

Análise exploratória de dados multitemporais de sensoriamento remoto usando mapas auto-organizáveis: estudo em áreas de expansão da cana-de-açúcar

Ramon Morais de Freitas
Marcos Adami
Bernardo Friedrich Theodor Rudorff
Yosio Edemir Shimabukuro
Reinaldo Roberto Rosa
José Demísio Simões da Silva
Gustavo Felipe Balué Arcoverde

Instituto Nacional de Pesquisas Espaciais – INPE
Caixa Postal 515 – 12245-970 – São José dos Campos - SP, Brasil
{ramon, adami, bernardo, yosio, gustavo }@dsr.inpe.br
{demisio, reinaldo}@lac.inpe.br

Abstract. This work shows an approach to use Self Organizing Map (SOM) to feature selection in multitemporal remote sensing data. Optical images derived from Thematic Mapper bands were used for this analysis. The objective was to improve and understanding about SOM operational capabilities for feature selection in remote sensing data applied to agricultural systems. The study area is the municipality of the Barretos located in the northern part of Sao Paulo State. The test site comprises several landscape types such as citrus, annual croplands fields, pasture and sugarcane with high land cover conversion rates the last years. Using Landsat TM multitemporal data and segmentation techniques we generated 1110 samples for the SOM analysis. The visual inspection of Kohonen maps were used for feature selection. Preliminary results indicated that there is a temporal dependence between the spectral bands and the land cover observation over time. We recommended that further studies should focus on the correlation of spectral and texture data using the SOM approach.

Keywords: remote sensing, SOM, feature selection sensoriamento remoto, Mapas auto-organizáveis, seleção de atributos.

1. Introdução

O cultivo da cana-de-açúcar é uma atividade importante na economia de países como Brasil, Índia, China e África do Sul. O monitoramento e mapeamento das áreas de cana-de-açúcar através de imagens de sensoriamento remoto permite estimar a área cultivada e a produtividade; além disso analisar os impactos sócio-econômicos e ambientais particularmente nas regiões de expansão do cultivo da cana-de-açúcar. Nesse contexto, as técnicas de sensoriamento remoto, geoprocessamento e reconhecimento de padrões são fundamentais para o pré-processamento, mapeamento, classificação e análise dos dados (Abdel-Rahman e Ahmed, 2008). No Brasil, a área cultivada e a expansão do cultivo da cana-de-açúcar têm sido mapeadas anualmente no Estado de São Paulo desde a safra 2003/2004 pelo projeto Canasat (Rudorff et al., 2005). Em resumo, este projeto utiliza imagens de sensores orbitais Landsat-TM e CBERS-CCD para mapeamento das áreas de cultivo de cana-de-açúcar. O mapeamento é realizado por fotointérpretes experientes utilizando recursos de edição de imagem no software SPRING. Além de fornecer, anualmente a área cultivada com cana-de-açúcar o projeto também informa a distribuição espacial dos talhões da cultura.

Do ponto de vista econômico e ambiental o mapeamento das classes de uso e cobertura da terra nas áreas de expansão da cultura permite identificar e entender as mudanças ocorridas em uma dada região geográfica. O mapeamento dessas áreas, especificamente das classes de pastagem, citros, vegetação natural, reflorestamento e outras culturas agrícolas não é uma tarefa trivial dada a complexidade da variação da resposta espectral dessas classes. Nesse contexto, a utilização de imagens adquiridas em diferentes datas é necessária pois permite diferenciar as classes em função da variação da resposta espectral das diferentes classes ao longo do tempo. A

classificação automática das áreas de cana-de-açúcar embora reportada por alguns autores como satisfatória (Tardin et al., 1992; Narciso e Schmidt, 1999) é dependente da área de estudo. Em regiões onde ocorrem vários tipos de culturas agrícolas e áreas de pastagem, a classificação automática utilizando estatística tradicional não é conveniente devido aos elevados erros de omissão e comissão da classificação. No entanto, novas abordagens de classificação de imagens baseadas em técnicas de inteligência artificial e reconhecimento de padrões buscam melhorar os resultados das classificações (Schowengerdt, 2006). Dentre outros métodos podemos citar as redes neurais artificiais, máquinas de vetores de suporte e as redes semânticas.

A etapa de seleção de atributos é uma importante análise a ser realizada antes do emprego de qualquer tipo de classificador. A seleção de atributos pode ser vista como um processo de busca em que o algoritmo usado deve encontrar o menor subconjunto de atributos com a melhor acurácia de classificação (Jain e Zongker, 1997). Em geral, dados com alta dimensionalidade no espaço de atributos estão sujeitos ao alto custo de processamento e ao fenômeno de Hughes (Hughes, 1968) que está ligado à degradação da acurácia da classificação com o aumento da dimensionalidade dos dados, mantendo constante o número de amostras de treinamento (Haertel e Landgrebe, 1998).

A taxonomia dos algoritmos de seleção de atributos proposta por Jain e Zongker (1997) divide inicialmente os métodos em duas classes: 1) reconhecimento de padrões estatístico e 2) redes neurais artificiais. Dentro das técnicas estatísticas várias abordagens são possíveis que vão desde a aplicação de algoritmos genéticos numa abordagem estocástica e de múltiplas soluções, até métodos determinísticos de solução simples como os algoritmos de busca de sequências. Dentro das redes neurais artificiais o mapa auto organizável (SOM - Self-Organizing Map) de Kohonen (Willshaw e Malsburg, 1976; Kohonen, 1982; Kohonen, 1989) é uma técnica inovadora tanto para a seleção de atributos quanto para a análise exploratória de dados (Vesanto, 2002).

Na área de sensoriamento remoto e geoprocessamento, a técnica SOM tem sido empregada para análise de dados espacializados (Silva, 2004), classificação de imagens de sensores remotos (Veronese, 2007; Gonçalves, 2008) e classificação de segmentos agrícolas para estimativa de área (Arco-verde, 2008). Embora a maioria dos trabalhos envolvendo o SOM está focada no problema específico da classificação de imagens, este trabalho tem o objetivo de avaliar a aplicação do SOM como um método de seleção de atributos e análise exploratória de dados de sensoriamento remoto. Como estudo de caso, selecionou-se imagens multitemporais do TM utilizadas no projeto Canasat e que representam a variabilidade espectro-temporal, dos alvos agrícolas.

2. Mapa Auto-Organizável de Kohonen

As bases teóricas do SOM estão relacionadas às observações neurobiológicas da região do córtex cerebral humano, especificamente nas questões de memória associativa e auto-organização (Kohonen, 1989). Estas observações indicam que o cérebro possui regiões específicas para uma determinada atividade sensorial (por ex. visual e motora). Além disso, observou-se que dada uma determinada ativação cerebral, o grau de ativação dos neurônios diminui à medida que aumenta a distância da região de ativação inicial (Haykin, 2001). O SOM é uma Rede Neural Artificial que possui como característica principal a arquitetura bi-dimensional dos neurônios (Mapa de Kohonen). O SOM é representado por uma grade discreta 2-D que mantém as relações topológicas e geométricas dos dados de entrada.

O SOM é uma rede neural artificial baseada em aprendizado competitivo, cooperativo e não supervisionado (Kohonen, 1989). A aprendizagem competitiva é um processo adaptativo em que os neurônios da rede tornam-se gradualmente sensíveis às diferentes categorias de entrada

e aos conjuntos específicos de amostras. Organizando esses neurônios, por um dado critério, pode-se obter a redução da dimensionalidade do espaço de atributos bem como a observação da formação de agrupamentos dos dados (*clusters*) (Haykin, 1999; Vesanto, 2002). Nesse contexto, o SOM tornou-se um exemplo tradicional de rede neural artificial que utiliza o aprendizado não supervisionado e produz uma visualização do espaço de atributos em menor dimensão que o dado de entrada, preservando suas relações topológicas (Haykin, 1999).

Do ponto de vista computacional, o algoritmo do SOM pode ser dividido em três passos. O primeiro passo é a fase competitiva em que os neurônios da camada de saída competem entre si segundo alguma métrica. Em geral é utilizada a distância euclidiana. Nesse passo, é possível encontrar um único neurônio vencedor, também chamado de BMU (*Best Match Unit*). No segundo passo ocorre a fase cooperativa, em que é definida a vizinhança deste neurônio. No último passo, ocorre a fase adaptativa, em que os vetores de código do neurônio vencedor e de sua vizinhança são ajustados. A relação de vizinhança entre os neurônios é estabelecida segundo alguma função de vizinhança, geralmente um envelope gaussiano. O objetivo desta função é controlar o nível de atuação dos neurônios em torno do neurônio vencedor do processo competitivo. Conforme o modelo neurobiológico o nível de atuação dos neurônios vizinhos decai à medida que se distanciam do BMU (Haykin, 1999; Silva, 2004).

A arquitetura do SOM, quanto rede neural, pode ser dividida em duas camadas: a camada de entrada I e a de saída U , em que os neurônios estão dispostos em um arranjo topológico bi-dimensional. A entrada da rede corresponde a um vetor p -dimensional, x , geralmente no espaço \mathbb{R}^p . Todas as p componentes do vetor de entrada alimentam cada um dos neurônios do mapa. Cada neurônio i pode ser representado por um vetor de pesos sinápticos $w_i = [w_{i1}, w_{i2}, \dots, w_{ip}]^T$, também no espaço p -dimensional. No treinamento seqüencial, para cada padrão de entrada x um neurônio é escolhido o vencedor (BMU), c , usando o critério de maior similaridade, ou seja, menor distância euclidiana:

$$\|x - w_c\| = \min\|x - w_i\| \quad (1)$$

em que $\|\cdot\|$ representa a distância euclidiana. Os pesos do neurônio vencedor, juntamente com os pesos dos seus neurônios vizinhos, são ajustados de acordo com a equação:

$$w_i(t+1) = w_i(t) + h_{ci}(t)[x(t) - w_i(t)] \quad (2)$$

em que t representa a iteração do processo de treinamento, $x(t)$ é o padrão de entrada e $h_{ci}(t)$ é o núcleo de vizinhança ao redor do neurônio vencedor c . Este último termo é uma função que decresce com o tempo e com a distância do neurônio i ao neurônio vencedor c , e geralmente corresponde ao produto de duas componentes: a taxa de aprendizado $\alpha(t)$ e a função de vizinhança $h(d, t)$:

$$h_{ci}(t) = \alpha(t) \cdot h(\|r_c \cdot r_i\|, t) \quad (3)$$

em que, r_i e r_c são, respectivamente, as posições do neurônio i e do neurônio vencedor no arranjo topológico da rede. Uma característica importante do SOM, derivada da quantização vetorial gerada pelo algoritmo, é preservar a ordenação topológica e a densidade dos dados. Isso possibilita obter informações dos agrupamentos de dados originais analisando as relações geométricas dos neurônios do mapa treinado (Haykin, 1999). A Figura 1 apresenta essa relação topológica e geométrica entre o conjunto de dados de entrada e os neurônios ajustados a esse conjunto de dados.

Uma forma alternativa para o processo de aprendizagem da rede é o treinamento em Lote que é caracterizado por não possuir taxa de aprendizagem. Os vetores de entrada no mapa de

Kohonen é gerado de forma particionada conforme as regiões convexas geradas pelo diagrama Voronoi da distância euclidiana entre os pesos dos neurônios e os vetores de entrada (Vesanto et al., 2000, Silva, 2004). A cada passo do treinamento a fase de cooperação é realizada para todo o conjunto de amostras do dado de entrada. Somente após a fase de cooperação é realizada a atualização dos pesos e no final dessas duas etapas é definida uma época (Vesanto et al., 2000).

3. Materiais e Métodos

3.1. Área de Estudo

A área de estudo utilizada nesse trabalho corresponde ao município de Barretos que está localizado na região norte do Estado de São Paulo. O principal motivo para escolha deste município foi a grande expansão da área de cana-de-açúcar ocorrida durante o período analisado. Entre 2004 e 2008 a área de expansão totalizou mais de 26.000 ha o que corresponde a aproximadamente 17% da área do município (156.470 ha) e um aumento de 100 % da área de cana existente em 2004. Além disso o município possui área com outras culturas agrícolas como soja, milho, algodão, citrus, pastagens e reflorestamento. Nesse contexto, a área de estudo é ideal para este tipo de análise, pois apresenta variabilidade espectral entre os alvos de interesse.

3.2. Dados

Um conjunto de 3 imagens Landsat-TM da órbita 221/ponto 74 adquiridas em 11 de outubro de 2002 (Figura 1-a), 04 de março de 2003 (Figura 1-b) e 07 de maio de 2003 (Figura 1-c) foram utilizadas para elaboração do espaço de atributos. No total foram utilizadas 18 bandas espectrais, 6 para cada data, correspondendo respectivamente as regiões espectrais do azul (0,45 - 0,52 μm , B1), verde(0,52 - 0,60 μm , B2), vermelho (0,63 - 0,69 μm , B3), infra-vermelho próximo (0,76 - 0,90 μm , B4), infra-vermelho médio (1,55 - 1,5 μm , B5) e infra-vermelho médio (2,08 - 2,35 μm , B7). O georeferenciamento entre todas as imagens foi realizado utilizando como referência uma imagem ortoretificada disponível no sítio do *Global Land Cover Facility* (GLCF, <http://glcfapp.umiacs.umd.edu:8080/esdi/index.jsp>). Outra base de dados utilizada no trabalho foi o mapa de uso das áreas de expansão da cana-de-açúcar (Figura 1-d) realizado através da interpretação visual das imagens TM. Este mapa possui as classes temáticas de pastagem, vegetação natural, citrus, cana-de-açúcar e outras culturas agrícolas. Os *softwares* SPRING e MATLAB foram utilizados para integração dos dados e processamento das imagens. A SOM Toolbox (Vesanto, 1999) desenvolvida pela Universidade de Tecnologia de Helsinki foi utilizada para análise exploratória dos dados e elaboração dos mapas de Kohonen.

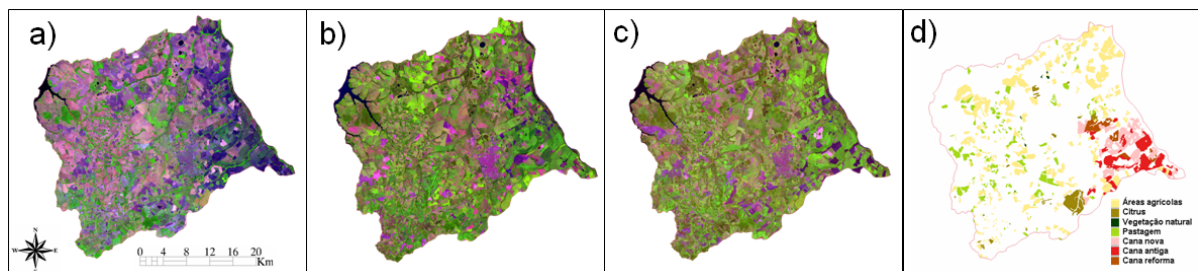


Figura 1: a) Composição 3B4G5R da imagem Landsat-TM de 11-10-2002; b) Composição 3B4G5R da imagem Landsat-TM de 04-03-2003; c) Composição 3B4G5R da imagem Landsat-TM de 07-05-2003; d) Imagem temática classificada

3.3. Métodos

O primeiro passo da metodologia consistiu na integração dos dados no software SPRING. A partir disso, foi realizado um recorte da área de estudo, representando as áreas de expansão da cultura de cana-de-açúcar. Este procedimento foi realizado mascarando as regiões fora dessa área de expansão nas 6 bandas espectrais das três imagens. A seguir foi realizada a segmentação destas 18 imagens, cujo objetivo principal foi homogeneizar as amostras utilizadas na rede SOM, para diminuir significativamente o número de amostras em relação a uma abordagem pontual ou por janelas de pixels. A segmentação de imagem é uma técnica na qual as regiões espacialmente adjacentes e com características similares podem ser agrupadas. Para realizar a segmentação no software SPRING é necessário definir dois limiares: de similaridade e de área. No presente trabalho o limiar de similaridade utilizado foi igual a 25, definido após vários testes. O limiar de área foi definido como 50, o que significa que a área mínima considerada na segmentação equivale a 45 ha no terreno. Quanto menor os valores dos limiares maior será o número de regiões geradas e mais detalhada será a segmentação (Bins et al., 1993).

Após esse procedimento, partiu-se para extração de atributos e rotulação dos polígonos. A rotulação dos polígonos foi realizada sob supervisão do foto-intérprete com auxílio das 3 imagens TM utilizadas e o mapa de uso das regiões de expansão da cana-de-açúcar (Figura 3). Para cada região gerada na segmentação foi atribuída uma única classe. As classes utilizadas nessa análise foram: área agrícola, representando as áreas de agricultura exceto cana-de-açúcar, pastagem, citrus, vegetação natural e cana-de-açúcar, em diferentes fases de desenvolvimento. Para cada polígono foi calculado o atributo de média em cada uma das 18 imagens. Além das informações espectrais também, foram calculados os atributos de área e posição x e y do centróide do polígono. No total foram gerados 1082 amostras, sendo 976 amostras para áreas de expansão e 106 amostras para áreas de cana-de-açúcar.

A partir do espaço de atributos rotulados partiu-se para análise exploratória dos dados utilizando as funcionalidades da SOM Toolbox. O primeiro passo para elaboração da rede SOM foi a normalização dos dados. Os dados espectrais, derivados das imagens, foram normalizadas entre 0 e 1 dividindo-se por 255. Já os valores de área e posição do centróide foram normalizados pelo máximo valor encontrado. Para avaliação da qualidade dos mapas gerados foram analisados os erros de quantização média e o erro topológico de cada mapa (Vesanto et al., 2000). Ressalta-se que a variabilidade do erro pode estar relacionada com o processo de normalização, uma vez que o SOM computa a distância euclidiana entre os atributos.

Após vários testes empíricos dentre as diversas configurações da rede SOM utilizadas, escolheu-se a topologia plana com grade hexagonal, função de vizinhança gaussiana e algoritmo de treinamento em lote. Esta topologia foi escolhida devido ao menor tempo de processamento e tamanho da grade de 20 x 20 neurônios. Ressalta-se que a definição de mapas de dimensões menores prejudica a interpretação dos resultados, uma vez que as classes com menor frequência não são representativas nos neurônios.

No total foram gerados 4 conjunto de mapas de Kohonen: 3 conjuntos de mapas utilizando individualmente as 6 bandas espectrais de cada data e 1 um conjunto de mapas utilizando todos os atributos, ou seja, as 18 imagens e as informações de área e posição do centróide do polígono. A partir dos mapas de Kohonen foi realizada a análise exploratória dos dados através da interpretação visual dos mapas rotulados.

4. Resultados e Discussão

A Figura 2 apresenta os mapas de Kohonen para as bandas espectrais da imagem TM de 11 de outubro de 2002. Nos planos de componentes podemos identificar as bandas que possuem maior correlação (linha tracejada em vermelho) são as bandas 1 e 2. A primeira matriz no canto

superior esquerdo é a matriz unificada U e representa a distância média euclidiana entre os neurônios e seus vizinhos. Através dessa matriz pode-se observar a formação de agrupamentos que podem ser utilizados em um eventual processo de classificação. A última matriz no canto inferior direito representa os rótulos das classes (conforme legenda da Figura 1-d). Os erros de quantização e topológicos para este mapa foram 0,02 e 0,028, respectivamente, o que indica um baixo erro na topologia dos mapas. Nos planos de componentes (mapas de Kohonen) pode-se observar que as bandas 1 e 2 estão correlacionadas indicando redundância de informação. Na matriz de rótulos, observa-se que não há formação de agrupamentos nessa matriz, o que indica a dificuldade na separação dos alvos utilizando somente os atributos espectrais desta data.

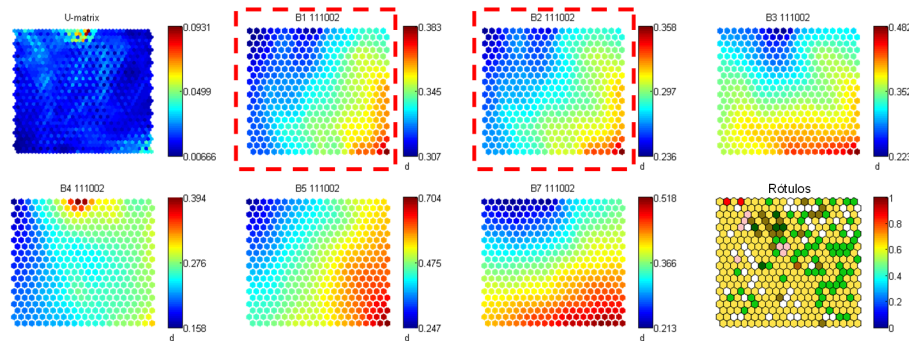


Figura 2: Mapas de Kohonen para 6 bandas espectrais de 11-10-2002

As Figuras 3 e 4 apresentam a os mapas de Kohonen para as imagens TM de 04 de março de 2003 e 07 de maio de 2003, respectivamente. Para todos os casos não há agrupamentos na matriz de rótulos, o que indica que qualquer processo de classificação que se utilize dos atributos espectrais apresentará erros de comissão e omissão acentuados, dada a dispersão dos dados. Para a data de 04 de março 2003 há correlação entre todas as bandas, exceto para o mapa da banda 4 (linha azul tracejada). O resultado corrobora com a análise visual das imagens, sendo que nessa época do ano há confusão espectral entre as culturas de milho e cana-de-açúcar. Além disso, as áreas de colheita se confundem com áreas de pastagem. Na data de 07 de maio de 2003 as correlações entre as bandas 3 e 7 (linhas vermelhas tracejadas), sendo que os valores mais altos dessas bandas indicam regiões de solo exposto, tipicamente presente nas áreas agrícolas (Figura 4).

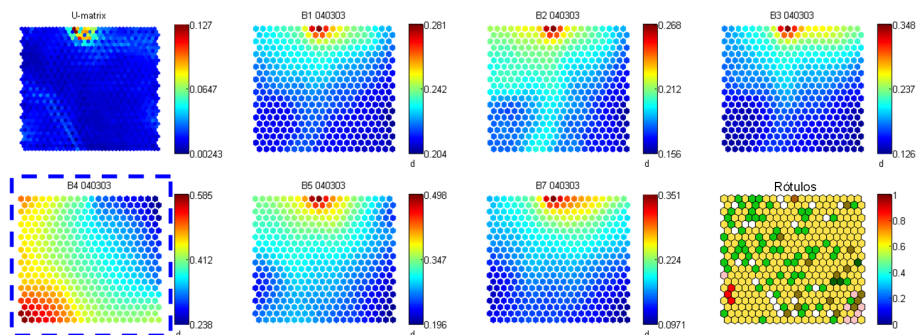


Figura 3: Mapas de Kohonen para 6 bandas espectrais de 04-03-2003

Na Figura 5 utiliza-se todos atributos espectrais, mais os atributos de área e posição (x,y) do centróide do polígono. Esses atributos adicionais permitem analisar a influência da posição geográfica dos polígonos, como é o caso das áreas de cultura de cana-de-açúcar que estão

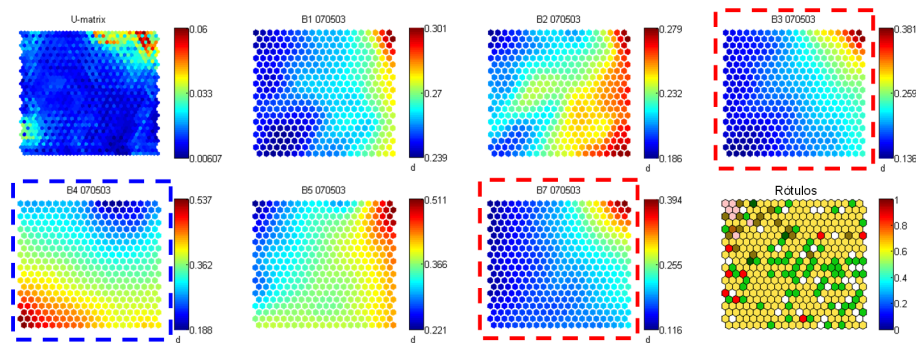


Figura 4: Mapas de Kohonen para 6 bandas espectrais de 07-05-2003

localizadas no sudeste da área de estudo. Tal informação pode ser útil na interpretação dos resultados e na análise de dependência espacial dos dados. Analisando todos os atributos, observa-se a importância da banda 4 para mapeamentos de uso e cobertura da terra, que está de-correlacionada em todas as datas (linha preta tracejada). Analisando as matrizes rotuladas e histogramas, percebemos que o SOM preserva a distribuição inicial dos dados, em que as áreas de vegetação natural, citrus e cana-de-açúcar apresentam menor frequência que as áreas de pastagem e outras áreas de agricultura.

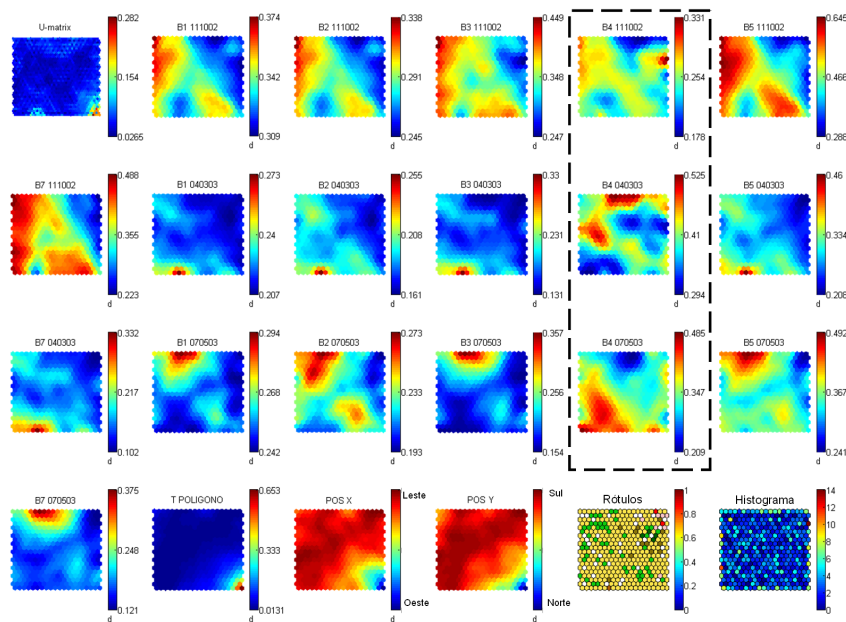


Figura 5: Mapas de Kohonen para 18 bandas espectrais, área dos polígonos e posição x e y

5. Considerações Finais

Este trabalho apresentou uma descrição geral de como os mapas auto-organizáveis de Kohonen podem ser úteis na análise exploratória de dados de sensoriamento remoto. Os resultados indicam que há bandas correlacionadas entre as datas analisadas. Quanto ao método de seleção de atributos, a análise visual dos mapas proporcionou uma visão integrada do conjunto de dados, permitindo identificar rapidamente atributos correlacionados. Recomenda-se que atributos texturais sejam inseridos para análise da variabilidade temporal. Ressalta-se no trabalho que a redução da dimensionalidade das amostras através da segmentação apresenta uma técnica alternativa para amostragem da imagem uma vez que a análise da imagem

como todo, em uma abordagem pontual, por pixels, é computacionalmente inviável, dada as dimensões das imagens e o número de atributos. No contexto do mapeamento da cana-de-açúcar e da análise da resposta espectral de diferentes alvos, o método SOM pode auxiliar na interpretação dos resultados uma vez que possui a capacidade de analisar todo o conjunto de dados de forma integrada.

Referências Bibliográficas

- Abdel-Rahman, E. M., Ahmed, F. B. The application of remote sensing techniques to sugarcane (*Saccharum spp. hybrid*) production: a review of the literature, **International Journal of Remote Sensing**, 29:13, 3753 - 3767, 2008.
- Arco-verde, G. F. B. **Estratificação espacial para estimativa da área de culturas agrícolas de verão com imagens de pré-plantio**. 2008-08-25. 179 p. (INPE-T/). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2008.
- Bins, L.S.; Fonseca, L.M.G.; Erthal, G.J. Ii, F.M. Satellite imagery segmentation: a region growing approach. **Anais...** Simpósio Brasileiro de Sensoriamento Remoto, 1996, 8, 221-223.
- Gonçalves, M. L., Netto, M. L. A., Costa, J. A. F., Zullo Junior, J., Costa, J. A. F. Classificação não-supervisionada de imagens de sensores remotos utilizando redes neurais auto-organizáveis e métodos de agrupamentos hierárquicos. **Revista Brasileira de Cartografia**, v.60, n1, 17-29. Abril 2008.
- Haertel, V. ; Landgrebe, D. On the Classification of Classes with Nearly Equal Spectral Responses in Remote Sensing Hyperspectral Image Data. **IEEE Transactions on Geoscience and Remote Sensing**, Vol. 37, No. 5, Part 2, 2374-2386, Sep. 1999.
- Haykin, S. **Neural Networks: A Comprehensive Foundation**. New York: Prentice-Hall, 2nd edition, 1999.
- Hughes, G. F. On the mean accuracy of statistical pattern recognizers, **IEEE Transactions on Information Theory**, Vol. IT-14, No. 1, January 1968.
- Jain, A. and Zongker, D. Feature Selection: Evaluation, Application, and Small Sample Performance. **IEEE Trans. Pattern Anal. Mach. Intell.** 19, 2, 1997.
- Kohonen, T. Self-organized formation of topologically correct feature maps. **Biological Cybernetics**, v. 43,59-69, 1982.
- Kohonen, T. **Self-organization and associative memory**. 3. ed. New York: Springer-Verlag, 1989.
- Narciso, G. and Schmidt, E.J. Identification and classification of sugarcane based on satellite remote sensing. **Proceedings of the South African Sugar Technologists' Association**, 73, 189-194, 1999.
- Rudorff, B. F. T. ; Berka, L. M. S. ; Moreira, M. A. ; Duarte, V. ; Xavier, A. C. ; Rosa, V. G. C. ; Shimabukuro, Y. E. Imagens de satélite no mapeamento e estimativa de área de cana-de-açúcar em São Paulo: ano safra 2003/04. **Agricultura em São Paulo**, São Paulo, v. 52, n. 1, 21-39, 2005.
- Schowengerdt R. A. **Remote Sensing: Models and Methods for Image Processing**, Academic Press, Inc., Orlando, FL, 2006
- Silva, M. A. S. **Mapas auto-organizáveis na análise exploratória de dados geoespaciais multivariados**. 2004-03-08. 117 p. (INPE-12434-TDI/996). Dissertação (Mestrado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2004.
- Tardin, A.T., Assuncao, G.V.; Soares, J.V. Preliminary analysis of TM data for coffee, citrus and sugarcane discrimination in Furnas region, MG, Brazil. **Pesquisa Agropecuária Brasileira**, 27, 1355-1361, 1992.
- Veronese, T. B. **Aplicação de mapa de Kohonen para classificação de texturas em imagens SAR em ambiente paralelo**. 2007-02-05. 110 p. (INPE-15239-TDI/1326). Dissertação (Mestrado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2007.
- Vesanto J. SOM-Based data visualization methods. **Intelligent data analysis**. v 3, n. 2, 111-126, 1999.
- Vesanto, J.; Himberg, J.; Alhoniemi, E.; Parhankangas, J. **SOM toolbox for Matlab 5**. Helsinki University of Technology, 2000.
- Vesanto, J. **Data exploration process based on the self-organizing map**. 2002. Tese em Tecnologia - Acta Polytechnica Scandinavica, Espoo, 2002.
- Willshaw, D. J.; Von Der Malsburg. How patterned neural connections can be set up by self-organization. In: **Proceedings of the Royal Society of London Series B**, v. 194, 431-445, 1976