



Paula Beatriz Cerqueira Leite

**Identificação de Tipos de Culturas Agrícolas a
partir de Seqüências de Imagens Multitemporais
Utilizando Modelos de Markov Ocultos**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre pelo Programa de Pós-
graduação em Engenharia Elétrica do Departamento de
Engenharia Elétrica da PUC-Rio.

Orientador: Prof. Raul Queiroz Feitosa
Co-orientador: Prof. Antônio Roberto Formaggio

Rio de Janeiro
Agosto de 2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



Paula Beatriz Cerqueira Leite

**Identificação de Tipos de Culturas Agrícolas a
partir de Seqüências de Imagens Multitemporais
Utilizando Modelos de Markov Ocultos**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Dr. Raul Queiroz Feitosa
Orientador

Departamento de Engenharia Elétrica - PUC-Rio

Dr. Antônio Roberto Formaggio
INPE

Dr. Álvaro de Lima Veiga Filho
Departamento de Engenharia Elétrica - PUC-Rio

Dr. Cláudia Maria de Almeida
INPE

Prof. José Eugenio Leal
Coordenador Setorial do Centro
Técnico Científico

Rio de Janeiro, 15 de agosto de 2008

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, da autora e do orientador.

Paula Beatriz Cerqueira Leite

Graduou-se em Engenharia Elétrica com ênfase em sistemas de apoio à decisão na PUC-Rio (Pontifícia Universidade Católica do Rio de Janeiro) em 2006.

Ficha Catalográfica

Leite, Paula Beatriz Cerqueira

Identificação de tipos de culturas agrícolas a partir de seqüências de imagens multitemporais utilizando modelos de Markov ocultos / Paula Beatriz Cerqueira Leite ; orientador: Raul Queiroz Feitosa ; co-orientador: Antônio Roberto Formaggio. – 2008.

79 f. ; 30 cm

Dissertação (Mestrado em Engenharia Elétrica)– Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2008.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Identificação de culturas agrícolas. 3. Modelos de Markov ocultos. 4. Análise multitemporal. 5. Sensoriamento remoto. 6. Landsat 5 TM. 7. Landsat 7 ETM. I. Feitosa, Raul Queiroz. II. Formaggio, Antônio Roberto. III. Pontifícia Universidade Católica do Rio de Janeiro. IV. Departamento de Engenharia Elétrica. V. Título.

CDD: 621.3

Aos meus pais.

Agradecimentos

Ao meu orientador Dr. Raul Queiroz Feitosa pela instrução, força e ajuda ao longo deste trabalho.

Ao meu co-orientador Dr. Antônio Roberto Formaggio pelo auxílio constante.

Ao CNPq e à PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Aos meus pais e a minha avó Adelina pelo carinho, amor e apoio incondicional em tudo.

Ao meu avô Cerqueira pelos conselhos sábios que guardarei para sempre comigo.

Ao Leo pela compreensão, incentivo e amor em todos os momentos.

Aos meus amigos do LVC pelo companheirismo e incentivo: Dário, Gilson, Marcelo, Vanessa, Ilana, Henrique, Cecília e Edmundo.

As minhas irmãs do ballet pelos momentos de descontração.

A Renata e a Fabiola por estarem sempre comigo, prontas para me ouvir e ajudar.

A minha família por estar, de um jeito ou de outro, presente em mais esta etapa da minha vida.

Aos meus queridos amigos pelo apoio e incentivo nos momentos difíceis.

A Deus.

Resumo

Leite, Paula Beatriz Leite; Feitosa, Raul Queiroz (Orientador). **Identificação de tipos de culturas agrícolas a partir de seqüências de imagens multitemporais utilizando modelos de Markov ocultos**. Rio de Janeiro, 2008. 79p. Dissertação de Mestrado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Esta dissertação propõe uma metodologia baseada em Modelos de Markov Ocultos (*Hidden Markov Models – HMM*) para a classificação de culturas agrícolas, explorando informações de seqüências temporais de imagens dos sensores TM e ETM⁺/Landsat. O método reconhece os diferentes tipos de culturas agrícolas analisando os perfis espectrais em uma seqüência temporal de imagens de satélite de média resolução espacial ($\cong 30\text{m}$). Nesta abordagem, o comportamento temporal de cada classe de cultura é modelado por um HMM específico. A classificação é feita segmento-a-segmento, descritos por um vetor de atributos calculado como as médias espectrais dos pixels contidos no segmento em cada banda da imagem. Os vetores de atributos do segmento em cada imagem da seqüência de imagens são subseqüentemente submetidos aos HMMs de cada classe de cultura. O segmento é então associado à cultura cujo HMM correspondente gera a maior probabilidade de emitir a seqüência de valores espectrais observada. Os experimentos para análise foram conduzidos utilizando-se um conjunto de 12 imagens LANDSAT co-registradas e corrigidas radiometricamente. As imagens cobrem uma área do estado de São Paulo, Brasil, com aproximadamente 124.100ha, entre 2002 e 2004. As seguintes coberturas vegetais foram consideradas: cana de açúcar, soja, milho, pastagem e mata-galeria. A avaliação do desempenho do método foi efetuada utilizando-se um conjunto de dados classificado visualmente por dois especialistas e validado por um extenso trabalho de campo. O desempenho do método de classificação multitemporal proposto foi comparado com o de um classificador monotemporal de máxima verossimilhança, e os resultados mostraram a superioridade notável do método baseado em HMM, o qual alcançou uma acurácia média de nada menos que 91% na identificação do tipo correto de cultura agrícola, para seqüências de dados contendo apenas uma única classe de cultura.

Palavras-chaves

Identificação de culturas agrícolas, modelos de Markov ocultos, análise multitemporal, sensoriamento remoto, Landsat 5 TM, Landsat 7 ETM⁺.

Abstract

Leite, Paula Beatriz Leite; Feitosa, Raul Queiroz (Advisor). **Crop type identification based on Hidden Markov Models using multitemporal image sequences**. Rio de Janeiro, 2008. 79p. MSc Dissertation – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This work proposes a Hidden Markov Model (HMM)-based methodology to classify agricultural crops, exploring information of temporal image sequences from TM and ETM⁺/Landsat sensors. HMMs are used to relate the varying spectral response along the crop cycle with plant phenology for different crop classes. The method recognizes different agricultural crops by analyzing their spectral profiles over a temporal sequence of medium resolution satellite images ($\cong 30\text{m}$). In our approach the temporal behaviour of each crop class is modelled by a specific HMM. A segment-based classification is performed using the average spectral values of the pixels in each image segment across an image sequence, which is subsequently submitted to the HMMs of each crop class. The image segment is assigned to the crop class, whose corresponding HMM delivers the highest probability of emitting the observed sequence of spectral values. Experiments were conducted upon a set of 12 co-registered and radiometrically corrected LANDSAT images. The images cover an area of the State of São Paulo, Brazil with about 124.100ha, between the years 2002 and 2004. The following classes were considered: sugarcane, soybean, corn, pasture and riparian forest. Performance assessment was carried out upon a data set classified visually by two analysts and validated by extensive field work. The performance of the proposed multitemporal classification method was compared to that of a monotemporal maximum likelihood classifier, and the results indicated a remarkable superiority of the HMM-based method, which achieved an average of no less than 91% accuracy in the identification of the correct crop, for sequences of data containing a single crop class.

Keywords

Crop class identification, hidden Markov models, multitemporal analysis, remote sensing, Landsat 5 TM, Landsat 7 ETM⁺.

Sumário

1 Introdução	13
1.1. Motivação	13
1.2. Objetivos	15
1.2.1. Objetivo Geral	15
1.2.2. Objetivos Específicos	16
1.3. Organização do restante da dissertação	16
2 Trabalhos anteriores	17
2.1. Interpretação multitemporal de imagens de sensoriamento remoto	17
2.2. Estimativas agrícolas a partir de imagens de sensoriamento remoto	19
2.3. Fenologia	20
2.4. Aplicações de Modelos de Markov Ocultos	21
3 Modelos de Markov Ocultos	23
3.1. Processos Estocásticos	23
3.2. Processos Markovianos	23
3.3. Cadeias de Markov	24
3.4. Modelos de Markov Ocultos	25
3.4.1. Elementos de um Modelo de Markov Oculto	26
3.4.2. Problemas básicos associados a HMMs	27
3.4.3. Soluções para os problemas 1 e 2 associados a HMMs	28
3.4.3.1. Solução do problema 1	28
3.4.3.2. Solução do problema 2	31
4 Método Proposto	34
4.1. Descrição geral do modelo	34
4.2. Ajustando o modelo à aplicação – Estimativa de parâmetros	36
4.2.1. Estimativa da matriz de transição de estado	38
4.2.2. Estimativa das probabilidades a priori de estados	39
4.2.3. Estimativa da emissão de símbolo	39
4.3. Classificação	41
5 Análise experimental	42
5.1. Culturas e respectivos ciclos fenológicos	42
5.1.1. Cana-de-açúcar	42

5.1.2. Culturas de ciclo curto (cereais)	43
5.1.3. Pastagem	46
5.1.4. Outras coberturas vegetais	46
5.2. Área geográfica	46
5.3. Imagens	48
5.4. Pré-processamento das imagens	48
5.5. Dados de referência	49
5.5.1. Segmentação	49
5.5.2. Processo de classificação visual	49
5.6. Experimentos	54
5.6.1. Experimento 1: Conjunto ótimo de atributos espectrais	54
5.6.2. Experimento 2: Desempenho do método para identificar as culturas agrícolas	55
5.6.3. Experimento 3: Desempenho do método para identificar os estágios fenológicos	56
5.6.4. Experimento 4: Avaliação da influência do comprimento das seqüências sendo utilizadas no desempenho do método	58
5.6.5. Experimento 5: Determinação dos estágios fenológicos que melhor caracterizam cada cultura	59
5.6.6. Experimento 6: Comparação do método multitemporal baseado em HMMs com uma abordagem monotemporal de máxima verossimilhança	62
6 Conclusão	64
7 Referências Bibliográficas	66
Apêndice I	70
Apêndice II	71
Apêndice III	78

Lista de figuras

Figura 1 - Cadeia de Markov com 3 estados (S_1 a S_3), onde as setas indicam as transições possíveis entre os estados.	24
Figura 2 - Exemplo de um modelo de Markov Oculto ($S_i \rightarrow$ estados, $v_k \rightarrow$ símbolos observáveis, $a_{ij} \rightarrow$ probabilidade de transição de estado, $b_{jk} \rightarrow$ probabilidade de emissão de símbolo).	27
Figura 3 - Ilustração das seqüências de operações necessárias para o cálculo da variável direta $\alpha_{t+1}(j)$.	29
Figura 4 - Procedimento de cálculo de $\alpha_{t+1}(j)$ em função dos instantes das observações t e dos estados S_i	30
Figura 5 - Ilustração das seqüências de operações necessárias para o cálculo da variável reversa $\beta_{t+1}(j)$.	31
Figura 6 - HMM usado neste trabalho para cana de açúcar, soja e milho (PP = pré-plantio, CR = crescimento, AD = fase adulta e PA = palhada).	34
Figura 7 - HMM usado neste trabalho para pastagem e mata-galeria (AD = fase adulta).	35
Figura 8 - Exemplo de áreas de pastagem e mata-galeria nas imagens de satélite utilizadas.	35
Figura 9 - Exemplo de áreas de soja, milho e cana de açúcar nas imagens de satélite utilizadas.	36
Figura 10 - Processo de estimativa dos valores dos parâmetros do HMM de cada cultura.	38
Figura 11 - Processo de classificação.	41
Figura 12 - Ciclos fenológicos da cana de açúcar de 12 meses (cana-de-ano) e 18 meses (cana-de-ano-e-meio). Figura obtida em (Gleriani, 2004).	43
Figura 13 - Ciclo fenológico da cultura de milho. Figura adaptada de (Gleriani, 2004).	44
Figura 14 - Ciclo fenológico da cultura de soja. Figura adaptada de (Gleriani, 2004).	45
Figura 15 - Área de estudo, estado de São Paulo, Brasil. Os pontos amarelos correspondem a 316 pontos amostrais de referência.	47

Figura 16 - Recorte de imagem Landsat na composição colorida RGB 453, e foto digital de campo de talhões de milho em diferentes estágios de crescimento. FONTE: (Sanches, 2004). 51

Figura 17 - Recorte de imagem Landsat na composição colorida RGB 453, e foto digital de campo de talhões de cana-de-açúcar em diferentes estágios de crescimento. FONTE: (Sanches, 2004). 52

Figura 18 - Recorte de imagem Landsat na composição colorida RGB 453, e foto digital de campo de talhões de soja em diferentes estágios de crescimento. FONTE: (Sanches, 2004). 53

Lista de tabelas

Tabela 1 - Imagens disponíveis no conjunto de dados utilizado no presente trabalho.	47
Tabela 2 - Conjunto ótimo de n atributos espectrais e o desempenho da classificação correspondente.	55
Tabela 3 - Taxas de acerto para a classificação de coberturas vegetais.	55
Tabela 4 - Matriz de confusão para a classificação de coberturas vegetais.	56
Tabela 5 - Taxas de acerto para a classificação de estados.	57
Tabela 6 - Matriz de confusão para a classificação de estados.	57
Tabela 7 - Taxas de acerto para seqüências de tamanhos variados.	58
Tabela 8 - Taxas de acerto para seqüências de soja com certos estados (indicados por ✓) obrigatoriamente presentes.	60
Tabela 9 - Taxas de acerto para seqüências de milho com certos estados (indicados por ✓) obrigatoriamente presentes.	61
Tabela 10 - Taxas de acerto para seqüências de cana de açúcar com certos estados (indicados por ✓) obrigatoriamente presentes.	62
Tabela 11 - Desempenho da abordagem monotemporal.	63

1

Introdução

1.1. Motivação

A agricultura historicamente tem grande importância socioeconômica para as populações.

Atualmente, agravada pela redução dos estoques mundiais, principalmente de cereais, o mercado internacional vem sofrendo significativo impacto, o que tem gerado aumento nos preços dos alimentos, alarmando dirigentes e líderes. Segundo a Organização para Agricultura e Alimentação das Nações Unidas (FAO, sigla em inglês), o preço geral dos alimentos aumentou cerca de 75% desde o início deste século e continuará subindo até 2017.

Além disso, diante da atual crise dos Estados Unidos, provocando a queda do dólar, vem ocorrendo especulação em torno do mercado de *commodities* como, o café, a soja, o trigo e outros produtos básicos.

A valorização dessas *commodities* tem efeitos distintos para os diversos atores. Obviamente, os países exportadores são os que se beneficiam com esta situação, conseqüentemente os seus produtores.

O Brasil, com um clima diversificado, chuvas regulares, energia solar abundante, alta disponibilidade de água doce, possui cerca de 388 milhões de hectares de terras agricultáveis, dos quais apenas 298 milhões são utilizados (MAPA, 2008), o que o torna um país ideal para atividades agropecuárias.

O agronegócio é atualmente responsável por um em cada três reais gerados. Assim, considerando-se a participação desta atividade no PIB nacional, fica ainda mais evidenciada a sua importância. A agropecuária representa aproximadamente 12% de seu total, considerando-se apenas o valor de produção. Se todo o complexo agroindustrial for considerado, esta parcela sobe para 35%. Os produtos agrícolas são ainda responsáveis por mais de 25% do total exportado pelo país (MAPA, 2008), sendo que o Brasil é um dos líderes mundiais na produção e

exportação de vários produtos agropecuários como café, açúcar, álcool, sucos de frutas e soja.

O Brasil tem se beneficiado com a alta dos preços. Já se vislumbra inclusive que, em função da forte demanda, a próxima safra bata um recorde histórico, com colheita prevista de cento e quarenta milhões de toneladas, o que representa aumento de quase 7,0% em relação à safra anterior (ICTSD, 2008). O milho é o principal responsável por este aumento, em função do desvio da produção americana para o etanol. A soja, no entanto, continua sendo o destaque da produção e exportação. Há uma expectativa de serem colhidas sessenta milhões de toneladas no ano.

A produção brasileira é quase 30% maior do que o consumo interno, ressaltando a vocação exportadora do país. Diante deste dado, é incontestável que o Brasil tem potencial suficiente para contribuir para a solução da crise mundial de alimentos.

Com informações exatas sobre o plantio de diferentes culturas, é possível traçar planos comerciais para regularização dos estoques internos, para a tomada de decisão sobre subsídios e para delimitar estratégias de negociação de *commodities* agrícolas em mercados financeiros. Produtores rurais, exportadores, importadores, indústria alimentícia, fornecedores de insumos, investidores e o governo são alguns dos setores interessados nessas informações.

Devido a essa crescente demanda, cada vez mais dados provenientes de sensoriamento remoto vêm sendo empregados em atividades relacionadas a estimativas agrícolas, já que estes dados são obtidos de forma objetiva e exata (Ippoliti-Ramilo, 1999). Alguns trabalhos como (Gleriani, 2004), (Sanches, 2004) e (Epiphanyo, 2007) são exemplos desta aplicação.

O uso de uma metodologia operacional que utiliza dados de sensoriamento remoto orbital para a classificação de culturas agrícolas pode apresentar problemas relacionados ao período de aquisição das imagens orbitais, já que em certos períodos do ano como, por exemplo, no verão, a pluviosidade é maior e a ocorrência de cobertura de nuvens mais frequente. Esse fato dificulta a utilização de dados provenientes de sensores ópticos (Formaggio et al., 2003).

Uma forma de contornar esse problema, aumentando a chance de obtenção de imagens limpas, livres de nuvens, em épocas como o verão, é a utilização de sensores de alta resolução temporal. Estes sensores passam por uma área

geográfica em intervalos curtos de tempo, o que possibilita a aquisição de várias imagens durante um mesmo período de safra.

Sabe-se, por outro lado, que a resolução temporal está inversamente relacionada à resolução espacial. Quando um sensor possui alta resolução temporal, em geral, possui resoluções espaciais mais baixas, podendo um pixel corresponder a centenas de metros quadrados. Para aplicações onde se deseja delimitar e classificar talhões agrícolas, sensores de baixa resolução espacial podem ter exatidão insuficiente.

Uma alternativa para isso é o uso de sensores de alta resolução temporal e resolução espacial média, como o Landsat usado neste trabalho, cuja resolução espacial é de 30m. Esses sensores, além de aumentarem a chance de obter imagens livres de nuvens em certas épocas do ano, ainda permitem que se delimitem os talhões agrícolas com boa exatidão, facilitando todo o processo de classificação.

O presente trabalho procura combinar dois campos de conhecimento que vêm tendo uma evolução significativa nos últimos anos: i) a pesquisa sobre técnicas de classificação multitemporal a partir de imagens de sensoriamento remoto e ii) sobre fenologia, que é o estudo do ciclo biológico das plantas e compreende o acompanhamento de fases como floração, frutificação, queda e brotamento foliar e a relação com o clima (Ibot, 2008). Aqui se encontra a inovação principal deste trabalho, uma vez que, existem poucos registros do uso de modelos fenológicos para auxiliar no processo de classificação das imagens (Aurdal et al., 2005).

1.2. Objetivos

1.2.1. Objetivo Geral

O objetivo da presente dissertação consiste em avaliar o potencial dos Modelos de Markov Ocultos para a classificação de culturas agrícolas a partir de seqüências temporais de imagens de sensoriamento remoto.

1.2.2. Objetivos Específicos

Desenvolver um método baseado em HMM para:

- identificar diferentes tipos de culturas agrícolas;
- identificar os estágios fenológicos em que as culturas se encontram ao longo de seus ciclos;
- avaliar o método proposto em uma sequência de imagens de satélite de média resolução;
- avaliar a contribuição do conhecimento multitemporal na acurácia da classificação automática em relação ao uso de informação monotemporal.

1.3. Organização do restante da dissertação

O restante desta dissertação está organizado da seguinte maneira: o próximo capítulo apresenta alguns trabalhos científicos produzidos anteriormente, utilizando metodologias baseadas em análise multitemporal de imagens de sensoriamento remoto, além da utilização desse tipo de análise para estimativas agrícolas, estudos sobre fenologia e algumas aplicações recentes dos Modelos de Markov Ocultos.

O Capítulo 3 explica a técnica de Modelos de Markov Ocultos, de forma a facilitar o entendimento do trabalho desenvolvido nesta dissertação.

No capítulo seguinte é apresentado o método proposto para a classificação de culturas agrícolas, baseado em Modelos de Markov Ocultos. Esse capítulo mostra os modelos escolhidos e como os parâmetros desses modelos foram estimados.

No Capítulo 5 é apresentada, de forma detalhada, a análise experimental. Depois de uma descrição da aplicação e do banco de dados disponível, os experimentos realizados são apresentados, assim como a discussão de seus resultados.

Finalmente, são apresentadas as conclusões obtidas ao longo deste trabalho e sugestões para continuidade desta pesquisa no Capítulo 6.

2

Trabalhos anteriores

Este capítulo apresenta alguns dos principais trabalhos anteriores sobre análise multitemporal de imagens de sensoriamento remoto visando estimativas de áreas agrícolas, estudos sobre fenologia e algumas aplicações recentes dos Modelos de Markov Ocultos.

2.1.

Interpretação multitemporal de imagens de sensoriamento remoto

A tecnologia de sensoriamento remoto fornece meios para o monitoramento de áreas extensas em termos de florestas, recursos hídricos, ecossistemas, áreas produtivas e espaços urbanos, entre outros. Apesar do aumento no número de novos sensores orbitais nos últimos anos e dos recentes avanços na tecnologia de software de apoio ao tratamento de imagens de sensoriamento remoto, a interpretação das imagens é ainda realizada de forma predominantemente manual, o que requer grande esforço, experiência e perícia do foto intérprete. Este conhecimento fornecido pelo especialista sobre a região analisada é de extrema importância e influencia decisivamente o resultado da interpretação sobre os objetos esperados na cena.

A maior parte dos algoritmos de processamento de imagens convencionais, disponíveis em softwares comerciais, produz resultados fragmentados, incompletos e até mesmo contendo erros, tornando necessária a execução de uma etapa posterior de correção manual. Neste processo de pós-edição, o foto intérprete usualmente utiliza informações adicionais, além das imagens, disponíveis em Sistemas de Informação Geográfica (SIG), assim como o seu conhecimento específico da região sob análise e a sua experiência. Devido à crescente necessidade de monitoramento via sensores remotos associada ao longo tempo gasto para a pós-edição visual, cresce a demanda por uma tecnologia capaz de aumentar o grau de automação da interpretação de imagens a fim de melhorar a produtividade e rapidez do processo.

Ultimamente, tem crescido em todo o mundo o interesse pelos chamados sistemas baseados em conhecimento. Estes são sistemas que procuram modelar em um ambiente computacional o conhecimento e o raciocínio do foto intérprete e emular sua capacidade de combinar dados de diferentes fontes e diferentes formatos na avaliação de imagens de sensores remotos, de modo a realizar automaticamente a tarefa do operador humano.

Os sistemas baseados em conhecimento podem utilizar inúmeras formas de conhecimento, entre elas está o multitemporal. Este tipo de conhecimento expressa o comportamento de uma classe de um objeto em instantes diferentes de tempo.

Em alguns trabalhos anteriores, os diagramas de transição de estado têm sido utilizados como ferramenta para modelar o conhecimento multitemporal. Em (Pakzad, 2001) e (Growe, 2001), as transições estão associadas a um valor que denota a probabilidade de cada transição ocorrer. Em (Campos et al., 2005) os diagramas de transição de estado são utilizados para expressar a relação entre a classe de um objeto em um dado momento e a classe do mesmo objeto em um instante anterior. Os valores de probabilidade de transição representados no diagrama são utilizados como parâmetros das funções discriminantes. Estes valores são estimados utilizando algoritmos genéticos a partir de dados de datas anteriores à da imagem analisada.

Em (Mota et al., 2007), uma abordagem possibilística é introduzida de forma a modelar transições de classes de cobertura do solo. Possibilidades de transição de classe são estimadas por meio de Algoritmos Genéticos, cuja função objetivo é a taxa de acerto médio baseada nas amostras de treinamento. A partir de uma classificação *crispy* anterior da região de interesse e os valores das possibilidades de transição de classe, uma classificação *fuzzy* é obtida e então combinada com a resposta de uma classificação *fuzzy* de uma imagem de uma data posterior.

Em (Pakzad et al, 2003), é descrito um procedimento para a interpretação multitemporal de áreas de vegetação no Brasil usando imagens landsat TM 5. Neste trabalho, o conhecimento multitemporal é introduzido no processo de classificação através de um diagrama de transição de estado que define as transições possíveis, de acordo com razões biológicas e limitações regidas por

leis. O conhecimento multitemporal reduz o espaço de busca e, conseqüentemente, os erros de classificação.

Em (Shao et al., 2002), dados multitemporais de RADARSAT, adquiridos entre 1996 e 1997, são utilizados para monitoramento e estimativa de produção de arroz na China através de métodos computacionais. Já em (Zhou et al., 2008), esse tipo de conhecimento é aplicado em uma classificação e pós-classificação (para detecção de mudanças) baseadas em objetos, utilizando imagens aéreas multitemporais de alta resolução espacial da região metropolitana de Baltimore de 1999 a 2004. Neste trabalho, é efetuada uma classificação baseada em regras para classificar objetos primitivos em uma de cinco classes de cobertura de solo.

2.2.

Estimativas agrícolas a partir de imagens de sensoriamento remoto

Com uma demanda crescente por informações exatas sobre o estado de diferentes culturas agrícolas, como foi mencionado no capítulo anterior, dados provenientes de sensores remotos têm sido cada vez mais empregados em atividades relacionadas a estimativas agrícolas.

Em (Gleriani, 2004), é investigada uma metodologia de classificação de cobertura agrícola baseada em redes neurais artificiais em dez imagens do sensor ETM+ do ano de 2002 da região de Miguelópolis (SP). As redes neurais utilizadas neste trabalho geram bons resultados de classificação, além de se mostrarem tolerantes a dados ausentes. Esses resultados são comparados com os de um algoritmo de máxima verossimilhança, o qual tem desempenho inferior.

Em (Sanches, 2004), é feita uma análise multitemporal do comportamento espectro-temporal de culturas agrícolas baseada em imagens Landsat, e a estimativa de área agrícola das principais culturas nos municípios de Guará, Ipuã e São Joaquim da Barra, SP (mesma área de estudo desta dissertação), é obtida. São feitos dois tipos de classificação, um utilizando o classificador de máxima verossimilhança, e outro utilizando o classificador Isoseg. De acordo com a taxa de acerto global, os resultados das duas classificações automáticas não são satisfatórios, sendo o resultado obtido com o Isoseg pior do que o obtido pela classificação por máxima verossimilhança.

Já em (Epiphany, 2007), o objetivo principal do trabalho é avaliar a capacidade das imagens do sensor MODIS (*MODerate resolution Imaging Spectroradiometer*) na estimativa da área plantada com a cultura da soja no estado do Mato Grosso na safra 2005/2006. A metodologia empregada é a da classificação por Superfície de Resposta Espectro-Temporal (SRET), para quantificar a área com a cultura da soja. A classificação individual de cada um dos segmentos de referência apresenta uma taxa de acerto global superior a 95%.

2.3. Fenologia

Fenologia é o ramo da Ecologia que estuda os fenômenos periódicos dos seres vivos e suas relações com as condições do ambiente, tais como temperatura, luz, umidade, etc. (Fina & Ravelo, 1973). Este ramo estuda o ciclo biológico das plantas, que compreende o acompanhamento de fases como a floração, frutificação e mudança foliar.

O estudo de fenologia é uma disciplina antiga em diversos países. Recentemente, a pesquisa neste campo se intensificou devido ao fato da fenologia estar fortemente ligada ao clima, podendo ser um indicador sensível de mudanças climáticas locais ou mesmo globais (Cleland et al., 2007). De forma a não apenas considerar aspectos locais da fenologia, mas sim em uma escala global, estudos de fenologia baseados em sensoriamento remoto têm atraído interesse de diversos pesquisadores, exemplos disso são (Reed et al., 2005), (Ahla et al., 2006) e (Törmä et al., 2007).

Existem diversos estudos na área de classificação de cobertura de solo, porém são poucos os que utilizam modelos fenológicos para ajudar no processo de classificação, ainda que o desenvolvimento desses modelos seja extensivamente mencionados na literatura, como por exemplo, em (Zhang et al., 2001).

Em (Sakamoto et al., 2005), um método é desenvolvido para a classificação de estágios fenológicos de plantações de arroz através de dados de sensoriamento remoto (MODIS) e com auxílio de técnicas como transformadas de Fourier e Wavelet. Em (Viovy & Saint, 1994), é sugerido o uso de Modelos de Markov Ocultos de fenologia para a identificação dos estágios fenológicos de vegetações nas savanas africanas.

A técnica de HMM foi escolhida para ser utilizada neste trabalho, principalmente para a identificação de culturas agrícolas no estado de São Paulo, assim como para identificar os estágios fenológicos ao longo dos ciclos dessas culturas.

2.4.

Aplicações de Modelos de Markov Ocultos

A teoria básica dos Modelos de Markov Ocultos foi publicada pela primeira vez na segunda metade da década de 1960 em (Baum & Petrie, 1966). Uma de suas primeiras aplicações foi em problemas de processamento de voz e vem sendo usada desde então como, por exemplo, em (Juang & Rabiner, 1998) e (Aarnio, 1999) para reconhecimento de voz, em (Maskey & Hirschberg, 2006) para resumo de texto falado sem qualquer tipo de registro escrito e em (Nogueiras et al. 2001) para reconhecimento de emoções através do discurso falado.

Na segunda metade dos anos 80, os HMMs começaram a ser usados na análise de seqüências biológicas, em particular o DNA e este uso se tornou cada vez mais freqüente com o passar do tempo (Birney, 2001). Alguns trabalhos merecem ser citados, como (Boys et al., 2000), onde HMMs foram aplicados na detecção de segmentos homogêneos em seqüências de DNA e (Conroy et al., 2001), onde HMMs foram usados na identificação de cromossomos.

Existem ainda outros tipos de aplicações dos Modelos de Markov Ocultos relacionados à área de processamento de imagens. Um exemplo é (Otsuka, 1997) onde é utilizado um método baseado em HMM para reconhecimento de expressões faciais de diversas pessoas e (Huang et al., 2008) onde o mesmo tipo de método é utilizado para a extração de caracteres de documentos.

Ainda na área de processamento de imagens, os HMM's vêm sendo empregados na classificação de tipos de vegetação. (Viovy & Saint, 1994) busca identificar estágios fenológicos de vegetação numa região africana. O princípio básico do método é construir um modelo do perfil temporal de radiometria baseado em um dado comportamento fenológico. Esta aplicação mostrou que com um modelo relativamente simples, informações importantes sobre a dinâmica da vegetação (início e duração do ciclo) em grande escala podem ser extraídas automaticamente.

Em (Aurdal et al., 2005), HMM's e modelos fenológicos são utilizados para a classificação de tipos de vegetação na Noruega. Assim como no trabalho descrito nesta dissertação, os autores utilizam técnicas mais comuns de classificação como máxima verossimilhança para fins de comparação. Em ambos os trabalhos estas técnicas geraram taxas de acerto menores do que as obtidas por métodos baseados em HMM.

3

Modelos de Markov Ocultos

3.1.

Processos Estocásticos

Um processo estocástico é definido como uma família de variáveis aleatórias $X(t)$, sendo t geralmente a variável tempo. $X(t)$ representa uma característica mensurável de interesse no tempo t . Exemplificando, $X(t)$ pode representar o nível de estoque de um produto no fim do dia t .

Processos estocásticos são usados para descrever o comportamento ao longo de um período de tempo de um sistema onde a incerteza é significativa. Em termos mais formais, a variável aleatória $X(t)$ representa o estado do sistema no instante t .

Os processos estocásticos podem ser classificados:

Em relação ao estado:

- Estado discreto: $X(t)$ é definido sobre um conjunto enumerável ou finito,
- Estado contínuo: caso contrário.

Em relação ao tempo:

- Tempo discreto: t é finito ou enumerável,
- Tempo contínuo: caso contrário.

3.2.

Processos Markovianos

Um processo estocástico é dito markoviano se a probabilidade do processo estar em um dado estado, em um instante qualquer no futuro, condicionada à sequência completa de estados até o presente permanece a mesma quando condicionada apenas ao estado presente.

Usando a notação q_t para designar a variável aleatória $X(t)$, as probabilidades condicionais $P(q_{t+1} = S_j | q_t = S_i)$ são denominadas probabilidades

de transição e representam, a probabilidade do estado q_{t+1} ser S_j no instante $t+1$ dado que o estado q_t é S_i no instante t .

3.3. Cadeias de Markov

Um processo markoviano é uma cadeia de Markov quando as variáveis aleatórias $X(t)$ estão definidas em um espaço de estados discreto.

Consideram-se sistemas que podem ser descritos como estando em qualquer instante de tempo em um estado de um conjunto de N estados distintos (S_1, S_2, \dots, S_N) como ilustrado na figura 1 (onde $N=3$ para simplificar a explicação).

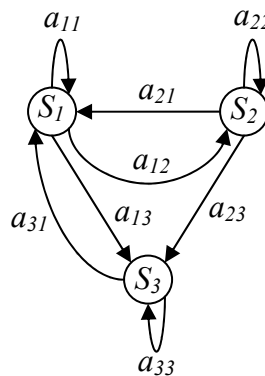


Figura 1 - Cadeia de Markov com 3 estados (S_1 a S_3), onde as setas indicam as transições possíveis entre os estados.

Para instantes de tempo igualmente espaçados, o sistema passa por uma mudança de estado (podendo retornar ao mesmo estado) de acordo com um conjunto de probabilidades de transição. Cada instante de tempo, onde ocorrem as mudanças de estado, é descrito por $t = 1, 2, \dots$, e o estado no tempo t é denotado com o símbolo q_t , considerando um espaço de tempo discreto.

Uma descrição probabilística completa do sistema requer a especificação do estado presente (no tempo t) e também dos estados anteriores. Para o caso especial de uma cadeia de Markov de primeira ordem, essa descrição pode ser truncada apenas no estado presente e no seu antecessor:

$$P(q_t = S_j | q_{t-1} = S_i, q_{t-2} = S_k, \dots) = P(q_t = S_j | q_{t-1} = S_i) \quad \text{Equação 1}$$

Considerando processos em que o lado direito da equação 1 é independente do tempo, tem-se um conjunto de probabilidades de transição de estado a_{ij} :

$$a_{ij} = P(q_t = S_j | q_{t-1} = S_i), \quad 1 \leq i, j \leq N. \quad \text{Equação 2}$$

As probabilidades de transição de estado têm as seguintes propriedades:

$$a_{ij} \geq 0 \quad \text{Equação 3a}$$

$$\sum_{i=1}^N a_{ij} = 1 \quad \text{Equação 3b}$$

Quando as probabilidades de transição não se alteram ao longo do tempo, diz-se que o processo é estacionário.

O processo estocástico mencionado acima poderia ser chamado de modelo de Markov observável uma vez que a saída do processo é um conjunto de estados em cada instante de tempo, onde cada estado corresponde a um evento físico observável.

3.4. Modelos de Markov Ocultos

Nos modelos de Markov descritos anteriormente, cada estado corresponde a um evento físico observável. Essa característica torna esses modelos muito restritos para solução de diversos problemas de interesse. Para estes casos aplica-se uma extensão chamada de modelos de Markov ocultos (do inglês: *Hidden Markov Models* - HMM), onde o estado não é diretamente observável, mas uma observação é uma função probabilística do estado.

Um HMM é um processo duplamente estocástico composto por um processo latente que não é observável (*hidden*), mas que se manifesta através de um outro processo estocástico o qual produz uma sequência de símbolos observáveis (Bunke & Caelli, 2001).

3.4.1. Elementos de um Modelo de Markov Oculto

Em um HMM, as observações (v_i) são símbolos emitidos por estados não observáveis (S_i) de acordo com determinadas funções probabilísticas, sendo que cada seqüência de estados é uma cadeia de Markov de primeira ordem. A figura 2 ilustra um HMM, sendo N o número de estados do modelo (o conjunto dos estados individuais é representado por $\mathbf{S} = \{S_1, \dots, S_N\}$ e o estado no tempo t pelo símbolo q_t) e M o número de símbolos distintos observáveis (o conjunto dos símbolos individuais é representado por $\mathbf{V} = \{v_1, \dots, v_M\}$).

Um HMM básico é constituído de 3 conjuntos de parâmetros principais:

1. as probabilidades de emissão de símbolo $B = \{b_{jk}\}$ – a probabilidade do símbolo v_k ser emitido pelo estado S_j , i.e.

$$b_{jk} = P(v_k \text{ em } t | q_t = S_j), \quad 1 \leq j \leq N \text{ e } 1 \leq k \leq M \quad \text{Equação 4}$$

2. as probabilidades de transição de estado $A = \{a_{ij}\}$ – a probabilidade de estar no estado S_j no instante de tempo subsequente dado que o estado atual é S_i , i.e.

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i), \quad 1 \leq i, j \leq N \quad \text{Equação 5}$$

3. A distribuição de probabilidade a priori $\pi = \{\pi_i\}$ do sistema estar em um dado estado S_i no instante inicial de tempo (não mostrado na Figura 2), i.e.

$$\pi_i = P(q_1 = S_i), \quad 1 \leq i \leq N \quad \text{Equação 6}$$

Com respeito à figura 2, se o sistema pode passar de um estado S_i para outro S_j , $a_{ij} > 0$ e se esses dois estados não são conectados, $a_{ij} = 0$. Apesar dos estados serem escondidos, em muitas aplicações práticas, estes possuem significado físico relevante para o problema em questão. Este é o caso da aplicação alvo do presente trabalho, em que os estados estão associados aos estágios fenológicos de culturas agrícolas.

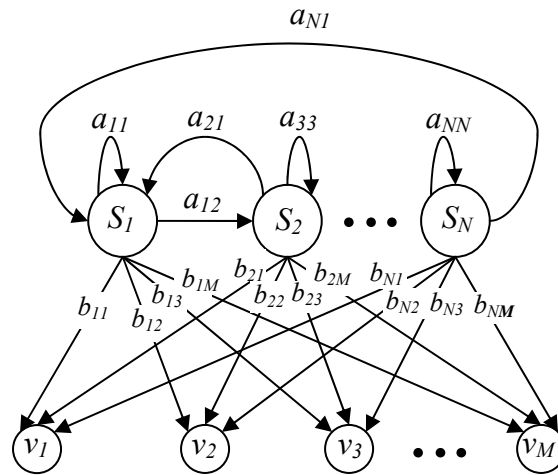


Figura 2 - Exemplo de um modelo de Markov Oculto ($S_i \rightarrow$ estados, $v_k \rightarrow$ símbolos observáveis, $a_{ij} \rightarrow$ probabilidade de transição de estado, $b_{jk} \rightarrow$ probabilidade de emissão de símbolo).

3.4.2. Problemas básicos associados a HMMs

A maioria das aplicações de HMMs envolve a solução de até três problemas principais:

- **Problema 1:** Considere uma seqüência de observações $O = \{O_1, \dots, O_M\}$. Dado o modelo $\lambda = (A, B, \pi)$, calcular $P(O | \lambda)$ (a probabilidade de que o modelo λ gere a seqüência de observações O).
- **Problema 2:** Para uma seqüência de observações $O = \{O_1, \dots, O_M\}$, dado o modelo $\lambda = (A, B, \pi)$, determinar a seqüência dos estados $Q = \{q_1, \dots, q_N\}$, de modo que $P(S, O | \lambda)$ (distribuição de probabilidade conjunta das seqüências de observações, O , e de estados S , dado o modelo λ) seja máxima.
- **Problema 3:** Estimar os parâmetros do modelo $\lambda = (A, B, \pi)$ de modo que $P(O | \lambda)$ ou $P(S, O | \lambda)$ seja maximizada.

O problema de classificação de culturas agrícolas, que é o tema do presente trabalho, corresponde ao problema 1. O problema 2 consiste em revelar a parte oculta do sistema, na aplicação alvo, determinar o estágio fenológico em que se encontra cada cultura em cada data. Ambos os problemas são, portanto relevantes neste trabalho.

O problema 3, em que se estimam os valores dos parâmetros do modelo, pressupõe a inexistência de dados de treinamento rotulados. Para a aplicação alvo deste trabalho está disponível um conjunto de dados em que o modelo (cultura), os estados (estágios fenológicos) e os símbolos emitidos (respostas espectrais) são conhecidos. A partir desta informação, podem-se estimar diretamente os parâmetros do modelo, como será descrito na Seção 4.2. Por isso o texto que se segue não apresenta a solução do problema 3. Detalhes sobre métodos de solução do problema 3 podem ser encontrados, por exemplo, em (Rabiner, 1989).

3.4.3.

Soluções para os problemas 1 e 2 associados a HMMs

3.4.3.1.

Solução do problema 1

Deseja-se calcular $P(O|\lambda)$, ou seja, a probabilidade de que o modelo dado λ produza a sequência de observações, $O = \{O_1 O_2 \dots O_T\}$, sendo T o número de observações na sequência. Esta probabilidade pode ser calculada através de um procedimento que consiste de duas etapas, denominadas direta e reversa.

A) A etapa direta:

Seja α_t a probabilidade (chamada usualmente na literatura de direta) da sequência parcial de observações, $O = \{O_1 O_2 \dots O_t\}$, até o tempo t e de que o sistema esteja no estado S_i no tempo t , dado o modelo λ , ou seja,

$$\alpha_t(i) = P(O_1 O_2 \dots O_t, q_t = S_i | \lambda) \quad \text{Equação 7}$$

O valor de $\alpha_t(i)$ pode ser calculado indutivamente, pelos seguintes três passos:

1. Inicialização:

$$\alpha_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \quad \text{Equação 8}$$

2. Indução:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(O_{t+1}), \quad 1 \leq i \leq T-1 \text{ e } 1 \leq j \leq N \quad \text{Equação 9}$$

3. Terminação:

$$P(O | \lambda) = \sum_{i=1}^N \alpha_T(i). \quad \text{Equação 10}$$

No primeiro passo a probabilidade direta é iniciada como a probabilidade conjunta do estado S_i e da observação inicial O_1 . O passo de indução, a etapa mais importante do cálculo, é ilustrada pela figura 3. Esta figura mostra que o estado S_j pode ser alcançado no instante $t+1$ partindo-se de qualquer um dos N estados possíveis, S_i , $1 \leq i \leq N$, no tempo t . Como $\alpha_t(i)$ é a probabilidade de $O_1 O_2 \dots O_t$ ser observada e o estado no tempo t ser S_i , o produto $\alpha_t(i) a_{ij}$ é então a probabilidade de $O_1 O_2 \dots O_t$ ser observada e o estado S_j ser alcançado no tempo $t+1$ via o estado S_i no tempo t . Somando este produto para todos os N estados possíveis S_i , $1 \leq i \leq N$, no tempo t obtém-se a probabilidade de estar em S_j no tempo $t+1$ com todas as observações parciais anteriores.

Tendo sido calculado este valor e conhecendo-se o estado S_j , é fácil notar que $\alpha_{t+1}(j)$ é obtido agregando-se a informação sobre a observação O_{t+1} à informação sobre o estado S_j , ou seja, multiplicando-se o valor do somatório pela probabilidade $b_j(O_{t+1})$. O cálculo da equação 9 é feito para todos os estados S_j , $1 \leq j \leq N$, e para $t = 1, 2, \dots, T-1$.

O cálculo de $P(O | \lambda)$ é então a soma das variáveis *diretas* finais $\alpha_T(i)$ já que $\alpha_T(i) = P(O_1 O_2 \dots O_T, q_T = S_i | \lambda)$. O cálculo da probabilidade *diretas* é ilustrado pela estrutura mostrada na figura 4.

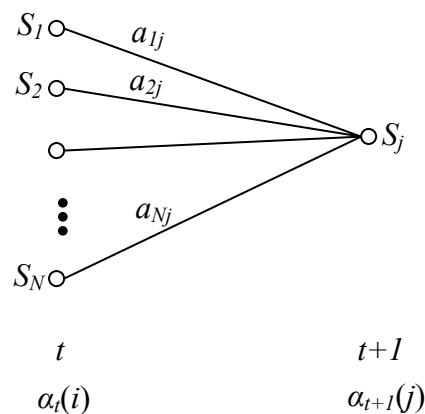


Figura 3 - Ilustração das seqüências de operações necessárias para o cálculo da variável direta $\alpha_{t+1}(j)$.

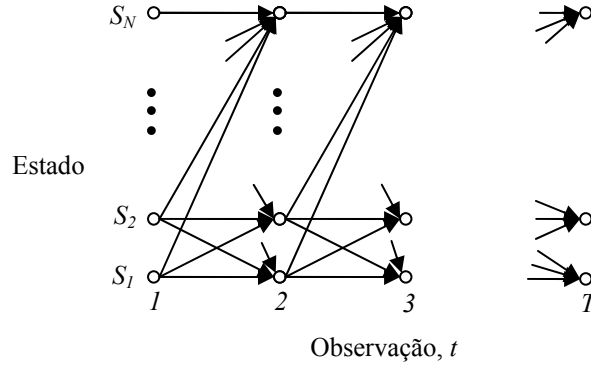


Figura 4 - Procedimento de cálculo de $\alpha_{t+1}(j)$ em função dos instantes das observações t e dos estados S_i

B) A etapa reversa:

De maneira similar ao procedimento descrito anteriormente, seja a variável $\beta_t(i)$ (chamada usualmente na literatura de reversa) definida como

$$\beta_t(i) = P(O_{t+1}O_{t+2}\dots O_T \mid q_t = S_i, \lambda), \quad \text{Equação 11}$$

i.e., $\beta_t(i)$ é a probabilidade da seqüência parcial de observações de $t+1$ até o fim, $\{O_{t+1}O_{t+2}\dots O_T\}$ dados o estado S_i no tempo t e o modelo λ . O valor de $\beta_t(i)$ pode ser calculado indutivamente, segundo os seguintes 3 passos:

1. Inicialização:

$$\beta_T(i) = 1, \quad 1 \leq i \leq N \quad \text{Equação 12}$$

2. Indução:

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(O_{t+1}) \beta_{t+1}(j), \quad t = T-1, T-2, \dots, 1 \text{ e } 1 \leq i \leq N \quad \text{Equação 13}$$

A etapa de inicialização define $\beta_T(i)$ arbitrariamente como 1 para todos os valores de i . A etapa de indução, ilustrada na figura 5, mostra que, para estar no estado S_i no tempo t tendo em vista a seqüência de observações do tempo $t+1$ em diante, é preciso considerar a possibilidade de que o sistema modelado estará em qualquer um dos possíveis estados S_j no tempo $t+1$. Isto é feito levando em conta a transição de S_i para S_j (o termo a_{ij}) para $1 \leq j \leq N$, assim como a observação O_{t+1} no estado S_j – o termo $b_j(O_{t+1})$ – e a seqüência parcial de observações restantes a

partir do estado S_j (o termo $\beta_{t+1}(j)$). $\beta_t(i)$ é calculado para $1 \leq t \leq T$ e $1 \leq i \leq N$. O cálculo da probabilidade *reversa* é ilustrado por uma estrutura similar à mostrada na figura 4.

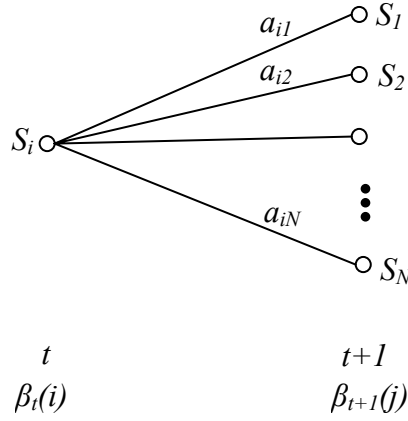


Figura 5 - Ilustração das seqüências de operações necessárias para o cálculo da variável reversa $\beta_{t+1}(j)$.

Neste trabalho foi escolhido o procedimento direto apesar da resposta dos dois procedimentos ser semelhante.

3.4.3.2. Solução do problema 2

Diferentemente do problema 1 para o qual uma solução exata pode ser obtida, existem diversas maneiras possíveis de resolver o problema 2. Uma técnica formal para achar a melhor seqüência de estados existente é o algoritmo de Viterbi.

Algoritmo de Viterbi:

Para achar a seqüência ótima de estados, $Q = \{q_1 q_2 \dots q_T\}$, tendo em vista uma dada seqüência de observações $O = \{O_1 O_2 \dots O_T\}$, define-se a variável $\delta_t(i)$:

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P(q_1 q_2 \dots q_t = S_i, O_1 O_2 \dots O_t \mid \lambda) \quad \text{Equação 14}$$

i.e., $\delta_t(i)$ é a probabilidade do caminho (seqüência de estados) mais provável que leva ao estado S_i no tempo t , considerando as primeiras t observações. Para indução tem-se

$$\delta_{t+1}(j) = \max_i [\delta_t(i) a_{ij}] b_j(O_{t+1}) \quad \text{Equação 15}$$

Para recuperar a seqüência de estados é necessário armazenar a informação sobre o argumento que maximizou a equação 15, para cada t e j . Isto é feito através da variável $\psi_t(j)$. O procedimento completo de obtenção da seqüência de estados pode ser resumido da seguinte forma:

1. Inicialização:

$$\delta_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \quad \text{Equação 16}$$

$$\psi_1(i) = 0. \quad \text{Equação 17}$$

2. Recursão:

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(O_t), \quad 2 \leq t \leq T \text{ e } 1 \leq j \leq N \quad \text{Equação 18}$$

$$\psi_t(j) = \arg \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}], \quad 2 \leq t \leq T \text{ e } 1 \leq j \leq N \quad \text{Equação 19}$$

3. Terminação:

$$P^* = \max_{1 \leq i \leq N} [\delta_T(i)] \quad \text{Equação 20}$$

$$q_T^* = \arg \max_{1 \leq i \leq N} [\delta_T(i)] \quad \text{Equação 21}$$

4. Levantamento da seqüência de estados:

$$q_t^* = \psi_{t+1}(q_{t+1}^*), \quad t = T-1, T-2, \dots, 1 \quad \text{Equação 22}$$

É importante destacar que a implementação do algoritmo de Viterbi é similar (exceto pela etapa de levantamento da seqüência de estados) ao cálculo das equações 8 a 10 (procedimento *direto*). A maior diferença é a maximização

(equação 18) dos estados anteriores que é usada no lugar do procedimento de soma na equação 9. Cabe ressaltar que a estrutura mostrada na figura 4 também ilustra o algoritmo de Viterbi.

Mais detalhes sobre soluções dos problemas relacionados aos HMMs e aos HMMs em si encontram-se em (Rabiner, 1989).

4 Método Proposto

4.1. Descrição geral do modelo

Neste trabalho, cada classe (cobertura vegetal) possui um HMM específico. Os estágios fenológicos correspondem a estados e os símbolos observáveis a vetores contendo os números digitais registrados pelos sensores orbitais em cada banda espectral, assim como o NDVI (Rouse et al, 1973) calculado a partir das bandas 4 (infravermelho próximo - NIR, *near infrared*) e 3 (vermelho - Red) de acordo com

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad \text{Equação 23}$$

O HMM básico mostrado na figura 6 foi escolhido para modelar o comportamento temporal da cana de açúcar, da soja e do milho. As setas ilustram como os estados estão relacionados ao longo do tempo. Os estados PP, CR, AD e PA correspondem aos estágios de pré-plantio, crescimento, fase adulta e palhada (pós-colheita) respectivamente.

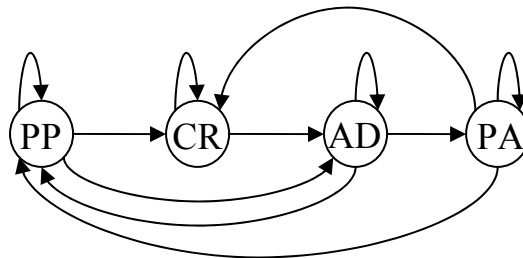


Figura 6 - HMM usado neste trabalho para cana de açúcar, soja e milho (PP = pré-plantio, CR = crescimento, AD = fase adulta e PA = palhada).

Para a pastagem e a mata-galeria, no período de tempo coberto pelas imagens disponíveis na base de dados, não ocorrem mudanças significativas nos atributos espectrais, isto porque essas coberturas vegetais não são plantadas nem colhidas e sim se encontram nas áreas em questão há muito tempo. Portanto um HMM específico foi escolhido para essas classes, tendo um único estado *AD* correspondente à fase adulta (figura 7).



Figura 7 - HMM usado neste trabalho para pastagem e mata-galeria (AD = fase adulta).

As figuras 8 e 9 mostram exemplos de áreas de cada cultura mencionada anteriormente (e estágios fenológicos), nas imagens de satélite (composição RGB 543). As amostras em questão encontram-se no centro das imagens nessas tabelas.

	Pastagem	Mata de Galeria
Fase Adulta		

Figura 8 - Exemplo de áreas de pastagem e mata-galeria nas imagens de satélite utilizadas.

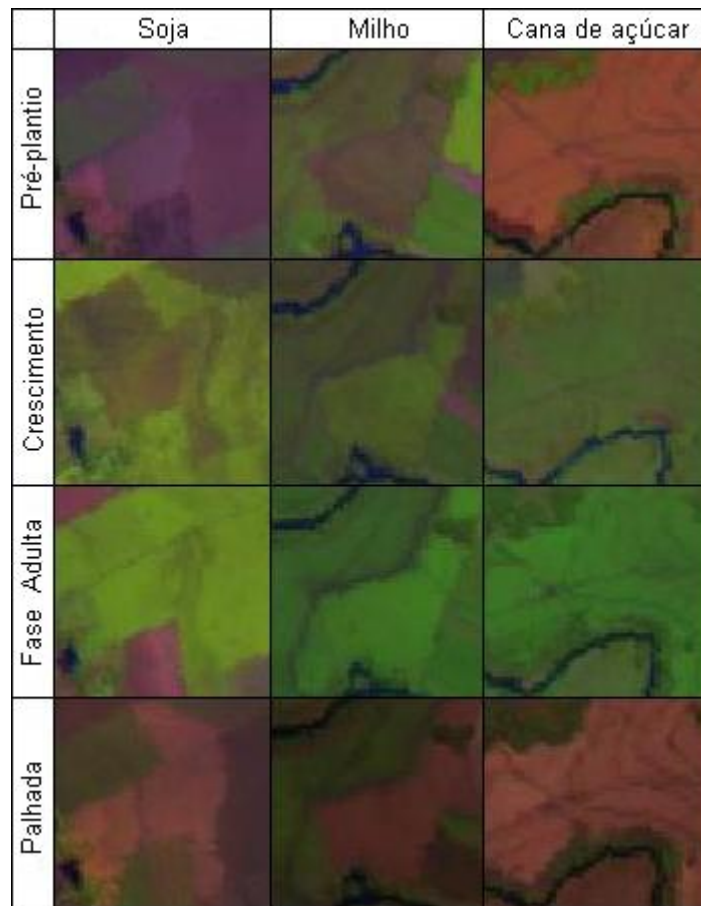


Figura 9 - Exemplo de áreas de soja, milho e cana de açúcar nas imagens de satélite utilizadas.

4.2.

Ajustando o modelo à aplicação – Estimativa de parâmetros

O problema considerado neste trabalho diverge em alguns aspectos da descrição básica dos HMMs apresentada nas seções anteriores. Em primeiro lugar, as probabilidades de emissão de símbolos (b_{jk}) dependem de efeitos sazonais, os quais não são totalmente compensados na etapa de pré-processamento das imagens. Segundo, a distribuição de probabilidade a priori (π_i) não é constante ao longo do ano (ver Capítulo 5, Seção 1). Isto pode ser explicado pelo fato de cada cobertura vegetal possuir meses preferenciais para o seu plantio e colheita, condicionados principalmente pelo clima e por suas características próprias. Portanto, a probabilidade a priori de uma determinada cultura estar em um dado estado varia ao longo do ano. Terceiro, o modelo básico descrito no Capítulo 3 assume que os símbolos são emitidos em intervalos de tempo regulares. Na maioria das aplicações reais nem todas as imagens adquiridas com uma frequência

fixa são utilizáveis, por exemplo, devido à cobertura de nuvens sobre a área geográfica de interesse. É importante mencionar também que o modelo básico representado na figura 6 mostra transições possíveis entre estados que, no caso de símbolos emitidos em intervalos regulares, não seriam plausíveis. Por exemplo, a transição de PP para AD, sem passar pelo estado CR. A transição de PP para AD pode existir entre duas imagens consecutivas do banco de dados, caso o tempo que decorre entre elas seja suficientemente longo, de modo que a etapa intermediária de crescimento não tenha sido capturada. Se a base de dados disponível for analisada, existem os pares de imagens consecutivas de Outubro de 2002 e Janeiro de 2003 e também de Outubro de 2003 e Janeiro de 2004 (Apêndice II). O intervalo Outubro-Janeiro é justamente a primeira metade do ciclo da soja. Com isso, existem diversas amostras nesses pares de datas onde na data inicial (Outubro) existe a presença de soja no pré-plantio e na data seguinte (Janeiro) soja na fase adulta.

Assim, um HMM para o problema tratado neste trabalho tem que considerar as probabilidades de emissão de símbolos, assim como matrizes de probabilidade de transição de estado distintas para cada par de imagens consecutivas disponíveis no conjunto de dados.

Em benefício da legibilidade, a notação adotada neste texto para as probabilidades a priori (π_i) de ocorrência de estado, para as probabilidades de emissão de símbolos (b_{jk}) e para as probabilidades de transição de estados (a_{ij}) omite a dependência do tempo.

Em relação às probabilidades de emissão de símbolos, admitiu-se, neste trabalho, que elas possuem distribuição Gaussiana. Dessa forma, a densidade de probabilidade de emissão de um símbolo x (vetor contendo as bandas espectrais e o NDVI) é dada por:

$$p = \frac{1}{(2\pi)^{d/2} |\Sigma_{cs}|^{1/2}} \exp \left[-\frac{(x - \mu_{cs})^T \Sigma_{cs}^{-1} (x - \mu_{cs})}{2} \right] \quad \text{Equação 24}$$

onde μ_{cs} e Σ_{cs} correspondem respectivamente ao vetor de média e a matriz de covariância para a cultura c e o estado s , d é a dimensão de x e p é a função densidade de probabilidade para o caso de x possuir distribuição normal

multivariada (que é usada neste trabalho para descrever a probabilidade de emissão de símbolos).

A estimativa dos valores dos parâmetros dos modelos é feita para cada tipo de cultura agrícola separadamente a partir de amostras de treinamento, seguindo os passos mostrados na figura 10.

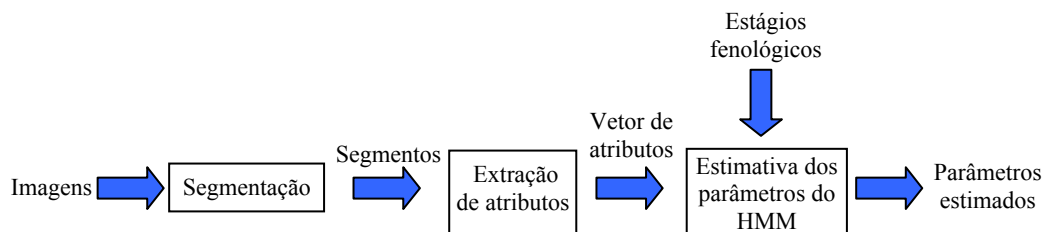


Figura 10 - Processo de estimativa dos valores dos parâmetros do HMM de cada cultura.

4.2.1. Estimativa da matriz de transição de estado

As probabilidades de transição de estado a_{ij} são estimadas para cada par de imagens consecutivas do conjunto de treinamento, em que apenas segmentos das imagens, correspondentes à cultura sendo modelada, são considerados.

Para esta etapa é utilizada a técnica *leave-one-out* (Brovelli et al., 2006).

Em primeiro lugar, uma matriz acumuladora é criada para cada par de datas por cobertura vegetal. Suas linhas e colunas correspondem respectivamente aos estados nas datas anterior e posterior. Em seguida, para todas as linhas e colunas, o elemento na linha i e na coluna j é incrementado para cada amostra de segmento nos estados S_i e S_j em datas consecutivas. Cabe ressaltar que as transições de estado correspondentes à sequência de teste não são levadas em consideração no cálculo da matriz acumuladora (o incremento não é realizado).

Uma vez tendo sido considerados todos os segmentos de treinamento para aquela cobertura vegetal e par de datas, os elementos da matriz acumuladora são analisados para correção. Os elementos nulos que correspondem a transições possíveis naquele par de datas são carregados com o valor “1” (um). Esta correção se justifica em vista da escassez de padrões disponíveis no conjunto de dados. Ocorre em alguns casos que o segmento retirado para teste na técnica *leave-one-out* é o único exemplo de transição entre dois estados. Sem esta correção a probabilidade de transição seria estimada como sendo nula, o que teria

impacto não desprezível na avaliação de desempenho do método. A informação quanto às transições possíveis é considerada neste trabalho com um conhecimento *a priori* fornecido por um foto intérprete conhecedor das práticas e coberturas vegetais analisadas.

Por último, a matriz acumuladora é normalizada, dividindo cada um de seus elementos pela soma da linha correspondente. O resultado é a estimativa da matriz de transição de estado para aquela cobertura vegetal e para aquele par de datas.

4.2.2.

Estimativa das probabilidades a priori de estados

O procedimento para estimar a distribuição de probabilidade a priori dos estados π_i é similar (utiliza-se também aqui a técnica *leave-one-out*).

Novamente, apenas segmentos das imagens correspondentes às classes da cobertura vegetal modelada naquela data são considerados. Um vetor acumulador é criado, contendo um elemento por estado. O *i-ésimo* elemento é incrementado para cada amostra de treinamento no estado S_i . Assim como na subseção anterior, a informação referente à sequência de teste não é levada em consideração no cálculo do vetor acumulador (o incremento não é realizado).

Como na seção anterior, o vetor acumulador sofre uma correção, ou seja, os elementos nulos deste vetor, que correspondem a um estado possível de ocorrer no par de datas em questão, é carregado com o valor “1”. Por último, o vetor acumulador é dividido pela soma de todos os seus elementos. O resultado é a estimativa da distribuição de probabilidade a priori dos estados para aquela cobertura vegetal e para aquela data.

4.2.3.

Estimativa da emissão de símbolo

Como foi mencionado anteriormente, foi assumido que as probabilidades de emissão de símbolos possuem distribuição Gaussiana; com isso, o problema de estimar esses parâmetros reduz-se a estimar os vetores média e as matrizes de covariância para cada estado de cada cultura em cada imagem.

Neste trabalho classificam-se segmentos. Cada segmento é descrito por um vetor que contém os valores médios dos pixels que compõem o segmento a ser

classificado nas seis bandas e do correspondente valor de NDVI. Assim, todos os pixels que compõem um mesmo segmento dão origem a um único padrão a ser classificado. Tem-se desta forma um número muito menor de padrões do que haveria no caso da classificação pixel-a-pixel, para um mesmo conjunto de imagens. Esta redução no número de padrões pode implicar um número insuficiente de padrões de treinamento para estimar o vetor média e especialmente a matriz de covariância que representam os símbolos emitidos por cada estado.

Para contornar este problema, o vetor média e a matriz de covariância foram estimados a partir dos valores espectrais dos pixels contidos em cada segmento. Embora na fase de teste se classifiquem segmentos descritos por seus valores espectrais médios, as estimativas do vetor média e da matriz de covariância é feita com base nos valores dos pixels internos aos segmentos de treinamento.

O procedimento adotado foi o seguinte: a segmentação das imagens foi realizada em duas etapas. Na primeira etapa, o conjunto de 12 imagens, cada uma com seis bandas foi empilhado de modo a formar uma imagem sintética contendo 72 bandas. Este conjunto foi segmentado como se fosse uma única imagem, utilizando o método baseado em *watershed* descrito em (Mota et al., 2007). Os valores dos parâmetros de segmentação foram escolhidos por tentativa-e-erro, até que o resultado da segmentação fosse considerado consistente em toda a sequência de imagens por um foto intérprete, ou seja, até que cada segmento produzido fosse homogêneo em cada uma das 12 imagens.

A segunda etapa de segmentação foi realizada para cada uma das imagens individualmente utilizando o mesmo algoritmo e o mesmo conjunto de valores dos parâmetros de segmentação da primeira etapa. De um modo geral, os segmentos resultantes desta segunda etapa são maiores e contêm os segmentos obtidos na primeira etapa.

Os segmentos assim obtidos são divididos em duas partes, contendo cada uma aproximadamente a metade dos pixels. As estimativas do vetor médias e matriz de covariância foram feitas considerando os pixels de uma das partes. Para a fase de teste, foram considerados os valores espectrais médios da outra parte do segmento.

As estimativas do vetor média são obtidas através da média dos vetores de treinamento que representam o tipo de cobertura vegetal e o estágio fenológico

naquela data particular. Similarmente, a covariância amostral (Johnson & Wichern, 1998) é usada para estimar a matriz de covariância.

4.3. Classificação

Uma vez que o HMM tenha sido definido e seus parâmetros estimados, a classificação da imagem é feita da seguinte forma (a figura 11 ilustra esse processo).

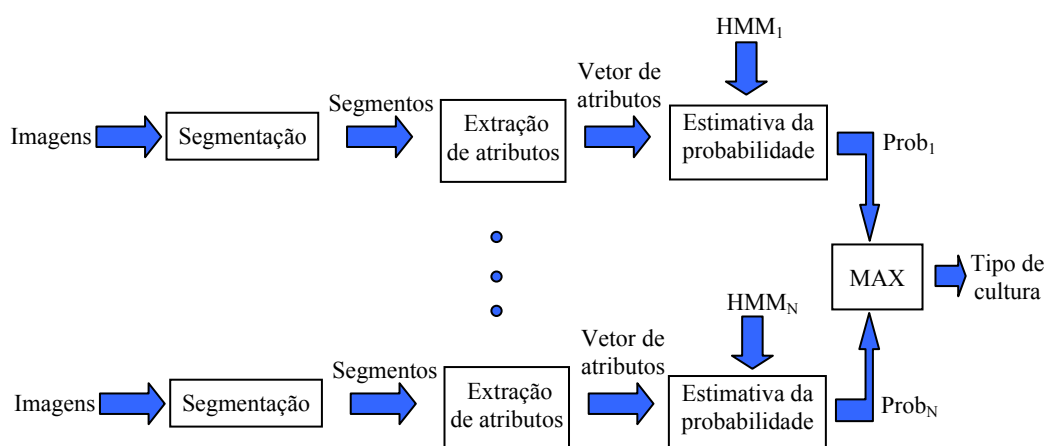


Figura 11 - Processo de classificação.

Cada segmento de teste é representado, em cada data, por um vetor composto pelos valores das bandas espectrais e do NDVI naquela data. Este vetor corresponde ao símbolo emitido pelo segmento naquela data.

Consideram-se nesta avaliação apenas seqüências de datas em que o segmento analisado pertence a uma única cultura. O classificador calcula, para cada modelo, a probabilidade da cultura correspondente emitir a seqüência de vetores de símbolos observada naquelas datas (solução do problema 1 mencionado na Seção 3.4.3.1). O segmento de teste é atribuído à classe cujo modelo retorna a maior probabilidade de emissão. Uma descrição detalhada de como as probabilidades de emissão das seqüências de símbolos são calculadas de modo eficiente encontra-se na Seção 3.4.3.1. A Seção 3.4.3.2 descreve o algoritmo que permite inferir, para cada modelo e para cada segmento, qual o estado mais provável em cada data que compõe a seqüência.

5

Análise experimental

5.1.

Culturas e respectivos ciclos fenológicos

Os ciclos e datas de plantio e colheita das principais culturas encontradas em uma área de estudo determinam as variações temporais da quantidade de área foliar, volume de fitomassa e cobertura do solo. O seu conhecimento fornece a base para o entendimento e a discussão dos comportamentos espectrais apresentados pelos talhões em determinado período do ano. Na região onde está a área de estudo, os ciclos das principais culturas anuais (cereais) estão, de modo geral, compreendidos entre Outubro e Abril, em cada ano agrícola.

5.1.1.

Cana-de-açúcar

Em São Paulo, o cultivo da cana-de-açúcar (*CA*) (*Saccharum spp.*) segue basicamente dois ciclos: um de 12 meses (“cana-de-ano”) e outro de 18 meses (“cana-de-ano-e-meio”). A cana-de-ano-e-meio é plantada entre Janeiro e Março e a cana-de-ano, entre Outubro e Novembro. Importante é destacar que cada talhão de cana pode ser colhido por cerca de cinco ou seis ciclos agrícolas seguidos. Desta forma, a cana tem um ciclo denominado “semi-perene”, que é diferente dos ciclos das culturas de grãos, tanto pela sua duração, quanto pela sua dinâmica fenológica.

Para as áreas novas (onde há cana recém-plantada), a massa verde da cana de ano-e-meio começa a recobrir completamente a superfície a partir de Outubro, quando há maior calor e mais precipitação pluviométrica; por outro lado, para as áreas novas da cana-de-ano, em Abril ou Maio deverá ocorrer o recobrimento verde total da superfície, e depois, a fitomassa verde tende a continuar aumentando o seu índice de área foliar até próximo da colheita. Em cada ano, o período de colheita inicia-se em Abril e se estende até Novembro e, desta forma, em uma mesma data de imagem satelitária, será possível encontrar: palha de cana

colhida, cana recém-plantada, cana em desenvolvimento vegetativo, cana adulta e também solos expostos preparados para o plantio.

A figura 12 ilustra os ciclos fenológicos da cana de 12 meses e da cana de 18 meses.

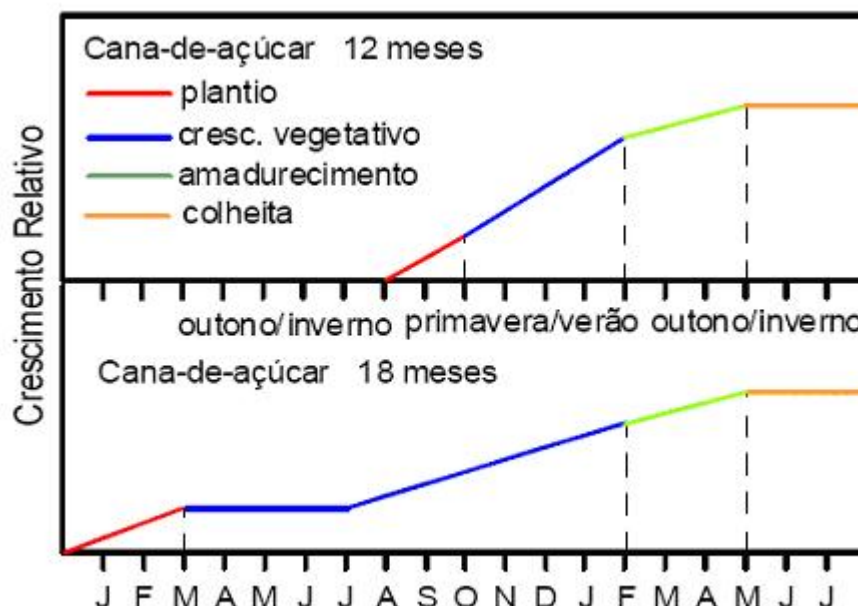


Figura 12 - Ciclos fenológicos da cana de açúcar de 12 meses (cana-de-ano) e 18 meses (cana-de-ano-e-meio). Figura obtida em (Gleriani, 2004).

5.1.2. Culturas de ciclo curto (cereais)

A soja (*SJ*) e o milho (*MI*) são culturas denominadas “anuais” ou “de ciclo curto”, uma vez que, em um período entre 110 e 140 dias, elas podem completar os seus respectivos ciclos fenológicos. São plantadas, em geral, no final de Outubro ou no início de Novembro. Germinam oito a dez dias após o plantio, iniciando então o seu desenvolvimento vegetativo, de modo a recobrir completamente a superfície cerca de sessenta dias após a germinação. Atingem um pico de fitomassa verde e, em seguida, iniciam o processo de enchimento dos grãos, quando começam a diminuir as folhas verdes, aumentando as folhas amareladas, que chegam a secar e a cair, passando então a expor novamente os solos de fundo até a colheita. A figura 13 ilustra o ciclo fenológico do milho, e a figura 14, o da soja.

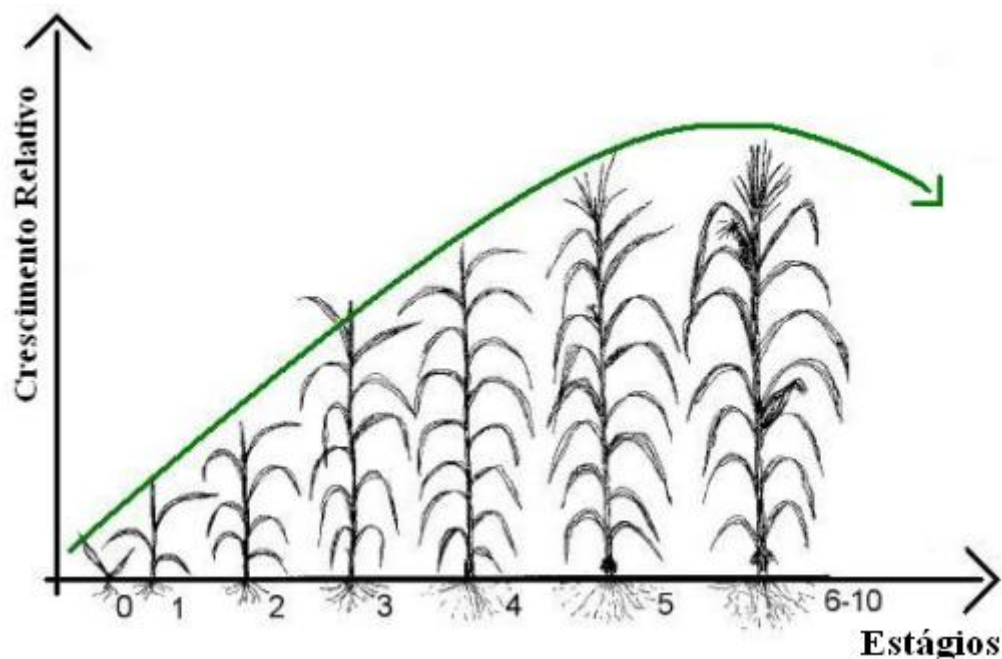


Figura 13 - Ciclo fenológico da cultura de milho. Figura adaptada de (Gleriani, 2004).

Na figura 13, a linha verde ilustra o perfil espectro-temporal do milho.

Segundo (Gleriani, 2004), o estágio 0 ocorre 4 a 5 dias após o plantio do milho e é quando há início de emergência. No estágio 1 (aproximadamente 15 dias após a emergência), fica visível o colar da quarta folha. No estágio seguinte, perto de 30 dias após a emergência, fica visível o colar da oitava folha. Já 40 dias depois da emergência, inicia-se o estágio 3, onde ocorre um aumento quase linear de massa seca e fica visível o colar da décima segunda folha. O estágio 4 ocorre de 55 a 58 dias após a emergência, e é nele que fica visível o colar da décima sexta folha. O estágio seguinte ocorre aproximadamente 65 dias após a emergência e é nele que aparecem as primeiras bonecas (espigas) entre a décima segunda e décima quarta folhas. O estágio 6 ocorre doze dias após 75% das plantas estarem embonecadas, e é nesse período que as espigas estão com o tamanho próximo do máximo. Inicia-se aqui a formação dos grãos. Doze dias após o estágio anterior, começa o estágio 7. Nesta fase, acontece um rápido desenvolvimento dos grãos. No estágio 8, doze dias após o estágio 7, aproximadamente 25% das plantas possuem grãos com dente. O estágio 9 inicia-se doze dias após a fase 8, e é nele que a taxa de acúmulo de matéria seca pela planta começa a diminuir e todos os grãos estão dentados. O estágio 10 ocorre doze dias após o estágio 9, e é quando há plantas com folhas senescentes e alta

perda de umidade. Após o estágio 10, inicia-se a senescência, quando as folhas secam e caem e então a resposta espectral começa a diminuir.

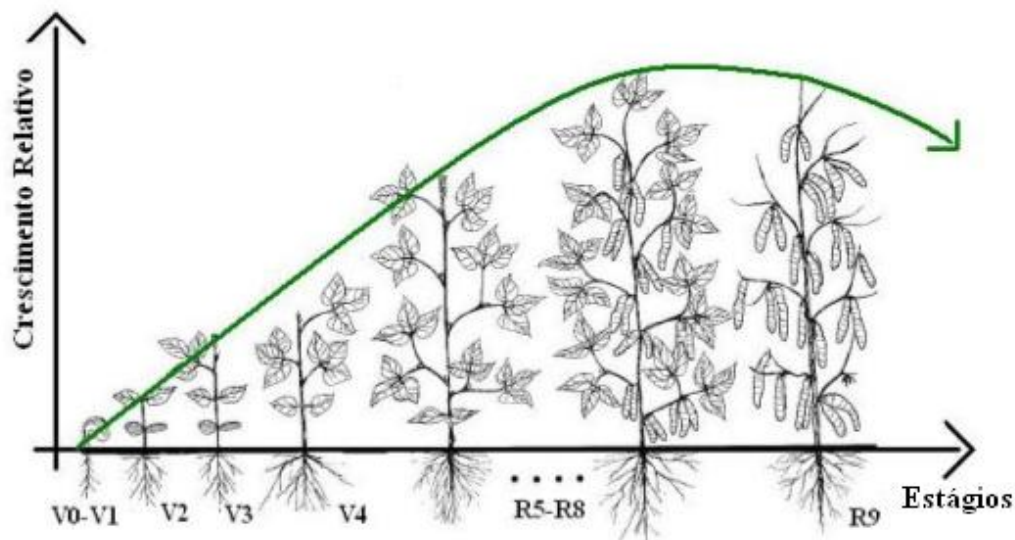


Figura 14 - Ciclo fenológico da cultura de soja. Figura adaptada de (Gleriani, 2004).

Na figura 14, a linha verde ilustra o perfil espectro-temporal da soja.

Segundo (Gleriani, 2004), a germinação (V0) ocorre de 4 a 5 dias após a sementeira. Já a emergência (V1) inicia-se quando os cotilédones aparecem em pelo menos 50% das plantas. O estágio das folhas primárias (V2) começa quando as folhas primárias de 50% das plantas estão desenvolvidas. A etapa da primeira folha trifoliada (V3) começa quando 50% das plantas estão com a folha trifoliada aberta. Por sua vez, o estágio da terceira folha trifoliada (V4) inicia-se quando 50% das plantas estão com 50% das terceiras folhas trifoliadas abertas. A pré-floração (R5) é o primeiro estágio reprodutivo da cultura. Considera-se o ingresso neste estágio quando 50% das plantas apresentam o primeiro botão floral. A floração (R6) ocorre quando 50% das plantas apresentam a primeira flor e a formação das vagens (R7) tem início quando 50% das plantas apresentam a primeira vagem. O estágio de enchimento de vagens (R8) inicia-se quando 50% das plantas começam a ter enchimento de grãos. No final desta fase a cor dos grãos deixa de ser verde e adquire a cor característica da variedade, e se inicia o desfolhamento das plantas. A maturação (R9), por sua vez, ocorre quando metade da população apresenta a primeira vagem descolorida e seca.

5.1.3. Pastagem

As pastagens (*PS*) apresentam uma dinâmica fenológica e espectral distinta em relação à dinâmica das culturas agrícolas citadas acima. Muitas vezes a dinâmica fenológico-espectral das pastagens depende dos tipos de manejos utilizados pelos pecuaristas, porém, em geral, as pastagens apresentam-se mais secas e ralas entre abril e setembro, quando, com a chegada da época chuvosa, iniciam seu revigoramento, com aumento do índice de área foliar, permanecendo então com vigor vegetativo verde entre novembro e março.

5.1.4. Outras coberturas vegetais

Além das coberturas vegetais mencionadas anteriormente, a mata-galeria (GA) também foi considerada neste trabalho. Esta, também conhecida como mata ciliar, é uma variante de cerrado, porém composta por vegetação mais frondosa. Ocupa áreas de vales úmidos ao longo de cursos de água, em solos aluvionados por conta da erosão.

Outras classes de cobertura de solo estão presentes na área de estudo: áreas urbanas, ruas, florestas e corpos d'água. Estas aparecem poucas vezes e praticamente não sofrem mudanças ao longo da sequência de imagens, e por isso, não foram tratadas neste trabalho.

5.2. Área geográfica

A área de estudo correspondeu a três municípios localizados ao norte do estado de São Paulo: Ipuã, Guará e São Joaquim da Barra (dentro de um retângulo definido pelas seguintes coordenadas: 20°16'30"S até 20°40'00"S x 47°37'36"W até 48°13'50"W), somando uma área de 124.100 ha (Figura 15). Trata-se de uma região essencialmente agrícola, muito representativa da agricultura praticada em grande parte das regiões sul e sudeste brasileiras, tendo como cultivos principais a cana-de-açúcar, a soja e o milho. Esta área apresenta relevos planos a suavemente ondulados, clima tropical com inverno seco, com temperatura média anual de 22,9°C e precipitação média anual de 1480 mm.

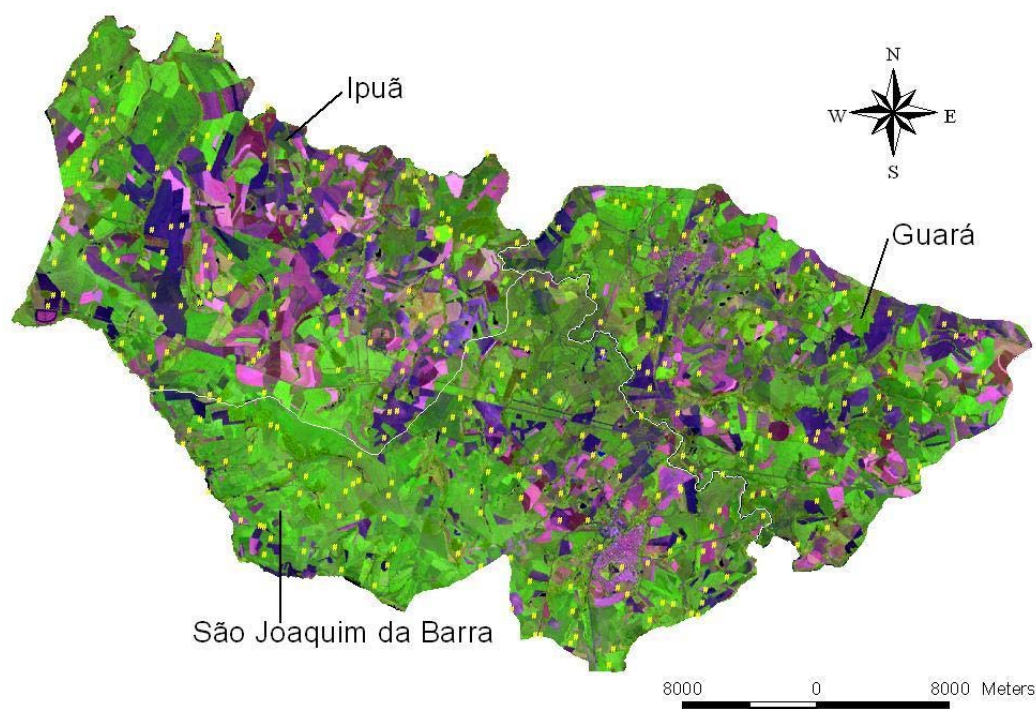


Figura 15 - Área de estudo, estado de São Paulo, Brasil. Os pontos amarelos correspondem a 316 pontos amostrais de referência.

Tabela 1 - Imagens disponíveis no conjunto de dados utilizado no presente trabalho.

	2002	2003	2004
Janeiro		(ETM+) 08/01/03	(TM) 19/01/04
Fevereiro		(ETM+) 09/02/03 - 25/02/03	
Abril		(ETM+) 14/04/03 - 30/04/03	
Maiο		(ETM+) 16/05/03	
Julho		(TM) 27/07/03	
Agosto		(TM) 12/08/03	
Setembro	(ETM+) 02/09/02		
Outubro	(ETM+) 20/10/02	(TM) 15/10/03	

5.3. Imagens

A base de dados contém um total de 12 imagens do satélite Landsat, órbita/ponto WRS 220/74, de 2002 a 2004 (Tabela 1), tanto do sensor TM/Landsat-5, como do ETM+/Landsat-7 (Sanches, 2004). As bandas espectrais destes sensores, cuja resolução espacial é de 30m, têm as seguintes denominações e larguras: Banda 1 (450-520 nm); Banda 2 (530-610 nm); Banda 3 (630-690 nm); Banda 4 (780-900 nm); Banda 5 (1550-1750 nm); Banda 7 (2090-2350nm), sendo que a banda termal não foi utilizada. Descrições detalhadas sobre os sensores TM e ETM+/Landsat podem ser encontradas em (Lillesand & Kiefer, 1994).

5.4. Pré-processamento das imagens

As imagens Landsat estavam no formato Geotiff e, para o georeferenciamento, foram utilizados treze pontos de controle obtidos com GPS em campo. O método de reamostragem aplicado foi o do vizinho mais próximo, por ser o que melhor preserva a radiometria da imagem original (Mather, 1993; Richards, 1995).

Foi feita uma correção para minimizar efeitos atmosféricos presentes nas imagens multitemporais, uma vez que a atmosfera, pelos fenômenos de espalhamento, absorção e refração da energia eletromagnética, afeta a radiância efetivamente medida pelos sensores orbitais. A técnica utilizada foi a de *Subtração do Pixel Escuro*, também conhecida como *Dark-object subtraction technique*, desenvolvida por (Chavez, 1988).

Como um mesmo objeto pode apresentar distintos valores digitais em imagens de diferentes datas de aquisição, em função de diferenças dos ângulos solares e do espalhamento atmosférico, conjuntos multitemporais de imagens precisam ser uniformizados radiométricamente. Assim, no presente trabalho, esta uniformização foi feita segundo metodologia proposta por (Gürtler, 2003).

Por outro lado, como os algoritmos de classificação tomam por base o comportamento espectral dos objetos a serem classificados em imagens de diferentes datas, os valores de níveis de cinza foram convertidos para valores de reflectância, que são valores com significado físico e, assim, representam

corretamente os diferentes objetos e as condições em que se encontravam no momento da aquisição de cada imagem. Esta transformação para reflectância foi feita segundo metodologia proposta por (Luiz et al., 2003).

É importante mencionar que a etapa de pré-processamento das imagens não foi executada no presente trabalho e sim em (Sanches, 2004), onde uma descrição mais detalhada desta etapa pode ser encontrada.

5.5.

Dados de referência

5.5.1.

Segmentação

Nesta seção acrescentam-se detalhes à descrição já apresentada na seção sobre o algoritmo de segmentação na Seção 4.2.3. As 12 imagens disponíveis para este trabalho foram “empilhadas” e segmentadas utilizando-se o método de *watershed*, conforme detalhado em (Mota et al., 2007). Este método foi escolhido por ser muito rápido.

Todos os algoritmos de segmentação envolvem ajuste de parâmetros para que objetos significativos, para uma determinada aplicação, sejam obtidos. Esta tarefa é geralmente executada por um foto intérprete, assim como neste trabalho, o qual sabe o que esperar do resultado da segmentação e, sendo assim, adapta os parâmetros do algoritmo selecionado à aplicação. Por esta razão, é de se esperar que a escolha do algoritmo de segmentação não afete significativamente o resultado final. Um algoritmo do tipo crescimento de região poderia ter sido usado nesta etapa, por exemplo.

Os valores dos parâmetros da segmentação foram escolhidos depois de um trabalho de tentativa e erro, até que o resultado fosse visualmente consistente, considerando todas as imagens no banco de dados.

5.5.2.

Processo de classificação visual

As coberturas vegetais consideradas neste estudo são as principais encontradas na área de estudo: cana-de-açúcar, soja, milho, pastagem e mata-galeria. Para as culturas de soja, milho e cana, o ciclo fenológico foi subdividido

em quatro fases: pré-plantio (época em que as superfícies aparecem como solos expostos nas imagens orbitais; sigla PP), crescimento (época em que as plantas recobrem completamente a superfície; sigla CR), adulta (época em que as plantas já se encontram em pleno vigor verde e podendo estar em início de senescimento; sigla AD), pós-colheita (época em que as áreas em que antes havia uma cultura agrícola, na imagem apresentam respostas espectrais típicas de palhaça seca remanescente da colheita; sigla PA).

As áreas de pastagem e mata-galeria apresentam uma dinâmica fenológico-espectral típica e são representadas por um único estado (AD).

Foram observados 316 pontos amostrais distribuídos pela área de estudo, estando representados em amarelo na figura 15. Esses pontos se situavam em segmentos distintos, tendo estes sido delimitados na etapa de segmentação. Cada um desses segmentos foi classificado, indicando seu tipo de cobertura vegetal e respectivo estágio fenológico, considerando cada data de passagem TM ou ETM+, segundo as classes acima indicadas por dois fotointérpretes experientes, tomando por base dois trabalhos de campo realizados (um em março e outro em agosto de 2003).

Primeiramente, os dois especialistas trabalharam independentemente e cada um interpretou todos os 316 segmentos de referência. Em seguida, foi feito um trabalho em conjunto de comparação de resultados a fim de gerar um resultado único para a classificação. Quando havia alguma discordância entre os dois, ambos os especialistas consideravam novamente a seqüência multitemporal de imagens e, assim, determinavam a classificação de consenso. Os segmentos foram então usados para construir os conjuntos de treinamento e teste. Cada segmento obtido foi dividido meio a meio, cada parte contendo aproximadamente a mesma quantidade de pixels. Uma parte foi usada para treinamento e a outra para teste.

As figuras 16, 17 e 18 mostram as culturas de milho, cana de açúcar e soja em diferentes estágios de crescimento. Nelas estão recortes de imagens Landsat na composição RGB 453 e fotos digitais obtidas durante o trabalho de campo, além da indicação dos respectivos estágios fenológicos. Estas figuras foram obtidas em (Sanches, 2004), e neste trabalho podem ser encontradas informações adicionais sobre o trabalho de campo.

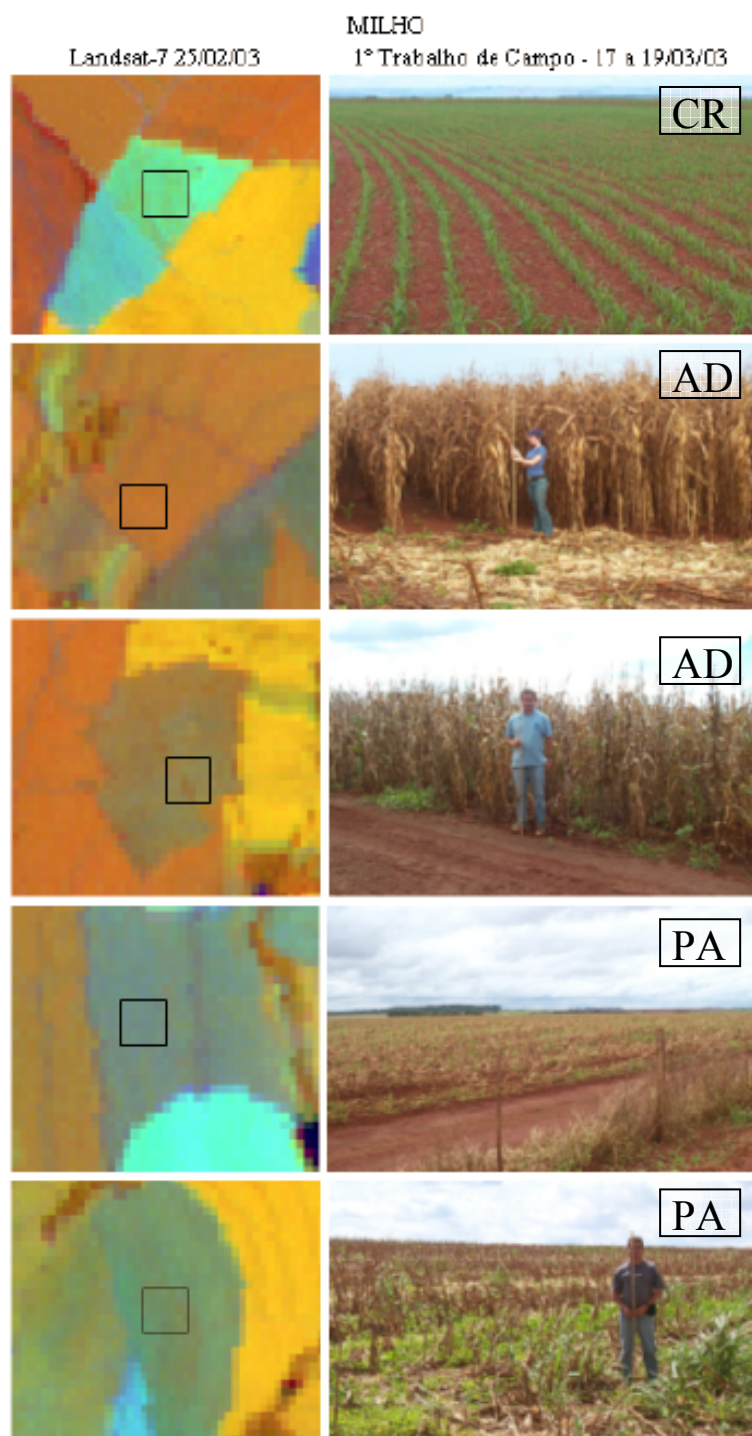


Figura 16 - Recorte de imagem Landsat na composição colorida RGB 453, e foto digital de campo de talhões de milho em diferentes estágios de crescimento. FONTE: (Sanches, 2004).

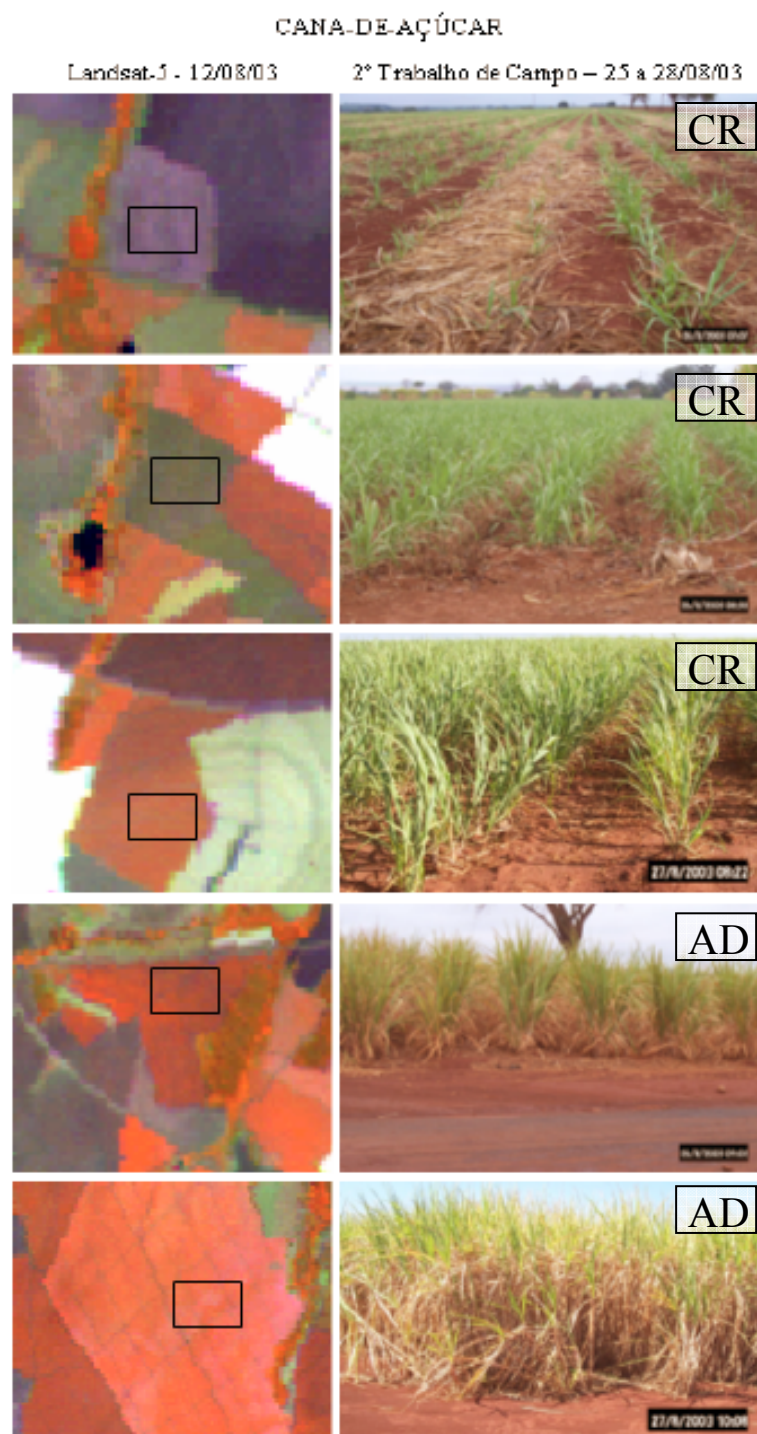


Figura 17 - Recorte de imagem Landsat na composição colorida RGB 453, e foto digital de campo de talhões de cana-de-açúcar em diferentes estágios de crescimento. FONTE: (Sanches, 2004).

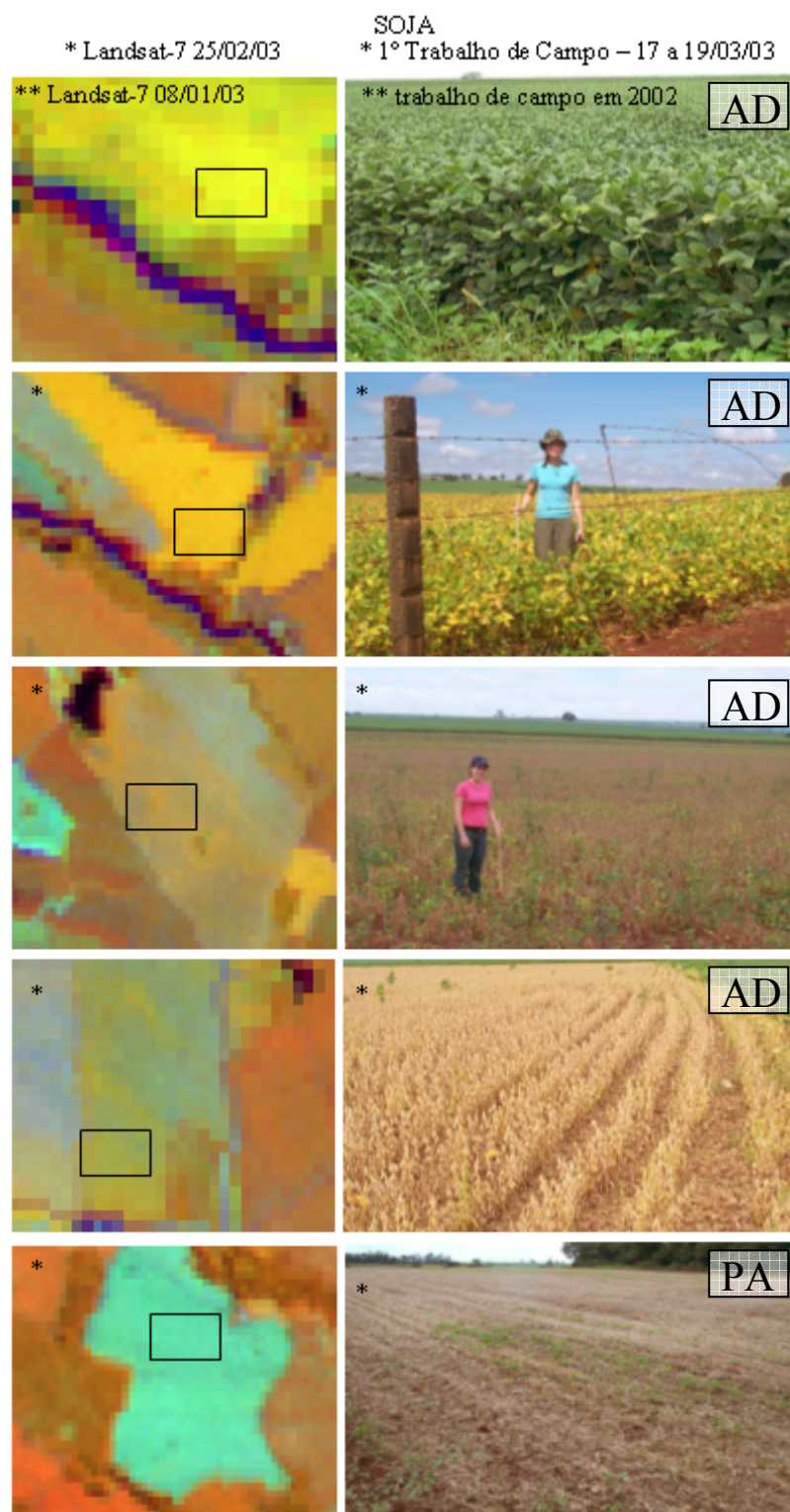


Figura 18 - Recorte de imagem Landsat na composição colorida RGB 453, e foto digital de campo de talhões de soja em diferentes estágios de crescimento. FONTE: (Sanches, 2004).

5.6. Experimentos

Neste trabalho foram executados seis experimentos para avaliar o potencial dos HMMs para a identificação de tipos de culturas agrícolas (e estágios fenológicos):

- Experimento 1: Conjunto ótimo de atributos espectrais.
- Experimento 2: Desempenho do método para identificar as culturas agrícolas.
- Experimento 3: Desempenho do método para identificar os estágios fenológicos.
- Experimento 4: Avaliação da influência do comprimento das seqüências sendo utilizadas no desempenho do método.
- Experimento 5: Determinação de quais estágios fenológicos melhor caracterizam cada cultura.
- Experimento 6: Comparação do método multitemporal baseado em HMMs com uma abordagem monotemporal de máxima verossimilhança.

Estes experimentos estão apresentados nas seções subseqüentes.

5.6.1. Experimento 1: Conjunto ótimo de atributos espectrais

Para os experimentos existia um total de sete atributos (seis bandas espectrais e NDVI) para a classificação do método baseado em HMMs. Um primeiro experimento foi conduzido para determinar o conjunto de n atributos que resultam no melhor desempenho da classificação para n variando de 1 a 7.

Os grupos de treinamento e teste foram divididos de acordo com o que foi mencionado no final da Seção 5.5.2.

O conjunto ótimo foi determinado através de uma busca exaustiva. A Tabela 2 mostra os resultados. O atributo isolado que resultou no melhor desempenho foi o NDVI, com uma taxa de acerto médio de aproximadamente 84%. Acrescentando os demais seis atributos espectrais ao NDVI, houve um aumento no desempenho de aproximadamente 7%. É interessante destacar que o maior

desempenho mostrado na Tabela 2 foi alcançado quando se utilizam seis atributos (somente os valores das bandas espectrais).

Tabela 2 - Conjunto ótimo de n atributos espectrais e o desempenho da classificação correspondente.

Número de atributos espectrais	Conjunto ótimo de atributos espectrais	Taxa de acerto médio (%)
1	NDVI	84,47
2	4, NDVI	86,51
3	4, 5, 7	89,92
4	2, 3, 5, NDVI	89,95
5	1, 2, 4, 5, 7	91,89
6	1, 2, 3, 4, 5, 7	92,72
7	Todos	91,22

5.6.2.

Experimento 2: Desempenho do método para identificar as culturas agrícolas

Este experimento visava determinar o desempenho do método baseado em HMM na identificação das culturas agrícolas nas datas das seqüências de teste. Só foram utilizadas seqüências completas (seqüências com todas as datas disponíveis, desde a primeira data, para uma dada cultura em um dado segmento, até a última; isso para todas as culturas e todos os segmentos) neste experimento, não importando o comprimento delas. Além disso, só foram utilizados seis atributos espectrais (bandas 1, 2, 3, 4, 5 e 7) para descrever os símbolos observados, já que este foi o conjunto ótimo de atributos encontrado no experimento anterior. Os grupos de treinamento e teste foram os mesmos do experimento 1.

Tabela 3 - Taxas de acerto para a classificação de coberturas vegetais.

Taxa de acerto por cobertura vegetal	
Coberturas vegetais	Taxas (%)
Soja (SJ)	95
Milho (MI)	83
Cana de açúcar (CA)	96
Pastagem (PS)	92
Mata-galeria (GA)	91
Taxa de acerto por cobertura vegetal global:	94
Taxa de acerto por cobertura vegetal médio:	91

Tabela 4 - Matriz de confusão para a classificação de coberturas vegetais.

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	95	0	4	1	0
MI	3	25	1	1	0
CA	3	1	191	4	0
PS	0	0	1	23	1
GA	0	0	2	1	29

As Tabelas 3 e 4 mostram as taxas de acerto e a matriz de confusão para a classificação de coberturas vegetais, respectivamente. A Tabela 3 mostra que o método consegue identificar bem todas as coberturas vegetais, com uma taxa média de acerto de 91%. A Tabela 4 ratifica esta conclusão uma vez que a matriz de confusão mostra valores muito baixos ou nulos fora da diagonal principal.

Ao observar a matriz de confusão, é fácil identificar os casos nos quais o método mais errou. No caso da soja, por exemplo, percebe-se que na maioria das vezes em que ocorreu erro de classificação ela foi confundida com a cana de açúcar. Já o milho foi classificado erroneamente, em pouquíssimos casos, como soja e cana de açúcar. A cana de açúcar foi mais confundida com a pastagem do que com as outras coberturas vegetais. A pastagem, por sua vez, foi confundida com a cana de açúcar e com a mata-galeria e esta última foi classificada como cana de açúcar em duas ocasiões e como pastagem uma única vez.

5.6.3.

Experimento 3: Desempenho do método para identificar os estágios fenológicos

Este experimento teve como objetivo determinar o desempenho do método baseado em HMM na identificação dos estágios fenológicos nas datas das seqüências de teste. Assim como nos experimentos 1 e 2, só foram utilizadas aqui seqüências completas, não importando o comprimento delas.

De acordo com o que foi mencionado no experimento 2, só foram utilizados seis atributos espectrais (bandas 1, 2, 3, 4, 5 e 7) para descrever os símbolos observados e seqüências corretamente identificadas. Não houve alterações nos grupos de treinamento e teste em relação ao experimento anterior.

As Tabelas 5 e 6 mostram os resultados encontrados. Mesmo a identificação do estágio fenológico não sendo o objetivo primário deste trabalho, os resultados

mostraram um bom desempenho do método para esse fim. Com exceção do estágio de crescimento, todos os outros foram bem classificados de uma forma geral. A Tabela 7 mostra que o experimento atingiu um acerto médio de 82%.

Tabela 5 - Taxas de acerto para a classificação de estados.

Estados	Taxas (%)
Pré-plantio (PP)	79
Fase de crescimento (CR)	55
Fase adulta (AD)	95
Palhada (PA)	97
Taxa de acerto por estado global:	88
Taxa de acerto por estado médio:	82

Tabela 6 - Matriz de confusão para a classificação de estados.

Matriz de confusão (estados)				
	PP	CR	AD	PA
PP	430	37	34	41
CR	54	220	119	7
AD	15	69	1838	4
PA	2	1	5	230

A maior parte dos erros ocorreu na identificação da fase de crescimento. Isso pode ser explicado pela evolução temporal das culturas ao longo de seus ciclos fenológicos. Durante os estágios de pré-plantio, fase adulta e palhada (pós-colheita), não ocorrem mudanças significativas nas respostas espectrais das culturas. Por outro lado, a resposta espectral da fase de crescimento muda continuamente, variando desde o pré-plantio até a fase adulta. Dessa forma, sua resposta espectral pode estar próxima da resposta destes dois estágios, ou adotar valores intermediários, dependendo do momento em que a imagem é adquirida, o que pode gerar erro de classificação. A Tabela 6 confirma essa interpretação, já que a matriz de confusão mostra que a fase de crescimento é confundida com frequência com a fase adulta e o pré-plantio. Essa variação ao longo do tempo não pode ser bem modelada por uma simples distribuição Gaussiana dos vetores contendo as respostas espectrais. Essa característica sugere que a fase de crescimento poderia ser melhor descrita a partir das derivadas temporais ao invés dos valores absolutos das respostas espectrais. Esta alternativa não foi, contudo, investigada neste trabalho.

5.6.4.

Experimento 4: Avaliação da influência do comprimento das seqüências sendo utilizadas no desempenho do método

Outro aspecto importante para a avaliação do método é o comprimento das seqüências de imagens consideradas. Este experimento teve como objetivo avaliar a influência do comprimento das seqüências sobre o desempenho do método proposto.

Foram consideradas seqüências de tamanhos variados, indo de 2 até 12 (comprimento máximo possível) tanto para o treinamento quanto para o teste. Em todos os casos, as seqüências de testes foram selecionadas observando a condição de que a cobertura vegetal do segmento analisado seja a mesma em todas as datas da seqüência. Ressalte-se que para este experimento foram novamente utilizados apenas os atributos espectrais 1, 2, 3, 4, 5 e 7, pois, como foi mostrado na Tabela 2, este é o conjunto ótimo de atributos que proporciona o melhor desempenho do método em questão

A Tabela 7 mostra o resultado deste experimento.

Tabela 7 - Taxas de acerto para seqüências de tamanhos variados.

Taxa de acerto por classe (%)						Taxa de acerto global	Taxa de acerto médio
Comprimento das seqüências	Coberturas vegetais						
	SJ	MI	CA	PS	GA		
2	84	76	79	69	72	78	76
3	88	83	86	76	74	84	81
4	90	88	92	77	75	88	84
5	91	90	95	78	74	91	86
6	90	94	97	82	74	92	87
7	91	88	98	85	74	93	87
8	92	100	99	85	76	93	90
9	92	X	99	85	77	93	88
10	91	X	99	87	80	94	89
11	93	X	99	88	84	94	91
12	92	X	100	90	89	96	93

Na área de estudo, em todos os talhões onde havia o plantio do milho, este era alternado com o plantio da soja. Como só estão disponíveis 12 imagens, não existem seqüências de milho de comprimento maior do que 9, o que explica o

símbolo X em algumas posições da tabela acima. Isso não acontece para a soja, cana de açúcar, pastagem e mata-galeria.

Os resultados da Tabela 7 mostram que quanto maior é a sequência, maior é a taxa de acerto global. Isto se dá devido ao fato de que, quanto mais longa a sequência, mais informação haverá no que basear a classificação, melhorando assim o seu desempenho.

5.6.5.

Experimento 5: Determinação dos estágios fenológicos que melhor caracterizam cada cultura

Um outro experimento foi conduzido para determinar quais estágios fenológicos melhor caracterizam cada tipo de cultura. Em outras palavras procurou-se determinar quais estágios fenológicos geram as maiores taxas de acerto para a classificação das culturas.

Neste experimento foram utilizados os mesmos atributos espectrais (bandas 1, 2, 3, 4, 5 e 7) para descrever os símbolos observados que nos experimentos anteriores.

As Tabelas 8, 9 e 10 mostram esses resultados. O símbolo ✓ indica quais os estados estão obrigatoriamente presentes nas sequências de teste. Assim, a primeira linha da Tabela 10 indica que em 94% das sequências de soja em que o estágio de pré-plantio (PP) esteve presente a cultura foi reconhecida corretamente. De modo semelhante, em 97% das sequências de soja em que tanto o estágio de pré-plantio (PP) quanto o estágio de palhada (PA) estão presentes (5ª linha da Tabela 8) o método identificou corretamente a cultura. Nessas tabelas, estão destacadas em amarelo as linhas correspondentes às combinações de estados que geraram as maiores taxas de acerto por cultura. É preciso ressaltar que o número de sequências disponíveis, quando se fazem as restrições de estados obrigatoriamente presentes, varia de uma cultura para a outra. O milho, por exemplo, é a cultura com menos sequências disponíveis. Este dado encontra-se na última coluna das Tabelas 8, 9 e 10.

Para a soja (Tabela 8) percebe-se que o estado mais característico é a palhada. Este estado está presente em todas as combinações de estados que geraram as taxas de acerto mais elevadas. Isto, de certa forma, pode ser explicado

pelo fato da quantidade de palhada na fase de pós-colheita ser fortemente influenciada pela cultura que estava plantada na região.

Depois da palhada, os estados mais característicos são a fase de crescimento e a fase adulta. A fase de crescimento apesar de enfrentar os problemas discutidos na Seção 5.6.2 é bem variada entre as culturas, uma vez que cada uma delas possui um ciclo fenológico distinto e passa pela etapa de crescimento de forma diferenciada. A fase adulta por sua vez, como também foi discutido na Seção 5.6.2, é um estado mais estável e característico de cada cultura.

O pré-plantio, estado menos característico, é na verdade o solo exposto, não variando muito de uma cultura para a outra, trazendo portanto pouca contribuição para identificar a cultura.

Para a soja, nota-se também que entre as combinações de estados obrigatoriamente presentes nas seqüências que geraram 100% de acerto está a seqüência com todos os estados presentes e apenas o CR, AD e PA. De um modo geral neste caso, com mais estados presentes na seqüência analisada, o método comporta-se melhor. É preciso destacar que na 10ª e 14ª linhas, apesar da taxa de acerto ter sido de 100%, o número de seqüências usadas para o teste era muito baixo o que torna as contribuições desses dados menos significativas.

Tabela 8 - Taxas de acerto para seqüências de soja com certos estados (indicados por ✓) obrigatoriamente presentes.

Estados Presentes				Taxa de acerto: SJ (%)	Número de seqüências disponíveis
PP	CR	AD	PA		
✓				94	85
	✓			94	34
		✓		96	93
			✓	98	44
✓			✓	97	30
✓	✓			93	28
✓		✓		95	79
	✓		✓	100	15
	✓	✓		96	27
		✓	✓	98	42
✓	✓		✓	100	9
✓		✓	✓	97	29
✓	✓	✓		95	22
	✓	✓	✓	100	13
✓	✓	✓	✓	100	8

Para o milho, a Tabela 9 mostra que os estágios mais característicos são a fase adulta e a palhada. Isto já era esperado em vista do que foi discutido na seção anterior, porém, as maiores taxas de acerto foram alcançadas quando esses estados estavam obrigatoriamente presentes nas seqüências de teste separadamente. Isto pode ser entendido, considerando a base de dados disponível para este trabalho (Apêndice II). Efetivamente, há poucas seqüências de milho disponíveis, sendo ainda em menor número as seqüências com combinações de três ou mais estados. Dessa forma, por exemplo, ao forçar que todos os estados estejam presentes reduzem-se substancialmente as amostras disponíveis para teste, como mostra a última coluna da Tabela 9, o que condiciona os resultados do método. Este fato verifica-se também com a cana de açúcar, conforme exposto na Tabela 10.

O milho é fortemente afetado pelo problema discutido na Seção 4.2, ou seja, algumas imagens consecutivas no banco de dados estão separadas por vários meses.

Tabela 9 - Taxas de acerto para seqüências de milho com certos estados (indicados por ✓) obrigatoriamente presentes.

Estados Presentes				Taxa de acerto: MI (%)	Número de seqüências disponíveis
PP	CR	AD	PA		
✓				77	13
	✓			75	16
		✓		88	25
			✓	87	23
✓			✓	80	10
✓	✓			71	7
✓		✓		82	11
	✓		✓	83	12
	✓	✓		82	11
		✓	✓	86	21
✓	✓		✓	80	5
✓		✓	✓	78	9
✓	✓	✓		80	5
	✓	✓	✓	80	10
✓	✓	✓	✓	75	4

Para a cana de açúcar não houve muita variação entre as taxas de acerto, como mostra a Tabela 10, sendo a diferença entre o valor mínimo e máximo igual a 5%. Apesar disso, percebe-se que nas linhas destacadas em amarelo, onde as

taxas de acerto são maiores, os estados de CR, AD e PA são igualmente distintivos desta cultura. Vale destacar que para essas combinações de estados a taxa de acerto variou apenas em um ponto percentual.

Tabela 10 - Taxas de acerto para seqüências de cana de açúcar com certos estados (indicados por ✓) obrigatoriamente presentes.

Estados Presentes				Taxa de acerto: CA(%)	Número de seqüências disponíveis
PP	CR	AD	PA		
✓				95	149
	✓			96	150
		✓		96	195
			✓	96	74
✓			✓	92	39
✓	✓			94	105
✓		✓		94	145
	✓		✓	97	59
	✓	✓		96	146
		✓	✓	96	73
✓	✓		✓	92	26
✓		✓	✓	92	38
✓	✓	✓		94	101
	✓	✓	✓	97	58
✓	✓	✓	✓	92	25

Para a pastagem e mata-galeria, este experimento não acrescenta nenhuma informação à análise do método uma vez que essas culturas se encontram em um único estágio fenológico, a fase adulta.

5.6.6.

Experimento 6: Comparação do método multitemporal baseado em HMMs com uma abordagem monotemporal de máxima verossimilhança

Um experimento final foi conduzido para fins de comparação. O objetivo foi comparar o método multitemporal proposto com uma abordagem monotemporal.

Pode-se demonstrar que, para seqüências de imagens de uma única data, o HMM concebido neste trabalho se torna um banco de classificadores de máxima verossimilhança, onde cada tipo de cobertura vegetal é representado por tantas classes quantos forem os estágios fenológicos.

A Tabela 11 mostra os resultados deste experimento. Quando foram consideradas amostras em qualquer estágio fenológico, a taxa de acerto médio foi de aproximadamente 70%.

O experimento foi repetido três vezes, retirando amostras nos estágios PP, PA e CR respectivamente. O melhor resultado obtido para esses experimentos foi de uma taxa de acerto médio de 74%.

Relembrando o resultado da Tabela 3, onde a taxa de acerto médio foi de 91%, fica clara a superioridade de uma abordagem multitemporal sobre a monotemporal.

Tabela 11 - Desempenho da abordagem monotemporal.

Estágios fenológico presentes na amostra	Taxa de acerto médio (%)
PP, CR, AD, PA	70
CR, AD, PA	69
CR, AD	71
AD	74

6 Conclusão

Este trabalho avaliou o potencial dos Modelos de Markov Ocultos para a classificação de culturas agrícolas. A avaliação experimental, baseada em uma sequência de 12 imagens Landsat para cinco tipos de coberturas vegetais, comprovou a eficiência do método para este tipo de aplicação, uma vez que ele obteve uma taxa de acerto médio de 91%

O estudo demonstrou que o melhor desempenho do método foi alcançado quando se utilizam seis atributos (somente os valores das bandas espectrais), e que o NDVI é o atributo espectral que, isoladamente, tem o maior poder de discriminação.

A abordagem por HMM também obteve bons resultados na identificação dos estágios fenológicos. Uma exceção foi o estágio de crescimento, que com frequência foi confundido com o solo exposto (fase de pré-plantio) e com a fase adulta. Essa observação sugere que os vetores de símbolos utilizados para caracterizar a fase de crescimento deveriam levar em conta não só os valores espectrais absolutos, mas também a sua variação ao longo do tempo.

Verificou-se também que o comprimento da sequência temporal avaliada influencia o resultado final da classificação. Quanto maior é a sequência, maior é a taxa de acerto global alcançada pelo método.

A presença de certos estágios fenológicos como fase adulta e a palhada nas sequências sendo avaliadas garante um melhor desempenho da classificação, uma vez que estes variam significativamente de uma cultura para outra, tornando mais fácil o processo de diferenciá-las.

Os resultados indicaram uma superioridade notável do método baseado em HMM em relação a uma abordagem monotemporal de classificadores de máxima verossimilhança.

Para este trabalho, apenas sequências de dados associadas a um único tipo de cultura foram consideradas. Está planejada, para trabalhos futuros, a análise do

comportamento do método considerando seqüências contendo amostras de mais de um tipo de cultura.

Em trabalhos futuros deseja-se também descrever a fase de crescimento a partir das derivadas temporais ao invés dos valores absolutos das respostas espectrais, uma vez que estas respostas mudam continuamente, variando desde o pré-plantio até a fase adulta e esta característica não pode ser bem modelada por uma simples distribuição Gaussiana dos vetores contendo as respostas espectrais, como se fez nos experimentos realizados.

Pretende-se também efetuar uma identificação automática do início das seqüências, que, por exemplo, pode ser marcado pelo estado de pré-plantio (solo exposto), que não varia muito de uma cultura para outra, tornando o método ainda mais robusto e menos dependente de informações disponíveis na base de dados sobre os estágios fenológicos de cada cultura.

Deseja-se também comparar o método proposto neste trabalho com outras abordagens multitemporais para a classificação de culturas agrícolas.

AARNIO, T. **Speech Recognition with Hidden Markov Models in Visual Communication**, Turku, Finlândia, Dissertação de Mestrado, Universidade de Turku, Abril 1999, 73 pp.

AHLA, D.E.; GOWERA, S.T.; BURROWSA, S.N.; SHABANOV, N.V.; MYNENIB, R.B.; KNYAZIKHIN, Y.; Monitoring spring canopy phenology of a deciduous broadleaf forest using MODIS, **Remote Sensing of Environment**, Volume 104, n. 1, September 2006, pp. 88-95.

AURDAL, L.; HUSEBY, R.B.; EIKVIL, L.; SOLBERG, R.; VIKHAMAR, D.; SOLBERG, A.; Use of Hidden Markov Models and phenology for multitemporal satellite image classification: Applications to mountain vegetation classification, **Proceedings of the Third International Workshop on the Analysis of Multi-Temporal Remote Sensing Images 2005**, no. 1469877, 2005, pp. 220-224.

BAUM, L.E.; PETRIE, T.; Statistical inference for probabilistic functions of finite state Markov chains, **The Annals of Mathematical Statistics**, Volume 37, 1966, pp 1554-1563.

BIRNEY, E.; Hidden Markov models in biological sequence analysis, **IBM Journal of Research and Development**, Volume 45, n. 3/4, May/July 2001, pp. 449-454.

BOYS, R. J.; HENDERSON, D.A.; WILKINSON, D.J.; Detecting homogeneous segments in DNA sequences by using hidden Markov models, **Journal of the Royal Statistical Society. Series C: Applied Statistics**, Volume 49, n. 2, 2000, pp. 269-285.

BROVELLI, M.A.; CRESPI, M.; FRATARCANGELI, F.; GIANNONE, F.; REALINI, E., Accuracy assessment of High Resolution Satellite Imagery by Leave-one-out method. **Proceedings of Accuracy 2006 - 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences**, Lisboa, Portugal, 5 - 7 July, Instituto Geografico Portugues, pp. 533-542, 2006 (ISBN 972-8867-27-1).

BUNKLE, H.; CAELLI, T.; Hidden Markov Models – applications in computer vision, **World Scientific**, Volume 45, 2001.

CAMPOS, V. DE O.; FEITOSA, R.Q.; MOTA, L.A.M.; PACHECO, M.A.C.; COUTINHO, H.L. DA C.; Um método para modelagem de conhecimento multitemporal no processo de classificação automática de imagens de sensores remotos, **XII Simpósio Brasileiro de Sensoriamento Remoto**, Goiânia, Brasil. Anais. São José dos Campos: INPE, 16-21 abril 2005, p. 4021-4028.,

CHAVEZ JR., P.S.; An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. **Remote Sensing of Environment**, V.24, n.9, p. 459-479, Set. 1988.

CLELAND, E.E.; CHUINE, I.; MENZEL, A.; MOONEY, H.A.; SCHWARTZ, M.D.; Shifting plant phenology in response to global change, **Trends in Ecology & Evolution**, Volume 22, n. 7, July 2007, pp. 357-365

CONROY, J.M.; BECKER, R.L.; JR. LEFKOWITZ, W.; CHRISTOPHER, K.L.; SURANA, R.B.; O'LEARY, T.; O'LEARY, D.P.; KOLDA, T.G.; Hidden Markov models for chromosome identification, **Proceedings of the 14th IEEE Symposium on Computer-Based Medical Systems**, Bethesda, MD, USA, 2001, pp. 473-477.

DE FINA, A.L.; RAVELO, A.C.; Fenologia, **Climatologia y Fenologia Agrícolas**, Buenos Aires, EUDEBA, 1973, p.201-209.

EPIPHANIO, R.D.V. **Avaliação da potencialidade das imagens MODIS na estimação da área de soja no estado do Mato Grosso**. 2007-05-07. 102 p. (INPE-15190-TDI/1299). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2007. Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m17@80/2007/07.18.12.27>>. Acesso em: 16 jun. 2008.

FORMAGGIO, A.R.; MOURA, V.; EPIPHANIO, J.C.N.; RICARDO SILVA, H. R.; FIORIO, P.R.; CAMPOS, R.C. Dados TM-Landsat na estimativa de áreas destinadas a culturas de verão, no estado de São Paulo. In: **Simpósio Brasileiro de sensoriamento remoto**, 11., 2003, Belo Horizonte. Anais... São José dos Campos: INPE, 2003. p. 93-100. CD-ROM. ISBN 85-17-00017-X. Disponível em: <<http://urlib.net/ltid.inpe.br/sbsr/2002/11.07.16.58>>. Acesso em: 26 jun. 2008.

GLERIANI, J.M.; **Redes neurais artificiais para classificação espectro-temporal de culturas agrícolas**. 2004-08-30. 212 p. (INPE-12908-TDI/1011). Tese (Doutorado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2004. Disponível em: <<http://urlib.net/sid.inpe.br/jeferson/2004/10.06.14.19>>. Acesso em: 16 jun. 2008.

GROWE, S.; **Wissensbasierte Interpretation multitemporaler Luftbilder**, Hannover, Alemanha, 2001, Tese de Doutorado, Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung - Universidade de Hannover.

GÜRTLER, S.; **Estimativa da área agrícola a partir de sensoriamento remoto e banco de pixels amostrais**. São José dos Campos, Brasil, 2003, Dissertação de Mestrado, INPE. 179 p. Disponível em: <sid.inpe.br/jeferson/2003/06.02.07.29> Acesso em: junho, 2003. (INPE-9774_TDI/858).

HUANG, S.; AHMADI, M.; SID-AHMED, M.A.; A hidden Markov model-based character extraction method, **The Journal of the Pattern Recognition society**, 9, 2008, p. 2890-2900

IPPOLITI-RAMILO, G.A.; EPIPHANIO, J.C.N.; SHIMABUKURO, Y.E. Utilização de sensoriamento remoto na previsão de área a ser plantada com culturas de verão em três municípios do Estado de São Paulo, **IX Simpósio Brasileiro de sensoriamento remoto**, São Paulo, 1999.

Instituto de Botânica - Jardim botânico de São Paulo - Ibot - **Termos e definições**. Disponível em <<http://www.ibot.sp.gov.br>>. Acesso em: jul. 2008.

International Centre for Trade and Sustainable Development - ICTSD - Disponível em < <http://ictsd.net/i/news/12431/>>. Acesso em: jul. 2008.

JOHNSON, R.A.; WICHERN, D.W.; **Applied Multivariate Statistical Analysis**, 4. ed. New Jersey: Prentice Hall. 816p., 1998. ISBN:0-13-834194-X

JUANG, B.H.; RABINER, L.R.; Hidden Markov Models for Speech Recognition; **Technometrics**, Volume 33, No. 3., August 1991, pp. 251-272.

LILLESAND, T.M.; KIEFER, R.W.; **Remote Sensing and Image Interpretation**, 3. ed. New York: John Wiley & Sons, 750 p., 1994. ISBN: 0-47-115227-7

LUIZ, A.J.B.; GÜRTLER, S.; GLERIANI, J.M.; EPIPHANIO, J.C.N.; CAMPOS, R.C.; Reflectância a partir do número digital de imagens ETM. **Simpósio Brasileiro de Sensoriamento Remoto**, 11., Belo Horizonte. Anais. São José dos Campos: INPE, 2003. p. 2071-2078. CD-ROM.

MASKEY, S.; HIRSCHBERG, J.; **Summarizing Speech Without Text Using Hidden Markov Models**, Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL, New York, June 2006, pp. 89-92.

MATHER, P.M.; **Computer processing of remotely-sensed images: an introduction**. 3.ed. Chichester: John Wiley & Sons. 352 p., 1993, ISBN: 0-47-198550-3

Ministério da Agricultura, Pecuária e Abastecimento – MAPA. **Estatísticas**. Disponível em: < <http://www.agricultura.gov.br/> >. Acesso em:jun. 2008.

MOTA, G.L.; FEITOSA, R.Q.; COUTINHO, H.L.C.; LIEDTKE, C.E.; MÜLLER, S.; PAKZAD, K.; MEIRELLES, M.S.P.; Multitemporal fuzzy classification model based on class transition possibilities. **ISPRS Journal of Photogrammetry and Remote Sensing**, 62, p.186-200, 2007.

NOGUEIRAS, A.; MORENO, A.; BONAFANTE, A.; MARINO, J.; Speech Emotion Recognition Using Hidden Markov Models, **Eurospeech 2001**, Poster Proceedings, Scandinavia, 2001, pp. 2679-2682.

OTSUKA, T.; OHYA, J.; Recognizing multiple persons' facial expressions using HMM based on automatic extraction of significant frames from image sequences, **Proceedings of the IEEE International Conference on Image Processing**, Volume 2, October 1997, pp. 546 - 549.

PAKZAD, K.; **Wissensbasierte Interpretation von Vegetationsflächen aus multitemporalen Fernerkundungsdaten**. Hannover, Alemanha, 2001. Tese de Doutorado, Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung -Universidade de Hannover.

PAKZAD, K.; MOTA, G.L.A.; MEIRELLES, M.; COUTINHO, H.; FEITOSA, R.Q.; Automatic Interpretation of Vegetation Areas in Brazil. In: Joint workshop of ISPRS working groups I/2, I/5, IC WG II/IV and the EARSeL special interest group 3D remote sensing “High resolution mapping from space 2002“, 2003. **Proceedings of the Joint Workshop of ISPRS Working Groups I/2, I/5, IC WG II/IV and the EARSeL Special Interest Group 3D Remote Sensing “High Resolution Mapping from Space 2002”**, Hannover, 2003.

RABINER, L.R.; A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. **Proceedings of the IEEE**, Volume 77, n.2, February 1989, pp. 257-286.

REED, B.C.; BROWN, J.F.; Trend Analysis of Time-series Phenology Derived from Satellite Data, **2005 International Workshop on the Analysis of Multi-Temporal Remote Sensing Images**, May 2005, pp. 166- 168

RICHARDS, J.A. **Remote sensing digital image analysis: an introduction**. 3. ed. Berlin:Springer-Verlag, 340 p, 1995.

ROUSE, J.W.; HAAS, R.H.; SCHELL, J.A. and DEERING, D.W. Monitoring vegetation systems in the great plains with ERTS, **Third ERTS Symposium**, NASA SP-351, vol. 1, pp.309-317., 1973.

SANCHES, I.D. **Sensoriamento remoto para o levantamento espectro-temporal e estimativa de área de culturas agrícolas**. 2004-03-19. 172 p. (INPE-10290-TDI/909). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2004. Disponível em: <<http://urlib.net/sid.inpe.br/jeferson/2004/05.14.10.59>>. Acesso em: 16 jun. 2008.

SAKAMOTO, T.; YOKOZAWA, M.; TORITANI, H.; SHIBAYAMA, M.; ISHITSUKA, N.; OHNO, H.; A crop phenology detection method using time-series MODIS data, **Remote sensing of environment**, Volume 96, n. 3-4, 2005, pp. 366-374.

SHAO, Y.; FAN, X.; LIU, H.; XIAO, J.; ROSS, S.; BRISCO, B.; BROWN, R.; STAPLES, G.; Rice monitoring and production estimation using multitemporal RADARSAT; **Remote Sensing of Environment**, Volume 76, Issue 3, June 2001, Pages 310-325.

TÖRMÄ, M.; RANKINEN, K.; HÄRMÄ, P.; Using Phenological Information Derived from MODIS-data to Aid Nutrient Modeling, **IEEE 2007 International Geoscience and Remote Sensing Symposium**, Barcelona, July 2007, pp. 2298-2301.

VIOVY, N.; SAINT, G.; Hidden Markov models applied to vegetation dynamics analysis using satellite remote sensing, **IEEE Transactions on Geoscience and Remote Sensing**, Volume 32 n. 4, 1994, pp. 906-917.

ZHANG, X.; HODGES, J. C. F.; SCHAAF, C. B.; FRIEDL, M. A.; STRAHLER, A. H.; GAO, F.; Global Vegetation Phenology from AVHRR and MODIS Data, **IEEE 2001 International Geoscience and Remote Sensing Symposium**, Sydney, Australia, 2001, pp. 2262-2264.

ZHOU, W.; TROY, A.; GROVE, MORGAN; Object-based Land Cover Classification and Change Analysis in the Baltimore Metropolitan Area Using Multitemporal High Resolution Remote Sensing Data; **Sensors** **2008**, 8, p. 1613-1636.

Apêndice I

Lista de símbolos:

PP – Pré-plantio

PA – Palhada

CA-CR – Cana de açúcar na fase de crescimento

CA-AD – Cana de açúcar na fase adulta

SJ-CR – Soja na fase de crescimento

SJ-AD – Soja na fase adulta

MI-CR – Milho na fase de crescimento

MI-AD – Milho na fase adulta

PS-AD – Pastagem na fase adulta

GA-AD – Mata-galeria na fase adulta

MA-AD – Mata na fase adulta

NV – Nuvem

UR – Região Urbana

FE – Feijão

TR – Trigo

ML – Milheto

AG – Corpo d'água

AE – Aeroporto

PO – Pousio

Apêndice II

Tabela com os dados sobre os segmentos de referência:

Segmento	02/09/02	20/10/02	08/01/03	09/02/03	25/02/03	14/04/03	30/04/03	16/05/03	27/07/03	12/08/03	15/10/03	19/01/04
1	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	PP	PP	PP	PP	SJ-AD
2	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-AD
3	PP	PP	SJ-AD	SJ-AD	PA	MI-CR	MI-AD	MI-AD	PA	PA	PP	SJ-CR
4	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	SJ-AD
5	PA	PA	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	XX	PA	PP	CA-CR
6	PP	PP	MI-CR	MI-AD	PA	PA	PA	PA	PP	PP	PP	SJ-AD
7	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP
8	CA-AD	PA	PP	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PP
9	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
10	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-CR	CA-CR
11	PO	PO	PO	PO	PO	NV	PO	PO	PO	PO	PO	PO
12	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PP	CA-CR
13	PP	PP	SJ-AD	SJ-AD	PP	MI-CR	MI-AD	MI-AD	PA	PP	PP	NV
14	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	NV
15	PP	PP	PP	PP	PP	PP	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD
16	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP
17	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
18	PA	PA	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR	CA-CR
19	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO
20	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
21	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO
22	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	NV
23	PA	PA	SJ-AD	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
24	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO
25	PP	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
26	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	PP	CA-CR
27	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	NV
28	PA	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
29	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-CR
30	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	NV
31	PA	PP	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	PA	PA	PP	NV
32	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
33	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
34	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PP	CA-CR
35	PA	PA	PP	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	PA	PA	PP	SJ-AD
36	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
37	CA-AD	CA-AD	PP	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
38	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
39	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
40	PA	PP	PP	PP	CA-CR	CA-CR	CA-AD	CA-AD	PP	PP	PP	SJ-AD
41	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PP	PP	CA-AD
42	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-CR	CA-AD

Segmento	02/09/02	20/10/02	08/01/03	09/02/03	25/02/03	14/04/03	30/04/03	16/05/03	27/07/03	12/08/03	15/10/03	19/01/04
43	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-CR	CA-CR	CA-AD
44	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PP	CA-CR	CA-CR	CA-AD
45	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
46	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	SJ-CR	SJ-CR	SJ-AD	PA	SJ-AD
47	PP	PP	SJ-AD	SJ-AD	SJ-AD	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
48	PP	PP	PP	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	PP	SJ-AD
49	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	NV
50	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	SJ-CR	SJ-AD	PP	PP	PP	NV
51	UR	UR	UR	UR	UR	UR	UR	UR	UR	UR	UR	UR
52	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PP	CA-AD
53	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-AD
54	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PP	NV
55	PA	PP	PP	PP	PP	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	PP
56	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO
57	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PP	CA-CR
58	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-CR	CA-CR
59	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
60	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-CR	CA-CR	CA-CR	CA-AD
61	PP	PP	CA-CR	CA-AD	CA-AD	PA	PP	PP	PP	PP	PP	SJ-AD
62	CA-AD	CA-AD	PP	PP	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
63	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO	PO
64	PA	PP	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-AD
65	PP	PP	SJ-AD	SJ-AD	PP	PP	MI-CR	MI-CR	PP	PP	PP	SJ-CR
66	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
67	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	PP	PP	PP	PP	SJ-AD
68	PP	PP	SJ-CR	SJ-CR	SJ-CR	PA	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD
69	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
70	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	PP	PP	PP	PP	SJ-AD
71	PP	PP	MI-AD	PA	PA	PA	PA	PA	ML	ML	PA	SJ-CR
72	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	PP	CA-CR	CA-AD	CA-AD	NV
73	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	NV
74	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-AD
75	PP	PP	PP	PP	PP	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD
76	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
77	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PP	NV
78	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	PP	SJ-AD
79	PP	PP	SJ-CR	SJ-AD	SJ-AD	PA	PP	PP	CA-CR	CA-AD	CA-AD	NV
80	PP	PP	SJ-AD	SJ-AD	SJ-AD	PA	PP	PP	PP	PP	PP	SJ-AD
81	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
82	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
83	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PP	CA-AD
84	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
85	PP	PP	SJ-AD	SJ-AD	PA	MI-CR	MI-AD	MI-AD	PA	PA	PP	SJ-AD
86	AE	AE	AE	AE	AE	AE	AE	AE	AE	AE	AE	AE
87	CA-CR	CA-AD	PA	PA	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-AD
88	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
89	PA	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	PA	CA-CR	CA-CR	CA-AD	CA-AD
90	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PA	CA-CR
91	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX
92	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	PP	CA-AD
93	PA	PA	PA	PA	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD

Segmento	02/09/02	20/10/02	08/01/03	09/02/03	25/02/03	14/04/03	30/04/03	16/05/03	27/07/03	12/08/03	15/10/03	19/01/04
94	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-CR	CA-AD
95	PA	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	PA	PP	CA-CR	CA-CR	CA-CR
96	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	SJ-CR	PP	PP	PP	SJ-AD
97	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
98	PP	PP	SJ-AD	SJ-AD	PO	PO	PP	PP	PP	PP	PP	SJ-CR
99	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-CR	NV
100	PA	PP	SJ-CR	SJ-AD	SJ-AD	NV	PP	PP	CA-CR	CA-CR	CA-AD	CA-AD
101	PA	PP	PP	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
102	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
103	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
104	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	PP	MI-AD	MI-AD	PP	MI-CR
105	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
106	PP	PP	SJ-AD	SJ-AD	PA	PP	MI-CR	MI-AD	PA	PA	PP	PP
107	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PP	CA-CR	CA-CR	CA-AD
108	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-CR	CA-AD	PP	CA-CR	CA-AD
109	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	NV	CA-AD	CA-AD	PP	PP	CA-CR	NV
110	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-CR	CA-AD	CA-AD
111	PP	PP	SJ-AD	SJ-AD	SJ-AD	PA	PA	PP	PP	PP	PP	SJ-CR
112	PP	PP	PP	PP	PP	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
113	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-CR	CA-AD
114	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
115	PP	PP	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	PP	NV
116	PP	PP	CA-CR	CA-AD	PP	PP	PP	PP	PP	PP	PP	SJ-AD
117	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	NV	CA-AD	CA-AD	PA	PA	CA-CR
118	PP	PP	SJ-AD	SJ-AD	SJ-AD	SR-CR	SR-CR	SR-AD	SR-AD	PA	PP	SJ-CR
119	PP	PP	CA-CR	CA-CR	CA-CR	NV	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-CR
120	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	SJ-AD
121	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
122	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-CR	CA-AD
123	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
124	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
125	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD
126	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-AD
127	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD	CA-AD
128	PA	PA	PP	PP	PP	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
129	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-CR	CA-CR
130	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
131	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	PP	CA-AD	CA-AD	CA-AD	NV
132	CA-AD	CA-AD	PA	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
133	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	MI-CR	MI-CR	PA	PA	PP	SJ-CR
134	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
135	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR	CA-AD
136	PA	PA	SJ-CR	SJ-CR	PA	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
137	PP	PP	PP	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	NV
138	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PP	PP
139	PP	PP	SJ-AD	SJ-AD	SJ-AD	PA	PP	PP	PP	PP	PP	SJ-AD
140	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	NV	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-CR
141	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-CR
142	CA-AD	PA	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	PA	SJ-PC
143	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX
144	PA	PA	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	PP	CA-AD

Segmento	02/09/02	20/10/02	08/01/03	09/02/03	25/02/03	14/04/03	30/04/03	16/05/03	27/07/03	12/08/03	15/10/03	19/01/04
145	PP	PP	SJ-AD	SJ-AD	PP	PP	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD
146	PA	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	NV
147	PA	PP	SJ-AD	SJ-AD	SJ-AD	PP	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD
148	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-CR	CA-AD
149	CA-CR	CA-CR	PA	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	PA	CA-CR
150	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD
151	PA	PP	SJ-AD	SJ-AD	PA	SJ-CR	SJ-CR	SJ-AD	PA	PA	PA	SJ-AD
152	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
153	CA-AD	PA	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
154	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
155	PA	PA	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	PA	PP	SJ-AD
156	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
157	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-CR	CA-AD
158	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
159	PA	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	PA	PP	CA-CR	CA-CR	CA-AD
160	CA-AD	CA-AD	CA-AD	PA	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
161	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	NV	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD
162	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX
163	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
164	PA	PA	SJ-AD	SJ-AD	SJ-AD	PP	PP	PP	PP	PP	PP	SJ-AD
165	CA-CR	CA-CR	PP	PP	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
166	PP	PP	SJ-AD	PP	PP	MI-AD	MI-AD	MI-AD	PA	PA	PP	SJ-AD
167	PP	PP	CA-AD	CA-AD	CA-AD	NV	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
168	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
169	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-AD	CA-AD
170	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	NV	PP	PP	PP	PP	CA-CR	NV
171	PP	PP	SJ-AD	SJ-AD	SJ-AD	PA	PA	PP	PP	PP	PP	SJ-AD
172	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
173	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
174	CA-AD	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	NV
175	CA-AD	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
176	PA	PA	SJ-CR	SJ-AD	PA	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
177	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
178	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PA	CA-AD
179	PA	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	PP	CA-AD
180	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-AD	NV
181	PA	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR	CA-CR	CA-CR	CA-AD
182	PP	PP	SJ-AD	PA	PA	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
183	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG	AG
184	CA-AD	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-AD	CA-AD
185	PA	PP	SJ-AD	SJ-AD	PA	MI-CR	MI-CR	MI-AD	PA	PA	PP	CA-AD
186	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
187	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX
188	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
189	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-CR	CA-CR	CA-AD
190	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	PP	CA-AD	CA-AD	CA-AD	NV
191	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
192	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
193	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
194	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	NV	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
195	PA	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR	CA-CR	CA-AD	CA-AD

Segmento	02/09/02	20/10/02	08/01/03	09/02/03	25/02/03	14/04/03	30/04/03	16/05/03	27/07/03	12/08/03	15/10/03	19/01/04
196	PP	PP	SJ-CR	SJ-AD	PA	MI-AD	MI-AD	PP	PP	PP	PP	SJ-AD
197	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
198	PA	PA	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR	NV
199	CA-CR	CA-CR	CA-AD	CA-AD	PA	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-CR	CA-AD
200	PP	PP	CA-AD	CA-AD	CA-AD	NV	CA-AD	CA-AD	PA	PA	PP	SJ-AD
201	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
202	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
203	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
204	PP	PP	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	PA	PA	PA	PA
205	UR	UR	UR	UR	UR	UR	UR	UR	UR	UR	UR	UR
206	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-AD	CA-AD
207	PA	PA	SJ-AD	SJ-AD	PA	SJ-CR	SJ-AD	SJ-AD	PA	PA	PA	SJ-AD
208	PP	PP	SJ-AD	SJ-AD	PA	PP	PP	PP	PP	PP	PP	SJ-AD
209	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
210	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR
211	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR	CA-AD
212	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	NV
213	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
214	PP	PP	SJ-AD	SJ-AD	SJ-AD	NV	MI-CR	MI-CR	PA	PA	PA	SJ-AD
215	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
216	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-AD	CA-AD
217	MI-AD	PP	SJ-AD	SJ-AD	PA	MI-AD	MI-AD	MI-AD	PA	PA	PP	SJ-AD
218	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
219	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	NV
220	PP	SJ-CR	SJ-CR	SJ-AD	SJ-AD	PP	PP	PP	PP	PP	MI-AD	PA
221	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-CR	CA-AD
222	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD
223	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PA	CA-CR
224	PA	PA	SJ-AD	PA	PP	MI-AD	MI-AD	MI-AD	PA	PP	PP	SJ-AD
225	PP	PP	SJ-AD	SJ-AD	SJ-AD	PP	MI-CR	MI-AD	PA	PA	PP	SJ-AD
226	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-CR
227	PA	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	PA	PA	NV
228	PA	PA	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-AD
229	PA	PA	SJ-AD	PA	PA	MI-AD	MI-AD	MI-AD	PA	PA	MI-AD	SJ-AD
230	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
231	PA	PP	SJ-AD	SJ-AD	SJ-AD	MI-CR	MI-CR	PP	PP	PP	PP	SJ-AD
232	PA	PA	SJ-AD	SJ-AD	PA	SJ-AD	SJ-AD	SJ-AD	PA	PA	PA	SJ-AD
233	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
234	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
235	PA	PA	PA	PA	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
236	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
237	MI-AD	PA	SJ-AD	PA	PA	MI-AD	MI-AD	MI-AD	PA	PA	PA	SJ-AD
238	PA	PA	SJ-CR	SJ-AD	SJ-AD	PA	PA	PA	PA	PA	SJ-CR	SJ-AD
239	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD	MA-AD
240	PP	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
241	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
242	PP	PP	SJ-CR	SJ-AD	SJ-AD	PA	PP	PP	CA-CR	CA-AD	CA-AD	NV
243	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
244	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PA	CA-CR	NV
245	PA	PA	PA	CA-CR	CA-AD	NV	CA-AD	CA-AD	PP	PP	PP	CA-AD
246	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR	CA-AD

Segmento	02/09/02	20/10/02	08/01/03	09/02/03	25/02/03	14/04/03	30/04/03	16/05/03	27/07/03	12/08/03	15/10/03	19/01/04
247	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
248	PA	PA	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR	CA-AD
249	PP	PP	PP	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	PP	PP	SJ-AD
250	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
251	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR
252	PP	PP	SJ-AD	SJ-AD	SJ-AD	PA	PP	PP	PP	PP	PP	SJ-AD
253	PP	PP	SJ-AD	PA	PA	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
254	CA-AD	PP	PP	PP	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
255	PA	PA	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	NV
256	CA-CR	CA-CR	PP	PP	PP	PP	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
257	PP	PP	SJ-CR	SJ-AD	SJ-AD	PA	PA	PA	PA	PA	PA	PA
258	MI-CR	MI-AD	PA	PA	PA	FE	PA	PA	TR	TR	PA	SJ-AD
259	PA	PA	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	NV
260	PP	PP	SJ-CR	SJ-AD	SJ-AD	PA	PP	PP	CA-AD	CA-AD	CA-AD	NV
261	PP	PP	SJ-AD	SJ-AD	SJ-AD	PA	MI-CR	MI-CR	PP	PP	PP	SJ-AD
262	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	NV	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
263	PA	PA	SJ-AD	SJ-AD	PA	PA	PA	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
264	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	NV
265	PP	PP	PP	SJ-CR	SJ-AD	MI-CR	MI-CR	MI-AD	MI-AD	MI-AD	PA	SJ-CR
266	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	NV	CA-AD	CA-AD	PA	CA-CR	CA-CR	NV
267	PA	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	PA	PA	CA-CR	CA-CR
268	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
269	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	CA-AD
270	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	PP	SJ-AD
271	PA	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PA	CA-CR	CA-AD
272	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	NV
273	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-CR	CA-CR	CA-CR	CA-AD
274	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	NV	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
275	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	NV
276	PA	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-AD
277	PP	PP	SJ-CR	SJ-AD	SJ-AD	PP	PP	PP	PP	CA-CR	CA-AD	NV
278	PA	PA	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	NV
279	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
280	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	NV	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
281	PP	PP	SJ-AD	PA	PA	PP	PP	PP	CA-CR	CA-CR	CA-AD	CA-AD
282	PP	PP	SJ-AD	PA	PA	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
283	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
284	PA	PA	SJ-AD	SJ-AD	PA	MI-AD	MI-AD	MI-AD	PA	PA	PA	SJ-AD
285	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	NV
286	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD
287	CA-AD	PA	CA-CR	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	PA	SJ-CR
288	PA	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-AD
289	PP	PP	SJ-AD	SJ-AD	PA	NV	PA	PP	PP	PP	PP	SJ-AD
290	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PP	SJ-AD
291	PA	PA	CA-CR	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	PA	CA-CR
292	PA	PA	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	PA	PP	CA-CR	CA-CR	NV
293	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
294	PP	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	PP	CA-CR	CA-CR
295	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
296	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX
297	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD

Segmento	02/09/02	20/10/02	08/01/03	09/02/03	25/02/03	14/04/03	30/04/03	16/05/03	27/07/03	12/08/03	15/10/03	19/01/04
298	PA	PA	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	NV
299	PP	PP	CA-AD	CA-AD	CA-AD	PA	PA	PA	CA-CR	CA-CR	PP	PA
300	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
301	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD	GA-AD
302	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD	PS-AD
303	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-CR	CA-CR	CA-AD	CA-AD
304	PA	SJ-CR	SJ-AD	PA	PP	MI-AD	NV	MI-AD	PA	PA	PA	NV
305	PP	PP	SJ-AD	PA	PP	CA-AD	CA-AD	CA-AD	PA	PP	PP	NV
306	PA	PA	SJ-AD	SJ-AD	SJ-AD	PA	MI-CR	MI-AD	MI-AD	MI-AD	PA	SJ-AD
307	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PA	PA	PA	NV
308	PA	PA	CA-CR	CA-CR	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	NV
309	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
310	PA	PA	SJ-AD	PA	PP	MI-AD	MI-AD	MI-AD	PA	PA	PA	SJ-AD
311	PA	PP	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	CA-AD	PP	CA-AD
312	PP	PP	SJ-AD	PA	PA	PP	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD	NV
313	PP	PP	SJ-AD	SJ-AD	SJ-AD	PP	PP	PP	PA	PA	PP	SJ-AD
314	PP	PP	SJ-AD	SJ-AD	PA	NV	PP	PP	PP	PP	PP	SJ-AD
315	PA	PA	SJ-CR	SJ-AD	PA	NV	PA	CA-CR	CA-AD	CA-AD	CA-AD	CA-AD
316	PP	PP	SJ-AD	SJ-AD	PA	NV	PP	PP	CA-AD	CA-AD	CA-AD	CA-AD

Apêndice III

Matrizes de confusão para o experimento 4 (seção 5.6.4):

Sequências de comprimento 2:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	371	24	28	15	5
MI	9	89	6	5	8
CA	114	58	1395	93	97
PS	14	12	19	171	31
GA	4	13	56	19	233

Sequências de comprimento 6:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	118	3	10	0	0
MI	1	17	0	0	0
CA	4	2	943	16	7
PS	4	0	8	120	15
GA	0	0	47	5	145

Sequências de comprimento 3:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	304	15	16	7	2
MI	7	73	3	1	4
CA	47	38	1346	63	64
PS	8	4	20	168	22
GA	2	4	56	14	217

Sequências de comprimento 7:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	94	0	9	0	0
MI	1	7	0	0	0
CA	1	0	778	11	5
PS	2	0	7	105	10
GA	0	0	36	7	124

Sequências de comprimento 4:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	234	6	12	5	2
MI	4	56	3	1	0
CA	13	21	1255	37	36
PS	5	1	19	151	21
GA	1	1	55	9	195

Sequências de comprimento 8:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	76	0	7	0	0
MI	0	2	0	0	0
CA	0	0	625	7	1
PS	1	0	6	88	8
GA	0	0	26	7	106

Sequências de comprimento 5:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	171	4	11	1	0
MI	3	36	0	1	0
CA	8	8	1112	21	17
PS	6	1	12	135	18
GA	0	0	53	6	170

Sequências de comprimento 9:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	58	0	5	0	0
MI	0	0	0	0	0
CA	0	0	474	5	1
PS	1	0	5	70	6
GA	0	0	20	5	86

Sequências de comprimento 10:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	41	0	4	0	0
MI	0	0	0	0	0
CA	0	0	330	4	0
PS	0	0	4	53	4
GA	0	0	14	3	66

Sequências de comprimento 11:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	25	0	2	0	0
MI	0	0	0	0	0
CA	0	0	199	2	0
PS	0	0	3	35	2
GA	0	0	6	3	46

Sequências de comprimento 12:

Matriz de confusão					
	SJ	MI	CA	PS	GA
SJ	12	0	1	0	0
MI	0	0	0	0	0
CA	0	0	88	0	0
PS	0	0	1	18	1
GA	0	0	2	1	24

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)