

Classificação orientada a objeto na identificação de áreas de reforma de cana-de-açúcar.

Elizabeth Goltz¹
Carolina Moutinho Duque de Pinho¹
Bernardo Friedrich Theodor Rudorff¹
Leila Maria Garcia Fonseca¹

¹Instituto Nacional de Pesquisas Espaciais – INPE
Caixa Postal 515 - 12245-970 - São José dos Campos - SP, Brasil
{goltz, bernardo}@dsr.inpe.br; {carolina, leila}@dpi.inpe.br

Abstract- Remote sensing images and geoprocessing techniques have been used yearly, since the crop year 2003/04, to map and estimate sugarcane area in São Paulo State, Brazil. The image classification is carried out visually by an experienced team of interpreters resulting in high quality thematic maps. However, this work is time-consuming and often quite tedious due to the large area covered by sugarcane which was 4.9 million ha in 2008 for São Paulo State. The annual update of the sugarcane maps consists basically of two steps: 1) to add new areas that are coming from the sugarcane cultivation expansion; and 2) to subtract sugarcane fields that are being renewed and consequently skipped for harvest during one crop year. In order find alternatives to automate the visual classification the goal of this work was to develop an oriented-based classification methodology using hierarchical net to identify sugarcane areas that are being renewed. Landsat-5 TM images acquired at favorable dates were selected to identify the renewed sugarcane fields using the Definies Developer software to perform segmentation, sample selection, hierarchical net construction (selection and customized features) and final classification. The result showed that the multiresolution segmentation was able to recognize the renewed sugarcane areas with images acquired at three dates. Comparing the visual interpreted map to the hierarchical net map it was verified that almost all renewed areas were correctly classified.

Palavras-chave: multitemporal, agriculture, hierarchical net, mutitemporal, agricultura, rede hierárquica.

1. Introdução:

Desde o ano safra 2003/04 o cultivo da cana-de-açúcar vem sendo mapeado anualmente no Estado de São Paulo através do Projeto Canasat utilizando-se imagens de sensoriamento remoto e técnicas de geoprocessamento (Rudorff e Sugawara, 2007).

Este mapeamento é realizado cuidadosamente através de interpretação visual de imagens na tela do computador. Isto permite obter um resultado de boa qualidade, mas o trabalho de interpretação visual de imagens para grandes áreas é demorado e chega a ser tedioso. A atualização do mapa do cultivo da cana a cada ano safra consiste basicamente em duas etapas: 1) acrescentar as novas áreas de plantio de cana que serão colhidas na safra em questão (cana expansão); e 2) subtrair as áreas de reforma com cana planta de ano-e-meio que não farão parte da colheita daquele ano safra.

Uma alternativa para automatizar a identificação das áreas de cana em reforma sem perder a qualidade alcançada na interpretação visual é o uso de sistemas baseados em conhecimento em que os conhecimentos do intérprete são modelados em um ambiente computacional emulando a sua capacidade de combinar dados na análise de imagens de sensores remotos (Feitosa et al., 2005).

Ainda nesta linha de pensamento, se destaca a análise orientada a objetos, a qual descreve que a informação semântica necessária para a interpretação de uma imagem não está presente no pixel, e sim em objetos e nas relações existentes entre eles. De acordo com Pinho (2005) estes objetos podem ser segmentos oriundos de segmentação ou polígonos previamente mapeados.

Além disso, a análise orientada a objeto permite a inserção do conhecimento do intérprete e a utilização de parâmetros de cor, forma, textura e contexto na classificação de imagens.

De acordo com Fockelmann (2007), a utilização da classificação orientada a objeto (*Definiens eCognition*) apresentou resultados satisfatórios na detecção de talhões (bordas) agrícolas em relação às técnicas de classificação não orientadas a objetos.

Lucas et al. (2007) realizaram um trabalho de mapeamento de áreas agrícolas com imagens multitemporais utilizando classificação baseada em regras (*Definiens eCognition*). Como resultado, esta classificação forneceu uma aproximação viável para o mapeamento e monitoramento das culturas agrícolas, sendo passível de ser empregado operacionalmente. Goltz et al. (2008), apresentaram resultados positivos na detecção de áreas já mapeadas de reforma, utilizando rede hierárquica baseada no brilho das imagens índice de vegetação.

Com base nestas pesquisas o presente trabalho tem por objetivo desenvolver uma metodologia de classificação orientada a objeto, através de rede hierárquica para automatizar o procedimento de identificação e classificação das áreas de reforma de cana-de-açúcar com cana planta de ano-e-meio.

2. Metodologia de Trabalho

A aplicação desta metodologia foi realizada na região oeste do estado de São Paulo referente a órbita/ponto 223/75 da imagem do sensor TM a bordo do satélite Landsat. As datas das imagens foram:

- 13 de setembro de 2007 (banda 3 e 4);
- 04 de fevereiro de 2008 (banda 3, 4 e 5);
- 23 de março de 2008 (banda 3, 4 e 5).

As imagens foram adquiridas em datas bem definidas e favoráveis para a identificação das áreas de reforma. Já as bandas foram selecionadas previamente para serem utilizadas tanto no processo de segmentação quanto na classificação. Por exemplo, a banda 3 é a que melhor destaca as áreas de reforma com solo exposto, utilizadas durante a segmentação ao longo das datas; já a combinação das bandas 3 e 4 geram o NDVI (Normalized Difference Vegetation Index), o qual é utilizado no atributo de brilho para classificação destas áreas de reforma; e a banda 5 é utilizada em combinação com a banda 3 para classificar áreas com cultura de verão. Todas as imagens foram convertidas para reflectância aparente, esta etapa foi realizada no software *ENVI 4.3*, utilizando os dados de conversão encontrados em Chandler e Markham (2003) e alguns parâmetros do cabeçalho das imagens.

As imagens (*tif*) e o vetor (*shapefile*) do mapeamento da cana-de-açúcar foram importados para o software de classificação orientada a objeto (*Definiens Developer*). Neste, foram realizadas as etapas de segmentações, seleções das amostras, criação da rede hierárquica (seleção e customização de atributos) e a classificação orientada a objeto.

A primeira segmentação realizada foi do tipo *Chessboard* com o objetivo de se respeitar os limites do vetor da área mapeada. Esta segmentação foi realizada em um primeiro nível sendo que os parâmetros utilizados foram 10.000 para o fator de escala, peso 1 para o *layer* temático (vetor da cana mapeada) e peso 0 para as imagens.

Já para na primeira classificação o atributo escolhido para diferenciar a classe CANA da classe OUTROS foi o de área. Os limiares utilizados para este atributo foram baseados nos valores das amostras de cada classe.

Uma segunda segmentação do tipo multiresolução foi realizada em um segundo nível abaixo do primeiro já criado. Para isto, utilizou-se o parâmetro de 100 para a escala, agregou-se peso 1

para as todas as bandas 3 de cada data. Esta segmentação foi realizada apenas dentro da classe CANA.

A segunda classificação foi baseada neste segundo nível de segmentação. Para isso, foi criada a classe REFORMA e a classe REFORMA 2 (soja).

Durante a criação da rede hierárquica alguns atributos foram customizados, dentre eles o NDVI e a razão entre as bandas 5 e 3 para diferenciar áreas de cultura de verão como por exemplo a soja.

3. Resultados e Discussão:

A Figura 1 mostra o resultado da primeira segmentação do tipo *Chessboard*, que resultou num total de 5.454 polígonos.

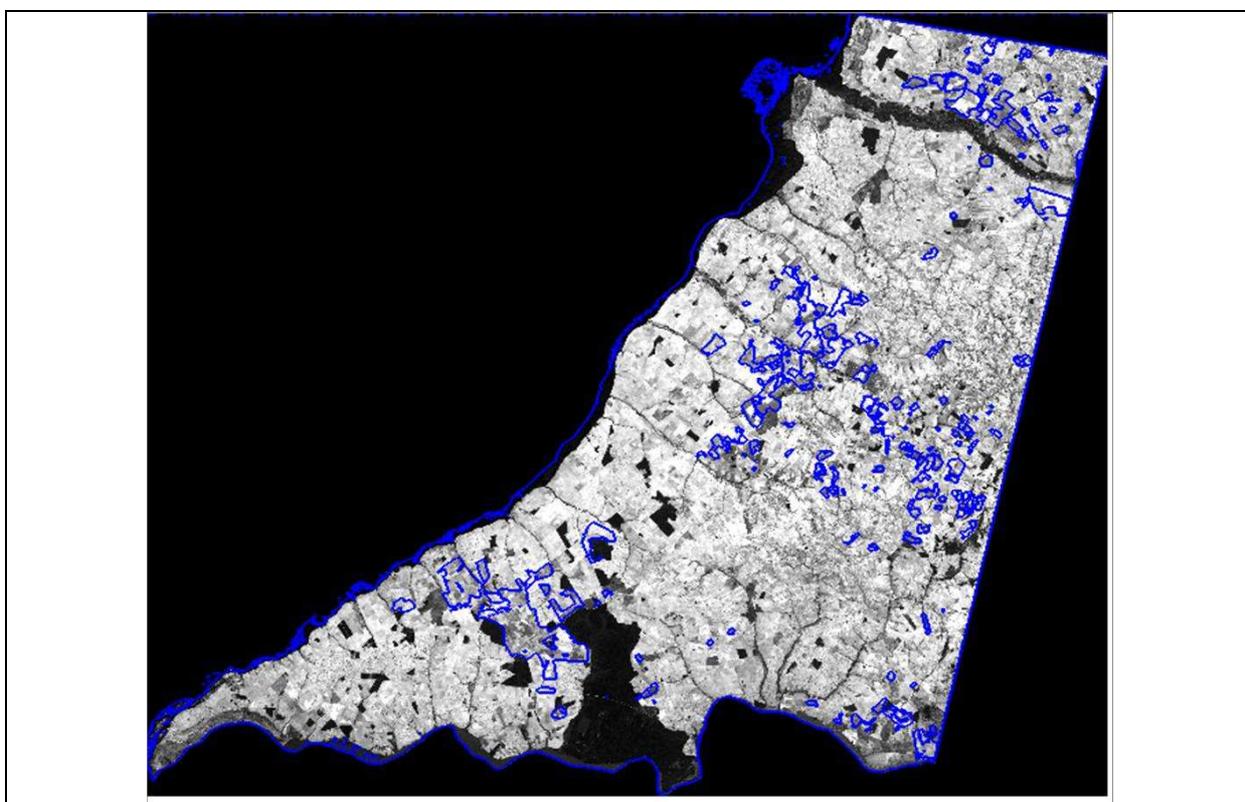


Figura 1- Resultado da segmentação *Chessboard* (1° nível).

O resultado desta segmentação foi positivo, pois ela respeitou bem os limites do vetor de entrada. Além disso, para este caso de identificação de limites já definidos, ela se mostrou bem mais rápida durante o processamento do dado do que outro tipo de segmentação.

Neste mesmo nível foi realizada uma primeira classificação para separar o que era cana-de-açúcar dos outros alvos que não eram de interesse criando-se as classes CANA e OUTROS. Algumas amostras de ambas as classes foram selecionadas a critério de observação do seu comportamento e para posterior seleção de atributos e definição de limiares para sua classificação. Esta primeira classificação foi necessária, pois, no próximo passo a segmentação multiresolução só seria realizada na classe CANA.

A segunda segmentação realizada foi a multiresolução. Esta segmentação teve o intuito de criar polígonos dentro dos polígonos já existentes da classe CANA. Desta forma, um polígono

novo foi criado para cada mudança espectral que ocorresse ao longo das datas. Como resultado esta nova segmentação originou 5715 polígonos, os quais podem ser vistos na Figura 2.

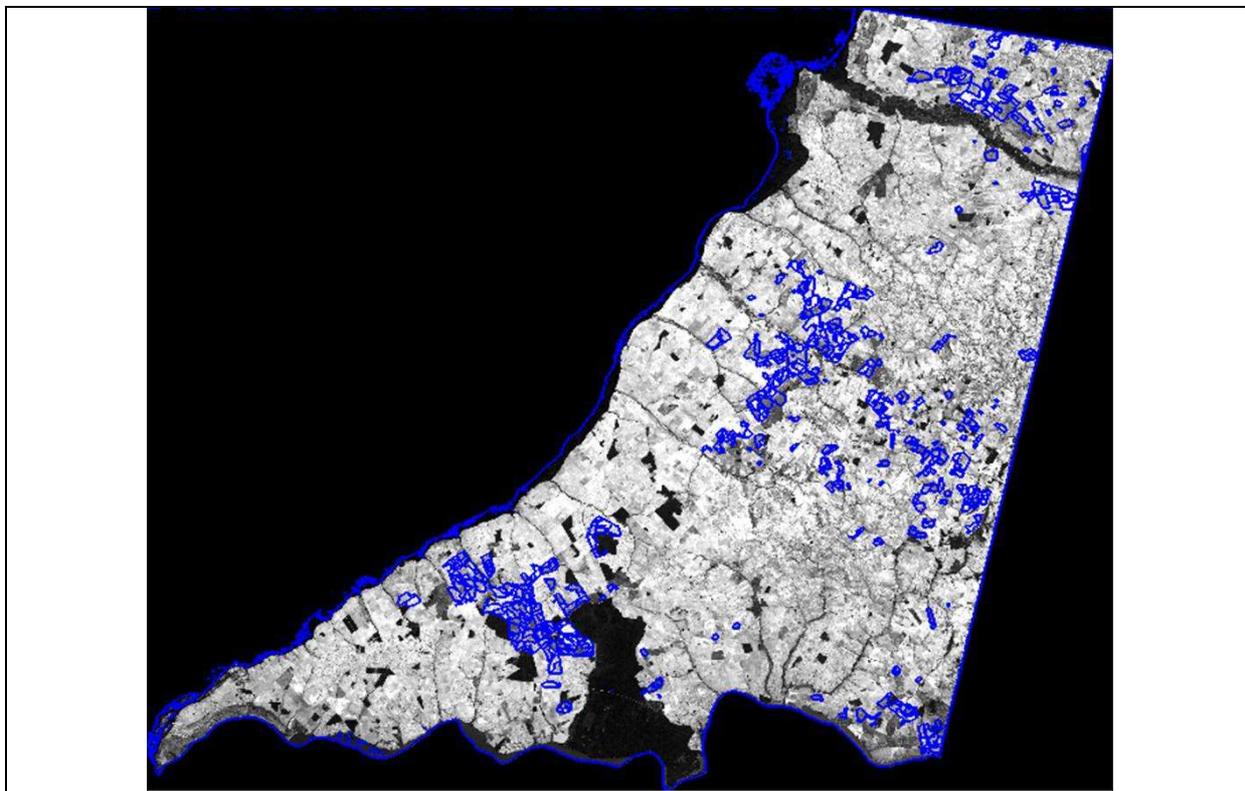


Figura 2 - Segmentação multiresolução (2º nível), somente na classe CANA.

Como resultado desta segmentação, observou-se a criação de vários novos polígonos de solo exposto dentro das áreas da classe CANA ao longo das três datas.

Após esta fase de identificação multitemporal das áreas novas, iniciou-se a fase de classificação final. Onde, os objetos com características em comum são agrupados em classes, e estas são organizadas em forma de rede hierárquica, onde as classes de um nível inferior herdam as características da classe do nível superior (Pinho, 2005).

Desta forma, foi criada mais uma classe na rede hierárquica denominada REFORMA. Com isso, algumas amostras de áreas de reforma foram selecionadas.

Para este caso o alvo de interesse, áreas de reforma, apresentava as seguintes condições no decorrer das datas:

- Setembro: solo exposto ou cana ainda em pé (porém com baixo valor NDVI);
- Fevereiro: solo exposto ou cultura de verão (geralmente soja);
- Março: solo exposto ou cultura de verão.

Já para as áreas de cana-de-açúcar durante as três datas, na maioria dos talhões, apresentavam a planta em pé (alto NDVI), aparecendo solo exposto devido à colheita em poucos casos (geralmente em setembro). Para delimitar esta situação foi criado um atributo entre a diferença do NDVI de setembro e março.

Com isso, tanto o atributo de diferença do NDVI quanto o atributo de brilho NDVI foram criados através da opção *Customized* no *Definies Developer* e selecionados para esta primeira separação entre as classes CANA e REFORMA. O atributo de brilho nada mais é do que a média

de NDVI das três datas. Com isso, áreas que apresentassem mais solo exposto teriam valores diferentes de brilho do que as áreas que apresentassem cana na maioria das datas. Os limiares utilizados também foram determinados a partir dos valores das duas amostras.

Porém, ainda havia um entrave nesta classificação, as áreas de cultura de verão, que aparecem em talhões de reforma. Seria difícil separá-las da classe de cana utilizando o atributo de brilho.

Analisando o comportamento da cultura de verão nas três datas, concluiu-se que:

- Setembro: solo exposto ou cana em pé.
- Fevereiro: cultura de verão.
- Março: cultura de verão ou solo exposto.

Neste caso, dois tipos de confusão poderiam ocorrer: Confusão 1 - a cultura de verão estivesse no campo em fevereiro e em março (alto NDVI); Confusão 2 - se em setembro ainda estivesse cana em pé e em uma das demais datas (fevereiro ou março) a cultura de verão.

Geralmente, a soja é a cultura de verão escolhida para rotação com a cana. Para resolver os dois tipos de confusão, o ideal seria separar as áreas de soja das áreas de cana. Sabendo-se que as áreas de soja se distinguem de forma muito fácil visualmente, aparecendo alaranjadas com a composição (RGB) 453, decidiu-se averiguar essa possibilidade de distinção entre elas.

Com isso, analisou-se algumas amostras de soja e cana observando o seu comportamento nas bandas 3, 4 e 5. Ao se analisar as bandas separadamente, verificou-se que na banda 5 a soja apresentava valores maiores do que a cana.

Desta forma, decidiu-se criar um índice normalizado entre as bandas 5 e 3 que mais se diferenciavam com o intuito de separar as áreas de soja das áreas de cana-de-açúcar.

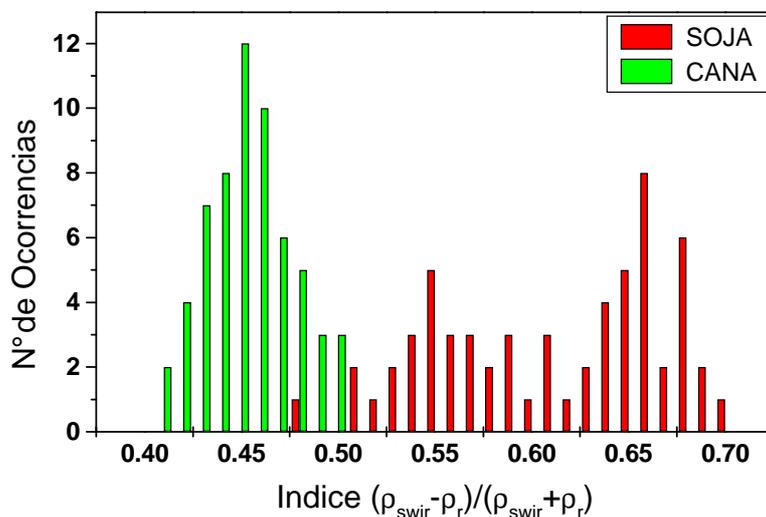


Figura 3 – Histograma das amostras de soja e cana para o índice criado pelas bandas 5 e 3.

Na Figura 3 pode-se verificar o histograma deste índice para algumas amostras dos dois alvos. Observa-se neste histograma que as duas culturas se diferenciaram, causando confusão apenas em uma das amostras. Pode-se ainda observar um limiar definido entre elas (0,5), o qual foi usado no trabalho para diferenciá-las. Vale salientar que este índice causou confusão entre a cana e o solo exposto, tendo de ser utilizado em conjunto com outra informação (banda ou outro índice) para diferenciá-las.

Após a criação deste atributo montou-se a rede hierárquica com base nas informações. A Figura 4 mostra a rede hierárquica utilizada na classificação final.

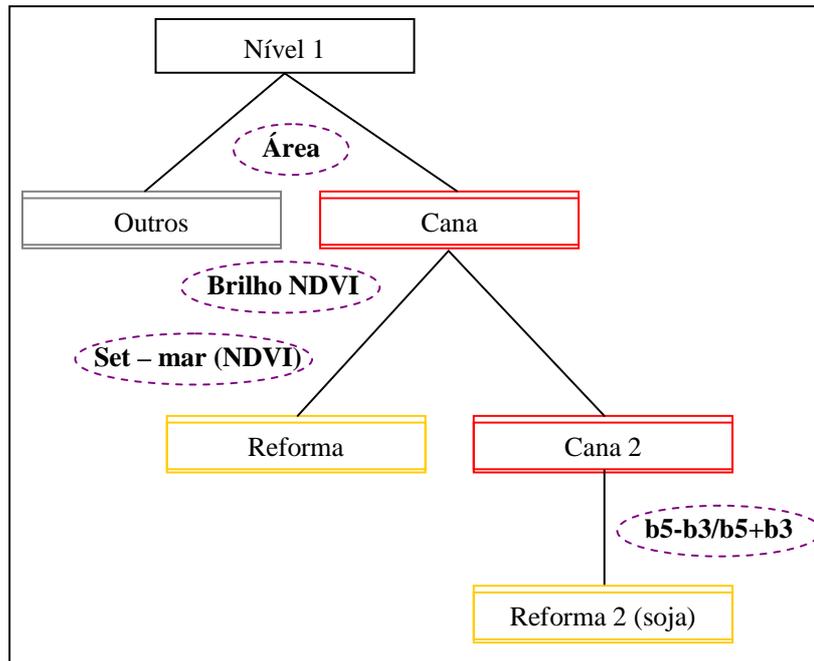


Figura 4- Rede hierárquica utilizada na classificação final.

A Figura 5 apresenta a classificação final gerada pela rede hierárquica e a Figura 6 mostra a classificação gerada por interpretação visual.

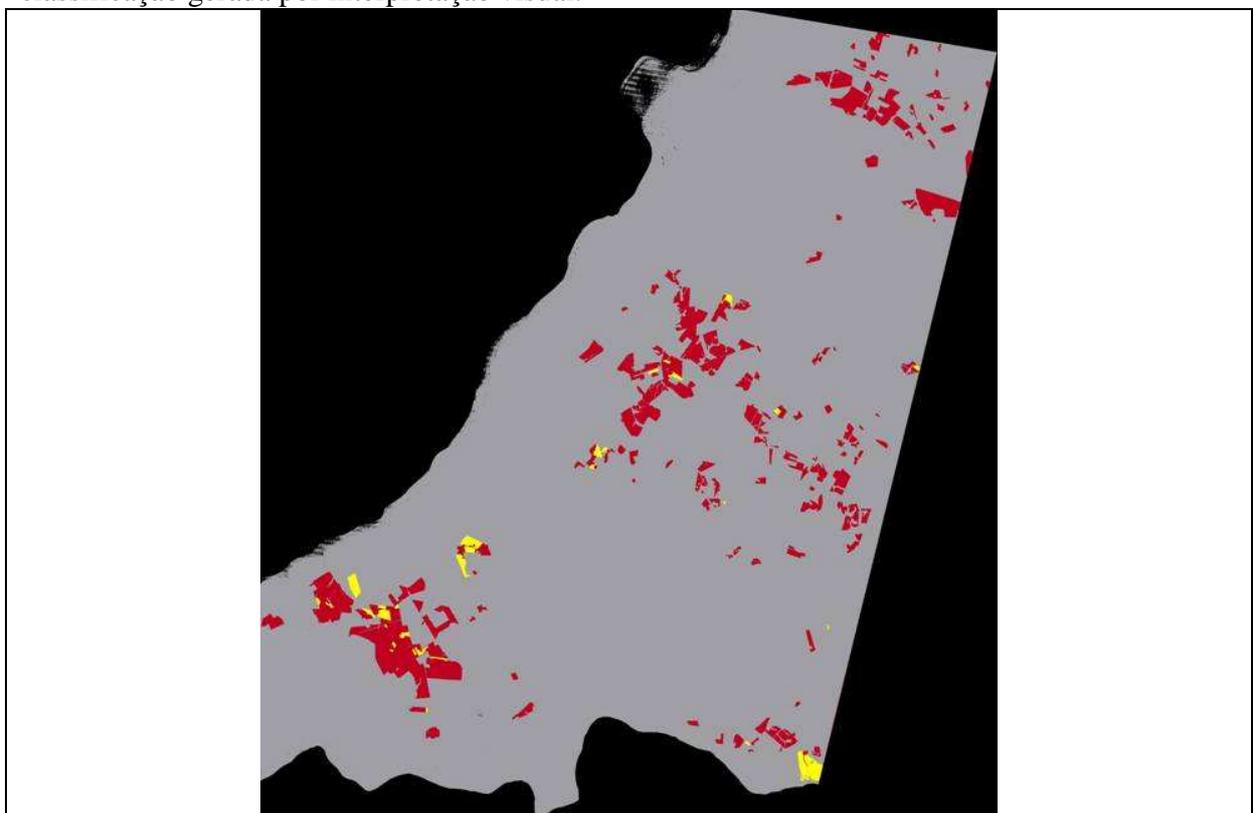


Figura 5- Classificação automática gerada pela rede hierárquica.

De maneira geral a maioria dos polígonos de reforma foi reconhecida na classificação automática.

Apareceram alguns casos de erros de inclusão cometidos pela classificação baseada na rede hierárquica. Estes casos ocorreram devido à presença de nuvens em uma das datas (março) e também, quando aparece cana com pouco vigor vegetativo em combinação com áreas de solo exposto. Nestas duas situações a média NDVI cai ao longo das datas, assim, considerando estas áreas como reforma.



Figura 5- Classificação gerada por edição matricial.

Durante a comparação entre os mapeamentos alguns polígonos pequenos não foram reconhecidos como reforma e eram reforma. Estes erros de comissão ocorreram devido a não identificação desses polígonos de reforma durante o processo da segmentação multiresolução.

4. Conclusões:

Verificou-se que a segmentação multiresolução reconheceu grande parte das áreas de reforma ao longo das datas. Isto foi possível devido à utilização da banda 3 na identificação de áreas de solo exposto. A utilização da razão entre as bandas 5 e 3 foi de grande ajuda na separação das classes de cana e soja.

Ao se comparar os mapeamentos gerados pela interpretação visual e pela rede hierárquica notou-se que a maioria dos polígonos das áreas de reforma foi classificada corretamente. A maioria dos erros foi de inclusão, destes alguns foram devido à presença de nuvens e outros devido à presença de cana com pouco vigor vegetativo em combinação com áreas de solo exposto.

Apareceram três casos de erros de comissão, que ocorreram devido a não segmentação de polígonos pequenos de reforma. Isto pode ser resolvido aplicando uma escala menor na segmentação multiresolução.

Vale salientar que o conhecimento do intérprete para a geração de uma rede hierárquica otimizada é peça fundamental no sucesso da classificação.

Referências Bibliográficas:

Chander, G.; Markham, B. Revised Landsat 5 TM radiometric calibration procedures and post-calibration dynamic ranges. 2003. Disponível em: <<http://www.dgi.inpe.br/html/radiometria-TM.htm>>. Acesso em: 01 out. 2008.

Feitosa, R. Q.; Mattos, C.; Santos, R. V.; Rego, L. F.G.; Pereira, A. A. C. Impacto de técnicas baseadas em conhecimento na análise de imagens de média resolução para mapeamento do desflorestamento na Amazônia. In: Simpósio Brasileiro de Sensoriamento Remoto, 12., 2005, Goiânia. **Anais...** São José dos Campos: INPE, 2005.p. 4065-4072. Disponível em: <<http://marte.dpi.inpe.br/col/ltid.inpe.br/sbsr/2004/11.16.18.14/doc/4065.pdf>>. Acesso em: 23 out. 2008.

Fockelmann R. Agricultural parcel detection with Definiens e-Cognition. Disponível em: <http://www.definiens.com/article.php?id=20&link_id=15&sublink_id=16>. Acesso em 8 out. 2008.

Goltz, E.; Pinho, C. M. D.; Fonseca, L. M. G.; Rudorff, B. F. T.; Aguiar, D. A. Detecção automática das áreas de reforma dentro das áreas de cana-de-açúcar (Projeto Canasat). In: Simpósio Latino Americano de Percepção Remota (Selper), 09, 2008, Havana – Cuba. Cd ROM.

Lucas, R.; Rowlands, A.; Brown, A.; Keyworth, S.; Bunting, P.. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 62, Issue 3, Agosto 2007, Pag. 165-185

Pinho, C. M. D. **Análise orientada a objeto de imagens de satélite de alta resolução espacial aplicada à classificação de cobertura do solo no espaço intra-urbano: o caso de São José dos Campos-SP**. 2005. 180 p. (INPE-14183-TDI/1095). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2005.

Rudorff, B. F. T.; Sugawara, L. M. Mapeamento da cana-de-açúcar na região Centro-Sul via imagens de satélites. **Informe Agropecuário**, v. 28, n. 241, p.79-86, 2007.