

Classificação de Texturas Invariante a Rotação Usando Matriz de Co-ocorrência

Rodrigo Hajime Ito, Hae Yong Kim, Walter Jaimes Salcedo
Dept. Eng. Sistemas Eletrônicos, Escola Politécnica
Universidade de São Paulo, Brasil
rodrigoito03@gmail.com, hae@lps.usp.br, wsalcedo@lme.usp.br

Resumo — Neste trabalho, modificamos o bem-conhecido método para classificação de texturas baseado em matriz de co-ocorrência para torná-lo invariante à rotação. Para isso, propomos duas novas matrizes de co-ocorrência: médias sobre circunferências (MC) e médias radiais (MR). Essas novas matrizes distinguem-se do original, pois consistem em comparar as médias de intensidades de pixels sobre determinadas geometrias, e não apenas sobre os tons de cinza de pares de pixels. A partir das duas novas matrizes, obtemos três métodos de classificação de texturas: utilizando somente MC, somente MR e conjuntamente MC/MR. Os experimentos realizados mostraram taxas de acerto de até 94% para o método MC, 95% para o MR e de 100% para o método conjunto MC/MR. Já o método original apenas apresentou uma taxa de acerto de 66%. Esses resultados mostram que o método proposto é altamente eficiente na classificação de texturas rotacionadas.

Palavras-chaves - Matriz de co-ocorrência; classificação de textura; invariância à rotação.

I. INTRODUÇÃO

As texturas são usadas naturalmente pela percepção humana no reconhecimento de objetos e de padrões e, portanto, são características importantes na análise de imagens. A análise de textura pode ser utilizada em várias aplicações, como a inspeção industrial, biomédica, classificação e análise de imagens de sensoriamento remoto e busca e identificação por imagens com determinadas características num banco de dados. Em geral a literatura reporta quatro métodos de análises de textura conhecidas como segmentação de textura, síntese de textura, determinação da forma a partir de texturas e a classificação de textura. A classificação de imagens por textura consiste em categorizar uma amostra de imagem de classe desconhecida a uma das classes de texturas conhecidas. Um problema na análise de textura no mundo real é frequentemente a não-uniformidade devido a mudanças de orientação, escala e outros aspectos visuais.

A comunidade científica em processamento de imagens não têm conseguido um consenso para definir de maneira única e universal a textura. Segundo Sklansky [1], uma região de imagem é definida como sendo de textura constante se um

conjunto de propriedades locais nessa região é constante, lentamente variante ou periódico. Normalmente, a classificação de texturas baseia-se na extração de características, de forma a descrevê-las simplificadamente. Haralick et al. [2] propuseram um método para extração de características de texturas baseado na relação espacial existentes entre os níveis de cinza da imagem, calculando as chamadas matrizes de co-ocorrência e, a partir delas, determinaram 14 parâmetros estatísticos que caracterizam as texturas. Devido a sua simplicidade e eficácia, este método tornou-se bastante popular na classificação de texturas, sendo considerado por muitos autores como “o método de referência” em relação ao qual os outros métodos são comparados. Segundo Baraldi e Parmiggiani [3], dentre as 14 medidas estatísticas propostas originalmente, as seguintes seis apresentam maior relevância: o segundo momento angular, a entropia, o contraste, a variância, a correlação e a homogeneidade.

Uma característica desejável de um método de classificação de textura é a invariância à rotação. Infelizmente, o método original de matriz de co-ocorrência é sensível à rotação de textura. No presente trabalho, propomos métodos alternativos para a construção das matrizes de co-ocorrência, de forma a tornar essas matrizes completamente invariantes à rotação da imagem em análise.

II. MATRIZ DE CO-OCORRÊNCIA

Os elementos de uma matriz de co-ocorrência descrevem a frequência com que ocorrem transições nos níveis de cinza em uma imagem [2, 4]. Ou seja, um elemento na linha i e coluna j com valor p nessa matriz indica que houve p variações do nível de cinza i para o nível de cinza j . Portanto, o número de linhas e colunas dessa matriz é proporcional à quantidade de níveis de cinza e independe das dimensões da imagem em análise. Para o cálculo dessa matriz, são definidas as relações espaciais entre os pixels em que serão feitas as análises da variação do nível de cinza.

A matriz de co-ocorrência é definida formalmente a seguir: Uma imagem I é uma função que associa um tom de cinza para cada pixel, ou seja, $I : D \rightarrow G$, onde D é o domínio da imagem e $G = \{1, 2, \dots, N_g\}$ é o conjunto dos níveis de cinza. Seja $S \subset D \times D$ um conjunto de pares ordenados de pixels.

S é o conjunto que define quais pares de pixels serão considerados na análise da variação dos níveis de cinza. Define-se a matriz de co-ocorrência $P(i,j)$ como:

$$P(i,j) = \#\{(u,v) \in S \mid I(u) = i \text{ e } I(v) = j\} \quad (1)$$

Após o cálculo da matriz de co-ocorrência, efetua-se sua normalização dividindo cada um de seus elementos pela soma da matriz:

$$p(i,j) = \frac{P(i,j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i,j)} \quad (2)$$

Haralick et al. [2] definiram S para determinados ângulos e distância entre pixels. Para isso, são criados dois parâmetros adicionais d e θ , que representam respectivamente a distância entre os dois pixels e o ângulo que a linha que passa por estes pixels faz com a horizontal. A maioria das técnicas que usam a matriz de co-ocorrência assume quatro valores para θ : 0° , 45° , 90° e 135° , conforme mostra a figura 1. Para essas relações espaciais entre pixels, quatro matrizes de co-ocorrência são construídas, uma para cada ângulo.

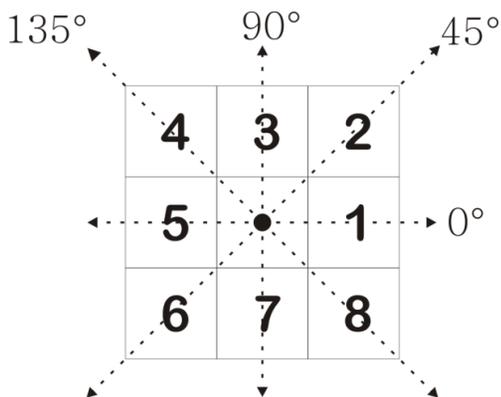


Figura 1: Ângulos utilizados no cálculo da matriz de co-ocorrência clássica [2].

III. CARACTERÍSTICAS DE TEXTURA

Para descrever as texturas, Haralick et al. propuseram 14 medidas estatísticas, calculadas a partir da matriz de co-ocorrência [2]. No entanto, segundo Baraldi e Parmiggiani [3], apenas seis delas são as mais relevantes: O segundo momento angular (3), a entropia (4), o contraste (5), a variância (6), a correlação (7) e a homogeneidade (8), definidas através das expressões abaixo. No Anexo I são descritos os significados “físico-intuitivos” dessas medidas.

$$f_{sma} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \{p(i,j)\}^2 \quad (3)$$

$$f_{ent} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log(p(i,j)) \quad (4)$$

$$f_{con} = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}_{|i-j|=n} \quad (5)$$

$$f_{var} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu)^2 p(i,j) \quad (6)$$

$$f_{cor} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (7)$$

$$f_{hom} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i-j)^2} p(i,j) \quad (8)$$

onde $p(i,j)$ é a matriz de co-ocorrência normalizada e $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i p(i,j)$; $p_x(i) = \sum_{j=1}^{N_g} p(i,j)$ e $p_y(j) = \sum_{i=1}^{N_g} p(i,j)$ são as distribuições marginais, μ_x , μ_y são seus valores médios, σ_x e σ_y são seus desvios padrões.

IV. MÉTODO PROPOSTO

Propomos neste trabalho duas novas definições de relações espaciais para o cálculo da matriz de co-ocorrência, com o objetivo de torná-la invariante à rotação. Um dos métodos consiste em calcular as médias dos níveis de cinza sobre duas circunferências (MC) concêntricas e de raios distintos (figura 2); o outro método consiste em calcular médias dos pixels sobre retas radiais (MR) em diversas orientações (figura 3). As médias sobre circunferências são inerentemente invariantes à rotação. As médias sobre linhas radiais não são inerentemente invariantes à rotação, mas os parâmetros estatísticos determinados a partir delas são invariantes à rotação. Estes métodos estão inspirados nas técnicas que utilizam as projeções circulares e radiais para efetuar o casamento de padrões [6, 7].

A. Média sobre circunferência (MC)

Dado um pixel $(j,k) \in D$, um raio interno r_i e um raio externo r_e , definem-se pelas expressões (9) e (10) os conjuntos dos pixels que formam a circunferência interna e externa centradas em (j,k) :

$$C_i(j,k) = \{(l,m) \in D \mid \lfloor \sqrt{(l-j)^2 + (m-k)^2} \rfloor = r_i\} \quad (9)$$

$$C_e(j,k) = \{(l,m) \in D \mid \lfloor \sqrt{(l-j)^2 + (m-k)^2} \rfloor = r_e\} \quad (10)$$

onde $\lfloor \cdot \rfloor$ indica que o argumento será arredondado para o inteiro mais próximo. As médias dos pixels sobre as circunferências de raios r_i e r_e centradas em (j,k) são calculadas através das seguintes equações:

$$F_i(j,k) = \frac{\sum_{(l,m) \in C_i(j,k)} I(l,m)}{\#C_i(j,k)} \quad (11)$$

$$F_e(j, k) = \frac{\sum_{(l,m) \in C_e(j,k)} I(l, m)}{\#C_e(j, k)} \quad (12)$$

Os elementos $P(i, j)$ da matriz de co-ocorrência MC são definidas como:

$$P(i, j) = \#\{(l, m) \in D \mid F_i(l, m) = i, F_e(l, m) = j\} \quad (13)$$

Os elementos dessa matriz são normalizados através da expressão (2), obtendo-se assim a matriz normalizada p .

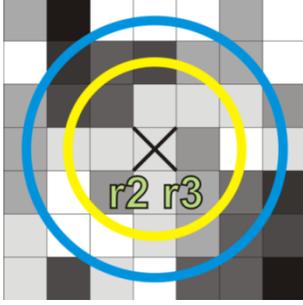


Figura 2: Representação da geometria para o cálculo das médias sobre circunferências de raios $r_i=2$ e $r_e=3$

B. Médias Radiais (MR)

Define-se o conjunto dos pixels pertencentes a uma direção radial θ , de comprimento r , centrado no pixel (j, k) como:

$$\rho_\theta(j, k) = \{\{\lfloor j - t \sin \theta \rfloor, \lfloor k + t \cos \theta \rfloor\}\} \quad (14)$$

onde t são números inteiros entre $0 \leq t \leq r$. Como antes, $\lfloor \cdot \rfloor$ denota o arredondamento para o inteiro mais próximo. Uma vez determinados estes pixels, calcula-se o valor médio de nível de cinza para os pixels radiais centrado em (j, k) e com direção θ pela expressão:

$$F'_\theta(j, k) = \frac{\sum_{(l,m) \in \rho_\theta(j,k)} I(l, m)}{\#\rho_\theta(j, k)} \quad (15)$$

A figura 3 ilustra as 24 direções radiais que utilizamos nos testes, para o raio cinco. Para obter maior robustez, em vez de usar diretamente as médias radiais, utilizamos uma média ponderada de quatro médias radiais consecutivas, dando mais peso para as duas médias centrais:

$$F_n(j, k) = \frac{F'_{(3n)15^\circ} + 2F'_{(3n+1)15^\circ} + 2F'_{(3n+2)15^\circ} + F'_{(3n+3)15^\circ}}{6} \quad (16)$$

sendo $n \in \{0, 1, \dots, 7\}$. A partir deste conjunto de valores, os elementos $P(i, j)$ da matriz de co-ocorrência MR são definidos:

$$P(i, j) = \#\{(l, m) \in D \mid F_k(l, m) = i, F_{k+1}(l, m) = j\} \quad (17)$$

sendo $k \in \{0, 1, \dots, 7\}$. A figura 3 indica, em azul, as direções que possuem peso 1 e, em vermelho, as direções que possuem

peso 2. Novamente, utiliza-se a expressão (2) para normalizar a matriz de co-ocorrência P , obtendo a matriz normalizada p .

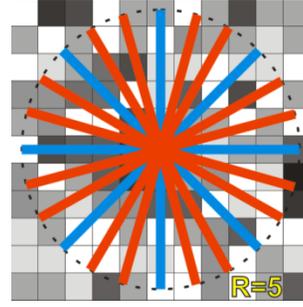


Figura 3: Representação geométrica para as regiões radiais para as médias radiais

V. PROCEDIMENTO EXPERIMENTAL

Para os testes, foram utilizadas 18 imagens de texturas 512×512 obtidas no site da USC (*University of Southern California*), sendo que 12 delas são originárias da biblioteca Brodatz. Os testes foram subdivididos em duas etapas: etapa de treinamento e etapa de reconhecimento. Para tal fim, foi definido um conjunto de amostras para ser utilizado no treinamento e outro conjunto de amostras para ser utilizado no reconhecimento.

A. Conjunto de amostras

O conjunto de amostras para a etapa de treinamento foi gerado a partir de todas as 18 texturas 512×512 utilizando apenas o quadrante superior esquerdo dessas imagens, gerando imagens de 256×256 , como ilustra a figura 4a.

Para o reconhecimento, as amostras também foram geradas a partir das 18 texturas 512×512 originais, sendo extraídas do setor interno dessas imagens submetidas a seis rotações utilizando a interpolação bilinear: $1 \times 360^\circ/7$, $2 \times 360^\circ/7$, $3 \times 360^\circ/7$, $4 \times 360^\circ/7$, $5 \times 360^\circ/7$ e $6 \times 360^\circ/7$. Foi gerado, portanto, um conjunto com 108 imagens 256×256 para a fase de reconhecimento.

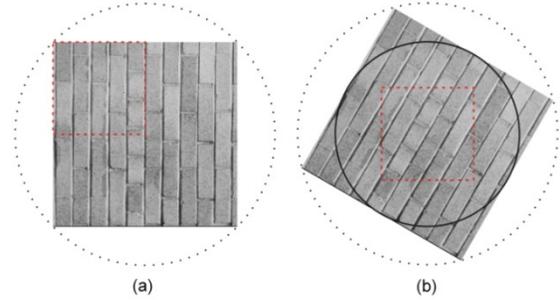


Figura 4: (a) geração das imagens para o treinamento; (b) geração das imagens para o reconhecimento

B. Fase de Treinamento

Para todas as imagens do conjunto de treinamento, calculamos as matrizes de co-ocorrência pela expressão (13) ou (17) (dependendo do método utilizado), normalizamos pela

expressão (2) e, em seguida, calculamos as seis características de textura pelas expressões (3) a (8). Assim obtivemos, para cada imagem de treinamento, um vetor de seis elementos. No total, obtivemos uma matriz com 18 linhas e 6 colunas, cada linha associada a uma amostra. Essas informações são normalizadas pela seguinte expressão:

$$y_i = \frac{x_i - \mu_i}{\sigma_i} \quad (18)$$

onde x_i é a i -ésima característica a ser normalizada, μ_i é a média dessas características e σ_i é o seu desvio padrão. Após a normalização, y_i , μ_i e σ_i são salvos para serem utilizados durante o reconhecimento na normalização das características.

C. Fase de Reconhecimento

A partir de cada uma das 108 imagens de reconhecimento, calculamos as matrizes de co-ocorrência pela expressão (13) ou (17) (dependendo do método utilizado), normalizamos pela

expressão (2) e calculamos as seis características através das expressões (3) a (8), da mesma forma que fizemos na fase de treinamento. Geramos um vetor com as seis características, cujos valores são normalizados pela expressão (18) com os mesmos parâmetros μ_i e σ_i da fase de treinamento. Então, comparamos esse vetor com cada linha da matriz gerada na fase de treinamento através da distância euclidiana. A linha que apresentar a menor distância é associada à textura em análise. Esse procedimento foi realizado para todas as 108 amostras do conjunto de reconhecimento.

D. Teste dos métodos em conjunto

Nos itens anteriores explicamos o funcionamento dos métodos MC e MR separadamente. No método conjunto MC/MR, a única diferença é que são geradas 2 matrizes de co-ocorrência (uma para cada método) e, portanto, são calculadas 12 características de texturas (6 de cada matriz). Assim cada imagem é representada por um vetor de 12 elementos.

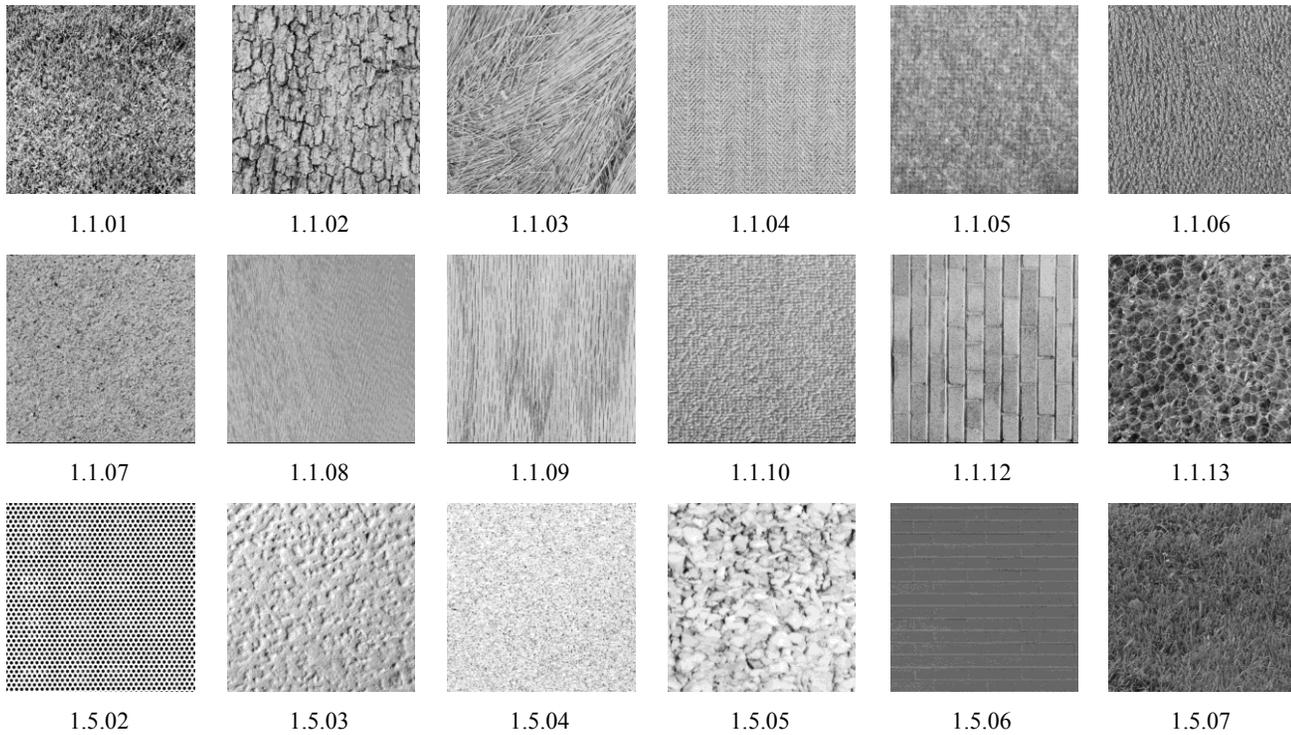


Figura 5: As 18 texturas utilizadas nos testes

VI. RESULTADOS

O método proposto foi implementado em C++ utilizando a biblioteca de rotinas para processamento de imagens Pro-Eikon, disponível no site:

<http://www.lps.usp.br/~hae/software>.

Como foi mencionado anteriormente, as matrizes de ocorrência dependem diretamente da quantidade de níveis de cinza da imagem. Portanto, em todos os métodos quantizamos as imagens utilizando 16, 32, 64, 128 e 256 tonalidades de cinza (a última escala corresponde às imagens originais).

A tabela 1 mostra as taxas de acerto obtidas pelo método das médias sobre circunferências. Foram realizados testes variando os raios interno e externo, conforme indicado na tabela. A maior taxa de acerto foi de 94%.

Na tabela 2 são apresentadas as taxas de acerto obtidas em testes utilizando o método das médias radiais, variando o

valor do raio de 2 a 6, obtendo a taxa de acerto máxima de 95%.

A tabela 3 mostra as taxas de acerto obtidas utilizando as características calculadas a partir dos dois métodos conjuntamente. Utilizando os parâmetros mais adequados, obtivemos até 100% de taxa de acerto. Mesmo variando um pouco os parâmetros em torno do ótimo, altas taxas de acertos foram observadas para este método.

Esses resultados mostram que ambos os métodos possuem eficiência semelhantes quando utilizados separadamente, mas melhoram seu desempenho quando são utilizados em conjunto, atingindo uma taxa de acerto de até 100%. Comparado com o método tradicional cuja maior taxa de acerto foi 66% (tabela 4), o método proposto apresenta uma taxa de acerto claramente superior para o reconhecimento de texturas submetidas a rotações.

TABELA 1: TAXA DE ACERTO (EM %) DO MÉTODO DAS MÉDIAS SOBRE CIRCUNFERÊNCIAS

Raios		Níveis de Cinza				
Interno	Externo	16	32	64	128	256
2	3	70,37%	71,30%	75,00%	77,78%	79,63%
2	4	93,52%	94,44%	94,44%	94,44%	94,44%
3	4	87,04%	85,19%	85,19%	84,26%	85,19%
2	5	94,44%	94,44%	94,44%	94,44%	94,44%

TABELA 2: TAXA DE ACERTO DO MÉTODO DAS MÉDIAS RADIAIS

Raio	Níveis de cinza				
	16	32	64	128	256
2	75,00%	79,63%	84,26%	84,26%	83,33%
3	87,04%	86,11%	84,26%	85,19%	86,11%
4	87,04%	90,74%	88,89%	89,81%	90,74%
5	86,11%	92,59%	94,44%	93,52%	89,81%
6	85,19%	91,67%	94,44%	95,37%	95,37%

TABELA 3: TAXA DE ACERTO UTILIZANDO O MÉTODO DAS MÉDIAS RADIAIS E O MÉTODO DAS MÉDIAS SOBRE CIRCUNFERÊNCIAS SIMULTANEAMENTE

MR	MC		Níveis de Cinza				
	r_i	r_e	16	32	64	128	256
4	2	3	85.19%	90.74%	86.11%	86.11%	87.96%
4	2	4	93.52%	94.44%	95.37%	96.30%	95.37%
4	2	5	93.52%	96.30%	99.07%	99.07%	100.00%
4	3	4	93.52%	92.59%	86.11%	86.11%	86.11%
5	2	3	85.19%	89.81%	88.89%	87.96%	87.04%
5	2	4	94.44%	94.44%	97.22%	97.22%	96.30%
5	2	5	93.52%	95.37%	98.15%	99.07%	99.07%
5	3	4	93.52%	92.59%	89.81%	86.11%	87.04%

TABELA 4: TAXA DE ACERTO DO MÉTODO TRADICIONAL

Níveis de cinza	Acerto
16	61,11%
32	62,96%
64	62,04%
128	59,26%
256	65,74%

VII. CONCLUSÃO

No presente trabalho, propusemos uma alteração no bem-conhecido método de matriz de co-ocorrência para classificação de texturas, para torná-lo invariante à rotação. Para isso, utilizamos médias circulares (MC) e médias radiais (MR) para obter matrizes de co-ocorrências insensíveis a rotações. Obtivemos taxas de acerto de aproximadamente 95% quando MC ou MR são utilizados isoladamente. Quando MC e MR são utilizados conjuntamente, a taxa de acerto chegou até 100%. Estes valores são significativamente superiores aos obtidos pelo método convencional (66%).

VIII. AGRADECIMENTOS

Os autores agradecem ao Conselho Nacional de desenvolvimento Científico e Tecnológico (CNPq) pela bolsa de iniciação científica recebida.

ANEXO I

Segundo momento angular: Esse parâmetro também é chamado de energia e uniformidade. O segundo momento angular mede a uniformidade da textura, ou seja, a quantidade de repetições de pares de pixels. Essa medida terá valores altos quando os valores da matriz de co-ocorrência estiverem concentrados, ou seja, quando houver grande repetição na variação dos níveis de cinza. Assume valores positivos menor ou igual a 1, para a matriz normalizada.

Entropia: Esse parâmetro mede, como o nome sugere, a desordem da imagem. Portanto, essa medida assume valores elevados quando a imagem possui textura não uniforme. A entropia está fortemente relacionada, porém de maneira inversa, com o segundo momento angular [3].

Contraste: A diferença entre o maior e o menor valor de um determinado conjunto contíguo de pixels é definida como frequência espacial. A expressão do contraste indica que uma imagem com valor baixo nesse parâmetro não é necessariamente caracterizada por uma distribuição de tons de cinza estreita, mas sim por uma baixa frequência espacial, possuindo

do valores da matriz de co-ocorrência concentrados próximos à diagonal principal.

Variância: Essa é uma medida da heterogeneidade da textura. O valor da variância aumenta quando os valores dos tons de cinza diferem de sua média.

Correlação: É a medida da dependência linear dos tons de cinza em uma imagem. Altos valores de correlação implicam uma relação linear entre os níveis de cinza dos pares de pixels.

Homogeneidade: Também chamado de Momento da Diferença Inversa, esse parâmetro mede, como o nome diz, a homogeneidade de uma imagem. Essa medida é sensível a valores próximos à diagonal da matriz de co-ocorrência e, portanto, ao baixo contraste da textura. A Homogeneidade está inversamente relacionada com o Contraste e a Energia da imagem.

REFERÊNCIAS

- [1] SKLANSKY, J. Image segmentation and feature extraction, 1978 – *IEEE Transactions on Systems, Man and Cybernetics*, 8, p.237-247.
- [2] HARALICK, R. M., SHANMUGAM, K. e DINSTEIN, I. Textural Features for Image Classification, 1973 – *IEEE Transactions on Systems, Man and Cybernetics*, 3 p.610-620.
- [3] BARALDI, A. e PARMIGGIANI, F. An Investigation of the Textural Characteristics Associated with Gray Level Co-occurrence Matrix Statistical Parameters, 1995 – *IEEE Transactions on Geoscience and Remote Sensing*, 33(2) p.293-304.
- [4] SCHWARTZ, W. R. Segmentação de imagens baseada em dependência espacial utilizando campo aleatório de Markov associado com características de texturas, 2005, Tese (Mestrado) – *Setor de Ciências exatas, Universidade Federal do Paraná*. cap. 2, p.19-23.
- [5] HARALICK, R. M., SHAPIRO, L. G. Glossary of Computer Vision Terms, 1991 – *Pattern Recognition*, 24(1) p.69-93.
- [6] H. Y. Kim and S. A. Araújo, “Grayscale Template-Matching Invariant to Rotation, Scale, Translation, Brightness and Contrast,” *IEEE Pacific-Rim Symposium on Image and Video Technology*, Lecture Notes in Computer Science, vol. 4872, pp. 100-113, 2007.
- [7] H. Y. Kim, “Rotation-Discriminating Template Matching Based on Fourier Coefficients of Radial Projections with Robustness to Scaling and Partial Occlusion,” *to appear in Pattern Recognition*, 2009.