

A DISCRIMINAÇÃO DE TEXTURAS USANDO UMA REPRESENTAÇÃO SIMBÓLICA PARA DADOS VISUAIS

Antonio Miguel Vieira Monteiro
Flávio Roberto Dias Velasco

Ministério da Ciência e Tecnologia - MCT
Instituto de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12201 - São José dos Campos - SP

1 - INTRODUÇÃO

A análise de textura tem-se mostrado fundamental para a percepção de regiões e superfícies em imagens. Por exemplo, nós facilmente distinguimos entre um tecido de lã e um de algodão, na mesma coloração. Na análise de imagens de satélite, propriedades de textura são sempre usadas pelos especialistas, direta ou indiretamente, quando classificando as diferenciadas regiões em uma imagem. Este fato levou a muitos estudos e algoritmos propondo reconhecimento e descrição de texturas por computador. No entanto, como poucas definições de textura têm sido dadas, e nenhuma precisamente define o problema, as técnicas de discriminação da textura são grandemente "ad hoc" (Brady, 1982).

Duas são as abordagens principais para a discriminação de textura: a estatística e a estrutural. As propriedades para a caracterização da textura, têm sido derivadas no domínio estatístico, usando principalmente matrizes de coocorrência. (Haralick et al. (1973) e Connors et al. (1984)).

Processos mais recentes sugerem que a extração de características locais de textura, e as relações entre as regiões que ficam determinadas por estas características, levariam a definição de elementos de textura, ou primitivas de textura. Estas primitivas teriam papel central na percepção de regiões e/ou superfícies de texturas diferenciadas.

A idéia deste trabalho é utilizar as primitivas definidas no Esboço Primário (Marr, 1978) como elementos de textura. As texturas, neste enfoque, são caracterizadas por estatísticas e relações entre as primitivas. A organização deste trabalho apresenta na secção 2 uma descrição sumária da proposta de Marr para descrição de imagens, na secção 3 alguns experimentos de discriminação de texturas são apresentados sendo então analisados na secção 4.

2 - O ESBOÇO PRIMÁRIO BRUTO

A representação mais comumente encontrada para imagens é uma matriz de inteiros, onde cada elemento é chamado pixel. Para realizar tarefas visuais por máquina esta é uma representação "pobre", na medida em que mistura as informações sobre a estrutura física da cena imageada.

Quando vemos uma cena, existe uma variação de energia luminosa que é "capturada" pelos elementos da retina. O objetivo de uma representação primária para dados visuais é então "capturar" estas variações de energia. Cada pixel possui um valor numérico que representa uma intensidade da energia capturada. Em uma cena existem vários fatores responsáveis por estes valores de intensidade, mas quatro (4) são os principais:

- (1) A geometria das superfícies visíveis imageadas;
- (2) A reflectância destas superfícies;
- (3) A iluminação sobre a qual está submetida a cena e
- (4) O ponto de visada do observador.

Na representação matricial todos estes fatores estão fundidos em um único valor dado pelo valor do pixel.

Na verdade, as variações de intensidade que ocorrem são devido a estas causas isoladamente ou a combinações destas. O objetivo básico da representação proposta por Marr é, então, definir quais variações são definidas por quais fatores e a partir daí criar uma representação que explicitamente os eventos físicos responsáveis pelas variações de intensidade, que em última análise, provocam a nossa "visão" da cena.

Para fazer isto se deve achar representações e métodos de detecção adequados para as variações e para as estruturas das superfícies existentes na cena imageada. A este estágio da representação é que chamamos ESBOÇO PRIMÁRIO BRUTO (EPB).

Na nossa implementação, a primitiva simbólica básica de representação para as variações de intensidade, em uma determinada escala (canal $\nabla^2 G$), é o SEGMENTO CRUZAMENTO POR ZERO (SCZ). Os SCZ's são caracterizados por seus atributos: POSIÇÃO, TAMANHO, CONTRASTE E ORIENTAÇÃO. A estrutura de dados para o SCZ é vista abaixo na Fig. 1. Para mais detalhes sobre o EPB veja Marr (1982) e Monteiro (1988).

SCZ	POSIÇÃO	TAMANHO	CONTRASTE	ORIENTAÇÃO	PONTOS	
	[x,y]				[x _i ,y _i]	[x _f ,y _f]

Fig 1 - Estrutura do SCZ

3 - O EXPERIMENTO

Nesta implementação, nenhuma primitiva específica para a textura foi estudada. No entanto, a textura se caracteriza por algum tipo de "regularidade" nos elementos primitivos que compõem a cena. A repetição ou quase-repetição dos elementos fundamentais, pode caracterizar uma região e/ou superfície. As características de organização destes elementos também podem fazê-lo.

Pensando textura desta forma, podemos considerar os elementos primitivos do ESBOÇO PRIMÁRIO BRUTO (EPB) como nossos elementos de textura. Usando, então, estatísticas simples (de primeira ordem) sobre o EPB, podemos discriminar texturas, baseando nossa análise nas relações espaciais e nas propriedades de cada primitiva básica do EPB, ou de um determinado canal ∇^2G .

Para testar estas considerações, utilizamos imagens de texturas obtidas com uma câmara de TV, a partir do Brodatz (1968), com tamanhos de (128 X 128) pixels. Fizemos dez (10) amostras de cada textura. Na Fig 2 estão apresentadas uma das amostras de cada uma das imagens trabalhadas.

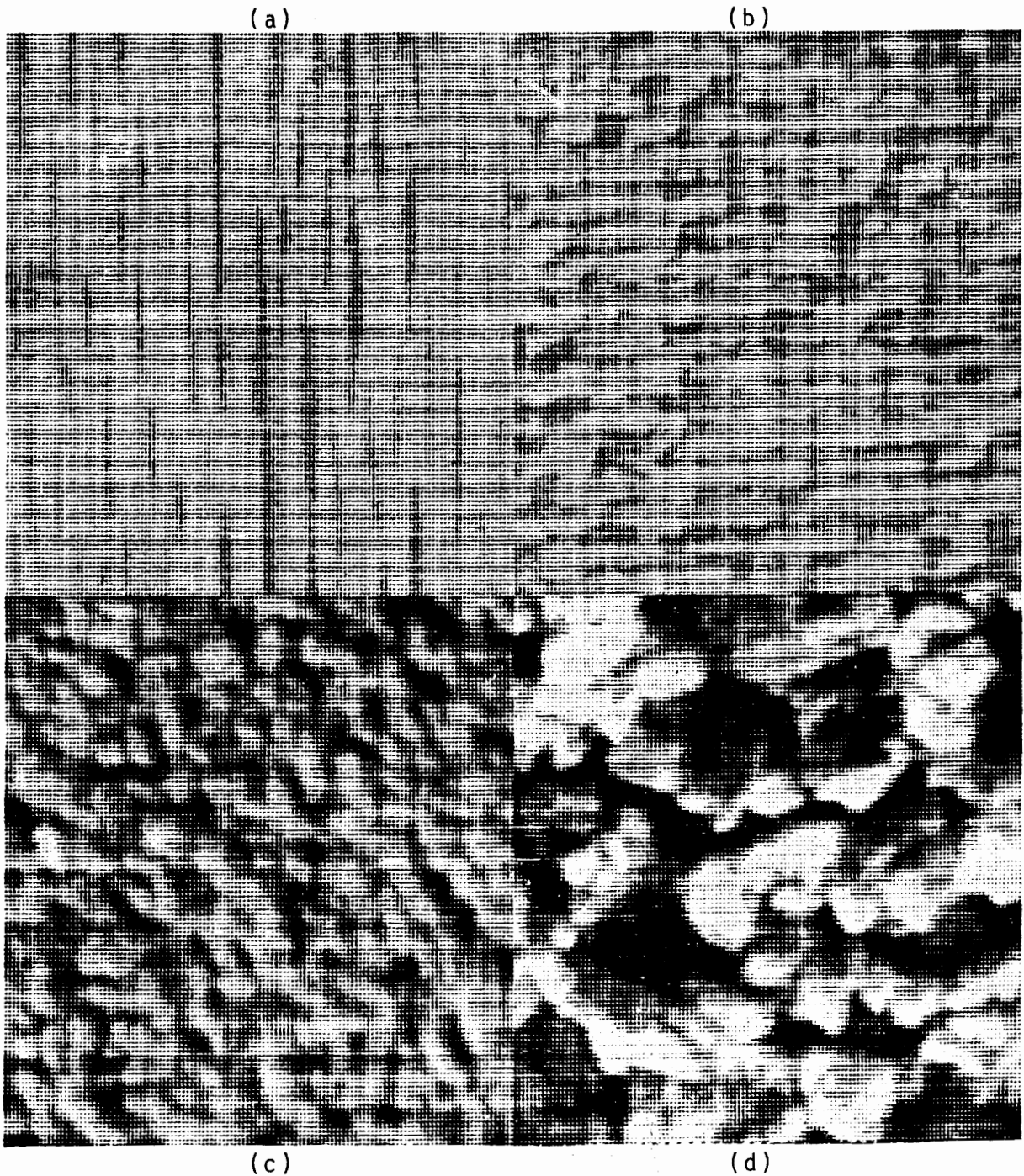


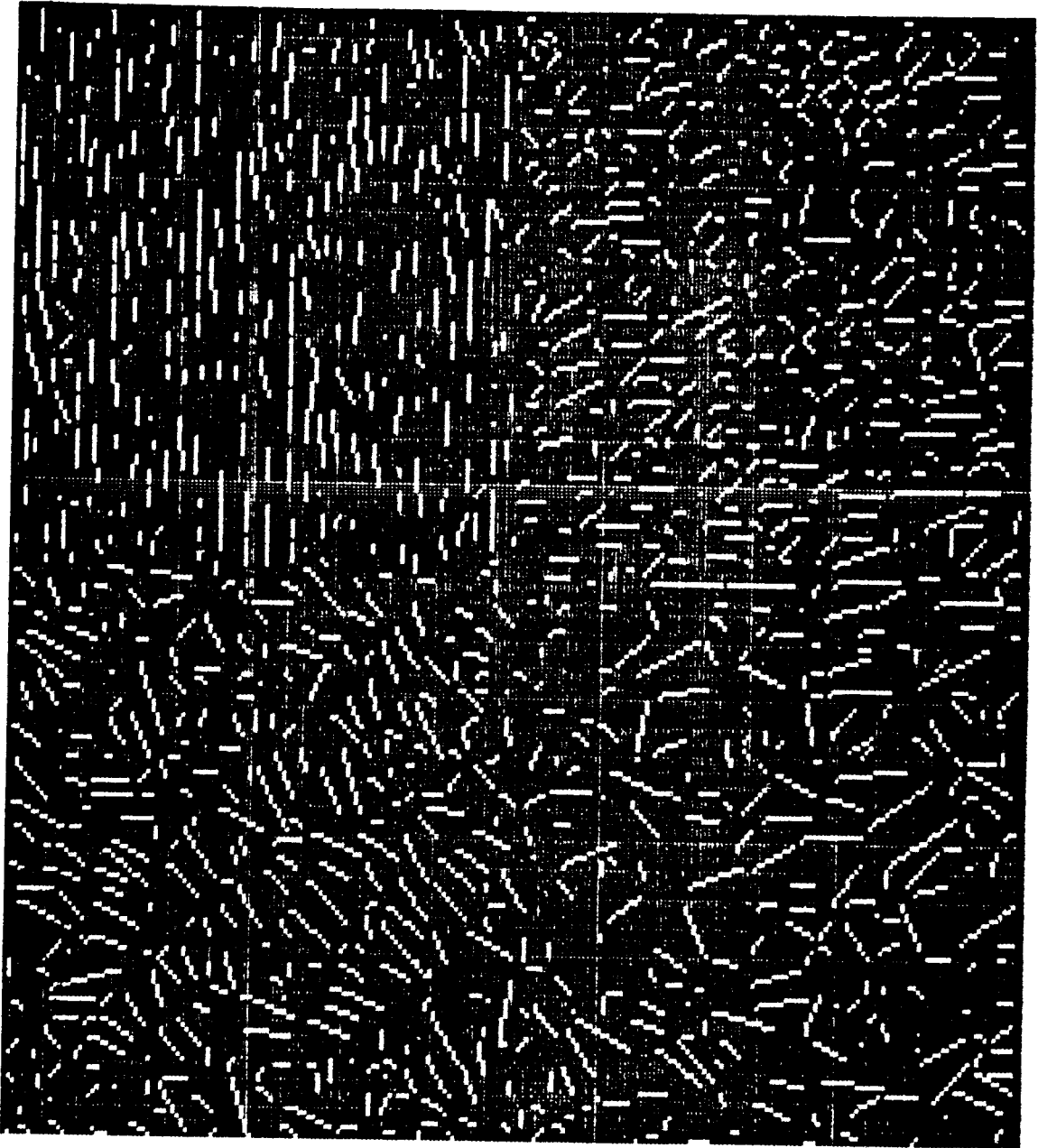
Fig 2 - As imagens originais, retiradas de Brodatz (1968) de (a) veio (b) rãfia, (c) cortiça e em (d) mica.

Para cada tipo de textura, uma descrição simbólica isolada para cada canal $\nabla^2 G$ com W (diâmetro do operador $\nabla^2 G$) de 4, 6, 8 e 10 pixels foi obtida. Escolhemos para trabalhar os canais com $W = 6$ e 8 pixels, visto que em $W = 4$ pixels, houve uma detecção bastante ruidosa e em $W = 10$ havia pouca informação. Sobre estas descrições, verificamos as distribuições de TAMANHO e ORIENTAÇÃO para as primitivas encontradas. A discriminação das texturas, baseia-se na análise destas distribuições.

No canal com $W = 6$ pixels, o veio e a rãfia são diretamente separáveis observando a distribuição do TAMANHO. O veio apresenta um histograma com pico em (18) em uma dada orientação preferencial (no caso $w90$ graus), enquanto a rãfia apresenta picos em (18) e (4) com duas orientações preferenciais defasadas de aproximadamente 180 graus. Estas configurações das primitivas, definindo estruturas diferenciadas para a rãfia e para o veio, discriminam visualmente as texturas e podem ser quantizadas. Veja na Fig 3 as descrições (conjunto de primitivas geradas) para o canal $W = 6$ pixels e na Fig 4 as distribuições para o veio e a rãfia.

(a)

(b)

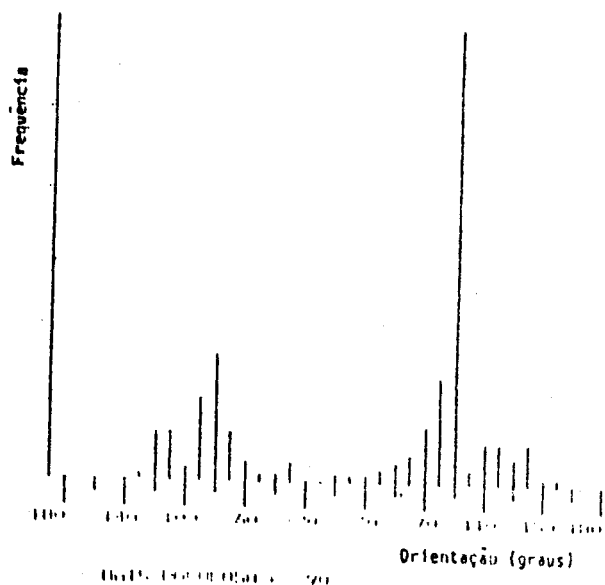
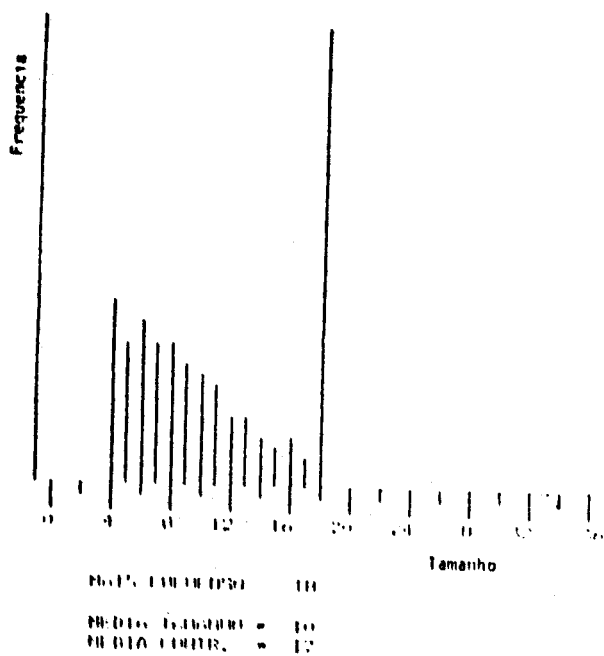


(c)

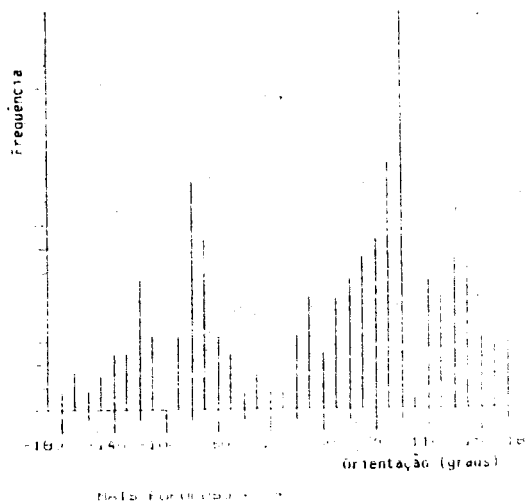
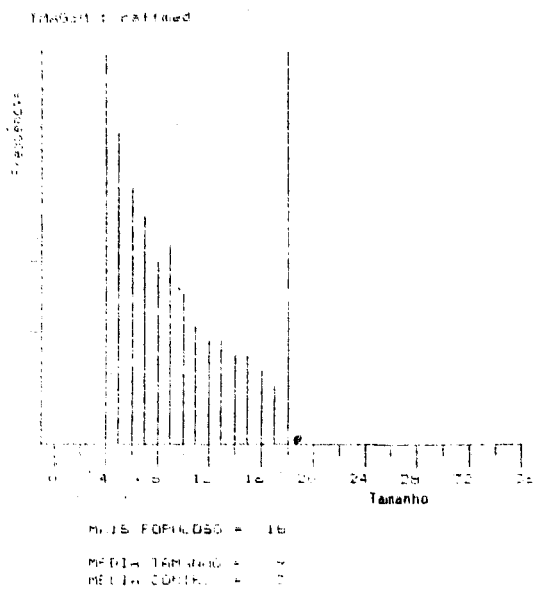
(d)

Fig 3 - Primitivas para (a), o veio, (b), a rãfia, (c) cortiça e, em (d), a mica

a



(a)



(b)

Fig 4 - Distribuições de TAMANHO e ORIENTAÇÃO para (a), o veio, e (b), a rãfia.

O veio e a rãfia se diferenciam diretamente da cortiça, também observando a distribuição de TAMANHO. A cortiça tem concentração de primitivas com tamanho (4), sem nenhuma orientação preferencial. As primitivas distribuem-se em todas as direções, não se podendo determinar alinhamentos direcionais. Esta organização caracteriza a textura, diferenciando-a do veio e da rãfia. Na Fig 5 apresentamos as distribuições de TAMANHO e ORIENTAÇÃO da cortiça para o canal $\nabla^2 G$ com $W = 6$ pixels.

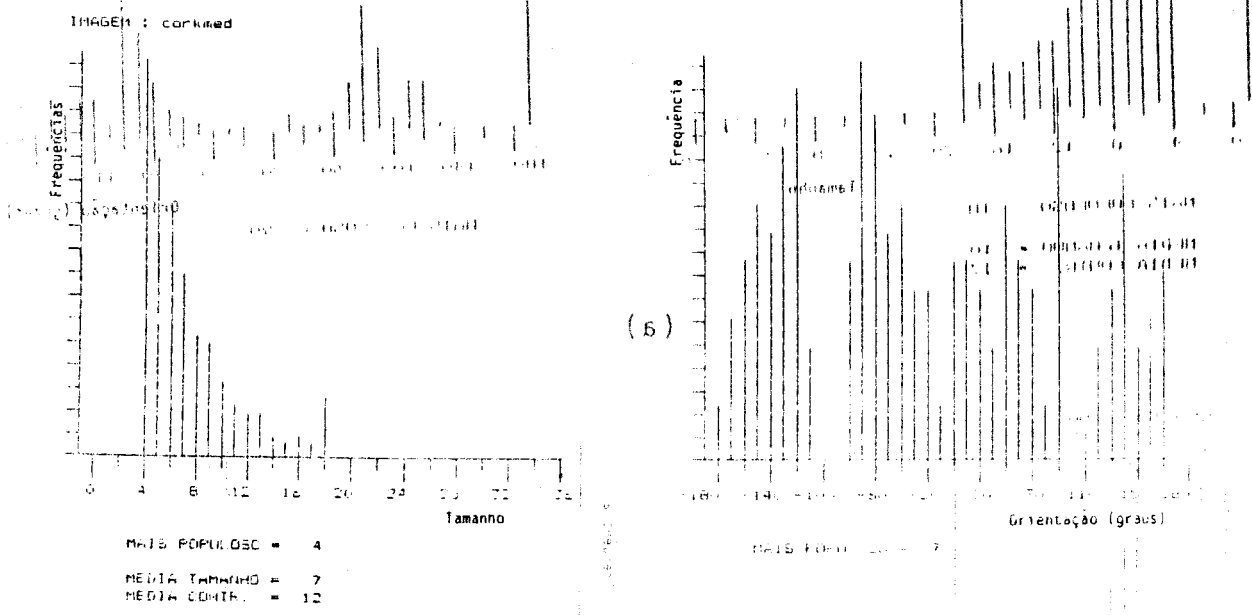


Fig 5 - Distribuição em TAMANHO e ORIENTAÇÃO.

A mica diferencia-se do veio e da rãfia pelas mesmas características apresentadas em relação a cortiça. Na Fig 6 tem-se as distribuições da mica para o canal $\nabla^2 G$ com $W = 6$ pixels.

(d)

Fig 4 - Distribuições de TAMANHO e ORIENTAÇÃO para o veio e a rãfia.

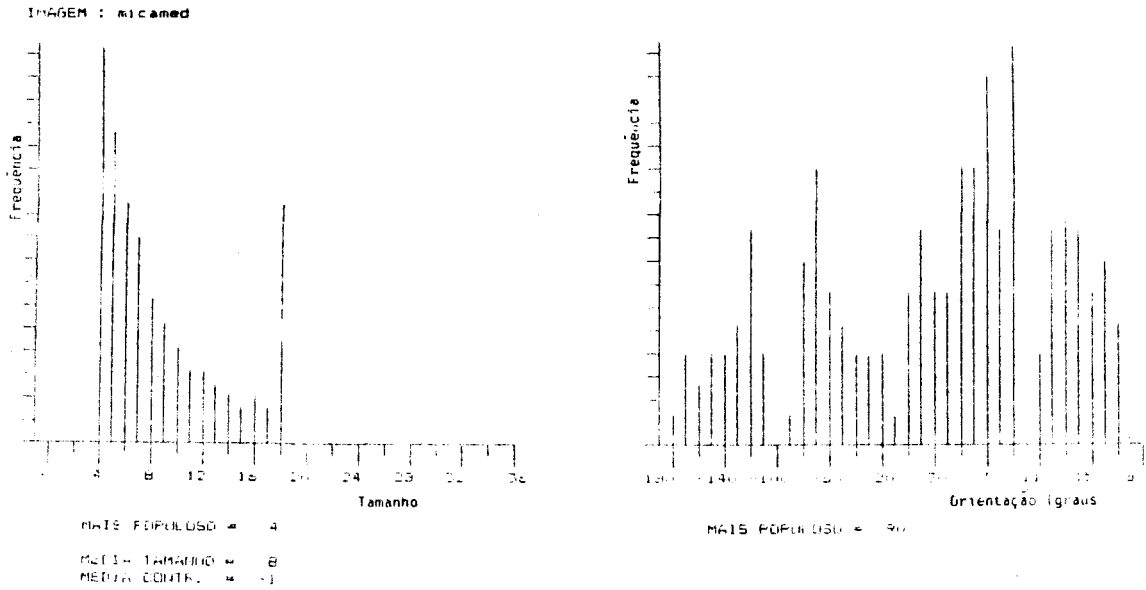


Fig 6 - Distribuição em TAMANHO e ORIENTAÇÃO.

Aparece, então, um problema. Como diferenciar a mica e a cortiça. Observando as imagens originais de ambas na Fig 1, é fácil verificar que a mica é semelhante a cortiça, só que em uma escala maior. Visto isto, tentamos estabelecer um novo canal $\nabla^2 G$ e verificar os resultados. Utilizando $\nabla^2 G$ com $W = 8$ pixels obtivemos os resultados apresentados na Fig 7.

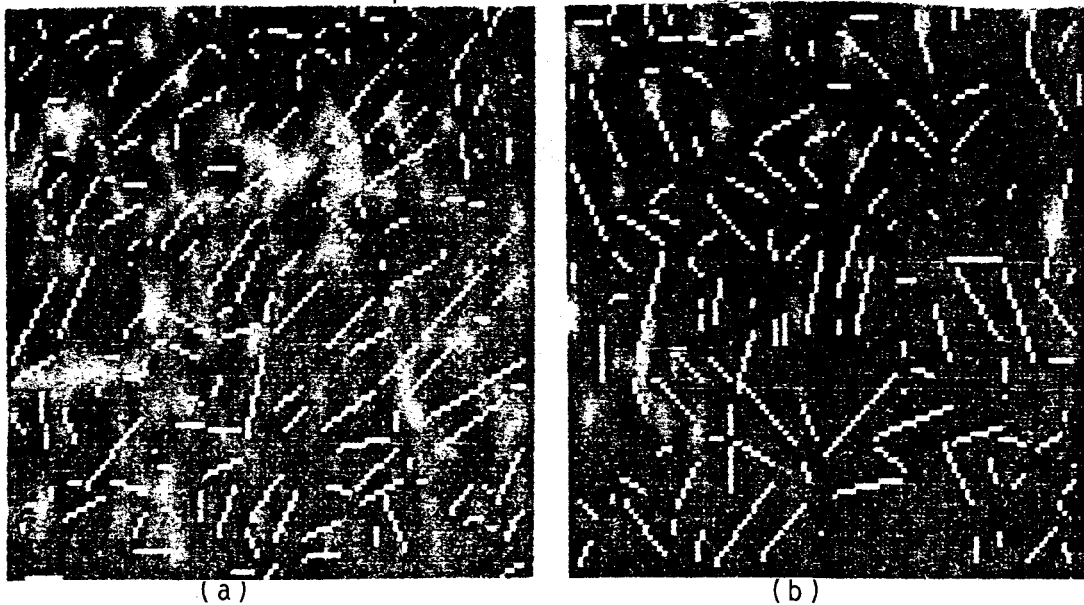
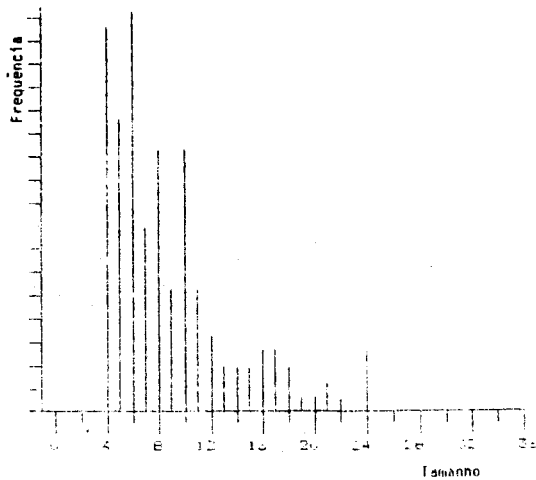
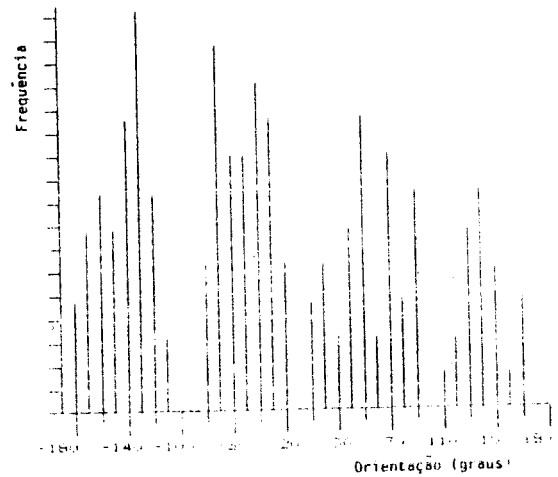


Fig 7 - Descrição com $W = 8$ pixels para $\nabla^2 G$ em (a) cortiça e (b) mica.

TRAGEM : cortiSWB

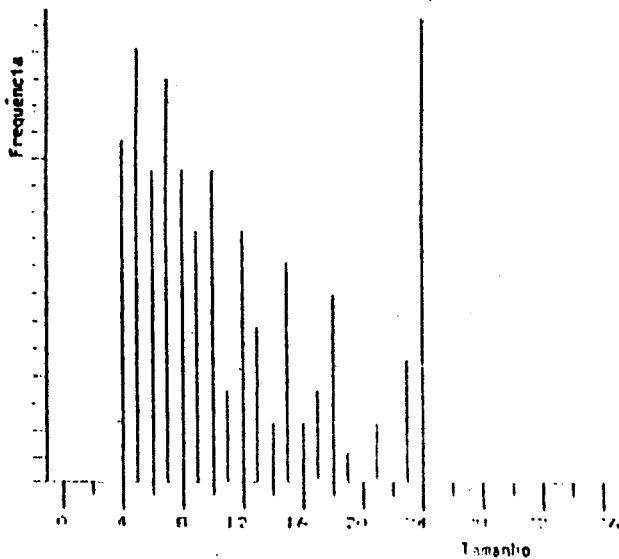


MAIS FREQUENTE = 4
 MEDIA TAMANHO = 6
 MEDIA COEFIC. = 17

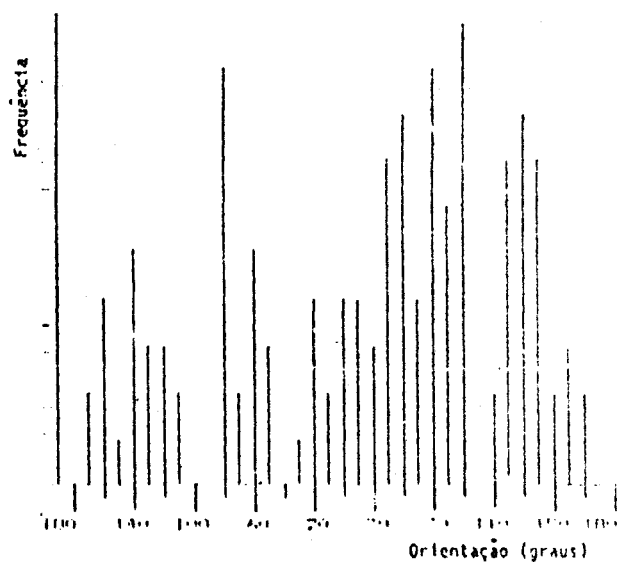


MAIS FREQUENTE = 140

(c)



MAIS FREQUENTE = 24
 MEDIA TAMANHO = 11
 MEDIA COEFIC. = 10



MAIS FREQUENTE = 20

(d)

Fig 7 - Distribuições de TAMANHO e ORIENTAÇÃO em (c) cortiça e (d) mica, para $W = 8$ pixels

Para este canal, conseguimos determinar uma nova caracterização para as texturas. A distribuição de TAMANHO para a mica, passou a definir uma organização com primitivas de tamanho (24), formando a base de percepção da cena. A cortiça, continuou definindo primitivas de tamanhos entre (4) e (6) como sua base perceptiva. Ambas sem preferências direcionais, portanto não definindo alinhamentos. Esta diferenciação, possibilitada pela nova escala do operador $\nabla^2 G$, torna possível discriminar a mica da cortiça.

4 - A ANÁLISE DOS RESULTADOS

No escopo deste trabalho, objetivamos com estes experimentos, estabelecer empiricamente, a possibilidade de uso de um modelo estrutural de representação para dados visuais para caracterizar texturas. Fazendo isto, trabalhando somente estatísticas simples sobre a representação simbólica da cena. Nosso objetivo posterior é trabalhar a segmentação da imagem, baseando nosso algoritmo na extração de características de textura da cena, utilizando esta abordagem aqui apresentada. Um EPB local na imagem, forneceria medidas baseadas nas distribuições, de maneira a caracterizar quantitativamente a textura na região. Trabalhando então com grafos, onde os nós representariam regiões com mesmas características de textura, buscaríamos definir regiões ou limiares entre regiões.

O experimento realizado indica a viabilidade em tentar esta abordagem para segmentação por textura. Com isto, passamos a utilizar uma representação simbólica e estrutural da cena para executar tarefas visuais, aproximando o trabalho do modelo de visão por máquina da forma de trabalho dos modelos de percepção visual biológicos.

REFERÊNCIAS BIBLIOGRÁFICAS

Brady, M.; 1982, **Computational approaches to image understanding**. Computing surveys, vol 14, n1, mar, p. 3 - 71

Connors, R.W.; Triveldi, M.M.; Harlow, C.A.; 1984, **Segmentation of a high-resolution urban scene using texture operators**. Computer vision, graphics and image process, vol 25, p. 273 - 310

Haralick, R.M.; Shanmugam, K.; Dinstein, I.; 1973, **Textural features for image classification**. IEEE Trans. sys. man. cyb., vol 3, p. 610 - 621

Marr, D.; 1978, **Representing a visual information - a computational approach**. Computer vision systems, ed. Hanson, A.R.; Riseman, E.M., Academic Press, New York, p. 61 - 80

Marr, D.; 1982, **Vision**. W. H. Freeman and company, San Francisco.

Monteiro, A.M.V.; 1988, **Processamento de imagens de satélites usando estruturas simbólicas do baixo nível da visão**, INPE-4638-TDL/337, Dissertação de Mestrado.