1. INTRODUÇÃO

A Morfologia Matemática [Banon (1994)] é uma teoria que estuda as decomposições de operadores (mapeamentos ou transformações) entre reticulados completos em termos de algumas famílias de mapeamentos simples, como a erosão, a dilatação e outras operações derivadas dessas. A aplicação dessa teoria se dá nas áreas onde o processamento digital de imagens possa atuar e exemplos disso podem ser encontrados em vasta bibliografia.

A Morfologia Matemática nasceu na França por volta de 1964 com o objetivo de analisar estruturas geométricas em imagens microscópicas de amostras de rochas ou metais. Seus mentores, George Matheron e Jean Serra, criaram duas classes de operadores elementares: as dilatações e as erosões que eram baseadas na soma e na subtração de Minkowski, respectivamente [Serra (1982)]. O nome "Morfologia " tem origem grega e significa *estudo das formas*.

Originalmente, as imagens trabalhadas pela Morfologia Matemática eram binárias. As aplicações em imagens de níveis de cinza passavam necessariamente pela binarização. Porém com a necessidade, os conceitos foram estendidos para as imagens em níveis de cinza, inicialmente por Sternberg [Serra (1982)] e melhor formalizado matematicamente por Heijmans (1991). A Morfologia Matemática também sofreu profundas alterações nos seus rumos e atualmente não se aplica somente ao estudo de formas. A sua maior formalização matemática fez com que despontassem novas aplicações nas quais não se reconhece mais a conotação semântica da erosão como "desbaste de formas" e a dilatação como "engorda de formas" (que se aplica quando a origem do domínio das imagens pertence ao elemento estruturante [Banon (1994)]).

O nosso objetivo é relatar as definições dadas para esses operadores elementares definidos sobre o domínio das imagens binárias e em níveis de cinza com a finalidade de mostrar que há a possibilidade de aplicá-las à Classificação de Imagens.

A organização do estudo consta das definições da dilatação, erosão, anti-dilatação e anti-erosão da Morfologia Matemática, seguidas de uma breve introdução aos conceitos mais básicos dos Conjuntos Nebulosos. Depois, a decomposição da função distância em operadores da Morfologia Matemática (o que se mostrará útil em classificação de imagens). Na seqüência, veremos como esses conceitos podem ser aplicados à Classificação de Imagens, partindo de um trabalho de Desachy (1991). Finalizando, apresentamos algumas conclusões e sugestões para trabalhos futuros.

Em diversas situações, diagramas esquemáticos resumirão partes expressivas dos métodos utilizados. O leitor pode visualizar a simbologia gráfica utilizada no decorrer desse trabalho no apêndice.

2. CONCEITOS DA MORFOLOGIA MATEMÁTICA

Para maior facilidade no entendimento do trabalho, nos limitaremos a tratar somente os conceitos da Morfologia Matemática que serão utilizados nesse trabalho.

Definição 1-1. [Birkhoff (1967)] Um conjunto parcialmente ordenado $(L_1, \leq_{L_1})$ que para quaisquer dois de seus elementos, denotados por a e b, possuir um maior limitante inferior (ínfimo), denotado por $a \wedge_1 b$ e um menor limitante superior (supremo), denotado por $a \vee_1 b$, é um *reticulado.*

Definição 1-2. [Birkhoff (1967)] Um reticulado L_1 é um *reticulado completo* quando cada um dos seus subconjuntos X tem um supremo e um ínfimo em L_1.

Definição 1-3. Seja $(K_1, \leq_{K_1})$ um conjunto parcialmente ordenado. Se a relação $\leq_{K_1}$ é uma relação de ordem total, então K_1 é uma *cadeia.*

Definição 1-4. Uma cadeia que possua um maior elemento e um menor elemento é uma *cadeia limitada.*

Exemplos de cadeias limitadas:
$K_1 = \{0, 1\}$
$K_2 = \{0, 1, 2, 3\}$
$K_3 = \{0, 1, ..., 255\}$
□

Proposição 1-1. Toda cadeia limitada é um reticulado.

A utilidade da cadeia limitada para nosso estudo reside no fato de que o contradomínio de uma imagens em níveis de cinza é uma cadeia limitada (há uma ordenação e limites bem definidos). Nos exemplos acima, K_1 representa o conjunto de níveis de cinza de uma imagem binária e K_3 é o conjunto de níveis de cinza das imagens usadas na maioria das aplicações.

Podemos assim, visualizar mapeamentos entre imagens em níveis de cinza, entre imagens binárias, ou mesmo de imagens em níveis de cinza para imagens binárias ou vice-versa, como mostra a figura 2-1.

Definição 1-5. Uma transformação Ψ entre imagens é *crescente* se:

$$\forall a, b \in L_1, a \leq_{L_1} b \Rightarrow \Psi(a) \leq_{L_2} \Psi(b)$$

Definição 1-6. Uma transformação Ψ entre imagens é *decrescente* se:

$$\forall a, b \in L_1, a \leq_{L_1} b \Rightarrow \Psi(b) \leq_{L_2} \Psi(a)$$

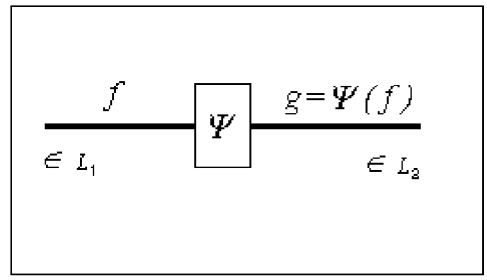

Figure 2-1 Transformação sobre imagens.

A definição axiomática é dada para algumas classes de operadores de L_1 em L_2 como se segue:

Definição 1-7. Seja Ψ, um operador de L_1 em L_2, sendo L_1 e L_2 dois reticulados finitos, com $i_1 = \sup L_1$ ($i_2 = \sup L_2$) e $o_1 = \inf L_1$ ($o_2 = \inf L_2$):

- Ψ é uma *dilatação* $\Leftrightarrow \Psi(a \vee_1 b) = \Psi(a) \vee_2 \Psi(b)$ e $\Psi(o_1) = o_2$
- Ψ é uma *erosão* $\Leftrightarrow \Psi(a \wedge_1 b) = \Psi(a) \wedge_2 \Psi(b)$ e $\Psi(i_1) = i_2$
- Ψ é uma *anti-dilatação* $\Leftrightarrow \Psi(a \vee_1 b) = \Psi(a) \wedge_2 \Psi(b)$ e $\Psi(o_1) = i_2$
- Ψ é uma *anti-erosão* $\Leftrightarrow \Psi(a \wedge_1 b) = \Psi(a) \vee_2 \Psi(b)$ e $\Psi(i_1) = o_2$

No caso K_1 e K_2 serem duas cadeias limitadas, a definição se simplifica, tornando-se um caso particular da proposição 1.4 de Banon (1995).

Definição 1-8. Sejam K_1 e K_2 duas cadeias limitadas. Seja também Ψ um mapeamento de K_1 para K_2, então:

- Ψ é uma *dilatação* $\Leftrightarrow \Psi$ é crescente e $\Psi(\min K_1) = \min K_2$
- Ψ é uma *erosão* $\Leftrightarrow \Psi$ é crescente e $\Psi(\max K_1) = \max K_2$
- Ψ é uma *anti-dilatação* $\Leftrightarrow \Psi$ é decrescente e $\Psi(\min K_1) = \max K_2$
- Ψ é uma *anti-erosão* $\Leftrightarrow \Psi$ é decrescente e $\Psi(\max K_1) = \min K_2$

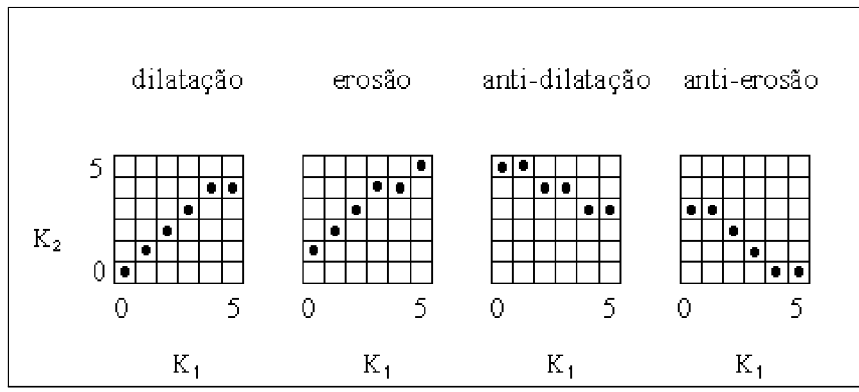

Figure 2-2 Exemplos de operadores morfológicos elementares.

Na figura 2-2, podemos visualizar exemplos de operadores elementares respectivamente de dilatação, erosão, anti-dilatação e anti-erosão, em acordo com as definições 1-7 e 1-8, usando $K_1 = K_2$.

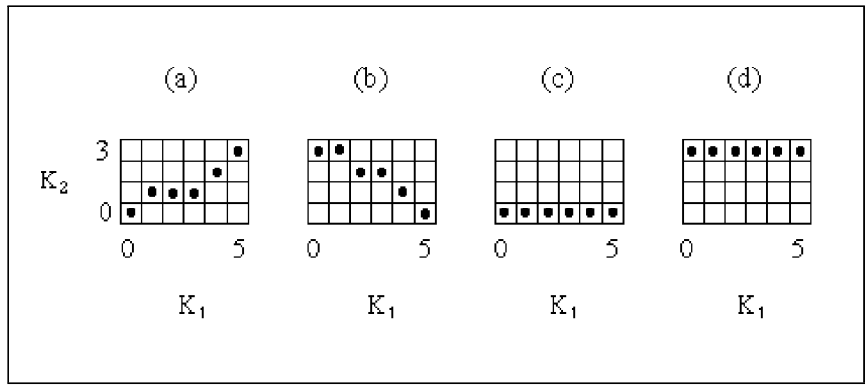

Figure 2-3 Exemplos de alguns operadores morfológicos de dupla função.

É importante salientar que um operador Ψ pode vir a não se enquadrar em apenas um tipo. A título de ilustração, mostramos na figura 2-3 quatro exemplos onde os operadores são, ao mesmo tempo: (a) dilatação e erosão, (b) anti-dilatação e anti-erosão, (c) dilatação e anti-erosão e (d) erosão e anti-dilatação. Fica fácil notar, pela definição 1-8, a duplicidade de tipos desses operadores.

As transformações elementares da Morfologia Matemática (dilatação, erosão, anti-dilatação e anti-erosão) entre imagens em níveis de cinza podem ser caracterizadas de maneira diferente por resposta impulsiva (função de espalhamento pontual, ou borrão, ou por elementos estruturantes, ou por família de ELUTs) [Banon (1995)]. As LUTs ("Look Up Tables") são mapementos entre níveis de cinza e as ELUTs ("Elementary Look Up Tables") são LUTs com propriedades de transformações elementares da Morfologia Matemática (dilatação ou erosão ou anti-dilatação ou anti-erosão).

Em particular, na Morfologia Matemática binária [Serra (1982)], as definições de dilatação, erosão podem ser dadas de formas equivalentes mais simples. Nesse caso, cabe algumas definições iniciais.

Definição 1-9. Seja E um espaço arbitrário e seja $P(E)$ a coleção de todos os seus subconjuntos X, ou seja: $P(E) = \{X : X \subset E\}$. A coleção $P(E)$ provida da relação de inclusão ($\subset$), forma um conjunto parcialmente ordenado, denotado por $(P(E), \subset)$. Essa ordenação garante a existência do maior conjunto (o próprio E) e do menor conjunto (o conjunto vazio, denotado por $\emptyset$).

Definição 1-10. Dado um conjunto X, sua translação por um vetor t, sua transposição (simetria com relação a origem), são denotados por: X_t, X^c e X^t, respectivamente:

$$\begin{aligned} X_t &= X + t = \{x : x - t \in X\}; \\ X^c &= \{x \in E : x \notin X\}; \\ X^t &= \{x : -x \in X\}. \end{aligned}$$

Nesse momento, é possível definir a dilatação e a erosão da Morfologia Matemática binária:

Definição 1-11. A dilatação do conjunto X por um borrão $B : E \to P(E)$ é definido por:

$$X \oplus B^t = \cup_{b \in B^t} X_b.$$

onde $X \oplus B$ é a clássica adição de Minkowski entre dois conjuntos [Serra (1982)].

Obs.: A equivalência entre dilatação por "borrão" e dilatação por "elemento estruturante" é dada em Banon (1995).

Definição 1-12. A erosão do conjunto X por um elemento estruturante B é definido por:

$$X \ominus B = \{y \in E : X \subset B + y\}.$$

onde $X \ominus B$ é a clássica subtração de Minkowski entre dois conjuntos [Serra (1982)].

Na prática, a grande maioria dos elementos estruturantes são simétricos, isto é, $B \equiv B^t$, e sua diferença pode ser ignorada, como no caso dos quadrados, discos, cruzes, etc.

As demais definições de anti-dilatação e anti-erosão binárias podem ser encontradas em Banon (1995).

3. CONJUNTOS NEBULOSOS ("FUZZY SETS")

3.1 Introdução

As origens dos Conjuntos Nebulosos, remontam ao ano de 1965, quando o professor Lofti A. Zadeh publicou um artigo que é considerado o marco-zero da teoria: "Fuzzy Sets", onde lança as bases matemáticas desse tipo de conjunto [Zadeh (1965)]. A partir daí, inúmeras aplicações e destinos têm sido dados a essa teoria, nos mais variados campos como estatística, inteligência artificial, reconhecimento de padrões e, programação linear, entre muitos outros.

Nosso objetivo é introduzir algumas definições básicas sobre os conjuntos nebulosos que serão úteis no decorrer de nossas discussões. A título meramente ilustrativo, forneceremos algumas operações sobre esse tipo de conjuntos.

3.2 O que são "fuzzy sets".

Um conjunto nebuloso é definido por uma função de pertinência μ, que associa a todo elemento x do universo de objetos X um valor real $\mu(x)$ no intervalo $[0,1]$. A função μ substitui a noção de função característica da Teoria de Conjuntos clássica que associa a todo elemento x do universo de objetos X um valor inteiro $\{0,1\}$.

Formalizando matematicamente, segue-se a definição [Zadeh(1965)]:

Definição 3-1. Seja X um espaço de pontos, com um elemento genérico de X denotado por x. Um *conjunto nebuloso* A em X é caracterizado por uma *função de pertinência* $\mu_A(x)$ que associa a cada ponto em X um número real no intervalo $[0,1]$, onde $\mu_A(x)$ representa para x o seu grau de pertinência em A.

De acordo com a definição acima, se $\mu_A(x) = 1$ significa que x é completamente compatível com o conjunto A em X e se $\mu_A(x) = 0$ significa que x é completamente incompatível com o conjunto A em X.

3.3 Operações com conjuntos nebulosos.

Definimos operações sobre os conjuntos nebulosos, similares aos da Teoria de Conjuntos clássica [Kandel (1986)]:

Definição 3-2. Um conjunto nebuloso é vazio se e somente se a sua função de pertinência é identicamente zero sobre X.

Definição 3-3. Dois conjuntos nebulosos A e B são iguais ($A = B$), se e somente se $\mu_A(x) = \mu_B(x)$ para todo x em X.

Definição 3-4. O complemento de um conjunto nebuloso A em X, denotado por $\bar{A}$ é definido por:

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \tag{3.1}$$

Definição 3-5. A união de dois conjuntos nebulosos A e B, com suas funções de pertinência $\mu_A(x)$ e $\mu_B(x)$ resulta um conjunto nebuloso C, cuja função de pertinência é dada por:

$$C = A \cup B \Leftrightarrow \mu_C(x) = \max\{\mu_A(x), \mu_B(x)\}, \forall x \in X. \tag{3.2}$$

Definição 3-6. A intersecção de dois conjuntos nebulosos A e B, com suas funções de pertinência $\mu_A(x)$ e $\mu_B(x)$ resulta um conjunto nebuloso C, cuja função de pertinência é dada por:

$$C = A \cap B \Leftrightarrow \mu_C(x) = \min\{\mu_A(x), \mu_B(x)\}, \forall x \in X. \tag{3.3}$$

OBS.: Podemos notar facilmente que $A \cap \bar{A} \neq \emptyset$ e portanto $A \cup \bar{A} \neq X$ (ver equações (3.2) e (3.3) a seguir).

Pode ser encontrada na literatura outras operações sobre os conjuntos nebulosos [Kandel (1986)], porém sem interesse ao trabalho proposto. O exemplo a seguir ilustra as definições acima:

Exemplo:

Dado um universo X, um conjunto nebuloso A pode ser denotado através de $A = \sum \mu(x)/x$, mostrando para cada x a sua pertinência correspondente.

Seja o universo:

$$X = \{1, 2, 3, 4, 5, 6, 7\};$$

e sejam dois conjuntos nebulosos A e B em X:

$$A = 0.8/3 + 1/5 + 0.6/6,$$

$$B = 0.7/3 + 1/4 + 0.5/6;$$

onde, segundo a notação, o conjunto nebuloso A, o elemento 3 possui um grau de pertinência 0.8, o elemento 5 possui um grau 1 e o elemento 6 possui um grau 0.6 e de forma similar para o conjunto nebuloso B.

Então:

$$\begin{aligned} A \cup B &= 0.8/3 + 1/4 + 1/5 + 0.6/6, \\ A \cap B &= 0.7/3 + 0.5/6; \\ \bar{A} &= 1/1 + 1/2 + 0.2/3 + 1/4 + 0.4/6 + 1/7, \end{aligned}$$

□

4. A FUNÇÃO DISTÂNCIA

Seja Z o conjunto dos inteiros e seja E um retângulo de Z^2 representando um subconjunto da grade quadrada. A função distância é uma transformação $\Psi : K_1^E \rightarrow K_2^E$, onde $K_1 = \{0, k_1\}$ e $K_2 = \{0, ..., k_2\}$ com $k_1, k_2 \in Z^+$, dada por [Barrera et al (1995)]:

$$\Psi_d(f)(x) = d(x, \{y \in E : f(y) = 0\})$$

onde

$$d(x, X) := \min \left\{ \left\lfloor d(x, y) + \frac{1}{2} \right\rfloor : y \in X \right\} \qquad X \subset E$$

onde $\lfloor a \rfloor$ retorna o menor valor inteiro maior que o valor a, e $d : E \times E \rightarrow R$, é uma distância. Ou seja, essa transformação associa a cada "pixel" de valor k_1 da imagem de entrada, sua distância ao mais próximo "pixel" com valor 0.

A função distância é frequentemente usada como ferramenta auxiliar para determinar transformações morfológicas como por exemplo, a erosão última e o esqueleto ou o SKIZ ("Skeleton by Influence Zones", ou esqueleto por zonas de influência). Dela pode ser igualmente obtida dilatações por disco de todos os diâmetros [Vincent (1990) pg 57]. Pode-se determinar a função distância do objeto ao fundo da imagem e depois limiarizar, o que produz erosões por discos. Com combinações dessas últimas operações pode-se chegar a aberturas e fechamentos por discos.

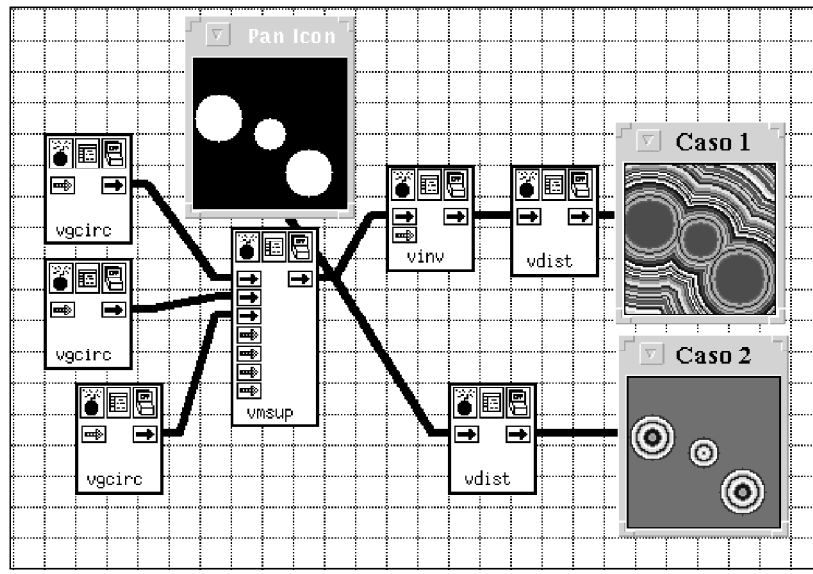

Figure 4-1 Aplicação da função distância a uma imagem binária com três círculos.

Na figura 4-1 pode-se ver o resultado da função distância aplicada a uma imagem binária com três círculos (onde o valor 0 é associado a cor preta e o valor 1 a branca). No caso 1, a imagem de entrrada é invertida e obtemos a distância fora do objeto (o valor 0 é associado a cor vermelha). No caso 2, a imagem não é invertida e obtemos a distância do centro do objeto às suas bordas (o valor zero é associado a cor azul).

Nesse trabalho, apresentamos uma construção da função distância em termos de operadores morfológicos - figura 4-2 (caso da função distância City Block). Nós observamos que a função distância é uma união em paralelo de composições de erosões por elementos estruturantes $B \subset E$ e, transformações por "Elementary Look Up Tables" (ELUTs) [Banon (1995)]. Nesse caso, os elementos estruturantes B são discos [Vincent (1990)]. Podemos notar na figura 4-2, o elemento estruturante nB, que é uma sucessão de $n-1$ composições do elemento estruturante B, isto é, $nB = (B \oplus B) ... \oplus B$, para $n \geq 1$ (quando $n = 0$, $nB = \{o\}$) [Barrera et al. (1995)]. As ELUTs são funções de K_1 em K_2 e essas ELUTs são funções de $\{0, k\}$ em $K = \{0, 1, ..., n+m\}$, onde k é um valor inteiro e a cadeia K depende das dimensões $n \times m$ (onde n é o número de linhas e m é o número de colunas) da imagem em questão.

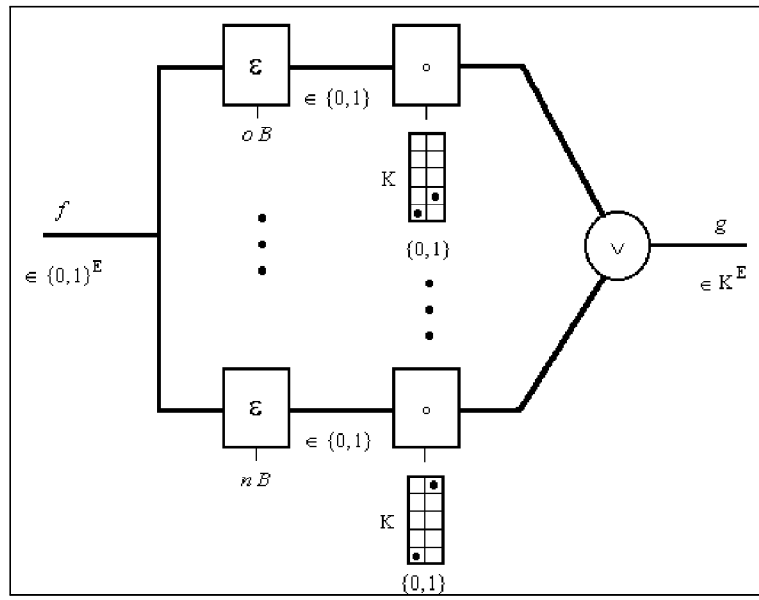

Figure 4-2 A função distância em termos de operadores morfológicos.

Além das aplicações mencionadas da função distância (dilatação, erosão, esqueleto, etc), uma outra em particular podem ser vista em [Desachy (1991)] em sistemas automáticos de classificação. Desachy (1991) utiliza-se da função distância para construir um dos parâmetros necessários para um sistema automático de classificação. Sob a inspiração desse último trabalho mencionado, será trabalhada a próxima seção.

5. A INTERPRETAÇÃO DE IMAGENS DE SATÉLITE

5.1 Introdução

Nessa seção, analisamos a possibilidade de se viabilizar um classificador de imagens assistido por um sistema especialista que se utiliza de operadores morfológicos agindo sobre imagens, inspirado no sistema ICARE de Desachy (1991).

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.* Dessa pré-classificação o sistema retém os três maiores valores resultantes da função de classificação para cada "pixel ", levando em consideração a radiometria e a textura. 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.

5.2 Apresentação Geral do Sistema

O sistema ICARE [Desachy (1991)] faz parte do diagrama esquemático da figura 5-1. Devemos notar que a fonte de dados do sistema ICARE também pode ser uma

classificação gerada anteriormente.

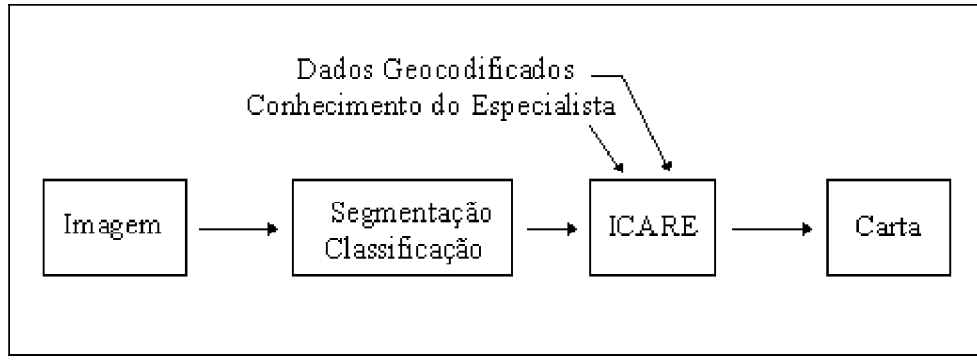

Figure 5-1 Diagrama geral de um sistema de interpretação de imagens.

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 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.

Tabela 5.1. 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).

As tabelas 5.1 e 5.2 mostram exemplos de conhecimento armazenados para uma determinada classe (zona ecoflorística, segundo Desachy (1991)). A tabela 5.1 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 5.2 trata do conhecimento sobre cinco classes por dados de localização, tipos cultivados e teledetecção (tipos de critérios, estação e melhor forma de detecção).

Tabela 5.2. Exemplo de conhecimento sobre uma determinada zona ecoflorística, por dados de vegetação.

	Vegetação				
			Teledetecção		
Tipos espontâneos	Localização	Tipos Cultivados	critérios	melhor forma	melhor estação
1. Floresta densa úmida sempervirente	- Vestígios nos fundos dos vales.		Textura densidade reflectância	Bandas 5 e 7	Dezembro à Março
2. Floresta densa úmida semi-décidue	- Vestígios		textura densidade	BC-FCC	Março
3. Floresta úmida degradada	- Vestígios		Textura tonali-dade		
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+ manidoca+ especiarias -Quaisquer refloresta-mentos de Eucaliptos -rizicultura		FCC	

Fonte: Desachy (1991, pg 41).

Devemos notar que em ambas as tabelas há informações de características meramente locais: os Ghâts e o lago Periyar, localizados na região centro-sul da Índia [Desachy (1991, pg 40)] - o que nos leva a concluir que se trata de uma Base de

Conhecimento bastante específica.

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).

5.2.1 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, fazemos 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 (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 (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.

A figura 5-2 mostra a versão morfológica do processo descrito. A entrada é composta pela imagem I, em suas bandas $KL1$ e $KL2$. A redução dos níveis de cinza é dada por uma transformação por uma ELUT para cada uma das bandas $KL1$ e $KL2$. Pode-se notar também que essa transformação é, ao mesmo tempo, uma dilatação e uma erosão (ver figura 2-3). Após essa transformação, são obtidos os parâmetros de textura pela metodologia de Haralick (1973), gerando 10 parâmetros de textura para cada banda de entrada.

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 (1973) para cada uma das bandas, totalizando 20 parâmetros). Assim, essa distribuição Gaussiana possui uma dimensão 22.

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

Podemos 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

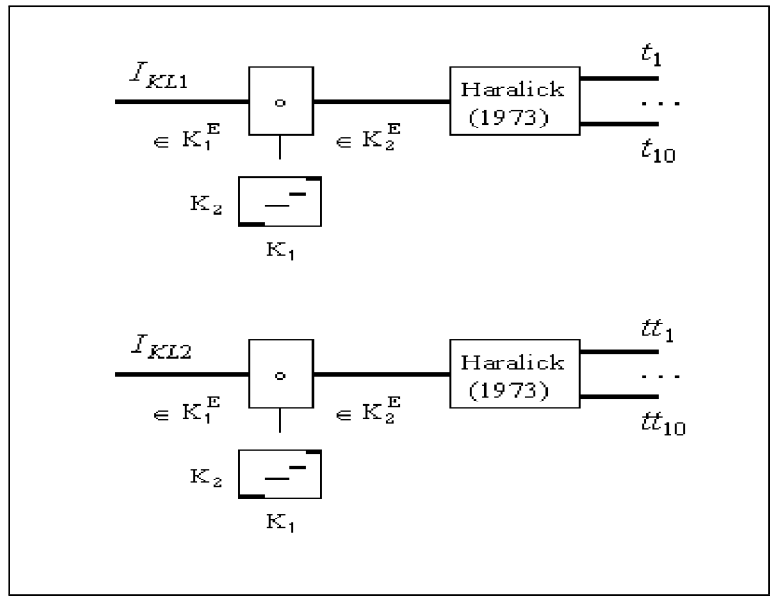

Figure 5-2 A obtenção dos parâmetros de textura.

somentes as variâncias de cada uma dessas variáveis. Como decorrência, os cálculos são mais rápidos.

5.2.2 Pré-Clasificação

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:

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

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

$$P_i(x,y) := 1 - \left(d_i(x,y)/S \right), i = 1,2,3 \qquad (x,y) \in E, \tag{5.1}$$

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.

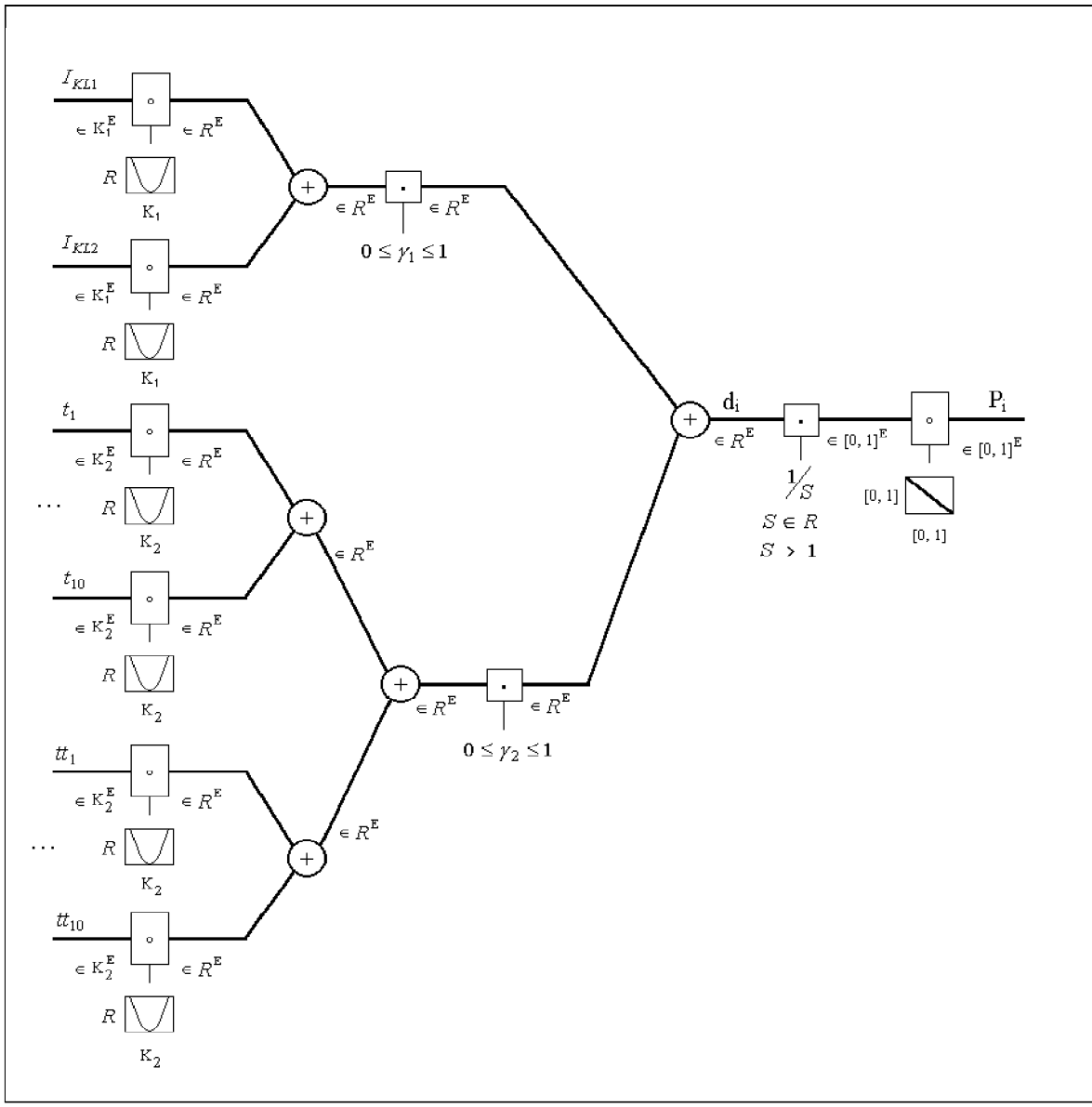

Figure 5-3 Diagrama esquemático da pré-classificação.

O diagrama esquemático da figura 5-3 mostra esses passos da pré-classificação. Podemos notar a profusão de transformações das imagens por LUTs na primeira "camada". Estas transformações podem ser decompostas em uniões de anti-erosões (anti-dilatações) e dilatações (ou erosões), conforme ilustra a figura 5-4.

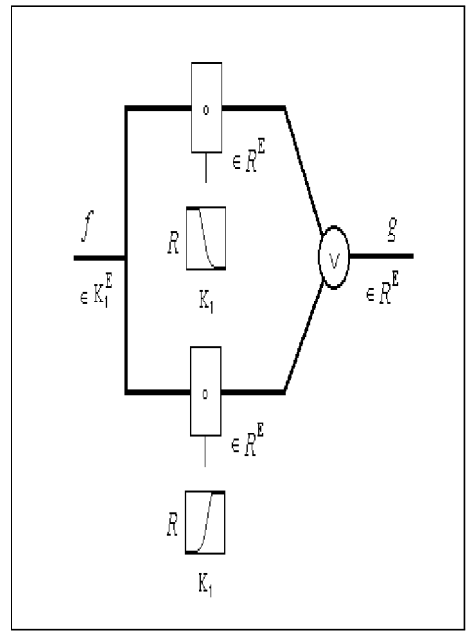

Figure 5-4 Decomposição de algumas LUTs da pré-classificação.

Notamos também, na figura 5-3, os fatores γ_1 e γ_2 que ponderam as contribuições de radiometria e textura na classificação. A operação de adição das imagens resultantes fornece uma imagem com as distâncias d entre a assinatura dos "pixels" e de uma classe dada. A operação de divisão da imagem pelo parâmetro S tem a função de rejeitar distâncias maiores que S para a classe dada. Ao final, uma transformação que é, ao mesmo tempo, uma anti-dilatação e uma anti-erosão, fornece uma imagem na qual o valor de cada "pixel" corresponde ao valor da função P calculado segundo a equação (5.1).

5.2.3 **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 5-5.

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

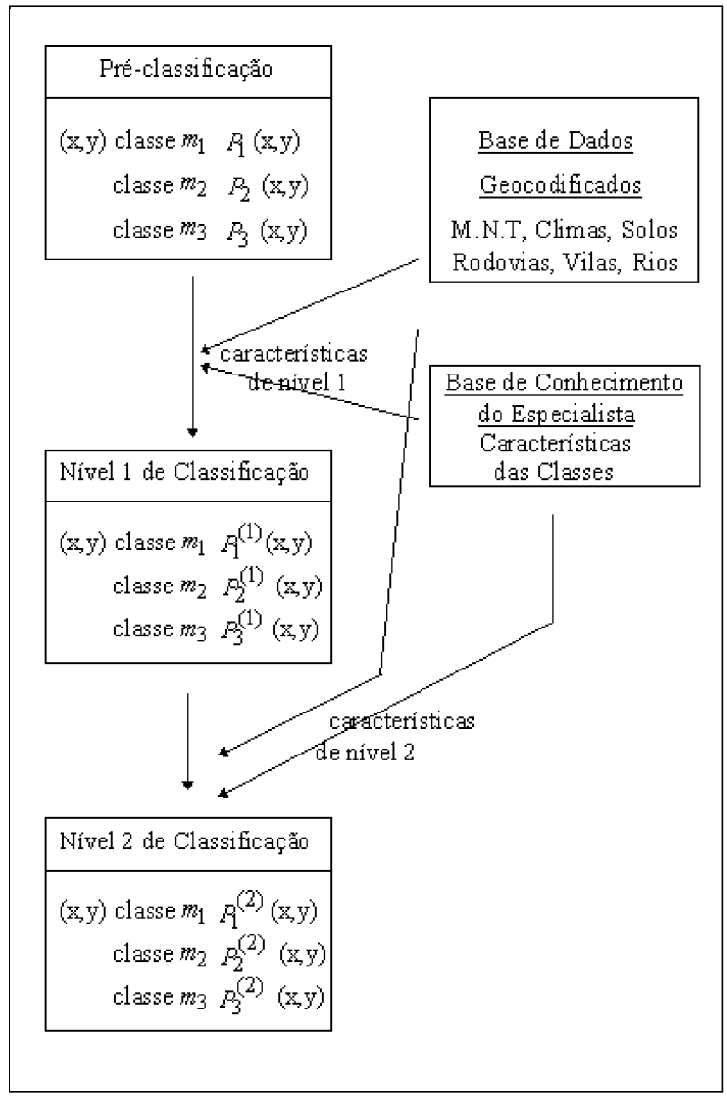

Figure 5-5 As três fases do sistema ICARE.

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

(5.2)

□

- 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:

$$P_i^{(1)}(x, y) := b_i(x, y) * P_i(x, y), \tag{5.3}$$

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:

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

□

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.

Com isso, chegamos ao final dessa fase de classificação, obtendo como resultado para cada "pixel", os três valores revisados:

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

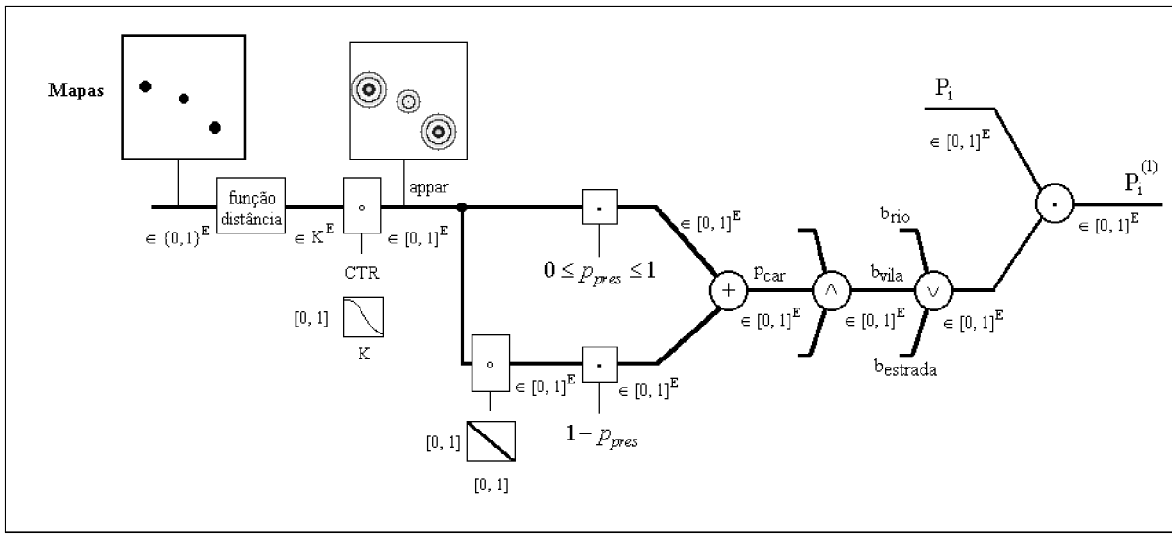

Figure 5-6 Diagrama esquemático do primeiro nível de classificação.

O diagrama esquemático da figura 5-6 mostra os passos do primeiro nível de classificação. Notamos a presença da *função distância* operando sobre objetos (por exemplo, as cidades ou vilas) na imagem obtida a partir de mapas. Depois, uma transformação por uma ELUT (que tem a conotação semântica de "perto", nessa figura) que é uma anti-erosão e antidilatação, ao mesmo tempo, resulta em uma imagem, onde cada "pixel" possui um valor *appar*. No cálculo de *pcar*, uma ELUT do mesmo tipo age sobre a imagem. O mínimo entre as imagens *pcar* para cada sentença ($\text{CTP}_1.\text{CTR}_1.\text{CTO}_1 \cap \text{CTP}_2.\text{CTR}_2.\text{CTO}_2 \cap ...$) resulta no valor de $b_j^{(1)}(x,y)$ para a classe dada, onde j são os atributos observados pelo sistema especialista, como vila, rio, estrada, etc e, o máximo entre todos os $b_j^{(1)}(x,y)$ resulta na imagem $b_i(x,y)$. Ao final, uma multiplicação entre as imagens P_i (da pré-classificação) e a imagem $b_i(x,y)$, resulta em uma imagem onde o valor de cada "pixel" corresponde ao valor da função $P_i^{(1)}$ calculado segundo a equação (5.3), encerrando o primeiro nível de classificação.

5.2.4 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:

$$P_i^{(2)}(x,y) := k_i(x,y) * P_i^{(1)}(x,y), \tag{5.4}$$

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:

$$p_{pres} := \begin{cases} 1, & \text{certeza a } 100\% \\ 0, & \text{falso a } 100\% \end{cases} \tag{5.5}$$

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

Suponhamos 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].
□

A imagem $appar(x,y)$ é calculada em função da imagem de superfícies S da região onde se encontra o ponto, em relação a dois limiares pré-definidos S_1 e S_2 (figura 5-7). 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.

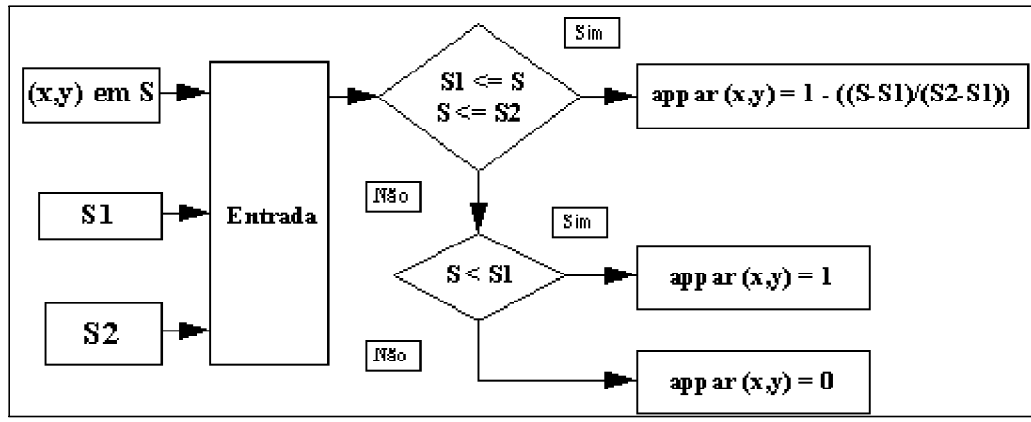

Figure 5-7 Diagrama de cálculo de appar(x,y).

Com isso, chegamos ao final dessa segunda fase de classificação, obtendo como resultado para cada "pixel", os três valores revisados:

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

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

$$\begin{aligned} P_i^{(2)}(x,y) \quad &: \quad = \max_k (P_k^{(2)}(x,y)) \\ \text{classe } (x,y) \quad &= \quad m_i \end{aligned} \tag{5.6}$$

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

O diagrama esquemático da figura 5-8 mostra os passos do segundo nível de classificação. Podemos notar que as imagens de terreno delineiam superfícies, para as quais o sistema estabelecerá relações com os pontos a serem classificados. A seguir, para o cálculo da imagem *appar*, observa-se a transformação por uma ELUT que é uma anti-erosão e uma anti-dilatação, ao mesmo tempo, dependente dos parâmetros S_1 e S_2. Neste ponto, cada "pixel" possui um valor *appar*. No cálculo de *pcar*, uma ELUT do mesmo tipo age sobre a imagem. O mínimo entre as imagens *pcar* para cada sentença ($\text{CTP}_1.\text{CTS}_1 \cap \text{CTP}_2.\text{CTS}_2 \cap ...$) resulta na imagem $k^{(2)}$ para a classe dada. O máximo entre essas imagens resulta na imagem k. A multiplicação entre as imagens $P_i^{(1)}$ (do primeiro nível de classificação) e a imagem k, resulta em uma imagem onde o valor de cada "pixel" corresponde ao valor da função $P_i^{(2)}$ calculado segundo a equação (5.4). Por fim, uma transformação não identificada na morfologia (a transformação dada por "?") faz a atribuição do pixel em questão a uma das classes i, j e l, segundo a equação (5.6).

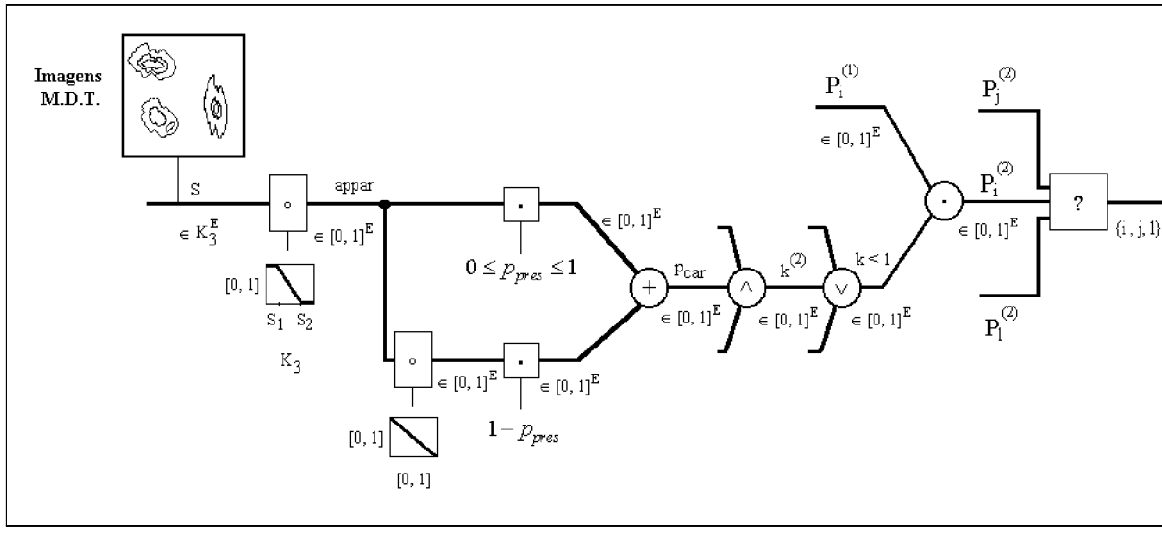

Figure 5-8 Diagrama esquemático do segundo nível de classificação.

6. CONCLUSÕES FINAIS E SUGESTÕES

Esse trabalho possui forte inspiração no trabalho de Desachy (1991). Sua metodologia de trabalho foi incorporada de forma bastante satisfatória e permitiu investigar a viabilidade de se obter um classificador de imagens utilizando transformações da Morfologia Matemática.

Podemos concluir que existe a viabilidade de se implantar um sistema de classificação através de operadores morfológicos. A viabilidade de implementar o próprio sistema especialista através de encadeamentos de transformações morfológicas também foi mostrada.

É interessante notar que somente uma das transformações não pôde ser expressa em termos de transformações morfológicas (a transformação que faz a atribuição do pixel em questão a uma das classes i, j e l), devido ao fato de que precisamos antes definir uma estrutura de reticulado para o conjunto de classes.

Podemos sugerir uma pesquisa mais aprofundada para descobrir transformações mais complexas da Morfologia Matemática que possam simplificar a metodologia apresentada (como aberturas, fechamentos, etc) e determinar quais transformações estão presentes na atribuição final da classe para o "pixel" em questão.

7. REFERÊNCIAS

G. J. F. Banon; Bases da Morfologia Matemática para a análise de imagens. Notas de aula do curso CAP-313, 3o. período de 1994.

G. J. F. Banon.; Characterization of Translation Invariant Elementary Operators for Gray-Level Morphology. 1995 (no prelo).

J. Barrera, G. J. F. Banon e R. A. Lotufo, "A Mathematical Morphology Toolbox for the KHOROS System: Specifications for Version 1.2b. Workshop'95 de Morfologia Matemática, Campinas, 1995.

J. Desachy; Interpretation Automatique d 'Images Satellite: Le Systeme ICARE. These de Doctorat. Institut de Recherche en Informatique de Toulouse. Université Paul Sabatier. France, 1991, 275 pp.

R. M. Haralick, K. Shanmugam and Its'Hak Dinstein; "Textural Features for Image Classification. IEEE Trans. on Syst., Man and Cyber. 3, 610-621, (1973).

H. J. A. M. Heijmans; Theoretical Aspects of Gray-Level Morphology. IEEE Trans. on Patt. Anal. and Mach. Intel. 13, 568-582 (1991).

A. Kandel, Fuzzy Mathematical Techinics with Applications. Addison-Wesley Publ. Co., USA, 1986.

J. Serra; Image Analysis and Mathematical Morphology. Academic Press, Londres, 1982.

L. Vincent; Algorithmes morphologiques a base de files d'attente et de lacets. Extension aux graphes. Tese de doutorado em Morfologia Matemática. L'Ecole Nationale Supérieure des Mines de Paris, France, 1990.

L. A. Zadeh, Fuzzy Sets, Information and Control, v8, 1965, pg 338-353.

8. APÊNDICE

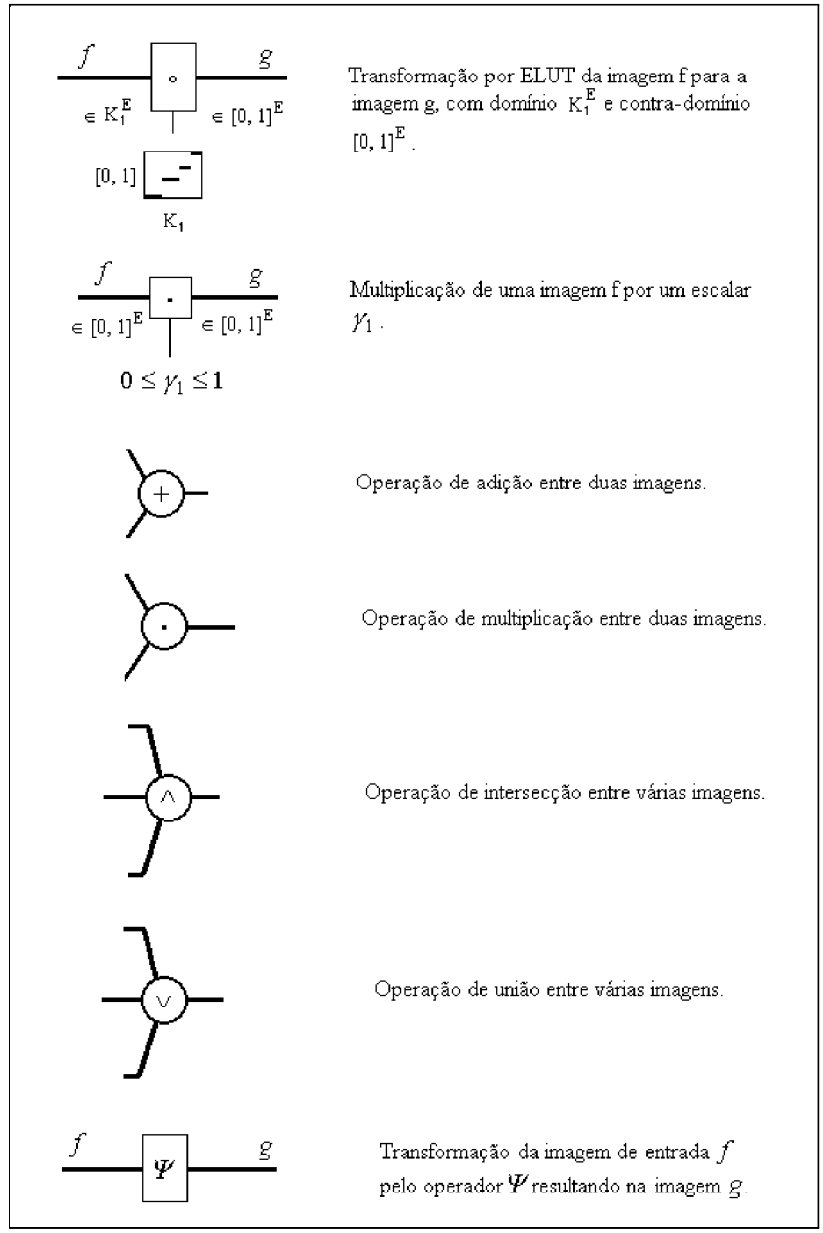

Figure 8-1 Notação gráfica adotada nesse trabalho.