2. Aspectos gerais

"Uma longa caminhada sempre começa
pelo primeiro passo"
Provérbio chinês.

Esse capítulo aborda resumidamente os tópicos teóricos que serão úteis na seqüência deste trabalho: conjuntos nebulosos, sistemas especialistas e morfologia matemática. Em seguida ele traz algumas considerações sobre o sistema Khoros e classificação de imagens.

2.1. Teoria dos Conjuntos Nebulosos ("Fuzzy Sets")

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, programação linear, entre muitos outros (Dubois e Prade, 1980, Kandel, 1986).

Esta seção visa introduzir algumas definições básicas sobre os conjuntos nebulosos que serão úteis nas discussões futuras.

2.1.1. O que são "Fuzzy Sets".

Na Teoria dos Conjuntos Clássica, cada conjunto A de um universo X pode ser definido através de uma função característica $\mu_A : X \longrightarrow \{0, 1\}$ que é tal que:

$$\mu_A(x) = \begin{cases} 1, & \text{se e somente se } x \in A \\ 0, & \text{caso contrário} \end{cases}$$

Um conjunto nebuloso pode ser visto como uma representação de um conjunto na Teoria dos Conjuntos Clássica, do qual só temos um conhecimento imperfeito. Nesse caso, a pertinência de um elemento a um conjunto é dada não mais por

um valor em $\{0,1\}$, mas por um valor qualquer no intervalo $[0,1]$, isto é, e a pertinência de um elemento a um conjunto pode ser parcial.

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

Definição 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. Quando $\exists x \in X$, tal que $\mu_A(x) = 1$, então A é dito ser normalizado.

Quando $\mu_A(x) = 1$ diz-se que x é completamente compatível com o conceito expresso por A em X e quando $\mu_A(x) = 0$ diz-se que x é completamente incompatível com o conceito expresso por A em X.

Em certos casos, pode ser necessário ou conveniente exibir elementos de X que pertençam a um conjunto nebuloso A, com valor de pertinência de pelo menos $\alpha \in (0,1]$. Nesses casos, utiliza-se o *corte* $-\ \alpha$ ou *corte de nível* (Zadeh, 1965).

Definição 2. O conjunto ordinário dos elementos x pertencentes a um conjunto nebuloso A com valores de pertinência de pelo menos $\alpha \in (0,1]$ é definido pelo conjunto de nível α, A_α de A, tal que:

$$A_\alpha = \{x \in X, \mid \mu_A(x) \geq \alpha\}.$$

A função de pertinência de um conjunto A pode ser expresso em termos de seus cortes de nível α de acordo com a expressão:

$$\mu_A(x) = \sup_{\alpha \in (0,1]} \min\{\alpha, \mu_{A_\alpha}(x)\},$$

onde:

$$\mu_{A_\alpha}(x) = \begin{cases} 1, & \text{se } x \ \in A_\alpha, \\ 0, & \text{caso contrário} \end{cases}$$

Exemplo:

Dado um universo X, um conjunto nebuloso A pode ser denotado através de $A = \sum \mu_A(x)/x$, mostrando para cada x a sua pertinência correspondente, para $\mu_A(x)$ não nulo. Sejam, o universo X dado por $X = \{1,2,3,4,5,6,7,8\}$ e um conjunto nebuloso A definido como:

$$A = 0.1/2 + 0.2/3 + 0.5/4 + 0.6/5 + 0.8/6 + 1.0/7;$$

os conjuntos de nível α de A são dados por:

$$
\begin{aligned}
A_{0.1} &= \{2,3,4,5,6,7\} \\
A_{0.2} &= \{2,3,4,5,6,7\} \\
A_{0.5} &= \{4,5,6,7\} \\
A_{0.6} &= \{5,6,7\} \\
A_{0.8} &= \{6,7\} \\
A_{1.0} &= \{7\}.
\end{aligned}
$$

2.1.2. Operações com Conjuntos Nebulosos.

Existem operações sobre os conjuntos nebulosos, similares aos da Teoria de Conjuntos clássica (Kandel, 1986).

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

Definição 4. 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 5. O complemento de um conjunto nebuloso A em X, denotado por $\neg A$ é definido por:

$$\mu_{\neg A}(x) = n(\mu_A(x)), \forall x \in X. \tag{2.1}$$

onde $n : [0,1] \rightarrow [0,1]$ é um operador de negação que satisfaz as seguintes propriedades:

5.1) $n(0) = 1$ e $n(1) = 0$
5.2) $n(a) \leq n(b)$ se $a > b$
5.3) $n(n(a)) = a$, $\forall x \in [0,1]$

e uma negação é estrita se ela é contínua e satisfaz a condição:

5.4) $n(a) < n(b)$ se $a > b$.

A principal operação de negação que satisfaz todas as quatro condições acima é $n(a) = 1 - a$.

Definição 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:

$$C = A \cap B \Leftrightarrow \mu_C(x) = \top\{\mu_A(x), \mu_B(x)\}, \forall x \in X, \tag{2.2}$$

onde o operador $\top : [0,1]^2 \rightarrow [0,1]$, chamado de T-norma, é associativo, comutativo, monotonicamente crescente e tal que $\forall x \in X, \top(x,1) = \top(1,x) = x$. Algumas T-normas estão listados no quadro 2.1.

Definição 7. 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:

$$C = A \cup B \Leftrightarrow \mu_C(x) = \perp\{\mu_A(x), \mu_B(x)\}, \forall x \in X, \tag{2.3}$$

onde o operador $\perp : [0,1]^2 \rightarrow [0,1]$, chamado de T-conorma, é associativo, comutativo, monotonicamente crescente e tal que $\forall x \in X, \perp(x,0) = \perp(0,x) = x$. Algumas T-conormas estão listados na tabela 2.1.
OBS.: Em particular, quando utiliza-se a t-norma min e a t-conorma max, pode-se verificar facilmente que $A \cap \neg A \neq \emptyset$ e portanto $A \cup \neg A \neq X$ (ver Equações 2.2, 2.3 e o exemplo a seguir).

TABELA 2.1 - T-NORMAS E T-CONORMAS DUAIS.

T-normas	T-conormas
$\min(a,b)$	$\max(a,b)$
$a.b$	$a+b-ab$
$\max(a+b-1,0)$	$\min(a+b,1)$
$\begin{cases} a, & \text{se } b=1 \\ b, & \text{se } a=1 \\ 0, & \text{senão} \end{cases}$	$\begin{cases} a, & \text{se } b=0 \\ b, & \text{se } a=0 \\ 1, & \text{senão} \end{cases}$

O exemplo a seguir ilustra as definições acima:
Exemplo:
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, quando a união dos conjuntos nebulosos A e B é dada por $A \cup B = \max(\mu_A, \mu_B)$; a intersecção, denotada por $A \cap B = \min(\mu_A, \mu_B)$ e a negação é dada por $\neg A = 1 - \mu_A$, temos:

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

□

Definição 8. A implicação entre dois conjuntos nebulosos A e B, com suas funções de pertinência $\mu_A(x)$ e $\mu_B(y)$ resulta um conjunto nebuloso C, cuja função de pertinência é dada por:

$$C = A \Rightarrow B \Leftrightarrow \mu_C(x, y) = \bigtriangledown\{\mu_A(x), \mu_B(y)\}, \forall x \in X, \forall y \in Y \qquad (2.4)$$

onde $\bigtriangledown : [0,1]^2 \rightarrow [0,1]$ é um operador de implicação que obedece às seguintes propriedades, para $\forall a, a', b, b' \in [0,1]$.

8.1) Se $b \leq b'$ então $I(a,b) \leq I(a,b')$

8.2) $I(0,b) = 1$

8.3) $I(1,b) = b$

As implicações "puras" obedecem também às propriedades:

8.4) Se $a \leq a'$ então $I(a,b) \geq I(a',b)$

8.5) $I(a, I(b,c)) = I(b, I(a,c))$

Existem vários operadores de implicação definidos na literatura e alguns dos principais estão descritos na tabela 2.2:

TABELA 2.2 - OPERADORES DE IMPLICAÇÃO

Operador de Implicação	Nome
$\max(1-a, b)$	Kleene-Diemes
$\min(1-a+b, 1)$	Lukasiewicz
$1-a+ab$	Reichenback
$\min(a, b)$	Mandani

As T-normas, embora não pertençam à classe das implicações "puras", são os operadores mais utilizadas na prática para modelar a implicação.

2.2. Sistemas Especialistas

A tarefa de um sistema de classificação consiste em examinar um conjunto de informações e ao final decidir entre um conjunto de classes conhecidas, qual é aquela que o conjunto de informações pertence. Os sistemas de classificação que utilizam conhecimento de especialistas para auxiliar na tarefa de classificação tem usualmente, a arquitetura de sistemas especialistas.

Os sistemas especialistas tratam problemas que normalmente são solucionados por "especialistas" humanos (Rich e Knight, 1993, pg631). Estes sistemas trabalham sobre o conhecimento de um dado domínio (por exemplo, conhecimento sobre as classes presentes em uma imagem), de forma a poder tratar de forma semelhante a um ser humano um conjunto de informações. Esse conhecimento deve ser *representado* ou codificado de forma a poder ser tratado por um computador. O conhecimento é geralmente caracterizado por um conjunto de regras e fatos. As regras são a tradução do conhecimento do especialista e os fatos são dados que podem ser adquiridos inicialmente ou posteriormente durante o processo de dedução (Buchanan e Shortliffe, 1985). Usando regras e fatos, novos fatos são obtidos. Para obter novos fatos, utilizam-se de uma ou mais formas de raciocínio. Em geral, os sistemas especialistas podem mostrar ao usuário os caminhos pelos quais trilharam em um processo decisório e adquirir novos e/ou modificar conhecimentos. É claro que o sistema também deve checar se um novo conhecimento não entrará em conflito com o conhecimento já armazenado, ou seja a consistência da base de conhecimento deve sempre ser mantida.

Como o conhecimento deve ser manipulado logicamente, cabe uma definição formal do que é uma *linguagem lógica de primeira ordem* (Bittencourt, 1996):

Definição 9. Uma *linguagem lógica de primeira ordem* - $L(\mathbf{P},\mathbf{F},\mathbf{C},\mathbf{V})$ - é determinada pela especificação dos seguintes conjuntos:

- um conjunto **P** de Símbolos de Predicado;
- um conjunto **F** de Símbolos de Função;
- um conjunto **C** de Símbolos de Constante;
- um conjunto **V** de Símbolos de Variável.

A cada símbolo de predicado e de função é associado um número de argumentos do predicado e da função. Estes conjuntos formam o alfabeto da linguagem lógica e com esse alfabeto, em associação aos operadores lógicos ($\neg$ é "não"; $\wedge$ é "e"; $\vee$ é "ou"; $\Longrightarrow$é "implica"; $\exists$ é "existe"; $\forall$ é "qualquer"), é possível a escrita de fórmulas, como:

$$((\text{Pai}(a,\, b))\wedge(\text{Pai}(b,\, c))\Longrightarrow(\text{Avô}(a,\, c));$$
$$((\text{Pai}(a,\, b))\wedge(\text{idade}(b)=x)\Longrightarrow(\text{Idade_aproximada}(a, x+20);$$

onde Pai, Avô e Idade_aproximada são predicados; idade é uma função que calcula a idade; a, b, c e x são variáveis e $+$ é a operação de adição.

Ainda que utilizada de uma forma "ad-hoc", a lógica de primeira ordem é a base dos sistemas especialistas.

Em um sistema especialista, a forma mais comum de se representar o conhecimento é através de regras, ou seja, as regras são representações de fórmulas lógicas. Em geral, essas regras são do tipo:

SE <*condição*> ENTÃO <*conclusão*>.

que podem ser vistos como uma implicação <*condição*> $\Longrightarrow$ <*conclusão*>.

Uma regra pode ser composta de conjunções e disjunções e pode levar a mais de uma conclusão. Por exemplo:

SE (<*condição 1*> OU <*condição 2*>) E (<*condição 3*>)
ENTÃO (<*conclusão 1*> E <*conclusão 2*>).

que é a representação da implicação (<*condição 1*> $\vee$ <*condição 2*>) $\wedge$ (<*condição 3*>) $\Longrightarrow$ (<*conclusão 1*> $\wedge$ <*conclusão 2*>).

A regra acima com mais de uma conclusão é equivalente a um conjunto de regras com somente uma conclusão, ou seja:

SE (<*condição 1*> OU <*condição 2*>) E (<*condição 3*>)
ENTÃO <*conclusão 1*>
SE (<*condição 1*> OU <*condição 2*>) E (<*condição 3*>)
ENTÃO <*conclusão 2*>

Uma regra simples em processamento de imagens, por exemplo, seria:

SE nível de cinza do "pixel" está no intervalo $[0, 13]$ na banda 4
ENTÃO classe é Rio.

As condições e conclusões de uma regra são usualmente representados por uma lista da forma:

(<objeto>, <atributo>, <valor>)

onde os atributos, através dos seus valores, descrevem as características do objeto em estudo (Bittencourt, 1996; Buchanan e Shortliffe, 1985). No caso desse trabalho, os **objetos** são os tipos sobre os quais o conhecimento atua; os **atributos** são as características presentes na imagem que podem se tornar referências para a utilização do conhecimento, os **valores** são os valores numéricos ou sintáticos que os atributos podem assumir. Um objeto pode ser, por exemplo um segmento de uma imagem, um de seus atributos pode ser a média dos níveis de cinza dos

"pixels" que o compõem com valores entre [0,255]. Um outro exemplo é dado pela lista ("pixel"-i, classe, "floresta"), onde classe é o atributo do objeto "pixel"-i que no caso tem como valor "floresta". Do ponto de vista lógico, usualmente a tripla (<objeto>, <atributo>, <valor>) pode ser codificado utilizando o predicado da igualdade, como (atributo(objeto)=valor), onde atributo é uma função, objeto é uma variável ou uma constante e valor é uma constante. Para que novos fatos sejam deduzidos, é preciso utilizar um mecanismo de inferência que combine fatos e regras.

Em geral há dois tipos de inferência em sistemas especialistas: para-frente e para-trás. No primeiro caso, a partir da verificação das premissas se concluem fatos que podem validar premissas de outras regras. No segundo caso, parte-se das conclusões possíveis e procura-se quais seriam as premissas que poderiam levar a elas e no final dessa seqüência para-trás, de posse de todo o "caminho" percorrido, ele é validado.

Podem haver casos nos quais os atributos não podem ser medidos com a exatidão necessária, ou o próprio especialista verifica que o conhecimento não é válido em qualquer situação. Então podem ser introduzidos os coeficientes de certeza para codificar essa incerteza na conclusão da regra. Nesses casos, a informação incerta pode ser tratada de através de probabilidades (Rich e Knight, 1993; James, 1985), crença e plausibilidade (Senne, 1988); possibilidades (Kruse et al., 1991; Dubois e Prade, 1996) ou "ad-hoc" (Desachy, 1991).

Sistemas especialistas onde regras e fatos envolvem o uso de conjuntos nebulosos são tratados pela chamada "lógica nebulosa", com aplicação notadamente em controle automático de processos (Driankov et al., 1993). Nesta lógica, para combinar regras e fatos lança-se mão de uma forma de raciocínio denominado raciocínio aproximado. Sejam A_i e $A'_i, i = 1, ..., n$, subconjuntos nebulosos em X_i, B um subconjunto nebuloso em Y com a funções de pertinência conhecidas, R uma regra e F um fato, onde:

$$R : \text{SE } \langle x_1 = A_1 \rangle \text{ E } (\langle x_2 = A_2 \rangle \text{ OU } \langle x_3 = A_3 \rangle) \text{ ENTÃO } \langle y = B \rangle .$$
$$F : \langle x_1 = A'_1 \rangle \text{ E } \langle x_2 = A'_2 \rangle \text{ E } \langle x_3 = A'_3 \rangle .$$

O processo geral de inferência segue os seguintes passos (Sandri, 1997):

a) Cria-se uma relação nebulosa A em $X_1 \times ... \times X_n$ entre os A_i e uma relação nebulosa A' em $X_1 \times ... \times X_n$ entre os A'_i.

b) Cria-se a relação nebulosa R entre a premissa dada pela relação nebulosa A e pela conclusão dada pelo conjunto nebuloso B,usando uma função de implicação.

c) Faz-se uma composição de R com A' e projeta-se o resultado sobre Y obtendo-se o subconjunto nebuloso B'.

No caso onde os A'_i representam fatos precisos, isto é, cada A'_i é tal que $\exists x^*_i \in X_i, \mu_{A'_i}(x^*_i) = 1$, o processo é simplificado e será descrito detalhadamente a seguir,

dada a sua relevância ao trabalho. Supondo que a base de conhecimento seja composta por n regras R_j do tipo:

R_j :Se $\langle x_{ij} = A_{ij} \rangle$ E ... E $\langle x_{nj} = A_{nj} \rangle$
Então $\langle y = B \rangle$;

o processo se divide em quatro fases:

1) Seja o valor de realização de cada x_i dado por $x_i^* \in X_i$, isto é x_i^* representa o fato A_i'. Determina-se o grau de compatibilidade de $x_i^*, 1 \leq i \leq n$, para cada A_{ij} da regra R_j, $1 \leq j \leq m$:

$$h_{ij} = \mu_{A_{ij}}(x_i^*);$$

2) Determina-se o grau de compatibilidade global da premissa de cada regra $R_j, 1 \leq j \leq m$, com uma T-norma $\top$:

$$h_j = \top(h_{1j}, ..., h_{nj});$$

Quando a regra envolve conectivos OU, este passo envolverá também o uso de uma T-conorma $\perp$.

3) Determina-se o conjunto nebuloso B_j', referente a cada regra R_j usando um operador de implicação $\bigtriangledown$:

$$\mu_{B_j'}(y) = \bigtriangledown \left(h_j, \mu_{B_j}(y) \right), \forall y \in Y;$$

4) Determina-se o conjunto nebuloso B', usando um operador $\diamondsuit$ para agregar os B_j' :

$$\mu_{B'}(y) = \diamondsuit \left(\mu_{B_1'}(y), ..., \mu_{B_m'}(y) \right); \forall y \in Y.$$

Exemplo:

Sejam, A, B e C conjuntos nebulosos e R_1 :Se $x = A$ então B e R_2 :Se $x = C$ então B, duas regras e três fatos precisos $x^* = x_i, i = 1, 2, 3$ e $x_i \in X$. Na Figura 2.1 pode-se observar o progresso do raciocínio aproximado simplificado para a regra R_1 quando a implicação adotada é $\bigtriangledown = \sim A \vee B$ e quando a implicação adotada é $\bigtriangledown = \min$. Nota-se que quando x_i é completamente compatível com o conjunto nebuloso em questão, os resultados obtidos são iguais para as duas formas de implicação. No entanto, quando x_i é completamente incompatível com o conjunto nebuloso em questão, a modelagem da ignorância sobre o ponto são diferentes. No caso da implicação $\bigtriangledown = \sim A \vee B$ todas as respostas possíveis recebem pertinência igual a 1 e no caso da implicação $\bigtriangledown = \min$, todas as respostas possíveis recebem pertinência igual a zero.

A Figura 2.2 ilustra o uso do operador $\diamondsuit = \max$ para a função de implicação $\bigtriangledown = \min$ utilizando-se as regras R_1 e R_2.

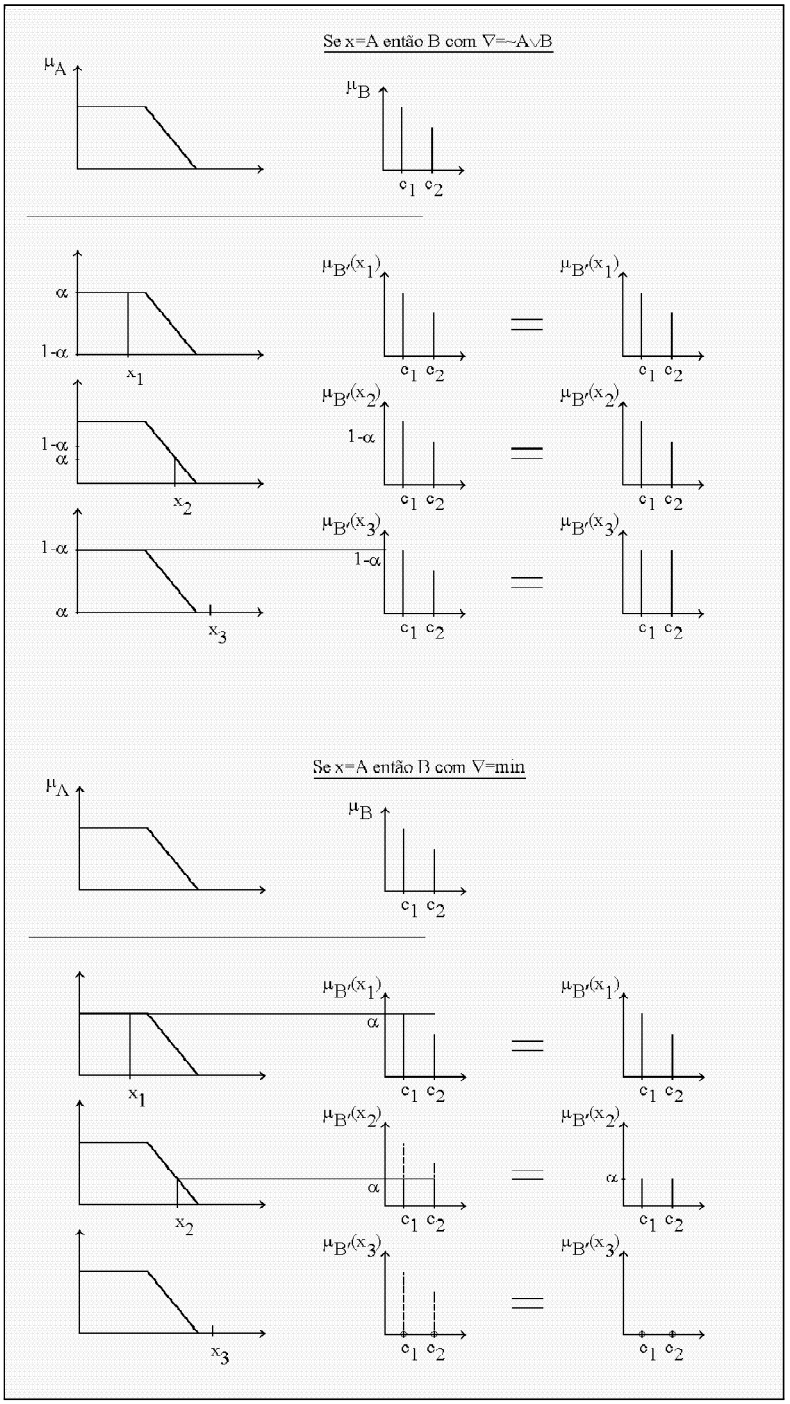

Fig. 2.1: Raciocínio nebuloso simplificado com uma regra e com implicação $\bigtriangledown = \sim A \vee B$ e $\bigtriangledown = \min$, e com exemplo de 3 fatos precisos $x^* = x_i$.

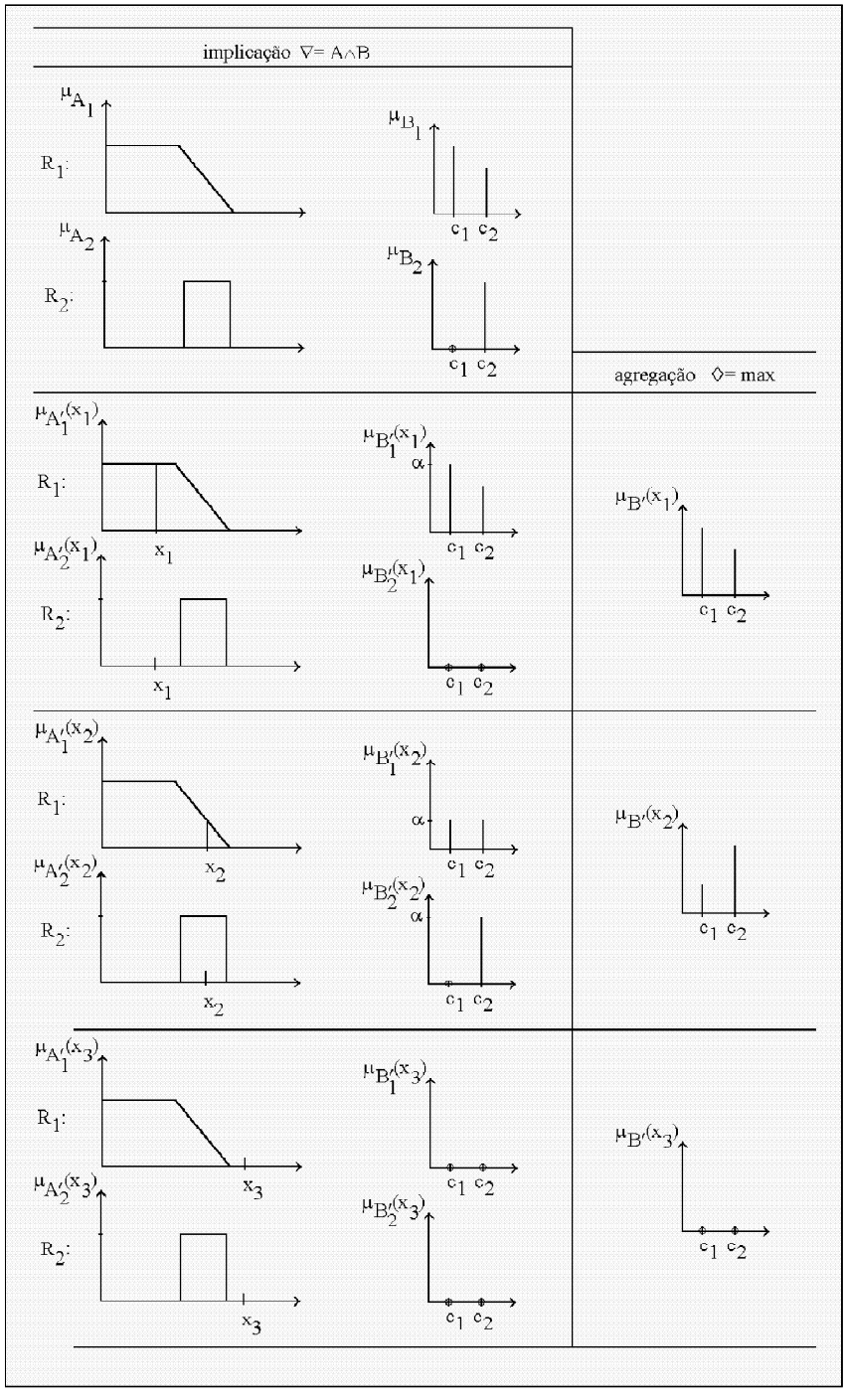

Fig. 2.2: Agregação com $\Diamond = \max$ para a função de implicação $\bigtriangledown = \min$.

□

O operador que implementa $\diamondsuit$ é usualmente escolhido em função da implicação $\triangledown$ usada. Quando $\triangledown(a,b) = \max(1-a,b)$, denotado por $\triangledown =\sim A \vee B$, escolhe-se para $\diamondsuit$ uma t-norma qualquer. Quando $\triangledown(a,b) = \top(a,b)$, denotado por $\triangledown = \top$, escolhe-se para $\diamondsuit$ uma t-conorma qualquer.

2.3. Conceitos da Morfologia Matemática

Para maior facilidade no entendimento do trabalho, apresenta-se a seguir alguns conceitos da Morfologia Matemática que serão utilizados nesse trabalho.

Definição 10. (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 11. (Birkhoff, 1967) Um reticulado L_1 é um *reticulado completo* quando cada um dos seus subconjuntos tem um supremo e um ínfimo em L_1.

Definição 12. 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 13. Uma cadeia que possua um maior elemento e um menor elemento é uma *cadeia limitada*.

Exemplos de cadeias limitadas, com a relação de ordem $\leq$ do conjunto dos números inteiros:

$$K_1 = \{0, 1\}$$
$$K_2 = \{0, 1, 2, 3\}$$
$$K_3 = \{0, 1, ..., 255\}$$

□

Proposição 1. Toda cadeia limitada é um reticulado.

A utilidade da cadeia limitada para nosso estudo reside no fato de que o contra-domínio de uma imagem 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 2.3.

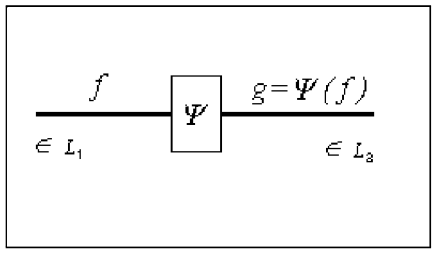

Fig. 2.3: Transformação sobre imagens.

Definição 14. Uma transformação $\Psi : L_1 \rightarrow L_2$ 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 15. Uma transformação $\Psi : L_1 \rightarrow L_2$ entre imagens é *decrescente* se:

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

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

Definição 16. 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 de 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 17. 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 $\Psi(\max K_1) = \min K_2$

Na Figura 2.4, podemos visualizar exemplos de operadores elementares respectivamente de dilatação, erosão, anti-dilatação e anti-erosão, em acordo com as definições 16 e 17, usando $K_1 = K_2$.

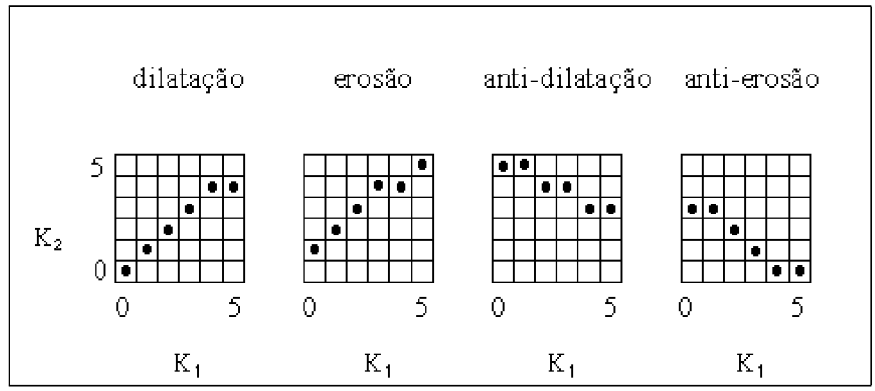

Fig. 2.4: Exemplos de operadores morfológicos elementares.

É importante salientar que um operador Ψ pode vir a não se enquadrar em apenas um tipo. A título de ilustração, mostramos na Figura 2.5 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 17, 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 e erosão podem ser dadas de formas equivalentes mais simples. Nesse caso, cabem algumas definições iniciais.

Definição 18. 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 19. Dado um conjunto X, sua translação por um vetor t, seu complemento e sua transposição (simetria com relação a origem), são denotados por: X_t, X^c e X^t, respectivamente, e dados por:

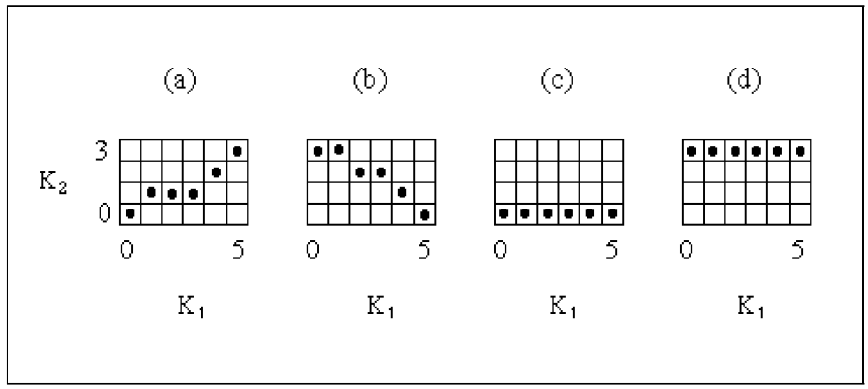

Fig. 2.5: Exemplos de alguns operadores morfológicos de dupla função.

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

Nesse momento, é possível construir a dilatação e a erosão invariante por translação da Morfologia Matemática binária:

Proposição 2. Seja $B \in P(E)$. O operador δ_B de $P(E)$ em $P(E)$ dado por:

$$\delta_B(X) = X \oplus B = \cup_{b \in B} X_b,$$

onde $X \oplus B$ é a soma de Minkowski entre os dois conjuntos X e B (Banon e Barrera, 1994), é uma dilatação de $(P(E), \subset)$ em $(P(E), \subset)$. $\delta_B(X)$ chama-se dilatação de X relativamente ao borrão B.

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

Proposição 3. Seja $B \in P(E)$. O operador ε_B de $P(E)$ em $P(E)$ dado por:

$$\varepsilon_B(X) = X \ominus B = \{y \in E : B + y \subset X\},$$

onde $X \ominus B$ é a diferença de Minkowski entre os dois conjuntos X e B (Banon e Barrera, 1994), é uma erosão de $(P(E), \subset)$ em $(P(E), \subset)$. $\varepsilon_B(X)$ chama-se erosão de X pelo elemento estruturante B.

Na prática, a grande maioria dos borrões e dos elementos estruturantes são simétricos, isto é, $B \equiv B^t$, como no caso dos quadrados, discos, cruzes, etc (quando centrados na origem).

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

A seguir serão dadas definições particulares de alguns operadores da Morfologia Matemática que são utilizados no decorrer deste trabalho.

2.3.1. 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 $\Psi_d : K_1^E \rightarrow K_2^E$, 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 maior valor inteiro menor ou igual 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.

Na Figura 2.6 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 entrada 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 dentro do objeto (o valor zero é associado a cor azul).

Na Figura 2.7 apresenta-se uma construção da função distância em termos de erosões de imagens binárias. Observa-se que a função distância é uma união de composições de erosões por elementos estruturantes $B \subset E$ e, transformações

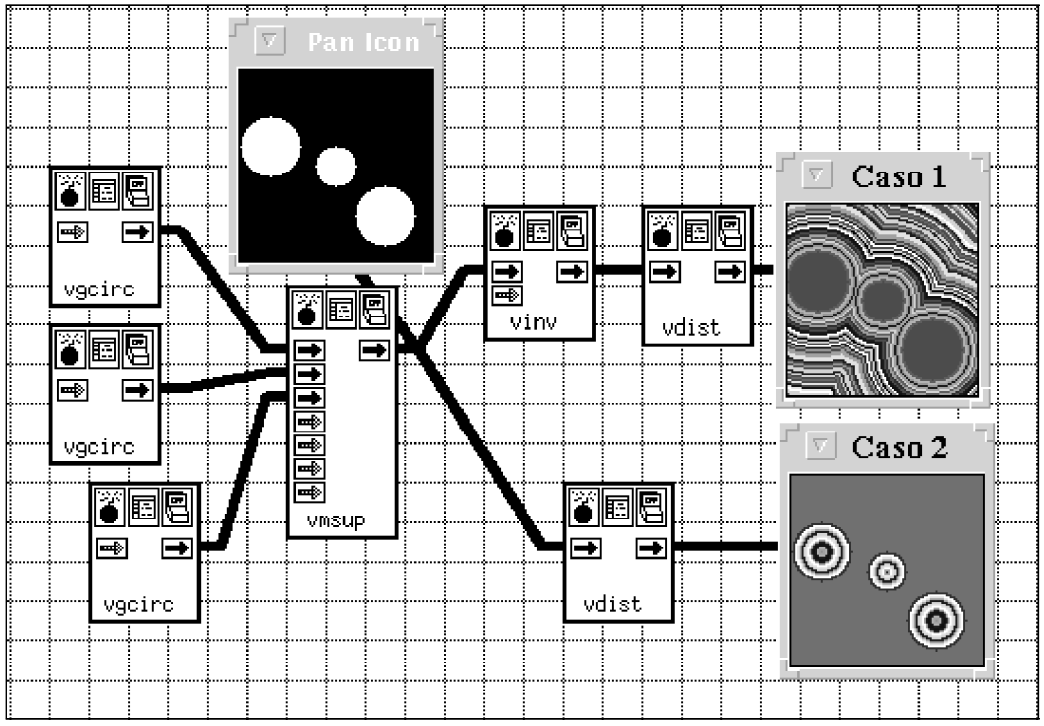

Fig. 2.6: Aplicação da função distância a uma imagem binária com três circulos no Khoros 1.0.5.

por "Elementary Look Up Tables" (ELUTs) (Banon 1995). Nesse caso, os elementos estruturantes B são discos (Vincent, 1990). Na Figura 2.7, o elemento estruturante nB é o resultado de uma sucessão de $n-1$ composições da adição de Minkowski aplicado a 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 $\{0,k\}$ em $K = \{0, 1, ..., n+m\}$, onde k é um valor inteiro e a cadeia K depende do número de linhas n e do número de colunas da imagem em questão.

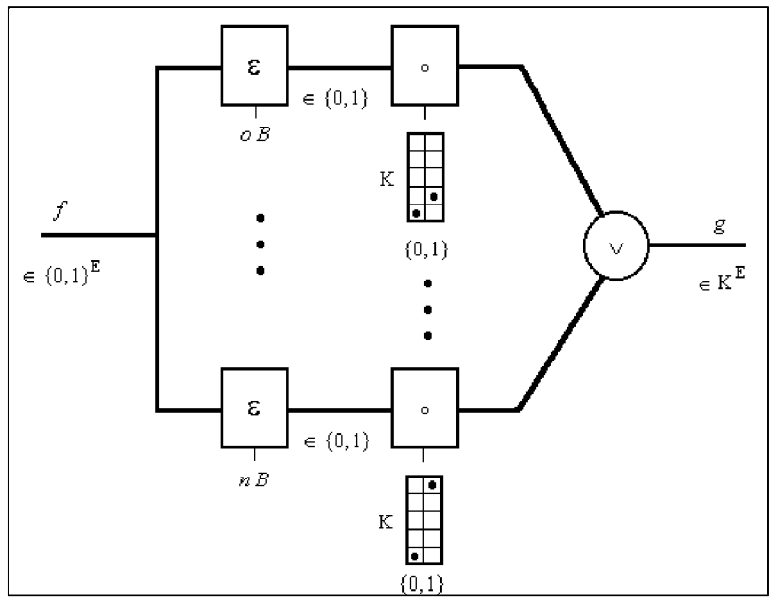

Fig. 2.7: A função distância em termos de operadores morfológicos.

Além das aplicações mencionadas da função distância (dilatação, erosão, esqueleto, etc), uma outra em particular pode 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.

2.3.2. A Limiarização

A limiarização de uma imagem $f : E \longrightarrow \{0, ..., 255\}$ relativo ao intervalo $[a, b]$ é o operador dado por:

$$\text{limiarização}_{[a,b]}(f)(x) = \begin{cases} 255, & \text{se } f(x) \in [a,b] \\ 0, & \text{caso contrário} \end{cases}$$

2.3.3. A Abertura por Reconstrução

Seja B um elemento da coleção $\mathcal{P}(E)$ e seja Y um subconjunto de E. O operador $\gamma_{B,Y}$ sobre $\mathcal{P}(E)$ é denominado abertura por reconstrução e é dado por (Banon e Barrera, 1994):

$$\gamma_{B,Y} = \bigcup_{n=1,...} \delta^n_{B,X}(Y), \text{ para qualquer } X \text{ em } \mathcal{P}(E).$$

2.3.4. O Sup-gerador

Sejam A e B dois elementos da coleção $\mathcal{P}(E)$, tais que $A \subset B$. O operador $\lambda_{A,B}$ sobre $\mathcal{P}(E)$ é denominado sup-gerador e é dado por (Banon e Barrera, 1994):

$$\lambda_{A,B} = \varepsilon_A \wedge \delta^a_{B^t}, \text{ para qualquer } X \text{ em } \mathcal{P}(E).$$

2.4. O Sistema Khoros

O sistema Khoros é um ambiente de aplicação e desenvolvimento de software para processamento, análise, exploração e visualização de dados multidimensionais (Khoral, 1996). Em particular, no caso específico do nosso trabalho, o Khoros proporciona uma boa variedade de programas para processamento de imagens podendo inclusive receber aplicações bastante diversas, como por exemplo controle de processos, análise de imagens médicas e análise numérica. Sua modularidade permite que ele seja utilizado como um conjunto de bibliotecas programadas visualmente em um ambiente específico denominado *Cantata*. Essas bibliotecas são denominadas "Toolboxes" e o Khoros, em sua versão original, possui as seguintes (Khoral, 1996):

- Bootstrap: contém o Khoros básico e ferramentas de portabilidade.
- Dataserv: contém bibliotecas para manipulação de grandes volumes de dados, além de permitir conversão de tipos de dados.
- Datamanip: contém operadores de manipulação de dados uni-dimensionais até n-dimensionais.
- Design: contém bibliotecas que tornam possível a programação visual e a criação de novas "Toolboxes" pelo usuário.
- Matrix e Image: oferecem operadores para análise numérica e o processamento de imagens.
- Envision e Geometry: contém operadores interativos para visualização e análise de dados 2-D e 3-D.
- Support: é uma coleção de programas utilitários para o ambiente de desenvolvimento.

◦ Migration: contém utilitários para a conversão de trabalhos realizados em versões anteriores do Khoros.

◦ Sampledata: contém exemplos de aplicação.

◦ Retro: É fornecida como cortesia para manter a compatibilidade a versões anteriores do Khoros com a versão atual.

Uma grande vantagem do Khoros sobre outras plataformas de Processamento de Imagens reside no fato dele ser uma plataforma aberta e disponível na Internet sem custos adicionais pelo endereço: http://www.khoral.com/. Por esses motivos, além dessas "Toolboxes", existem outras também disponibilizadas na Internet. Nos deteremos em particular em uma que auxiliou a execução desse trabalho: a "MMach Toolbox".

A "MMach Toolbox" ("Morphological Machine Toolbox") é um desenvolvimento em parceria do INPE, USP e UNICAMP. Implementa basicamente operadores da Morfologia Matemática e atualmente está na versão 1.4. É a ferramenta principal deste trabalho. Também está disponível para "download" pela Internet pelo endereço: ftp://ftp.dca.unicamp.br/pub/Khoros/Khoros2/contrib/toolboxes/mmach.

O trabalho começou a ser desenvolvido utilizando-se a versão 1.3. Com o lançamento da versão 1.4 em fevereiro de 1997, foi executado um trabalho de migração para a nova "release" e o desenvolvimento passou a ser nessa última versão.

2.5. Os Tipos de Classificação de Imagens

Os métodos de classificação dividem-se em dois grandes grupos principais: os não-supervisionados e os supervisionados (Saebo et al., 1985). Os não-supervisionados dividem as classes automaticamente e podem ser baseados no princípio de Fisher (Johnson e Wichern, 1982 e Fisher, 1936) da década de 30. Alguns métodos podem ser baseados também em funções discriminantes matemáticas, como por exemplo a Teoria dos Grafos e relações nebulosas (Dunn, 1974); abordagens estatísticas assintóticas (Cooper e Freeman, 1970), ou ainda modelos de pertinência nebulosa (Kent e Mardia,1988). Segundo Saebo et al. (1985), os supervisionados são mais precisos por levarem em consideração informações fornecidas pelo usuário, a *verdade terrestre*. Podem ser divididos em dois sub-grupos: os *clássicos* e os *contextuais*. Os primeiros usam a informação contida no próprio "pixel " para sua classificação do mesmo e o segundo atribui uma das classes disponíveis de acordo com a análise do "pixel" e dos seus vizinhos. Para mais detalhes de diversos métodos de classificação, pode-se citar (Swain e Davis, 1978) e (Fukunaga, 1990) sobre classificação clássica e (Fu e Yu, 1980) e (Saebo et al., 1985) a respeito de classificação contextual. Com relação a esta última, (Moraes, 1992) faz

comentários a respeito da implementação de uma metodologia contextual e sua comparação direta com uma metodologia de Máxima Verossimilhança no sistema SITIM. Em Baridó (1989), vários métodos - paramétricos e não-paramétricos - são apresentados e comparados a partir de simulações. Neste mesmo trabalho, encontra-se também a versão não-paramétrica da metodologia contextual. A Teoria de Conjuntos Nebulosos também fornece técnicas nebulosas para classificação supervisionada (Bezdek et al., 1986, Béreau e Dubuisson, 1991 e Kissiov e Hajitodorov, 1992). No vasto campo de pesquisa dos classificadores, foram explorados também a utilização de múltiplos classificadores (Ho et al., 1994) para melhoria da acurácia. Recentemente, técnicas baseadas em Redes Neurais tem ganhado força principalmente pela sua característica de aprendizado (Simpson, 1990) e capacidade de adaptação a diversos problemas (Mustavi et al., 1994).

Há ainda o conceito de pós-classificação ou "classificação em fases", onde uma metodologia é empregada para aumentar a acurácia da classificação gerada por outra. Nesse caso específico, podemos citar o trabalho de (Desachy, 1991), onde um pós-classificador inteligente age sobre uma pré-classificação estatística. Existem outros tipos, com características contextuais (vizinhança do "pixel"), como o MAXVER-ICM do Spring (Frery, 1993; INPE, 1997).

2.5.1. A Classificação Visual de Imagens

Esta é uma metodologias de análise para quem não dispõe de nenhum recurso computacional. Trata-se da Classificação ou Interpretação Visual da Imagem (I.V.I.) (Rudorff et al., 1986). A imagem é então adquirida, geralmente em papel, como uma composição colorida ou em níveis de cinza. Este tipo de material é fornecido pelo Instituto Nacional de Pesquisas Espaciais (INPE) podendo ser esta imagem de uma determinada época do ano, mono ou multiespectral, etc; ou seja, há várias opções de modo que o usuário deve procurar a que melhor se adeque à sua necessidade. Tendo a imagem em mãos, o passo seguinte é o de reconhecer o que existe na imagem e extrair dela a informação desejada.

A interpretação visual requer do usuário conhecimentos prévios. Por exemplo, há a necessidade do conhecimento prévio da época adequada do ano da passagem do satélite para o fim desejado, tipo de produto (colorido ou em níveis de cinza, já que nem sempre o colorido é que traz mais informações), as bandas a serem escolhidas (cada banda reflete uma determinada faixa do espectro - ver Apêndice A ou para trabalhos com monitoramento ambiental, por exemplo, pode-se encontrar boas indicações em Tucker (Tucker, 1978)). Dependendo da finalidade do estudo, a I.V.I. pode requerer mais conhecimentos da área "in loco", o que nem sempre é possível, ou associações com fotos aéreas, entre outros.

Os fotointérpretes podem ser classificados em três grupos: 1) os que possuem

conhecimento de campo (possuem dados que refletem a realidade do local em estudo); 2) os que não possuem este conhecimento e, por fim; 3) os que não possuem este conhecimento, mas que possuem fontes de informações adicionais. Os do primeiro grupo tem o trabalho simplificado pelo conhecimento que possuem. Os do terceiro grupo podem ter dificuldades em cruzar as informações de que dispõem com a imagem. Já os do segundo grupo, dificilmente conseguirão fazê-lo sem o uso de *chaves de interpretação* - conhecimentos utilizados que são estruturados como se fossem formulações para se resolver um problema. É de grande valia que o uso das chaves de interpretação seja feita de forma invariante, ou seja, não possa mudar de um fotointérprete para outro. Para isso, criam-se legendas que vão funcionar como critérios de separação das diversas classes. Além de fixar a interpretação, as legendas auxiliam uma leitura mais rápida.

Feita a classificação, parte-se para extrair da imagem as informações que se procura desde o início. Tomando um exemplo agrícola de estimativa de áreas plantadas, pode-se agir de duas formas. Na primeira, usa-se uma malha de pontos proporcionais a escala para se estimar a área, como um método grosseiro de integração. Logicamente que os resultados podem ser induzidos a erros, que podem comprometer todo o trabalho. Uma solução seria trabalhar com uma malha de pontos bastante densa, o que ainda assim proporciona um risco de proporções razoáveis se o terreno em questão tiver um relevo acidentado. Uma segunda forma, já bem mais sofisticada, requer um computador em associação a uma mesa digitalizadora. Os contornos da região são digitalizados com o auxílio da mesa e remetidos a um computador. Os dados podem ser associados a um mapa da região, incorporando informações sobre o relevo. A seguir, um simples processo de integração numérica fornece a estimativa da área com erros bastante aceitáveis. Esta forma, porém é bastante onerosa, mas permite que se leve em consideração a tridimensionalidade. Com a tecnologia atual - a utilização de "scanners" pode transformar o trabalho ainda mais preciso, pois evita o contato manual - a transferência da imagem é feita por um processo digital de leitura. A grande desvantagem destes dois últimos processos (mesa digitalizadora e integração numérica) é a necessidade de "software" específico, nem sempre disponível com facilidade para este tipo de aplicação.

2.5.2. Classificação Assistida por Sistema Especialista

Os classificadores assistidos por sistemas especialistas surgiram como alternativa para baratear o custo computacional dos classificadores numéricos. Basicamente, eles tentam embutir no processo de classificação o conhecimento humano de "como se faz". Como grande vantagem, pode-se salientar que o processo decisório pode ser suficientemente rico e envolver uma quantidade razoavelmente pequena

de informação. Como conseqüência interessante, temos que o conhecimento pode ser armazenado em termos de *regras*, ser explicável e poder ser tratado de forma simplificada (James, 1985). Como desvantagem, temos que ele não necessariamente irá responder corretamente fora dos padrões para os quais foi delineado, em oposição aos métodos exclusivamente numéricos que em geral respondem.

A idéia é bastante simples: um conjunto de regras do tipo "SE *condição* ENTÃO *conclusão*" armazena o conhecimento humano sobre determinado fenômeno, visto pela ótica de um especialista. No entanto, nem sempre um especialista tem plena certeza de que tal regra valha em todas as situações, ou que as variáveis possam ser medidas com total acurácia. Essa incerteza pode ser medida usando-se a probabilidade, as funções de pertinência, possibilidades e necessidades (Granger, 1988), ou ainda as probabilidades nebulosas. O problema de "como" combinar as medidas de incerteza para dar uma resposta final, pode ser resolvido de diferentes formas, de acordo com o caminho escolhido. Em (James, 1985), (Senne, 1988) e (Kruse et al., 1991), podem ser vistas explanações sobre o tratamento desses casos.