

MÉTODOS DE AVALIAÇÃO DE MODELOS DE CLASSIFICAÇÃO DE IMAGENS DIGITAIS

ANTONIO JOSÉ FERREIRA MACHADO E SILVA

CCRio - Centro Científico Rio - IBM Brasil
Av. Presidente Vargas 824 - 22º andar
20071-001, Rio de Janeiro, RJ, Brasil
amachado@riovmsc.vnet.ibm.com
amachado@riosc

Abstract. This paper describes some methodologies used in the evaluation of parametric and gaussian classifiers. In the first step, a theoretical study is presented, and synthetical gaussian images are used to confirm the studies. After this, the Amazon Region was selected to test the models, and five images with visual interpretation were used to evaluate the classifiers.

1. Introdução

O Instituto Nacional de Pesquisas Espaciais (INPE), órgão ligado à Secretaria de Ciência e Tecnologia, é um dos responsáveis pelo monitoramento da Amazônia. Para essa tarefa, o INPE faz uso das modernas técnicas de sensoriamento remoto, empregando imagens do sensor TM-LANDSAT.

Até o presente ano, o INPE emprega técnicas de foto-interpretação para realizar o monitoramento. São geradas centenas de imagens analógicas, em composição colorida, que são analisadas visualmente por especialistas (engenheiros florestais, agrônomos, geólogos, biólogos, etc.). Existe, no entanto, a vontade de se saber que tipo de ajuda o computador, e mais especificamente, as técnicas de processamento digital de imagens, podem dar para agilizar as tarefas, mantendo, ou quando possível melhorando, a qualidade e precisão dos resultados.

Em função disso, o INPE e a IBM Brasil, através do Centro Científico Rio (CCRio), firmaram um acordo, para desenvolvimento de um projeto em parceria, voltado para o estudo ambiental da Amazônia. Uma das etapas

deste projeto contempla pesquisas de algoritmos e metodologias para extração de informações a partir das imagens digitais do sensor TM-LANDSAT.

A fase de extração de informações, ou de classificação de imagens, é a responsável por segmentar a região Amazônica em classes homogêneas, identificando as florestas, os desflorestamentos, os cerrados, as regiões aquáticas, urbanas, etc. O objetivo é conseguir automatizar este processo que hoje é feito visualmente, substituindo as imagens analógicas por imagens digitais. Caso não seja possível um método totalmente automático, que se consiga um processo assistido por computador.

Para esta fase, procurou-se levantar o estado da arte, avaliando-se as principais tendências para classificação de imagens. No entanto, surgiu um problema relacionado com a avaliação dos modelos de classificação. Num primeiro passo, decidiu-se por avaliar os modelos paramétricos gaussianos, começando pelo questionamento do fundamento destes modelos: "Terão as classes um comportamento gaussiano?". Por enquanto, a base do estudo está concentrada em exemplos obtidos na região amazônica.

O processo de avaliação é difícil, principalmente pela necessidade de verificação da verdade terrestre. É comum encontrar avaliações relativas que comparam métodos entre si. Neste trabalho, procura-se buscar formas alternativas de avaliação dos modelos, fazendo uso de imagens sintéticas que apresentem características semelhantes às imagens reais, bem como resultados de interpretação visual, quando o conhecimento do especialista é levado em consideração.

2. A Área de Estudo

O trabalho concentra seus resultados em exemplos obtidos na região amazônica. Foram selecionadas 5 áreas que representam a diversidade de situações próprias da região: grandes áreas de florestas; áreas de desflorestamento, como as áreas conhecidas como espinha de peixe em Rondônia; regiões de cerrado (não floresta); regiões em torno de grandes represas (água), etc. As imagens utilizadas estavam distribuídas pelos estados do Pará, Acre, Rondônia, Mato Grosso, e a represa escolhida foi a de Tucuruí. Os codinomes das imagens, bem como as dimensões delas são: PA(768,768), AC(768,768), RO(768,768), MT(1000,1000) e TU(512,512).

As imagens, conforme citado anteriormente, são do sensor TM-LANDSAT, e nesta primeira fase apenas as bandas 3, 4 e 5 foram utilizadas, por serem as mesmas empregadas até aqui pelo INPE. Para estas imagens, o INPE forneceu "gabaritos" de classificação, que constam de imagens segmentadas de forma automática e rotuladas por especialistas, a partir de interpretação visual.

Foram definidas 7 classes, sendo 4 denominadas "limpas" e 3 ruidosas. As classes "limpas" são: floresta (FL), não

floresta (NF), água (AG) e desflorestamento (DF). As classes ruidosas são: nuvem (NV), sombra (SO) e bruma (BR). Para cada segmento, o especialista atribuiu uma nota de 1 a 5 em relação a cada possível classe. A nota 1 corresponde a possibilidade total e a 5 a que não há nenhuma possibilidade. Por exemplo: (1,5,5,5,5,5,5) significa que o segmento é floresta com alto grau de possibilidade, enquanto que (2,4,5,5,5,5,5) já indica que há possibilidade do segmento ser da classe não floresta. O conjunto de notas (5,3,5,3,5,5,5) indica dúvida em relação às classes não floresta e desflorestamento. Só foram analisados os segmentos relativos às classes limpas, e que apresentaram alto grau de possibilidade.

3. O Paradigma das Classes Gaussianas

Diversos modelos paramétricos fazem uso da premissa que as funções densidade de probabilidade relativa a alvos em imagens de sensoriamento remoto são gaussianas. Para avaliar esta premissa, levantou-se o histograma relativo a cada classe, por banda (histograma unidimensional), e procurou-se a gaussiana que melhor se ajustava a este histograma no sentido do mínimo erro quadrático. Na realidade o ajuste foi realizado em cima da função distribuição de probabilidade.

$$\hat{F}_x(X) = \sum_i^{\bar{X}} h(i), \text{ onde}$$

$$\bar{X} = \text{maior inteiro} \leq X$$

$$h(i) = \text{freq. relat. do nív. de cinza } i$$

$$F_x(X) = \int_{-\infty}^X f_x(\alpha) d\alpha$$

$$f_x(X) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(X-m)^2}{2\sigma^2}\right\}$$

$$\text{Erro} = \int_{-\infty}^{\infty} (F_x(X) - \hat{F}_x(X))^2 dX$$

No processo de minimização do erro médio quadrático, determina-se novos parâmetros de média (m) e variância (σ^2), que são então comparados com os obtidos das imagens.

Para melhor compreender os resultados, foram gerados dados sintéticos gaussianos com as médias e variâncias de cada classe, por banda (12 pares). Estes dados passaram pelo mesmo processo de ajuste e cálculo de erro. A tabela 1 apresenta os resultados para alguns casos (BiCj significa banda i e classe j).

	Δm	$\Delta\sigma^2\%$	Erro
C1B4	0,00	0,4	0,22
C2B4	0,00	0,3	0,68
C3B3	0,00	0,6	0,39
C4B5	0,11	0,6	0,13

Observando os dados reais para os mesmos casos, nota-se uma piora no desempenho, representada pela maior discrepância em todos os parâmetros: média, variância e erro. A tabela 2 resume estes resultados.

	Δm	$\Delta\sigma^2\%$	Erro
C1B4	-0,20	-4,7	0,44
C2B4	0,14	-8,9	1,28
C3B3	-0,44	-4,0	10,73
C4B5	-0,08	0,1	0,73

No geral, os resultados são satisfatórios, mas ainda não são

conclusivos quanto à eficiência de considerá-los gaussianos a título de classificação.

4. Estudo Teórico da Probabilidade de Erro

Qualquer processo não determinístico de tomada de decisão tem um risco associado. O risco pode ser tomado como o valor esperado do custo. No caso de classificação de imagens de satélite, uma imagem é dividida em NC classes, e para cada pixel tem-se que decidir por uma das classes (na realidade os algoritmos permitem que nenhuma decisão seja tomada). Seja $P(W_j|W_i)$ a probabilidade de se decidir pela classe W_j e o pixel pertencer à classe W_i . Um erro ocorrerá sempre que i for diferente de j . Supondo que os custos sejam nulos quando não ocorra erro e unitário nos demais casos, o risco passa a se confundir com a probabilidade de erro ($P(\varepsilon)$). Utilizando a regra de Bayes, e sabendo que a probabilidade de erro é igual a diferença entre 1 e a probabilidade de acerto, tem-se:

$$P(\varepsilon) = 1 - \sum_i^{NC} P_i \cdot \int_{Z_i} f_{x/w_i}(X) dX,$$

onde $f_{x/w_i}(X)$ é a função densidade de probabilidade de um pixel, dado que ele pertença à classe W_i . Deve-se destacar que neste ponto o pixel X pertence ao espaço NB -dimensional, onde NB é o número de bandas das imagens. Z_i corresponde a partição do citado espaço NB -dimensional onde a função discriminante relativa à classe W_i é maior que todas as demais, enquanto que P_i é a probabilidade de um pixel pertencer à classe W_i .

As funções discriminantes d_i são baseadas nas próprias funções densidade

de probabilidade condicionadas às classes. No caso dos classificadores gaussianos paramétricos supervisionados a função discriminante ótima é aquela que corresponde a decidir pela classe W_i sempre que:

$$P(W_i|X) > P(W_j|X) \quad \forall j \neq i$$

Utilizando a regra de Bayes, a expressão anterior é similar a:

$$P(W_i) \cdot f_x(X|W_i) > P(W_j) \cdot f_x(X|W_j)$$

No caso de classes gaussianas tem-se:

$$f_x(X|W_i) = \frac{\exp\{-\frac{1}{2} \cdot (X - m_i)' \Sigma_i^{-1} (X - m_i)\}}{(2 \cdot \pi)^{\frac{NB}{2}} \cdot |\Sigma_i|^{\frac{1}{2}}}$$

Quando existe conhecimento sobre as probabilidades a priori de cada classe a função discriminante ($d_1(X)$) está associada à decisão pelo critério de máxima probabilidade a posteriori (MAP). Quando não é possível estabelecer tais probabilidades, recai-se no critério de Máxima Verossimilhança (MaxVer: $d_2(X)$). Simplificações comumente utilizadas consideram a matriz covariância como diagonal (componentes descorrelatados: $d_3(X)$) e até matriz covariância diagonal e com todas as variâncias idênticas ($d_4(X)$). Nestes casos, as respectivas funções discriminantes são:

$$d_{1i}(X) = \ln(P(W_i)) - \frac{1}{2} \ln(|\Sigma_i|) - \frac{1}{2} (X - m_i)' \Sigma_i^{-1} (X - m_i)$$

$$d_{2i}(X) = -\ln(|\Sigma_i|) - (X - m_i)' \Sigma_i^{-1} (X - m_i)$$

$$d_{3i}(X) = -\sum_{j=1}^{NB} \ln(\sigma_{ij}) - \sum_{j=1}^{NB} \frac{(X_j - m_{ij})^2}{2 \cdot \sigma_{ij}^2}$$

$$d_{4i}(X) = -\sum_{j=1}^{NB} (X_j - m_{ij})^2$$

Em todos os casos, decide-se pela classe W_i sempre que $d_{K_i}(X) > d_{K_j}(X)$, $\forall i \neq j$. A última função discriminante corresponde à tomada de decisão pela mínima distância euclidiana.

É importante procurar compreender o porquê de certas simplificações. Os processos de classificação são custosos do ponto de vista computacional, pois para cada pixel requerem uma grande quantidade de operações. As simplificações visam diminuir os custos e agilizar o processo. Outras simplificações podem ocorrer por desconhecimento de alguns parâmetros, como por exemplo as probabilidades a priori. A tabela 3 apresenta o número de operações de multiplicação (*) e de adição (+) inerentes a cada função discriminante.

	+	*
d1	NB ² + NB	NB ² + 2.NB
d2	NB ² + NB	NB ² + 2.NB
d3	2.NB - 1	3.NB
d4	NB	NB

Tabela 3: número de operações por função discriminante.

Para situações similares a das 4 classes "limpas" em questão, cujos vetores média e matrizes de covariância para as bandas 3, 4 e 5 são apresentados

nas tabelas 4 a 7 (primeira linha refere-se à média e as demais à matriz covariância) realizou-se a determinação teórica do erro, para as diversas funções discriminantes, já levando em consideração a discretização do processo em valores inteiros de 0 a 255. Considerou-se a quantização como linear. Os pesos de cada classe também respeitaram a situação real. O melhor (PA) e o pior (MT) resultados estão apresentados na tabela 8, em termos de probabilidade de erro.

m	20,4	54,2	45,6
σ	2,2	-0,4	3,8
σ	-0,4	27,6	3,0
σ	3,8	3,0	26,1

Tabela 4: vetor média e matriz covariância da classe FL.

m	26,4	44,5	65,4
σ	8,7	-2,3	24,3
σ	-2,3	11,7	-10,5
σ	24,3	-10,5	101,5

Tabela 5: vetor média e matriz covariância da classe NF.

m	19,6	30,2	22,7
σ	35,9	123,5	97,4
σ	123,5	550,7	403,4
σ	97,4	403,4	368,5

Tabela 6: vetor média e matriz covariância da classe AG.

m	36,2	48,0	99,9
σ	57,4	23,9	122,3
σ	23,9	36,6	44,1
σ	122,3	44,1	388,5

Tabela 7: vetor média e matriz covariância da classe DF.

	MT	MT	PA	PA
	Teor.	Sint.	Teor.	Sint.
d1	5,9	5,9	1,5	1,5
d2	6,2	6,2	1,7	1,7
d3	7,8	7,9	5,1	5,7
d4	8,5	8,7	5,5	6,2

Tabela 8: probabilidade de erro: estudo teórico e imagem sintética.

Os resultados alcançados foram surpreendentemente bons, devidos sobretudo aos baixos valores de variância encontrados. A probabilidade de erro diminui com o aumento das distâncias entre as médias das classes e também com a redução dos valores de variância. Como uma breve ilustração, considere-se o caso unidimensional com 2 classes apresentando mesma variância e equiprováveis. A probabilidade de erro é dado por $Q(d/2\sigma)$, onde Q é uma função decrescente, ou seja quando d (distância entre as médias das classes) cresce ou quando a a variância diminui, a probabilidade de erro decresce.

Para observação do estudo teórico, optou-se pela geração de imagens sintéticas compostas por classes gaussianas de médias e matrizes covariância iguais às encontradas nas imagens. O processo de geração destas imagens (X) é inverso ao empregado na transformação de principais componentes. Inicialmente, gera-se dados gaussianos

(Z) de média nula e matriz covariância igual a identidade. Estipula-se o vetor média (m) e a matriz covariância que se deseja para os dados e, para esta matriz calcula-se o vetor de autovalores (λ) e a matriz de autovetores (Δ). Daí tem-se:

$$Y = Z \cdot \lambda \quad \text{tq} \quad Y_i = Z_i \cdot \lambda_i$$

$$X = \Delta^{-1} \cdot Y + m$$

O peso de cada classe é rigorosamente o mesmo das imagens reais, e os resultados mostrados na tabela 8 falam por si só: muito próximos aos valores teóricos e rigorosamente os mesmos para os modelos MAP e MaxVer.

Estes resultados indicam que se as classes nas imagens reais tiverem um comportamento gaussiano, os modelos estudos serão eficientes para classificar a imagem.

5. Avaliação das Imagens

Para avaliar os modelos sobre as imagens, considerou-se apenas os segmentos com alto grau de possibilidade para uma das 4 classes limpas. Infelizmente, nenhuma das 5 imagens apresentou as 4 classes simultaneamente com tais características.

Para determinação dos parâmetros necessários, optou-se por determiná-los a partir de todos os pixels relativos aos segmentos em questão, eliminando a influência do treinamento. Posteriormente serão analisadas as influências do treinamento, bem como a classificação de toda imagem, incluindo as regiões ruidosas.

As tabelas 9 a 12 apresentam os resultados para a imagem sobre o estado de Mato Grosso (MT), onde as classes presentes são: floresta (1, peso 51,9%), não floresta (2, peso 24,5%) e desflorestamento (3, peso 23,6%). As tabelas de 13 a 16 apresentam os

resultados para a imagem sobre o estado do Pará (PA). Neste caso, as classes presentes são: floresta (1, peso 61,4%), água (2, peso 20,3%) e desflorestamento (3, peso 18,3%). Um ponto na linha I e coluna J destas tabelas indica a percentagem de pixels que estão na classe I na imagem "gabarito" e na classe J na imagem classificada. Assim sendo, a quarta e última coluna apresenta os pesos de cada classe na imagem "gabarito", enquanto que a quarta e última linha indica os mesmos pesos na imagem classificada. O somatório dos pontos da diagonal principal (I=J) indica a percentagem de pixels que não sofreram alterações na classificação.

50,3	1,4	0,2	51,9
0,8	22,1	1,6	24,5
0,1	3,1	20,4	23,6
51,2	26,6	22,2	100,0

Tabela 9: migração de pixels na imagem MT (discriminante d1). Erro = 7,2%

49,7	2,1	0,1	51,9
0,4	22,3	1,8	24,5
0,1	3,1	20,4	23,6
50,2	27,5	22,3	100,0

Tabela 10: migração de pixels na imagem MT (discriminante d2). Erro = 7,6%

49,9	2,0	0,0	51,9
1,4	21,6	1,5	24,5
0,1	4,0	19,5	23,6
51,4	27,6	21,0	100,0

Tabela 11: migração de pixels na imagem MT (discriminante d3). Erro = 9,0%

50,6	1,3	0,0	51,9
1,5	21,3	1,7	24,5
0,1	4,8	18,7	23,6
52,2	27,4	20,4	100,0

Tabela 12: migração de pixels na imagem MT (discriminante d4). Erro = 9,4%

60,1	1,0	0,3	61,4
4,2	16,0	0,1	20,3
1,6	0,1	16,6	18,3
65,9	17,1	17,0	100,0

Tabela 16: migração de pixels na imagem PA (discriminante d4). Erro = 7,6%

60,1	0,7	0,6	61,4
1,0	19,0	0,3	20,3
0,7	0,2	17,4	18,3
61,8	19,9	18,3	100,0

Tabela 13: migração de pixels na imagem PA (discriminante d1). Erro = 3,5%

59,6	0,8	1,0	61,4
0,9	19,1	0,3	20,3
0,5	0,1	17,7	18,3
61,0	20,0	19,0	100,0

Tabela 14: migração de pixels na imagem PA (discriminante d2). Erro = 3,6%

60,0	1,0	0,4	61,4
3,9	16,2	0,2	20,3
1,4	0,1	16,8	18,3
65,3	17,3	17,4	100,0

Tabela 15: migração de pixels na imagem PA (discriminante d3). Erro = 7,0%

Os resultados encontrados são plenamente satisfatórios, e corroboram a premissa de considerar as classes como gaussianas. É claro, que os resultados mostrados se detiveram nos segmentos limpos, mas deve-se destacar que estes representam de 80 a 95% das imagens.

6. Considerações Finais

Este trabalho apresenta resultados preliminares de um estudo mais amplo, que visa determinar os parâmetros das diversas classes, bem como o número de classes, por um processo de ajuste de um somatório ponderado de gaussianas ao histograma da imagem (semelhante ao processo apresentado na seção 3). Desta forma, eliminar-se-ia a etapa de treinamento.

De qualquer forma, as etapas futuras contemplarão o estudo da influência do treinamento nestes resultados. As classes ruidosas, bem como as de indefinição, serão incorporadas ao processo, que neste caso passará a poder responder que está indeciso entre duas ou mais classes, no momento da classificação. As demais bandas também serão consideradas, principalmente se o estudo teórico dos erros assim o indicar.

Outro ponto importante a ser analisado é o dos custos de uma classificação incorreta. Talvez não seja justo considerar que o custo de classificar floresta como desflorestamento seja o mesmo que classificar não floresta como desflorestamento.

É importante destacar, ainda, que o "gabarito" nasce de uma segmentação automática, que com certeza apresenta suas distorções. O ideal seria eliminar etapas automáticas na geração deste "gabarito".

Os resultados obtidos nas imagens reais estiveram tão próximos dos valores teóricos e dos valores encontrados em imagens sintéticas, e com magnitude tão pequena que não houve necessidade de comparação direta entre os diversos modelos de classificação.

7. Referências

- Duda, R.O.; Hart, P.E. Pattern Classification and Scene Analysis. John Wiley and Sons, 1973.
- Fukunaga, K. Introduction to Statistical Pattern Recognition. Academic Press, 1972.
- Richards, J.A. Remote Sensing Digital Image Analysis: An Introduction. Springer-Verlag, 1986.
- Van Trees, H.L. Detection, Estimation, and Modulation Theory. John Wiley and Sons, 1968.
- Wozencraft, J.M.; Jacobs, I.M. Principles of Communication Engineering. John Wiley and Sons, 1965.