

## Investigação sobre o uso de dados auxiliares no processo de classificação de imagens digitais empregando princípios da teoria da evidência

Rodrigo Lersch<sup>1</sup>  
Vitor Haertel<sup>1</sup>  
Yosio Shimabukuro<sup>2</sup>

<sup>1</sup>Universidade Federal do Rio Grande do Sul, UFRGS  
C.P. 15044, Porto Alegre, RS, 91501-970, Brasil  
(rodrigo.lersch,victor.haertel)@ufrgs.br

<sup>2</sup>Instituto Nacional de Pesquisas Espaciais, INPE  
C.P. 515, São José dos Campos, SP, 12201-970, Brasil  
yosio@dsr.inpe.br

### **Abstract**

In this study we investigate the use of the concept of belief as proposed by the Theory of Evidence to increment the accuracy in the digital image classification process. We propose a methodology starting with the selection of the auxiliary data which is relevant to the problem under consideration, followed by the estimation of the amount of uncertainty and evidence as conveyed by each individual variable and the production of the layer of belief. Finally we proposed a quantitative way to use the information conveyed by belief to filter out mislabeled pixels in the thematic map. An example is provided to illustrate the proposed methodology.

### **Resumo**

No presente estudo investiga-se a utilização do conceito de confiança (*belief*), conforme proposto pela Teoria da Evidência com a finalidade de aumentar a acurácia no processo de classificação de imagens digitais. A metodologia aqui proposta inicia pela seleção das variáveis relevantes a cada problema em particular, seguida pela estimação dos valores de incerteza e de evidência associados a cada variável individual, e pela produção do *layer* de confiança. Nesta metodologia é então proposto uma abordagem quantitativa para utilizar a informação fornecida pela confiança para fins de remoção de pixels erroneamente rotulados na imagem temática produzida pelo classificador.

**Palavras-chave:** evidence, belief, digital image classification, evidência, confiança, classificação de imagens digitais.

### **1. Introdução**

No processo de classificação de imagens digitais em Sensoriamento Remoto, diversos fatores podem contribuir para um decréscimo na acurácia da imagem temática produzida pelo classificador. Entre aqueles que mais freqüentemente ocorrem, pode-se mencionar entre outros, a presença de classes espectralmente muito semelhantes entre si, a não normalidade nas distribuições das classes individuais, a não uniformidade na irradiância solar ao longo da cena, devido ao relevo do terreno. Diferentes propostas vêm sendo estudadas por vários pesquisadores com a finalidade de minorar a influência negativa destes fatores na acurácia da imagem temática produzida. Neste contexto, uma possibilidade que tem sido bastante investigada consiste na introdução de dados auxiliares no processo de classificação. Hutchinson (1982) agrupa as várias abordagens propostas em três categorias: (i) introdução de dados auxiliares antes da classificação, (ii) durante o processo de classificação e (iii) após a classificação. Como exemplo da primeira categoria, pode-se mencionar a segmentação da cena de estudo, em áreas ou estratos mais homogêneos. Esta estratégia oferece o potencial de se reduzir a variância das classes em cada segmento, contribuindo assim para a redução do

erro de classificação. Os segmentos individuais são então processados separadamente e os resultados parciais agrupados posteriormente para produzir a imagem temática final. Dados auxiliares podem também ser incorporado durante o processo de classificação. Neste caso, uma abordagem frequentemente utilizada consiste em organizar os dados auxiliares na forma de “bandas” adicionais que são então agregadas ao conjunto das bandas espectrais disponíveis. Uma outra proposta consiste na utilização dos dados auxiliares para fins de uma estimativa mais confiável das probabilidades *a priori*. Na abordagem pós-classificação, a imagem temática produzida pelo classificador é refinada com o auxílio dos dados auxiliares, buscando-se identificar pixels erroneamente rotulados. Dentro desta última abordagem, alguns autores relatam o emprego de dados auxiliares como o modelo digital de elevação, declividade e aspecto para incrementar a acurácia da imagem temática (Bruzzone et al. 1997). No presente estudo é investigado o uso dos princípios propostos na Teoria da Evidência de Dempster-Shafer como forma de introduzir dados auxiliares, oriundos de fontes diversas, no processo de classificação. Esta abordagem genérica, investigada neste estudo, foi anteriormente estudada por alguns pesquisadores. Lee et al. (1987) testaram o emprego de dados de evidência utilizando imagens Landsat MSS, utilizando as bandas na região visível do espectro e aquelas na região do infra-vermelho, como duas fontes independentes de dados. Os princípios da Teoria da Evidência foram também empregados por Moon (1990) para integrar dados de natureza geológica e geofísica. Gong (1996) desenvolveu trabalhos comparando a abordagem que emprega dados de evidência com aquela que emprega técnicas de redes neurais. Em seu estudo, aquele autor utilizou dados de fontes diversas como dados Landsat-TM, dados aero-magnéticos, radiométricos e gravimétricos.

O objetivo deste estudo consiste em investigar a utilização de conceitos propostos pela Teoria da Evidência no processo de classificação de imagens digitais em Sensoriamento Remoto. Na abordagem proposta, dados auxiliares de fontes diversas são organizados na forma de *layers*, à semelhança de um SIG, para fins de geração do *layer* de confiança (*belief*). Neste processo, são aqui propostas metodologias visando abordagens quantitativas na ponderação dos dados multi-espectrais e das variáveis auxiliares individuais bem como na utilização do *layer* de confiança para fins de detecção de erros de inclusão na imagem temática produzida a partir dos dados multi-espectrais. A metodologia proposta é aqui ilustrada por meio de uma aplicação envolvendo uma área teste com verdade terrestre conhecida.

## 2. Metodologia

Nesta seção é inicialmente feita uma rápida revisão dos princípios básicos desenvolvidos na Teoria da Evidência. No contexto deste estudo, as duas características mais relevantes da Teoria da Evidência são: (1) a possibilidade de utilizar-se dados de fontes diversas, inclusive dados de natureza qualitativa, isto é, dados em forma originalmente não numérica, e (2) a possibilidade de implementar o conceito de incerteza ou ignorância (Lee et al., 1987, Richards and Jia, 1999, Eastman, 1999).

Inicialmente, as variáveis oriundas de várias fontes são individualmente organizadas na forma de *layers*, expressando o caráter de variáveis espacialmente distribuídas. Para que as várias variáveis possam ser combinadas entre si, é necessário que os vários *layers* sejam previamente registrados entre si, ou ainda melhor, geo-referenciados. O passo inicial consiste em gerar para cada variável individual o respectivo *layer* de evidência. A evidência  $m(X)$  estima o grau de conhecimento disponível, que permite rotular um pixel  $X$  como pertencente a cada uma das classes individuais presentes na cena. A evidência é normalmente estimada fazendo-se uso de funções de probabilidade ou, quando esta opção não é viável, de funções de pertinência. Nesta abordagem, para cada pixel  $X$  e em cada uma das variáveis estarão

associados tantos graus de evidência quantas são as classes em consideração. Assim, o grau de evidência em favor de cada classe individual fica compreendido no intervalo  $[0,1]$ . Em situações reais, ocorre frequentemente um certo grau de incerteza ou ignorância com relação a pertinência de um pixel as classes presentes na cena. Este fato pode ser levado em consideração, restringindo-se o valor máximo que o grau de evidência pode assumir, para um valor menor, por exemplo 0.8. Neste exemplo, 20% da evidência total fica reservada como incerteza ou ignorância. O somatório das evidências mais a incerteza é, portanto igual à unidade. Na Teoria da Evidência, além da evidência as seguintes variáveis adicionais são utilizadas: confiança (*belief*), plausibilidade e intervalo de confiança. A confiança  $bel(X)$ , estima a evidência concreta total em favor de rotular um pixel como membro de uma determinada classe, enquanto que a plausibilidade  $plau(X)$  estima a quantidade máxima de evidência disponível em favor de cada possível rótulo para um pixel. Assim, a plausibilidade estima o grau com que um determinado rótulo parece estar correto, mesmo que esteja faltando evidência concreta em seu favor. Confiança e plausibilidade podem ser entendidas respectivamente como os valores mínimo e máximo de evidência disponível em favor de um rótulo em particular. Finalmente, intervalo de confiança, também conhecido como intervalo de evidência, é definido como a diferença entre plausibilidade (limiar superior) e confiança (limiar inferior) e pode ser interpretado como uma estimativa da incerteza existente em aceitar ou rejeitar um determinado rótulo para pixels individuais.

Os conceitos acima colocados podem ser melhor explicados por meio de um exemplo. Considere um problema de classificação envolvendo quatro classes:  $\omega_1$ ,  $\omega_2$ ,  $\omega_3$  e  $\omega_4$ . Estas classes formam um conjunto exaustivo, o que significa que na prática uma das quatro classes pode ser rotulada como “outros” (Lee et al., 1987). Suponhamos que para um determinado pixel a probabilidade de pertinência para as quatro classes foi estimada como proporcional a 2:2:1:1. Suponha-se ainda que existe um grau de incerteza na atribuição do rótulo ao pixel. Foi ainda estimado que, a confiança no processo de classificação é de 90%, ficando os restantes 10% associados com características de incerteza ou de ignorância. Neste caso, a evidência disponível pode ser expressa pelas seguintes massas de evidência:

$$m(\omega_1)=0.3 \quad m(\omega_2)=0.3 \quad m(\omega_3)=0.15 \quad m(\omega_4)=0.15 \quad m(U)=0.1$$

onde U representa a incerteza no processo de rotulação do pixel. Estes conceitos podem ser ainda tornados mais gerais, ao se assumir que a classe  $\omega_3$  representa na verdade a união das classes  $\omega_1$  e  $\omega_2$  ( $\omega_1 \cup \omega_2$ ). Isto significa que para o pixel em consideração o valor de 0.15 atribuídos a  $m(\omega_3)$  traduz o grau adicional de evidência disponível indicando que o pixel pertence ou a classe  $\omega_1$  ou a classe  $\omega_2$ , embora não seja possível precisar a qual das duas. Propostas diferentes para estimação das massas de evidência são encontradas na literatura. Gong (1996) menciona dois procedimentos distintos: tabelas de frequência e o modelo de distribuição normal (similar ao classificador máxima verossimilhança). Lee et al. (1987) propõe o seguinte procedimento: na estimação do valor da massa de evidência e cada pixel em particular, primeiro estima-se o valor da incerteza no processo de classificação (o erro média na classificação, por exemplo). O restante disponível será então atribuído como evidência, de uma forma proporcional à verossimilhança de cada rótulo ou classe em particular. O passo seguinte consiste na combinação dos valores estimados para massa de evidência oriundas de duas ou mais fontes de dados. Neste estudo é adotado o método conhecido com *soma ortogonal*. Neste processo os valores para massa de evidência originados por fontes de dados independentes são combinados dois a dois conforme a equação (1):

$$m_{a,b}(z) = \frac{\sum_{x \cap y = z} m_a(x) \cdot m_b(y)}{\sum_{x \cap y \neq \emptyset} m_a(x) \cdot m_b(y)} \quad (1)$$

onde  $m_a(x)$  e  $m_b(x)$  representam as massas de evidência para duas fontes independentes ( $a$ ,  $b$ ) e  $m_{a,b}(z)$  representa a massa de evidência combinada,  $x$ ,  $y$  e  $z$  representam rótulos ou classes. A soma ortogonal (1) pode ser aplicada de uma forma repetitiva, sempre ocorrem mais de duas fontes de dados.

Confiança ( $bel$ ) e plausibilidade ( $plau$ ) conforme definidos acima podem então ser estimado por:

$$bel(X) = \sum m(Y) \forall Y \subseteq X \quad (2)$$

$$plau(X) = 1 - bel(NOT(X))$$

Utilizando os dados do exemplo anterior, os valores estimados para confiança são:

$$bel(\omega_1) = 0.3, \quad bel(\omega_2) = 0.3, \quad bel(\omega_3) = 0.75$$

$$plau(\omega_1) = 0.55, \quad plau(\omega_2) = 0.55, \quad plau(\omega_4) = 0.25$$

Neste estudo é investigado o potencial apresentado pelo conceito de confiança para fins de incremento na acurácia da imagem temática produzida por um classificador. Conforme descrito anteriormente, o valor de confiança serve para estimar o valor mínimo de evidência disponível de que o pixel em questão pertence a uma determinada classe. A informação relativa pode então ser empregada na detecção de erros de inclusão na imagem temática produzida.

A metodologia proposta neste estudo pode ser sumarizada pelas seguintes etapas:

1- Analisando-se a cena a ser classificada, decide-se quais as variáveis que apresentam um potencial de discriminar as classes presentes. Estas variáveis devem ser então estruturadas em *layers* individuais, todos registrados entre si e com os dados imagem da cena.

2- Nesta etapa estima-se o grau de incerteza associado a cada uma das classes presentes na cena, com relação a cada variável individual. Neste particular, adotou-se neste estudo o critério proposto por Lee et al. (1987). Procedeu-se o processo de classificação da cena, empregando-se cada variável individualmente. A incerteza é então estimada por  $1 -$  (valor médio da acurácia do para aquela variável). Deve-se aqui recordar que a acurácia do usuário estima o grau de confiança na imagem temática produzida pelo classificador (Congalton, 1991). As evidências estimadas na etapa anterior são então ajustadas de forma a incluir o valor da incerteza.

3- A etapa seguinte consiste em estimar o valor da evidência para cada pixel, para cada classe, em função de cada variável individual. A evidência pode ser estimada de duas maneiras gerais: (a) descreve-se o comportamento de cada classe para cada variável individual em termos de uma função de probabilidade. A evidência é associada a esta função de probabilidade, (b) na impossibilidade de descrever-se este comportamento por meio de uma função de probabilidade, utiliza-se uma função de pertinência, com base no conhecimento que o analista tem da cena.

- 4- Uma vez obtidos os *layers* de evidência para cada variável individual, procede-se a estimação da evidência combinada, empregando-se para esta finalidade o algoritmo da soma ortogonal (1), iterativamente, de forma a abranger todas as variáveis.
- 5- Com os valores de evidência globais, combinados por todas as variáveis, estimam-se os valores de confiança (*bel*), organizados na forma de um *layer* adicional.
- 6- Utiliza-se o *layer* de confiança para identificar na imagem temática erros de inclusão, que serão então excluídos.

Neste estudo é proposto um critério para escolha de um valor para o limiar. Aplicado ao *layer* de confiança, este limiar servirá para filtrar pixels erroneamente rotulados. Na medida em que o valor atribuído ao limiar aumenta, um número crescente de pixels rotulados na imagem temática, mais especificamente aqueles que apresentam um valor para a confiança inferior ao estabelecido pelo limiar, são removidos desta, restando somente pixels com um valor mais alto para a confiança. O problema que surge na seleção do limiar pode ser resumida da seguinte forma: na medida em que o valor do limiar aumenta, somente pixels rotulados e com um valor crescente para a variável confiança permanecem na imagem temática. Este fato resulta em um crescente aumento da acurácia do usuário. Este incremento no limiar resulta entretanto em um crescente número de pixels corretamente rotulados que passam a ser eliminados da imagem temática, resultando em um decréscimo no valor da acurácia do produtor. Neste estudo propõe-se que o valor adotado para o limiar seja aquele que maximiza média entre as acurácias do usuário e do produtor.

### 3. Experimento

A metodologia acima proposta foi avaliada utilizando uma área teste localizada na região nordeste do Estado do Rio Grande do Sul. O software necessário para a implementação da metodologia proposta foi desenvolvido em ambiente MATLAB. Nesta região são encontradas áreas cobertas por Mata Nativa com Araucária. Experiências anteriores mostraram que esta classe de cobertura do solo é difícil de ser identificada com boa acurácia, sempre que dados multi-espectrais de Sensoriamento Remoto são empregados, ocorrendo confusão com outras classes. Neste experimento visou-se, portanto implementar a metodologia proposta para aumentar a acurácia na identificação de manchas de Mata Nativa com Araucária. Utilizando-se as seis bandas refletivas coletadas pelo sistema Landsat-TM, procedeu-se a uma classificação supervisionada da cena, empregando-se o classificador Máxima Verossimilhança Gaussiana (MVG). Amostras de teste disponíveis permitiram estimar o valor da acurácia do usuário em 72.09%. Uma simples inspeção visual na imagem temática mostrou a presença de vários pixels indevidamente rotulados como Mata Nativa com Araucária, resultado em um valor estimado reduzido para a classe em questão. A metodologia descrita foi aplicada aqui com a finalidade de aumentar a acurácia na imagem temática produzida. No problema foram consideradas duas classes: Mata Nativa com Araucária ( $\omega_1$ ) e “outros” ( $\omega_2$ ). Seguindo a metodologia proposta, foram inicialmente identificadas as variáveis com poder discriminante em relação à classe  $\omega_1$ : elevação do terreno, temperatura média anual, precipitação média anual, classe de solos, NDVI, fração de sombra e fração de vegetação. No contexto deste problema, dados de textura são relevantes. Com esta finalidade a imagem foi segmentada e atributos de textura estimados para cada segmento. Neste processo foi utilizado neste estudo o software eCognition com as seguintes variáveis: contraste, correlação, dissimilaridade, entropia, homogeneidade, média, momento angular e desvio padrão. Estas 16 variáveis foram digitalizadas na forma de *layers* e registrados com a imagem Landsat-TM. A etapa dois consistiu em estimar o valor da incerteza associada a cada variável. Nesta etapa foi aplicado o critério proposto por Lee et al. (1987). Empregando-se cada uma das variáveis individualmente, foi feita uma classificação da cena e estimada em cada caso o valor da

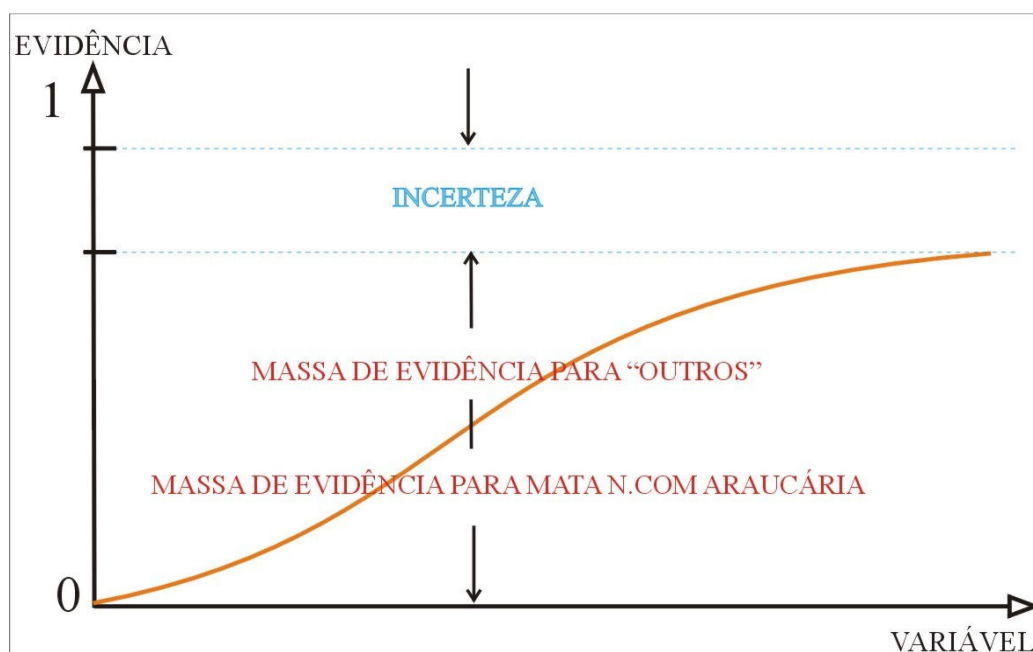
acurácia do usuário. Os resultados estão listados na Tabela 1. Esta etapa permite portanto a estimação do valor de incerteza associado a cada variável individual.

#### Acurácia do Usuário para as Variáveis Utilizadas

Variáveis	Acurácia do Usuário	Variáveis	Acurácia do Usuário
Dados Multi-espectrais	0.897	Contraste	0.338
Elevação do Terreno	0.510	Correlação	0.377
Temperatura Média Anual	0.515	Dissimilaridade	0.372
Precipitação Média Anual	Desprezível	Entropia	0.413
Classes de Solos	Desprezível	Homogeneidade	0.454
NDVI	0.463	Valor Média	0.613
Fração de Sombra	0.513	Momentum Angular	Desprezível
Fração de Vegetação	0.352	Desvio Padrão	0.360

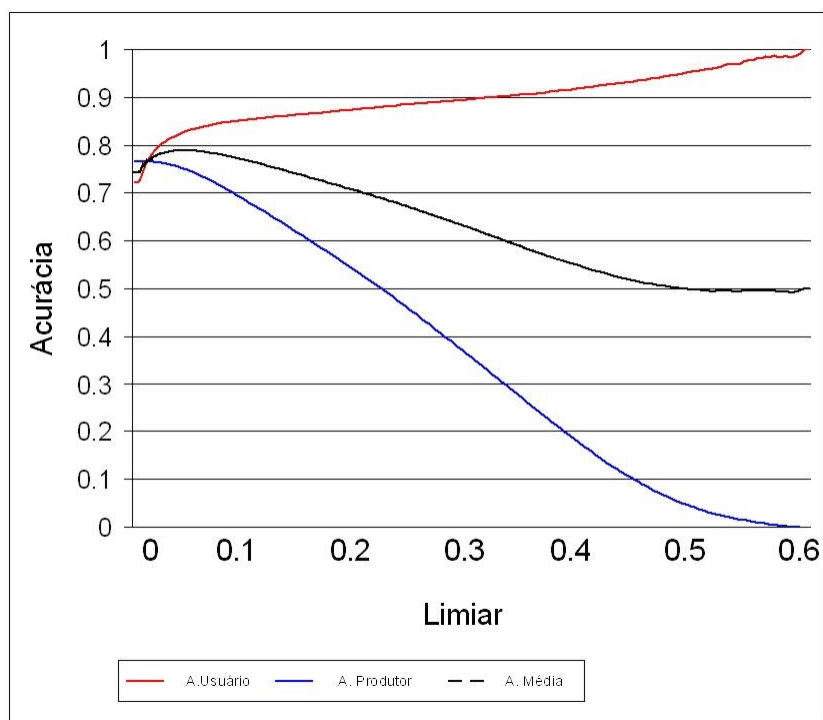
**Tabela 1**

Pode-se observar que três variáveis (precipitação média anual, classe de solo e momento angular) apresentaram um valor desprezível para a acurácia do usuário, o que equivale a um valor para a incerteza de 100%. Esta três variáveis foram portanto descartadas. Nesta etapa foi ainda incluída a variável “dados multi-espectrais”, com a finalidade de permitir a estimação da evidência associada com os dados Landsat-TM empregados. A terceira etapa consistiu portanto na estimação das evidências apoiando a classe Mata Nativa com Araucária. As variáveis para as quais a classe  $\omega_1$  pode ser razoavelmente modelada pela distribuição normal, a função de Gauss foi utilizada. Para as demais (elevação do terreno e temperatura média anual), foi adotada funções de pertinência utilizando o conhecimento do analista. Em ambos os casos, os processo de estimação das evidências associadas as classes  $\omega_1$  e  $\omega_2$  encontra-se ilustrado na Figura 1. Uma vez formados os *layers* com as evidências associadas as duas classes, o algoritmo da soma ortogonal (1) é aplicado iterativamente, de modo a se obter a evidência combinada de todas as variáveis. A etapa 5, que consiste na elaboração do *layer* de confiança pode então ser efetivada conforme (2).



**Figura 1**

O *layer* de confiança foi então utilizado para fins de filtragem de pixels erroneamente rotulados na imagem temática produzida pelo classificador MVG. Foi feito um estudo envolvendo diferentes valores para o limiar e as resultantes acurácias do usuário e do produtor foram estimadas a partir de um segmento da cena reservado como área de teste. As curvas ilustrando os valores estimados para as acurácias do usuário, do produtor e média, estão ilustradas na Figura 2. Conforme o esperado, pode-se observar que um valor baixo para o limiar permite a manutenção na imagem temática de um numero grande de erros de inclusão, o que se traduz por um valor baixo para a acurácia do usuário. Na medida em que o valor do limiar aumenta, pixels com baixo valor para confiança e, portanto possivelmente rotulados erroneamente como Mata Nativa com Araucária são removidos resultando em um incremento na acurácia do usuário. O aumento no valor do limiar aumenta, entretanto a possibilidade de que pixels corretamente rotulados como Mata Nativa com Araucária sejam também eliminados, resultando em um decremento na acurácia do produtor. Em outras palavras, o aumento no valor do limiar diminui o número de erros de inclusão, aumentando assim a acurácia do usuário, mas aumenta o número de erros de omissão, diminuindo assim a acurácia do produtor.



Limiares para a acurácia do usuário, produtor e média

**Figura 2**

A escolha de um valor adequado para o limiar de confiança a ser aplicado à totalidade da imagem temática deve seguir um certo critério. Os dados ilustrados na Figura 2 podem, portanto ser utilizados pelo analista com vistas a uma decisão final que melhor atenda as finalidades que se tem em vista. Neste estudo é proposto o valor para o limiar que maximiza a acurácia média. No presente experimento, este valor para o limiar encontra-se em torno de 0.041 (Figura 2) resultando em uma acurácia do usuário igual a 82.64%, um acurácia do produtor igual a 75.19% e uma acurácia média igual a 78.91%. Os valores correspondentes para estas três acurácias conforme a imagem temática produzida diretamente pelo classificador MVG são respectivamente iguais a 72.30%, 76.65% e 74.48%. Obteve-se,

portanto um significativo incremento na acurácia do usuário e na acurácia média enquanto que um pequeno decréscimo ocorreu na acurácia do produtor.

## **Conclusão**

No presente estudo foi descrita e testada uma metodologia que implementa conceitos propostos pela Teoria da Evidência para incrementar a acurácia no processo de classificação digital de imagens. A contribuição maior encontra-se na proposta de uma abordagem mais quantitativa quanto à utilização da informação fornecida pela variável confiança. Os resultados obtidos no experimento servem para ilustrar a aplicabilidade da metodologia proposta.

## **Referências:**

- Bruzzone, L., Conese, C., Maselli F., and Roli, F., 1997. Multisource classification of complex rural areas by statistical and neural network approaches. *Photogrammetric Engineering and Remote Sensing*, 63(5), pp. 523-533.
- Congalton, R., 1991. A review of assessing the accuracy of classification of remotely sensed data. *Remote Sensing of Environment*, 37, pp. 35-46
- Eastman, J. R., 1999. *Guide to GIS and Image Processing*, vol. 2, Clark Labs, Clark University.
- Gong, P., 1996. Integrated analysis of spatial data from multiple sources: using evidential reasoning and artificial neural network techniques for geological mapping. *Photogrammetric Engineering & Remote Sensing*, 62(5), pp. 513-523.
- Hutchinson, C.F., 1982. Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing*, 48(1), pp. 123-130.
- Lee, T., Richards J. A., and Swain, P. H., 1987. Probabilistic and evidential approaches for multisource data analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 25(3), pp. 283-293.
- Moon, W. M., (1990). Integration of geophysical and geological data using evidential belief function, *IEEE Transactions on Geoscience and Remote Sensing*, 28(4), pp. 711-720.
- Richards, J. A. and Jia, X., 1999. *Remote Sensing Digital Image Analysis, an Introduction*, 3<sup>rd</sup> edition, Springer, Berlin.