	0,399	0,367	0,381	0,382	0,350	0,305	0,318	0,285
10	0,372	0,354	0,392	0,374	0,354	0,297	0,306	0,283

A matriz de erro obtida ao usar o descritor madograma, apresentada na Tabela 7, indica que todas as classes apresentaram comportamento semelhante ao observado ao usar o descritor variograma. A classe milho tendeu a apresentar maior porcentagem de acerto e, por conseguinte, menor confundimento com as demais. As classes leiteira e capim milhã apresentaram-se confundidas e a classe timbete apresentou maior número de observações classificadas incorretamente à medida que a distância aumentou.

Tabela 7 - Matriz de erro de classificação obtida ao usar o descritor madograma para a imagem ExV nas distâncias 1, 5 e 10 e ângulo 0° para as 600 amostras de treinamento

Distância 1						
<i>Classe</i>	<i>Milho</i>	<i>Leiteira</i>	<i>Capim Milhã</i>	<i>Timbete</i>	<i>Total</i>	
<i>Milho</i>	541	54	13	1	609	
<i>Leiteira</i>	54	359	218	37	668	
<i>Capim Milhã</i>	3	148	224	157	532	
<i>Timbete</i>	2	39	145	405	591	
<i>Total</i>	600	600	600	600	2400	

Distância 5						
<i>Classe</i>	<i>Milho</i>	<i>Leiteira</i>	<i>Capim Milhã</i>	<i>Timbete</i>	<i>Total</i>	
<i>Milho</i>	549	57	14	0	620	
<i>Leiteira</i>	46	385	183	58	672	
<i>Capim Milhã</i>	4	130	200	171	505	
<i>Timbete</i>	1	28	203	371	603	
<i>Total</i>	600	600	600	600	2400	

Distância 10						
<i>Classe</i>	<i>Milho</i>	<i>Leiteira</i>	<i>Capim Milhã</i>	<i>Timbete</i>	<i>Total</i>	
<i>Milho</i>	537	92	13	0	642	
<i>Leiteira</i>	54	344	180	153	731	
<i>Capim Milhã</i>	3	70	115	174	362	
<i>Timbete</i>	6	94	292	273	665	
<i>Total</i>	600	600	600	600	2400	

4.3.3. Variograma Cruzado

Na Tabela são apresentados os índices kappa obtidos ao usar o descritor variograma cruzado avaliados nas distâncias de 1,2, ... ,10 e nos ângulos 0, 45, 90 e 135° utilizando combinações das bandas RxG, BxG e IVxG. A combinação RxG tendeu a apresentar valores de kappa menores que zero. Isso significa que a classificação foi pior que uma classificação ao acaso. Já as combinações BxG e IVxG apresentaram um desempenho de classificação variando entre ruim e razoável (Tabela 8).

Tabela 8 - Índice Kappa obtido ao usar o descritor variograma cruzado considerando as distâncias de 1 a 10 e ângulos 0, 45, 90 e 135° para as combinações RxG, BxG e índice de vegetação da diferença de verde normalizado (GNDVI).

<i>Distâncias</i>	<i>Ângulos</i>							
	<i>0°</i>	<i>45°</i>	<i>90°</i>	<i>135°</i>	<i>0°</i>	<i>45°</i>	<i>90°</i>	<i>135°</i>
	RxG				BxG			
<i>1</i>	<i>0,066</i>	<i>0,087</i>	<i>0,063</i>	<i>0,043</i>	<i>0,221</i>	<i>0,206</i>	<i>0,228</i>	<i>0,226</i>
<i>2</i>	<i>0,054</i>	<i>0,064</i>	<i>0,059</i>	<i>0,040</i>	<i>0,238</i>	<i>0,222</i>	<i>0,239</i>	<i>0,239</i>
<i>3</i>	<i>0,039</i>	<i>0,042</i>	<i>0,051</i>	<i>0,031</i>	<i>0,253</i>	<i>0,226</i>	<i>0,236</i>	<i>0,242</i>
<i>4</i>	<i>0,021</i>	<i>0,028</i>	<i>0,042</i>	<i>0,020</i>	<i>0,266</i>	<i>0,221</i>	<i>0,242</i>	<i>0,229</i>
<i>5</i>	<i>0,008</i>	<i>0,007</i>	<i>0,033</i>	<i>0,013</i>	<i>0,247</i>	<i>0,192</i>	<i>0,230</i>	<i>0,209</i>
<i>6</i>	<i>0,002</i>	<i>0,077</i>	<i>0,020</i>	<i>0,051</i>	<i>0,236</i>	<i>0,177</i>	<i>0,214</i>	<i>0,184</i>
<i>7</i>	<i>-0,008</i>	<i>0,087</i>	<i>0,016</i>	<i>0,071</i>	<i>0,220</i>	<i>0,181</i>	<i>0,193</i>	<i>0,198</i>
<i>8</i>	<i>-0,011</i>	<i>0,094</i>	<i>0,063</i>	<i>0,082</i>	<i>0,206</i>	<i>0,180</i>	<i>0,183</i>	<i>0,199</i>
<i>9</i>	<i>0,021</i>	<i>0,106</i>	<i>0,073</i>	<i>0,090</i>	<i>0,196</i>	<i>0,186</i>	<i>0,172</i>	<i>0,184</i>
<i>10</i>	<i>0,033</i>	<i>0,110</i>	<i>0,082</i>	<i>0,100</i>	<i>0,188</i>	<i>0,089</i>	<i>0,184</i>	<i>0,178</i>

IVxG				
<i>Distâncias</i>	<i>0°</i>	<i>45°</i>	<i>90°</i>	<i>135°</i>
<i>1</i>	<i>0,332</i>	<i>0,312</i>	<i>0,295</i>	<i>0,291</i>
<i>2</i>	<i>0,336</i>	<i>0,309</i>	<i>0,307</i>	<i>0,314</i>
<i>3</i>	<i>0,340</i>	<i>0,309</i>	<i>0,318</i>	<i>0,322</i>
<i>4</i>	<i>0,339</i>	<i>0,301</i>	<i>0,313</i>	<i>0,323</i>
<i>5</i>	<i>0,336</i>	<i>0,314</i>	<i>0,329</i>	<i>0,331</i>
<i>6</i>	<i>0,308</i>	<i>0,314</i>	<i>0,330</i>	<i>0,319</i>
<i>7</i>	<i>0,313</i>	<i>0,313</i>	<i>0,327</i>	<i>0,314</i>
<i>8</i>	<i>0,322</i>	<i>0,323</i>	<i>0,317</i>	<i>0,303</i>
<i>9</i>	<i>0,331</i>	<i>0,308</i>	<i>0,319</i>	<i>0,298</i>
<i>10</i>	<i>0,334</i>	<i>0,297</i>	<i>0,316</i>	<i>0,295</i>

4.3.4. Pseudo variograma Cruzado

Na Tabela são apresentados os índices kappa obtidos para o descritor pseudo variograma cruzado avaliados nas distâncias de 1,2, ... ,10 e nos ângulos 0, 45, 90 e 135° ao utilizar combinações das bandas RxG, BxG e IVxG.

Os resultados apresentados (Tabela 9Tabela) seguiu comportamento semelhante ao apresentado pelo variograma cruzado.

Tabela 9 - Índices kappa obtidos ao usar o descritor pseudo variograma cruzado considerando as distâncias de 1 a 10 e ângulos 0, 45, 90 e 135° para as combinações RxG, BxG e índice de vegetação da diferença de verde normalizado (GNDVI).

<i>Distâncias</i>	<i>Ângulos</i>							
	<i>0°</i>	<i>45°</i>	<i>90°</i>	<i>135°</i>	<i>0°</i>	<i>45°</i>	<i>90°</i>	<i>135°</i>
	RxG				BxG			
<i>1</i>	-0,083	-0,088	-0,084	-0,077	0,382	0,348	0,370	0,368
<i>2</i>	-0,066	-0,073	-0,076	-0,062	0,366	0,328	0,350	0,353
<i>3</i>	-0,066	-0,067	-0,069	-0,062	0,352	0,323	0,342	0,335
<i>4</i>	-0,063	-0,067	-0,063	-0,064	0,348	0,306	0,334	0,332
<i>5</i>	-0,066	-0,069	-0,063	-0,075	0,344	0,308	0,336	0,326
<i>6</i>	-0,074	-0,081	-0,069	-0,087	0,332	0,292	0,342	0,318
<i>7</i>	-0,076	-0,091	-0,074	-0,097	0,328	0,287	0,327	0,308
<i>8</i>	-0,083	-0,101	-0,074	-0,101	0,324	0,273	0,321	0,291
<i>9</i>	-0,087	-0,110	-0,079	-0,103	0,317	0,260	0,317	0,280
<i>10</i>	-0,091	-0,119	-0,084	-0,111	0,311	0,255	0,302	0,273
	IVxG							
	<i>Distâncias</i>	<i>0°</i>	<i>45°</i>	<i>90°</i>	<i>135°</i>			
	<i>1</i>	0,222	0,228	0,229	0,229			
	<i>2</i>	0,235	0,241	0,236	0,251			
	<i>3</i>	0,251	0,242	0,242	0,253			
	<i>4</i>	0,254	0,250	0,244	0,253			
	<i>5</i>	0,253	0,245	0,252	0,241			
	<i>6</i>	0,253	0,133	0,250	0,242			
	<i>7</i>	0,253	0,241	0,247	0,243			
	<i>8</i>	0,259	0,235	0,253	0,247			
	<i>9</i>	0,254	0,226	0,250	0,239			
	<i>10</i>	0,236	0,218	0,243	0,238			

4.4. Descritores de Haralick

Esta seção teve como objetivo principal avaliar se os descritores de Haralick fornecem bons resultados na classificação. Além disso, teve o propósito de avaliar para cada descritor se o ângulo tem tendência de efeito no desempenho da classificação. Adicionalmente avaliou-se qual tipo de imagem tende a proporcionar melhores resultados na classificação.

Na Tabela 10 são apresentados os valores do índice kappa obtidos para os descritores de Haralick considerando os ângulos 0, 45, 90 e 135°, e usando as

imagens RGB e falsa cor infravermelha. Para corrigir a variação de iluminação, foram obtidas a partir destas imagens, respectivamente, a imagem do excesso de verde (ExV) e a imagem do índice de vegetação da diferença de verde normalizado (GNDVI).

Ao analisar a Tabela 10, observa-se que para cada descritor os valores do índice kappa tenderam a ser os mesmos para cada um dos tipos de imagem nos quatro ângulos avaliados. Este resultado indica que o ângulo possivelmente não influenciou a exatidão. Além disso, a classificação usando imagem ExV tendeu a apresentar maiores valores de kappa do que a imagem GNDVI para os descritores (1, 4, 6 e 7). O oposto ocorreu para os demais descritores.

Segundo Shearer (1987), a técnica de análise textural com uso da matriz de co-ocorrência além de possuir várias informações texturais, não depende da orientação da imagem. Isso pode ser constatado nos resultados obtidos na Tabela 10. Onde os descritores mostram uma tendência de apresentarem valores semelhantes para os quatro diferentes ângulos.

Tabela 10 - Valores do índice kappa obtidos para os descritores de Haralick (1 - momento angular, 2 - média, 3 - variância, 4 - entropia, 5 - correlação, 6 - momento do produto, 7 - momento inverso da diferença e 8 - medida de correlação) considerando os ângulos 0, 45, 90 e 135° para as imagens de excesso de verde (ExV) e índice de vegetação da diferença de verde normalizado (GNDVI).

Ângulos	Imagens	Descritores							
		1	2	3	4	5	6	7	8
0	ExV	0,304	0,267	0,365	0,275	0,368	0,504	0,458	0,157
	GNDVI	0,123	0,302	0,403	0,123	0,403	0,207	0,202	0,198
45	ExV	0,325	0,264	0,365	0,271	0,362	0,522	0,452	0,154
	GNDVI	0,160	0,301	0,404	0,121	0,408	0,226	0,224	0,217
90	ExV	0,301	0,264	0,367	0,274	0,361	0,482	0,444	0,143
	GNDVI	0,139	0,300	0,404	0,118	0,408	0,148	0,179	0,203
135	ExV	0,319	0,264	0,365	0,271	0,364	0,502	0,447	0,133
	GNDVI	0,119	0,301	0,404	0,121	0,403	0,210	0,212	0,214

4.5. Comparação dos descritores geoestatísticos com os de Haralick

Esta seção teve como objetivo comparar o desempenho dos descritores geoestatísticos com os de Haralick com relação aos kappas obtidos na classificação. Para cada descritor geoestatístico avaliado, foi utilizada a imagem ou combinação de bandas que proporcionou kappa acima de 0,4. Desta maneira, para os descritores variograma, madograma e Haralick foi utilizada a imagem ExV por ter sobressaído em relação às demais combinações. Tais resultados podem ter ocorrido pelo realce do componente verde, executado pela Equação 18. Este, que por sua vez aumenta o contraste dos pixels, caracterizando melhor a textura em relação às demais (ORLANDO, 2003). Resultados semelhantes também foram encontrados por Orlando (2003) ao comparar descritores obtidos a partir de imagem ExV com outros descritores obtidos a partir de imagens monocromáticas e falsa cor. Os demais descritores geoestatísticos não são apresentados pois não obtiveram valores de kappa igual ou superior a 0,4.

A comparação do desempenho dos descritores foi realizada pela estatística Z. Os resultados estão apresentados na Tabela 11. O descritor de Haralick utilizado na comparação foi o momento do produto, ângulo 45°. No madograma utilizou-se a distância 4 e ângulo 0°, e para o variograma a distância 1 e ângulo 0°.

Tabela 11 - Teste Z para comparação dos descritores geoestatísticos com os de Haralick

Descritores	Teste Z unilateral à direita	
	Kappa	Variância de Kappa
Haralick	0.52167* a	0.00017056
Madograma	0.51611* a	0.00017004
Variograma	0.44889* b	0.00017303

* Diferente de zero pelo teste Z a 1% de probabilidade

Os índices kappa seguidos de uma mesma letra não diferiram pelo teste Z a 1% de probabilidade

No teste Z para a comparação dos descritores observa-se que, tanto o descritor de Haralick quanto o madograma apresentaram-se estatisticamente

iguais a 1% de probabilidade. Sendo estes, maiores que o variograma. E todos diferente de zero a 1% de probabilidade.

4.6. Avaliação da combinação de descritores texturais

Na seção anterior foram identificados os descritores que apresentaram melhores desempenhos na classificação. Com o objetivo de obter descritores que proporcionem melhores desempenhos, foram obtidos novos descritores resultantes de combinações entre os descritores selecionados. Espera-se que a combinação destes descritores consiga extrair conjuntamente informações que os mesmos não extraem separadamente. As combinações de descritores avaliadas foram:

- Madograma ângulo 0°;
- Madograma ângulo 0° e o descritor de Haralick momento do produto;
- Madograma ângulo 0° e os descritores de Haralick momento do produto, variância e correlação;
- Madograma ângulo 0° e todos os descritores de Haralick citados na Tabela 10.

O desempenho da classificação obtido por estas combinações é apresentado na Tabela 12.

Tabela 12 - Avaliação da classificação por meio das combinações madograma, variograma e Haralick

Combinações	Exatidão		
	Global	Kappa	Variância kappa
Madograma ângulo 0°	0.75750	0.67667* a	0.00013596
Madograma ângulo 0° e momento do produto	0.79750	0.73000* a	0.00011973
Madograma ângulo 0° e momento do produto, variância e correlação	0.86458	0.81944	0.86721000
Madograma ângulo 0° e todos os descritores de Haralick	0.89625	0.86167	0.68872000

* Diferente de zero pelo teste Z a 1% de probabilidade

Os índices kappa seguidos de uma mesma letra não diferiram pelo teste Z a 1% de probabilidade

Os índices kappa do classificador desenvolvido com os valores das duas últimas combinações da Tabela 12 foram iguais a zero ($P > 0,01$), pelo teste Z, a 1% de probabilidade. Isso equivale a dizer que estas classificações foram estatisticamente iguais a uma classificação ao acaso. Já as duas primeiras combinações apresentaram índices kappa diferente de zero. Sendo 67,66% e 73% melhor que uma classificação ao acaso respectivamente. Estas combinações apresentaram índices kappa estatisticamente iguais ($P > 0,01$) (Tabela 12). Sendo assim, a primeira combinação foi escolhida para compor o vetor de característica a ser usado na validação. A justificativa do seu uso é, além de ser estatisticamente igual a segunda combinação, usa menos informação e consegue atingir o mesmo resultado.

4.7. Validação

O próximo passo, na avaliação de um descritor é validar o seu desempenho ao usar padrões não utilizados na seleção do vetor de características. A validação visa verificar a habilidade do classificador no que diz respeito à generalização para amostras diferentes das utilizadas no treinamento. Na etapa anterior foi selecionado o vetor obtido pela combinação das 10 distâncias no ângulo 0° do madograma. Para validar o desempenho deste vetor no classificador, foram utilizados 300 blocos de imagens para cada espécie, sendo estes blocos diferentes dos utilizados no treinamento para escolha do descritor.

Sendo assim, ao usar estas imagens destinadas para a validação, a exatidão global alcançada foi de 80,66% com índice kappa igual a 0,7422. Sendo este índice kappa diferente de zero a 1% de probabilidade pelo teste Z. Que por sua vez, é 74,22% melhor que uma classificação ao acaso. De acordo com a Tabela 3 este valor de Kappa indica um desempenho de classificação muito bom.

Orlando (2003), ao utilizar descritores de Haralick obtidos de imagem excesso de verde em um classificador estatístico e outro neuronal em diferentes condições de iluminação, obteve exatidões globais variando de 75,3% a 86,5%.

5. CONCLUSÕES

Diante os resultados alcançados a hipótese do trabalho é portanto aceita. Descritores texturais geoestatísticos são tão eficientes em reconhecer padrões quanto descritores não-geoestatísticos. Tendo em vista os objetivos específicos, os resultados obtidos indicaram as seguintes conclusões:

- a imagem que tendeu a apresentar os melhores resultados de exatidão global foi a imagem excesso de verde.

- dentre os descritores geoestatísticos o madograma apresentou maiores valores de kappa variando de 0,271 a 0,516. Sendo este, estatisticamente igual ao descritor de Haralick.

Com o objetivo de suprir as necessidades de estudos relacionados a reconhecimento de padrões com uso de descritores geoestatísticos, sugere-se para trabalhos futuros os seguintes estudos:

- usar como característica na classificação os parâmetros do variograma,
- uso de Wavelets (função que descompõe outras funções no domínio da frequência),
- uso de um método Step Forward para checar quais as melhores variáveis discriminantes geoestatísticas.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALVES, E.A., KHOURY JUNIOR, J.K., PINTO, F.A.C., QUEIROZ, D.M. **Classificação de plantas daninhas utilizando características texturais em imagens digitais.** In: 2º SIMPÓSIO INTERNACIONAL DE AGRICULTURA DE PRECISÃO – SIAP, Viçosa, MG. 2002.

ANTUNIASSI, W.R.; GARDANHA JUNIOR, C.D. **Aplicação localizada de produtos fitossanitários.** In: BORÉM, A.; GIÚDICE, M.P.; QUEIRÓZ, D.M.; MANTOVANI, E.C.; FERREIRA, L.R.; VALLE, F.X.R.; GOMIDE, R.L. (Ed.) Agricultura de precisão. Viçosa: Giúde & Borén, 2000. p.181-202.

ATKINSON, P.M.; LLOYD, C.D., **Mapping precipitation in Switzerland with ordinary and indicator kriging.** Journal of Geographic Information and Decision Analysis, v.2, n.2, p.72-86, 1998.

BERBEROGLU, S., LLOYD, C.D., ATKINSON, P.M., CURRAN, P.J. The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. **Computers & Geosciences.** Vol. 26, 385-396 p, 2000.

BERNARDI, H.V.F.; DZEDZEJ, M.; CARVALHO, L.M.T.; JÚNIOR, F.W.A. **Classificação digital do uso do solo comparando os métodos “pixel a pixel” e orientada ao objeto em imagem QuickBird.** In: XIII Simpósio Brasileiro de Sensoriamento Remoto, 2007, Florianópolis. Anais... Florianópolis: Simpósio Brasileiro de Sensoriamento Remoto. P 5595-5602. 2007.

BURKS, T.F.; SHEARER, S.A.; GATES, R. S.; DONOHUE, K. D. Backpropagation neural network design and avaluation for classifying weed species using color image texture. **Transactions of the ASAE,** v. 43, nº 4, p. 1029-1037, 2000.

CAMPOS, E.T. **Técnicas de seleção de atributos e de classificação para reconhecimento de faces.** São Paulo: USP - IME- Departamento de Ciência da Computação. Exame de Qualificação. 73p. 2000.

CARD, D.H. **Using known map category marginal frequencies to improve estimates of thematica map accuracy.** Photogrammetric Engineering and Remote Sensing. V 48, n. 3, p. 431-439. 1982.

CARR, J. R., Spectral and Textural Classification of single and multiple band digital images. **Computers & Geosciences,** vol. 22, nº 8, pp. 849-865, 1996.

CHANCELLOR, W. J.; GORONEA, M. A. Effects of spatial variability of nitrogen, moisture, and weeds on the advantages of site-specific application for wheat. **Transactions of the ASAE**, v. 37, n.º 3, p. 717-724, 1993.

CONGALTON, R.G.; MEAD, R.A. **A quantitative method to test for consistency and correctness in photointerpretation**. Photogrammetric engineering & remote sensing, Bethesda, v.49 n. 1, p.69-74, 1983.

CONGALTON, R. G. A review of assessing the accuracy of classification of remotely sensed data. **Remote Sensing of Environment**, v. 37, n. 1, p. 35-46, 1991.

CRESSIE, N. A. C. **Statistics for spatial data**. New York: John Wiley. 900p. 1993.

DAVID, MICHEL. **Geoestatística ore reserve estimation**. New York, Elsevier Scientific, 1977. 363p.

DEUTSCH, C.V., JOURNAL, A.G. GLSIB. **Geostatistical Software Library and User's**. Oxford University Press, New York 340 p. 1992.

FAN, G.; ZHANG, N.; PETERSON, D. E.; LOUGHIN, T. M. Real-time weed detection using machine vision. **ASAE annual International Meeting**. Orlando, Florida. Paper n.º 983032, 17 p, 1998.

FELGUEIRAS, C. A. **Modelagem ambiental com tratamento de incertezas em Sistemas de informação geográfica: O paradigma geoestatístico por indicação**. São José dos Campos. Tese (Doutorado)-Ministério da Ciência e Tecnologia, Instituto Nacional de Pesquisas Espaciais – INPE. 181p. 1999.

FIGUEIRA L. B.; COLAFEMINA J. F.; ROQUE, A.C. **Diagnóstico de patologias do sistema vestibular utilizando redes neurais na análise dos movimentos sacádicos**. Disponível em : [HTTP://ftp.inf.pucpcaldas.br/CDs/SBC2003/pdf/arq0172.pdf](http://ftp.inf.pucpcaldas.br/CDs/SBC2003/pdf/arq0172.pdf). Acesso em 20 set. 2006.

FIGUEIREDO, G.C.; VIEIRA, C.A.O. **Estudo do comportamento dos índices de exatidão global, kappa e tau, comumente usados para avaliar a classificação de imagens do sensoriamento remoto**. In: XIII Simpósio Brasileiro de Sensoriamento Remoto, 2007, Florianópolis, INPE. Anais... Florianópolis: Simpósio Brasileiro de Sensoriamento Remoto. p 5755-5762. 2007.

GANAN, J.R.; ROCHA, J.V.; MERCANTE, E.; ANTUNES, J.F.G. Mapeamento da cultura da soja com imagens Landsat 5/TM utilizando algoritmos de classificação supervisionada. **AnaisXII** Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16-21, abril 2005, INPE, p. 549-555.

GONÇALVES, A. C. A. Variabilidade espacial de propriedades físicas do solo para fins de manejo da irrigação. Piracicaba, 1997. 118p. Tese – (Doutorado) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo.

GONZALES, R.C.; WOODS, R.E. **Processamento de imagens digitais**. Editora Edgard Blücher Ltda, São Paulo, 509p, 2000.

GUERRA, P.A.G. **Geoestatística operacional**. Brasília: Ministério das Minas e Energia, 1988. 145p.

HARALICK, R. M., SHANMUGAM, K., “Textural features for image classification,” **IEEE Trans. Syst., MAN, Cybern.**, vol. SMC-3, pg 610-621, Nov. 1973.

HERZFELD, U.C.; CHRIS, A.H. automated geoestatística se floor classification-principles, parameters, feature vectors, and discrimination criteria. *Computers & Geosciences*. v.22, n.1, p.35-52. 1996.

HUDSON, W.D.; RAMM, C.W. Correct formulation of the kappa coefficient os agreement. *Photogrammetric engineering & remote sensing*, Bethesda, v.53 n.4, p.421-422, 1987.

HUMMEL, J. W.; STOLLER, E. W. On-the-go weed sensing and herbicide application for the Northern Cornbelt. **The 2002 ASAE Annual International Meeting**. Chicago, IL. Paper nº 021021, 13p, 2002.

JAIN, A.K.; ROBERTH, P.W.; et. al. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):4-37, Janeiro 2000.

JAYAS, D.S.; PALIWAL, J.; VISEN, N.S. Multi-layer neural networks for image analysis of agricultural products. **Journal of agricultural engineering research**, London, v.2 n.77, p.119-128, 2000.

JOURNEL, A. G.; HUIJBREGTS, C. J. *Mining Geostatistics*, Centre de Geoestatistique, Fontainebleu, France, New York: Centre de Geoestatistique, Academic Press, 1978. 600p.

KITANIDIS, P. K. **Introduction to geostatistcs applications to hydrology**. Cambridge: University Press, 1997. 249p.

KOHN, A.F. Reconhecimento de padrões, uma abordagem estatística. Apostila, Escola Politécnica, USP, 1998.

LACAZE, B., Rambal, S., Winkel, T., 1994. Identifying spatial patterns of Mediterranean landscapes from Geostatistical analysis of remotely-sensed data. *International Journal of Remote Sensing* 15, 2437±2450.

LARK, R.M. Geostatistical description of texture on an aerial photograph for discriminating classes of land cover. **International Journal of Remote Sensing**. Vol. 17, pp 2115-2133, 1996.

LEÃO, C.; KRUG, L.A.; KAMPEL, M.; FONSECA, L.M.G. Avaliação de métodos de classificação em imagens TM/Landsat e CCD/CBERS para o mapeamento do uso e cobertura da terra na região costeira do extremo sul da Bahia. **Anais XIII** Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Brasil, 21-26, Abril 2007, INPE, p. 939-946.

LUTMAN, P.J.W.; PERRY, N.H. Methods of weed patch detection in cereal crops. In: THE 1999 BRIGHTON CONFERENCE – WEEDS, Brighton, 1999a. **Proceedings**. Brighton: BCPC, 1999. P.627-634.

MATSUYAMA, T., 1980. Structural analysis of natural textures by Fourier transformation. *Computer Vision, Graphics and Image Processing* 12, 286-308.
MATHERON, G. The theory of regionalized variables and its applications. **Les cahiers du CMM**, n.5, 211p. 1971.

MIRANDA F. P., Macdonald J. A., Can J. R., 1992, Application of the semivariogram textural classifier (STC) for vegetation discrimination using SIR-B data of Borneo, *International Journal of Remote Sensing*, Vol. 13, No. 12, 2349-2354.

MIRANDA, F.P.; FONSECA, L.E.N.; TARANIK, J.V. Application of the Semivariogram textural classifier (STC) for vegetation discrimination using JERS-1SAR data of northwestern Brazil, In: **ISPRS Commission VII Symposium – Resource and Environmental Monitoring**, Rio de Janeiro, Sept. 26-30, p 151-158, 1994.

MIRANDA, F.P.; FONSECA, L.E.N., CARR, J.R.; TARANIK, J.V. Analysis of JERS-1SAR data for vegetation discrimination in northwestern Brazil using the Semivariogram textural classifier, (STC). In: **International Journal of Remote Sensing**, 1995.

MIRANDA, F.P., FONSECA, L.E. N., CAN J. R., RARANIK, J.V. Analysis of JERS-1 (Fuyo-1) SAR data for vegetation discrimination in northwestern Brazil using the semivariogram textural classifier (STC). In: *International Journal of Remote Sensing*, Vol. 17, No. 17, p3523-3529, 1996.

MIRANDA, R.P., FONSECA, L.E.N., CARR, J.R. Semivariograma textural classification of JERS-1 (Fuyo-1) SAR data obtained over a flooded area of the Amazon rainforest. **International Journal of Remote Sensing**. Vol.19, p 549-556, 1998.

OLIVEIRA, M.C.N. Métodos de estimação de parâmetros em modelos geostatísticos com diferentes estruturas de covariâncias: uma aplicação ao teor de cálcio no solo. Piracicaba: Escola Superior de Agricultura Luiz de Queiroz, 2003. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura Luiz de Queiroz, 2003.

OLMO, M.C.; HERNÁNDEZ, F.A. Computing Geostatistical image texture for remotely sensed data classification. **Computers & Geosciences**. Vol. 26, pp 373-383, 2000.

ORLANDO, R. C. Sistema de visão artificial para discriminação de plantas daninhas e milho. Viçosa, MG: UFV, 2003. 97p. (Tese Doutorado).

PARK, B.; CHEN Y.R. Co-occurrence matrix texture feature of multi-spectral images on poultry carcasses. **Journal Agricultural Engineering Research**, v. 2, n°.78, p. 127-139, 2001.

RAMSTEIN, G., RAFY, M., 1989. Analysis of the structure of radiometric remotely-sensed images. *International Journal of Remote Sensing* 10, 1049-1073.
RAO, A.R. A Taxonomy for texture description and identification. Nova York: Springer-Verlag.

RIBEIRO JUNIOR, P.J. Métodos geoestatísticos no estudo da variabilidade espacial de parâmetros do solo. Piracicaba. Dissertação (Mestrado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo. P99. 1995.

ROSENFELD, G. H.; FITZPATRICK, K. A coefficient of agreement as a measure of thematic classification accuracy. **Photogrammetric Engineering & Remote Sensing**, v. 52, n. 2, Feb.. p. 233-227. 1986.

SAMPER, F.J.S.; RAMIREZ, J.C. Geoestadística: **Aplicaciones a la hidrogeología subterránea**. Barcelona: Gráficas Torres, 1990. 484p.

SAS Institute. **SASOnline Doc**. Version 8, SAS Institute Inc., Cary, USA, 1999.

SENA JUNIOR, D.G. **Utilização de técnicas de visão artificial para ajuste da adubação nitrogenada em trigo**. Viçosa: Universidade Federal de Viçosa. 2005. 160p. Tese (Doutorado em Engenharia Agrícola) – Universidade Federal de Viçosa. 2005.

SHEARER, S. A.; HOLMES, R. G. Plant identification using color co-occurrence matrices. **Transactions of the ASAE**, v. 33, nº 6, p. 2037-2044, 1990.

SILVA, A.C.; CARVALHO, P.C.P. Analysis of spatial variability using Geostatistical functions for diagnosis of lung nodule in computerized tomography images. In: *Pattern Anal Applic*, 7. 2004, p. 227-234. London.

SOARES, A. **Geoestatística para ciência da terra e do ambiente**. Lisboa: Instituto Superior Técnico, 206p, 2000.

STAFFORD, J.V.; MILLER, P.C.H. Spatially variable treatment of weed patches. In: INTERNATIONAL CONFERENCE ON PRECISION AGRICULTURE, 3., Minneapolis. 1996. **Proceedings**. Wisconsin: ASA, CSSA, SSSA, 1996. P.465-474.

STEWARD, B. L., TIAN, L. F. Real-time machine vision weed-sensing. **The 1998 ASAE Annual International Meeting**. Orlando, Florida. Paper nº 983033. 17p. 1998.

TANG, L.; TIAN, L. F.; STEWARD, B. L.; REID, J. F. Texture-based weed classification using Gabor Wavelets and Neural Network for real-time selective herbicide applications. **ASAE/CSAE-SCGR Annual International Meeting**. Toronto, Ontario. Paper nº 993036, 10 p, 1999.

THEODORIDIS, S. *Pattern Recognition*. Academic Press, USA, 1st edition, 1999.

TIAN, L.; REID, J.F.; HUMMEL, J.W. Development of a precision sprayer for site-specific weed management. **Transactions of the ASAE**. St. Joseph, MI. v. 42, nº.4, p. 893 – 900. 1999.

TOU, J.T. e GONZALES, R.C. *Pattern recognition principles*: Addison-Wesley Publishing Company, Reading, p 377, 1974.

TRANGMAR, B. B.; YOST, R. S.; UEHARA, G. Application of geostatistics to spatial studies of soil properties. **Advanced in Agronomy**, v.38, p.45-94, 1985.
VALENTE, J.M.G.P. *Geomatématica. Lições de Geoestatística*. Ouro Preto: Fundação Gorceix, v.3, p.511-713. 1989.

VARELLA, C.A.A.; KHOURY JR., J.K.; PINTO, F.A.C.; FRANÇA, G.E. Estimativa da deficiência de nitrogênio na cultura do milho utilizando características de imagens digitais para uso em sensoriamento remoto. *Anais XI SBSR*, Belo Horizonte, p. 2391 – 2398, 2003.

VIEIRA, C.A.O. Accuracy of remotely sensing classification of agricultural crops: a comparative study. Ph.D. Thesis, University of Nottingham, Nottingham, 323p. 2000.

VIEIRA, S. R.; HATFIELD, J.L.; NIELSEN, D. R.; DIGGAR, J. W. Geostatistical theory and application to variabilidade of some agronomical properties. **Hilgardia**, v.51, n.3, p.1-75, 1983.

WATANABE, S. *Pattern Recognition: Human and Mechanical*. New York: Wiley, 1985.

WEBB, A. **Statistical pattern recognition**. 2.ed. England: John Wiley & Sons, 2004. 491p.

WILLIS, B. D.; STOLLER, E. W. Weed suppression for vegetation management in corn and soybeans. **Proc. N. Central Weed. Sci. Soc.** v. 45, n.9. 1990.

WOODCOCK, C. E. et al., 1988a, The use of variograms in remote sensing: I real digital images, *Remote Sensing of Environment* 25:323-348.

WOODCOCK, C. E. et al., 1988b, The use of variograms in remote sensing: II real digital images, *Remote Sensing of Environment* 25:349-379.

YANG, C. C.; PRASHER, S. O.; LANDRY, J. A. Application of artificial neural networks to image recognition in precision farming. **ASAE Annual International Meeting**. Orlando Florida. Paper n° 983039, 10 p, 1998.

YANG, C. C.; PRASHER, S. O.; LANDRY, J. A. Weed recognition in precision farming. **ASAE Annual International Meeting**. Toronto, Ontario. Paper n° 993115, 10 p, 1999.

YANG, C.C.; PRASHER, S.O. e LANDRY, J.A. 2002. Weed recognition in corn fields using back-propagation neural network models. *Canadian Biosystems Engineering/Le genie des biosystèmes au Canada* 44: 7.15-7.22.

ZHANG, N.; CHAISATTAPAGON, C. Effective criteria for weed identification in wheat fields using machine vision. **Transactions of the ASAE**. St. Joseph, MI v. 38, n°3, p. 965-974. 1995.

APÊNDICES

Apêndice A

Código-fonte do programa de computador implementado no MATLAB para recorte dos sub-blocos das imagens

```
% Algoritmo para corte dos sub-blocos
% Autor DANILO PEREIRA BARBOSA & ANTONIO CARLOS FAVA DE BARROS
% marco de 2008

clear
clc

[nome, caminho]=uigetfile('*.tif','imagem a processar');
conta=0;

if ~ischar(nome)
    warndlg('Imagem nao Fornecida!!');
else
    arquivo=fullfile(caminho,nome);
    ima=imread(arquivo);
end

imshow(ima);
[col lin r]=impixel(ima)
local=strcat(nome, '.txt')
fid=fopen(local, 'w')
arq=[col lin]
fprintf(fid, '%4.0f %4.0f\n', arq)
conta = size(col,1)
fclose(fid)
xlswrite('planilha_arquivo', arq);
arq=xlsread('planilha_arquivo');

lin_inicial=arq(1,2);
col_inicial=arq(1,1);
col_final=arq(2,1);
lin_final=arq(3,2);

lin=lin_inicial;
col=col_inicial;
i=1;
while lin<(lin_final-67)
    while col<(col_final-67)
        amostra=ima(lin:lin+67,col:col+67,:);
        nome=strcat('bloco',num2str(i),'.tif');
        imwrite(amostra,nome);
        col=col+67;
        i=i+1;
    end
    col=col_inicial;
    lin=lin+67;
end
end
```

Apêndice B

Código-fonte do programa de computador implementado no MATLAB para o cálculo da semivariância nos ângulos 0, 45, 90 e 135°

```
% algoritmo para calcular semivariância usando "greenness method"  
% DANILO PEREIRA BARBOSA & ANTONIO CARLOS FAVA DE BARROS
```

```
function [y_0, y_45, y_90, y_135] = semivariance(dados, p)
```

```
[lin,col]=size(im2double(dados));  
for lag=p % distância
```

```
    n=0;  
    total=0;  
    % calculando pra ângulo 0  
    for i=1:lin  
        for j=1:(col-lag)  
            if ((dados(i,j)~=0) && (dados(i,j+lag)~=0))  
                total= total + (dados(i,j)-dados(i,j+lag))^2; %  
                n= n+1;  
            end  
        end  
    end  
    y_0=total/(2*n);
```

```
    clear n total
```

```
% calculando pra ângulo 90  
n=0;  
[lin,col]=size(dados);  
total=0;  
for j=1:(col)  
    for i=1:(lin-lag)  
        if ((dados(i,j)~=0) && (dados(i+lag,j)~=0))  
            total= total + (dados(i,j)-dados(i+lag,j))^2;  
            n= n+1;  
        end  
    end  
end  
y_90=total/(2*n);
```

```
clear n total
```

```
% calculando pra ângulo 45
```

```
n=0;  
total= 0;  
for i= 1:lin  
    for j= 1:col  
        mlin=i+lag;  
        mcol=j-lag;  
        if((mlin<=lin)&&(mlin>0)&&(mcol<=col)&&(mcol>0))
```

```

                if ((dados(i,j)~=0) && (dados(mlin,mcol)~=0))
                    total=total+(dados(i,j)-dados(mlin,mcol))^2;
                    n=n+1;
                end
            end
        end
    end
    y_45= total/(2*n);

    clear n total
    % calculando pra ângulo 135

    n=0;
    total= 0;
    for i= 1:lin
        for j= 1:col
            mlin=i+lag;
            mcol=j+lag;
            if((mlin<=lin)&&(mlin>0)&&(mcol<=col)&&(mcol>0))
                if ((dados(i,j)~=0) && (dados(mlin,mcol)~=0))
                    total=total+(dados(i,j)-dados(mlin,mcol))^2;
                    n=n+1;
                end
            end
        end
    end
    y_135= total/(2*n);

    clear n total

end

```

Apêndice C

Código-fonte do programa de computador implementado no MATLAB para obtenção das semivariâncias usando “Greenes Method” e excesso de verde.

```
% Algoritmo para obtenção das semivariâncias
% Autor DANILO PEREIRA BARBOSA
% marco de 2008

clear
clc

lag=1 % distância para o calculo da semivariância

for n=1:600

    % lendo as amostras
    A=imread(strcat('D:\execução da
dissertação\bloco',num2str(n),'.tif'));

    [lin,col]=size(A(:,:,1));
    dados= zeros(size(A(:,:,1)));
    % faz o calculo do excesso de verde somente para pixels que contém
    % valores maiores na banda G do que nas outras duas bandas R e B.
    Para os demais pixels atribui-se valor zero (0).
    for i= 1:lin
        for j= 1:col
            if ((A(i,j,2)>A(i,j,1)) && (A(i,j,2)>A(i,j,3)))
                dados(i,j)=(double(A(i,j,2)).*2-double(A(i,j,1))-
double(A(i,j,3)))./(double(A(i,j,1))+double(A(i,j,2))+double(A(i,j,3))
);
                else
                    dados(i,j)= 0;
                end
            end
        end
    end

    % salva as semivariâncias com seus respectivos ângulos
    % grava y_0° em y(:,1), y_45° em y(:,2), etc...
    [y(n,2), y(n,3), y(n,4), y(n,5)] = semivariance(dados,lag)

end

nome_salva=(strcat('matriz_semivariância_',num2str(lag),'.xls'));
xlswrite(nome_salva,y) ;
```

Apêndice D

Código-fonte do programa de computador implementado no MATLAB para o cálculo do madograma nos ângulos 0, 45, 90 e 135°

```
% algoritmo para calcular o madograma usando "greenness method"
% DANILO PEREIRA BARBOSA

function [y_0, y_45, y_90, y_135] = madogram(dados, p)

[lin,col]=size(dados);
for lag=p

    n=0;
    total=0;
%calculando para o ângulo 0
    for i=1:lin
        for j=1:(col-lag)
            if ((dados(i,j)~=0) && (dados(i,j+lag)~=0))
                total= total + abs(dados(i,j)-dados(i,j+lag)); %
calcula e totaliza
            end
            n= n+1;
        end
    end
    y_0=total/(2*n);

    clear n total

%calculando para o ângulo 90
    n=0;
    [lin,col]=size(dados);
    total=0;
    for j=1:(col)
        for i=1:(lin-lag)
            if ((dados(i,j)~=0) && (dados(i+lag,j)~=0))
                total= total + abs(dados(i,j)-dados(i+lag,j)); %
calcula e totaliza
            end
            n= n+1;
        end
    end
    y_90= total/(2*n);

    clear n total

%calculando para o ângulo 45
    n=0;
    total= 0;
    for i= 1:lin
        for j= 1:col
            mlin=i+lag;
```

```

        mcol=j-lag;
        if((mlin<=lin)&&(mlin>0)&&(mcol<=col)&&(mcol>0))
            if ((dados(i,j)~=0) && (dados(mlin,mcol)~=0))
                total=total+abs(dados(i,j)-dados(mlin,mcol));
                n=n+1;
            end
        end
    end
end
end
y_45= total/(2*n);

clear n total
%calculando para o ângulo 135
n=0;
total= 0;
for i= 1:lin
    for j= 1:col
        mlin=i+lag;
        mcol=j+lag;
        if((mlin<=lin)&&(mlin>0)&&(mcol<=col)&&(mcol>0))
            if ((dados(i,j)~=0) && (dados(mlin,mcol)~=0))
                total=total+abs(dados(i,j)-dados(mlin,mcol));
                n=n+1;
            end
        end
    end
end
end
y_135= total/(2*n);

clear n total

end

```

Apêndice E

Código-fonte do programa de computador implementado no MATLAB para obtenção do madograma usando “Greenness Method” e excesso de verde.

```
% Algoritmo para obtenção das semivariâncias
% Autor DANILO PEREIRA BARBOSA
% marco de 2008

clear
clc

lag=1 % distância para o calculo do madograma

for n=1:600

    % lendo as amostras
    A=imread(strcat('D:\execução da
dissertação\bloco',num2str(n),'.tif'));

    [lin,col]=size(A(:,:,1));
    dados= zeros(size(A(:,:,1)));
    % faz o calculo do excesso de verde somente para pixels que contém
    % valores maiores na banda G do que nas outras duas bandas R e B.
    Para
    % os demais pixels atribui-se valor zero (0).
    for i= 1:lin
        for j= 1:col
            if ((A(i,j,2)>A(i,j,1)) && (A(i,j,2)>A(i,j,3)))
                dados(i,j)=(double(A(i,j,2)).*2-double(A(i,j,1))-
double(A(i,j,3)))./(double(A(i,j,1))+double(A(i,j,2))+double(A(i,j,3))
);
                else
                    dados(i,j)= 0;
            end
        end
    end

    % salva o madograma com seus respectivos ângulos (y_0° em y(:,1),
y_45°
% em y(:,2), etc...)
    [y(n,2), y(n,3), y(n,4), y(n,5)] = madogram(dados,lag) % grava

end

nome_salva=(strcat('matriz_madograma_',num2str(lag),'.xls'));
xlswrite(nome_salva,y) ;
```

Apêndice F

Código-fonte do programa de computador implementado no MATLAB para o cálculo do variograma cruzado nos ângulos 0, 45, 90 e 135°

```
% algoritmo para calcular o variograma cruzado usando "greenness
method"
% DANILO PEREIRA BARBOSA

function [y_0, y_45, y_90, y_135] = cross_variogram(dados, p)

% combinação BxG
dados1= im2double(dados(:,:,3));%dados1 é a banda B de dados
dados2= im2double(dados(:,:,2));%dados2 é a banda G de dados

% combinação RxG
% dados1= im2double(dados(:,:,2));%dados1 é a banda G de dados
% dados2= im2double(dados(:,:,1));%dados2 é a banda R de dados

[lin,col]=size(im2double(dados1));
for lag=p

    n=0;
    total=0;
% calculando para o ângulo 0
    for i=1:lin
        for j=1:(col-lag)
            if ((dados1(i,j)~=0) && (dados1(i,j+lag)~=0) &&
(dados2(i,j)~=0) && (dados2(i,j+lag)~=0))
                total= total + (dados1(i,j)-
dados1(i,j+lag))*(dados2(i,j)-dados2(i,j+lag)); % calcula e totaliza
                n= n+1;
            end
        end
    end
    y_0=total/(2*n);
    clear n total

% calculando para o ângulo 90
    n=0;
    total=0;
    for j=1:(col)
        for i=1:(lin-lag)
            if ((dados1(i,j)~=0) && (dados1(i+lag,j)~=0) &&
(dados2(i,j)~=0) && (dados2(i+lag,j)~=0))
                total= total + (dados1(i,j)-
dados1(i+lag,j))*(dados2(i,j)-dados2(i+lag,j)); % calcula e totaliza
                n= n+1;
            end
        end
    end
```

```

end
y_90= total/(2*n);

clear n total

% calculando para o ângulo 45
n=0;
total= 0;
for i= 1:lin
    for j= 1:col
        mlin=i+lag;
        mcol=j-lag;
        if((mlin<=lin)&&(mlin>0)&&(mcol<=col)&&(mcol>0))
            if ((dados1(i,j)~=0) && (dados1(mlin,mcol)~=0) &&
(dados2(i,j)~=0) && (dados2(mlin,mcol)~=0))
                total=total+(dados1(i,j)-
dados1(mlin,mcol))*(dados2(i,j)-dados2(mlin,mcol));
                n=n+1;
            end
        end
    end
end
y_45= total/(2*n);

clear n total
% calculando para o ângulo 135
n=0;
total= 0;
for i= 1:lin
    for j= 1:col
        mlin=i+lag;
        mcol=j+lag;
        if((mlin<=lin)&&(mlin>0)&&(mcol<=col)&&(mcol>0))
            if ((dados1(i,j)~=0) && (dados1(mlin,mcol)~=0) &&
(dados2(i,j)~=0) && (dados2(mlin,mcol)~=0))
                total=total+(dados1(i,j)-
dados1(mlin,mcol))*(dados2(i,j)-dados2(mlin,mcol));
                n=n+1;
            end
        end
    end
end
y_135= total/(2*n);

clear n total

end

```

Apêndice G

Código-fonte do programa de computador implementado no MATLAB para obtenção variograma cruzado usando “Greenness Method”

```
% Algoritmo para obtenção das semivariâncias
% Autor DANILO PEREIRA BARBOSA
% marco de 2008

clear
clc

lag= 1 % distância para o calculo variograma cruzado

for n=1:600

    % lendo as amostras
    A=imread(strcat('D:\execusão da
dissertação\bloco',num2str(n),'.tif'));

    [lin,col]=size(A(:,:,1));
    dados= zeros(size(A(:,:,1)));
    % faz o calculo do excesso de verde somente para pixels que contém
    % valores maiores na banda G do que nas outras duas bandas R e B.
    Para
    % os demais pixels atribui-se valor zero (0).
    for i= 1:lin
        for j= 1:col
            if ((A(i,j,2)>A(i,j,1)) && (A(i,j,2)>A(i,j,3)))
                dados(i,j,1)=A(i,j,1);
                dados(i,j,2)=A(i,j,2);
                dados(i,j,3)=A(i,j,3);
            else
                dados(i,j,:)= 0;
            end
        end
    end

    % salva o variograma cruzado com seus respectivos ângulos (y_0 em
y(:,1),
    % y_45 em y(:,2), etc...)
    [y(n,2), y(n,3), y(n,4), y(n,5)] = cross_variogram(dados,lag)

end

nome_salva=(strcat('matriz_cross_variogram_',num2str(lag),'.xls'));
xlswrite(nome_salva,y) ;
```

Apêndice H

Código-fonte do programa de computador implementado no MATLAB para o cálculo do pseudo variograma nos ângulos 0, 45, 90 e 135°

```
% algoritmo para calcular o pseudo variograma cruzado usando
"greenness method"
% DANILO PEREIRA BARBOSA

function [y_0, y_45, y_90, y_135] = pseudo_cross_variogram(dados, p)

% combinação BxG
dados1= im2double(dados(:,:,3));%dados1 é a banda B de dados
dados2= im2double(dados(:,:,2));%dados2 é a banda G de dados

% combinação RxG
% dados1= im2double(dados(:,:,2));%dados1 é a banda G de dados
% dados2= im2double(dados(:,:,1));%dados2 é a banda R de dados

[lin,col]=size(im2double(dados1));
for lag=p

    n=0;
    total=0;
% calculando para o ângulo 0
    for i=1:lin
        for j=1:(col-lag)
            if ((dados1(i,j)~=0) && (dados2(i,j+lag)~=0))
                total= total + (dados1(i,j)-dados2(i,j+lag))^2; %
calcula e totaliza
            end
            n= n+1;
        end
    end
    y_0=total/(2*n);
    clear n total

    n=0;
% calculando para o ângulo 90
    total=0;
    for j=1:(col)
        for i=1:(lin-lag)
            if ((dados1(i,j)~=0) && (dados2(i+lag,j)~=0))
                total= total + (dados1(i,j)-dados2(i+lag,j))^2;
                n= n+1;
            end
        end
    end
    y_90= total/(2*n);

    clear n total
```

```

% calculando para o ângulo 45
n=0;
total= 0;
for i= 1:lin
    for j= 1:col
        mlin=i+lag;
        mcol=j-lag;
        if((mlin<=lin)&&(mlin>0)&&(mcol<=col)&&(mcol>0))
            if ((dados1(i,j)~=0) && (dados2(mlin,mcol)~=0))
                total=total+(dados1(i,j)-dados2(mlin,mcol))^2;
                n=n+1;
            end
        end
    end
end
y_45= total/(2*n);

clear n total
% calculando para o ângulo 135
n=0;
total= 0;
for i= 1:lin
    for j= 1:col
        mlin=i+lag;
        mcol=j+lag;
        if((mlin<=lin)&&(mlin>0)&&(mcol<=col)&&(mcol>0))
            if ((dados1(i,j)~=0) && (dados2(mlin,mcol)~=0))
                total=total+(dados1(i,j)-dados2(mlin,mcol))^2;
                n=n+1;
            end
        end
    end
end
y_135= total/(2*n);

clear n total

end

```

Apêndice I

Código-fonte do programa de computador implementado no MATLAB para obtenção pseudo variograma cruzado usando “Greenness Method”

```
% Algoritmo para obtenção das semivariâncias
% Autor DANILO PEREIRA BARBOSA
% marco de 2008
clear
clc

lag=1 % distância para o calculo do pseudo variograma cruzado

for n=1:600

    % lendo as amostras
    A=imread(strcat('D:\execução da
dissertação\bloco',num2str(n),'.tif'));

    [lin,col]=size(A(:,:,1));
    dados= zeros(size(A(:,:,1)));
    % faz o calculo do excesso de verde somente para pixels que contém
    % valres maiores na banda G do que nas outras duas bandas R e B.
Para
    % os demais pixels atribui-se valor zero (0).
    for i= 1:lin
        for j= 1:col
            if ((A(i,j,2)>A(i,j,1)) && (A(i,j,2)>A(i,j,3)))
                dados(i,j,1)=A(i,j,1);
                dados(i,j,2)=A(i,j,2);
                dados(i,j,3)=A(i,j,3);
            else
                dados(i,j,:)= 0;
            end
        end
    end
end

    % salva o pseudo variograma cruzado com seus respectivos ângulos
    (y_0° em y(:,1), y_45° em y(:,2), etc...)

    [y(n,2),          y(n,3),          y(n,4),          y(n,5)]      =
pseudo_cross_variogram(dados,lag)

end

nome_salva=(strcat('matriz_pseudo_cross_variogram',num2str(lag),'.xls'
));
xlswrite(nome_salva,y) ;
```

Apêndice J

Código-fonte do programa de computador implementado no MATLAB para obtenção dos descritores de Haralick

```
% Programa para calcular a matriz de co-ocorrendia da matriz I
% OBS. So serve para distancia entre pixels igual a 1. Dessa maneira a
% matriz de co-ocorencia e simétrica. O algoritmo gera os parâmetros
texturais nas direções 0° 45° 90° e 135°. No desenvolvimento da
pesquisa utilizou-se somente 0°
% ROBERTO CARLOS ORLANDO & HENRIQUE ANASTÁCIO ALVES
% marco de 2003

clear
for n=1:600
    n
    % lendo as amostras
    A=imread(strcat('D:\execusão da
dissertação\classe_4_blocos_t\bloco',num2str(n),'.tif'));

    % fazendo o excesso de verde
    Im=(double(A(:,:,2)).*2-double(A(:,:,1))-
double(A(:,:,3)))./(double(A(:,:,1))+double(A(:,:,2))+double(A(:,:,3))
);
    Ii=(Im+1).*255./1;
    I=fix(Ii);
    clear Ii A;

    ngl=256; % numero de niveis de verde
    n_linhas = size(I,1); % numero de linhas
    n_colunas = size(I,2); %numero de colunas
    P=zeros(ngl); % zera a matriz de co-ocorrendia
    P_45=zeros(ngl);
    P_90=zeros(ngl);
    P_135=zeros(ngl);

    % Matriz P para teta igual a zero
    for i=1:n_linhas
        for j=1:n_colunas
            if (j==1)
                P(I(i,1)+1,I(i,2)+1)=P(I(i,1)+1,I(i,2)+1)+1;
            else if (j==n_colunas)
                P(I(i,n_colunas)+1,I(i,n_colunas-
1)+1)=P(I(i,n_colunas)+1,I(i,n_colunas-1)+1)+1;
            else
                P(I(i,j)+1,I(i,j+1)+1)=P(I(i,j)+1,I(i,j+1)+1)+1;
                P(I(i,j)+1,I(i,j-1)+1)=P(I(i,j)+1,I(i,j-1)+1)+1;
            end
        end
    end
end

%Matriz P para teta igual a 45
for i=1:n_linhas
    for j=1:n_colunas
```

```

        if (i==1)
            if (j~=1)
                P_45(I(i,j)+1,I(i+1,j-1)+1)=P_45(I(i,j)+1,I(i+1,j-1)+1)+1;
            end
        else if (i==n_linhas)
            if (j~=n_colunas)
                P_45(I(i,j)+1,I(i-1,j+1)+1)=P_45(I(i,j)+1,I(i-
1,j+1)+1)+1;
            end
            else if (j==1)
                P_45(I(i,j)+1,I(i-1,j+1)+1)=P_45(I(i,j)+1,I(i-
1,j+1)+1)+1;
            else if (j==n_colunas)
                P_45(I(i,j)+1,I(i+1,j-1)+1)=P_45(I(i,j)+1,I(i+1,j-
1)+1)+1;
            else
                P_45(I(i,j)+1,I(i+1,j-1)+1)=P_45(I(i,j)+1,I(i+1,j-
1)+1)+1;
                P_45(I(i,j)+1,I(i-1,j+1)+1)=P_45(I(i,j)+1,I(i-
1,j+1)+1)+1;
            end
        end
    end
end
end
end

```

```

% matriz P para teta igual a 90

```

```

for i=1:n_linhas
    for j=1:n_colunas
        if (i==1)
            P_90(I(1,j)+1,I(2,j)+1)=P_90(I(1,j)+1,I(2,j)+1)+1;
        else if (i==n_linhas)
            P_90(I(n_linhas,j)+1,I(n_linhas-
1,j)+1)=P_90(I(n_linhas,j)+1,I(n_linhas-1,j)+1)+1;
        else
            P_90(I(i,j)+1,I(i+1,j)+1)=P_90(I(i,j)+1,I(i+1,j)+1)+1;
            P_90(I(i,j)+1,I(i-1,j)+1)=P_90(I(i,j)+1,I(i-1,j)+1)+1;
        end
    end
end
end
end

```

```

% matriz P para teta igual a 135

```

```

for i=1:n_linhas
    for j=1:n_colunas
        if (i==1)
            if (j~=n_colunas);
                P_135(I(i,j)+1,I(i+1,j+1)+1)=P_135(I(i,j)+1,I(i+1,j+1)+1)+1;
            end
            else if (i==n_linhas);
                if (j~=1);
                    P_135(I(i,j)+1,I(i-1,j-1)+1)=P_135(I(i,j)+1,I(i-1,j-
1)+1)+1;
                end
            else if (j==1);
                P_135(I(i,j)+1,I(i+1,j+1)+1)=P_135(I(i,j)+1,I(i+1,j+1)+1)+1;
            else if (j==n_colunas);

```

```

        P_135(I(i,j)+1,I(i-1,j-1)+1)=P_135(I(i,j)+1,I(i-1,j-
1)+1)+1;
        else
P_135(I(i,j)+1,I(i+1,j+1)+1)=P_135(I(i,j)+1,I(i+1,j+1)+1)+1;
        P_135(I(i,j)+1,I(i-1,j-1)+1)=P_135(I(i,j)+1,I(i-1,j-
1)+1)+1;
        end
        end
        end
        end
        end
end

% Normalização da matriz
p = P./sum(sum(P));
p_45 = P_45./sum(sum(P_45));
p_90 = P_90./sum(sum(P_90));
p_135 = P_135./sum(sum(P_135));

% probabilidade marginal da matriz
px=sum(p);
px_45=sum(p_45);
px_90=sum(p_90);
px_135=sum(p_135);

% momento angular

f1=sum(sum(p.^2));
f1_45=sum(sum(p_45.^2));
f1_90=sum(sum(p_90.^2));
f1_135=sum(sum(p_135.^2));

% valor medio

f2=0;
f2_45=0;
f2_90=0;
f2_135=0;
for i=0:ngl-1
    f2=(i*px(i+1))+f2;
    f2_45=(i*px_45(i+1))+f2_45;
    f2_90=(i*px_90(i+1))+f2_90;
    f2_135=(i*px_135(i+1))+f2_135;
end

% variância
f3=0;
f3_45=0;
f3_90=0;
f3_135=0;
for i=0:ngl-1
    f3=((i-f2)^2)*px(i+1)+f3;
    f3_45=((i-f2_45)^2)*px_45(i+1)+f3_45;
    f3_90=((i-f2_90)^2)*px_90(i+1)+f3_90;
    f3_135=((i-f2_135)^2)*px_135(i+1)+f3_135;
end

% correlação
f4=0;

```

```

f4_45=0;
f4_90=0;
f4_135=0;

for i=0:ngl-1
    for j=0:ngl-1
        f41=i*j*p(i+1,j+1)+f4;
        f41_45=i*j*p_45(i+1,j+1)+f4_45;
        f41_90=i*j*p_90(i+1,j+1)+f4_90;
        f41_135=i*j*p_135(i+1,j+1)+f4_135;
    end
end
if f3==0
    f3=0.00001;
end
if f3_45==0
    f3_45=0.00001;
end
if f3_90==0
    f3_90=0.00001;
end
if f3_135==0
    f3_135=0.00001;
end
f4=(f41-f2^2)/f3;
f4_45=(f41_45-f2_45^2)/f3_45;
f4_90=(f41_90-f2_90^2)/f3_90;
f4_135=(f41_135-f2_135^2)/f3_135;

% momento do produto
f5=0;
f5_45=0;
f5_90=0;
f5_135=0;
for i=0:ngl-1
    for j=0:ngl-1
        f5=((i-f2)*(j-f2))*p(i+1,j+1)+f5;
        f5_45=((i-f2_45)*(j-f2_45))*p_45(i+1,j+1)+f5_45;
        f5_90=((i-f2_90)*(j-f2_90))*p_90(i+1,j+1)+f5_90;
        f5_135=((i-f2_135)*(j-f2_135))*p_135(i+1,j+1)+f5_135;
    end
end

% Diferença inversa do momento
f6=0;
f6_45=0;
f6_90=0;
f6_135=0;
for i=0:ngl-1
    for j=0:ngl-1

        f6=(p(i+1,j+1)/(1+(i-j)^2))+f6;
        f6_45=(p_45(i+1,j+1)/(1+(i-j)^2))+f6_45;
        f6_90=(p_90(i+1,j+1)/(1+(i-j)^2))+f6_90;
        f6_135=(p_135(i+1,j+1)/(1+(i-j)^2))+f6_135;
    end
end

% entropia
f7=0;

```

```

f7_45=0;
f7_90=0;
f7_135=0;
for i=0:ngl-1
    for j=0:ngl-1
        if p(i+1,j+1)==0
            p(i+1,j+1)=0.000001;
        end
        f7=p(i+1,j+1)*log(p(i+1,j+1))+f7;

        if p_45(i+1,j+1)==0
            p_45(i+1,j+1)=0.000001;
        end
        f7_45=p_45(i+1,j+1)*log(p_45(i+1,j+1))+f7_45;

        if p_90(i+1,j+1)==0
            p_90(i+1,j+1)=0.000001;
        end
        f7_90=p_90(i+1,j+1)*log(p_90(i+1,j+1))+f7_90;

        if p_135(i+1,j+1)==0
            p_135(i+1,j+1)=0.000001;
        end
        f7_135=p_135(i+1,j+1)*log(p_135(i+1,j+1))+f7_135;
    end
end
end

```

```

% medidas de correlação
hx=0;
hx_45=0;
hx_90=0;
hx_135=0;
for i=1:ngl
    if px(i)==0
        px(i)=0.00001;
    end
    hx= -px(i)*log(px(i))+hx;

    if px_45(i)==0
        px_45(i)=0.00001;
    end
    hx_45= -px_45(i)*log(px_45(i))+hx_45;

    if px_90(i)==0
        px_90(i)=0.00001;
    end
    hx_90= -px_90(i)*log(px_90(i))+hx_90;

    if px_135(i)==0
        px_135(i)=0.00001;
    end
    hx_135= -px_135(i)*log(px_135(i))+hx_135;
end
end

```

```

hxy1=0;
hxy1_45=0;
hxy1_90=0;
hxy1_135=0;
for i=0:ngl-1

```

```

for j=0:ngl-1
    pcx1=px(i+1)*px(j+1);
    pcx1_45=(px_45(i+1)*px_45(j+1));
    pcx1_90=(px_90(i+1)*px_90(j+1));
    pcx1_135=(px_135(i+1)*px_135(j+1));
    if pcx1==0
        pcx1=0.0001;
    end
    hxy1=-(p(i+1,j+1)*log(pcx1))+hxy1;

    if pcx1_45==0
        pcx1_45=0.00001;
    end
    hxy1_45=-(p_45(i+1,j+1)*log(pcx1_45))+hxy1_45;

    if pcx1_90==0
        pcx1_90=0.00001;
    end
    hxy1_90=-(p_90(i+1,j+1)*log(pcx1_90))+hxy1_90;

    if pcx1_135==0
        pcx1_135=0.00001;
    end
    hxy1_135=-(p_135(i+1,j+1)*log(pcx1_135))+hxy1_135;
end
end

hxy2=0;
hxy2_45=0;
hxy2_90=0;
hxy2_135=0;
for i=0:ngl-1
    for j=0:ngl-1
        pcx2=px(i+1)*px(j+1);
        pcx2_45=(px_45(i+1)*px_45(j+1));
        pcx2_90=(px_90(i+1)*px_90(j+1));
        pcx2_135=(px_135(i+1)*px_135(j+1));

        if pcx2==0
            pcx2=0.00001;
        end
        hxy2= -(px(i+1)*px(j+1))*log(pcx2)+hxy2;
        if pcx2_45==0
            pcx2_45=0.00001;
        end
        hxy2_45= -(px_45(i+1)*px_45(j+1))*log(pcx2_45)+hxy2_45;

        if pcx2_90==0
            pcx2_90=0.00001;
        end
        hxy2_90= -(px_90(i+1)*px_90(j+1))*log(pcx2_90)+hxy2_90;

        if pcx2_135==0
            pcx2_135=0.00001;
        end
        hxy2_135= -(px_135(i+1)*px_135(j+1))*log(pcx2_135)+hxy2_135;
    end
end
end

f10=0;

```

```

f10_45=0;
f10_90=0;
f10_135=0;
f10=(f7-hxy1)/hx;
f10_45=(f7_45-hxy1_45)/hx_45;
f10_90=(f7_90-hxy1_90)/hx_90;
f10_135=(f7_135-hxy1_135)/hx_135;

% salva o resultado da imagem
salva=[f1 f1_45 f1_90 f1_135 f2 f2_45 f2_90 f2_135 f3 f3_45 f3_90
f3_135 f4 f4_45 f4_90 f4_135 f5 f5_45 f5_90 f5_135 f6 f6_45 f6_90
f6_135 f7 f7_45 f7_90 f7_135 f10 f10_45 f10_90 f10_135];
fed=fopen('teste_256.txt','a');
fprintf(fed,'%10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t
%10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t
%10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t
%10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\t
%10.4f\t %10.4f\t %10.4f\t %10.4f\t %10.4f\n',salva);
fclose(fed);
end

```

Apêndice K

Código-fonte do programa de computador implementado no MATLAB para classificação e avaliação do classificador

```
% algoritmo para classificação e avaliação de classificador
estatístico
% a avaliação aqui é feita pelo índice kappa
% DARLY GERALDO SENNA JUNIOR
% 4 Classes

clear
clc
% carrega os dados
[nomearq,caminho] = uigetfile('*.xls','Selecione dados de origem');
dados=xlsread(strcat(caminho,nomearq));
parc=dados(:,1);
mat=dados(:,2:size(dados,2));

indices=size(mat,2);
v=[1:indices];
% inicializa a matriz resultados do kappa
res_kappa=zeros(1,size(v,2)+2);
% inicializa o contador de combinações
tot_comb=0;

for n=1
    comb=combnk(v,n); % quando o n da linha anterior valer 1, serão
    criadas as combinações tomadas 1 a 1... Se n=2, combinações de 2 a
    2...
    disp(n);

    % monta a matriz com a combinação da vez a partir das combinações

    for i=1:size(comb,1)
        comb_vez=comb(i,:);
        tot_comb=tot_comb+1;
        for k=1:size(comb_vez,2)
            p=comb_vez(k);
            mat_vez(:,k)=mat(:,p);
        end
        mat_erro=zeros(4,4);

        for nc=1:size(mat_vez,1)
            conta_1=1;
            conta_2=1;
            conta_3=1;
            conta_4=1;
            for c=1:size(mat_vez,1)
                if c ~= nc
                    if parc(c,1)==1
                        mat_1(conta_1,:)=mat_vez(c,:);
                        conta_1=conta_1+1;
                    elseif parc(c,1)==2
                        mat_2(conta_2,:)=mat_vez(c,:);
```

```

        conta_2=conta_2+1
    elseif parc(c,1)==3
        mat_3(conta_3,:)=mat_vez(c,:)
        conta_3=conta_3+1
    elseif parc(c,1)==4
        mat_4(conta_4,:)=mat_vez(c,:)
        conta_4=conta_4+1
    end
end
end
% calcula media e matriz de covariancia
med_1=mean(mat_1);
med_2=mean(mat_2);
med_3=mean(mat_3);
med_4=mean(mat_4);

cov_1=cov(mat_1);
cov_2=cov(mat_2);
cov_3=cov(mat_3);
cov_4=cov(mat_4);

Sc=(conta_1-2)*cov_1 + (conta_2-2)*cov_2 + (conta_3-
2)*cov_3 + (conta_4-2)*cov_4 ;
Sc=Sc/((conta_1-2) + (conta_2-2) + (conta_3-2) + (conta_4-
2));

d1=-0.5*(mat_vez(nc,:)'-med_1')'*inv(Sc)*(mat_vez(nc,:)'-
med_1')-0.5*log(det(Sc));
d2=-0.5*(mat_vez(nc,:)'-med_2')'*inv(Sc)*(mat_vez(nc,:)'-
med_2')-0.5*log(det(Sc));
d3=-0.5*(mat_vez(nc,:)'-med_3')'*inv(Sc)*(mat_vez(nc,:)'-
med_3')-0.5*log(det(Sc));
d4=-0.5*(mat_vez(nc,:)'-med_4')'*inv(Sc)*(mat_vez(nc,:)'-
med_4')-0.5*log(det(Sc));
%
% "escolhe" o maior "d"
[dmax,classe]=max([d1;d2;d3;d4])
% monta a matriz de erros
if parc(nc,1)==1 & classe==1
    mat_erro(1,1)=mat_erro(1,1)+1
elseif parc(nc,1)==1 & classe==2
    mat_erro(1,2)=mat_erro(1,2)+1
elseif parc(nc,1)==1 & classe==3
    mat_erro(1,3)=mat_erro(1,3)+1
elseif parc(nc,1)==1 & classe==4
    mat_erro(1,4)=mat_erro(1,4)+1
elseif parc(nc,1)==2 & classe==1
    mat_erro(2,1)=mat_erro(2,1)+1
elseif parc(nc,1)==2 & classe==2
    mat_erro(2,2)=mat_erro(2,2)+1
elseif parc(nc,1)==2 & classe==3
    mat_erro(2,3)=mat_erro(2,3)+1
elseif parc(nc,1)==2 & classe==4
    mat_erro(2,4)=mat_erro(2,4)+1
elseif parc(nc,1)==3 & classe==1
    mat_erro(3,1)=mat_erro(3,1)+1
elseif parc(nc,1)==3 & classe==2
    mat_erro(3,2)=mat_erro(3,2)+1
elseif parc(nc,1)==3 & classe==3
    mat_erro(3,3)=mat_erro(3,3)+1

```

```

elseif parc(nc,1)==3 & classe==4
    mat_erro(3,4)=mat_erro(3,4)+1
elseif parc(nc,1)==4 & classe==1
    mat_erro(4,1)=mat_erro(4,1)+1
elseif parc(nc,1)==4 & classe==2
    mat_erro(4,2)=mat_erro(4,2)+1
elseif parc(nc,1)==4 & classe==3
    mat_erro(4,3)=mat_erro(4,3)+1
elseif parc(nc,1)==4 & classe==4
    mat_erro(4,4)=mat_erro(4,4)+1

end
% limpa as matrizes da validacao cruzada
clear mat_1 mat_2 mat_3 mat_4
end
[kappa,var_kappa]=kappa_certo(mat_erro)
% armazena kappa e var_kappa
res_kappa(tot_comb,1:size(comb_vez,2))=comb_vez
res_kappa(tot_comb,10)=trace(mat_erro)/sum(sum(mat_erro))
res_kappa(tot_comb,11)=kappa
res_kappa(tot_comb,12)=var_kappa
res_kappa(tot_comb,13:16)=mat_erro(1,1:4)
res_kappa(tot_comb,18:21)=mat_erro(2,1:4)
res_kappa(tot_comb,23:26)=mat_erro(3,1:4)
res_kappa(tot_comb,28:31)=mat_erro(4,1:4)

clear mat_vez
end
end
nome_salva=(strcat('resultado',nomearq))
dlmwrite(nome_salva,res_kappa,'\t')
dlmwrite(nome_salva,res_kappa,'\t')

```

Apêndice L

Código-fonte do programa de computador implementado no MATLAB para o cálculo do índice Kappa

```
% algoritmo para calcular indice kappa
% Gislaine Pacheco Tormen

function [khat,var_k] = kappa_certo(me)
% exatidao global
eg=trace(me)/sum(me(:));% teta 1
%
exp_a=sum(me,1)*sum(me,2)/sum(sum(sum(me,2)*sum(me,1)));%teta 2
khat=(eg-exp_a)/(1-exp_a);%k= (tetal - teta2)/ (1 - teta2)
%
n=sum(sum(me));
%
mp=sum(me,2)*sum(me,1);
tetal=eg;
teta2=exp_a;
% teta 3= somatorio(me(ii)*(linha_marg(i)+col_marg(i))/n*n
somat3=0;
mg_li=sum(me,2)';
mg_co=sum(me,1);
soma_mg=mg_li+mg_co;
for i=1:size(me)
    somat3=somat3+me(i,i)*(soma_mg(1,i));
end
teta3=somat3/n^2;
% teta4=somatorio(me(i,j)*(linha_marg(i)+col_marg(i))^2)/n^3
somat4=0;
for j=1:size(me,1)
    for i=1:size(me)
        somat4=somat4+me(i,j)*(mg_li(1,j)+mg_co(1,i))^2;
    end
end
teta4=somat4/n^3;
% calcula variancia kappa
%
parc1=tetal*(1-tetal)/(1-teta2)^2;
parc2=2*(1-tetal)*(2*tetal*teta2-teta3)/(1-teta2)^3;
parc3=(1-tetal)^2*(teta4-4*teta2^2)/(1-teta2)^4;
var_k=(parc1+parc2+parc3)/n;
```