

Information System using Computer-Aided Diagnosis and Content-Based Image Retrieval

Elias Borges Macena Junior, Ronaldo Martins da Costa,
Leandro Luis Galdino de Oliveira, Rogerio Lopes Salvini
Instituto de Informática
Universidade Federal de Goiás – UFG
{jrlinked, ronaldomc12, leandroluis.inf,
rogerio.salvini}@gmail.com

Fátima de Lourdes dos Santos Nunes Marques
Escola de Artes, Ciências e Humanidades
Universidade de São Paulo – USP
fatima.nunes@usp.br

Abstract—To improve the process of diagnosis with improved quality several studies center have focused their research on the development of information systems applying powerful computer-aided diagnosis techniques (CAD). In this context, the creation of databases to images with appropriate structures, capable of store images from different imaging systems with different characteristics and specifications is the first step to develop an efficient CAD system. This work proposes the validation of storage techniques and recovery with a hybrid CBIR method based on 2D medical images. The results of the techniques applied indicate a gain of 41% in image retrieval speed, opening way for the development of more efficient information systems to support diagnosis.

Keywords—Computer-Aided Diagnosis, Content-Based Retrieve

I. INTRODUÇÃO

Bancos de imagens médicas constituem uma necessidade fundamental para a avaliação e desenvolvimento de sistemas de Diagnóstico Auxiliado por Computador (CAD). A utilização de técnicas de CBIR podem auxiliar na composição deste sistemas pois proveem aos profissionais da saúde casos similares a um determinado modelo fornecido como argumento de consulta[1][2][3].

Na tentativa de aprimorar o processo do diagnóstico, vários centros de pesquisa têm direcionado seus trabalhos para o desenvolvimento de sistemas CAD [4]. Segundo [3] e [5], aplicações de radiologia têm impulsionado o desenvolvimento de sistemas CAD e *Content Based Image Retrieval* (CBIR). A validação dos algoritmos de tais sistemas exige bases de imagens com casos variados e que contenham estruturas de interesse procuradas pelos algoritmos. Porém em sua grande maioria, os bancos de imagens possuem limitações como ferramentas de busca, imagens de baixa qualidade, pouca quantidade de imagens ou poucas variações de casos [4].

Neste contexto, a criação de um banco de imagens com estruturas capazes de armazenar adequadamente imagens providas de diferentes sistemas de captura, com diferentes características e especificações, é o primeiro passo para o desenvolvimento de um eficiente sistema CAD. [6] afirma ser fundamental para sistemas CAD e de CBIR que a base de imagens seja definida de forma eficiente capaz de armazenar informações relevantes das imagens para posterior recuperação.

Existem diversas técnicas para armazenar informações de uma imagem médica. [6] aplicou com sucesso Non-Negative Tensor Factorization (NTF), Filtros Gabor com Non-Negative Tensor Factorization (NMF) e Principal Components Analysis

(PCA) para extrair características de imagens e armazená-las em um banco de imagens bidimensionais (2D).

Também é possível observar cada vez mais presente nos centros médicos, centros de pesquisa e hospitais a utilização de equipamentos capazes de gerar imagens 3D que podem aumentar a precisão do diagnóstico. Paralelamente é possível observar, porém em menor escala, trabalhos que avaliam características 3D em imagens médicas [7][8][9]. No entanto, a grande maioria dos trabalhos que realizam armazenamento ou busca em imagens 3D passam pelo processo de reconstrução das mesmas o que não é uma operação simples[10][11][12][13].

Dada esta carência e demanda por bancos de imagens médicas, sejam eles 2D ou 3D, justifica a condução de estudos que possam sugerir especificações para a construção de bancos de imagens médicas 2D ou 3D. Uma correta definição pode impulsionar a implementação de bancos de imagens para aplicações médicas e também o desenvolvimento de sistemas CAD mais eficientes.

Com base nessas premissas, o foco deste trabalho é desenvolver e validar técnicas de armazenamento e principalmente tempo de recuperação utilizando CBIR em uma base com imagens médicas 2D. Este estudo pode contribuir para o desenvolvimento de sistemas CAD mais eficientes que auxiliem os profissionais de saúde.

II. METODOLOGIA

A. Armazenamento

Bancos de dados atualmente trabalham com arquivos multimídia, usando um campo do tipo BLOB (Binary Large Object - Grande Objeto binário). Trata-se um campo criado para o armazenamento de qualquer tipo de informação em formato binário.

Os bancos de dados contêm campos para tratar os arquivos multimídia, como as imagens, tornando o armazenamento uma etapa trivial [14][15][16][17]. A grande dificuldade e foco dos principais trabalhos publicados atualmente está em como recuperar essas informações de maneira eficiente.

Para solucionar esta dificuldade, desenvolvemos uma metodologia que a cada vez que uma nova imagem é adicionada ao banco, todas as características que serão utilizadas para recuperação serão extraídas, calculadas e armazenadas no banco, isto é feito para evitar que as mesmas operações sejam realizadas repetidas vezes, a cada nova busca, impactando diretamente na velocidade de recuperação das imagens. Além delas, outras informações poderão ser armazenadas junto da imagem, como nome do paciente, sexo,

idade, e qualquer informação que venha a ser útil para a recuperação.

B. Recuperação

A maneira usada para recuperar imagens em bancos de dados é conhecida como busca por similaridade na qual se inclui técnicas de CBIR. Funciona usando uma imagem como parâmetro de busca para recuperar as que se assemelham a ela. O CBIR funciona através de 3 passos:

- Extração de Características (a imagem é analisada e as características de interesse são extraídas através de algum descritor).
- Indexação de Características (Vetor de Características) Uma vez extraídas, é construído um vetor de características com as informações extraídas (ex: cor, dimensão, formas, etc.) da imagem.
- Comparação (É realizada a comparação entre os diversos vetores de características extraídos, nessa etapa que são definidas as imagens semelhantes que devem ser recuperadas).

Existem várias técnicas de CBIR, cada uma baseada em algum tipo de descritor, ou característica da imagem foco. Para implementar uma técnica de CBIR eficiente é preciso conhecer o tipo de imagem foco, nesse caso, serão imagens radiográficas do pulmão, com ou sem pneumonia. Na Fig. 1 pode ser visto um exemplo de imagem que será usada no desenvolvimento da técnica de CBIR a ser usada no banco de imagens. Uma vez estabelecida a base pretende-se expandi-la para outros tipos de imagens.



Fig. 1. Exemplo de imagem usada nos testes

Segundo [18] e [19] a melhor forma de trabalhar essas imagens é separando o pulmão esquerdo e direito e analisando-os separadamente. Isto é feito não apenas para descartar áreas sem informações importantes, mas também para gerar subimagens quadradas e tornar possível a aplicação da Transformada Wavelet de Haar. A separação é feita de forma automática pelo algoritmo.

A Transformada de Haar permite extrair informações da imagem tanto no domínio de tempo como no domínio de frequência, permitindo que sejam analisadas informações globais como também detalhes da imagem. A transformação de uma imagem consiste nas sucessivas passagens de filtros

passas-altas e passas-baixas, decompondo os sinais em detalhes e aproximações. A Fig.2 apresenta como uma imagem é decomposta pela transformada.

Cada imagem é decomposta em 3 direções, horizontal, vertical e diagonal, uma nova imagem chamada de coeficiente de aproximação (CA_1) é gerada. Essas 4 imagens formam o nível de D1 de decomposição da imagem original. Para mais detalhes, basta repetir o processo, usando a imagem CA_1 criando assim o nível de decomposição D2, com mais 4 imagens, D2 Horizontal, D2 Vertical, D2 Diagonal e CA_2 . Cada imagem é decomposta nos níveis D1, D2 e D3 após serem divididas em duas subimagens. A Fig. 3 mostra a imagem original, e as duas subimagens resultantes após a divisão enquanto a Fig. 4 mostra a Fig. 3(b) e a Fig. 3(c) no nível D1 de decomposição.

Nenhum outro tipo de pré-processamento é necessário além de dividir as imagens em duas partes e da decomposição pela Transformada de Haar.

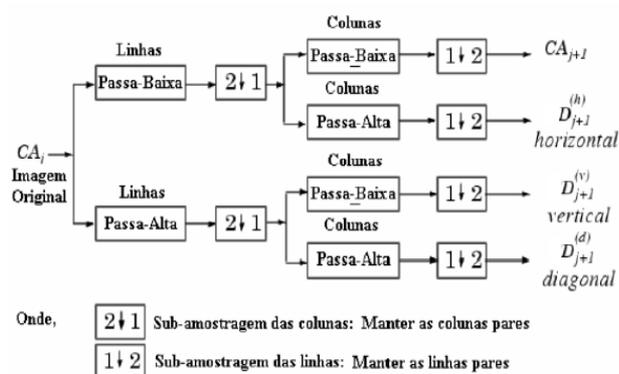


Fig. 2. Funcionamento da Transformada

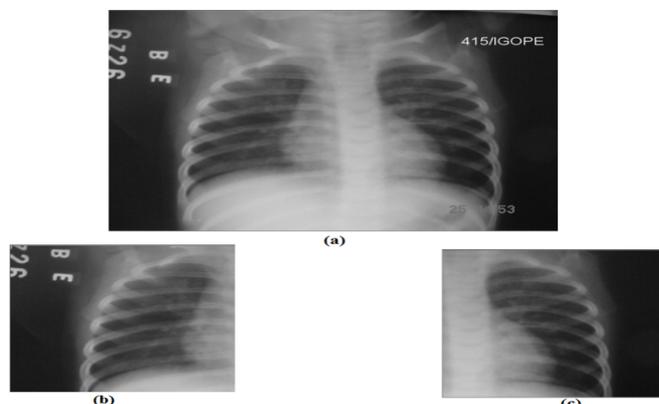


Fig. 3. (a) Imagem original (b) Subimagem do pulmão esquerdo (c) Subimagem do pulmão direito

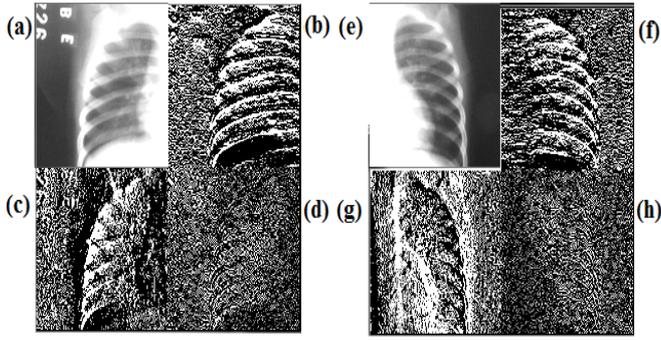


Fig. 4. (a) CA₁ (b) D1 Horizontal (c) D1 Vertical (d) D1 Diagonal referentes ao pulmão esquerdo (e) CA₁ (f) D1 Horizontal (g) D1 Vertical (h) D1 Diagonal referentes ao pulmão direito

Uma vez aplicada a transformada, executam-se as equações matemáticas que produzem informações estatísticas dos coeficientes em cada sub-imagem. As equações matemáticas apresentadas na Tab. I estão baseadas no trabalho de [20].

TABLE I. EQUAÇÕES UTILIZADAS PARA CONSTRUÇÃO DO VETOR DE CARACTERÍSTICAS

Descrição	Equação
Coefficiente de variação	$CV = \frac{\sigma}{\mu}$
Contraste	$Contraste = \sum_i \sum_j (i - j)^2 * p(i,j)$
Correlação	$Correlação = \frac{\sum_i \sum_j (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
Desvio Médio	$DesvMedio = \frac{\sum_i \sum_j p(i,j) - \bar{x} }{N}$
Variância	$V = \frac{\sum_i \sum_j (i,j) - \bar{x} }{N}$
Somatória da variância	$SumVariância = \sum_i (i - SumEntropia)^2 p_{x+y}(i)$
Diferença da variância	$difVariância = V(p_{x-y})$
Entropia	$Entropia = \sum_i \sum_j p(i,j) \log(p(i,j))$
Somatória da entropia	$SumEntropia = - \sum_i p_{x+y}(i) \log(p_{x+y}(i))$
Diferença da entropia	$difEntropia = - \sum_i p_{x-y}(i) \log(p_{x-y}(i))$
Energia	$Energia = \sum_i \sum_j p(i,j)^2$
Energia Média	$EnergiaMedia = \frac{Energia}{N}$
Momento de diferenças inversas	$MDI = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i,j)$
Resíduo médio	$ResMedio = \sum_i \sum_j ((i - j) - \bar{x})^2$
Soma média	$SumMedia = \sum_i i p_{x+y}(i)$
Suavidade	$SV = 1 - \frac{1}{1 + \mu^2}$
Desvio padrão	$\sigma = \sqrt{V}$

Essas informações são indexadas no vetor de características de cada imagem. Cada uma delas terá 18 valores (9 de cada pulmão, direito e esquerdo).

Uma vez que as características sejam armazenadas, é realizada o cálculo da distância de Manhattan, dado pela equação (1) entre cada imagem, isto é, a distância entre cada valor de cada característica. A imagem mais próxima da imagem de consulta é retornada. A distância de Manhattan foi escolhida por ser mais barata computacionalmente, pois usa apenas somas e subtrações para realizar o cálculo.

$$DistManhattan = |x_1 - y_1| + |x_2 - y_2| + \dots + |x_n - y_n| \quad (1)$$

Para cada nova imagem que fosse usada, esse cálculo seria refeito. Ao final a imagem com a menor distância calculada seria a imagem mais semelhante a original.

Para cada imagem na base, serão geradas 18 sub-imagens com 17 características cada uma, gerando um total de 306 valores a serem usados, o que tem um alto custo computacional. Numa base com 1000 imagens, seriam necessários 306 mil comparações. Apenas implementar o CBIR não é eficiente, é preciso aumentar a eficiência da técnica e diminuir os custos computacionais.

Uma maneira de tornar o método mais eficiente é mesclando buscas textuais com as técnicas de CBIR. Segundo [14] buscas textuais produzem resultados melhores do que buscas por similaridade. [14] sugere que combinar as duas produziriam melhores resultados e diminuiriam os custos computacionais. Usando esse conceito como base é criado uma etapa que precede o CBIR com um limiar textual, que reduz o número de imagens nas quais serão aplicados os cálculos para extração de características e cálculo de distância.

Dentro do banco de imagens existem radiografias do pulmão e radiografias de outras partes do corpo, inseridas propositalmente. Comparar essas radiografias seria um desperdício. Um limiar textual que defina o foco apenas para exames do pulmão reduz o escopo de comparações tornando o algoritmo mais eficiente.

Um limiar textual também permite a mudança do algoritmo a ser utilizado, maximizando a taxa de acerto. Desta maneira é possível definir um algoritmo para cada tipo de exame armazenado na base.

III. RESULTADOS

Foram usadas 80 imagens radiográficas e outras 80 imagens aleatórias com o mesmo tamanho, mas conteúdo completamente diferente. Na Fig.5 Temos um exemplo de imagem aleatória. Isto serve para mostrar que o método é capaz de ignorar imagens que não tem relação com a imagem foco do teste, permitindo que outros tipos de imagens sejam adicionados sem comprometer a eficiência, pois não entram na média de tempo. Para testar a eficiência da técnica de CBIR, todas as imagens em algum momento foram usadas como parâmetro de busca. O algoritmo apresentou 100% de acerto na busca da imagem. Em 100% dos casos, a segunda imagem mais semelhante era uma variação da imagem original, porém do mesmo caso/paciente/diagnóstico. Isto é verificado devido as informações associadas a imagem que também são

recuperadas.



Fig. 5. Exemplo de imagem aleatória retirada da internet.

Comparando a utilização do algoritmo aplicado diretamente na base de imagens ou utilizando um limiar textual, obteve-se melhora de desempenho. A etapa de comparação e cálculo de distância entre as imagens recuperadas também teve uma melhora de desempenho. Num total de 400 testes, cada imagem foi verificada cinco vezes, os resultados podem ser vistos na Tab. II.

TABLE II. TEMPOS E TAXA DE ACERTO DA METODOLOGIA

Técnica	Tempo sem Limiar	Tempo com Limiar	Melhora
Recuperação das informações da base	0.035 Segundos	0.015 Segundos	58%
Comparação entre vetores de característica e Exibição das Imagens	0.016 Segundos	0.015 Segundos	7%
Total	0.051 Segundos	0.03 Segundos	41%
Taxa de Acerto	100%	100%	

Os resultados exibidos na Tab. II foram obtidos usando um notebook com processador i7-3630QM de 2.4GHz com 8GB de memória RAM.

A Referência [14], usando um limiar dinâmico, conseguiu aumentos de eficiência de no máximo 22.9% em relação ao melhor tipo de busca, em seu estudo a busca textual simples comparada a busca em dois estágios (limiar + CBIR), usando a base de imagens da Wikipédia. Em nosso trabalho, os tempos de recuperação da metodologia empregada atingiram 41% de melhora no tempo na recuperação.

Dessa maneira, como mencionado, é possível que outros tipos de imagens sejam adicionados na base, sem que as técnicas já implementadas sejam prejudicadas, abrindo um escopo cada vez maior no tipo de imagens suportadas.

IV. CONCLUSÃO

A base de imagens médicas, e a técnica de CBIR implementada para recuperação das mesmas tem como maior contribuição o fato de auxiliar no diagnóstico e tratamento de pessoas que apresentem exames semelhantes aos armazenados na base, uma vez que a comparação visual dos exames feito

pelo CBIR poderá ser útil ao avaliar as condições de cada paciente e o tipo de tratamento a ser utilizado.

Este trabalho abre possibilidades de construção de um sistema de informação CAD que pode ser utilizado por clínicas ou hospitais na busca por casos semelhantes em exames baseados em imagens. Tal sistema de informação ofereceria suporte aos profissionais tornando os diagnósticos mais precisos, pois seria possível identificar rapidamente casos semelhantes com base nas imagens.

Pretende-se ainda implementar outros algoritmos para outros tipos de imagem, onde a técnica usada para recuperação seria escolhidos através de um limiar textual, cobrindo uma quantidade cada vez maior de tipos de imagem sem perder a eficiência da técnica.

REFERENCES

- [1] Müller, H., Michoux, N., Bandon, D., and Geissbuhler, A. (2004). A review of content-based image retrieval systems in medical applications—Clinical benefits and future directions. *International Journal of Medical Informatics*, 73(1), p. 1–23.
- [2] Siadat, M., Soltanian-Zadeh, H., Fotouhi, F., and Elisevich, K. (2005). Content-based image database system for epilepsy. *Computer Methods and Programs in Biomedicine*, 79(3), p. 209–226.
- [3] Kim, J., Cai, W., Feng, D., and Wu, H. (2006). A new way for multidimensional medical data management: Volume of interest (VOI)-based retrieval of medical images with visual and functional features. *IEEE Transactions on Information Technology in Biomedicine*, 10(3), p. 598–606.
- [4] Matheus, B. R. N. and Schiabel, H. (2010). Online Mammographic Images Database for Development and Comparison of CAD Schemes. *Journal of Digital Imaging*, p. 1618-727X.
- [5] Kachouri, R., Djemal, K., Maaref, H., Masmoudi, D.S. and Derbel, N. (2008). Feature extraction and relevance evaluation for heterogeneous image database recognition. *Image Processing Theory, Tools and Applications. IPTA 2008. First Workshops on*. p. 1-6.
- [6] Yuan, K., Tian, Z., Zou, J., Bai, Y. and You, Q. (2011). Brain CT image database building for computer-aided diagnosis using content-based image retrieval. *Information Processing and Management. Vol.47(2)*, p. 176-185.
- [7] Frank H., Olaf K., Horst K. H. and Heinz-Otto P. (2011). Interactive 3D medical image segmentation with energy-minimizing implicit functions. *Visual Computing in Biology and Medicine - Computers & Graphics. Volume 35, Issue 2*, p. 275–287.
- [8] Danzhou L., Hua, K.A. and Sugaya, K. (2006). A Framework for Web-Based Interactive Applications of High-Resolution 3D Medical Image Data. *Computer-Based Medical Systems. CBMS 2006. 19th IEEE International Symposium on. Digital Object Identifier: 10.1109/CBMS.2006.12*. p. 119-124.
- [9] Chen T., Wang Y., Zhang H. and Xiao C. (2010). An embedded 3D medical image processing and visualization platform based on dual-core processor. *Intelligent Control and Automation (WCICA), 8th World Congress on. Digital Object Identifier: 10.1109/WCICA.2010.5554388*, p. 2936-2941.
- [10] Bergamasco, L.C.C., Nunes, F.L.S. (2013). Applying Distance Histogram to retrieve 3D cardiac medical models, In: *The American Medical Informatics Association Annual Symposium*.
- [11] Chen, Z. Q., Zou, K. S., Ip, W. H. and Chan, C. Y. (2011). 3D model retrieval based on fuzzy weighted shape distributions. In: *Advanced Materials Research. [S.l.]: Trans Tech Publications. v. 201-203*, p. 1678–1681.
- [12] Vranic, D. V. (2001). 3D Model Retrieval. Tese (PhD Thesis) — University of Leipezig, Germany.

- [13] Yang, F. and Leng, B. (2007). OFS: A feature selection method for shape-based 3D model retrieval. In: Proceeding of 10th IEEE International Conference on Computer-Aided Design and Computer Graphics. Beijing, China: IEEE Computer Society. p. 114–119.
- [14] Arampatzis, A., Zagoris, K., and Chatzichristofis, S. A. (2011). Dynamic two-stage image retrieval from large multimedia databases. In Information Processing and Management.
- [15] Borges, E. N., Cony, C. A., and de Freitas, A. L. C. (2005). Utilização de algoritmos de busca por similaridade em sgbd's comerciais.
- [16] Muneesawang, P. and L. Guan, M. (2006). Database retrieval: A human-centered approach. In IEEE Computer Society.
- [17] Samet, H. (2010). Techniques for similarity searching in multimedia databases. In The 36th International Conference on Very Large Data Bases.
- [18] Oliveira, L. L. G., e Silva, S. A., Ribeiro, L. H. V., de Oliveira, R. M., Coelho, C. J., and Andrade, A. L. S. S. (2008). Computer-aided diagnosis in chest radiography for detection of childhood pneumonia. In International Journal of Medical Informatics.
- [19] de Macedo, S. O. (2012). Desenvolvimento de um sistema de auxílio ao diagnóstico de pneumonia na infância utilizando visão computacional. Mestrado, Instituto de Informática, Universidade Federal de Goiás.
- [20] Haralick, R., Shanmugam, K and Dinstein, I. (1973). Textural Features for Image Classification. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, vol. SMC-3, NO. 6.

APÊNDICE I

Coefficiente de variação: Tem como foco analisar o peso do desvio-padrão em relação a média dos valores dos pixels das imagens. Dado pela equação:

$$CV = \frac{\sigma}{\mu}$$

Onde σ é o desvio padrão e μ é a média dos coeficientes.

Contraste: O contraste caracteriza-se pela diferença entre os tons de cinza. Baixo contraste ocorre quando há pequena diferença entre os níveis de cinza dos pixels localizados em uma região contígua da imagem. É dada pela equação:

$$Contraste = \sum_i \sum_j (i - j)^2 * p(i, j)$$

Onde $p(i, j)$ é (i,j)-ésima probabilidade de ocorrência de um nível de cinza na imagem.

Correlação: A correlação mede a dependência linear entre os tons de cinza presentes em uma imagem. Valores altos de correlação dados pela equação, próximos de um, indicam a existência de relação linear entre os pares de níveis de cinza.

$$Corrrelação = \frac{\sum_i \sum_j (ij)p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

Onde $p(i, j)$ é (i,j)-ésima probabilidade de ocorrência de um nível de cinza na imagem e $\sigma_x, \sigma_y, \mu_x$ e μ_y são os desvios padrão e médias em x e y respectivamente.

Desvio médio: O Desvio Médio é uma medida da dispersão dos dados em relação à média de uma sequência, o “afastamento” em relação a essa média. Esta medida representa a média das distâncias entre cada elemento da amostra e seu valor médio. Descrito pela equação:

$$DesvMedio = \frac{\sum_i \sum_j |p(i, j) - \bar{x}|}{N}$$

Onde N é o número total de pixels na imagem e \bar{x} é a média dos valores das linhas da matriz que representa a imagem.

Variância: Descreve quanto os valores de tons de cinza estão dispersos em torno da média. É dada pela equação:

$$V = \frac{\sum_i \sum_j |(i, j) - \bar{x}|}{N}$$

Onde V é a variância de $P(i, j)$

Diferença da variância:

$$difVariância = V(p_{x-y})$$

$p_{x-y}(k)$ é $\sum_i \sum_j p(i, j)$ onde $k = |i - j|$

Diferença da entropia:

$$difEntropia = - \sum_i p_{x-y}(i) \log\{p_{x-y}(i)\}$$

Energia: Também conhecida como segundo momento angular, expressa a uniformidade de uma textura, assumindo o valor máximo 1 e é expressa pela equação:

$$Energia = \sum_i \sum_j p(i, j)^2$$

Energia média: É a energia da imagem dividida pelo número total de pixels. Dado pela equação:

$$EnergiaMedia = Energia/N$$

Entropia: A medida de entropia expressa a desordem contida na textura. Quando uma imagem não é uniforme, as entradas apresentam valores próximos de zero o que resulta em valores altos de entropia. É dada pela equação:

$$Entropia = \sum_i \sum_j p(i, j) \log(p(i, j))$$

Momento de diferenças inversas: Quando a concentração dos valores na diagonal da matriz de co-ocorrência for máxima, o Momento da Diferença Inverso (MDI) atinge o valor máximo. É dado pela equação:

$$MDI = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j)$$

Resíduo médio: É a diferença média entre os dados observados e os valores ajustados pelo modelo (calculados). É dado pela equação:

$$ResMedio = \sum_i \sum_j ((i - j) - \bar{x})^2$$

Onde \bar{x} é a média dos valores das linhas da matriz que representa a imagem.

Soma média:

$$SumMedia = \sum_i ip_{x+y}(i)$$

$p_{x+y}(k)$ é $\sum_i \sum_j p(i, j)$ onde $k = i + j$

Somatório da entropia:

$$SumEntropia = - \sum_i p_{x+y}(i) \log(p_{x+y}(i))$$

Somatório da variância:

$$SumVariancia = \sum_i (i - SumEntropia)^2 p_{x+y}(i)$$

Suavidade:

$$SV = 1 - \frac{1}{1 + \mu^2}$$

Desvio padrão: É a raiz quadrada da variância, dado pela equação:

$$\sigma = \sqrt{V}$$