

# SELEÇÃO DE BANDAS PARA CLASSIFICAÇÃO DE IMAGENS

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@rioivmsc.vnet.ibm.com

amachado@riosc.bitnet

**Abstract.** This paper describes some methodologies used for select the best combination of 3 spectral bands of TM-LANDSAT for image classification in the Amazon region. The mean vector, the covariance and correlation matrices, the entropy and the histogram of each class are the parameters used to make the decision. The distance and divergence between the classes and the error probability for the maximum a priori probability classifier are calculated in order to make easy this decision.

## 1. Introdução

As imagens de satélite são hoje uma importante ferramenta para o estudo do meio ambiente. É comum que estas imagens sejam multiespectrais, o que significa que de uma mesma região são geradas diversas imagens, em diferentes faixas do espectro eletromagnético (bandas espectrais).

Por sua vez, a classificação de imagens é um dos processos mais importantes para análise de informações contidas nas imagens digitais. Muitos dos algoritmos de classificação têm o número de operações aritméticas proporcional ao número de bandas (distância euclidiana) ou até mesmo ao quadrado do número de bandas (máxima verossimilhança).

É sabido, no entanto, que algumas bandas têm forte correlação, decorrente da redundância de informação contida nelas. Existem transformações que procuram diminuir, e até eliminar, esta redundância, tais como transformadas canônicas, transformada de Karhunen-Loeve (principais componentes), Tasseled-Cap, etc. (Richards, 1986).

Ou seja, a seleção apropriada de bandas pode representar uma enorme redução nos custos computacionais, sem que isso acarrete uma queda acentuada na qualidade do classificador.

O Instituto Nacional de Pesquisas Espaciais (INPE) vem realizando, anualmente, o monitoramento ambiental da Amazônia, a partir de imagens do sensor TM-LANDSAT. Este sensor gera imagens em 7 bandas espectrais, cobrindo a região do visível (bandas

1, 2 e 3) e do infra vermelho (bandas 4, 5, 6 e 7). A banda 6, na região do infra vermelho termal tem uma resolução espacial maior, e portanto não será utilizada. Maiores informações sobre a configuração do sensor TM-LANDSAT pode ser visto em Novo (1988).

O INPE, no seu monitoramento, utiliza as bandas 3 (visível - vermelho), 4 (infra vermelho próximo) e 5 (infra vermelho médio). Uma redução desta natureza (de 6 para 3 bandas) reduz o número de operações aritméticas em 75%, no caso de classificadores que têm o número destas operações proporcional ao quadrado do número de bandas, como por exemplo o classificador de máxima probabilidade a posteriori (Machado e Silva, 1993).

O Centro Científico Rio (CCRio) da IBM Brasil firmou um projeto de parceria com o INPE para pesquisar novas tecnologias que possam ser incorporadas ao processo de monitoramento ambiental da Amazônia. Uma das linhas de pesquisa é a de extração de informações. Dentro deste escopo, este trabalho procura verificar quais seriam as 3 bandas mais indicadas para serem utilizadas no processo de classificação de imagens na Amazônia.

Deve-se ter em mente que os critérios aqui discutidos para a seleção de bandas estão relacionados aos classificadores que utilizam os parâmetros de média, matriz covariância e peso das classes. Como exemplo, pode-se citar os classificadores de máxima verossimilhança, de máxima probabilidade a posteriori, de mínima distância euclidiana, o K-médias e o

ISODATA. Atributos geométricos, de textura, ou outros, não são considerados neste trabalho.

Numa etapa inicial, pretende-se dividir a região em 4 classes: floresta (FL), não floresta (NF), água (AG) e desflorestamento (DF). Além dessas, encontram-se outras classes espúrias, tais como: nuvem, sombra e bruma.

A análise para seleção de bandas será feita em três partes. Na primeira, as bandas são vistas no global, sem se considerar as características próprias de cada classe. Deste modo, parâmetros importantes são: entropia e variância das bandas, matriz de correlação e o peso de cada banda nas principais componentes. Na fase seguinte, as características das classes, em cada banda, são levadas em consideração para avaliar o poder de discriminação de cada banda. Neste ponto, média, variância e o histograma de cada classe, relativo a cada banda, são os parâmetros observados. Na última fase, observa-se os conjuntos de bandas 3 a 3 (são 20 combinações possíveis). As distâncias entre os centros das classes e a probabilidade de erro, dado que as classes são representadas por amostras com distribuição gaussiana, são analisados, agregando as informações obtidas anteriormente. Nesta fase, utiliza-se o vetor média, a matriz de covariância e o peso de cada classe.

## 2. Observando as bandas

Nesta primeira fase, foram selecionadas duas cenas sobre a Amazônia, e observadas a média, as matrizes de covariância e de correlação, bem como a entropia de cada banda (tabelas 1 a 3 e 6 a 8).

Observando os valores encontrados, nota-se uma certa proporcionalidade entre o logaritmo da variância e a entropia. Deve-se ressaltar que a entropia é uma das principais medidas de informação para imagens digitais. Estes parâmetros dão destaque a banda 5, seguida das bandas 4, 7 e 3. As bandas 1 e 2 são as que apresentam menor quantidade de informação (menor entropia e variância).

A matriz de correlação permite verificar se a informação de uma banda é relevante, ou se é redundante. As tabelas 3 a 8 indicam uma correlação muito forte entre as bandas 1, 2 e 3

(região do visível) bem como entre as bandas 5 e 7 (região do infra vermelho médio). Além disso, fica nítido que existe uma forte correlação entre as bandas destes dois grupos. A banda 4 (infra vermelho próximo) ganha importância por ser a que apresenta os índices de correlação mais baixos com as demais (inclusive variando do positivo para o negativo).

Ou seja, numa primeira análise pode-se inferir que o conjunto de três bandas deve ser composto por uma banda da região do visível (1, 2 ou 3, com destaque até aqui para a banda 3), a banda 4 e uma banda da região do infra vermelho médio (5 ou 7, com destaque para a banda 5).

Outra boa fonte de subsídios para seleção das bandas é a matriz da transformada de Karhunen-Loeve (ou transformada de principais componentes). Esta transformação descorrelata as bandas (matriz de covariância resultante é diagonal), concentrando as informações relevantes nas primeiras bandas, conforme pode ser observado na variância de cada componente. As tabelas 4 e 9 mostram os coeficientes relativos a cada banda para gerar as principais componentes, enquanto que as tabelas 5 e 10 mostram as variâncias de cada uma destas componentes.

Observando as matrizes, verifica-se que a banda 5 é a de maior peso para a componente principal (de maior variância: 78,7 e 93,0% do somatório das variâncias) seguido das bandas 4 e 7. Considerando conjuntamente as 2 primeiras componentes (97,8 e 98,1% de informação) a banda 4 passa a ser a de maior contribuição, seguida da banda 5. E, finalmente, quando se olha o conjunto da 3 principais componentes (99,5 e 99,2% de informação) a banda de maior peso ainda é a 4, seguida das bandas 5 e 3.

## 3. Observando as classes

Nesta etapa, procura-se avaliar o poder de cada banda em separar (discriminar) as classes. Para isto, são levados em consideração as médias, variâncias e os histogramas unidimensionais. Blahut (1990), propõe uma medida de divergência entre duas distribuições de probabilidade, obtidas a partir dos histogramas normalizados:  $h_1$  e  $h_2$ .

$$D = \sum_{i=0}^{255} h_1(i) \log \left( \frac{h_1(i)}{h_2(i)} \right) + h_2(i) \log \left( \frac{h_2(i)}{h_1(i)} \right)$$

Esta medida tem o inconveniente de apresentar problemas quando um nível de cinza ocorrer numa classe e não ocorrer em outra. Neste caso a divergência iria para infinito, independente do comportamento dos demais níveis de cinza. Quando os dois histogramas fossem idênticos, e somente neste caso, a divergência seria zero.

Desta forma, procurou-se definir uma nova medida de divergência que mantivesse a característica de ser zero se e somente se os histogramas fossem iguais, e fosse 1 quando os níveis de cinza das classes fossem disjuntos. No caso contínuo, supondo  $p(x)$  a função densidade de probabilidade, tem-se:

$$Dc = \frac{1}{2} \int_{-\infty}^{\infty} |p_1(x) - p_2(x)| dx$$

No caso de imagens digitais, as funções densidade de probabilidade podem ser substituídas pelos histogramas normalizados, e a integral pode ser substituída pelo somatório:

$$Dd = \frac{1}{2} \sum_{i=0}^{255} |h_1(i) - h_2(i)|$$

As classes são observadas 2 a 2, gerando 6 combinações possíveis, para cada banda. As tabelas 11 e 12 mostram os resultados da divergência, respectivamente para os casos contínuo e discreto. No caso contínuo, considerou-se distribuições gaussianas univariadas.

Observando os resultados, as bandas 5 e 7 aparecem como as que têm maior poder de discriminar as classes. A banda 3 apresenta resultado bastante satisfatório, vindo logo a seguir a banda 4. As bandas 1 e 2 são as que apresentam os piores resultados.

#### 4. Observando as bandas 3 a 3

Uma vez analisados os casos em que as bandas são observadas isoladamente, e as classes são

vistas por banda, procura-se agora analisar a melhor combinação de 3 bandas para classificar imagens da Amazônia em floresta, não floresta, água e desflorestamento. Neste caso são utilizados os vetores de média, as matrizes de covariância e os pesos de cada classe.

Uma primeira observação é realizada em função das distâncias entre os centros (médias) das classes. São 6 distâncias, para 20 combinações de 3 bandas possíveis. A tabela 13 apresenta os resultados. Quanto maiores forem as distâncias, mais fácil é separar as classes. Volta-se a enfatizar que supõe-se que a média das amostras de cada classe está sempre presente na função discriminante.

Uma análise nas distâncias permite dizer que todas as combinações que incluem as bandas 4 e 5 apresentam resultados semelhantes e bastante superior aos demais. As combinações que não incluem nem a banda 4 nem a banda 5 apresentam resultados bastante ruins. Estes resultados reforçam a presença das bandas 4 e 5 no melhor conjunto de 3 bandas.

Uma última análise é realizada em função da probabilidade de erro do classificador de máxima probabilidade a posteriori. 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  $j$ . 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  $W_i$ .

A tabela 15 apresenta os resultados da probabilidade de erro para as situações de probabilidade de ocorrência de cada classe apresentada na tabela 14. Assumiu-se que as classes são representadas por amostras com distribuição gaussiana multivariada.

As combinações que incluem as bandas 1 e 2 apresentam resultados fracos quando comparados com aquelas que incluem as bandas 4 e 5. A combinação que apresenta melhor resultado é a 345, seguida da combinação 347. Surpreendentemente, a combinação 245 também apresenta resultado satisfatório.

### 5. Considerações Finais

Conforme previsto desde o início, a combinação 345 é a que apresenta melhores resultados para discriminar classes na Amazônia. As bandas 1 e 2 podem ser totalmente excluídas da composição, e as demais combinações (347, 357 e 457) apresentam desempenho satisfatório.

A grande vantagem da combinação 345 é agregar uma banda da região do visível (3), uma da região do infra vermelho próximo (4) e uma da região do infra vermelho médio (5).

Analisando isoladamente, a banda 4 deve constar de qualquer combinação que se deseje formar, pois é a banda que apresenta os menores índices de correlação com as demais, associado a um grande poder de discriminação.

Numa próxima etapa, pretende-se comparar o desempenho da classificação usando 3 e 4 bandas, bem como a classificação a partir de bandas resultantes de transformadas, como a de Karhunen-Loeve, por exemplo.

### 6. Referências

- Blahut, R.E. Principles and Practice of Information Theory. Addison Wesley Publishing Company, 1990.
- Machado e Silva, A.J.F. Métodos de Avaliação de Modelos de Classificação de Imagens Digitais. Anais do VII Simpósio Brasileiro de Sensoriamento Remoto, Curitiba, PR, 1993.
- Novo, E.M.L.M. Sensoriamento Remoto: Princípios e Aplicações. Editora Edgard Blucher, 1988.
- Richards, J.A. Remote Sensing Digital Image Analysis: An Introduction. Springer-Verlag, 1986.

**Imagem Pará (768,768)**

	B1	B2	B3	B4	B5	B7
m	58,8	24,8	21,6	68,6	58,9	12,9
H	3,74	3,87	4,20	6,03	6,40	4,53
$\sigma^2$	17,40	17,60	43,78	493,82	603,38	60,02

Tabela 1: média, entropia e variância.

	B1	B2	B3	B4	B5	B7
B1	17,40	15,42	24,91	19,85	72,51	26,03
B2	15,42	17,60	25,23	39,32	82,57	27,48
B3	24,91	25,23	43,78	26,29	117,06	43,32
B4	19,85	39,32	26,29	493,82	353,58	69,23
B5	72,51	82,57	117,06	353,58	603,38	175,37
B7	26,03	27,48	43,32	69,23	175,37	60,02

Tabela 2: matriz de covariância.

	<b>B1</b>	<b>B2</b>	<b>B3</b>	<b>B4</b>	<b>B5</b>	<b>B7</b>
<b>B1</b>	1,00	0,88	0,90	0,21	0,71	0,81
<b>B2</b>	0,88	1,00	0,91	0,42	0,80	0,85
<b>B3</b>	0,90	0,91	1,00	0,18	0,72	0,85
<b>B4</b>	0,21	0,42	0,18	1,00	0,65	0,40
<b>B5</b>	0,71	0,80	0,72	0,65	1,00	0,92
<b>B7</b>	0,81	0,85	0,85	0,40	0,92	1,00

Tabela 3: matriz de correlação.

	<b>B1</b>	<b>B2</b>	<b>B3</b>	<b>B4</b>	<b>B5</b>	<b>B7</b>
<b>PC1</b>	0,08	0,10	0,13	0,60	0,75	0,20
<b>PC2</b>	-0,16	-0,12	-0,28	0,77	-0,47	-0,27
<b>PC3</b>	0,41	0,42	0,68	0,19	-0,39	0,10
<b>PC4</b>	0,29	0,13	0,05	-0,08	0,25	-0,91
<b>PC5</b>	0,81	-0,01	0,55	0,01	-0,06	0,21
<b>PC6</b>	0,25	-0,89	0,39	0,05	0,00	-0,03

Tabela 4: matriz da transformada de Karhunen-Loeve.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>	<b>PC6</b>
$\sigma^2$	972,40	235,82	20,69	3,47	2,32	1,30
%	78,7	19,1	1,7	0,3	0,2	0,1

Tabela 5: variância das principais componentes.

**Imagem Mato Grosso (1000,1000)**

	<b>B1</b>	<b>B2</b>	<b>B3</b>	<b>B4</b>	<b>B5</b>	<b>B7</b>
<b>m</b>	56,0	24,4	25,7	49,9	63,6	16,7
<b>H</b>	3,63	3,37	4,33	4,69	6,16	4,80
$\sigma^2$	12,14	8,87	50,58	42,19	547,46	104,61

Tabela 6: média, entropia e variância.

	<b>B1</b>	<b>B2</b>	<b>B3</b>	<b>B4</b>	<b>B5</b>	<b>B7</b>
<b>B1</b>	12,14	9,09	22,35	-4,91	71,51	31,28
<b>B2</b>	9,09	8,87	19,82	-2,58	61,00	26,83
<b>B3</b>	22,35	19,82	50,58	-11,00	155,57	68,85
<b>B4</b>	-4,91	-2,58	-11,00	42,19	-41,69	-19,89
<b>B5</b>	71,51	61,00	155,57	-41,69	547,46	233,04
<b>B7</b>	31,28	26,83	68,85	-19,89	233,04	104,61

Tabela 7: matriz de covariância.

	<b>B1</b>	<b>B2</b>	<b>B3</b>	<b>B4</b>	<b>B5</b>	<b>B7</b>
<b>B1</b>	1,00	0,88	0,90	-0,22	0,88	0,88
<b>B2</b>	0,88	1,00	0,94	-0,13	0,88	0,88
<b>B3</b>	0,90	0,94	1,00	-0,24	0,93	0,95
<b>B4</b>	-0,22	-0,13	-0,24	1,00	-0,27	-0,30
<b>B5</b>	0,88	0,88	0,93	-0,27	1,00	0,97
<b>B7</b>	0,88	0,88	0,95	-0,30	0,97	1,00

Tabela 8: matriz de correlação.

	<b>B1</b>	<b>B2</b>	<b>B3</b>	<b>B4</b>	<b>B5</b>	<b>B7</b>
<b>PC1</b>	0,12	0,10	0,25	-0,07	0,88	0,38
<b>PC2</b>	0,03	0,07	0,05	0,99	0,06	-0,02
<b>PC3</b>	0,30	0,32	0,71	-0,03	-0,43	0,34
<b>PC4</b>	0,32	0,22	0,30	-0,07	0,20	-0,84
<b>PC5</b>	0,88	-0,05	-0,45	0,01	-0,05	0,15
<b>PC6</b>	0,15	-0,91	0,38	0,04	0,00	-0,05

Tabela 9: matriz da transformada de Karhunen-Loeve.

	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>	<b>PC6</b>
$\sigma^2$	712,53	39,08	8,17	3,40	1,83	0,84
%	93,0	5,1	1,1	0,4	0,2	0,1

Tabela 10: variância das principais componentes.

	<b>C12</b>	<b>C13</b>	<b>C14</b>	<b>C23</b>	<b>C24</b>	<b>C34</b>
<b>B1</b>	0,13	0,34	0,89	0,33	0,90	0,74
<b>B2</b>	0,42	0,50	0,93	0,59	0,86	0,83
<b>B3</b>	0,94	0,60	0,93	0,62	0,57	0,69
<b>B4</b>	0,98	0,84	0,24	0,79	0,97	0,81
<b>B5</b>	0,41	0,79	0,91	0,87	0,80	0,95
<b>B7</b>	0,76	0,69	0,94	0,85	0,63	0,94

Tabela 11: divergência para distribuição gaussiana contínua.



	<b>C12</b>	<b>C13</b>	<b>C14</b>	<b>C23</b>	<b>C24</b>	<b>C34</b>
<b>B1</b>	0,16	0,41	0,88	0,31	0,89	0,72
<b>B2</b>	0,43	0,60	0,92	0,60	0,85	0,83
<b>B3</b>	0,96	0,68	0,94	0,63	0,52	0,66
<b>B4</b>	0,96	0,85	0,26	0,77	0,94	0,80
<b>B5</b>	0,42	0,79	0,95	0,92	0,76	0,97
<b>B7</b>	0,77	0,65	0,95	0,91	0,61	0,96

Tabela 12: divergência para distribuição discreta.

	<b>C12</b>	<b>C13</b>	<b>C14</b>	<b>C23</b>	<b>C24</b>	<b>C34</b>
<b>B123</b>	8,1	1,7	18,5	7,4	12,8	18,0
<b>B124</b>	31,4	45,6	13,1	14,5	28,0	41,8
<b>B125</b>	10,2	32,7	40,8	42,8	31,1	72,8
<b>B127</b>	7,0	5,8	19,8	12,8	14,5	24,9
<b>B134</b>	32,4	45,6	17,6	15,8	28,0	42,8
<b>B135</b>	12,8	32,7	42,5	43,3	31,1	73,4
<b>B137</b>	10,4	5,8	23,1	14,2	14,4	26,6
<b>B145</b>	33,0	56,1	40,4	45,0	39,9	82,5
<b>B147</b>	32,1	46,0	19,0	18,9	28,8	46,1
<b>B157</b>	12,1	33,2	43,1	44,5	31,8	75,4
<b>B234</b>	32,4	45,7	17,3	16,0	27,3	43,1
<b>B235</b>	12,9	32,7	42,4	43,3	30,4	73,5
<b>B237</b>	10,5	5,9	22,9	14,4	12,9	26,9
<b>B245</b>	33,0	56,1	40,3	45,1	39,3	82,7
<b>B247</b>	32,1	46,0	18,8	19,1	28,1	46,4
<b>B257</b>	12,2	33,2	43,0	44,6	31,2	75,5
<b>B345</b>	33,9	56,1	42,3	45,5	39,3	83,2
<b>B347</b>	33,1	46,0	22,1	20,1	28,0	47,3
<b>B357</b>	14,5	33,2	44,6	45,0	31,1	76,1
<b>B457</b>	33,6	56,4	42,6	46,7	39,9	84,9

Tabela 13: distância entre as classes para combinações de 3 bandas.

	<b>P1</b>	<b>P2</b>	<b>P3</b>	<b>P4</b>
<b>C1</b>	0,25	0,50	0,70	0,50
<b>C2</b>	0,25	0,20	0,15	0,30
<b>C3</b>	0,25	0,10	0,05	0,05
<b>C4</b>	0,25	0,20	0,10	0,15

Tabela 14: distribuição de probabilidades para as classes Ci.

	<b>P1</b>	<b>P2</b>	<b>P3</b>	<b>P4</b>
<b>B123</b>	14,7	9,0	5,3	6,0
<b>B124</b>	8,5	5,2	3,1	3,6
<b>B125</b>	14,1	15,0	12,7	17,1
<b>B127</b>	12,2	10,4	7,7	10,4
<b>B134</b>	6,5	4,2	2,4	2,9
<b>B135</b>	4,2	3,6	2,5	3,2
<b>B137</b>	6,4	5,0	3,3	4,1
<b>B145</b>	3,4	3,0	2,1	2,4
<b>B147</b>	3,8	3,1	2,2	2,5
<b>B157</b>	9,1	8,4	6,3	8,9
<b>B234</b>	8,5	5,0	2,9	3,4
<b>B235</b>	5,2	4,4	3,1	4,2
<b>B237</b>	7,4	5,8	3,9	5,2
<b>B245</b>	3,4	2,6	1,8	2,1
<b>B247</b>	4,2	3,0	2,0	2,3
<b>B257</b>	11,1	9,8	7,3	10,8
<b>B345</b>	2,9	2,3	1,5	1,8
<b>B347</b>	3,8	2,8	1,7	2,1
<b>B357</b>	7,9	6,5	4,4	6,2
<b>B457</b>	6,4	4,5	3,0	3,3

Tabela 15: probabilidade de erro - MAP com combinações de 3 bandas.