

CLASSIFICAÇÃO SUPERVISIONADA DE IMAGENS

Ronei Marcos de Moraes

Departamento de Estatística

Universidade Federal da Paraíba

nov/1999

Introdução

Os Classificadores Supervisionados Clássicos estudados serão:

- Método do Paralelepípedo
- Mínima Distância
- Método dos K-Vizinhos Próximos - KNN
- Máxima Verossimilhança
- Redes Neurais

E os Classificadores Supervisionados Contextuais:

- Método Geral
- Método p,q,r

Amostras de Treinamento

Os Métodos Supervisionados utilizam-se de áreas ou pontos de referência da imagem para obter o número e determinados características de natureza estatística das classes [4]. Essas características dependerão do tipo do Método utilizado.

É crucial que essas informações *a priori* sejam factíveis. A acurácia da classificação supervisionada depende basicamente de dois fatores:

- A representatividade das estimativas do número e das características estatísticas das classes presentes na imagem;
- O grau de acerto das restrições (hipóteses) sobre as quais a técnica de classificação é baseada.

Essas hipóteses variam de uma técnica para outra e, em geral, quanto mais sofisticada a técnica mais restrições ela contém.

A validade das estimativas dependem também de dois fatores: tamanho e representatividade das amostras. Idealmente, quanto maior o tamanho da amostra, melhor, porém a questão do tamanho das amostras depende também da questão custo. Essas amostras também não são obtidas de forma aleatória, mas representam uma referência conhecida para uma determinada classe.

Em geral, as amostras são localizadas no campo de trabalho por fotografias aéreas ou por fotointerpretação ou por coordenadas obtidas em campo por GPS. É muito importante para a localização das amostras que a imagem seja corrigida geometricamente.

Análise das Amostras

Existem várias formas de se avaliar a qualidade das amostras coletadas.

Usualmente, a qualidade pode ser analisada por uma matriz de confusão, que é uma matriz cuja diagonal principal mostra o número de pontos classificados corretamente para cada classe e os elementos fora dela mostram a quantidade de pontos classificados erroneamente, baseando-se numa classificação gerada somente sobre as amostras.

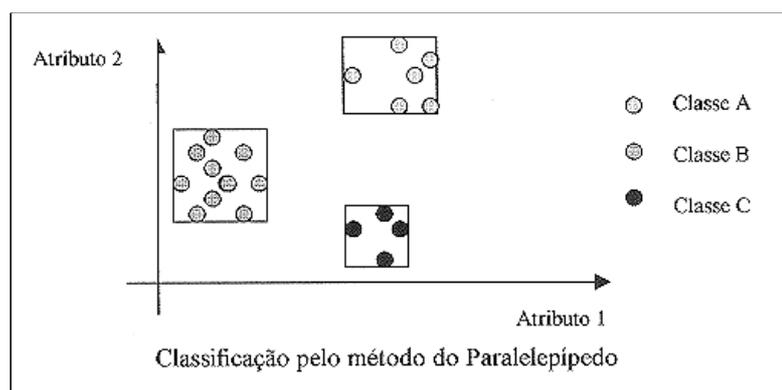
Quanto menores forem os valores fora da diagonal principal, melhor será o conjunto de amostras. Consideremos a matriz de confusão:

Imagem	Classif.							
Referên.	Á. Ant.	Fl. Omb.	Fl. Sec.	Rio	C. Av.	Fl. Al.	Á. Cont.	Á. Mist.
Á. Ant.	14499	1532	5673	201	201	200	99	2332
Fl. Omb.	367	73884	5170	0	0	0	0	0
Fl. Sec.	2326	4294	35396	16	38	55	17	0
Rio	547	0	1054	75035	0	3	2797	150
C. Av.	383	0	71	1	395	0	0	0
Fl. Al.	385	91	1584	0	0	1776	0	1153
Á. Cont.	120	0	331	1194	0	0	6091	167
Á. Mist.	2070	0	6	75	0	247	19	8156

Método do Paralelepípedo

Classificador muito simples que se utiliza dos histogramas das amostras de treinamento de cada componente espectral e obtendo limiares máximo e mínimo para cada classe. Juntos, esses limiares descrevem um paralelepípedo no espaço multidimensional.

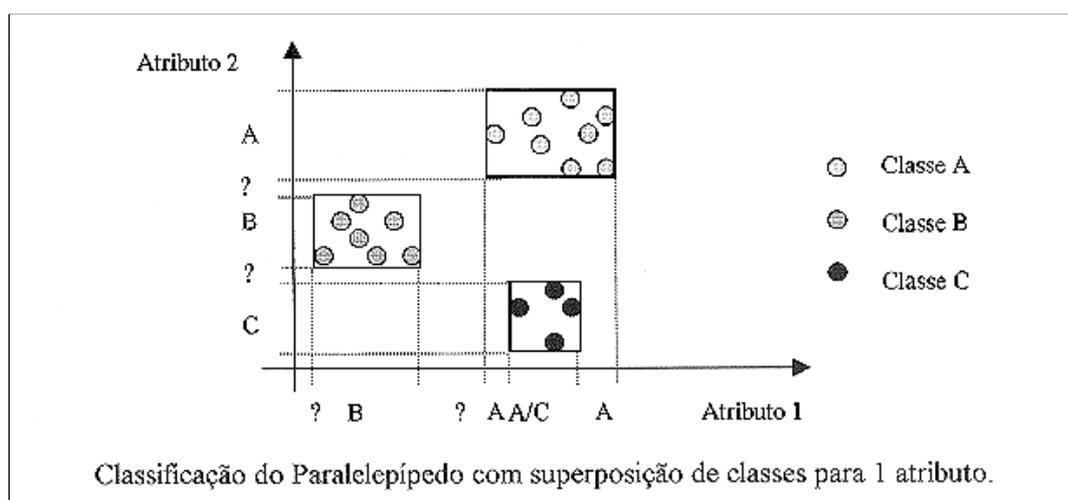
"Pixel" X é atribuído à classe i se todos os elementos do vetor característico X estiver entre os máximos e mínimos que caracterizam a classe i .



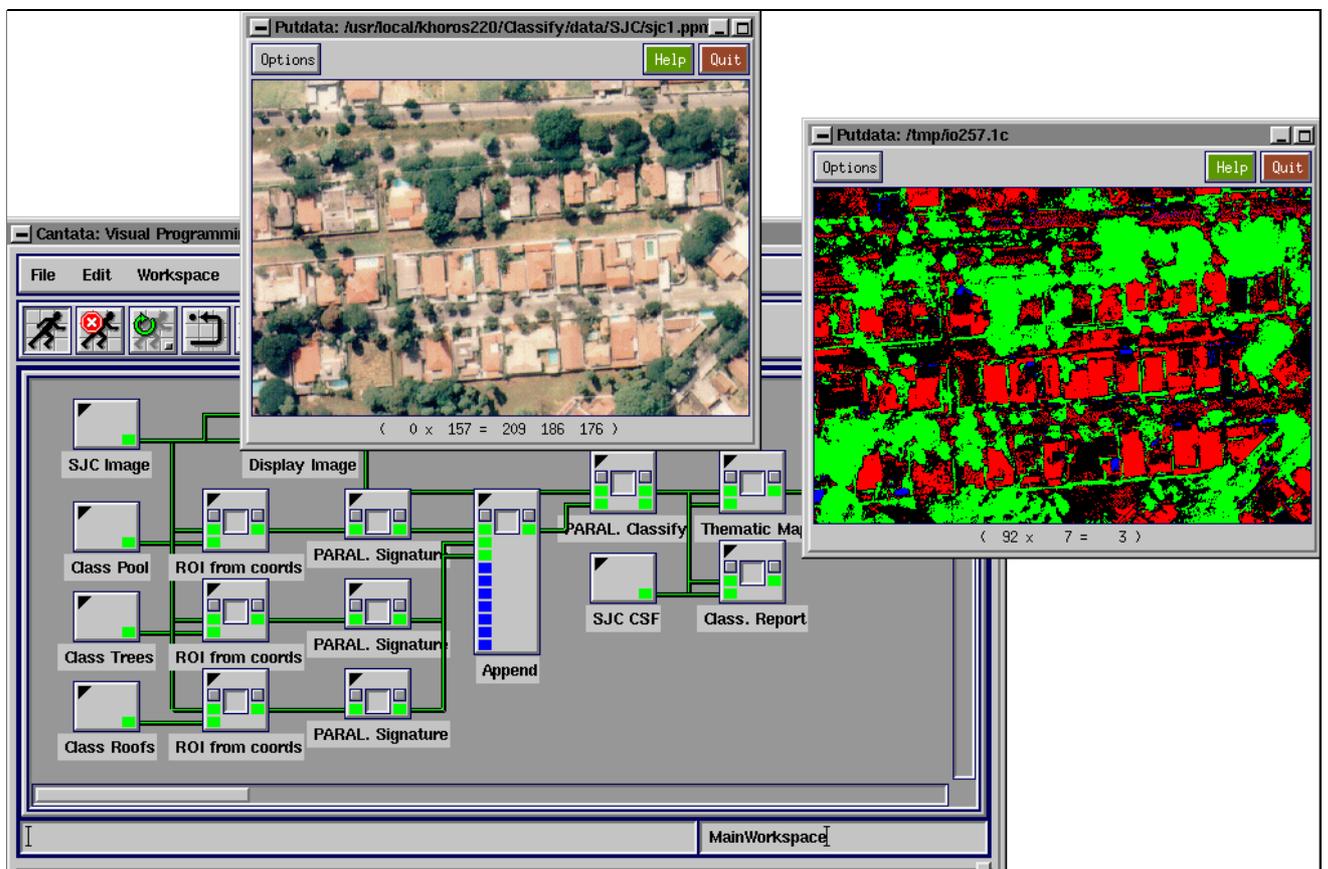
Inclui naturalmente uma opção de rejeição e como vantagens, é rápido e simples.

Como desvantagens, pode-se citar:

- "Pixels" que estejam nas regiões "vazias" entre os paralelepípedos não serão classificados
- Podem ocorrer sobreposições de classes para algum ou vários atributos.



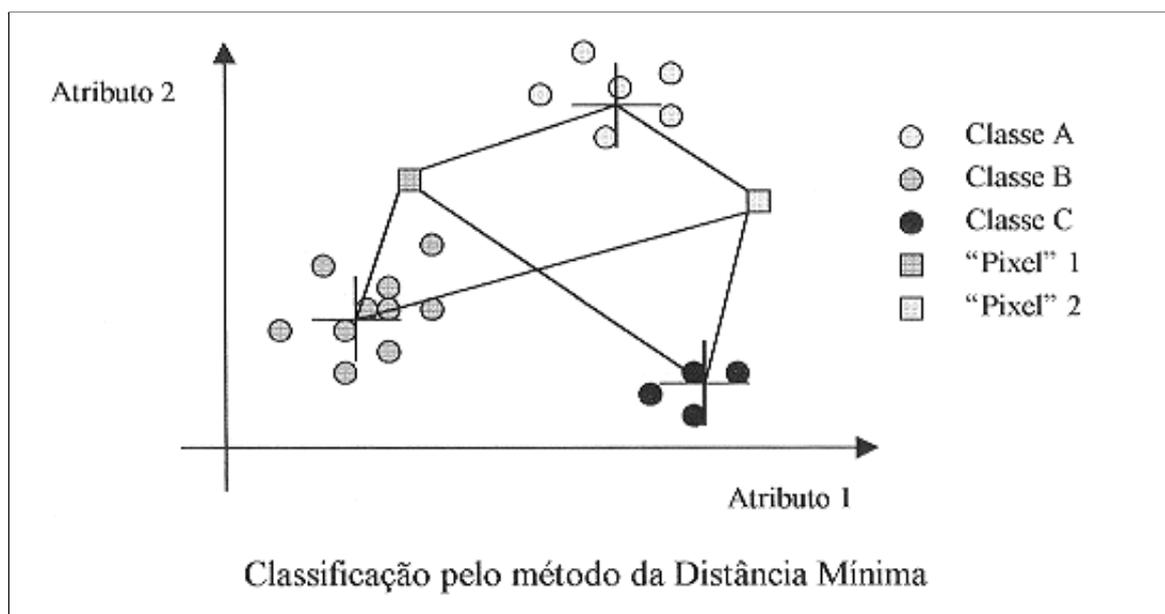
A figura mostra uma classificação de uma área de São José dos Campos pelo Método do Paralelepípedo no Khoros 2.2 [5]



Método da Mínima Distância

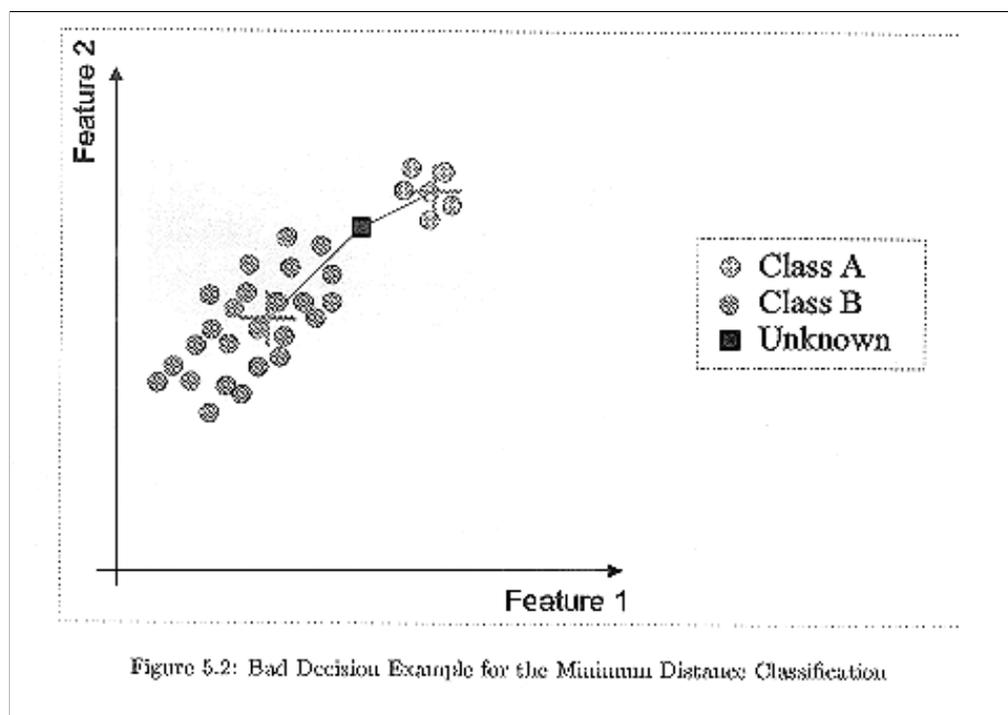
Utiliza as amostras para determinar os pontos médios das N classes μ_i , $i = 1, \dots, N$.

"Pixel" X é atribuído à classe i se a Distância Euclideana entre X e μ_i for a menor para todas as classes.

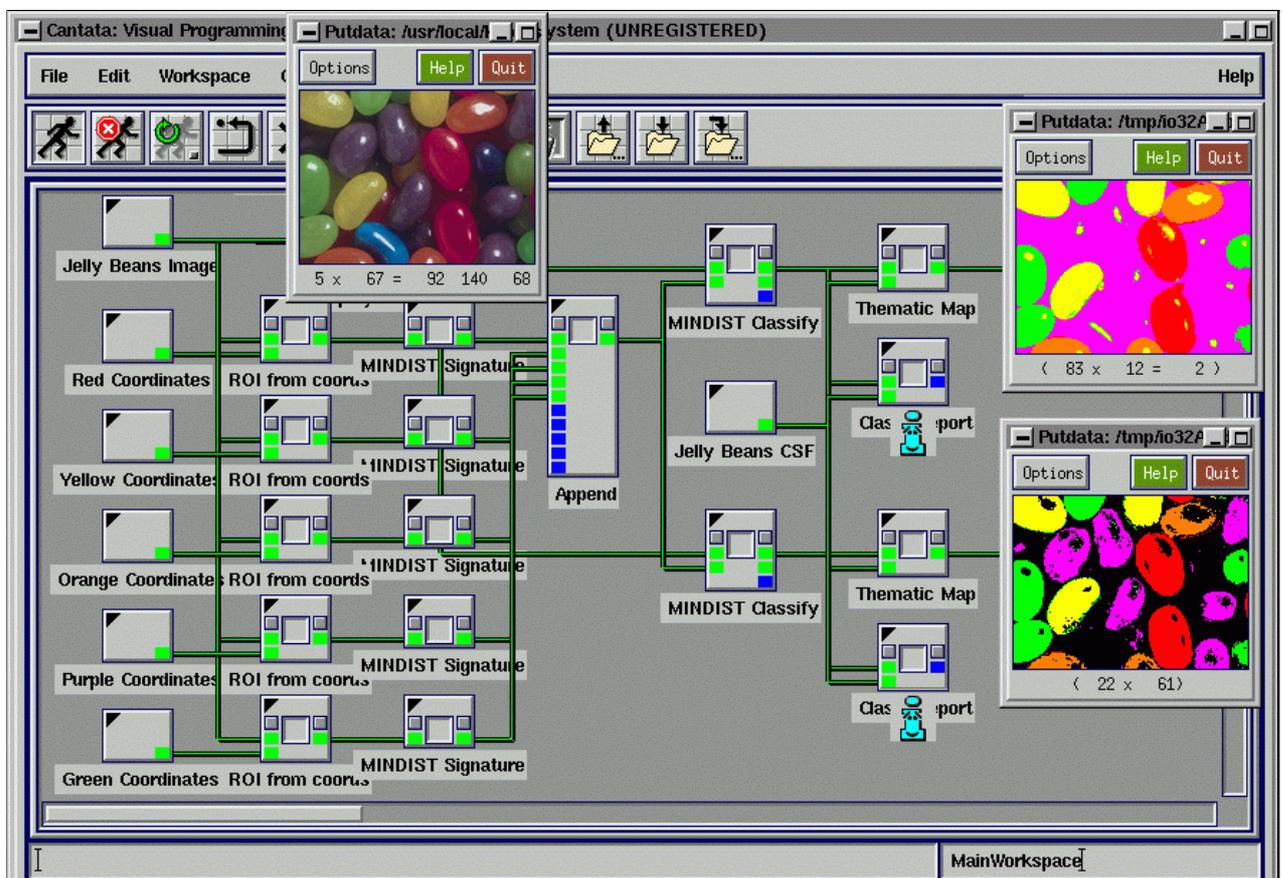


Se for utilizado um limiar, esse método descreve para cada classe uma esfera no espaço multidimensional (opção de rejeição).

É uma técnica rápida, porém não muito eficiente [2].



Exemplo de Classificação de uma imagem de "jujubas" pelo Método da Mínima Distância no Khoros 2.2 [5].



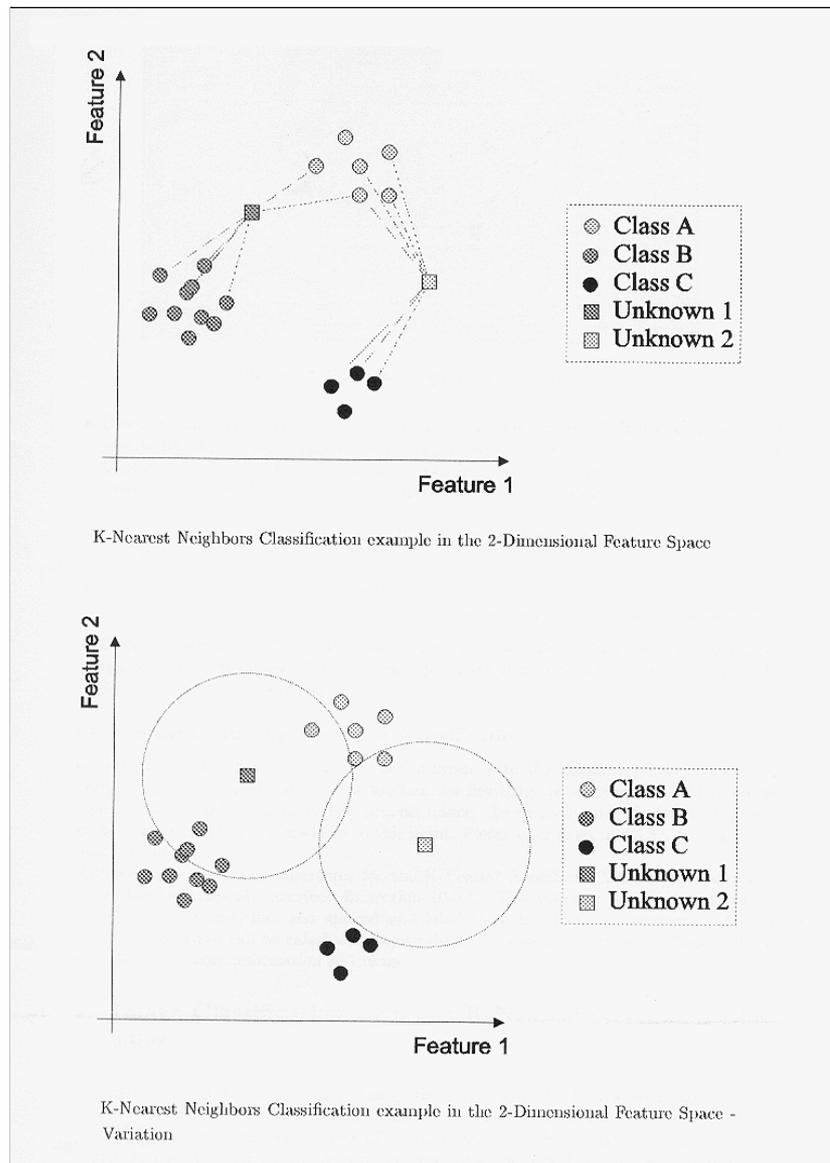
Método dos K-Vizinhos Próximos - KNN

É um método não-paramétrico: os parâmetros utilizadas são os próprios "pixels" amostrados para cada uma das N classes.

Usuário define a vizinhança a ser pesquisada: valor de K .

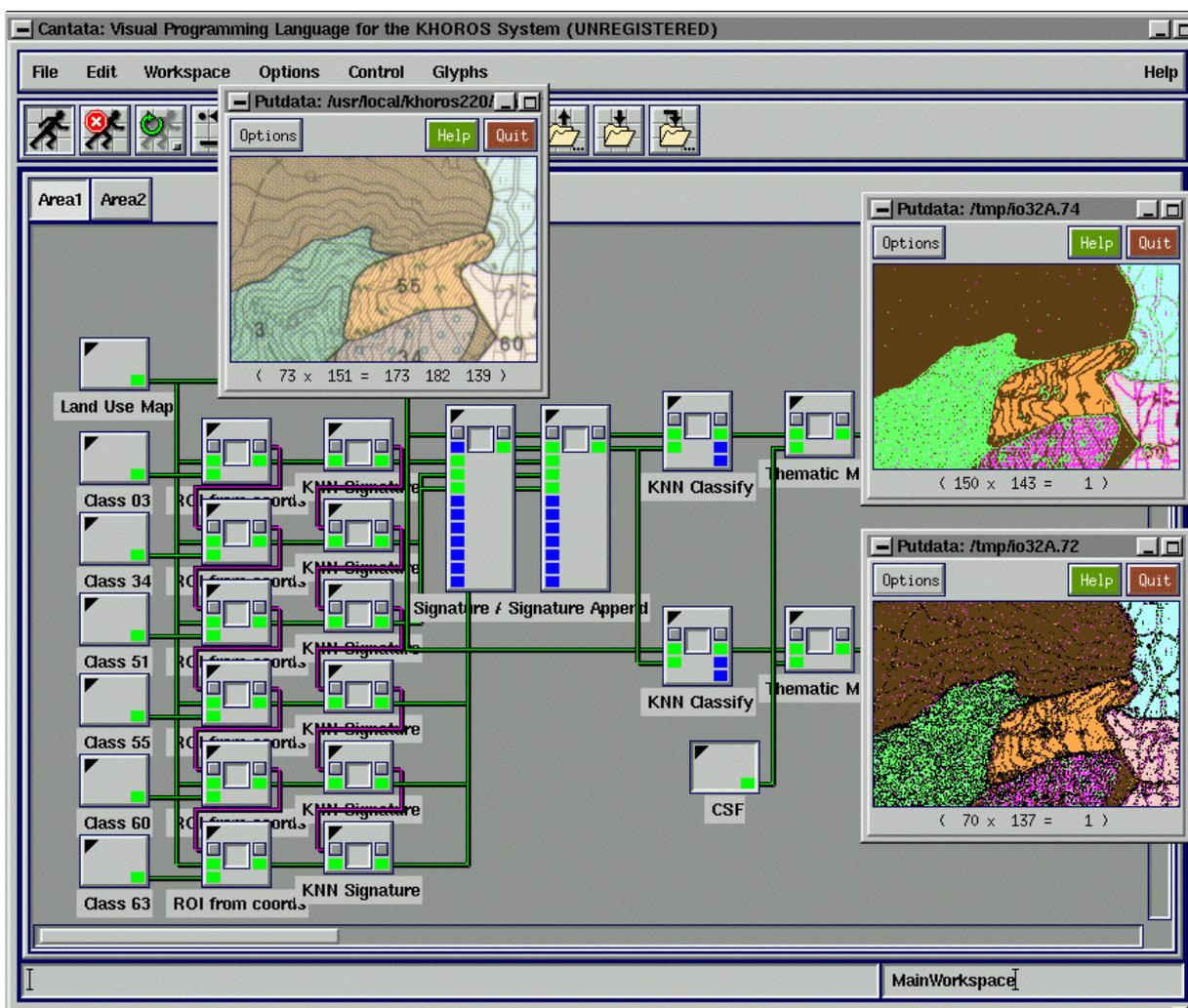
"Pixel" X é atribuído à classe i se os K vizinhos mais próximos são amostras da classe i , utilizando a distância Euclideana.

Se for utilizado um limiar, esse método utiliza uma distância fixa (descrevendo para cada classe uma esfera no espaço multidimensional) e não mais um valor fixo para K (opção de rejeição).



Reflete bem a distribuição das amostras coletadas, mas é uma técnica que envolve muitos cálculos de distância para encontrar os K vizinhos e é mais lento se o número de pontos amostrais é grande [1].

Exemplo de Classificação das cores de um mapa pelo Método dos K-Vizinhos Próximos no Khoros.2.2 [5]:



Máxima Verossimilhança

Utiliza as amostras para determinar o vetor de médias, e matriz de covariâncias das N classes $\mu_i, i = 1, \dots, N$.

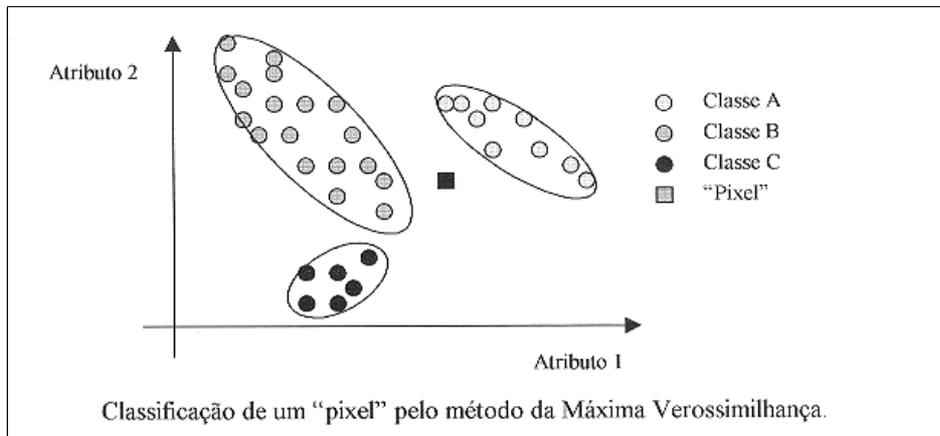
O vetor de médias e matriz inversa de covariâncias delimitam um elipsóide no espaço multidimensional.

"Pixel" X é atribuído à classe i se a Verossimilhança entre X e a classe i for a maior para todas as classes:

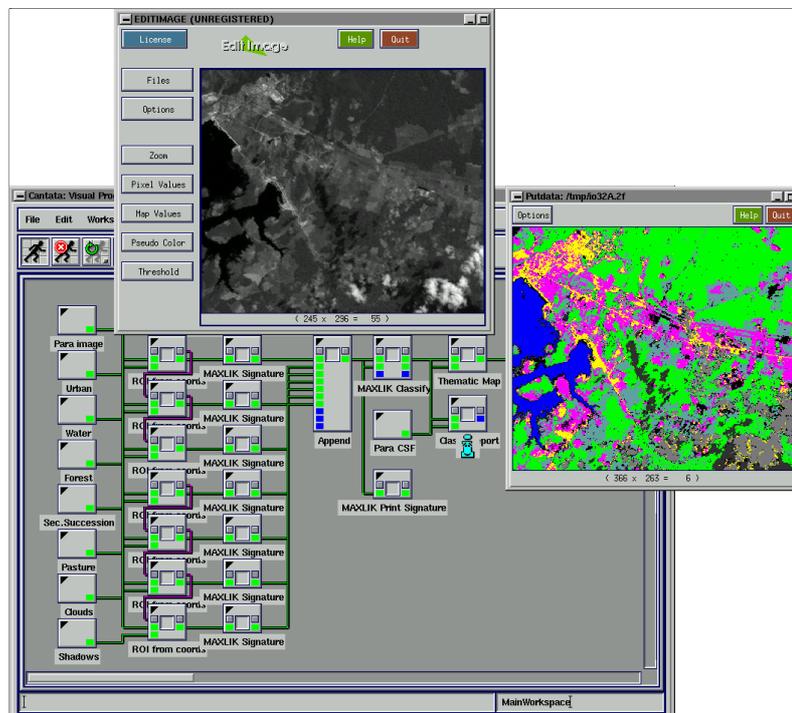
$$g(X) = \ln P(X|\text{classe}_i) + \ln P(\text{classe}_i)$$

Limiares podem ser aplicados à g_i de modo a rejeitar pontos.

É uma técnica robusta para modelar classes unimodais e é relativamente rápida [2].



Classificação de uma imagem Landsat 5 pelo Método da Máxima Verossimilhança no Khoros.2.2 [5]:



Classificador da Distância de Mahalanobis

É uma particularização do Método da Máxima Verossimilhança quando se pode admitir a igualdade das probabilidades à priori e iguais covariâncias.

Utiliza as amostras para determinar o vetor de médias, e matriz de covariâncias das N classes $\mu_i, i = 1, \dots, N$.

O vetor de médias e matriz inversa de covariâncias delimitam um elipsóide no espaço multidimensional.

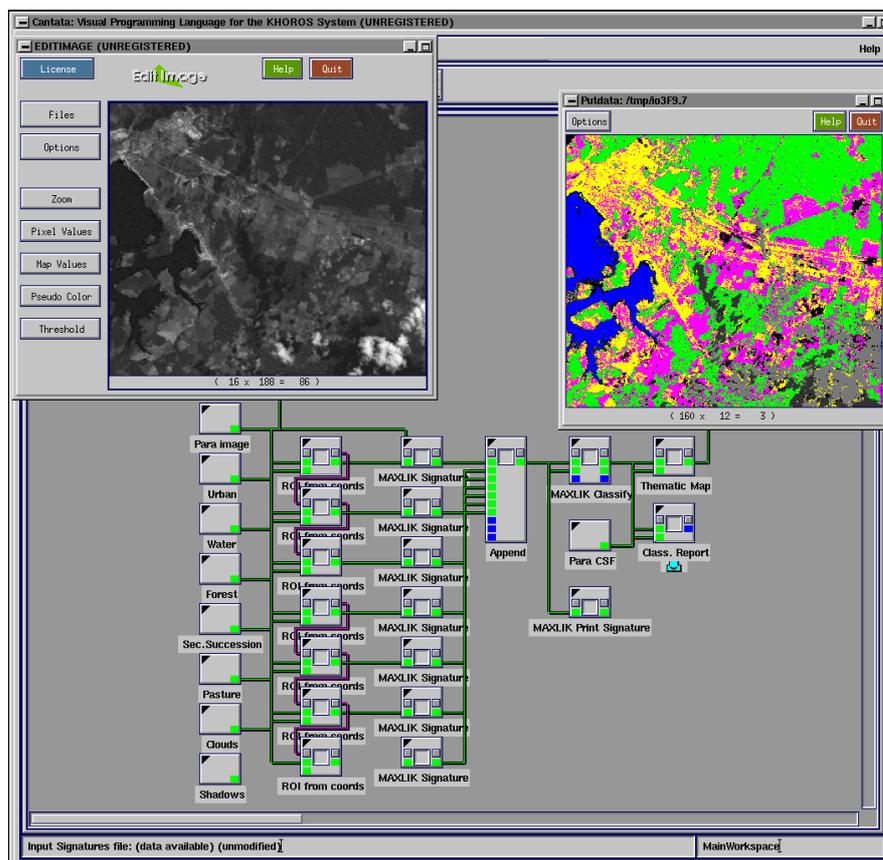
"Pixel" X é atribuído à classe i se a Distância de Mahalanobis entre X e a classe i for a menor para todas as classes:

$$d(X, \text{classe}_i) = (X - \mu_i)^t \Sigma_i^{-1} (X - \mu_i)$$

Limiares podem ser aplicados às distâncias de modo a rejeitar pontos.

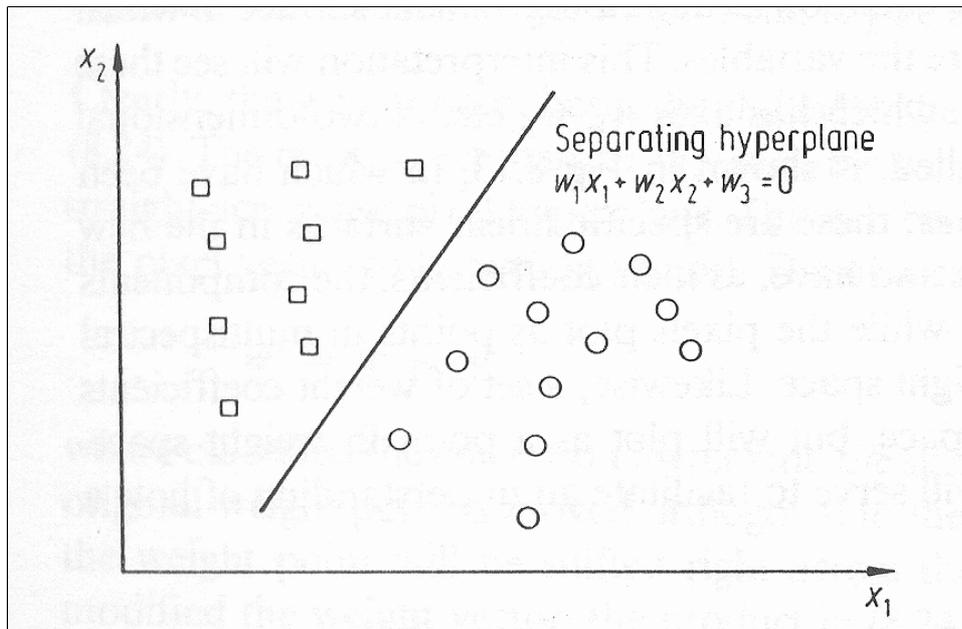
É uma técnica robusta para modelar classes unimodais e é relativamente rápida [2], mas as hipóteses iniciais restringem sua utilização.

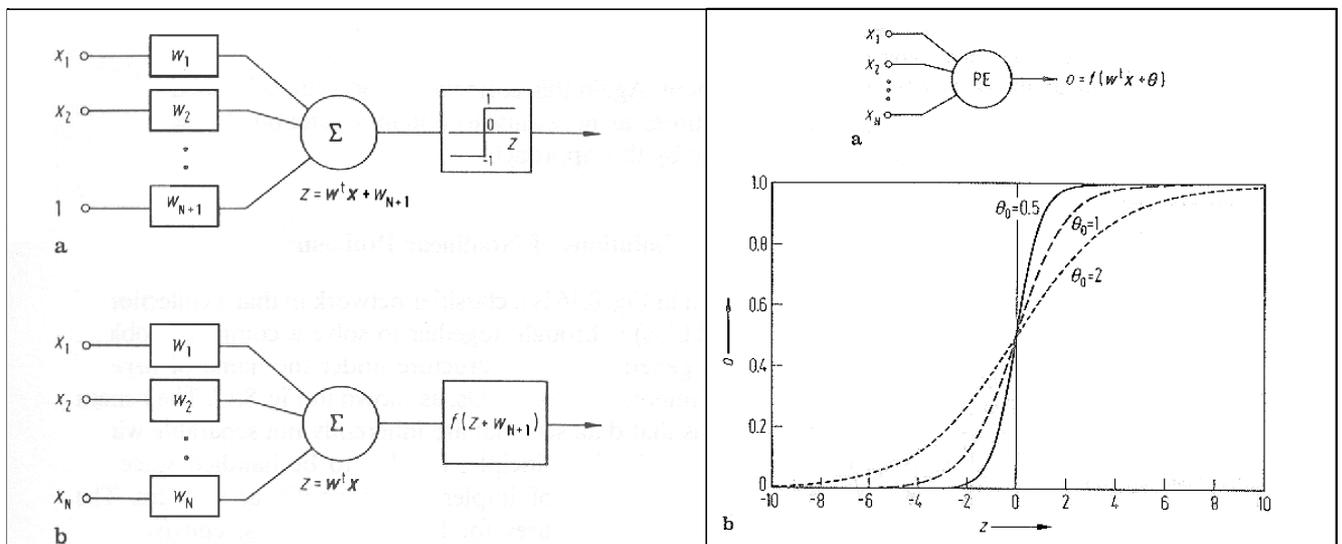
Exemplo de Classificação de uma imagem Landsat 5 pelo Método de Mahalanobis no Khoros.2.2 [5]:



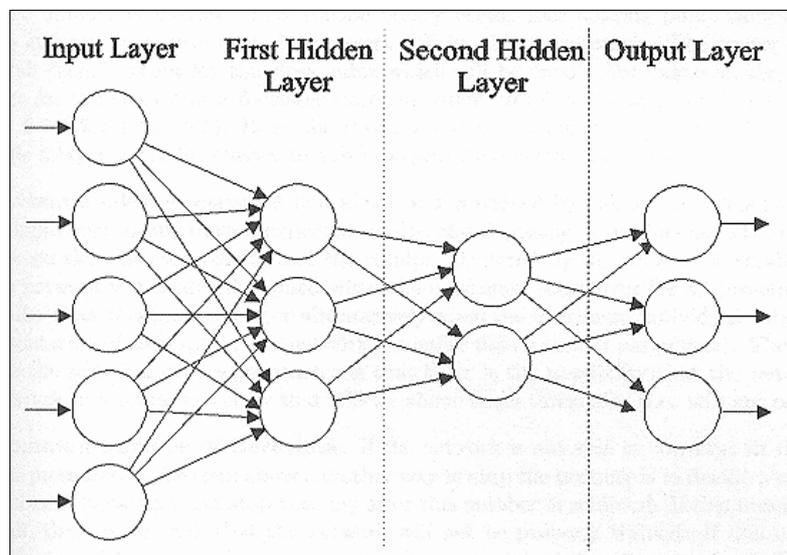
Classificação Usando Redes Neurais

É uma técnica que utiliza uma concepção biológica do neurônio humano, traduzida por uma função linear seguida por uma função discriminante não linear.





Sua concepção pode variar constituindo múltiplas camadas



Os cinco neurônios da camada de entrada correspondem a cinco atributos para um "pixel" na entrada. Duas camadas escondidas, respectivamente com três e dois neurônios formam a primeira e a segunda camada escondida. A saída é formada por três neurônios para a classificação do "pixel" em três classes [6].

As conexões trabalham como somatórias de valores ponderados que são limiarizados para produzir a ativação para algum neurônio na saída.

As amostras da imagem (chamadas de dados de treinamento) são utilizadas para determinar os pesos da função linear e o limiar para cada uma das N classes $\mu_i, i = 1, \dots, N$.

A técnica mais comum de treinamento é a "Backpropagation" que é uma complexa técnica de aprendizado para balancear os pesos e o limiar [6].

Resumidamente, a rede é inicializada com pesos arbitrários. Os "pixels" de treinamento são apresentados à rede, um a cada vez. Para cada "pixel", a saída da rede é computada utilizando um conjunto de equações [5]. Se a saída da rede não é igual ao rótulo do "pixel" na sua amostra, os pesos são recalculados obtendo-se novas ponderações.

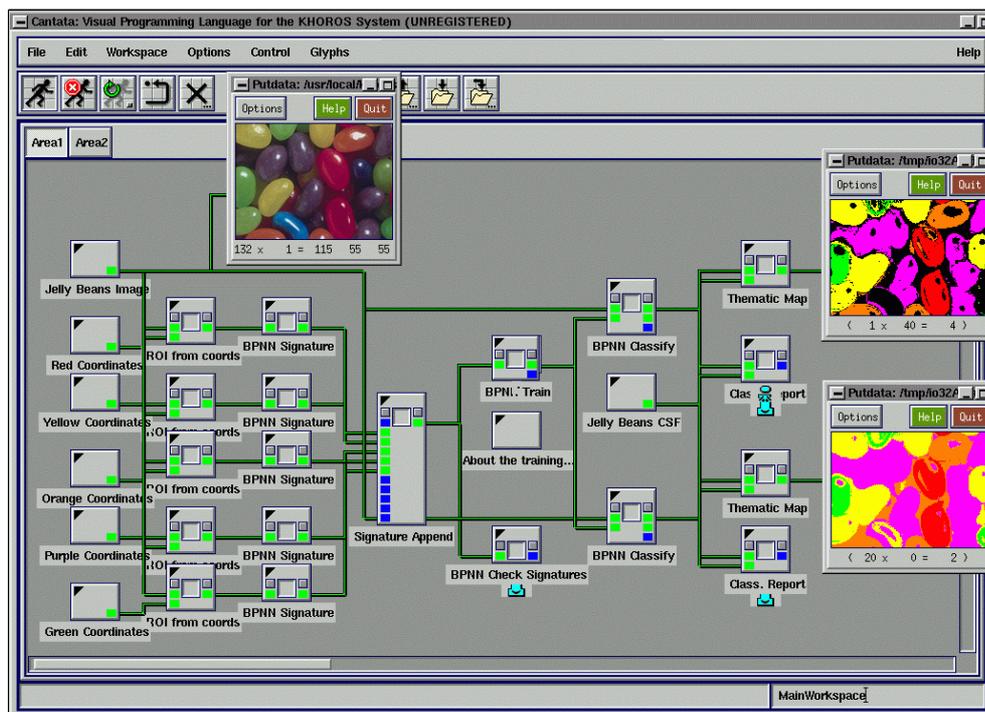
Dessa forma, o erro é propagado de volta através de todas as camadas da rede (retropropagação), no sentido de corrigir os seus pesos. Esse processo é repetido até que a rede classifique todos os pontos amostrais corretamente, ou quando os erros de classificação atingem um nível aceitável.

Limiares podem ser aplicados à rede neural de modo a rejeitar pontos.

É uma técnica robusta porém lenta no treinamento (retropropagação) e sua acurácia é muito dependente

das amostras de treinamento, da estrutura da rede e dos parâmetros de aceitação [2].

Exemplo de Classificação de uma imagem de "jujubas" por Redes Neurais no Khoros 2.2 [5]:



Método Supervisionado Contextual

Os métodos contextuais analisam o ponto a ser classificado e uma vizinhança pré-determinada a ele. Temos K classes possíveis para classificação, com probabilidades a priori p_1, p_2, \dots, p_k , com densidade da classe k dada por $f_k(x)$. Com a possibilidade de se classificar um "pixel" de forma duvidosa, denotada por D , o espaço das decisões será: $\Omega = \{1, 2, \dots, K, D\}$.

A perda por se estar em dúvida seja $e \in [0, 1 - 1/K]$, com soma unitária.

A regra de classificação será atribuir um "pixel" a classe k se

$$P_i(k|\mathcal{D}_i) = \frac{p_k f_k(x_i) R_k(x_{i1}, \dots, x_{in})}{P_i(\mathcal{D}_i)}$$

é a maior para todas as classes e superior ao limiar e ; onde: \mathcal{D}_i é a vizinhança para o "pixel" i ; $P_i(k|\mathcal{D}_i)$ é a probabilidade condicional da classe k dado \mathcal{D}_i e:

$$R_k(x_{i1}, \dots, x_{in}) = \sum_{K_1} \dots \sum_{K_n} h(x_{i1}, \dots, x_{in} | x_i, k, k_1, \dots, k_n) g(k_1, \dots, k_n | k)$$

A função h descreve a densidade conjunta dos vetores característicos em \mathcal{D}_i dado a configuração particular: k, k_1, \dots, k_n .

A função g descreve essa configuração particular dado que o "pixel" central pertence à classe k .

As várias somatórias descreve o total das possíveis configurações. Para um exemplo com 4 vizinhos:

A função g é:

$$g(k_1, k_2, k_3, k_4 \setminus k) = \text{PR} \left(\begin{array}{c} \boxed{k_1} \\ \boxed{k_3} \quad \boxed{} \quad \boxed{k_2} \\ \boxed{k_4} \end{array} \left| \begin{array}{l} \text{pixel central} \\ \text{é da classe} \\ k \end{array} \right. \right)$$

A função h é :

$$h(x_{i1}, x_{i2}, x_{i3}, x_{i4} \mid x, k, k_1, k_2, k_3, k_4) =$$

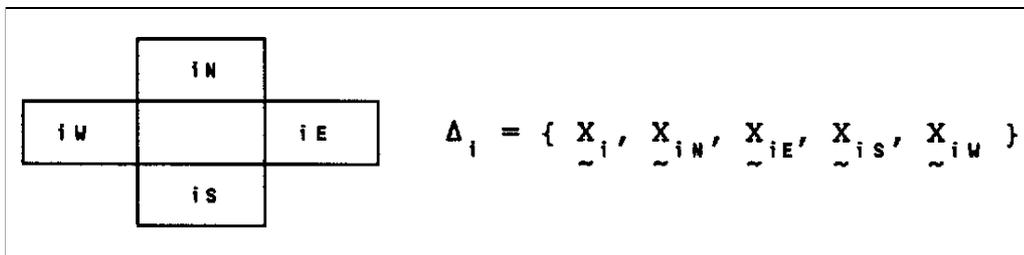
$$= PR \left(\begin{array}{ccc} & x_{i1} & \\ x_{i3} & & x_{i2} \\ & x_{i4} & \end{array} \mid \begin{array}{ccc} & k_1 & \\ k_3 & x_i & k_2 \\ & k_4 & \end{array} \right)$$

Como vantagens, essa metodologia é mais acurada que as vistas anteriormente, porém o esforço computacional para realizar a classificação é grande e faz com que o método seja lento.

Mais detalhes com respeito a essa metodologia pode ser encontrada em Moraes [7].

Modelo Contextual p,q,r

Visando reduzir o custo computacional envolvido no método Contextual geral, várias abordagens foram propostas. Caso de uma vizinhança-4 em cruz:



Procura-se modelar as probabilidades g e h , de modo a tentar reduzir e simplificar os cálculos

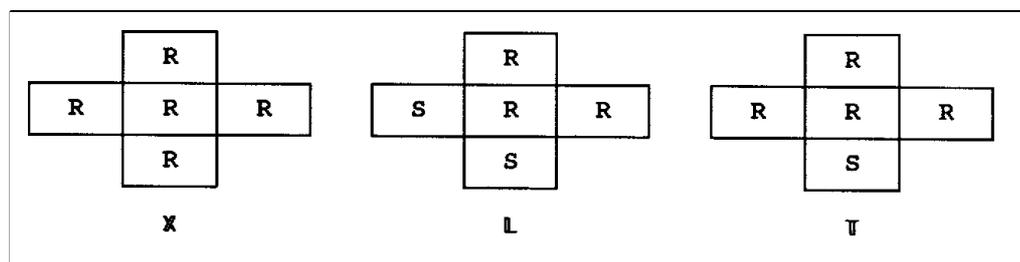
Para h , vários autores assumem os vetores \vec{X} como sendo condicionalmente independentes dado as classes:

$$f(\mathcal{D}_i | k, a, b, c, d) = f_k(\vec{X}_i) f_k(\vec{X}_{iN}) f_k(\vec{X}_{iE}) f_k(\vec{X}_{iS}) f_k(\vec{X}_{iW})$$

$$h(\mathcal{D}_i | k, a, b, c, d, \vec{X}_i) = f_k(\vec{X}_{iN}) f_k(\vec{X}_{iE}) f_k(\vec{X}_{iS}) f_k(\vec{X}_{iW})$$

Para g , Owen and Switzer [8] assumem um modelo de probabilidade geométrica ("pixels" são pequenos em relação ao tamanho das classes) e Hjort [9] assume que o número de classes presentes na cruz será de no máximo duas.

A figura ilustra os três tipos possíveis de cruzes com probabilidade maior (ou eventualmente igual) a zero:



Para \mathcal{X} existe somente um padrão e para \mathcal{L} e \mathcal{T} , quatro padrões possíveis pela rotação dos padrões R e S.

Adotando esta sistemática, o número de termos sofrerá uma redução de K^4 para $8K - 7$.

Referências

- [1] Santos, Rafael; *Introdução à Classificação de Imagens*. Tutorial SIBGRAPI'99
- [2] Richards, J. A.; *Remote Sensing Digital Image Analysis - An Introduction*. Springer-Verlag, 1995.
- [3] Kandel, A.; *Fuzzy Techniques in Pattern Recognition*. John Wiley & Sons, 1982.
- [4] Mather, P. M.; *Computer Processing of Remotely-Sensed Images*. John Wiley & Sons, 1987.
- [5] Santos, R.; Ohashi, T.; Yoshida T. and Ejima, T; *Supervised image classification with Khoros - the Classify Toolbox Manual*. Obtido pela Internet pelo endereço: <http://www.mickey.ai.kyutech.ac.jp/~khoros21/>. 1997, 105 pp, em 13/03/97.
- [6] Simpson, P. K.; *Artificial Neural Systems*. Pergamon Press, New York, 1990.

[7] Moraes, R. M.; *Implementação de um sistema contextual de classificação de imagens orbitais compatível com o sistema SITIM*. Dissertação de Mestrado em Engenharia Elétrica. Universidade Federal da Paraíba, 1992.

[8] A. Owen and P. Switzer, A neighbourhood based classifier for LANDSAT Data : Technical Report, Dep. of Statistics, Stanford University, 1982.

[9] N.L. Hjort, Neighbourhood Based Classification of Remotely Sensed Data based on Geometric Probability Models. in Saebo, H.V. et al; Contextual Classification of Remotely Sensed Data: Statistical Methods and Development of a System; Technical Report of Norwegian Computing Center, n.768, 1985, Oslo, Enclosure 4.

Exercícios para nota no Khoros:

Não Superv.: K-Médias, Max-min e Fuzzy C-Médias

Superv.: Paral, MínimaDist., KNN, MáxVer., Mahalanobis e BPNN

Grupo 1: Fuzzy C-Médias, MaxVer, Max-Min.

Grupo 2: KNN, MinDist, K-Médias.

Grupo 3: BPNN, Paralelepípedo, Mahalanobis

Para todos a imagem de teste será a imagem do exemplo do Class. Máx. Verossimilhança.

Entregar: Relatório com pelo menos 10 simulações de cada Método variando seus parâmetros, e as amostras (caso supervisionado) e disquetes com as workspaces, arquivos de amostras e as imagens resultantes (.gif)