

Classificação Multitemporal de Imagens Utilizando Cadeias de Markov Nebulosas

Maria Clara de Oliveira Costa¹
Raul Queiroz Feitosa^{1,2}
Vanessa de Oliveira Campos¹

¹ Pontifícia Universidade Católica do Rio de Janeiro – PUC-Rio
Rua Marquês de São Vicente 225, Gávea, Rio de Janeiro, CEP: 22453-900, RJ, Brasil
{mclara, raul, vanessa}@ele.puc-rio.br

² Universidade do Estado do Rio de Janeiro – UERJ
Rua São Francisco Xavier 524, Maracanã, Rio de Janeiro, CEP: 20550-900, RJ, Brasil

Abstract. This work presents a multitemporal classification method for remotely sensed images. The method involves the monotemporal classification of two images of the same geographical area obtained at two different dates. The classification result relative to the earlier image is transformed to the date of the later image by applying a Fuzzy Markov Chain reasoning and further combined with the monotemporal classification of the later image to produce a single classification result. An experimental evaluation was carried out upon a set of Landsat-5 TM images of the Rio de Janeiro State, Brazil, acquired at five dates separated by approximately four years. Results indicate that the use of multitemporal information as modeled by the proposed method brings an expressive improvement of the classification performance, when compared to a pure monotemporal approach.

Palavras-chave: multitemporal classification, knowledge based classification, genetic algorithm.

1. Introdução

Métodos de classificação “em cascata” de imagens multitemporais de sensores remotos têm sido investigados há quase três décadas (Swain (1978)). Estes métodos utilizam toda a informação contida numa seqüência de imagens, e procuram explorar a correlação temporal entre elas. Trabalhos mais recentes (Pakzad (2003), Mota (2004), Campos et. al. (2005), Campos (2005)) demonstram que o uso da informação multitemporal contribui significativamente para o aumento do desempenho da classificação automática.

Em particular, o modelo proposto em Campos et. al. (2005) e Campos (2005) utiliza como dados multitemporais, além da imagem a ser classificada, a informação quanto à “verdade de campo” da mesma área geográfica numa data anterior. Porém, na maioria das aplicações práticas, esta informação não está disponível. O presente trabalho visa contornar esta limitação e apresenta uma generalização daquele modelo, em que a informação sobre o passado pode ser meramente a imagem da mesma área numa data anterior. O método descrito em Campos et. al. (2005) e Campos (2005) é, portanto, um caso particular da presente proposta. O modelo aqui apresentado utiliza conceitos da Teoria dos Conjuntos Nebulosos, (Zadeh (1965)) e de Cadeias de Markov Nebulosas (Avrachenkov (2002)).

Outra importante contribuição do presente trabalho é um novo método para estimativa das possibilidades de transição de classes. O método usa os valores relativos a um intervalo para estimar as possibilidades de transição em intervalos mais largos de tempo, usando as propriedades das Cadeias de Markov Nebulosas.

Este trabalho está organizado da seguinte maneira: a seção 2 apresenta o modelo proposto. Na seção 3, são descritos os experimentos de avaliação do modelo e seus resultados. Por fim, são apresentadas as conclusões.

2. Método de Classificação

2.1. Formulação do Problema

O problema de classificação multitemporal pode ser formulado da seguinte maneira. Seja t o instante de aquisição da imagem a ser classificada e $t-\Delta t$ um instante anterior. Sejam \mathbf{I}_t e $\mathbf{I}_{t-\Delta t}$ duas imagens de uma mesma área geográfica obtidas respectivamente nos instantes t e $t-\Delta t$. Denota-se com o símbolo $\Omega = \{\omega_1, \dots, \omega_n\}$ o conjunto das n classes da legenda. O classificador multitemporal identifica, para cada objeto da imagem \mathbf{I}_t , a classe a que o objeto pertence no instante t baseando-se nos valores dos atributos do objeto em ambas as imagens. A saída do classificador tem a forma de um vetor unitário $\mathbf{C}_t = [0, \dots, 1, \dots, 0]^T$, que contém “1” na i -ésima componente e “0” nas demais, quando a classe identificada é ω_i .

2.2. Descrição Geral do Método

Os termos monotemporal e multitemporal são usados daqui em diante para designar classificadores cujas entradas são respectivamente valores de atributos dos objetos em uma única ou em mais de uma data. O método multitemporal proposto consiste essencialmente em combinar os resultados de dois classificadores monotemporais de duas datas distintas, produzindo um resultado final único, conforme descrito a seguir.

A **Figura 1** ilustra o procedimento de classificação. Um classificador monotemporal produz, para cada objeto da imagem \mathbf{I}_t , um vetor $\mathbf{A}_t = [\alpha_1, \dots, \alpha_n]^T$, contendo as pertinências a cada uma das classes da legenda. Um segundo classificador monotemporal produz o vetor de pertinências $\mathbf{B}_{t-\Delta t} = [\beta_1, \dots, \beta_n]^T$ para o objeto correspondente na imagem e $\mathbf{I}_{t-\Delta t}$. Não se exige que estes classificadores sejam iguais. Estes classificadores monotemporais podem ter como entrada qualquer tipo de atributos.

Uma transformação temporal produz o vetor $\mathbf{T}_t = [\tau_1, \dots, \tau_n]^T$, que expressa as pertinências do objeto no instante t baseando-se na classificação $\mathbf{B}_{t-\Delta t}$ do instante anterior. As duas classificações relativas ao instante t são combinadas por uma função f que faz a fusão dos resultados \mathbf{A}_t e \mathbf{T}_t e produz o vetor de pertinência $\mathbf{M}_t = [\mu_1, \dots, \mu_n]^T$. O resultado final da classificação é produzido por um *defuzzificador*, e tem a forma de um vetor \mathbf{C}_t que possui a i -ésima componente igual a “1” se $\mu_i > \mu_j$ para todo $i, j = 1, \dots, n$, e $i \neq j$, e todas as demais componentes iguais a “0”.

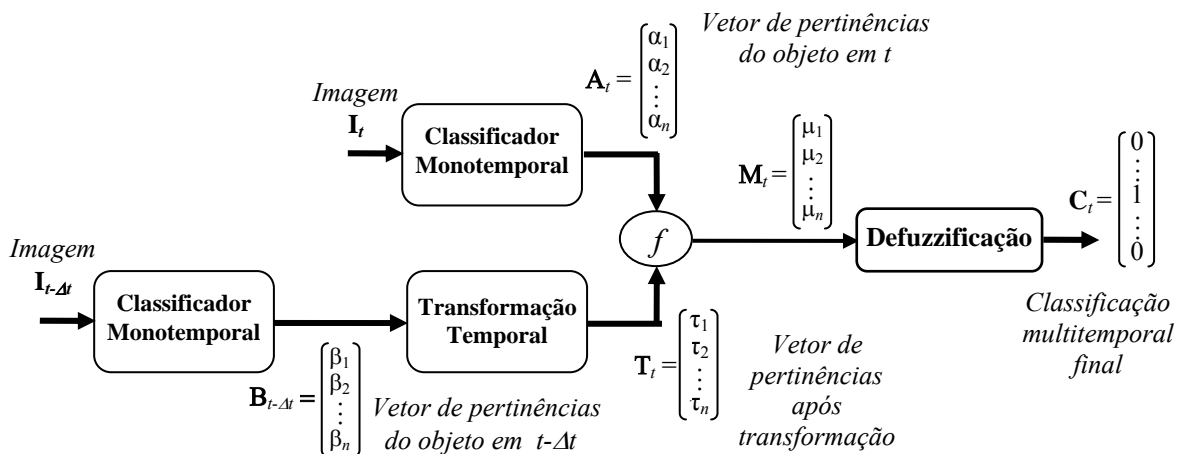


Figura 1 – Esquema geral do classificador multitemporal proposto.

2.3. Transformação Temporal

Neste trabalho, a transformação temporal baseia-se na possibilidade de um objeto pertencer a uma determinada classe num instante t , considerando a classe à qual o objeto pertencia num instante anterior $t-\Delta t$. Esta transformação pode ser representada por um diagrama de transições de classes, como o apresentado na **Figura 2**.

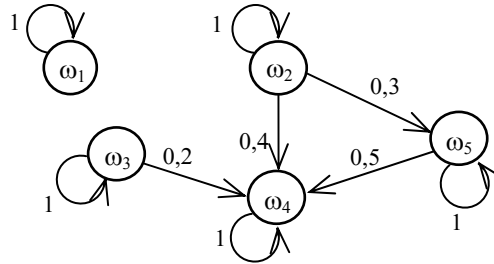


Figura 2 – Exemplo de Diagrama de Transição de Classes.

Neste diagrama, os nós correspondem a classes e os arcos a transições possíveis entre classes no intervalo de tempo Δt considerado. A cada transição, é atribuído um valor p_{ij} no intervalo $[0,1]$, que denota a possibilidade de que um objeto pertencente à classe ω_i num determinado instante $t-\Delta t$ passe a pertencer à classe ω_j no instante posterior t , sendo $i, j = 1, \dots, n$, e n o número de classes da legenda.

A transformação temporal baseia-se no conceito de Cadeias de Markov Nebulosas (CMN), conforme apresentado em Avrachenkov (2002), e pode ser formulada da seguinte forma.

Seja $\Omega = \{\omega_1, \dots, \omega_n\}$ o espaço finito de classes da legenda. Define-se um conjunto de pertinências nebulosas \mathbf{B} mapeado sobre Ω no intervalo de possibilidades $[0,1]$, representado pelo vetor $\mathbf{B}_{t-\Delta t} = [\beta_{t-\Delta t,1}, \beta_{t-\Delta t,2}, \dots, \beta_{t-\Delta t,n}]^T$ em que cada $\beta_{t-\Delta t,i}$ corresponde ao grau de pertinência do objeto à classe ω_i em um instante $t-\Delta t$.

Define-se também uma relação *nebulosa* \mathbf{P} , como um conjunto nebuloso sobre o produto cartesiano $\Omega \times \Omega$, que pode ser representada por uma matriz $\mathbf{P} = \{p_{ij}\}$, de dimensão $n \times n$, chamada, neste texto, de matriz de possibilidades de transição (MPT). A linha i corresponde à classe (ω_i) a que pertence o objeto no instante $t-\Delta t$ e a coluna j à classe (ω_j) a que pertence o objeto num instante posterior t . Todos os elementos da matriz \mathbf{P} estão no intervalo $[0,1]$ e pelo menos um dos elementos em cada linha tem valor igual a “1”. A transformação temporal utilizada nesta proposta se baseia na lei de transição própria das CMN, que, para a relação *nebulosa* \mathbf{P} , toma a forma dada pela **equação 1**

$$\mathbf{T}_t^T = \mathbf{B}_{t-\Delta t}^T * \mathbf{P} = [\tau_1, \tau_2, \dots, \tau_n], \quad (1)$$

onde o operador $*$ corresponde ao máximo dos produtos. Ou seja, a transformação do vetor de pertinência $\mathbf{B}_{t-\Delta t} = [\beta_{t-\Delta t,1}, \beta_{t-\Delta t,2}, \dots, \beta_{t-\Delta t,n}]^T$ do instante $t-\Delta t$ para o instante t resulta no vetor $\mathbf{T}_t = \{\tau_k\}$, onde

$$\tau_k = \max \{ [\beta_{t-\Delta t,1} p_{1k}, \beta_{t-\Delta t,2} p_{2k}, \dots, \beta_{t-\Delta t,n} p_{nk}] \} \quad (2)$$

com $k = 1, 2, \dots, n$, sendo n o número de classes da legenda.

2.4. Fusão dos resultados

No passo chamado fusão, os dois resultados relativos ao instante t , \mathbf{A}_t e \mathbf{T}_t são agregados (Klir (1995)) num único vetor de pertinências $\mathbf{M}_t = [\mu_1, \mu_2, \dots, \mu_n]^T$. A média geométrica foi adotada como função de agregação (f) neste trabalho. Assim, o vetor de pertinências resultante da agregação é dado por:

$$\mathbf{M}_t = f(\mathbf{A}_t, \mathbf{T}_t) = [(\alpha_1 \tau_1)^{1/2}, (\alpha_1 \tau_2)^{1/2}, \dots, (\alpha_n \tau_n)^{1/2}]^T \quad (3)$$

2.5. Regra de decisão

A etapa final do método proposto consiste em transformar o vetor \mathbf{M}_t no vetor \mathbf{C}_t , que conterá o resultado final da classificação. Esta operação é chamada *defuzzificação*, que, neste trabalho, atribui o objeto à classe ω_i para a qual μ_i é máximo. Logo, o vetor \mathbf{C}_t será dado pela **Equação 4**:

$$\mathbf{C}_t = [c_{t,1}, c_{t,2}, \dots, c_{t,n}] \text{ onde } \begin{cases} c_{t,i} = 1, & \text{caso } \mu_i = \max_j \{\mu_j\} \\ c_{t,i} = 0, & \text{caso contrário} \end{cases} \quad (4)$$

2.6. Estimativa dos parâmetros do modelo

Uma questão central nesta proposta é a estimativa dos parâmetros do modelo. Isto inclui os classificadores monotemporais, e muito especialmente as possibilidades de transição p_{ij} . Para isso, será necessário reunir um conjunto de objetos da imagem para os quais se conheça a priori os atributos e a “verdade de campo” nos instantes t e $t-\Delta t$. Com base nestes dados de treinamento, determinam-se primeiro os valores dos parâmetros dos classificadores monotemporais que maximizam o desempenho de cada um isoladamente nas datas correspondentes, tendo em vista a classificação dos objetos de treinamento previamente conhecida.

Uma vez tendo sido ajustados os parâmetros dos classificadores monotemporais, estimam-se em seguida as possibilidades de transição de classes que maximizam a medida de desempenho selecionada com base na classe conhecida dos objetos de treinamento no instante t . Neste trabalho, foram usados Algoritmos Genéticos (AG) para este fim, seguindo o mesmo procedimento de treinamento descrito em Costa (2006). Cabe salientar que o modelo de classificação proposto não impõe o uso de AG nesta etapa, e admite outras técnicas de otimização para estimar as possibilidades de transição.

Adotou-se como função objetivo para o AG a *taxa de reconhecimento média por classe* definida como o percentual médio dos segmentos de cada classe corretamente classificados pelo modelo.

A matriz de possibilidades de transição \mathbf{P} possui n^2 elementos a serem estimados. Este número pode ser consideravelmente reduzido explorando algum conhecimento prévio sobre a área de teste. O fotointérprete pode indicar as transições impossíveis, cujas possibilidades serão consideradas constantes e iguais a zero. Atribui-se ainda o valor unitário à transição mais provável de cada linha, o que reduz em n o número de parâmetros a serem estimados.

O procedimento de treinamento descrito acima requer que todas as transições de classe possíveis estejam adequadamente representadas no conjunto de treinamento. Reunir tais amostras pode ser uma tarefa difícil para o analista. Em muitas aplicações há dados de outras duas datas t' e $t'-\Delta t$ ($t' < t$), também separadas por Δt . Se é razoável admitir que as

possibilidades de transição são aproximadamente estáveis entre t' e t , pode-se lançar mão destes dados para estimar a MPT.

2.7. Classificação para intervalos maiores

A alternativa apresentada na seção precedente pressupõe um mesmo intervalo de tempo, Δt , separando as duas imagens utilizadas no treinamento e as duas imagens envolvidas na classificação. Porém, por diversas razões, nem sempre se dispõe de um par de imagens de treinamento que atendam a esta condição.

Este trabalho utiliza a propriedade de convergência finita das Cadeias de Markov e propõe um método para estimar a MPT relativa a um intervalo $\lambda\Delta t$, com λ inteiro, a partir da outra MPT relativa a intervalo igual a Δt , conforme descrito a seguir.

Admite-se que a MPT é invariante no tempo, ou seja, independe das datas das imagens consideradas, desde que estas sejam separadas por Δt . A transformação temporal para um intervalo $\lambda\Delta t$ pode ser vista como a replicação do modelo básico apresentado na seção 0, λ vezes. Isto leva à **Equação 5**

$$\mathbf{T}_t = \mathbf{B}_{t-\lambda\Delta t} * \mathbf{P}^\lambda \quad (5)$$

em que a potência da matriz \mathbf{P} é definida a partir do produto de matrizes, em que a álgebra de máximo de produtos substitui a álgebra convencional de soma de produtos.

3. Avaliação Experimental

3.1 Base de dados e legenda

O modelo proposto foi avaliado tendo como área de teste o município do Rio de Janeiro, e como base de dados as imagens digitais multiespectrais obtidas pelo sensor TM do satélite LANDSAT 5, órbita 217 ponto 76. Foram imagens adquiridas em cinco datas diferentes: 30 de Maio de 1984, 18 de Janeiro de 1988, 18 de Abril de 1992, 16 de Junho de 1996 e 12 de Agosto de 1999.

Todas as cinco imagens foram co-registradas e classificadas visualmente por um foto-intérprete experiente. Uma imagem SPOT-5 e dados de SIG foram utilizados como subsídio. Essa classificação visual serviu como referência ou “padrão ouro” na análise de desempenho que se seguiu. Consiste de um mapa temático para cada data definindo a classe de cada pixel.

Nos experimentos realizados, os objetos da imagem a serem classificados foram segmentos multi-pixel e não pixels individuais. A resposta espectral média do segmento nas bandas 3, 4 e 5 foi usada para formar o vetor de atributos de cada objeto. Os segmentos foram extraídos a partir das classificações de referência, como descrito a seguir.

As classificações de cada ano foram “empilhadas” de modo a representar cada pixel por um vetor composto de cinco rótulos, cada um relativo à classe a que o pixel pertence em cada data. Os segmentos foram definidos por componentes conectados contendo pixels com o mesmo vetor de rótulos. Para reduzir o efeito de possíveis imperfeições no registro das imagens, segmentos muito pequenos e muito alongados foram descartados.

A legenda é constituída pelas seguintes sete classes: *afloramento rochoso*, *água*, *área úmida*, *área urbana*, *campo*, *floresta*, *solo exposto*. A Tabela 1 apresenta o número de objetos das imagens distribuídos por ano e por classe.

Tabela 1 – número de objetos do banco de dados utilizado nos experimentos.

ANO	CLASSE						
	<i>afloramento rochoso</i>	<i>água</i>	<i>área úmida</i>	<i>área urbana</i>	<i>campo</i>	<i>floresta</i>	<i>solo exposto</i>
1984	12	49	91	320	352	581	75
1988	12	49	83	358	337	560	81
1992	12	49	83	386	333	533	84
1996	12	49	81	463	325	494	56
1999	12	49	87	514	329	447	42

3.2. Procedimento Experimental e Resultados

Nos experimentos realizados para avaliar a presente proposta, os classificadores monotemporais que compõem o classificador multitemporal são idênticos ao classificador descrito em Mota (2004). Trata-se de um classificador paramétrico definido a partir do vetor média e da matriz de covariância para cada uma das classes do problema. Os valores destes parâmetros foram estimados aplicando o método de máxima verossimilhança sobre 10% dos objetos disponíveis no banco de dados. Os objetos não utilizados para esta estimativa foram usados na fase de teste.

Os experimentos consideraram ainda um segundo tipo de classificador monotemporal que foi aplicado apenas para a imagem da data anterior. Trata-se de um classificador ideal com 100% de acerto. Tal classificador foi simulado em nossos experimentos atribuindo aos vetores de pertinência $\mathbf{B}_{t-\Delta t}$ o vetor $\mathbf{C}_{t-\Delta t}$ formado a partir da classificação de referência para a imagem nesta data. Esta situação pode na verdade ser realizada na prática, quando a verdade de campo da data anterior é conhecida. Este caso corresponde ao modelo de classificação multitemporal proposto em Campos et. al. (2005). O propósito de considerar este tipo de classificador para a data anterior é avaliar a relevância da exatidão do classificador monotemporal sobre o desempenho global do classificador multitemporal proposto. Mesmo nestes casos, o vetor de pertinências \mathbf{A}_t relativo à data da imagem que está sendo classificada é também produzido pelo classificador descrito em Mota (2004).

Os valores de possibilidade de transição foram estimados utilizando a totalidade dos dados disponíveis em duas datas, denotadas respectivamente por t_1 e $t_1-\Delta t$. A MPT foi estimada por meio de AG de modo a maximizar a taxa de reconhecimento média por classe nos objetos relativos a t_1 usando como referência a classificação visual. O método foi avaliado classificando uma imagem obtida em t_2 ($t_2 \neq t_1$) utilizando a informação de uma imagem obtida numa data anterior $t_2-\Delta t$.

Os resultados obtidos são apresentados na **Tabela 2**. As duas colunas mais à esquerda indicam as datas dos pares de imagem utilizados para o treinamento e para o teste do modelo. Em todos os casos, o lapso de tempo separando as imagens em cada par é de aproximadamente quatro anos. As três colunas mais à direita apresentam as taxas médias de reconhecimento calculadas em cada experimento. A coluna “monotemporal” apresenta os resultados obtidos a partir da classificação monotemporal de cada imagem usando como entrada apenas os atributos da imagem em t_2 . As duas colunas mais à direita contêm o desempenho da classificação multitemporal sobre a imagem de t_2 em duas situações. A coluna indicada com “real” corresponde ao caso em que se utiliza o classificador monotemporal paramétrico descrito em Mota (2004), tanto para a imagem que está sendo classificada como para a imagem anterior. A coluna indicada com “ideal” diz respeito ao caso em que se simula um classificador ideal para a imagem passada, conforme esclarecido anteriormente.

Por serem estocásticos, os algoritmos genéticos podem resultar em diferentes valores para cada execução realizada. Os resultados de taxa de reconhecimento apresentados na **Tabela 2** são os valores médios medidos ao longo de 20 experimentos realizados para cada imagem.

Tabela 2: Taxa de reconhecimento média por classe para $\Delta t \cong 4$ anos

<i>Datas dos Conjuntos de Dados</i>				<i>Taxa de Reconhecimento (%)</i>		
Treinamento		Teste		Monotemporal	Multitemporal	
$t_1 - \Delta t$	t_1	$t_2 - \Delta t$	t_2		real	ideal
84	88	88	92	54,4	65,3	88,4
88	92	92	96	58,3	63,9	90,4
92	96	96	99	57,4	60,0	81,5

A **Tabela 2** indica que, em todos os casos considerados, o método multitemporal proposto produziu resultados melhores do que o método puramente monotemporal. Para os casos em que a imagem da data anterior foi classificada pelo classificador paramétrico usado em Mota (2004), o ganho de desempenho ficou entre 2,5% e 11%. Já para o caso em que se usou o conhecimento da classe a que pertenciam os objetos na data anterior, o método multitemporal foi entre 24% e 34%, superior ao método monotemporal. Estes resultados indicam que, em relação aos métodos monotemporais, o método multitemporal proposto pode trazer um aumento de desempenho entre moderado e excepcional, dependendo da exatidão do classificador monotemporal que compõe o classificador multitemporal.

Um outro aspecto analisado nos experimentos é a aplicação do método exposto na seção 2.7, que permite estimar a MPT de intervalos de tempo múltiplos inteiros de Δt a partir da MPT relativa a um intervalo igual a Δt . Os resultados são apresentados na **Tabela 3**.

Tabela 3: Taxa de reconhecimento média por classe para $\Delta t \geq 4$ anos.

<i>Datas dos Conjuntos de Dados</i>				<i>Taxa de Reconhecimento (%)</i>		
Treinamento		Teste		Monotemporal	Multitemporal	
$t_1 - \Delta t$	t_1	$t_2 - \lambda \Delta t$	t_2		real	ideal
92	96	96	99	57,4	60,0	81,5
88	92	92	99	57,4	59,9	78,3
84	88	88	99	57,4	59,7	75,9

Nestes experimentos, a MPT foi estimada a partir de imagens de duas datas (t_1 e $t_1 - \Delta t$) separadas por aproximadamente quatro anos. O valor assim estimado foi corrigido para intervalos maiores utilizando a **Equação 5**. Em todos os casos, a imagem classificada foi a de 1999 (t_2), em que se usou como informação anterior imagens de datas ($t_2 - \lambda \Delta t$) cada vez mais remotas, nomeadamente de 1996, 1992 e 1988. Os resultados da **Tabela 3** indicam em todos os casos um ganho de desempenho em relação à classificação monotemporal. Este ganho decai à medida que se busca a informação anterior de uma data cada vez mais distante no passado. Ainda assim, os resultados da **Tabela 3** corroboram a validade do método descrito na seção 2.7 para estimar as possibilidades de transição em intervalos mais longos a partir de valores relativos a intervalos mais curtos.

4. Comentários Finais

O presente trabalho apresentou um modelo multitemporal para classificação de imagens de sensores remotos baseado em Cadeias de Markov Nebulosas. O modelo proposto é uma generalização de um método anterior introduzido em Campos (2005), e pode operar usando

como informação do passado apenas uma imagem da mesma área geográfica numa data anterior.

O método foi avaliado experimentalmente utilizando uma seqüência de imagens LANDSAT-5 do município do Rio de Janeiro, adquiridas em datas separadas por aproximadamente quatro anos. O método multitemporal proposto foi superior à abordagem monotemporal em todos os experimentos, em alguns casos muito superior. Verificou-se que o ganho de desempenho que se pode alcançar depende primordialmente de dois fatores. O primeiro fator é a exatidão dos classificadores monotemporais que compõem o esquema multitemporal. O segundo fator é o lapso de tempo que separa as imagens multitemporais: quanto mais recente é a informação passada maior sua contribuição para a classificação.

Como trabalhos futuros, prevê-se a investigação de métodos automáticos alternativos aos algoritmos genéticos para a estimativa da matriz de transição de classes e a avaliação do desempenho do modelo de classificação para imagens de alta resolução.

Agradecimento

À CAPES e ao DAAD (programa PROBRAL), e à FAPERJ pelo apoio recebido para a realização desta pesquisa.

Referências

- Avrachenkov, K.E.; Sanchez, E. Fuzzy Markov Chains and Decision-Making, **Fuzzy Optimization and Decision Making**, v. 1, n. 2, p. 143 – 159, 2002.
- Campos, V. O.; Feitosa, R. Q.; Mota, G. L. A.; Pacheco, M. A. C.; COUTINHO, H. L. C. Um método para modelagem do conhecimento multitemporal no processo de classificação automática de imagens de sensores remotos. **Revista Brasileira de Cartografia, Brasil**, v. 57, n. 01, p. 28-35, 2005.
- Campos, V. **Um Método de Modelagem do Conhecimento Multitemporal para a Interpretação Automática de Imagens de Sensores Remotos**. 2005. 172p. (PUC-CDD:621.3). Dissertação (Mestrado em Métodos de Apoio à Decisão) - Departamento de Engenharia Elétrica, PUC-RIO, Rio de Janeiro. 2005.
- Cazes, T.B.; Feitosa, R.Q.; Mota, G.L.A. Automatic Selection of Training Samples for Multitemporal Image Classification. **Proceedings of International Conference on Image Analysis and Recognition**. 2004.
- Costa, M.C. **Um Modelo Nebuloso para Classificação Multitemporal de Imagens**. 2006. 72p. (PUC-CDD:621.3). Dissertação (Mestrado em Métodos de Apoio à Decisão) - Departamento de Engenharia Elétrica, PUC-RIO, Rio de Janeiro. 2006.
- Heipke, C. Some Requirements for Geographic Information Systems: A Photogrammetric Point of View. In: **Photogrammetric Engineering & Remote Sensing, Vol. 70, No. 2**. 2004.
- Holland J.H. **Adaptation in natural and artificial system**. Ann Arbor: The University of Michigan Press, 1975.
- Klir, G.I., Yuan, B. **Fuzzy Sets and Fuzzy Logic: Theory and Applications**. New Jersey: Prentice Hall, 1995. 342p.
- Mota, G. L. A. **Interpretação baseada em conhecimento aplicada a imagens multitemporais de satélite de baixa resolução**. Tese (Doutorado em Engenharia Elétrica), Departamento de Engenharia Elétrica, PUC-RIO, Rio de Janeiro. 2004.
- Pakzad, K.; Mota, G. L. A.; Meireles, M.; Coutinho, H.; Feitosa, R. Q. Automatic Interpretation of Vegetation Areas in Brazil. **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**, 2003.
- Swain, P.H. Bayesian classification in time-varying environment. *IEEE Trans. Syst. Man Cybern.*, vol. 8, pp. 880–883, 1978.
- Zadeh, L. A., Fuzzy Sets. *Journal of Information and Control*, v. 8, p. 338-353, 1965.