

Reclassificação controlada de *pixels* isolados em mapas de cana-de-açúcar

Luciana Miura Sugawara Berka¹
Bernardo Friedrich Theodor Rudorff¹
Ramon Morais de Freitas¹

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

Abstract. Thematic maps generated from remote sensing images provide valuable and updated information. However, during the updating process some errors are introduced leaving isolated pixels or group of pixels that need to be correctly reclassified. Standard procedures are not efficient to correct for this source of error and may even enhance the problem. Therefore, a specific algorithm was developed to minimize errors from isolated pixels on a sugarcane thematic map by reclassifying these pixels on a controlled and automated way. Prior to the reclassification process a mask with isolated groups of pixels (<1 ha), was generated using a conventional pixel cleaning procedure and map subtraction operation. Both map algebra and statistic approaches were used during the decision criteria procedure of reclassification. No significant difference was observed for area comparison between estimates prior and after the reclassification procedure. However, the map quality was significantly improved; furthermore, error propagation is avoided during yearly map updating.

Palavras-chave: remote sensing, mapping, sugarcane, map algebra, sensoriamento remoto, mapeamento, cana-de-açúcar, álgebra de mapa.

1. Introdução

O sensoriamento remoto (SR) tem a capacidade de fornecer dados para grandes extensões territoriais a custos relativamente baixos, se comparado com levantamentos feitos a campo. Além disso, o SR possibilita a obtenção de dados em freqüentes intervalos de tempo. Estas características permitem a realização de estudos de uso da terra, detecção de mudanças, previsão e planejamento (Prenzel, 2004; Rogan & Chen, 2004).

Em estudos de detecção de mudanças é importante conhecer os erros associados e a propagação dos mesmos, pois isso interfere na decisão a ser tomada no manejo de uma determinada área. Uma ampla literatura pode ser encontrada com a descrição destes erros e os métodos para detectá-los, tais como os trabalhos de Lunetta et al. (1991), Goodchild & Gopal (1992), Lanter & Veregin (1992), Congalton & Green (1999) e Fuller et al. (2003).

Neste trabalho foi abordado apenas um tipo de erro ocorrido no mapeamento da cana-de-açúcar no Estado de São Paulo. A metodologia para este mapeamento vem sendo desenvolvida pelo Instituto Nacional de Pesquisas Espaciais (INPE) num projeto executado em parceria com o Centro de Estudos Avançados em Economia Aplicada (CPEA) e a União da Agroindústria Canavieira de São Paulo (UNICA). O mapeamento foi realizado para as safras 2003/2004 e 2004/2005, empregando imagens dos satélites Landsat 5 e 7 (Rudorff et al., 2004a; Rudorff et al., 2004b).

Na safra 2004/2005 o mapa da cana-de-açúcar foi atualizado tomando como base o mapeamento do ano safra anterior. A cana-de-açúcar é uma cultura bastante dinâmica de uma safra para outra, principalmente em função das áreas de reforma. Com a observação das imagens Landsat, verificou-se que estas áreas poderiam ser extraídas através de uma técnica de classificação automática não supervisionada (Rudorff et al, 2004b).

Congalton & Green (1999) comentaram que nem mesmo a técnica mais eficiente de classificação é capaz de gerar um resultado sem erros. O uso de classificação automática fez com que ocorresse a presença de *pixels* isolados no mapa da safra 2004/2005. Um fator que pode ter contribuído para este erro foi a ineficiência do algoritmo do classificador na detecção

de bordas das áreas de reforma de cana-de-açúcar (Lunetta et al., 1991; Veregin, 1992). Outro fator que influenciou a geração de *pixels* isolados foi o erro, dentro dos limites aceitáveis, no registro das imagens dos dois anos safra mapeados. Além disso, o intérprete pode ter deixado de corrigir pequenos erros que permaneceram ou foram introduzidos durante a etapa de conferência e correção dos mapas.

Na **Figura 1** é mostrado um detalhe dos mapas das safras 2003/2004 (a) e 2004/2005 (b) em que as setas vermelhas indicam os *pixels* isolados presentes no mapa de 2004 e que são fonte de erro.

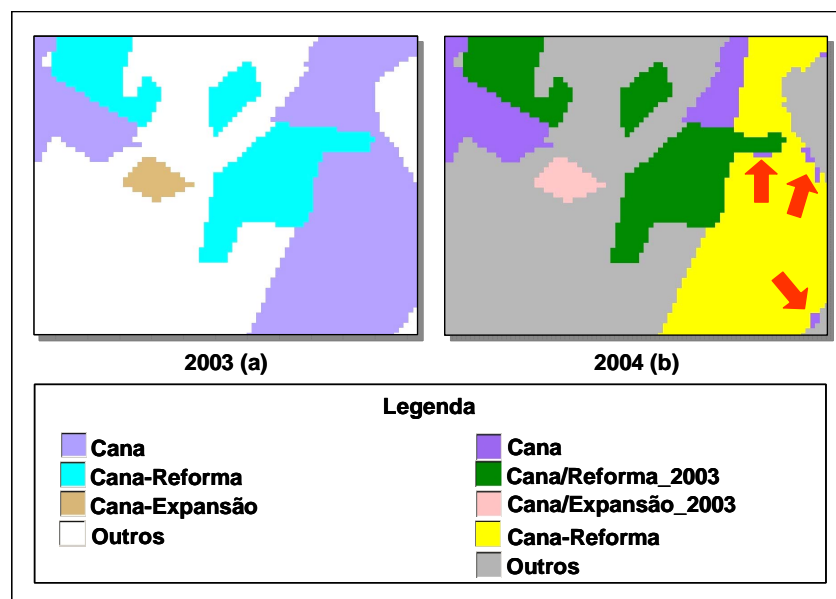


Figura 1: Detalhe do mapa de cana-de-açúcar das safras 2003/2004 (a) e 2004/2005 (b).

Fonte: Rudorff et al., 2004a; Rudorff et al., 2004b

Os erros gerados pela presença de *pixels* isolados podem não ser significativos no cálculo de área, mas os erros acumulados ao longo dos anos podem ser problemáticos, pois a localização das áreas se tornará cada vez mais incerta.

As abordagens utilizadas atualmente para a correção desses erros não conseguem atribuir classes aos *pixels* isolados de forma controlada, causando até mesmo o aumento dos erros. Em função disso, foi necessário desenvolver um método de reclassificação destes *pixels* isolados com o uso de critérios de decisão.

O objetivo deste trabalho é desenvolver um método de reclassificação controlada e automática dos *pixels* isolados com o uso de programação em Linguagem Espacial para Geoprocessamento Algébrico (LEGAL) para o mapa da cana-de-açúcar.

2. Material e Métodos

A metodologia foi desenvolvida na área canavieira do Estado de São Paulo abrangida pela órbita/ponto 220/74 do sistema de referência mundial do satélite Landsat (**Figura 2**).

O trabalho foi realizado com o uso do programa Sistema de Processamento de Informações Georreferenciadas (SPRING), versão 4.1 e o seu módulo de programação denominado LEGAL, que é uma linguagem de consulta e manipulação espacial que realiza operações sobre dados do tipo: mapa temático, modelo numérico de terreno e imagem (INPE, 2004). O microcomputador utilizado possui um processador Pentium IV de 2,4GHz, memória de acesso aleatório (RAM) de 1 Gigabytes e disco rígido de 120Gigabytes.

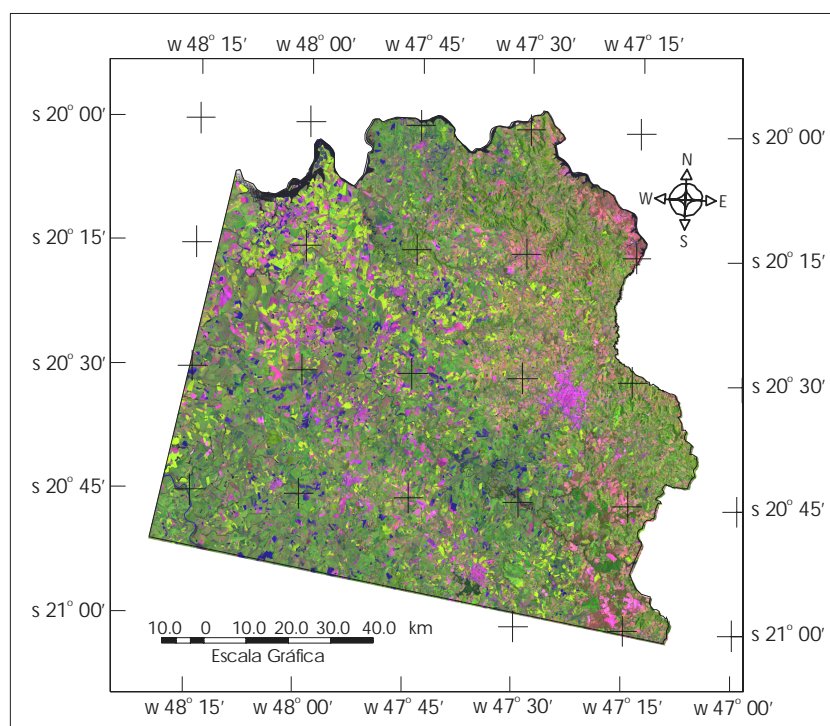


Figura 2: Área de estudo - Composição colorida RGB 543 do Landsat 7 ETM+ de 25/02/2003 - Cena 220/74.

O mapa de cana a ser corrigido foi gerado no trabalho de Rudorff et al. (2004b). As classes temáticas presentes no mapa da cana são: **Cana**, **Cana/Reforma_2003**, **Cana/Expansão_2003**, **Cana-12meses**, **Cana-Reforma**, **Cana-Muda**, **Cana-Dúvida** e **Outros**.

A primeira etapa do processo consistiu em duplicar o Plano de Informação (PI) contendo o mapa da cana-de-açúcar (**Figura 3a**). O PI duplicado foi transformado do formato matriz para vetor, sem a suavização de arcos, com a finalidade de utilizar a técnica de limpeza de vetor implementada no SPRING, eliminando polígonos menores que 1 ha. Este limiar de 1 ha foi determinado através de testes com diversos valores de limiar aplicados à técnica de limpeza de vetor. Após a aplicação desta técnica foi gerada uma nova matriz para este PI (**Figura 3b**).

Na **Figura 3**, no destaque 1, representado por um círculo vermelho, a classe **Outros** passou a ser classe **Cana** após a limpeza de vetor, porém não está correto, pois deveria permanecer como classe **Outros**. No destaque 2 a classe **Cana** passou a ser classe **Cana/Reforma 2003**, o que também não foi correto, pois deveria ser classificada como **Cana Reforma**. No destaque 3 a limpeza de vetor reclassificou de forma correta. Ou seja, não foi possível determinar para qual classe os *pixels* seriam reclassificados.

No passo seguinte um programa em linguagem LEGAL foi executado para criar um PI contendo as diferenças existentes entre o PI original e o PI com a limpeza de vetor. Este PI foi denominado de “máscara” (**Figuras 3c**) e determinou que apenas os *pixels* contidos nela fossem reclassificados.

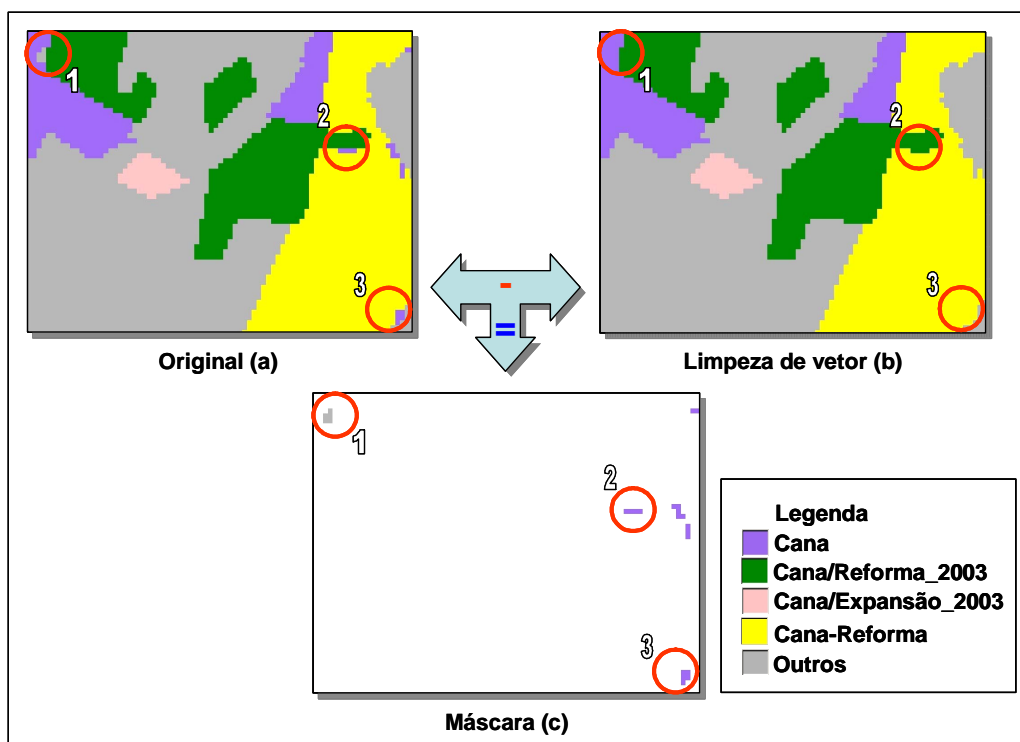


Figura 3: Obtenção da máscara.

A determinação da classe para a qual os *pixels* isolados foram reclassificados foi baseada em operações condicional, booleana e de vizinhança, com o auxílio de comandos de controle (“enquanto”) e o uso do cálculo da moda (Figura 4).

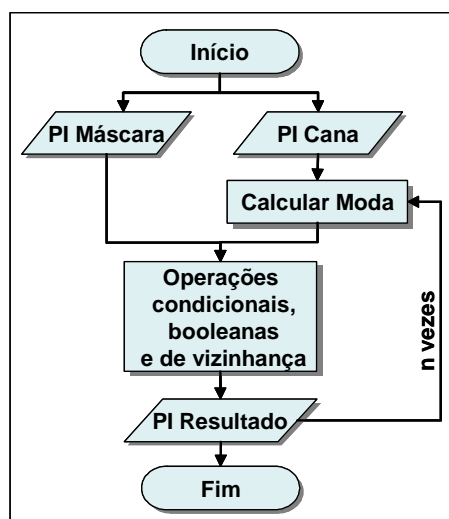


Figura 4: Fluxograma do programa em LEGAL.

Loops foram necessários para que a reclassificação fosse feita completamente dentro da área delimitada pela máscara. A determinação do número de *loops* foi realizada numa área teste de 113 linhas e 132 colunas, que continha a pior situação possível para o processamento. Esta situação é determinada pela forma como os *pixels* são lidos, pela quantidade de *pixels* conexos e pelo seu formato. No SPRING o processamento é feito *pixel a pixel* no modo *row-wise binding*, ou seja, do canto superior esquerdo para o canto inferior direito, sempre iniciando a leitura de uma linha pelo lado esquerdo. O número máximo de *pixels* conexos foi

determinado pelo limiar utilizado na limpeza de vetor. Outro dado obtido nessa área teste foi o tempo aproximado de processamento em função do aumento do número de *loops*.

A classe atribuída ao *pixel* analisado depende de sua vizinhança, de acordo com as classes existentes no mapa. Além disso, para que o programa gerado reclassificasse corretamente, foi seguido uma seqüência de prioridades de classes. Por exemplo, se a máscara contém a classe **Cana**, na reclassificação a classe **Cana-Reforma** tem prioridade sobre as demais classes, ou seja, a primeira possibilidade de reclassificação é para a classe **Cana-Reforma**. Caso as condições não sejam satisfeitas, o programa seguirá percorrendo as linhas de comando até que encontre aquela que as satisfaça. A **Figura 5** apresenta um trecho do fluxograma do algoritmo utilizado no programa em linguagem LEGAL.

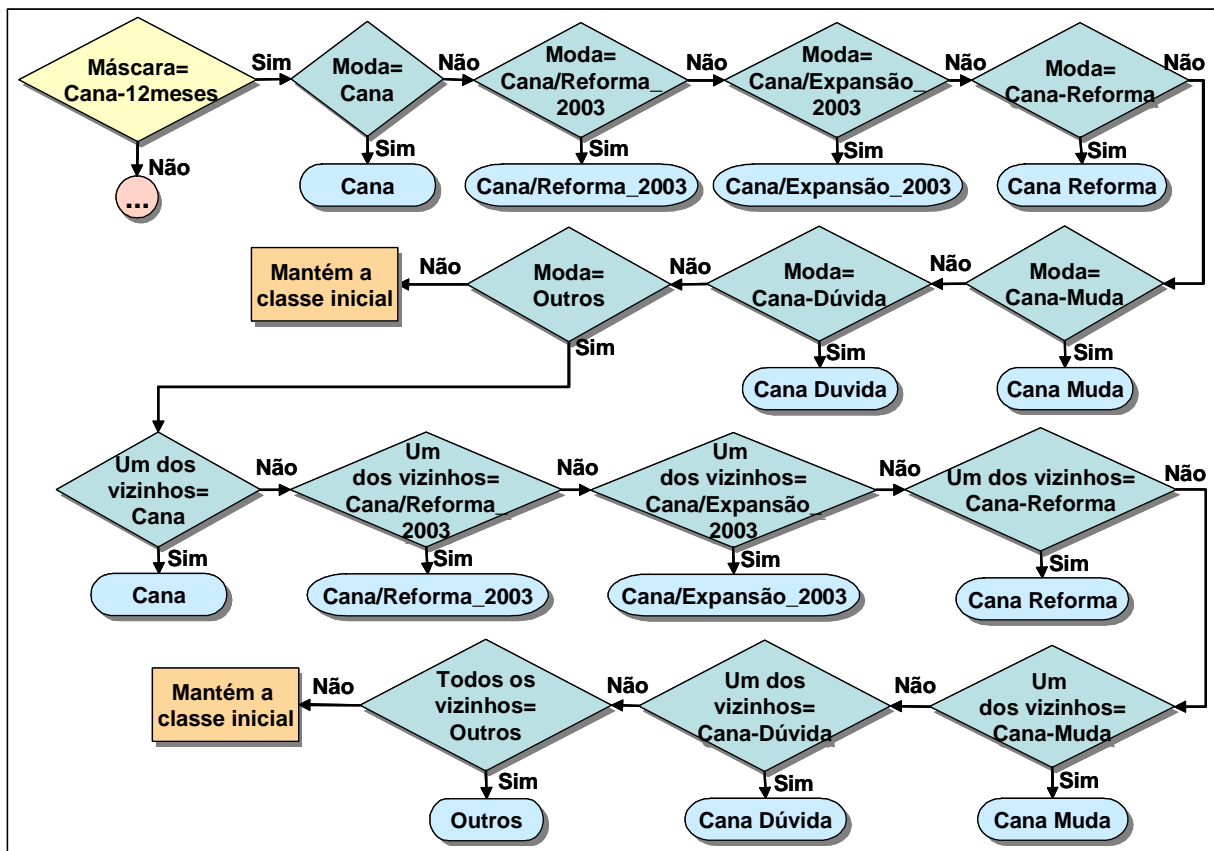


Figura 5: Trecho do fluxograma dos critérios de decisão utilizados no programa em linguagem LEGAL.

O programa em linguagem LEGAL não foi apresentado neste trabalho devido ao número de linhas de comando exceder o espaço concedido à elaboração do trabalho, mas ele se encontra disponível na internet no seguinte endereço: http://geocities.yahoo.com.br/lmiura2002/legal/Pos_Class.txt.

3. Resultados e Discussão

O tempo de processamento do programa em linguagem LEGAL é descrito por uma função linear (**Figura 6a**), já que o processamento é feito *pixel a pixel* pelo SPRING. Desta forma, com a obtenção do tempo de processamento para a uma pequena área é possível ter uma idéia do tempo para toda a área. Este tempo pode ser impreciso em função do comportamento do computador durante o processamento, o que pode retardar a execução do programa. Porém, o

indicativo de tempo de processamento é importante para o planejamento em trabalhos operacionais.

A pior situação para a reclassificação da área de estudo deste trabalho é apresentada na **Figura 6b**, representado pelos *pixels* da classe **Cana/Reforma 2003**. Caso o formato dos *pixels* conexos fosse ao contrário, com a parte mais larga na parte de baixo, o número de *loops* poderia ser reduzido.

O limiar na limpeza de vetor é um valor crítico e deve ser determinado através de teste com diversos valores. Veregin (1992) comentou este fato no uso de um modelo de correção de polígonos espúrios gerados pela sobreposição de PI's. No caso da cana, o valor de 1 ha mostrou ser satisfatório, pois com valores maiores poderia ocorrer a inclusão de áreas que estão classificadas corretamente e valores menores deixariam muitos *pixels* sem reclassificação.

A **Figura 6b** ilustra ainda a determinação do número de *loops* necessários. A quantidade de *pixels* reclassificados aumenta a cada acréscimo de um *loop* até se tornar constante, o que indica que não necessita de mais *loops* para a reclassificação dos *pixels* isolados. A área de estudo deste trabalho necessitou de seis *loops*.

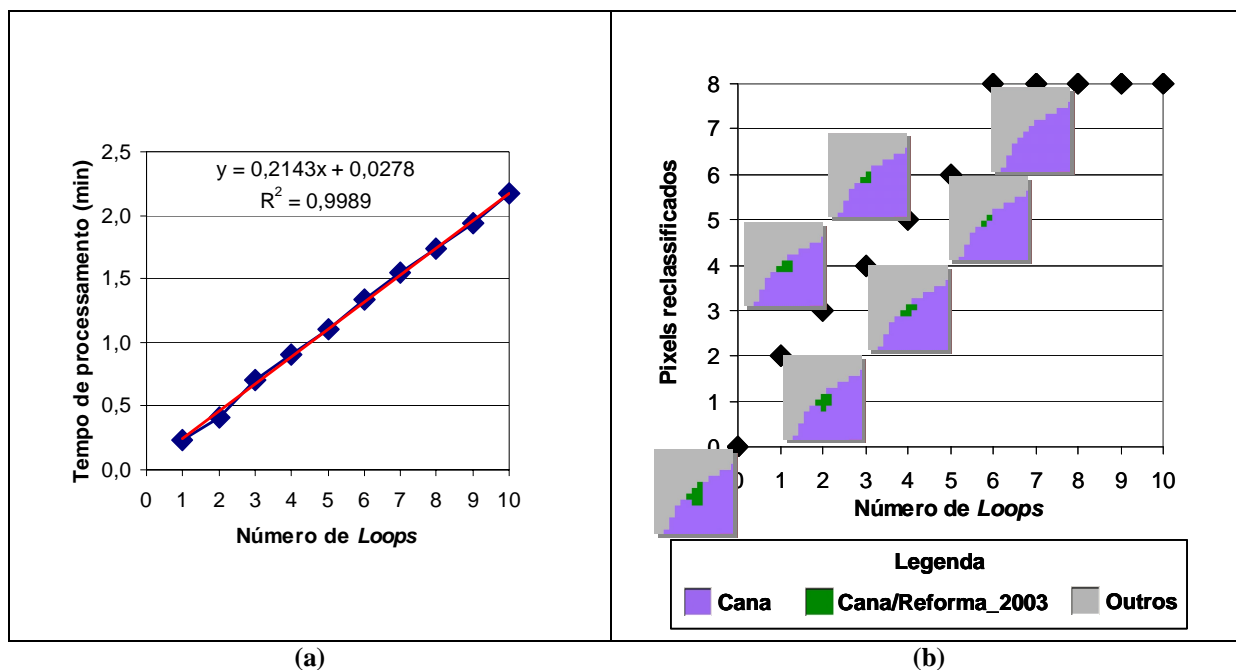


Figura 6: Aumento do tempo de processamento em função do número de *loops* (a) e estimativa do número de *loops* (b).

O mapa final da reclassificação comparado ao mapa inicial pode ser visto na **Figura 7**. O algoritmo executado através de programação em LEGAL foi satisfatório. Como era esperado, o destaque 1 permaneceu classificado como **Outros** e os destaques 2 e 3 foram reclassificados para **Cana-Reforma**.

Nem todos os problemas de *pixels* isolados foram resolvidos, devido ao limiar de área utilizado, mas foram minimizados a um ponto em que será possível realizar a correção manual com maior facilidade do que no mapa inicial (**Figura 7a**).

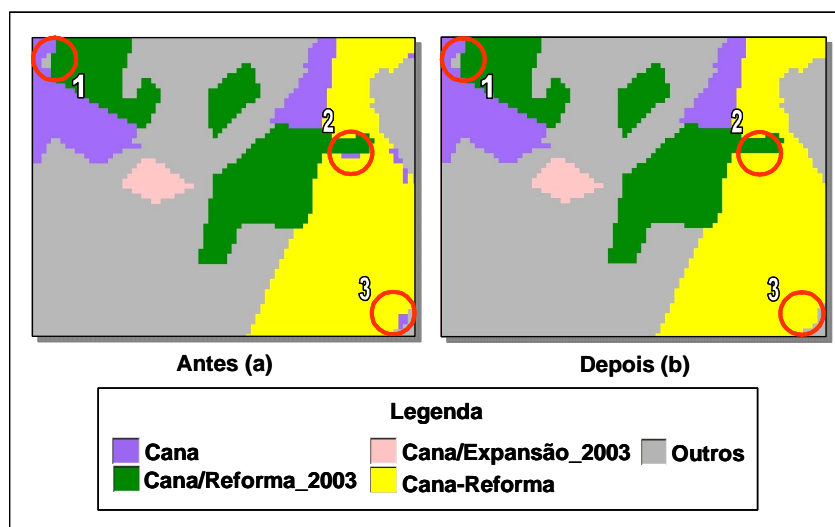


Figura 7: Mapa inicial (a) e o resultado da reclassificação dos *pixels* isolados (b).

A **Tabela 1** mostra os resultados divididos por classe no mapa da cana. As diferenças existentes entre os mapas antes e depois do processamento são mínimas. Isto comprova que quantitativamente o mapeamento não sofre influência em função dos *pixels* isolados presentes. Mas visualmente e em termos de propagação dos erros o mapa é melhorado com o processamento realizado.

Tabela 1: Área, diferenças absoluta e relativa do mapa resultante

Classe	Antes (ha)	Depois (ha)	Diferença absoluta (ha)	Diferença Relativa (%)
Cana	280.624,32	280.446,93	-177	-0,1
Cana/Reforma_2003	49.056,75	49.039,92	-17	-0,03
Cana/Expansão_2003	1.992,96	1.992,96	0	0,0
Cana-12meses	10.545,93	10.574,73	29	0,3
Cana-Reforma	36.138,15	36.297,99	160	0,4
Cana-Muda	521,55	527,31	6	1,1
Cana-Dúvida	15,21	14,94	-0,3	-1,8
Outros	680.879,88	680.879,97	0,09	0,00001

O modelo de reclassificação gerado é baseado em álgebras de mapas e estatística, o que foi possível devido à característica matricial do mapa de cana. Ou seja, um PI no formato matricial pode ser tratado como se fosse uma imagem e os *pixels* isolados podem ser comparados aos ruídos presentes em imagens de SR. Contudo, a abordagem para a resolução do problema não pode ser a mesma aplicada a uma imagem, pois como foi descrita na metodologia, é preciso executar uma série de critérios de decisão para que um bom resultado seja alcançado.

A reclassificação manual de *pixels* isolados no SPRING é onerosa, exige concentração do intérprete, não é operacional e é mais uma fonte de erros. Portanto, a reclassificação automática e controlada é uma ferramenta indispensável para reduzir o tempo e o custo da geração de mapas temáticos.

A eliminação total de erros é impraticável (Verigin, 1992), mas as medidas possíveis devem ser tomadas para minimizá-los. Embora não seja o escopo deste trabalho, é importante

salientar que os erros associados aos mapas devem ser investigados e divulgados, pois são importantes para os tomadores de decisão (Lunetta et al., 1991).

4. Conclusão e considerações finais

O método desenvolvido para reclassificação controlada e automática de *pixels* isolados por meio da programação em Linguagem Espacial para Geoprocessamento Algébrico (LEGAL) para o mapa da cana-de-açúcar foi eficiente e pode ser adaptado para a reclassificação de *pixels* isolados para qualquer mapa no formato matricial. É importante salientar que este método não consiste na mera eliminação de ruídos ou possíveis erros conforme é feito no algoritmo de limpeza de vetor implementado no aplicativo SPRING, mas sim na utilização criteriosa de um procedimento automático e interativo de tomada de decisão que visa alocar os *pixels* isolados na classe à qual supostamente devem pertencer. Já era de se esperar que o resultado da estimativa de área da classificação temática não sofreria alterações significativas em função da compensação dos erros encontrados nas diferentes classes temáticas. Contudo, o mapeamento foi significativamente melhorado; além disso, evita-se a propagação dos erros nos mapeamentos subsequentes que visam atualizar as estatísticas agrícolas a cada ano.

Referências

- Congalton, R. G.; Green, K. **Assessing the accuracy of remotely sensed data: principles and practices**. London: Lewis, 1999. 137 p.
- Fuller, R. M.; Smith, G. M.; Devereux, B. J. The characterisation and measurement of land cover change through remote sensing: problems in operational applications? **International Journal of Applied Earth Observation and Geoinformation**, v. 4, n. 3, p. 243-253, 2003.
- Goodchild, M.; Gopal, S. **Accuracy of spatial databases**. New York: Taylor & Francis, 1992. 290 p.
- Instituto Nacional de Pesquisas Espaciais (INPE). **Sistema de Processamento de Informações Georreferenciadas (SPRING)**. São José dos Campos. Disponível em: <www.dpi.inpe.br/spring>. Acesso em: 01 outubro 2004.
- Lanter, D. P.; Veregin, H. A research paradigm for propagating error in layer-based GIS. **Photogrammetric Engineering Remote Sensing**, v. 58, n. 6, p. 825-833, 1992.
- Lunetta, R.S., Congalton, R.G., Fenstermaker, L.K., Jensen, J.R., McGuire, K.C., Tinney, L.R.. Remote sensing and geographical information systems data integration: error sources and research issues. **Photogrammetric Engineering Remote Sensing**, v. 57, n. 6, p. 677-687, 1991.
- Prenzel, B. Remote sensing-based quantification of land-cover and land-use change for planning. **Progress in Planning**, v. 61, n. 4, p. 281-299, 2004.
- Rogan, J.; Chen, D. Remote sensing technology for mapping and monitoring land-cover and land-use change, **Progress in Planning**, v. 61, n. 4, p. 301-325, 2004.
- Rudorff, B. F. T.; Berka, L. M. S.; Xavier, A. C.; Moreira, M. A.; Duarte, V.; Rosa, V. G. C.; Shimabukuro, Y. E. **Estimativa de área plantada com cana-de-açúcar em municípios do estado de São Paulo por meio de imagens de satélites e técnicas de geoprocessamento: ano safra 2003/2004**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2004a. 47 p. (INPE-10791-RPQ/759).
- Rudorff, B.F.T.; Berka, L.M.S.; Moreira, M.A.; Duarte, V.; Rosa, V.G.C. **Estimativa de área plantada com cana-de-açúcar em municípios do Estado de São Paulo por meio de imagens de satélites e técnicas de geoprocessamento: ano safra 2004/2005**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2004b. 39p. (INPE-11421-RPE/762).
- Veregin, H. Error modeling for the map overlay operation. In: Goodchild, M.; Gopal, S. (ed.) **Accuracy of spatial databases**. New York: Taylor & Francis, 1992. p. 3-18.