3. Sistemas especialistas para classificação de imagens

"Um dia o homem olhará para o céu
e verá que tão pequena quanto aparentam as estrelas no céu
é a sua vã existência. Só aí então, tomará consciência
da grandeza da vida que ele desperdiça"
Autor anônimo.

Várias implementações práticas de classificadores existem na literatura. Como exemplos disso podemos citar:

• o sistema ICARE (Desachy, 1991) que procura gerar um mapa para uma região, com um sistema especialista "nebuloso" agindo sobre uma pré-classificação estatística, mapas e classificações anteriores da região de interesse;

• O trabalho de Haverkamp et al (1994) que combina métodos numéricos e heurísticas para a classificação de blocos de gelo no mar por imagens de radar. É um classificador de 3 passos, onde, no primeiro passo é feita uma pré-classificação estatística simples, no segundo passo extrai-se atributos para no terceiro passo aplicar regras sobre eles e sobre a classificação inicial.

• O trabalho de Mascarilla (1995) que é uma versão do sistema ICARE com conhecimento armazenado em uma rede neural nebulosa, podendo tratar conhecimento expresso em linguagem natural.

• O trabalho de Wendling et al (1995) que faz a interpretação de imagens de satélite por conhecimento acerca das classes procuradas, similar ao ICARE (Desachy está entre os autores). Também existe uma pré-classificação que agora é realizada por uma técnica nebulosa.

• O trabalho de Mathieu-Marni et al (1996) que utiliza um sistema especialista que gerencia uma biblioteca de programas para computação de misturas de classes e classificação de imagens Spot e Landsat.

• O trabalho de Roux e Desachy (1997) que usam informações de um classificador nebuloso e de um sistema especialista dado por uma rede neural nebulosa. A fusão das informações dessas duas fontes é usada para a classificação.

• O trabalho de Queiroz et al (1997) que utiliza um pequeno sistema especialista nebuloso para classificação de axônios, cujas descrições são obtidas por operadores da Morfologia Matemática. Apesar de não ser um classificador de imagens, é interessante observar que esse sistema poderia ser utilizado para tal, com algumas modificações.

3.1. O Sistema de Interpretação de Imagens ICARE

Como o sistema proposto tem uma inspiração no trabalho inicial de Desachy (1991), será dada aqui uma breve descrição do seu trabalho.

O sistema proposto por Desachy é um classificador automático de imagens, simulando a interpretação visual de imagens, denominado ICARE ("Image Cartographie Expert"), com o objetivo de gerar novas cartas topográficas de uma certa região. O número de classes é limitado em 14 e Desachy as chamou de classes *ecoflorísticas*.

O algoritmo de classificação é dividido em três fases, sendo a primeira denominada de *pré-classificação*. Nessa pré-classificação o sistema obtém os valores numéricos para todas as classes possíveis de um "pixel", levando em consideração a radiometria e a textura. Para cada "pixel", a pré-classificação retém apenas os três maiores valores resultantes da função de classificação. A segunda fase, denominada *fase 1 da classificação*, é auxiliada por um sistema especialista, utilizando-se de dados de Modelos Numéricos de Terreno (MNT). Na terceira fase, denominada *fase 2 da classificação*, o sistema especialista utiliza-se de informações sobre Cartas Topográficas anteriores, para aprimorar a classificação. Nas duas últimas fases o sistema mantém três valores resultantes da função de classificação e somente ao final decide pela classe que produz o maior valor da função de classificação.

Nas fases 1 e 2 de classificação, são necessários dados referentes a área em estudo, tais como mapa de estradas, rios, cidades, isolinhas, etc. Uma boa fonte desse tipo de informação são os Sistemas Geográficos de Informação (SIG). Para manipular as informações oriundas do SIG de forma inteligente, Desachy utiliza um sistema especialista. Esse sistema pode, por exemplo, determinar se um "pixel", estando próximo a uma estrada, favorece alguma classificação em especial. Essa determinação se faz com o auxílio da teoria dos conjuntos nebulosos (no caso, o subconjunto *perto)* aplicada a função distância.

O sistema ICARE (Desachy, 1991) faz parte do diagrama esquemático da Figura 3.1. Devemos notar que a fonte de dados do sistema ICARE também pode ser uma classificação gerada anteriormente.

Duas bases de dados são utilizadas por Desachy: a Base de Conhecimento (construída a partir de descrições das zonas florísticas) e, a Base de Dados. Na Base de Conhecimento encontram-se informações sobre a vegetação característica

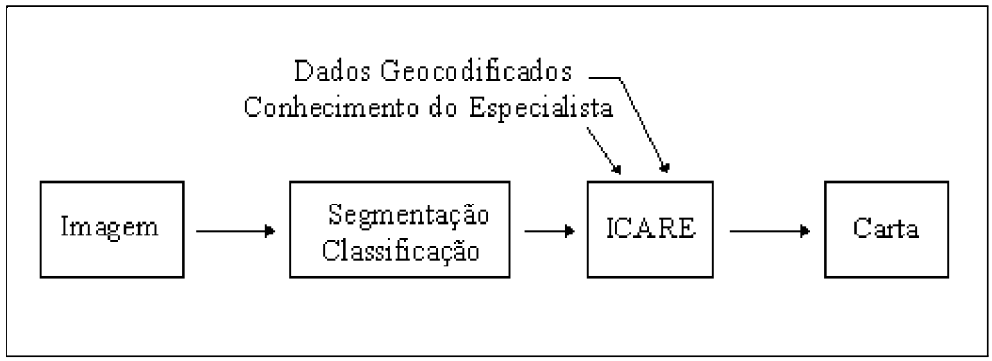

Fig. 3.1: Diagrama geral de um sistema de interpretação de imagens.

das regiões por tipo ou utilização do solo, clima e topografia. Na Base de Dados encontram-se modelos numéricos de cartas climáticas, de solos e topográficas.

As Tabelas 3.1 e 3.2 mostram exemplos de conhecimento armazenados para uma determinada classe (zona ecoflorística, segundo Desachy (1991)). A Tabela 3.1 trata do conhecimento sobre uma classe por dados ecológicos (dados de situação ou localização espacial, clima e tipos de solos) e dados florísticos (área de domínio e tipos de árvores características). A Tabela 3.2 trata do conhecimento sobre cinco classes por dados de localização, tipos cultivados e detecção (tipos de critérios, estação e melhor forma de detecção).

TABELA 3.1 - EXEMPLO DE CONHECIMENTO SOBRE UMA DETERMINADA ZONA ECOFLORÍSTICA POR DADOS ECOLÓGICOS E FLORÍSTICOS.

		Dados	Ecológicos	Dados	Florísticos
Situação	Clima	Solos	Ocupação Humana	Domínio Florístico	Árvores Características
Versante Oeste dos Ghâts ocidentais entre 900 a 1000 m de altitude	Tropical úmido $T = 26^oC$ $T_m = 20^oC$ $dt = 3$ a 4^oC $P > 2000$m 1 à 3 m. s.	ferralíticos às vezes com couraças laterais	Densidade da população muito alta nas partes baixas (600 hab/km^2)	Domínio Ocidental setor Travancore de baixa altitude	dipterocarpaceae cullenia exarillata mesua ferrea disoxylum malabaricum alstonia scholaris

Fonte: Desachy (1991, pg 41).

Note-se que em ambas as tabelas há informações de características locais: os Ghâts e o lago Periyar, localizados na região centro-sul da Índia (Desachy,1991, pg 40).

O processo de classificação utilizado contém três fases, ou etapas, sendo que as duas últimas fazem uso das bases de Conhecimento e de Dados. O objetivo é aperfeiçoar uma classificação inicial (pré-classificação) utilizando um sistema especialista que manipula o conhecimento armazenado na Base de Conhecimento. Esse conhecimento em certas ocasiões depende de dados que estão armazenados na Base de Dados (como altitude, tipo de clima, tipo de solo, etc).

TABELA 3.2 - EXEMPLO DE CONHECIMENTO SOBRE UMA DETERMINADA ZONA ECOFLORÍSTICA POR DADOS DE VEGETAÇÃO.

	Vegetação				
			Detecção		
Tipos espontâneos	Localização	Tipos Cultivados	critérios	melhor forma	melhor estação
1. Floresta densa úmida	- Vestígios nos fundos dos vales.		Textura densidade reflectância	Bandas 5 e 7	Dezembro à Março
2. Floresta densa úmida	- Vestígios		textura densidade	BC-FCC	Março
3. Floresta úmida degradada	- Vestígios		Textura tonalidade		
4. Savana arbórea	- Reservas perto do lago Periyar		Textura	FCC	Indiferente
5. Mosaico de culturas e de vegetação degradada	- Tipo largamente dominante	Mosaico de culturas e de vegetação degradada: - quaisquer coqueirais -pomares+ mandioca+ especiarias -Quaisquer refloresta- mentos de Eucaliptos -rizicultura		FCC	

Fonte: Desachy (1991, pg 41).

Definição das Classes

A metodologia trabalha com informações oriundas de uma imagem multi-espectral, com bandas B_1, B_2, ..., B_b. Porém, é muito oneroso computacionalmente trabalhar com todas as bandas de uma imagem e por esse motivo, é feita uma transformação de Karhünen-Loève, gerando apenas duas bandas $KL1$ e $KL2$. Essas duas bandas contém apenas informações radiométricas. Para obter informações de textura a partir da imagem radiométrica, utilizamos a metodologia de Haralick et al. (1973).

A partir dessas duas bandas $KL1$ e $KL2$, há uma redução dos níveis de cinza (por exemplo 256) para quatro. São obtidos os parâmetros de textura, extraídos de matrizes de co-ocorrências de Haralick et al. (1973) de níveis de cinza, sendo utilizados 10 parâmetros de textura (t_1 a t_{10} para a banda $KL1$ e tt_1 a tt_{10} para a banda $KL2$). Seu objetivo é embutir em cada pixel a informação da textura proveniente de uma vizinhança de 8 pixels ao pixel dado.

Supõe-se que as classes têm distribuição Gaussiana e cada classe é definida pelas médias e variâncias radiométricas das bandas $KL1$ e $KL2$ (2 parâmetros) e pelas médias e variâncias dos parâmetros de textura, também para cada uma das bandas $KL1$ e $KL2$ (são os 10 parâmetros obtidos pelo método de Haralick et al. (1973) para cada uma das bandas, totalizando 20 parâmetros). Assim, essa distribuição Gaussiana possui uma dimensão 22.

$$\begin{aligned} \mu &= \{\ \mu(KL1), \mu(KL2), \mu(\text{texturas em } KL1), \mu(\text{texturas em } KL2)\} \\ \sigma^2 &= \{\ \sigma^2(KL1), \sigma^2(KL2), \sigma^2(\text{texturas em } KL1), \sigma^2(\text{texturas em } KL2)\} \end{aligned}$$

Pode-se notar que não foi mencionada nenhuma matriz de covariâncias para as variáveis de textura. Desachy (1991) assume a independência dessas variáveis, fazendo com que a matriz de covariâncias torne-se uma matriz diagonal, na qual estão presentes somentes as variâncias de cada uma dessas variáveis. Como decorrência, os cálculos são mais rápidos.

Pré-Classificação

A pré-classificação é efetuada utilizando-se um método geométrico de classificação que mede a distância $d(x, y)$ entre a "assinatura" espectral do pixel e cada classe. Uma classe m_i é válida para um determinado "pixel", se a distância d entre as assinaturas espectrais do "pixel" e a classe m_i é menor ou igual a S, onde $S \in R$ é um limiar ($S > 1$), onde:

$$d_i(x,y) = \gamma_1 \left[\left(\frac{I_{KL1}(x,y) - \mu_{IKL1}}{\sigma_{IKL1}} \right)^2 + \left(\frac{I_{KL2}(x,y) - \mu_{IKL2}}{\sigma_{IKL2}} \right)^2 \right] + \\ + \gamma_2 \left[\sum_{j=1}^{10} \left(\frac{t_j - \mu_{tj}}{\sigma_{tj}} \right)^2 + \sum_{j=1}^{10} \left(\frac{tt_j - \mu_{ttj}}{\sigma_{ttj}} \right)^2 \right]$$

Para cada pixel, retemos três classes que possuam as menores distâncias d e que estas sejam inferiores a um limiar S. Calcula-se então uma função $P : [0,1] \rightarrow [0,1]$ (que Desachy (1991) denomina probabilidade em seu trabalho) para cada uma dessas três classes:

$$P_i(x,y) := 1 - (d_i(x,y)/S), i = 1,2,3 \qquad (x,y) \in E, \tag{3.1}$$

que serão retidas pelo sistema para aquele ponto (x,y).

É digno de nota que para o cálculo de $d_i(x,y)$, leva-se em consideração alguns fatores auxiliares: γ_1 e γ_2 (assumindo valores no intervalo $[0,1]$) (Desachy, 1991). Esses fatores ponderam a contribuição das informações radiométricas e texturais na classificação.

Primeiro Nível de Classificação

O conhecimento do especialista intervirá nas duas últimas fases da classificação, fazendo com que o valor da função P_i de uma classe m_i de um pixel (x,y) permaneça inalterada, diminua menos ou mais, caso sua localização seja totalmente favorável, favorável ou desfavorável a essa determinada classe. Isso ocorre em duas etapas, de acordo com o tipo de característica tratada (características de nível 1 ou de nível 2) - Figura 3.2.

A construção da base de conhecimento deve satisfazer a dois critérios:

1) Facilitar a expressão do conhecimento pelo especialista sobre as classes

2) Facilitar o cálculo para cada ponto do seu contexto para cada classe possível.

Segundo Desachy (1991), isso o obriga a definir de 4 tipos de características, a partir das quais o conhecimento será representado:

- CTO (características do tipo de objeto): indica qual entidade (objeto) está sendo considerada, dentro de um determinado contexto de classe.

Exemplo: vila, estrada, etc.

□

- CTR (características do tipo relacional): descreve a relação entre a classe e o objeto definido pelo CTO.

Exemplo: perto, longe, etc.

□

- CTP (características do tipo de presença): define um grau de incerteza sobre a descrição de uma classe, que é quantificado em p_{pres}. Desachy (1991) apresenta esse grau p_{pres} de forma tabelada e o denomina "probabilidade de presença".

Exemplo:

CTP	sempre	principalmente	predominante	frequente	presente	jamais
p_{pres}	1.00	0.85	0.70	0.55	0.40	0.00

(3.2)

□

- CTS (características de tipo espacial): exprime certas características de classes de nível 2 (relacionadas ao contexto, como superfície, forma, etc).

A partir dessas 4 características representa-se o conhecimento sobre cada classe, dependendo do nível em que atua:

Para o nível 1:
((CTP.CTR.CTO) E (CTP.CTR.CTO) E ...)
OU
((CTP.CTR.CTO) E (CTP.CTR.CTO) E ...)
OU ...

Exemplo:

((COMUM.	PERTO.	VILA)	OU	(COMUM.	PERTO.	RUA))
CTP	CTR	CTO		CTP	CTR	CTO

□

Para o nível 2:
((CTP.CTS) E (CTP.CTS) E ...)
OU
((CTP.CTS) E (CTP.CTS) E ...)

Exemplo:
"Todos os cumes altos antigos ou fundos de vales antigos"
para o nível 1: (todos.altos.cumes) OU (todos.fundos.vales)
para o nível 2: (todos.antigos)

□

As características CTP, CTR e CTO são utilizadas no primeiro nível de classificação, para melhorar a pré-classificação. Essas características participam na construção do fator $b(x, y)$ que é um fator de correção. Esse fator modifica o valor da função P de cada "pixel". O princípio utilizado é aumentar a diferença entre os três valores da função P para um dado "pixel" (x, y) se o contexto é favorável a uma das classe dada e diminuí-la caso contrário.

Sendo assim, a função de classificação $P_i^{(1)}(x, y)$ para o nível 1, é dada por:

$$P_i^{(1)}(x,y) := b_i(x,y) * P_i(x,y), \tag{3.3}$$

onde $b_i(x,y)$ é o fator de correção global para o ponto (x,y) e classe m_i.

O fator de correção global $b_i(x,y)$ depende da característica global de nível 1 da classe m_i (conhecimento armazenado na base de conhecimento) e da localização do "pixel" em relação à base de dados geocodificados. Sua forma de cálculo é dada a seguir:

A partir de cada característica parcial da forma $(CTP_i.CTR_i.CTO_i)$, produz-se um fator de correção parcial $pcar_i(x,y)$. CTP_i define uma função de presença (que no trabalho de Desachy (1991) é denominada "probabilidade de presença"), já mencionada. CTR_i introduz um nível de pertinência nebuloso ao objeto CTO_i, $appar(x,y)$ que pode ser interpretado como o grau de realização (valor no intervalo $[0,1]$) de uma característica "proximidade".

Exemplo:

Se um "pixel " (x,y) situa-se a uma distância de 200 metros de uma rua, o grau de realização da característica "perto de rua" é fornecido por uma função de pertinência de um sub-conjunto nebuloso "perto de rua". Por essa função, necessariamente, quanto mais próximo o ponto for da rua, maior será o seu grau de realização e a medida em que a distância for maior, o grau diminui. Em termos práticos, isso pode ser obtido por uma transformação sobre a *função distância* (vista na seção anterior) de modo a fornecer altos graus de realização quanto menor for a distância entre a rua e o ponto em questão para o referido exemplo. □

Por fim, o fator de correção parcial $pcar_i(x,y)$ para a característica particular $(CTP_i.CTR_i.CTO_i)$ é definido por:

$$pcar_i(x,y) := p_{pres} * appar(x,y) + (1 - p_{pres}) * (1 - appar(x,y))$$

Dessa forma, quando $appar(x,y) = 1$ temos que $pcar_i(x,y) = p_{pres}$ e quando $appar(x,y) = 0$ temos que $pcar_i(x,y) = 1 - p_{pres}$ ou seja:

$$pcar_i(x,y) = \begin{cases} p_{pres} & \text{se} \quad (CTR_i.CTO_i) \text{ é verdadeiro a } 100\% \\ 1 - p_{pres} & \text{se} \quad (CTR_i.CTO_i) \text{ é falso a } 100\% \end{cases} .$$

O fator de correção global $b_i(x,y)$ é calculado com base em todos os fatores de correção parcial $pcar_i(x,y)$. Para cada composição de interseções das características particulares, calcula-se um fator intermediário $b^{(1)}(x,y)$ que é definido por:

$$b^{(1)}(x,y) := \min_k(pcar_k(x,y))$$

onde k é a quantidade de características particulares que estão presentes na composição de interseções.
Exemplo:

$$\underset{b_1^{(1)}(x,y) := \min_{k_1}(pcar_{k_1}(x,y))}{(CTP_1.CTR_1.CTO_1 \cap CTP_2.CTR_2.CTO_2 \cap ...)} \quad \cup \quad \underset{b_2^{(1)}(x,y) := \min_{k_2}(...)}{(...)} \quad \cup \quad ...$$

□

Para compor toda a informação, é necessário analisar a composição como um todo, ou seja, analisar também as uniões. O cálculo é dado por:

$$b_i(x,y) := \max_j(b_j^{(1)}(x,y))$$

para a classe m_i e j são os atributos observados pelo sistema especialista, como vila, rio, estrada, etc.

Assim, finaliza-se essa fase de classificação, obtendo como resultado para cada "pixel", os três valores revisados:

$$\begin{aligned} P_1^{(1)}(x,y) &= b_1(x,y) * P_1(x,y) \quad \text{para a classe } m_1 \\ P_2^{(1)}(x,y) &= b_2(x,y) * P_2(x,y) \quad \text{para a classe } m_2 \\ P_3^{(1)}(x,y) &= b_3(x,y) * P_3(x,y) \quad \text{para a classe } m_3 \end{aligned}$$

Segundo Nível de Classificação

Para o nível 2, são utilizadas as características CTP e CTS das classes. No entanto, os cálculos são análogos ao nível 1, inclusive a própria nomenclatura atribuída por Desachy permanece inalterada. O cálculo da função $P_i^{(2)}(x,y)$ é dado por:

$$P_i^{(2)}(x,y) := k_i(x,y) * P_i^{(1)}(x,y), \tag{3.4}$$

onde $k_i(x,y)$ é o fator de correção global no segundo nível de classificação, para o ponto (x,y) e classe m_i.

As características globais do nível 2 diferem das do nível 1 pela utilização do CTS e CTP e, são do tipo:

$$(CTP_1.CTS_1) \quad \cap \quad (CTP_2.CTS_2) \quad \cap ...) \quad \cup \quad (CTP...)$$

Para cada característica parcial $(CTP.CTS)$ calcula-se um fator de correção parcial:

$$pcar_i(x,y) := p_{pres} * appar(x,y) + (1 - p_{pres}) * (1 - appar(x,y))$$

onde : p_{pres} descreve a confiança (CTP_i) que se dá à afirmação CTS_i, ou seja:

$$p_{pres} := \begin{cases} 1, & \text{certeza a } 100\% \\ 0, & \text{falso a } 100\% \end{cases} \tag{3.5}$$

e, $appar(x, y)$ é o grau de realização de CTS_i no ponto (x, y).
Exemplo:

Suponha-se que uma formação rochosa em questão é antiga. No sistema ICARE, "antigos" será considerado como "sempre antigos" e então $p_{pres} = 1$ (Desachy, 1991, pg84).
□

O grau $appar(x, y)$ é calculado em função de uma superfície S da região onde se encontra o ponto, em relação a dois limiares pré-definidos S_1 e S_2 (Figura 3.3). Um possível exemplo, é a de classes comuns em determinadas faixas de altitudes que ficam evidenciadas nas imagens dos Modelos Digitais de Terreno.

O mínimo entre todas as imagens $pcar$ resulta na imagem $k^{(2)}$:

$$k_i^{(2)} = \min_j(pcar_j(x, y))$$

para a classe m_i e j é a quantidade de características particulares que estão presentes na composição de interseções.

Novamente, para compor toda a informação, é necessário analisar a composição como um todo, analisando também as uniões, se elas existirem. O cálculo é dado por:

$$k_i(x, y) := \max_j(k_j^{(2)}(x, y))$$

onde j são os atributos observados pelo sistema especialista para a classe m_i.

Finaliza-se então a segunda fase de classificação, obtendo como resultado para cada "pixel", os três valores revisados:

$$\begin{aligned} P_1^{(2)}(x, y) &= k_1(x, y) * P_1^{(1)}(x, y) \quad \text{para a classe } m_1 \\ P_2^{(2)}(x, y) &= k_2(x, y) * P_2^{(1)}(x, y) \quad \text{para a classe } m_2 \\ P_3^{(2)}(x, y) &= k_3(x, y) * P_3^{(1)}(x, y) \quad \text{para a classe } m_3 \end{aligned}$$

E, finalmente, é feita a atribuição do "pixel" (x, y) à uma classe m_i das três classes inicialmente selecionadas, de forma que:

$$\begin{aligned} P_i^{(2)}(x, y) \quad &: \quad = \max_k(P_k^{(2)}(x, y)) \\ \text{classe } (x, y) \quad &= \quad m_i \end{aligned} \tag{3.6}$$

onde k denota as três classes selecionadas para a maximização da função $P^{(2)}$ para cada "pixel".

3.2. Considerações sobre o sistema ICARE

O trabalho de Desachy (1991) deixa transparecer a abordagem "ad-hoc". A fase de pré-classificação utiliza-se de uma metodologia complexa de matrizes de textura, enquanto a primeira e segunda fases se encarregam de utilizar dados provenientes de bases de dados e de conhecimento para aperfeiçoar a classificação inicial gerada.

O sistema é pouco formal e as medidas ditas "probabilidades", na realidade não podem ser vistas como tal. O tratamento numérico não é uniforme, misturando "probabilidades" e fatores de correção com interpretações diferenciadas. Isso torna o algoritmo bastante complexo e de difícil compreensão. Por isso o ICARE utiliza-se de uma linguagem para a composição das regras para auxiliar o usuário a interagir com esse complexo sistema.