

HBCNB) em relação ao histograma que trata da textura (HDCIG). Nezamabadi-Pour e Kabir (NEZAMABADI-POUR e KABIR, 2004) trabalharam com o valor 0,5 para α_1 e α_2 .

Para mostrar o resultado da aplicação da técnica dos histogramas de cores e de variação da magnitude do gradiente nas imagens da Figura 3.54, as distâncias cruzadas entre elas são apresentadas na Tabela 3.17.

Tabela 3.17 – Distâncias entre as imagens da Figura 3.54.

<i>Distâncias nominais normalizadas</i>	<i>buildings.0004</i>	<i>buildings.0005</i>	<i>food.0006</i>	<i>food.0008</i>
<i>buildings.0004</i>		0,25	0,39	0,39
<i>buildings.0005</i>	0,25		0,39	0,39
<i>food.0006</i>	0,39	0,39		0,19
<i>food.0008</i>	0,39	0,39	0,19	

Pode-se verificar na Tabela 3.17 que os resultados foram conforme o esperado. Entretanto, a discriminação entre as imagens de grupos distintos, *buildings* e *food*, não é tão grande como suposto inicialmente.

3.8.3 O descritor de padrões de bordas locais

A utilização combinada de histogramas de cor e de textura foi a linha seguida por K. Chen e S. Chen para a segmentação de imagens coloridas (CHEN e CHEN, 2002). Eles propuseram a quantização adaptativa do espaço de cores HSV utilizando padrões celulares hexagonais para a construção do histograma de cores da imagem e uma construção denominada por eles de padrões de bordas locais, LEP (*local edge patterns*).

Os padrões de bordas locais seguem o conceito de padrões binários locais, LBP, introduzido por Ojala e outros (OJALA *et alii*, 1996) e apresentado na seção 3.6.4 deste capítulo. Para o cálculo do LEP, primeiramente calcula-se a imagem de bordas da imagem colorida quantizada. Um *pixel* é considerado como pertencente a uma borda se o rótulo de sua cor for diferente do rótulo de algum dos *pixels* pertencentes a sua vizinhança de quatro. Portanto, a imagem de bordas é uma imagem binária em que os *pixels* de borda possuem

valor um e os demais *pixels* possuem valor zero. Uma estrutura elementar 3 x 3 permite calcular o valor do LEP pela soma dos resultados da multiplicação dos valores da imagem de borda por potências de dois associadas aos elementos da estrutura. A Figura 3.58 mostra os passos desses cálculos.

IMAGEM QUANTIZADA			IMAGEM DE BORDAS			POTÊNCIAS DE 2		
6	6	5	0	1	1	1	2	4
6	6	5	0	1	1	8	16	32
6	5	5	1	1	0	64	128	256

$$LEP = 0 \times 1 + 1 \times 2 + 1 \times 4 + 0 \times 8 + 1 \times 16 + 1 \times 32 + 1 \times 64 + 1 \times 128 + 0 \times 256 = 246$$

Figura 3.58 – Cálculo do padrão local de bordas, LEP, de um elemento da imagem.

O histograma dos LEP da imagem é montado fazendo-se o elemento 3 x 3 deslizar sobre a imagem. Para cada posição ocupada pelo elemento, o valor do LEP é calculado e computado no respectivo *bin* do histograma. Como passo final, o histograma é normalizado.

A partir dos histogramas de cor e de LEP de duas imagens, ou de duas regiões, foi definida uma medida de homogeneidade, que é a média ponderada das homogeneidades de cor e de textura, calculadas pela interseção dos respectivos histogramas das imagens, ou regiões. A Equação (3.89) define a homogeneidade entre duas regiões.

$$H(A, B) = 0,6 \times H_c(A, B) + 0,4 \times H_e(A, B)$$

onde :

A, B : imagens comparadas;

(3.89)

$$H_c(A, B) = \sum_{i \in \text{cores}} \min(hc_i(A), hc_i(B))$$

$$H_e(A, B) = \sum_{i=1}^{512} \min(he_i(A), he_i(B))$$

hc_i, he_i : valores dos histogramas de cor e de LEP.

A segmentação da imagem é feita pelo processo de divisão hierárquica seguida da aglomeração de vizinhos semelhantes. O objetivo da divisão hierárquica é obter um conjunto de pequenas regiões com textura razoavelmente homogênea. O objetivo da aglomeração de regiões vizinhas semelhantes é concluir o processo de segmentação, produzindo as diferentes regiões finais com texturas homogêneas.

A divisão hierárquica é um processo iterativo pelo qual uma região da imagem é sub-dividida em quatro regiões de mesmo tamanho segundo um critério de homogeneidade. Para verificar se uma região deve ou não ser dividida, os histogramas de cor e LEP para cada uma das quatro sub-regiões são construídos e são calculados os seis valores de homogeneidade correspondentes às comparações das quatro sub-regiões, duas a duas. A divisão ocorre se a razão entre o maior e o menor valor de homogeneidade, entre os seis calculados, superar um patamar definido empiricamente como 1,1.

A aglomeração é também um processo iterativo que, em cada etapa, seleciona um par de regiões vizinhas para verificar se serão aglutinadas, ou não. As duas regiões são determinadas pelo cálculo do valor *merger importance* (MI), obtido pela Equação (3.90).

$$MI = \frac{1}{\sqrt{p}} \times H$$

onde :

(3.90)

p : número de *pixels* da menor das regiões;

H : homogeneidade entre as duas regiões.

As duas regiões que apresentam o maior valor de MI são aglutinadas. Quando duas regiões são aglutinadas, os histogramas de cor e de LEP da nova região são obtidos pela soma normalizada dos histogramas de cor e de LEP das duas regiões originais. O processo de aglutinação prossegue até que um dos critérios de parada mostrados na Equação (3.91) seja atingido.

$$MIR1 = \frac{MI_{atual}}{MI_{min}} < Y$$

ou

(3.91)

$$MIR2 = \frac{MI_{atual}}{MI_{max}} < Z$$

Y, Z : determinados empiricamente

Os valores MI_{min} e MI_{max} são os valores mínimo e máximo dos valores MI correspondentes às várias etapas de aglutinação. O valor MI_{atual} é o valor de MI da aglutinação mais recente.

Esse processo de segmentação da imagem em regiões homogêneas levou à uma outra técnica, mais elaborada, para a recuperação de imagens semelhantes em bancos de dados baseada na comparação de histogramas de características entre regiões de imagens. Essa técnica apóia-se na segmentação da imagem em regiões homogêneas apresentada anteriormente nesta seção. A determinação da semelhança entre duas imagens é realizada

pela comparação entre as regiões dessas imagens, e não mais pela comparação das imagens inteiras. Isso permite melhores resultados do que os obtidos pelas técnicas anteriormente apresentadas porque, com ela, torna-se possível detectar regiões semelhantes que ocorrem em posições diferentes nas imagens comparadas. O artigo de Y. Cheng e S. Chen apresenta essa técnica em detalhes (CHENG e CHEN, 2003).

A utilização da técnica LEP para recuperação de imagens semelhantes em bancos de dados apresenta algumas dificuldades de ordem prática tendo em vista a grande quantidade de cálculos necessários. As regiões de duas imagens devem ser comparadas duas a duas e, dependendo do número de regiões da imagem de busca, do número de imagens do banco de dados e da quantidade média de regiões das imagens do banco, o processamento envolvido pode tornar-se proibitivo.

Para contornar essa dificuldade, Cheng e Chen (CHENG e CHEN, 2003) definiram duas métricas de distância: a distância, ou dissimilaridade entre duas imagens e entre regiões de duas imagens. A distância entre duas imagens serve como filtro para eliminar candidatos muito diferentes, reduzindo o escopo de comparação pela distância entre regiões, operação mais onerosa.

A distância entre imagens é calculada a partir do histograma de cores e do histograma de LEP, já apresentados nesta seção. Ambos sofreram algumas alterações visando ao melhor funcionamento. O histograma de cores da imagem é calculado quantizando-se uniformemente o espaço RGB em 64 níveis, quatro níveis para cada plano de cor. O padrão local de bordas também sofreu ligeira alteração no cálculo da imagem de bordas e na atribuição dos pesos do elemento 3 x 3. A imagem de bordas deve ser obtida pelo cálculo da magnitude do resultado da aplicação dos detectores de bordas horizontais e verticais de Sobel à componente Y da imagem, no espaço YCrCb. Os *pixels* para os quais a magnitude do gradiente é superior a 100 são considerados *pixels* de borda e têm valor um. Os demais têm valor zero. Os novos pesos dos *pixels* do elemento 3 x 3 são mostrados na Figura 3.59. Pode-se perceber nessa figura que o *pixel* central recebe o peso de 256.

IMAGEM QUANTIZADA	IMAGEM DE BORDAS	POTÊNCIAS DE 2																											
<table> <tr><td>6</td><td>6</td><td>5</td></tr> <tr><td>6</td><td>6</td><td>5</td></tr> <tr><td>6</td><td>5</td><td>5</td></tr> </table>	6	6	5	6	6	5	6	5	5	<table> <tr><td>0</td><td>1</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>0</td></tr> </table>	0	1	1	0	1	1	1	1	0	<table> <tr><td>1</td><td>2</td><td>4</td></tr> <tr><td>8</td><td>256</td><td>16</td></tr> <tr><td>32</td><td>64</td><td>128</td></tr> </table>	1	2	4	8	256	16	32	64	128
6	6	5																											
6	6	5																											
6	5	5																											
0	1	1																											
0	1	1																											
1	1	0																											
1	2	4																											
8	256	16																											
32	64	128																											

$$LEP = 0 \times 1 + 1 \times 2 + 1 \times 4 + 0 \times 8 + 1 \times 16 + 1 \times 32 + 1 \times 64 + 0 \times 128 + 1 \times 256 = 374$$

Figura 3.59 – O cálculo do LEP modificado.

A distância entre histogramas de cor é calculada pela norma L1. A mesma métrica é utilizada para o cálculo da distância entre histogramas LEP. Um novo detalhe foi introduzido: a partição do histograma LEP em dois histogramas LEP condicionais. O primeiro, chamado de histograma LEP de borda, quando o *pixel* central é um *pixel* de borda e o segundo, chamado histograma LEP de não-borda, quando o *pixel* central não é *pixel* de borda. O primeiro caso contabiliza os valores de LEP maiores ou iguais a 256 e o segundo caso contabiliza os valores de LEP menores do que 256. As definições para esses histogramas condicionais são apresentadas na Equação (3.92).

$$he_m^{(e)} = \frac{n_{m+256}}{N^{(e)}}, \quad m = 0, 1, \dots, 255$$

$$he_m^{(ne)} = \frac{n_m}{N - N^{(e)}}, \quad m = 0, 1, \dots, 255$$

onde: (3.92)

n_m : número de *pixels* cujo valor LEP é m ;

N : número total de elementos 3x3; e

$N^{(e)}$: número de elementos com *pixel* central de borda.

O cálculo da distância entre histogramas LEP segue as expressões mostradas na Equação (3.93).

$$d_{he}^{(c)}(A, B) = \frac{1}{2} \{d_{he}^{(e)}(A, B) + d_{he}^{(ne)}(A, B)\}$$

onde :

(3.93)

$$d_{he}^{(e)}(A, B) = \sum_{m=0}^{255} |he_m^{(e)}(A) - he_m^{(e)}(B)|$$

$$d_{he}^{(ne)}(A, B) = \sum_{m=0}^{255} |he_m^{(ne)}(A) - he_m^{(ne)}(B)|$$

Com base nas expressões da distância entre histogramas de cor e de LEP, a distância entre duas imagens é calculada segundo a expressão apresentada na Equação (3.94).

$$D_{imagem}(A, B) = \frac{w_{hc} \times d_{hc}(A, B) + w_{he} \times h_{he}^{(c)}(A, B)}{w_{hc} + w_{he}}$$

onde :

(3.94)

w_{hc}, w_{he} : pesos para ajuste; e

$d_{hc}, h_{he}^{(c)}$: distância entre histogramas de cor e de LEP modificado.

A Equação (3.94) é a expressão da distância entre imagens utilizada como filtro para a eliminação de imagens muito dessemelhantes. As imagens que passarem por essa filtragem são submetidas à comparação entre regiões.

A comparação entre regiões homogêneas é realizada levando-se em consideração seis características regionais: o histograma de cor da região, o histograma LEP da região, a cor média da região, a taxa de bordas da região, a posição do centro da região e a cardinalidade de cores da região. Para cada uma dessas seis características é definida uma métrica de dissimilaridade.

A definição de cada uma das seis características é apresentada de forma simplificada, a seguir. Uma definição formal e detalhada é apresentada por Cheng e Chen (CHENG e CHEN, 2003). O histograma de cor e o histograma LEP já foram abordados. A diferença de cálculo para o histograma LEP de regiões é que o número de *bins* é reduzido de 256 para 32 pelo agrupamento dos *bins* oito a oito. A cor média da região é obtida pela média de cada plano no sistema RGB. A distância entre cores médias é encontrada pela média das diferenças absolutas entre as componentes. A taxa de bordas é calculada pela razão do número de *pixels* de borda pelo número total de *pixels* da região. A distância de taxas de bordas é obtida pela diferença absoluta entre elas. A posição do centro da região é encontrada pela média das coordenadas dos *pixels* da região, normalizadas pelas dimensões da imagem. A distância entre posições centrais é calculada pela distância euclidiana entre elas. A cardinalidade de cores da região é obtida pelo número de cores que qualificam mais de 1% dos *pixels* da região. A distância entre cardinalidade de cores é encontrada pela diferença absoluta das cardinalidades, dividida pela maior delas.

O cálculo da distância entre duas imagens pela comparação de suas regiões é realizado em duas etapas: o casamento de regiões semelhantes e o cálculo da distância das regiões casadas.

O casamento de regiões semelhantes visa determinar a região (ou regiões) da imagem alvo mais semelhante a cada uma das regiões da imagem de busca. Para isso, define-se uma métrica que é a média ponderada das distâncias entre as seguintes características: cor média, taxa de bordas, posição do centro e cardinalidade de cor. Cada região da imagem de busca é comparada com cada região da imagem alvo. O par, ou pares, de regiões que apresentam menor distância são considerados casados. Ressalta-se que uma região da imagem de busca pode casar-se com uma ou mais regiões da imagem alvo.

O cálculo da distância entre as regiões casadas é definido como a média ponderada das distâncias entre as seguintes características: histograma de cor, histograma LEP, taxa de bordas e cardinalidade de cores.

O cálculo da distância total entre as regiões é a média aritmética das distâncias entre todas as regiões casadas.

A distância final entre duas imagens é calculada pela soma da distância entre as imagens e a distância total entre as regiões. A operação de busca por imagens semelhantes em um banco de dados é efetuada pela determinação das imagens alvo que apresentem menor distância total da imagem de busca.

Os resultados dos testes efetuados por Cheng e Chen (CHENG e CHEN, 2003) reportaram resultados médios de 90% de acerto para a imagem mais próxima, o que permite classificar a técnica como bastante efetiva.

3.8.4 A distância normalizada da informação

As técnicas de determinação de semelhança apresentadas anteriormente neste capítulo obedecem a uma mesma seqüência de etapas. A imagem sofre algum tipo de pré-processamento, seguido de uma segmentação opcional. Certas características são extraídas dos segmentos e posteriormente combinadas para formar um conjunto de dados que pretendem descrever a imagem como um todo. Define-se uma métrica no espaço de características capaz de comparar esses dados e produzir, como resposta, um valor pertencente a uma escala, que permite avaliar o grau de semelhança, ou dessemelhança entre duas imagens. Esse paradigma fundamenta-se fortemente na etapa intermediária de extração das características e grande parte das pesquisas dedica-se a desenvolver, testar e comparar esses mecanismos de extração.

Gondra e Heisterkamp propuseram uma abordagem diferente, baseada na determinação de semelhança diretamente a partir do conteúdo das imagens, sem a necessidade da etapa de extração de características (GONDRA e HEISTERKAMP, 2008). Essa proposta escapa dos objetivos deste trabalho porque não utiliza as distribuições de características para sua implementação. No entanto, é apresentada com a finalidade de ilustrar o fato de que o campo de pesquisas nessa área permanece aberto a novas indagações.

A abordagem de Gondra e Heisterkamp (GONDRA e HEISTERKAMP, 2008) fundamenta-se no conceito de complexidade de Kolmogorov. Segundo eles, a distância entre dois segmentos de informação definidos como a seqüência de números que os

expressa, pode ser medida pelo tamanho do menor programa que transforme a primeira seqüência de números na segunda. Toda a informação referente a uma imagem está contida no valor dos *pixels* que a compõe. Essa afirmação, bastante óbvia a princípio, é fundamental para instruir a linha de raciocínio da técnica.

A complexidade de Kolmogorov $K(x)$ de uma imagem x expressa pelo valor de seus *pixels* é definida como o tamanho da menor seqüência binária que efetivamente a descreve. Isso pode ser traduzido como o tamanho do menor programa que, sem qualquer dado de entrada, produz a imagem x como saída. De forma análoga, a complexidade $K(x|y)$, de uma imagem x relativa a outra imagem y é definida como o tamanho do menor programa que quando executado tendo y como entrada, produz x como saída. A informação sobre x contida em y , representada por $I(x : y)$ é definida pela expressão $I(x : y) = K(x) - K(x|y)$. Como pode-se afirmar, segundo Gondra e Heisterkamp (GONDRA e HEISTERKAMP, 2008), que $I(x : y) = I(y : x)$, exceto por uma constante aditiva, deduz-se que $K(x) - K(y) = K(x|y) - K(y|x)$. Esse resultado, seguido de algumas manipulações, sugere uma medida de distância entre imagens que deságua na proposta deles de uma medida de informação normalizada expressa pela Equação (3.95).

$$d(x, y) = \frac{\max\{K(x|y), K(y|x)\}}{\max\{K(x), K(y)\}} \quad (3.95)$$

A proposta de distância de Gondra e Heisterkamp (GONDRA e HEISTERKAMP, 2008) apresenta um inconveniente: a complexidade de Kolmogorov não é computável. Isso significa que não existe algoritmo que seja capaz de calculá-la em um tempo finito. Para contornar essa severa limitação, Gondra e Heisterkamp propuseram uma aproximação para essa distância baseada em algoritmos de compressão de dados.

O objetivo dos algoritmos de compressão de dados é a supressão de redundâncias existentes em sua expressão, produzindo uma representação de menor tamanho dos dados originais. Espera-se que o resultado da compressão de um conjunto de dados mais complexo seja maior do que o obtido a partir de um conjunto de dados menos

complexo. Portanto, existe uma conexão entre o conceito de complexidade de Kolmogorov e os algoritmos de compressão. Assim, os valores de $K(x)$ e $K(y)$ presentes no denominador da Equação (3.95) podem ser aproximados pelo tamanho dos resultados da compressão das respectivas imagens x e y . Cada uma das complexidades condicionais do numerador podem ser aproximadas como $K(x|y) = K(xy) - K(y)$, onde $K(xy)$ representa a complexidade conjunta das duas imagens x e y .

A complexidade conjunta pode ser aproximada pela compressão das imagens x e y em conjunto. Uma primeira abordagem é a de produzir um entrelaçamento dos *pixels* de ambas as imagens. Entretanto, como as técnicas de compressão operam normalmente pela busca de cadeias numéricas semelhantes, isso não é efetivo. A proposta de Gondra e Heisterkamp (GONDRA e HEISTERKAMP, 2008) foi no sentido de particionar as imagens em um conjunto de blocos de mesmo tamanho e testar o resultado da compressão para vários arranjos de entrelaçamento entre eles. Como valor da distância, vale o menor valor obtido para os vários entrelaçamentos.

A técnica proposta por Gondra e Heisterkamp (GONDRA e HEISTERKAMP, 2008) utiliza uma abordagem bastante diferente das linhas de pesquisa mais freqüentemente seguidas. No entanto, os testes comparativos efetuados por eles produziram resultados reportados como significativamente próximos àqueles obtidos pela aplicação de técnicas mais tradicionais.

3.9 Considerações Finais deste Capítulo

Este capítulo apresentou várias técnicas de extração de características de imagens digitais com a finalidade de determinar a semelhança entre elas. Essas técnicas foram organizadas, quando possível, pela ordem cronológica como foram propostas para poder ilustrar o desenvolvimento dos trabalhos nessa área. Essa ordem foi rompida quando houve a necessidade de se enfatizar alguma técnica mais importante como no caso da matriz de co-ocorrência de níveis de cinza, ou quando se considerou mais apropriado mostrar um conjunto de técnicas fundamentalmente semelhantes, embora propostas em momentos distintos, como no caso das técnicas de estimação de energia.

As pesquisas inicialmente foram sobre a extração de características em imagens de satélite visando a automatização da classificação de regiões segundo seu padrão de tonalidades. Esse interesse fundamentou-se na dificuldade de classificação manual de um número de imagens cada vez mais crescente. Dos descritores mais simples até os mais elaborados, um longo percurso foi trilhado na análise de texturas. O aumento de resolução das imagens disponíveis e a possibilidade de manipulação, análise e classificação de imagens médicas impuseram novos desafios aos pesquisadores.

Posteriormente, com o aumento do interesse na adoção da cor como uma característica importante a ser considerada, observou-se um grande aumento no número de trabalhos. Esse aumento coincidiu com a disseminação do uso da internet e com a necessidade, cada vez maior, de agilizar e facilitar as buscas por conteúdo nesse novo meio de intercâmbio de informações. Nesse contexto surgiu a iniciativa de se combinar as técnicas já consolidadas em um padrão de descrição de conteúdo multimídia, o MPEG-7.

No entanto, os desafios dessa área do conhecimento ainda estão longe de serem completamente vencidos. Novos problemas e novas propostas continuam a fomentar os trabalhos e instigar o interesse dos pesquisadores. As soluções propostas, apresentadas neste capítulo, não cobrem exaustivamente as vertentes existentes. Porém, fornecem uma mostra significativa do que já foi feito, possibilitando que novos interessados possam se colocar a par dos conceitos e técnicas fundamentais que servem de base aos trabalhos futuros.

O próximo capítulo trata da implementação e comparação das principais técnicas apresentadas neste capítulo, visando completar este estudo com um trabalho prático. A análise dos resultados obtidos permitirá verificar o âmbito da aplicabilidade de cada uma das técnicas testadas.

4 OS RESULTADOS OBTIDOS

4.1 Introdução

Este capítulo mostra os testes realizados, os resultados obtidos e apresenta as conclusões sobre esses resultados. Primeiramente são mostradas algumas informações contextuais importantes. A seguir, são descritas as bases de imagem utilizadas nos testes e são mostradas as técnicas testadas em cada experimento. O capítulo prossegue com uma abordagem sobre os instrumentos utilizados na medição e na análise dos resultados e com a descrição do equipamento e do *software* utilizados para a execução dos testes. Finalmente, são realizadas conclusões sobre os resultados obtidos.

4.2 Algumas Informações Contextuais

Os testes realizados dividiram-se em duas categorias. O objetivo da primeira categoria é verificar quantitativamente o desempenho das técnicas utilizadas e a finalidade da segunda é avaliar qualitativamente algumas outras técnicas. Para medir quantitativamente o desempenho foi aplicado o paradigma definido no padrão MPEG-7. Segundo esse paradigma, realiza-se a busca por semelhança para cada imagem da base, produzindo-se uma lista de imagens semelhantes composta por todas as demais. Essa lista é classificada em ordem decrescente de similaridade e permite calcular a posição de retorno modificada e normalizada, NMRR. Calcula-se a média de todas as NMRR calculadas, resultando na ANMRR, que é a medida de desempenho utilizada para comparar as técnicas de busca.

4.2.1 As bases de imagens utilizadas nos testes

Os testes de avaliação de semelhança de texturas foram realizados em um banco de imagens composto por um subconjunto de texturas em níveis de cinza do álbum de Brodatz (TEXTURAS BRODATZ, 2008). Esse conjunto é composto de 13 imagens originais de dimensão 512 x 512 *pixels*. As texturas são: *bark* (D12), *brick* (D94), *bubbles* (D112), *grass* (D9), *leather* (D24), *pigskin* (D92), *raffia* (D84), *sand* (D29), *straw* (D15),

water (D38), *weave* (D16), *wood* (D68) e *wool* (D19). O código entre parênteses ao lado das denominações corresponde ao número da página do álbum de Brodatz (BRODATZ, 1968) do qual as texturas foram extraídas. Além dessas imagens originais, cada uma delas foi rotacionadas por ângulos de 30° , 60° , 90° , 120° , 150° e 200° , resultando em sete versões, totalizando um universo de 91 imagens. Para utilizar uma notação mais curta, as imagens originais, não rotacionadas, serão tratadas neste texto por imagens 0° . Cada uma das 91 imagens foi subdividida em imagens menores não sobrepostas e de tamanhos iguais, gerando um conjunto de testes. O conjunto é composto por retalhos de 128×128 *pixels* totalizando $16 \times 91 = 1456$ sub-imagens.

Esse conjunto de imagens é adequado porque seu arranjo determina automaticamente o conjunto-verdade para cada uma das imagens testadas. O conjunto-verdade para as texturas de dimensão 512×512 *pixels* é composto por 13 classes correspondentes a cada uma das texturas originais. Para os conjuntos de sub-imagens de dimensão 128×128 *pixels*, existem duas possibilidades. A primeira possibilidade são os testes em que as técnicas utilizadas não são invariantes à rotação e a segunda são aqueles em que as técnicas usadas são invariantes à rotação. Para as técnicas que não são invariantes à rotação, o conjunto verdade é composto pelos 16 retalhos de dimensão 128×128 *pixels* obtidos de uma mesma imagem original. Para as técnicas que são invariantes à rotação, o conjunto-verdade é composto pelos $16 \times 7 = 112$ retalhos (128×128 *pixels*) extraídos das sete versões rotacionadas da textura original.

A Figura 4.1 mostra as sete rotações da textura *bark*, uma das 13 texturas de Brodatz utilizadas nos testes, e os 16 retalhos 128×128 *pixels* da textura *bark* 200° .

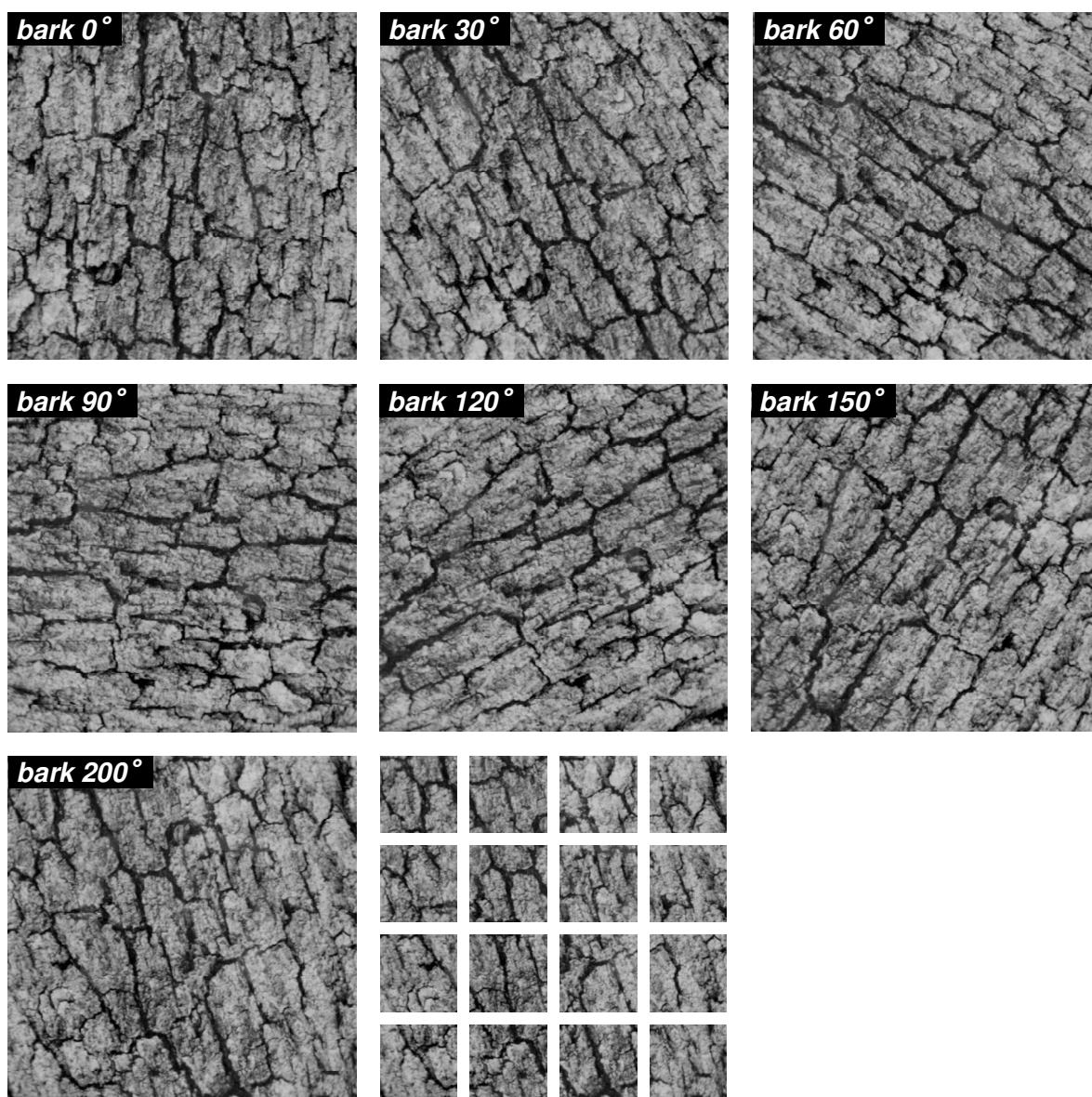


Figura 4.1 –Versões rotacionadas da textura *bark* e os 16 retalhos 128 x 128 *pixels* da versão rotacionada de 200°.

A Figura 4.2 ilustra a imagem original das outras 12 texturas da base de dados utilizada nos testes.

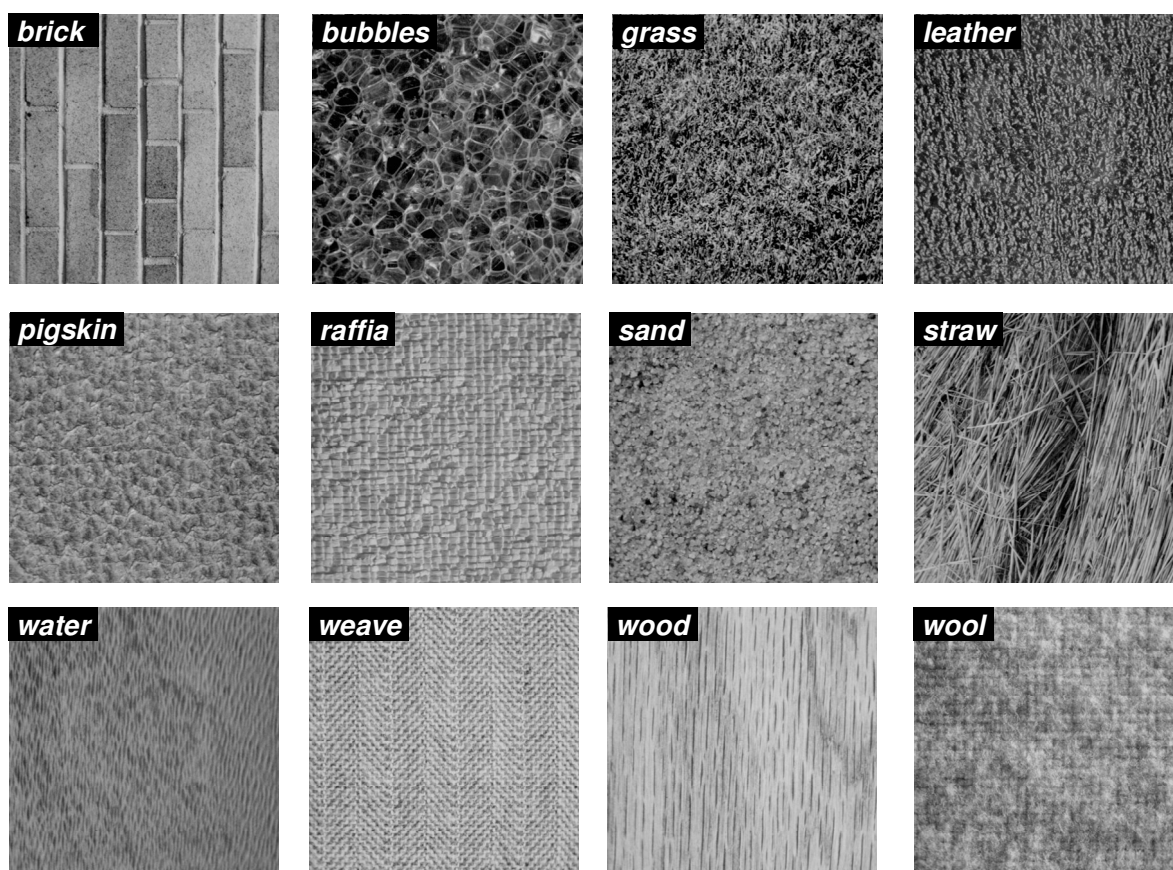


Figura 4.2 – As outras 12 texturas originais de dimensão 512 x 512 *pixels* utilizadas para compor a base de texturas dos testes deste trabalho.

Os testes de avaliação quantitativa dos algoritmos de semelhança por análise de cor foram realizados utilizando um banco de imagens disponibilizado pela Pennsylvania State University, o SIMPLIcity (SIMPLIcity, 2008). Esse banco é composto por 1000 imagens coloridas de dimensão 256 x 384 *pixels* grosseiramente classificadas em 10 classes de imagens: *indígenas*, *praia*, *construção*, *ônibus*, *dinossauro*, *elefante*, *flor*, *cavalo*, *montanha* e *comida*. Cada uma dessas classes possui 100 imagens.

A Figura 4.3 mostra duas amostras de cada uma das 10 classes da base de imagens coloridas da base SIMPLIcity.



Figura 4.3 – Duas amostras de cada classe da base SIMPLIcity: indígenas, praia, construção, ônibus, dinossauro, elefante, flor, cavalo, montanha e comida.

A classificação das imagens da base SIMPLIcity é grosseira porque não particiona o conjunto de imagens em sub-conjuntos disjuntos sob o ponto de vista da semelhança visual. Por exemplo, uma imagem membro da classe *praia*, à primeira vista, pode ser bastante similar a imagens das classes *construção* e *montanha*. Pode-se observar em todo o conjunto de imagens da base SIMPLIcity várias semelhanças interclasse.

Espera-se que o desempenho da busca por semelhança baseados em uma classificação grosseira sejam mais pobres. Esse desempenho é calculado pela medida da posição média de retorno modificada normalizada, ANMRR, definida pelo padrão MPEG-7. Essa medida calcula a posição média de retorno ocupada pelas imagens da mesma classe que a imagem de busca, na sequência de imagens recuperadas. Se existem imagens

semelhantes em classes distintas, as imagens de classe diferente, porém semelhantes à imagem de busca, estarão na seqüência de imagens recuperadas. Eventualmente, as imagens de outra classe ocuparão posições menores (pois são mais semelhantes) que algumas imagens da mesma classe que a da imagem de busca. Como consequência, a posição média de retorno será maior e o valor da ANMRR também será maior.

Os testes da avaliação qualitativa do desempenho utilizaram imagens coloridas da base de imagens VisTex (VisTex, 2008), disponibilizada pelo MediaLab do MIT. Essa base é composta por 167 imagens coloridas de dimensão 128 x 128 *pixels*. A avaliação meramente qualitativa foi feita porque, na ocasião dos testes, o conjunto-verdade para essa base de imagens não estava disponível. Como o número de imagens é razoavelmente grande, os procedimentos para a geração do conjunto-verdade envolveriam um elevado dispêndio de recursos, acima das possibilidades desse trabalho. Mesmo assim, considera-se que a avaliação subjetiva contribui significativamente para os objetivos dessa dissertação.

4.2.2 As técnicas testadas

Neste trabalho, as técnicas testadas quantitativamente para texturas foram: a média e a variância da intensidade, a matriz de co-ocorrência de níveis de cinza (MCNC), os padrões binários locais uniformes (PBLIRU16), o descritor de texturas homogêneas, o descritor de histograma de bordas (EHD), o histograma da direção e magnitude do gradiente e o conjunto de histogramas de cores e da direção da variação da magnitude do gradiente (HDCIG, HUCUB e HBCNB). Essas técnicas foram aplicadas ao conjunto de imagens de Brodatz.

Nesta dissertação, as técnicas testadas quantitativamente para imagens coloridas foram: o descritor de leiaute de cores (CLD), o descritor de estrutura de cores (CSD), o descritor de cores dominantes (DCD), o descritor escalável de cores (SCD) e o conjunto de histogramas de cores e da direção de variação da magnitude do gradiente (HDCIG, HUCUB e HBCNB). Essas técnicas foram aplicadas à base de imagens SIMPLIcity. Essas mesmas técnicas foram analisadas qualitativamente para a base de imagens VisTex.

Nos testes qualitativos foram realizadas análises visuais das sete imagens mais próximas que foram recuperadas para cada imagem de busca.

4.2.3 Os instrumentos de avaliação de desempenho utilizados

Além de utilizar a medida ANMRR na análise dos resultados dos testes quantitativos, foram geradas as matrizes de confusão que indicam numericamente as classificações corretas e incorretas para cada imagem de teste. Essas matrizes são bastante úteis para a análise dos resultados, pois mostram se existe uma tendência nos casos de classificação incorreta e possibilitam o questionamento das razões dessas incorreções.

A ANMRR é uma medida normalizada que apresenta valores no intervalo $[0,1]$. O valor zero para a ANMRR indica uma recuperação perfeita e o valor um indica uma recuperação totalmente falha. Para o cálculo da ANMRR, devem-se definir três parâmetros: o número de categorias em que as imagens são classificadas, NCAT, o número de imagens de cada categoria, NG, que equivale ao número de elementos do conjunto verdade da categoria e o número de posições relevantes na lista de retorno, K. Esse número de posições relevantes indica a posição máxima, na lista de retorno, que uma imagem do conjunto verdade deve ocupar para que seja considerada como acerto. Se uma imagem do conjunto verdade ocupar uma posição maior do que aquela definida por K, considera-se que houve uma falha da busca. A rigor, o valor de K não exerce uma grande influência no valor da ANMRR, por que essa medida é normalizada em relação à posição média de retorno. Os testes executados utilizaram o valor de K igual a 2.

Para avaliar um limite superior mais palpável para a ANMRR, foram realizadas simulações de cálculo para essa medida considerando uma técnica hipotética de busca que produz como lista de retorno uma permutação aleatória do conjunto de imagens do banco de dados. Uma técnica como essa é absolutamente ineficaz e não atende ao propósito de busca de imagens semelhantes. O valor da ANMRR esperado para o retorno de imagens aleatórias é menor do que o limite teórico igual a um. Isso porque o valor um para a ANMRR implica que, em todas as buscas, as imagens do conjunto verdade devem ser retornadas em posições maiores do que K. Entretanto, em uma lista de retorno aleatória, é provável

que algumas imagens do conjunto verdade sejam retornadas em posições significativas, o que reduz o valor da ANMRR.

As simulações do cálculo da ANMRR foram realizadas para quatro situações. Em todas elas, a ANMRR foi calculada como a média de 100 repetições do conjunto de buscas composto pelas buscas por cada imagem da respectiva base. A primeira situação simulada utilizou a base Brodatz 512 x 512 *pixels*. Para essa situação foram utilizados os seguintes dados: número de categorias NCAT = 13, número de imagens por categoria NG = 7 e K = 2. O valor resultante da simulação foi ANMRR = 0,9322. A segunda situação utilizou a base Brodatz 128 x 128 *pixels* para técnicas de busca invariantes à rotação. Para essa situação os dados utilizados foram: NCAT = 13, NG = 112 e K = 2. A simulação resultou em ANMRR = 0,9236. A terceira situação utilizou ainda a base Brodatz 128 x 128 *pixels* para as técnicas de busca que não apresentam invariância à rotação. Para essa situação os dados utilizados foram: NCAT = 91, NG = 16 e K = 2. A simulação teve como resultado ANMRR = 0,9898. A última situação utilizou a base SIMPLIcity de imagens coloridas. Para essa situação os dados utilizados foram: NCAT = 10, NG = 100 e K = 2. A simulação teve como resultado ANMRR = 0,9008.

As simulações foram repetidas para K igual ao número de imagens da base. O objetivo dessa simulação foi avaliar a influência do valor de K no valor da ANMRR. Os resultados dessa simulação foram praticamente iguais aos obtidos nas simulações que utilizaram K = 2, com diferença para mais ou para menos na casa dos milésimos. Os resultados dessas simulações permitem assumir o valor da ANMRR = 0,9 como limite superior prático para o mau desempenho das técnicas. Quanto mais próximo desse limite, pior é a técnica. Se o resultado de um teste utilizando uma técnica de busca resultar em um valor de ANMRR próximo, ou maior do que esse limite (ANMRR = 0,9), isso significa que essa técnica é absolutamente inaceitável por apresentar um desempenho similar, ou pior do que um resultado aleatório.

A outra ferramenta utilizada para a avaliação do resultado dos testes das técnicas de busca são as matrizes de confusão. Elas são apresentadas, neste trabalho em duas versões: as matrizes simples e as matrizes percentuais. Os elementos da matriz

simples expressam quantas imagens pertencentes a uma classe foram retornadas entre as primeiras K posições, nas buscas por imagens pertencentes a uma outra classe. Os elementos da matriz de confusão percentual apresentam o mesmo resultado na forma percentual. Relembrando, K indica o número de posições significativas na lista de imagens retornadas.

Nas matrizes de confusão, o cabeçalho da coluna expressa a classe da imagem de busca e o rótulo da linha expressa a classe da imagem retornada. A Figura 4.4 (a) mostra uma matriz de confusão simples resultante de buscas em uma base de dados em que o número de posições significativas (K) é igual a 200. Nessa matriz, o valor em destaque indica que 2255 imagens da classe *montanha* foram retornadas entre as 200 primeiras, nas buscas por imagens da classe *praia*.

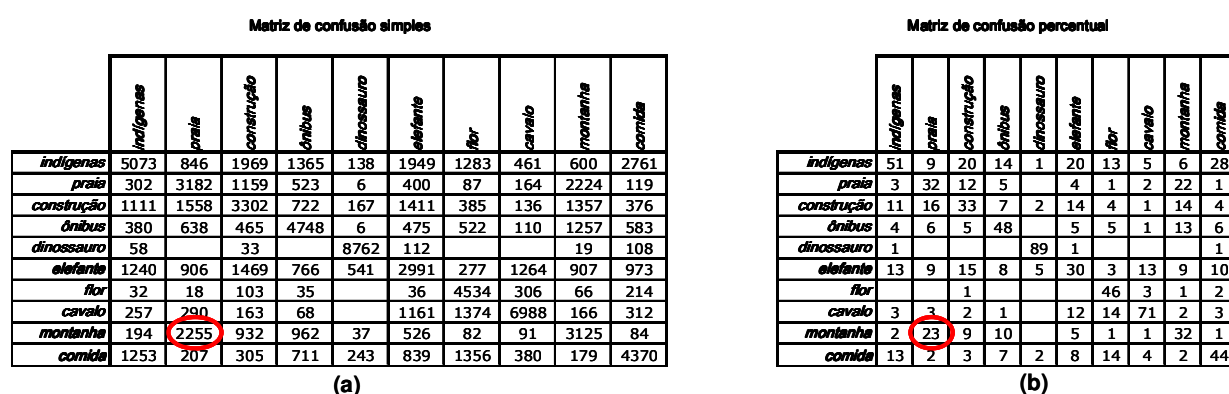


Figura 4.4 – (a) Matriz de confusão simples; (b) matriz de confusão percentual.

A Figura 4.4 (b) apresenta a versão percentual da matriz de confusão mostrada na Figura 4.4 (a). O valor em destaque indica que 23% das imagens significativas (nesse exemplo são 200) retornadas nas buscas por imagens da classe *praia* pertenciam à classe *montanha*.

4.2.4 Outras informações contextuais

O equipamento utilizado nos testes realizados foi um computador *laptop* padrão modelo HP Pavilion dv9500, com processador AMD Turim 64 X2 Mobile Technology TL-58,

clock de 1.9 GHz, com 1.5 GB de memória RAM e disco de 140 GB equipado com o sistema operacional Windows Vista Home Premium. Os programas foram desenvolvidos e testados utilizando Matlab versão 7. Como o objetivo desse trabalho foi o de verificar o desempenho das técnicas e não o desempenho de sua implementação, não foram efetuados testes quantitativos de velocidade de execução. Entretanto, foram apresentadas considerações subjetivas de desempenho das técnicas testadas.

4.3 Os Testes Quantitativos da Base de Texturas de Brodatz

4.3.1 A média e variância da intensidade

A utilização da média e da variância para a identificação de imagens é bastante precária. Esses valores medem a intensidade média dos *pixels* e a variação da intensidade em torno da média, ou contraste, respectivamente.

O vetor de características da média e da variância foi calculado para cada uma das imagens da base. Posteriormente foram calculadas as distâncias cruzadas entre esses vetores utilizando a métrica L1. Duas variações foram testadas: a distância simples e a distância normalizada pelo desvio padrão de cada elemento do vetor de características. A tabela de distâncias cruzadas foi classificada em ordem crescente para cada imagem de busca e o valor da ANMRR foi calculado para o conjunto de todas as buscas.

A técnica da média e da variância é claramente invariante à rotação. Por isso, o conjunto-verdade utilizado para o cálculo da ANMRR é composto das 13 texturas matrizes da base de Brodatz.

4.3.1.1 Base de texturas Brodatz 512 x 512

Para o cálculo da ANMRR foram utilizados: $NG = 6$ e $K = 12$, onde NG é o número de elementos do conjunto-verdade para cada classe e K é o número de posições significativas na lista de retorno.

A Tabela 4.1 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando a técnica da média e da variância com distância calculada pela norma L1, sem normalização. O valor da ANMRR encontrada foi 0,067686.

Tabela 4.1 – Matriz de confusão para Brodatz 512 x 512, utilizando a técnica da média e da variância sem normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	42								3				
<i>brick</i>		37											4
<i>bubbles</i>			33		3								
<i>grass</i>				36									
<i>leather</i>			9	1	39								
<i>pigskin</i>						41							
<i>raffia</i>							41						5
<i>sand</i>		2				1		41					5
<i>straw</i>				5					39				
<i>water</i>										42			
<i>weave</i>											42		
<i>wood</i>												42	
<i>wool</i>		3					1	1					28

A Tabela 4.2 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando a técnica da média e da variância com distância calculada pela norma L1, com normalização. O valor da ANMRR encontrada foi 0,070712.

Tabela 4.2 – Matriz de confusão para Brodatz 512 x 512, utilizando técnica da média e da variância com normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	42								3				
<i>brick</i>		42											
<i>bubbles</i>			29		10								
<i>grass</i>				37									
<i>leather</i>			13		32								
<i>pigskin</i>						42							
<i>raffia</i>							41						4
<i>sand</i>								41					7
<i>straw</i>				5					39				
<i>water</i>										42			
<i>weave</i>											42		
<i>wood</i>												42	
<i>wool</i>							1	1					31

A utilização da normalização pelo desvio das características calculado para as imagens da base mostrou um comportamento interessante. O valor da ANMRR foi ligeiramente menor quando não se utilizou a normalização. No entanto, analisando-se as matrizes de confusão, pode-se verificar que nas situações de pouca confusão (valores mais altos na diagonal principal) o desempenho manteve-se, ou melhorou ao se utilizar a normalização. Contudo, nas situações de mais confusão, o desempenho piorou.

As texturas que apresentaram pior resultado foram: o par *bubbles/leather*, que apresentou confusão recíproca e as texturas *grass*, confundida com *straw*; e *wool*, confundida com *sand* e *raffia*, principalmente. O par *bubbles/leather* apresenta tonalidade mais escura, com algumas raias mais claras. A média e variância são semelhantes. As texturas *wool*, *raffia* e *sand* apresentam variância menor e tonalidade mais clara semelhante.

Outros dois testes foram realizados. No primeiro apenas a média foi utilizada no cálculo da distância. No segundo teste utilizou-se apenas a variância.

A Tabela 4.3 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando apenas a média com distância calculada pela norma L1. O valor da ANMRR encontrada foi 0,19685.

Tabela 4.3 – Matriz de confusão para Brodatz 512 x 512, utilizando apenas a média, sem normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	40								1	6			
<i>brick</i>		24						1					24
<i>bubbles</i>			34		2								
<i>grass</i>				26	7								
<i>leather</i>			8	15	33								
<i>pigskin</i>						30		13					
<i>raffia</i>							42						2
<i>sand</i>		4				12		28					1
<i>straw</i>				1					41				
<i>water</i>	2									36			
<i>weave</i>											42		
<i>wood</i>												42	
<i>wool</i>		14											15

A Tabela 4.4 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando apenas a variância com distância calculada pela norma L1. O valor da ANMRR encontrada foi 0,38653.

Tabela 4.4 – Matriz de confusão para Brodatz 512 x 512, utilizando apenas a variância, sem normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	21		9	3					21				
<i>brick</i>		42											
<i>bubbles</i>	4		13		15				6				
<i>grass</i>				32					1				
<i>leather</i>			14		27								
<i>pigskin</i>						41						5	
<i>raffia</i>							7	8			7		11
<i>sand</i>							7	28					3
<i>straw</i>	17		6	7					14				
<i>water</i>										42			
<i>weave</i>							13	3			20		18
<i>wood</i>						1						37	
<i>wool</i>							15	3			15		10

Pode-se observar nas matrizes de confusão das Tabelas 4.3 e 4.4 que a média consegue distinguir as texturas *bark*, *raffia*, *straw*, *weave* e *wood*. A variância consegue distinguir as texturas *brick*, *pigskin* e *water*.

Comparando-se as Tabelas 4.3 e 4.4 com a Tabela 4.2, pode-se observar também que a variância interfere na classificação de *brick*, *grass*, *pigskin*, *sand*, *water* e *wool*. A média só não interfere na classificação das texturas *brick*, *pigskin* e *water*.

4.3.1.2 Base de texturas Brodatz 128 x 128

Para o cálculo do ANMRR foram utilizados: NG = 111 e K = 222.

A Tabela 4.5 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a técnica da média e da variância com distância calculada pela norma L1, sem normalização. O valor da ANMRR encontrada foi 0,5133.

Tabela 4.5 - Matriz de confusão para Brodatz 128 x 128, técnica da média e da variância sem normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	6109	44	59	1605	101			45	3470				
<i>brick</i>	457	3338				1109	1408	2241	922	68	239	5	1301
<i>bubbles</i>	63		5333	1190	4447				1002				
<i>grass</i>	1591		709	5858	702			3	1933				
<i>leather</i>	78		5292	1483	6288				1046				
<i>pigskin</i>		1153				5704	66	2327	6	2053			1942
<i>raffia</i>	40	2231				49	6458	488	135		694	32	2992
<i>sand</i>	101	2990				2651	883	4539	545	169			2642
<i>straw</i>	3989	369	1039	2296	894		40	140	3242				14
<i>water</i>		8				584		76		9917			94
<i>weave</i>		340					414				8360	4251	110
<i>wood</i>		14					31				3031	8131	21
<i>wool</i>	4	1945				2335	3132	2573	131	225	108	13	3316

A Tabela 4.6 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a técnica da média e da variância com distância calculada pela norma L1, com normalização. O valor da ANMRR encontrada foi 0,46774.

Tabela 4.6 - Matriz de confusão para Brodatz 128 x 128, técnica da média e da variância com normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	6392	38	100	1769	126				3673				
<i>brick</i>	65	3988				698	1247	2610	834	17	258	2	1038
<i>bubbles</i>	120		5464	642	4697				976				
<i>grass</i>	1863		455	6804	451				1976				
<i>leather</i>	167		5400	762	6371				964				
<i>pigskin</i>		622				6782	37	1758		1439			1871
<i>raffia</i>		2190				26	6623	542	111		732	64	2924
<i>sand</i>		3292				2124	1023	4852	198	7			2692
<i>straw</i>	3825	450	1013	2455	787		26	68	3689				
<i>water</i>						196				10946			
<i>weave</i>		485					346				9300	3252	58
<i>wood</i>		5					87				2127	9075	93
<i>wool</i>		1362				2606	3043	2602	11	23	15	39	3756

Para a base de retalhos de 128 x 128 *pixels*, o resultado inverso foi obtido. A utilização da normalização resultou em ANMRR menor, portanto em um desempenho

melhor. As principais confusões ocorreram com os pares *bark/straw*, *brick/sand*, *bubbles/leather*, e com o conjunto *wool/pigskin/raffia/sand*. O caso do par *bubbles/leather* repetiu a confusão das imagens de 512 x 512 *pixels*. As texturas *bark* e *straw* apresentam uma variação regional visível da tonalidade média. O conjunto *wool/pigskin/raffia/sand* apresenta tonalidades médias próximas e variância pequena. O par *brick/sand* apresenta tonalidade média próxima e variância pequena. O que distingue visualmente as duas texturas são as raias claras entre os tijolos de *brick* e a granularidade localizada mais grossa de *sand*. Entretanto, nenhuma dessas características é capturada pela técnica da média e da variância.

Outros dois testes foram realizados. No primeiro apenas a média foi utilizada no cálculo da distância. No segundo teste utilizou-se apenas a variância.

A Tabela 4.7 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando apenas a média com distância calculada pela norma L1. O valor da ANMRR encontrada foi 0,66897.

Tabela 4.7 – Matriz de confusão para Brodatz 128 x 128, utilizando apenas a média, sem normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	3016	626	26	1246	58	1182	59	968	1788	2209			494
<i>brick</i>	630	2163				1959	1964	1808	625	669	258	16	2032
<i>bubbles</i>	37		4714	2167	3370				1073				
<i>grass</i>	1403		2475	3841	3106			4	1769	66			
<i>leather</i>	97		4092	3287	4765				1250				
<i>pigskin</i>	1247	1978		6		2624	313	2446	1056	1683			1428
<i>raffia</i>	49	2069				244	5874	492	102	26	663	23	2863
<i>sand</i>	991	1925		13		2413	654	3022	830	1128			1977
<i>straw</i>	1963	637	1125	1696	1133	1031	110	787	1562	1778			504
<i>water</i>	2497	661		174		1647	24	1100	1863	4257			599
<i>weave</i>		263					478				7494	5114	148
<i>wood</i>		26					17				3816	7270	5
<i>wool</i>	502	2084		2		1332	2939	1805	514	616	201	9	2382

A Tabela 4.8 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando apenas a variância com distância calculada pela norma L1. O valor da ANMRR encontrada foi 0,68672.

Tabela 4.8 – Matriz de confusão para Brodatz 128 x 128, utilizando apenas a variância, sem normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	3351	3	2746	2402	2077				2467				
<i>brick</i>		4332	116		201	198	1339	2342	343		1362	38	673
<i>bubbles</i>	2736	198	2953	1196	3061			6	2783				
<i>grass</i>	1942	3	976	6038	818				1416				
<i>leather</i>	1995	357	2971	963	3508			15	2870				
<i>pigskin</i>		200				3154	1650	820		608	996	3291	2309
<i>raffia</i>		1597			8	1731	2236	2022	29		2273	709	2191
<i>sand</i>		2810	10		18	845	2043	2690	65		2527	299	1538
<i>straw</i>	2408	400	2660	1833	2741		9	36	2451				
<i>water</i>						258				8926		1562	13
<i>weave</i>		1676				982	2375	2689	4		3126	119	2040
<i>wood</i>		32				2878	593	247		2820	110	4922	1231
<i>wool</i>		824				2386	2187	1565	4	78	2038	1492	2437

Pode-se observar nas matrizes de confusão das Tabelas 4.7 e 4.8 que a média consegue distinguir as texturas *raffia*, *weave* e *wood*. A variância consegue distinguir as texturas *grass* e *water*.

Comparando-se as Tabelas 4.7 e 4.8 com a Tabela 4.6, pode-se verificar também que a variância interfere mais na classificação de *bark*, *grass*, *pigskin* e *water*. A média interfere mais na classificação das texturas *bark*, *bubbles*, *leather*, *pigskin*, *raffia*, *weave* e *wood*.

4.3.2 A matriz de co-ocorrência de níveis de cinza (MCNC)

Para cada imagem da base de texturas foram calculadas quatro matrizes de co-ocorrência de níveis de cinza com distância $d = 1$ *pixel* e direções determinadas pelos ângulos: 0° , 45° , 90° e 135° . A partir das MCNC, foram calculadas 11 características: de f_1 a f_{11} . Para tornar o método invariante à rotação, calculou-se a média de cada uma das características para as quatro MCNC, obtendo-se um vetor contendo 11 características, a saber: energia (f_1), contraste (f_2), correlação (f_3), variância (f_4), homogeneidade (f_5), média da soma (f_6), variância da soma (f_7), entropia da soma (f_8), entropia (f_9), variância da diferença (f_{10}), entropia da diferença (f_{11}).

O vetor de características de cada imagem foi normalizado subtraindo-se, de cada característica, a média e dividindo-se o resultado pelo desvio padrão. A média e desvio padrão de cada característica foram calculados sobre todas as imagens da base de teste. Essa normalização foi utilizada para que o peso de cada característica no cálculo das distâncias fosse o mesmo, já que os valores das características são muito variáveis. Por exemplo, os valores da energia (f_1) pertencem à faixa $[0,1]$, enquanto que os valores do contraste (f_2) podem variar no intervalo $[0,256^2]$ e, para as imagens testadas, os valores da entropia (f_9) variam aproximadamente entre oito e nove.

Foram calculadas as distâncias cruzadas L1 com normalização para todos os pares de imagens. Para cada imagem de busca, ordenou-se o vetor de distâncias por ordem crescente e o valor da ANMRR foi calculado considerando-se o número de classes igual a 13.

4.3.2.1 Base de texturas Brodatz 512 x 512

Para o cálculo do ANMRR foram utilizados: $NG = 6$ e $K = 12$.

A Tabela 4.9 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando a técnica MCNC com distância calculada pela norma L1, com normalização. O valor da ANMRR encontrada foi de 0,03233.

Tabela 4.9 – Matriz de confusão para Brodatz 512 x 512, utilizando a técnica da MCNC com normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	42		6						2				
<i>brick</i>		42											
<i>bubbles</i>			36										
<i>grass</i>				30									
<i>leather</i>				12	42								
<i>pigskin</i>						42							
<i>raffia</i>							42						1
<i>sand</i>								42					
<i>straw</i>									40				
<i>water</i>										42			
<i>weave</i>											42		
<i>wood</i>												42	
<i>wool</i>													41

Três outros testes foram realizados. O primeiro deles considerou apenas as nove primeiras características (f_1 a f_9), o segundo teste considerou as cinco primeiras características (f_1 a f_5) e o último considerou apenas as três primeiras características (f_1 a f_3).

Para o teste em que apenas as nove primeiras características (f_1 a f_9) foram utilizadas, o valor da ANMRR calculada foi de 0,03233.

Para o teste em que apenas as cinco primeiras características (f_1 a f_5) foram utilizadas, o valor da ANMRR calculada foi de 0,031056.

Para o teste em que apenas as três primeiras características (f_1 a f_3) foram utilizadas, o valor da ANMRR calculada foi de 0,10893.

A pequena melhoria obtida com a utilização de apenas as cinco primeiras características (f_1 a f_5) não é significativa. O seu resultado na matriz de confusão foi apenas remover a única confusão da textura *wool*, que passou a ter índice de acerto de 100%.

Os resultados das buscas foram bastante satisfatórios.

4.3.2.2 Base de texturas Brodatz 128 x 128

Para o cálculo do ANMRR foram utilizados: NG = 111 e K = 222.

A Tabela 4.10 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a técnica MCNC com distância calculada pela métrica L1, com normalização. O valor da ANMRR encontrada foi de 0,33548.

Tabela 4.10 - Matriz de confusão para Brodatz 128 x 128, utilizando a técnica da MCNC com normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	9189	27	3756	12	15				2103				
<i>brick</i>	10	4637	124		2	443	1208	1797	45	47	11	445	400
<i>bubbles</i>	2322	48	8243	100	20			1	410				
<i>grass</i>	4		14	8173	1703				790				
<i>leather</i>				3385	8694				2655				
<i>pigskin</i>		892				6462	698	1801	3	1362	167	83	4055
<i>raffia</i>		3447	6			566	7265	938	11	3	15	378	1694
<i>sand</i>		2478	27		3	1284	1360	6969	97	115	329	73	951
<i>straw</i>	907	19	257	762	1993		5	22	6188		19		2
<i>water</i>		6				239		2		10559		704	
<i>weave</i>					2	4	2	45	96		11229	119	165
<i>wood</i>		261				10				303		10629	
<i>wool</i>		617	5			3424	1894	857	34	43	662	1	5165

Três outros testes foram realizados. O primeiro deles considerou apenas as nove primeiras características (f_1 a f_9), o segundo considerou as cinco primeiras características (f_1 a f_5) e o último considerou apenas as três primeiras características (f_1 a f_3).

Para o teste em que apenas as nove primeiras características (f_1 a f_9) foram utilizadas, o valor da ANMRR calculada foi de 0,33536.

Para o teste em que apenas as cinco primeiras características (f_1 a f_5) foram utilizadas, o valor da ANMRR calculada foi de 0,38515.

Para o teste em que apenas as três primeiras características (f_1 a f_3) foram utilizadas, o valor da ANMRR calculada foi de 0,50554. Com a finalidade de permitir a comparação, a Tabela 4.11 mostra a tabela de confusão para esse caso.

Tabela 4.11 - Matriz de confusão para Brodatz 128 x 128, utilizando a técnica da MCNC com normalização, considerando-se apenas as características f_1 a f_3 .

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	7939	295	5602						819				
<i>brick</i>	207	4782	754			453	1557	846	58	47	32	134	1041
<i>bubbles</i>	4033	1348	5985	19			17	4	106				7
<i>grass</i>			2	7040	1689				747				
<i>leather</i>				4383	7765				3007		14		
<i>pigskin</i>		669				3764	2170	2155	30	1072	544	1431	3011
<i>raffia</i>	49	2820	48			2439	5009	1233	33	76	1	398	3004
<i>sand</i>	44	862	27	4	6	1765	796	4334	436	381	1708	794	1867
<i>straw</i>	155	2	8	972	2811			16	6585		38		2
<i>water</i>		5				190	1	76		7244	315	2532	43
<i>weave</i>		22		14	160	413	1	1177	500	534	8587	1424	310
<i>wood</i>		16				265	1	214	16	2669	753	4992	80
<i>wool</i>	5	1611	6		1	3143	2880	2377	95	409	440	727	3067

Pode-se observar na Tabela 4.10 que os piores desempenhos ocorreram para as texturas *brick* e *wool*. Uma explicação possível é a grande variação nos padrões de *pixels* entre os retalhos. Por exemplo, a textura *brick* apresenta alguns tijolos com maior contraste e outros mais homogêneos. Além disso, o rejunte entre os tijolos varia muito, ora apresentando sombras mais escuras e largas, ora apresentando sombras mais claras e estreitas. A textura *wool* apresenta comportamento análogo. Isso faz com que varie bastante a distribuição dos valores elevados pelos elementos da MCNC entre os retalhos.

Pode-se observar na Tabela 4.11 que os desempenhos pioraram sensivelmente e, além das texturas *brick* e *wool*, as texturas *pigskin* e *sand* apresentaram o pior desempenho.

4.3.3 Os padrões binários locais uniformes (PBLIRU16)

Os testes com os padrões binários invariantes à rotação e uniformes na vizinhança 16 (PBLIRU16) iniciaram-se com o cálculo do vetor de características que é um histograma normalizado de 18 *bins*. Da mesma forma que os experimentos descritos anteriormente neste capítulo, as distâncias cruzadas de todos os pares de imagem da base de testes foram calculadas utilizando-se a métrica L1, sem e com normalização pelo desvio

padrão das características na base de testes. Para cada imagem de busca, o vetor de distâncias foi classificado em ordem crescente e a ANMRR foi calculada.

4.3.3.1 Base de texturas Brodatz 512 x 512

Para o cálculo do ANMRR foram utilizados: $NG = 6$ e $K = 12$.

A Tabela 4.12 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando a técnica da PBLIRU16 com distância calculada pela métrica L1, sem normalização. O valor da ANMRR encontrada foi de 0,039975.

Tabela 4.12 – Matriz de confusão para Brodatz 512 x 512, utilizando a técnica da PBLIRU16 sem normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	42												
<i>brick</i>		42											
<i>bubbles</i>			42										
<i>grass</i>				42									
<i>leather</i>					42								
<i>pigskin</i>						42							
<i>raffia</i>							42						
<i>sand</i>								42					
<i>straw</i>									35			9	
<i>water</i>										37		3	
<i>weave</i>											42		
<i>wood</i>									7	5		30	
<i>wool</i>													42

A Tabela 4.13 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando a técnica da PBLIRU16 com distância calculada pela métrica L1, com normalização. O valor da ANMRR encontrada foi de 0,0087594.

Tabela 4.13 – Matriz de confusão para Brodatz 512 x 512, utilizando a técnica da PBLIRU16 com normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	42												
<i>brick</i>		42											
<i>bubbles</i>			42										
<i>grass</i>				42									
<i>leather</i>					42								
<i>pigskin</i>						42							
<i>raffia</i>							42						
<i>sand</i>								42					
<i>straw</i>									40	1		3	
<i>water</i>										41			
<i>weave</i>											42		
<i>wood</i>									2			39	
<i>wool</i>													42

O desempenho resultante da utilização de características normalizadas foi excelente, como pode ser observado na Tabela 4.13

4.3.3.2 Base de texturas Brodatz 128 x 128

Para o cálculo do ANMRR foram utilizados: NG = 111 e K = 222.

A Tabela 4.14 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a técnica PBLIRU16 com distância calculada pela métrica L1, sem normalização. O valor da ANMRR encontrada foi de 0,32559.

Tabela 4.14 - Matriz de confusão para Brodatz 128 x 128, utilizando a técnica da PBLIRU16 sem normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	6490	290	267	116	1612	2535	230	1199			397		345
<i>brick</i>		8166	664	1	187				41	62		273	77
<i>bubbles</i>		482	10765		18		18		10	50			
<i>grass</i>	9	42	17	5396	576	608		32					3677
<i>leather</i>	861	1723	575	615	8147	248			1	1		1	608
<i>pigskin</i>	3131	64	79	1679	685	5990	41	1710			402		1484
<i>raffia</i>	5	1	22				11915			8	108		
<i>sand</i>	1350		7	160		1647	3	8001			1681		173
<i>straw</i>		27	1						7599	1130		2114	
<i>water</i>		112	1						1371	8472		2192	
<i>weave</i>	377					298	225	1312			9844		
<i>wood</i>		348							3410	2709		7852	
<i>wool</i>	209	1177	34	4465	1207	1106		178					6068

A Tabela 4.15 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a técnica PLBIRU16 com distância calculada pela métrica L1, com normalização. O valor da ANMRR encontrada foi de 0,25721.

Tabela 4.15 - Matriz de confusão para Brodatz 128 x 128, utilizando a técnica da PBLIRU16 com normalização.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	7242	199	287	42	1598	2330	174	1284			397		286
<i>brick</i>		9632	323		42				28	30		246	110
<i>bubbles</i>		188	11121		7		32		1	103		5	
<i>grass</i>	1		45	6353	151	260		7	1	1			2080
<i>leather</i>	804	690	541	375	9217	228	9		7	10		7	476
<i>pigskin</i>	2573	17	2	1629	648	6987	8	1282			425		1783
<i>raffia</i>	1		60				12096			92	17		
<i>sand</i>	1208			111	3	1277		8510			1263		343
<i>straw</i>		54	5						8143	1606		1231	
<i>water</i>		22	17						1424	8762		870	
<i>weave</i>	466					276	113	1085			10322		2
<i>wood</i>		409	2						2827	1828		10073	
<i>wool</i>	137	1221	29	3922	766	1074		264	1		8		7352

Pode-se observar nas Tabelas 4.14 e 4.15 que o desempenho deste método foi muito bom. Melhor ainda quando se aplica a normalização às características. Nesse caso

apenas a textura *grass* apresentou uma confusão mais elevada, formando com a textura *wool* um par de confusão visível. Outro par de confusão foi o *bark/pigskin*. O par *grass/wool* apresenta um padrão de variação local muito semelhante, que a diferença de contrastes mascara quando da observação visual dos retalhos. Como o contraste não é relevante para a técnica PBLIRU16, essa é, certamente, a razão da confusão. A confusão do par *bark/pigskin* é menor, mas argumento análogo pode ser utilizado na análise. A configuração das depressões de ambas é aproximada formando um padrão reticulado irregular que é detectável pela técnica PBLIRU16. A diferença está na largura dos sulcos, que faz com que a confusão seja menor.

4.3.3.3 Teste combinando PBLIRU16 com MCNC e com a Média e a Variância

Com o objetivo de verificar a influência de outras características adicionais no histograma de padrões binários locais uniformes, foram realizados três testes adicionais. No primeiro, adicionaram-se ao vetor de características do PBLIRU16, a média e a variância da intensidade de nível de cinza das imagens. No segundo, adicionaram-se ao vetor de características do PBLIRU16, as onze características extraídas da MCNC e no terceiro teste, as características dos três métodos foram combinadas em um único vetor. Para uniformizar a contribuição das características dos três métodos, foi utilizada a métrica L1 normalizada pelos desvios padrão.

A Tabela 4.16 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a combinação das técnicas PBLIRU16 e a média e a variância com distância calculada pela métrica L1. O valor da ANMRR encontrada foi de 0,18076.

Tabela 4.16 - Matriz de confusão para Brodatz 128 x 128, utilizando a combinação de PBLIRU16 e da média e da variância.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	8990	92	238	264	1627	663	22	415			10		31
<i>brick</i>		9911	29		6				60	24		265	689
<i>bubbles</i>	3	103	11170	11	25				76	46			
<i>grass</i>	3		61	8392	337	4			7				355
<i>leather</i>	1009	84	919	1366	10155	35			30				119
<i>pigskin</i>	1344	87		829	106	8067	52	1799			303		2540
<i>raffia</i>	1						12241			93	19		
<i>sand</i>	957			68	2	1984		9244			882		468
<i>straw</i>		6	15						9549	1094		151	
<i>water</i>		14							1361	9813		559	
<i>weave</i>	104					141	117	662			11202		4
<i>wood</i>		251							1348	1362		11457	
<i>wool</i>	21	1884		1502	174	1538		312	1		16		8226

A Tabela 4.17 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a combinação das técnicas PBLIRU16 e MCNC com a distância calculada pela métrica L1. O valor da ANMRR encontrada foi de 0,13072.

Tabela 4.17 - Matriz de confusão para Brodatz 128 x 128, utilizando a combinação de PBLIRU16 e MCNC.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	11489		1569	307	697			21	16				
<i>brick</i>		10055	14						61	2		162	481
<i>bubbles</i>	73	38	10763	29					304				
<i>grass</i>	17		27	9084	466				4				
<i>leather</i>	499		2	3002	11265				429				
<i>pigskin</i>	88	110		4		8674	114	1919	2		2		3250
<i>raffia</i>	1	2	3			1	12242			12	10		
<i>sand</i>	252			5	4	2118	30	10041			649		322
<i>straw</i>		5	50						11373	37			
<i>water</i>									143	10445		798	
<i>weave</i>	4					25	46	372			11771		
<i>wood</i>		121							83	1936		11472	
<i>wool</i>	9	2101	4	1		1614		79	17				8379

A Tabela 4.18 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a combinação das técnicas PBLIRU16, MCNC e da média e da variância com a distância calculada pela métrica L1. O valor da ANMRR encontrada foi de 0,12295.

Tabela 4.18 - Matriz de confusão para Brodatz 128 x 128, utilizando a combinação de PBLIRU16, MCNC e da média e da variância.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	11608		1526	311	631			24	32				
<i>brick</i>		10041	1						94	2		179	473
<i>bubbles</i>	83	18	10773	37					346				
<i>grass</i>	22		36	9151	503				7				
<i>leather</i>	440		15	2932	11295				489				
<i>pigskin</i>	64	134				8770	102	2016	4				3230
<i>raffia</i>	1	5	1			1	12245	2	1	11	10		
<i>sand</i>	204	2			3	2101	44	10128			375		356
<i>straw</i>			79						11322	31			
<i>water</i>									92	10948		450	
<i>weave</i>						6	40	163			12047		
<i>wood</i>		105							23	1440		11803	
<i>wool</i>	10	2127	1	1		1554	1	99	22				8373

A Tabela 4.19 mostra o resumo dos resultados obtidos nos testes realizados utilizando a média e a variância, MCNC, PBLIRU16 e suas combinações.

Tabela 4.19 – Resumo dos resultados obtidos nos testes realizados com a média e a variância, MCNC e PBLIRU16 e suas combinações.

Teste	ANMRR (melhor caso)
Média e Variância (MedVar)	0,468
MCNC	0,335
PBLIRU16	0,257
PBLIRU16 + MedVar	0,181
PBLIRU16 + MCNC	0,131
PBLIRU16 + MCNC + MedVar	0,123

Pode-se observar na Tabela 4.19 que a combinação de características de técnicas distintas teve uma melhoria sensível no desempenho. O melhor resultado foi obtido com a combinação das características das três técnicas. Essa melhoria pode ser verificada pela análise da matriz de confusão da Tabela 4.18. O pior caso ocorreu com a textura *wool*, que foi bastante confundida com a textura *pigskin*. Mesmo nesse caso, quando a imagem de

busca pertencia à categoria *wool*, considerando-se as primeiras 111 imagens retornadas 67,4% delas pertenciam também à categoria *wool*. O melhor caso ocorreu para a categoria *weave*, com uma taxa de acerto de 96,9%. Nos testes realizados, apenas três categorias apresentaram índice de acerto inferior a 80%.

4.3.4 O histograma da direção e magnitude do gradiente

Vários testes foram realizados aplicando-se o histograma de direção e magnitude do gradiente. Nesses testes foram considerados o comprimento do histograma, a métrica utilizada e a normalização.

O histograma de direção com 360 *bins* foi calculado para cada imagem. Posteriormente, cada histograma foi suavizado, foi calculada sua FFT e foram armazenados os valores absolutos dos coeficientes.

As distâncias cruzadas foram calculadas e, para cada imagem de busca, o vetor de distâncias foi classificado em ordem crescente e finalmente foi calculado o valor da ANMRR.

Os testes comparativos realizados utilizaram as métricas L1, L2 e EMD, sem e com normalização pelo desvio da base de testes. Os melhores resultados obtidos foram utilizando-se a métrica L2, sem normalização.

Também foram realizados testes comparativos variando-se o número de coeficientes utilizados como características no intervalo de 1 a 90 para ambas as bases de busca: Brodatz 512 x 512 e Brodatz 128 x 128. A Figura 4.5 apresenta o gráfico de variação da ANMRR com o número de coeficientes para a base de busca Brodatz 512 x 512 e a Figura 4.6 apresenta esse gráfico para a base de busca Brodatz 128 x 128.

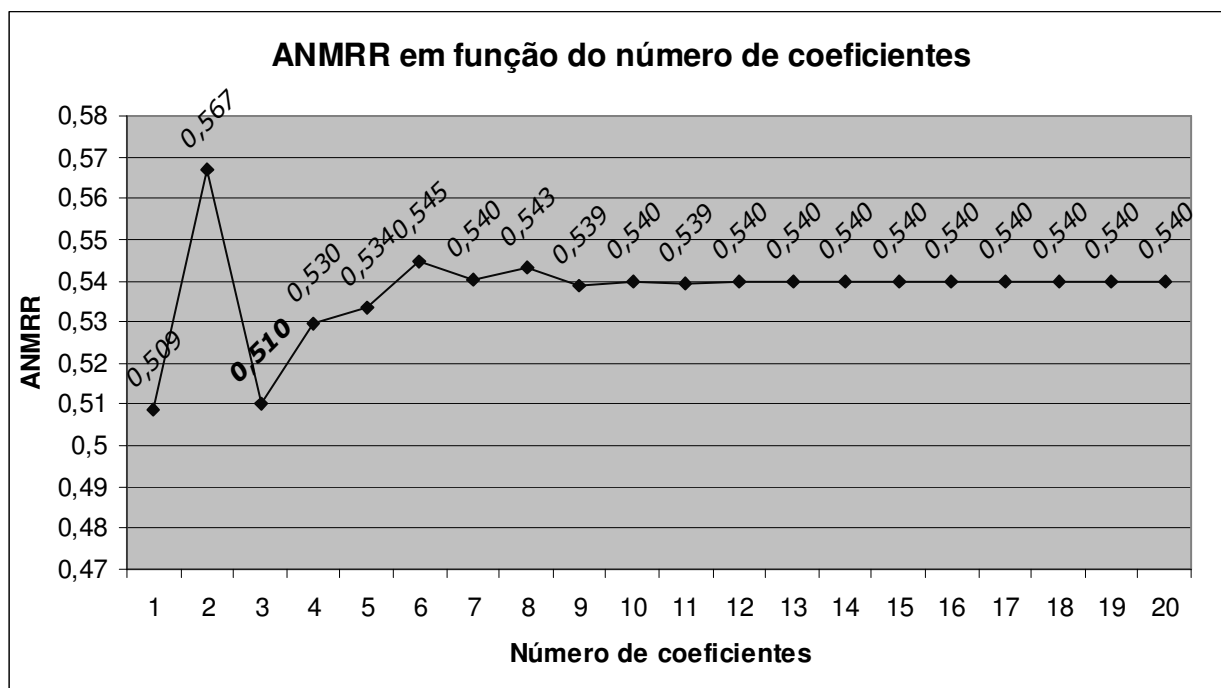


Figura 4.5 - Gráfico de variação da ANMRR em função do número de coeficientes para a base de busca Brodatz 512 x 512.

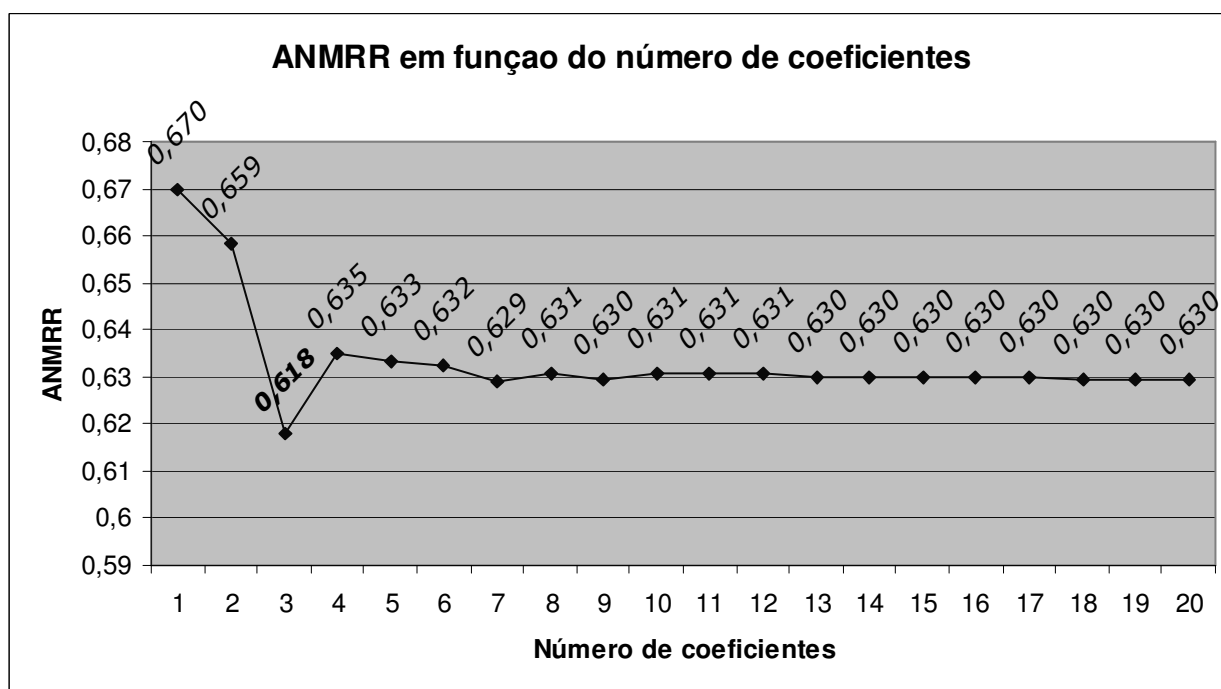


Figura 4.6 - Gráfico de variação da ANMRR em função do número de coeficientes para a base de busca Brodatz 128 x 128.

Pode-se observar na Figura 4.5 que o comportamento da ANMRR para a base Brodatz 512 x 512 oscilou até 10 coeficientes e estabilizou para um número maior de coeficientes. Os melhores valores de ANMRR foram obtidos com 1 e também com 3 coeficientes. Esse resultado parece estranho a princípio e, por isso, os testes com 1 e 3 coeficientes são apresentados.

Pode-se observar na Figura 4.6 que o comportamento da ANMRR para a base Brodatz 128 x 128 também oscilou até 10 coeficientes e estabilizou para um número maior de coeficientes. O melhor valor de ANMRR foi obtido com 3 coeficientes.

4.3.4.1 Base de texturas Brodatz 512 x 512

Para o cálculo do ANMRR foram utilizados: $NG = 6$ e $K = 12$.

A Tabela 4.20 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando a técnica do histograma da direção e magnitude do gradiente, com 3 coeficientes e com distância calculada pela métrica L2, sem normalização. O valor da ANMRR encontrada foi de 0,51027.

Tabela 4.20 - Matriz de confusão para Brodatz 512 x 512, utilizando a técnica do histograma de direção e magnitude do gradiente com 3 coeficientes.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	11			11	3	9		2			10		5
<i>brick</i>		32							8	2			
<i>bubbles</i>			15	2		1		16					4
<i>grass</i>	9		2	7	1	9		4					11
<i>leather</i>					26		3						
<i>pigskin</i>	10		1	8		8		4			2		8
<i>raffia</i>					12		39						
<i>sand</i>	2		19	2		6		12					6
<i>straw</i>		10							13	21			
<i>water</i>									21	19		3	
<i>weave</i>	6			1		1					30		
<i>wood</i>												39	
<i>wool</i>	4		5	11		8		4					8

Também foi realizado o teste utilizando apenas 1 coeficiente. A Tabela 4.21 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando a técnica do histograma da direção e magnitude do gradiente, com 1 coeficiente e com distância calculada pela métrica L2, sem normalização. O valor da ANMRR encontrada foi de 0,50868.

Tabela 4.21 - Matriz de confusão para Brodatz 512 x 512, utilizando a técnica do histograma de direção e magnitude do gradiente com 1 coeficiente.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	10			7		5		4			9		4
<i>brick</i>		32							7				
<i>bubbles</i>			19			1		17					4
<i>grass</i>	7		2	6		13		2			7		9
<i>leather</i>					33		5						
<i>pigskin</i>	7		2	12		7		1			10		7
<i>raffia</i>					7		37						
<i>sand</i>	4		16	2				9					9
<i>straw</i>		10							23	10			
<i>water</i>									12	32			
<i>weave</i>	9			7	2	9					11		4
<i>wood</i>												42	
<i>wool</i>	5		3	8		7		9			5		5

Pode-se observar-se nas Tabelas 4.20 e 4.21 que a utilização de apenas 1 coeficiente diminuiu a confusão para as texturas *bubbles*, *leather*, *straw* e *water*. Entretanto, aumentou bastante a confusão para a textura *weave*. Nenhuma explicação ainda foi encontrada para esse fato.

4.3.4.2 Base de texturas Brodatz 128 x 128

Para o cálculo do ANMRR foram utilizados: NG = 111 e K = 222.

A Tabela 4.22 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a técnica histograma de direção e magnitude do gradiente, com 3 coeficientes e com distância calculada pela métrica L2, sem normalização. O valor da ANMRR encontrada foi de 0,63075.

Tabela 4.22 - Matriz de confusão para Brodatz 128 x 128, utilizando a técnica do histograma de direção e magnitude do gradiente com 3 coeficientes.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	2100	5	1696	2086	1188	1920	434	1323	0	0	1659	0	1745
<i>brick</i>	3	6464	1	0	408	1	1544	0	2571	1081	1	80	0
<i>bubbles</i>	1703	4	2274	1480	278	1922	165	2017	0	0	2379	0	1517
<i>grass</i>	1851	1	1308	2084	1094	1770	103	1409	0	0	545	0	2120
<i>leather</i>	470	290	70	566	5303	202	1267	36	56	0	50	0	220
<i>pigskin</i>	2006	2	1952	2056	608	2095	149	2064	0	0	1214	0	2075
<i>raffia</i>	232	1672	96	75	2717	68	7958	0	124	0	599	0	9
<i>sand</i>	1216	0	1883	1420	95	1834	1	3357	0	0	323	0	2079
<i>straw</i>	0	2380	0	0	59	0	45	0	3511	3774	0	2144	0
<i>water</i>	0	1462	0	0	0	0	0	0	3857	7182	0	419	0
<i>weave</i>	1333	4	1857	570	167	881	749	289	0	0	5340	0	314
<i>wood</i>	0	148	0	0	0	0	0	0	2313	395	0	9789	0
<i>wool</i>	1518	0	1295	2095	515	1739	17	1937	0	0	322	0	2353

Pode-se observar nas matrizes de confusão da Tabela 4.20, para a base Brodatz 512 x 512, e da Tabela 4.22, para a base Brodatz 128 x 128, que os resultados são muito inferiores aos demais métodos. Algumas poucas texturas obtiveram resultados aceitáveis. Para a base Brodatz 512 x 512, as texturas *brick*, *leather*, *raffia*, *water* e *wood* apresentaram confusões menores. Para a base Brodatz 128 x 128, as texturas que apresentaram menor confusão foram: *brick*, *leather*, *raffia*, *water*, *weave* e *wood*. Todas elas, em termos do aspecto visual, apresentam forte direcionalidade. O valor da ANMRR foi indesejavelmente alto, maior do que o obtido no teste da média e da variância. Como os resultados obtidos foram ruins, não foram realizados estudos mais aprofundados sobre essa técnica.

4.3.5 O descritor de texturas homogêneas (HTD)

O descritor de texturas homogêneas não é invariante à rotação. Apesar disso, foram realizados testes utilizando, tanto a base Brodatz 512 x 512, quanto a base Brodatz 128 x 128. Os procedimentos nesses testes foram os mesmos realizados nos testes mostrados anteriormente neste capítulo.

4.3.5.1 Base de texturas Brodatz 512 x 512

Para o cálculo do ANMRR foram utilizados: $NG = 6$ e $K = 12$.

Nestes testes, a métrica recomendada é a L1 e é requerida a normalização pelo desvio da base de dados.

A Tabela 4.23 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando a técnica HTD com distância calculada pela métrica L1, com normalização. O valor da ANMRR encontrada foi de 0,58305.

Tabela 4.23 - Matriz de confusão para Brodatz 512 x 512, utilizando a técnica do descritor de texturas homogêneas.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	28			1	12	8	3		2				
<i>brick</i>		16											15
<i>bubbles</i>			42	6									
<i>grass</i>				18	12	7			5				8
<i>leather</i>				2	6	3			15	5			
<i>pigskin</i>	1			15	10	10	7		9				
<i>raffia</i>							6		4	8	5	9	
<i>sand</i>	13				2	14	11	42			18		
<i>straw</i>									4	6		10	
<i>water</i>									2	4		10	
<i>weave</i>							15			10	19	10	
<i>wood</i>									1	9		3	
<i>wool</i>		26											19

Pode-se observar na matriz de confusão da Tabela 4.23 um comportamento interessante. As texturas que causam confusão menor são aquelas que apresentam um aspecto visual com um menor grau de direcionalidade. Isto ocorre porque a técnica não é invariante à rotação. Uma busca que usa como exemplo uma imagem fortemente direcional retorna como imagens mais semelhantes aquelas que apresentam a mesma direção que ela.

4.3.5.2 Base de texturas Brodatz 128 x 128

Para o cálculo do ANMRR foram utilizados: $NG = 15$ e $K = 30$. Devido à não invariância à rotação, apenas as 16 texturas provenientes da mesma versão rotacionada são comparadas. Como uma delas é a imagem de busca, restam outras 15 no conjunto-verdade.

Nos testes, a métrica recomendada é a L1 e é requerida a normalização pelo desvio da base de dados.

A matriz de confusão para esse teste tem dimensão 91 x 91. Por ser bastante grande, decidiu-se por apresentar uma matriz de confusão reduzida que condensa os resultados para categorias da mesma textura mãe.

A Tabela 4.24 mostra a matriz de confusão reduzida para o teste utilizando a técnica do descritor de texturas homogêneas com distância calculada pela métrica L1, com normalização. O valor da ANMRR foi de 0,38658.

Tabela 4.24 - Matriz de confusão reduzida para Brodatz 128 x 128, utilizando a técnica do descritor de texturas homogêneas.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	1258	78	205	24	17	36	1	8	2				7
<i>brick</i>		1105				1			1				30
<i>bubbles</i>	65	3	1454		4	6			11				1
<i>grass</i>	8	11	1	1406	99	20			8				14
<i>leather</i>	4	13	4	106	1366	43			316				5
<i>pigskin</i>	73	75	16	54	127	1178		46	1				207
<i>raffia</i>		4					1669		10		6		
<i>sand</i>	272	6		88	56	352		1626	2		14		1
<i>straw</i>					11				1095	9		11	
<i>water</i>									85	1509		26	
<i>weave</i>							10				1660		
<i>wood</i>									149	162		1643	
<i>wool</i>		385		2		44							1415

Pode-se verificar na Tabela 4.24 que o desempenho do método do descritor de texturas homogêneas é bastante razoável. As confusões mais perceptíveis entre as texturas são: retornar *sand* quando busca por *bark*, *wool* por *brick*, *bark* por *bubbles*, *sand* por *pigskin*, *leather* por *straw*, *wood* por *water* e *pigskin* por *wool*.

Quando se analisa a matriz de confusão inteira, principalmente os blocos de confusão entre versões rotacionadas de uma mesma textura, verificam-se três comportamentos distintos: baixa confusão, confusão entre as rotações adjacentes e alta confusão. A Tabela 4.25 apresenta um exemplo de baixa confusão que ocorre com a textura *brick*. Outros casos de baixa confusão ocorrem para as texturas *leather*, *raffia*, *straw*, *water*, *weave* e *wood*. Todas elas apresentam um grau de direcionalidade elevado.

Tabela 4.25 – Bloco de confusão para as rotações da textura *brick*, apresentando baixo grau de confusão.

	0°	30°	60°	90°	120°	150°	200°
0°	178	0	0	0	0	0	0
30°	0	101	0	0	0	0	65
60°	0	0	135	0	0	0	0
90°	0	0	0	152	0	0	0
120°	0	0	0	0	145	0	0
150°	0	0	1	0	0	139	0
200°	2	104	0	0	0	0	83

A Tabela 4.26 mostra um exemplo de confusão entre as rotações adjacentes que ocorre com a textura *bark*. Outras texturas que apresentam o mesmo comportamento são: *grass*, *pigskin* e *wool*. Essas texturas (com exceção da *grass*) apresentam algumas linhas mais predominantes, porém, essas linhas não apresentam uma direção muito definida. O fato de existir confusão entre os ângulos adjacentes decorre do próprio método que utiliza um banco de filtros direcionais orientados em múltiplos de 30°. Esse é o mesmo diferencial angular apresentado pelas versões rotacionadas.

Tabela 4.26 – Bloco de confusão para as rotações da textura *bark*, apresentando confusão para as rotações adjacentes.

	0°	30°	60°	90°	120°	150°	200°
0°	97	5	1	2	2	12	35
30°	7	112	15	0	2	0	59
60°	0	5	102	10	1	3	1
90°	3	1	16	148	12	5	5
120°	4	3	2	14	136	4	4
150°	17	2	7	3	6	144	3
200°	63	81	2	7	5	3	87

A Tabela 4.27 mostra um exemplo de alto grau de confusão entre as versões que ocorre com a textura *grass*. A textura *sand* também apresenta o mesmo comportamento. A explicação para esse comportamento baseia-se nas características das texturas. A textura *bubbles* é composta de contornos claros aproximadamente circulares sobre um fundo escuro. As linhas existentes não têm qualquer orientação definida. Portanto, a resposta aos filtros direcionais devem ser aproximadamente as mesmas. Para a textura *sand* a explicação é outra. Não existe, nessa textura, qualquer informação direcional pois ela assemelha-se a um ruído. Assim, a resposta em qualquer direção é também aproximadamente a mesma.

Tabela 4.27 – Bloco de confusão para as rotações da textura *grass*, apresentando alto grau de confusão.

	0°	30°	60°	90°	120°	150°	200°
0°	38	39	17	13	13	23	51
30°	41	57	48	24	12	21	54
60°	11	34	54	37	23	17	26
90°	10	20	38	63	43	20	14
120°	19	14	20	43	57	36	12
150°	39	18	17	16	40	63	21
200°	45	37	18	12	14	19	33

4.3.6 O descritor de histograma de bordas (EHD)

O descritor de histograma de bordas não é invariante à rotação, por isso, foi testado na base Brodatz 128 x 128. Os procedimentos de teste são iguais aos dos testes realizados anteriormente neste capítulo.

Para o cálculo do ANMRR foram utilizados: NG = 15 e K = 30. Por causa da não invariância à rotação, apenas as 16 texturas da mesma versão rotacionada são comparadas. Como uma delas é a imagem de busca, restam outras 15 no conjunto-verdade.

A métrica recomendada é a L1 e não é requerida a normalização pelo desvio da base de dados.

A matriz de confusão para esse teste tem dimensão 91 x 91. Por ser bastante grande, decidiu-se por apresentar uma matriz de confusão reduzida que condensa os resultados para categorias da mesma textura mãe.

A Tabela 4.28 mostra a matriz de confusão reduzida para o teste utilizando a técnica do descritor de histograma de bordas com distância calculada pela métrica L1, sem normalização. O valor da ANMRR encontrada foi de 0,69283.

Tabela 4.28 - Matriz de confusão reduzida para Brodatz 128 x 128, utilizando a técnica do descritor de histograma de bordas.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	381		91	116	214	24	33	24	1		173		46
<i>brick</i>		800							1	1		5	
<i>bubbles</i>	150	3	144	3	8	69	28	58			101		44
<i>grass</i>	35			1075	454				8		10		2
<i>leather</i>	119	1	2	468	914		1		175		6		4
<i>pigskin</i>	211	88	422	1	11	470	234	447			116		378
<i>raffia</i>	183	47	223	1	28	232	806	260	2		78		217
<i>sand</i>	336	61	589	1	5	583	416	655	1		169		439
<i>straw</i>		2			7				1165	21			
<i>water</i>		96							313	1058		481	
<i>weave</i>	22		3	1	1			1			966		1
<i>wood</i>		436							12	600		1194	
<i>wool</i>	243	146	206	14	38	302	162	235	2		61		549

Pode-se observar na Tabela 4.28 que apenas as texturas que apresentam um grau de direcionalidade mais elevado conseguem apresentar resultados de recuperação razoáveis na própria categoria. Essas texturas são: *brick*, *leather*, *raffia*, *straw*, *water*, *weave* e *wood*. A exceção é a textura *grass*, que apresenta um resultado razoável, mas no aspecto visual ela não se distingue pela direcionalidade.

Uma explicação para isso pode ser encontrada na análise da característica de bordas não direcionais. Os valores do histograma de bordas não direcionais para as texturas *grass*, *leather* e *straw* são os mais elevados dentre todas as texturas. Os valores dos histogramas de bordas verticais, horizontais e das duas diagonais também são muito

semelhantes para as texturas *grass* e *leather*, mas são bastante diferentes se comparados com os valores da textura *straw*. Devido a isso as texturas *grass* e *leather* se confundem, mas não são confundidas na mesma intensidade com a textura *straw*. A mesma explicação pode ser considerada para a confusão entre as texturas *pigskin* e *sand*, para a confusão de *bubbles* tanto com *pigskin* quanto com *sand* e para a confusão entre *water* e *wood*. Ambas apresentam valores de histograma de bordas muito semelhantes para todos os cinco tipos de bordas. Esses detalhes podem ser observados na Figura 4.7, que mostra um gráfico comparativo dos valores médios dos histogramas de bordas para cada tipo de borda e para cada tipo de textura.

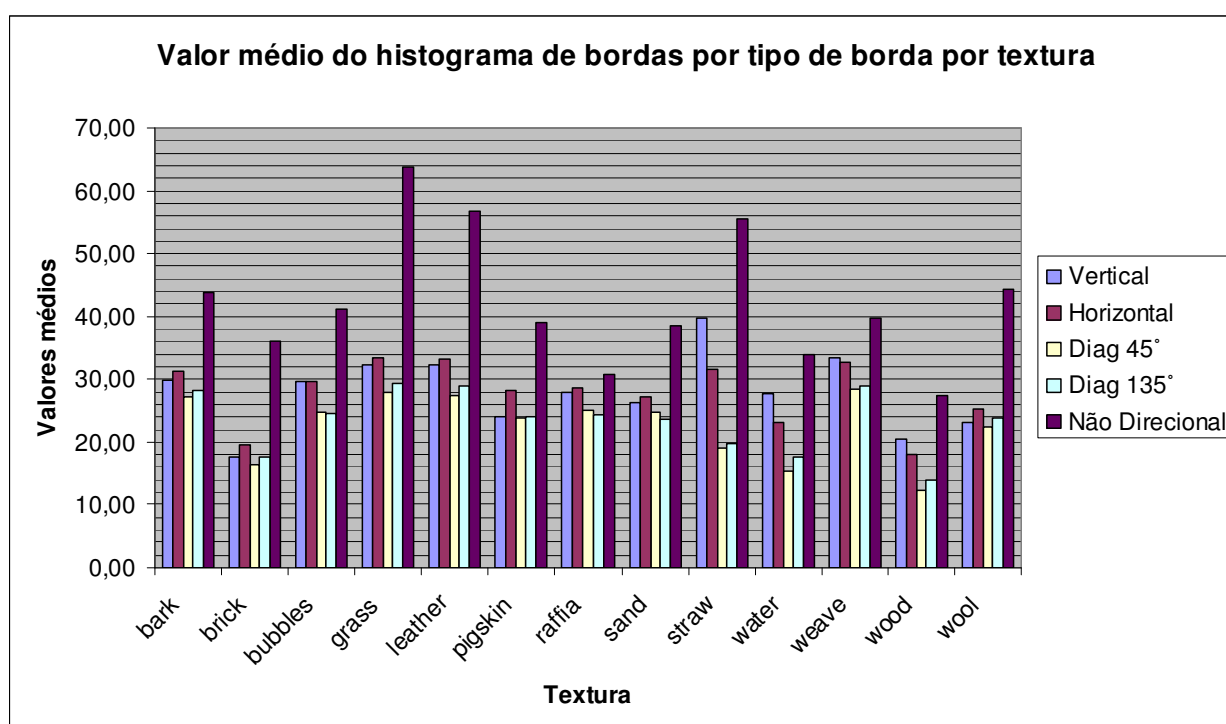


Figura 4.7 – Gráfico comparativo dos valores médios dos histogramas de bordas para as texturas Brodatz 128 x 128.

Pode-se verificar na Figura 4.7 que mesmo as texturas fortemente direcionais apresentam nível de confusão elevado, como, por exemplo, existe um grande retorno de *wood* na busca por *brick*, de *grass* na busca por *leather*, de *sand* por *raffia*, de *water* por *straw* e de *wood* por *water* e vice-versa.

Outro detalhe interessante de ser observado na Figura 4.7 é a recuperação significativa da textura *wool* quando se busca por retalhos das texturas *bark*, *brick*, *bubbles*, *pigskin*, *raffia* e *sand*. Quando se analisa o gráfico dessa figura, pode-se verificar que a relação entre os valores médios para cada tipo de borda obedece ao mesmo padrão: bordas verticais menores do que as horizontais, ambas maiores do que as diagonais; bordas diagonais aproximadamente iguais e bordas não direcionais significativamente maiores do que as demais. Esse padrão recorrente pode explicar a elevada taxa de confusão nos casos mencionados.

Finalmente, a Tabela 4.29 mostra esquematicamente os blocos de confusão para o teste do descritor do histograma de bordas.

Tabela 4.29 - Matriz de confusão reduzida para Brodatz 128 x 128, utilizando a técnica do descritor de histograma de bordas. Os blocos de confusão são apresentados esquematicamente.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>													
<i>brick</i>													
<i>bubbles</i>													
<i>grass</i>													
<i>leather</i>													
<i>pigskin</i>													
<i>raffia</i>													
<i>sand</i>													
<i>straw</i>													
<i>water</i>													
<i>weave</i>													
<i>wood</i>													
<i>wool</i>													



Baixo grau de confusão



Confusão entre as versões com ângulos adjacentes



Alto grau de confusão



Recuperação pertencente às versões rotacionadas da mesma textura

Na Tabela 4.29, as cores dos blocos de confusão indicam se os ângulos das texturas recuperadas são compatíveis com os ângulos das texturas de busca correspondentes. Os blocos verdes indicam compatibilidade entre os ângulos. Os blocos amarelos indicam que existe confusão entre os ângulos adjacentes. Os blocos vermelhos indicam que o grau de confusão é alto e não existe um padrão de relação entre os ângulos das texturas buscadas e recuperadas.

As Figuras 4.8 a 4.10 mostram exemplos de buscas. Cada linha apresenta os sete retalhos mais semelhantes ao retalho de busca localizado na primeira coluna. As setas vazadas marcam os casos de confusão entre os ângulos adjacentes.

Figura 4.8 – Exemplo de resultados de busca de texturas levemente direcionais utilizando o descritor de histograma de bordas.

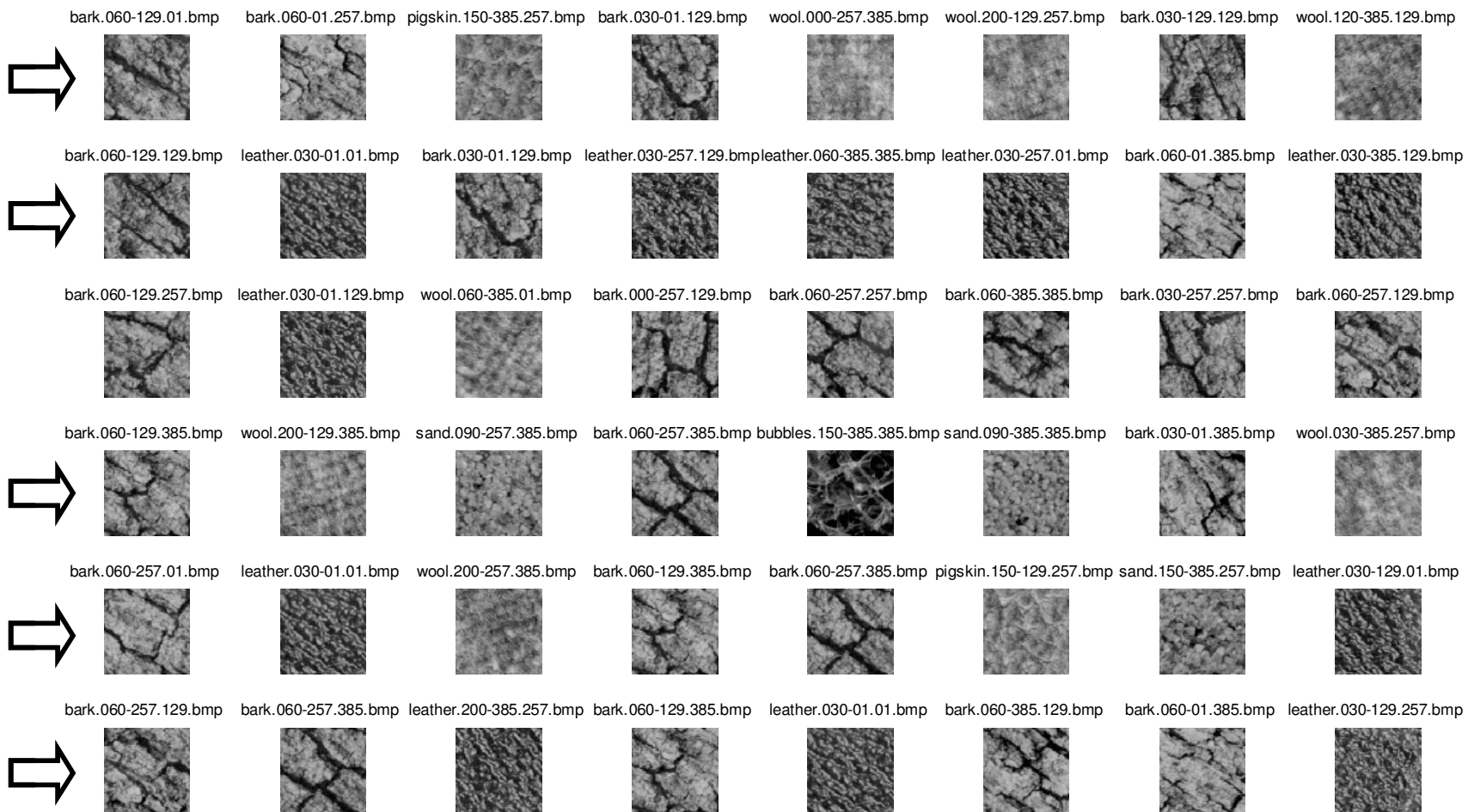
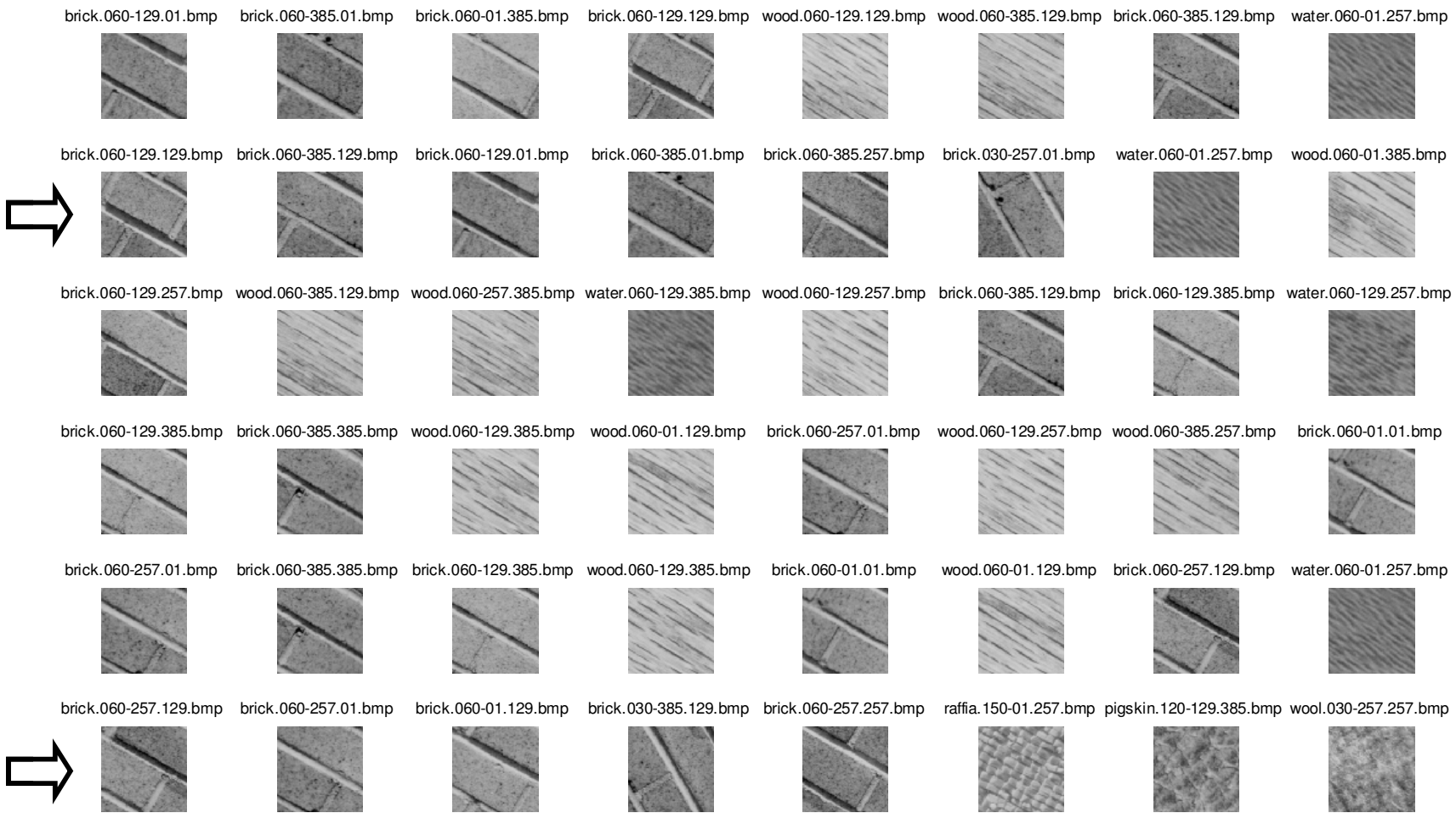


Figura 4.9 - Exemplo de resultados de busca de texturas com direcionalidade mais acentuada utilizando o descritor de histograma de bordas.



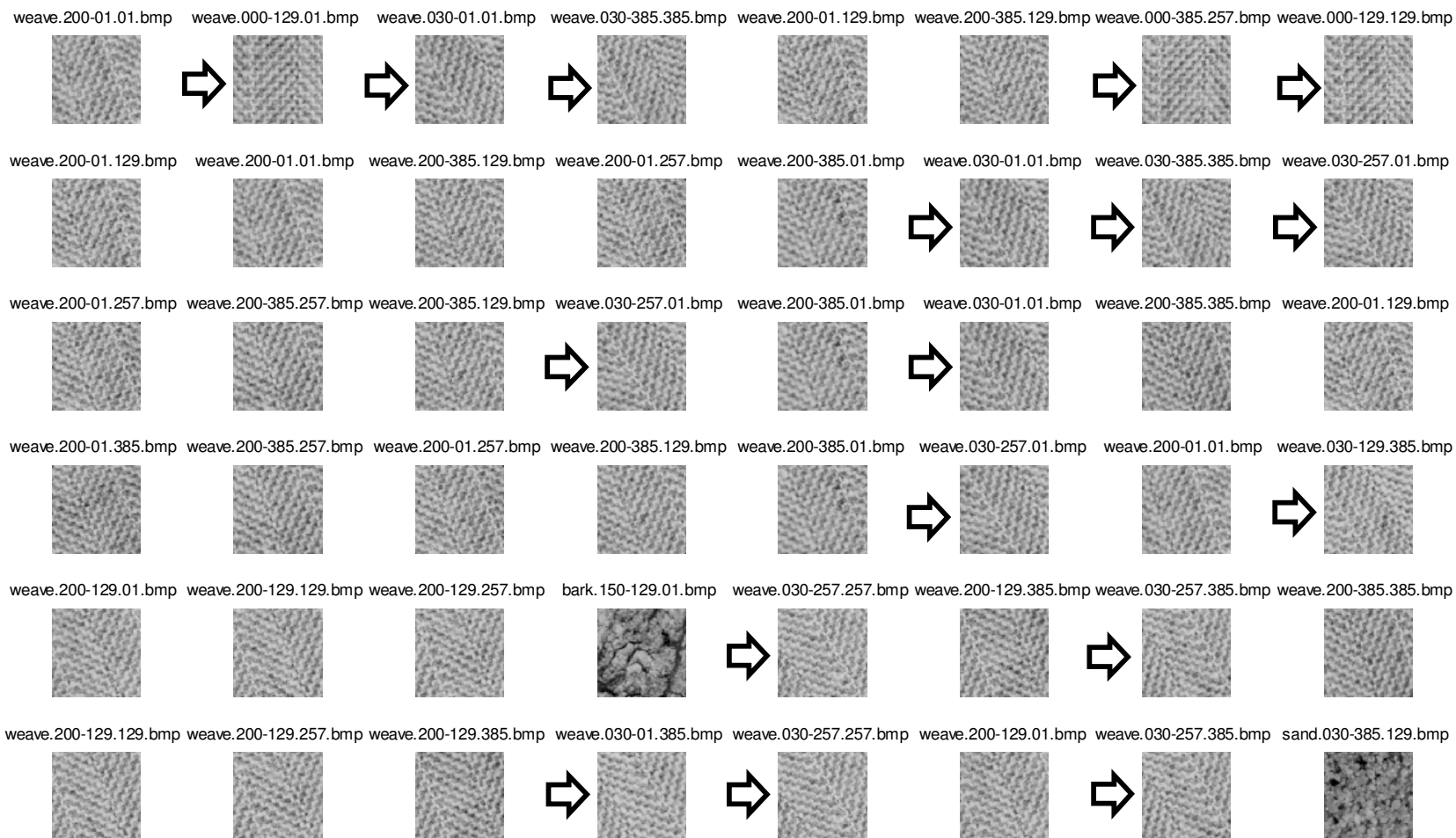


Figura 4.10 – Outro exemplo de resultados de busca de texturas direcionais utilizando o descritor de histograma de bordas. As confusões de ângulos adjacentes estão marcadas.

O descritor de histograma de bordas analisa os ângulos de bordas em incrementos de 45° a partir da horizontal. Entretanto, os ângulos de rotação das texturas analisadas obedecem a incrementos de 30° . Isso explica os blocos amarelos próximos à diagonal principal da Tabela 4.29. Esses blocos são consequência da confusão entre as rotações de 30° e 60° , por exemplo, ou de 200° , 0° e 30° que são comuns nas texturas direcionais. Essa confusão está claramente apresentada nos casos marcados das Figuras 4.9 e 4.10.

4.3.7 O conjunto de histogramas de cores e da direção da variação da magnitude do gradiente

Os testes com o conjunto de histogramas de cores e da direção da variação da magnitude do gradiente (HDCIG + HUCUB + HBCNB) iniciaram-se com o cálculo dos três vetores de características. Da mesma forma que os testes descritos anteriormente neste capítulo, as distâncias cruzadas de todos os pares de imagem da base de testes foram calculadas e os resultados foram classificados em ordem crescente de distância para cada imagem de busca testada. Foi utilizada a métrica L1 para cada vetor e os resultados foram combinados conforme determinado pela técnica. Finalmente, a ANMRR foi calculada. Os testes foram realizados nas bases de busca Brodatz 512 x 512 e Brodatz 128 x 128.

4.3.7.1 Base de texturas Brodatz 512 x 512

Para o cálculo da ANMRR foram utilizados: $NG = 6$ e $K = 12$.

A Tabela 4.30 mostra a matriz de confusão para a base Brodatz 512 x 512 utilizando a técnica dos histogramas HDCIG + HUCUB + HBCNB. O valor da ANMRR encontrada foi de 0,17471.

Tabela 4.30 – Matriz de confusão para Brodatz 512 x 512, utilizando a técnica dos histogramas HDCIG + HUCUB + HBCNB.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	42								2				
<i>brick</i>		41								6		4	
<i>bubbles</i>			42										
<i>grass</i>				26									
<i>leather</i>				16	42								
<i>pigskin</i>						23		1					25
<i>raffia</i>						1	23	8					
<i>sand</i>						2	19	32					
<i>straw</i>									40				
<i>water</i>		1								36			
<i>weave</i>											42		
<i>wood</i>												38	
<i>wool</i>						16		1					17

4.3.7.2 Base de texturas Brodatz 128 x 128

Para o cálculo da ANMRR foram utilizados: NG = 111 e K = 222.

A Tabela 4.31 mostra a matriz de confusão para a base Brodatz 128 x 128 utilizando a técnica dos histogramas HDCIG + HUCUB + HBCNB. O valor da ANMRR encontrada foi de 0,53028.

Tabela 4.31 - Matriz de confusão para Brodatz 128 x 128, utilizando a técnica dos histogramas HDCIG + HUCUB + HBCNB.

	<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>bark</i>	6988		1450	4	136	13	77	119	2895		156	45	39
<i>brick</i>	2	4842	1			226				3385		4471	245
<i>bubbles</i>	503		8738	21	21				1051			77	
<i>grass</i>	77		116	7170	4418				868		73		
<i>leather</i>	254		119	4380	7356				680		1		
<i>pigskin</i>	190	393				3116	2682	2959	33	1156		72	3187
<i>raffia</i>	369	13				2280	3944	3227	106	176		8	1980
<i>sand</i>	733	22	26			2912	3616	3960	65	248		1	2366
<i>straw</i>	3034		1979	804	501	8	56	26	4562		3093	5	17
<i>water</i>	4	4260				903	106	140		4703		2071	1098
<i>weave</i>	35			53					2134		9109	20	
<i>wood</i>		2482				44				1442		5546	6
<i>wool</i>	243	420	3			2930	1951	2001	38	1322		116	3494

A técnica do conjunto de histogramas de cores e da direção da variação da magnitude do gradiente (HDCIG + HUCUB + HBCNB) foi desenvolvida para a busca em banco de imagens coloridas. O objetivo do teste dessa técnica em um banco de texturas com níveis de cinza foi verificar a aplicação dos histogramas de cores de blocos uniformes e blocos bicolores para a descrição das texturas. Esperava-se um desempenho intermediário entre a técnica da média e da variância e as técnicas melhores, como a MCNC e os PBLIRU16. Contudo, os resultados dos testes em ambas as bases de busca, Brodatz 512 x 512 e Brodatz 128 x 128, indicaram um desempenho ainda pior do que aquele da média e da variância. A análise da matriz de confusão ilustrada na Tabela 4.31 mostra que o método dos histogramas HDCIG + HUCUB + HBCNB não proporciona resultados compatíveis com o tempo de processamento que ele dispende, nem com o espaço de armazenamento que ele utiliza.

4.3.8 Comentários sobre os testes de semelhança de texturas

Os testes de semelhança de texturas aplicados nas imagens da base Brodatz 512 x 512 mostraram que a melhor técnica testada foi a do histograma de padrões binários locais invariantes à rotação e uniformes de vizinhança 16 (PBLIRU16).

Em relação a ela (PBLIRU16), a técnica da matriz de co-ocorrência de níveis de cinza (MCNC) apresentou desempenho inferior, porém próximo. O desempenho das técnicas do conjunto de histogramas da direção de variação da magnitude do gradiente (HDCIG), histograma da direção e magnitude do gradiente (HistGrad) e do descritor de texturas homogêneas (HTD) foi bastante inferior. As técnicas HDCIG e HTD não são invariantes à rotação e, portanto, não são adequadas ao problema. Porém, o desempenho da técnica HistGrad, que propõe ser invariante à rotação, ficou muito abaixo do esperado quando se consideram os resultados reportados por Fountain e Tan (FOUNTAIN e TAN, 1998).

A técnica da média e da variância (MedVar) apresentou um desempenho acima do esperado, devido à sua simplicidade. A justificativa para isso está no conjunto de texturas escolhidas, que pode ser particionado pela média da intensidade das texturas. Isso nem sempre ocorre porque texturas bastante distintas podem apresentar intensidades médias e contrastes semelhantes.

A Tabela 4.32 mostra o resumo dos resultados obtidos para a base de busca Brodatz 512 x 512. As técnicas estão classificadas em ordem decrescente de desempenho. A técnica que apresentou melhor desempenho está em destaque nessa tabela.

Tabela 4.32 – Comparação dos resultados obtidos nos testes efetuados nas imagens da base de busca Brodatz 512 x 512.

<i>Técnica</i>	<i>ANMRR</i>
<i>PBLIRU16</i>	0,009
<i>MCNC</i>	0,031
<i>MedVar</i>	0,068
<i>HDCIG</i>	0,175
<i>HistGrad</i>	0,510
<i>HTD</i>	0,583

A Figura 4.11 mostra um gráfico comparativo dos resultados dos testes realizados nas imagens da base de busca Brodatz 512 x 512.

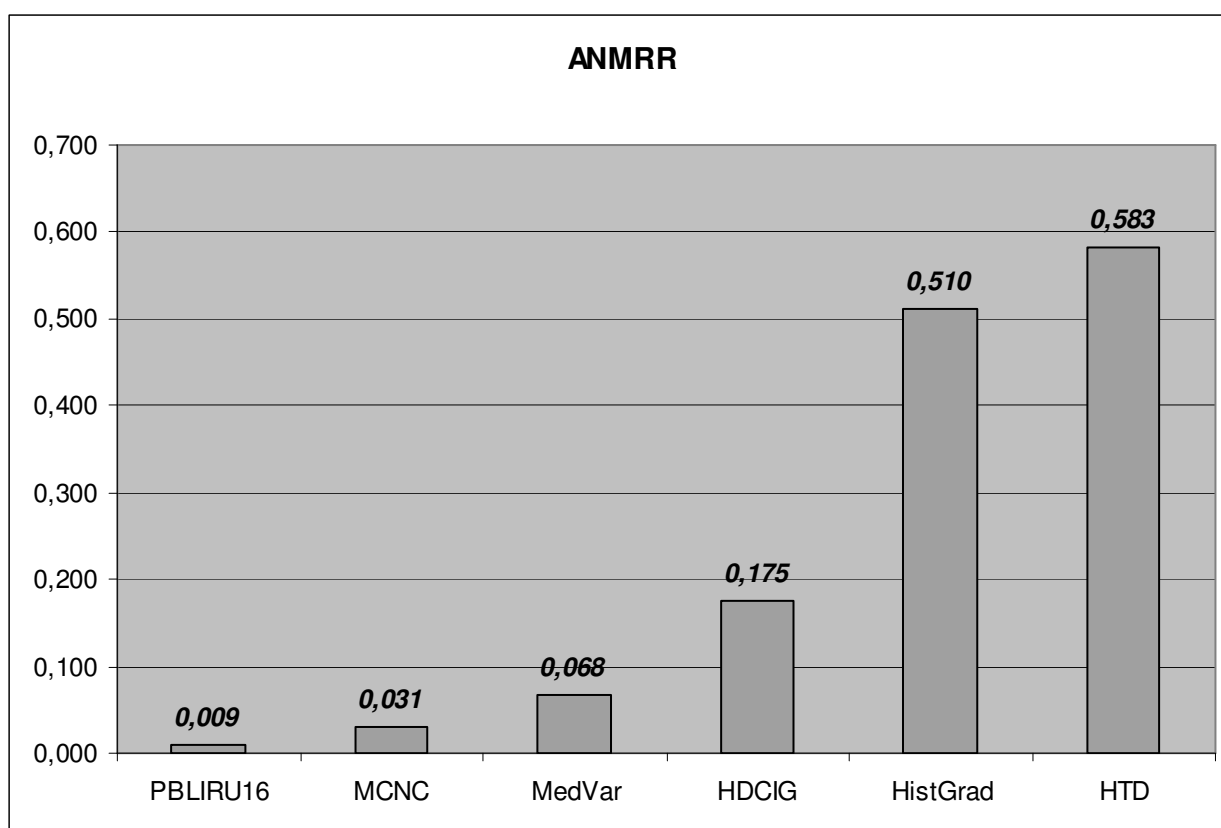


Figura 4.11 – Gráfico comparativo dos resultados dos testes realizados nas imagens da base de busca Brodatz 512 x 512.

Pode-se observar no gráfico da Figura 4.11 três patamares de desempenho. Os melhores resultados foram obtidos pela utilização das técnicas PBLIRU16, MCNC e MedVar.

Em um patamar intermediário localiza-se a técnica HDCIG. As técnicas HistGrad e HDT ocupam o patamar dos piores resultados. O desempenho ruim da técnica HTD era esperado porque ela não é invariante à rotação e a base Brodatz 512 x 512 é aplicável aos algoritmos invariantes à rotação. O desempenho ruim da técnica HistGrad não era esperado. A proposta dessa técnica é interessante, seu mecanismo é lógico e faz sentido, ao menos teoricamente, mas os resultados obtidos apresentaram uma discrepância muito grande com o resultado esperado. Esse desacordo chegou a levantar a suspeita sobre a incorreção da implementação que, no entanto, foi dissipada pela análise dos gráficos dos histogramas para os casos mais discrepantes.

A Tabela 4.33 mostra a comparação dos resultados obtidos nos testes aplicados nas imagens da base de busca Brodatz 128 x 128. As técnicas estão classificadas em ordem decrescente de desempenho. A técnica que apresentou melhor desempenho está em destaque nessa tabela.

Tabela 4.33 – Comparação dos resultados dos testes efetuados nas imagens da base de busca Brodatz 128 x 128.

<i>Técnica</i>	<i>ANMRR</i>
<i>PBLIRU16 + MCNC + MedVar</i>	0,123
<i>PBLIRU16</i>	0,257
<i>MCNC</i>	0,335
<i>MedVar</i>	0,468
<i>HDCIG</i>	0,530
<i>HistGrad</i>	0,631

A Figura 4.12 mostra um gráfico comparativo dos resultados dos testes realizados nas imagens da base de busca Brodatz 128 x 128.

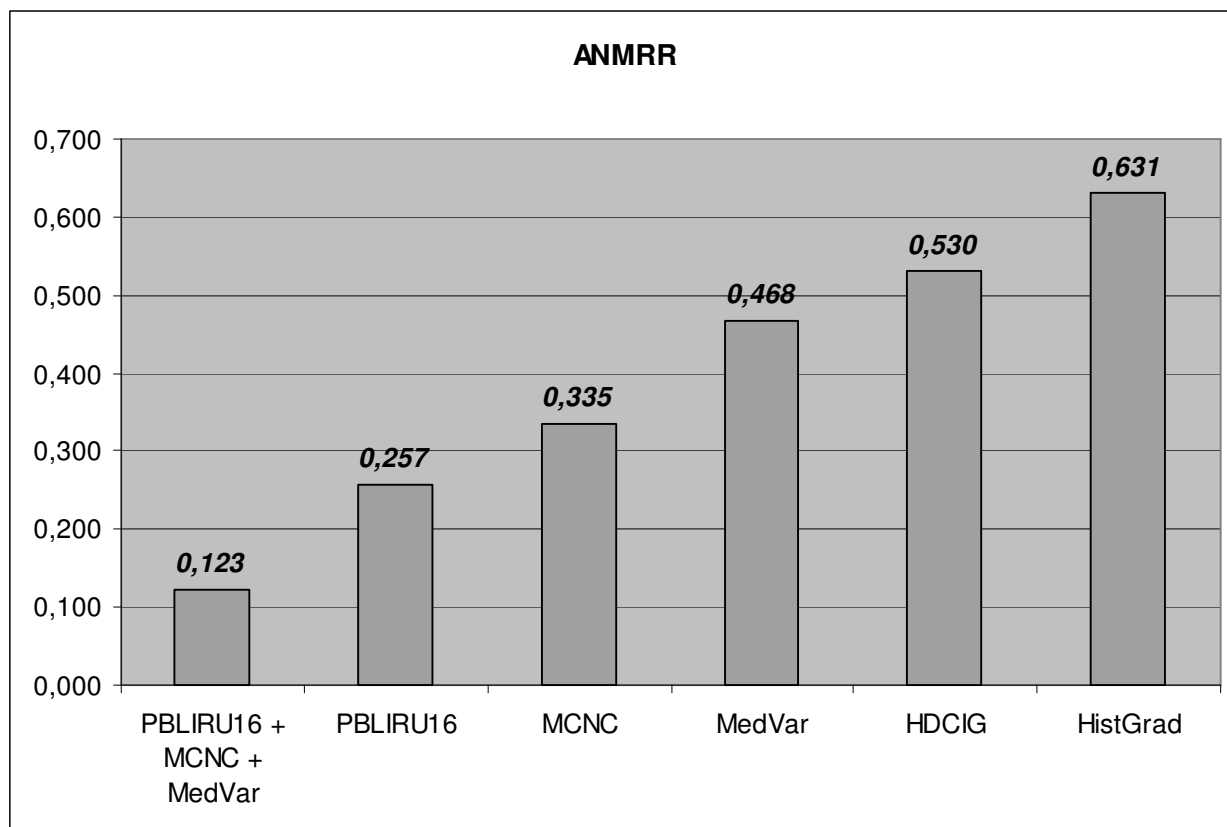


Figura 4.12 – Gráfico comparativo dos resultados dos testes realizados nas imagens da base de busca Brodatz 128 x 128.

Pode-se observar na Figura 4.12 que o mau desempenho da técnica HistGrad confirmou-se também na base de busca Brodatz 128 x 128. O desempenho da técnica HDCIG continuou fraco devido ao fato da não invariância à rotação. A técnica MedVar teve seu desempenho distanciado das melhores técnicas do grupo. Isso confirma a justificativa anteriormente apresentada nesta seção. Como os retalhos de uma mesma textura apresentam variação de tonalidade média e mesmo de contraste, o número de confusões aumentou consideravelmente, o que refletiu negativamente no valor da ANMRR.

O bom desempenho da técnica MCNC surpreendeu bastante. Essa técnica foi uma das primeiras desenvolvidas, datando do início da década de 1970. Não obstante, mostrou-se razoavelmente eficaz nos testes realizados. O tempo de processamento observado foi bastante razoável, não tendo se mostrado como um fator negativo. Na verdade, o tempo de cálculo das características foi um dos menores da categoria para o tipo de aplicação testado.

O método PBLIRU16 apresentou desempenho ligeiramente melhor do que o MCNC. Entretanto, o tempo de processamento gasto nos testes foi bastante maior. O destaque deve ser considerado para a combinação das características do PBLIRU16, MCNC e MedVar, que conseguiu o melhor resultado. A junção das características texturais do PBLIRU16 com as informações de intensidade e contraste do MedVar, aliadas às informações estatísticas de segunda ordem do MCNC forneceram um bom efeito.

Os métodos do descritor de texturas homogêneas (HTD) e do descritor de histograma de bordas (EHD) não são invariantes à rotação e, por isso, foram tabelados separadamente. A ANMRR foi calculada de forma diversa, considerando como conjunto-verdade apenas os retalhos 128 x 128 originados por cada versão rotacionada de cada textura. Os resultados obtidos são mostrados na Tabela 4.34. A técnica EHD apresentou desempenho ruim. A técnica de melhor desempenho foi o HTD. Seu desempenho foi bastante razoável, numericamente comparável ao das técnicas MCNC e PBLIRU16, e melhor do que o desempenho da técnica HDCIG. Isso mostra que a aplicação dos bancos de filtros direcionais de Gabor tem um lugar de destaque na avaliação de semelhança entre as texturas.

Tabela 4.34 – Comparação dos resultados dos testes não invariantes à rotação realizados nas imagens da base de busca Brodatz 128 x 128.

<i>Técnica</i>	<i>ANMRR</i>
<i>HTD</i>	0,387
<i>EHD</i>	0,693

Para finalizar essa análise a Tabela 4.35 mostra uma comparação entre as matrizes de confusão dos principais métodos testados como invariantes à rotação. Os elementos dessa tabela foram convertido para valores percentuais e destacados conforme seu valor com o objetivo de facilitar a análise dos dados.

Tabela 4.35 – Comparação dos valores percentuais das matrizes de confusão para as técnicas testadas como invariantes à rotação nas imagens da base de busca Brodatz 128 x 128.

		<i>bark</i>	<i>brick</i>	<i>bubbles</i>	<i>grass</i>	<i>leather</i>	<i>pigskin</i>	<i>raffia</i>	<i>sand</i>	<i>straw</i>	<i>water</i>	<i>weave</i>	<i>wood</i>	<i>wool</i>
<i>PBLUR16 + MCNC + MedVar</i>	<i>bark</i>	93		12	3	5								
	<i>brick</i>		81							1			1	4
	<i>bubbles</i>			87						3				
	<i>grass</i>				74	4								
	<i>leather</i>				24	91				4				
	<i>pigskin</i>		1	1			71	1	16					26
	<i>raffia</i>							98						
	<i>sand</i>		2				17		81			3		3
	<i>straw</i>			1						91				
	<i>water</i>									1	88		4	
	<i>weave</i>								1			97		
	<i>wood</i>			1							12		95	
	<i>wool</i>			17			13		1					67
<i>PBLUR16</i>	<i>bark</i>	52	2	2	1	13	20	2	10			3		3
	<i>brick</i>		66	5		2							2	1
	<i>bubbles</i>			87										
	<i>grass</i>				43	5	5							30
	<i>leather</i>		7	14	5	5	66	2						5
	<i>pigskin</i>		25	1	1	14	6	48		14		3		12
	<i>raffia</i>							96				1		
	<i>sand</i>		11			1	13		64			14		1
	<i>straw</i>									61	9		17	
	<i>water</i>			1						11	68		18	
	<i>weave</i>		3				2	2	11			79		
	<i>wood</i>			3						27	22		63	
	<i>wool</i>		2	9		36	10	9		1				49
<i>MCNC</i>	<i>bark</i>	74		30						17				
	<i>brick</i>		37	1			4	10	14				4	3
	<i>bubbles</i>		19		66	1				3				
	<i>grass</i>				66	14				6				
	<i>leather</i>				27	70				21				
	<i>pigskin</i>			7			52	6	14		11	1	1	33
	<i>raffia</i>		28				5	58	8				3	14
	<i>sand</i>			20			10	11	56	1	1	3	1	8
	<i>straw</i>		7		2	6	16			50				
	<i>water</i>						2				85		6	
	<i>weave</i>									1		90	1	1
	<i>wood</i>			2							2		85	
	<i>wool</i>			5			28	15	7			5		42
<i>MedVar</i>	<i>bark</i>	51		1	14	1				30				
	<i>brick</i>		32				8	10	21	7		2		8
	<i>bubbles</i>			44	5	38				8				
	<i>grass</i>		15		4	55	4			16				
	<i>leather</i>			43	6	51				8				
	<i>pigskin</i>			5			55		14		12			15
	<i>raffia</i>			18				53	4	1		6	1	24
	<i>sand</i>		26				17	8	39	2				22
	<i>straw</i>		31	4	8	20	6			1	30			
	<i>water</i>						2				88			
	<i>weave</i>			4				3				75	26	
	<i>wood</i>							1				17	73	1
	<i>wool</i>			11			21	24	21					30
<i>HDClG</i>	<i>bark</i>	56		12		1		1	1	23		1		
	<i>brick</i>		39				2				27		36	2
	<i>bubbles</i>		4		70					8			1	
	<i>grass</i>			1	58	36				7		1		
	<i>leather</i>		2		1	35	59			5				
	<i>pigskin</i>		2	3				25	22	24	6		1	26
	<i>raffia</i>		3					18	32	26	1	1		16
	<i>sand</i>		6					23	29	32	1	2		19
	<i>straw</i>		24		16	6	4				37		25	
	<i>water</i>			34				7	1	1		38	17	9
	<i>weave</i>										17	73		
	<i>wood</i>			20								12	45	
	<i>wool</i>		2	3				24	16	16		11	1	28

4.4 Os Testes Quantitativos da Base de Imagens Coloridas SIMPLicity

A base de imagens coloridas SIMPLicity é composta por 1000 imagens de dimensão 256 x 348 *pixels*, tanto em formato de retrato quanto em formato de paisagem. As imagens dessa base são grosseiramente classificadas em 10 categorias, cada uma contendo 100 cenas naturais ou desenhos coloridos. Conforme citado na seção 4.2 deste capítulo, essa classificação grosseira exerce influência negativa nos valores da ANMRR para as técnicas testadas. Entretanto, os recursos disponíveis para a realização deste trabalho não permitiram uma classificação melhor. Por isso, os resultados devem ser considerados como valores relativos para a comparação entre as técnicas, e não como escores absolutos de desempenho dos métodos testados.

Todos os testes obedeceram aos mesmos procedimentos descritos na seção 4.2 deste capítulo. Os vetores de características foram determinados para todas as imagens da base. Todas as distâncias cruzadas foram calculadas e, para cada imagem de busca, as distâncias entre ela e as demais imagens foram classificadas em ordem crescente.

Para o cálculo da ANMRR foram utilizados os seguintes valores de parâmetros: $NG = 99$ e $K = 198$. Em todos os testes deste trabalho utilizou-se o número de posições significativas igual ao dobro do número de elementos do conjunto-verdade correspondente. Os detalhes específicos de cada técnica constam na seção correspondente a cada uma delas.

4.4.1 O descritor de leiaute de cores (CLD)

A Tabela 4.36 mostra a matriz de confusão para a base SIMPLicity utilizando a técnica do descritor de leiaute de cores. O valor da ANMRR encontrada foi de 0,63527.

Tabela 4.36 – Matriz de confusão para SIMPLIcity, utilizando a técnica do descritor de leiaute de cores.

	<i>índigenas</i>	<i>praia</i>	<i>construção</i>	<i>ônibus</i>	<i>dinossauro</i>	<i>elefante</i>	<i>flor</i>	<i>cavalo</i>	<i>montanha</i>	<i>comida</i>
<i>índigenas</i>	2492	815	1264	1708	66	1165	1746	1161	669	2031
<i>praia</i>	447	2530	1448	941	57	739	119	378	1750	323
<i>construção</i>	1114	1811	1754	1410	36	1009	720	762	1940	805
<i>ônibus</i>	1141	1103	1146	2065	50	859	463	267	1314	661
<i>dinossauro</i>	353	156	323	290	9203	1704		173	95	457
<i>elefante</i>	1117	951	1011	1300	401	2415	136	619	937	820
<i>flor</i>	746	93	361	299		63	4150	255	82	1280
<i>cavalo</i>	1081	712	771	368	5	848	790	5615	362	1262
<i>montanha</i>	265	1524	1380	1041	1	648	82	103	2627	123
<i>comida</i>	1144	205	442	478	81	450	1694	567	124	2138

O desempenho do descritor de leiaute de cores não foi muito bom. Ele é calculado reduzindo-se a imagem a um quadrado de dimensão 8×8 *pixels* em que cada *pixel* tem a cor calculada pela média das cores dos *pixels* da região da imagem da qual ele é o centro. Isso realizado, as características são calculadas pelos coeficientes de frequência mais baixa da DCT do quadrado 8×8 . A componente da intensidade (Y) é enfatizada em relação às componentes cromáticas (Cr e Cb) porque o vetor de características produzido é composto por seis coeficientes da componente Y e apenas três coeficientes de cada uma das componentes cromáticas.

Pode-se concluir que o CLD mede a variação espacial da cor na imagem, com ênfase maior na intensidade. Por isso a excepcional taxa de acertos na categoria *dinossauro*, seguida pelas categorias *cavalo* e *flor*. Isso pode ser explicado pela característica marcante das imagens pertencentes a essas categorias. Elas são compostas por um motivo central de cor uniforme sobre um fundo mais ou menos homogêneo. As imagens da categoria *dinossauro* seguem rigidamente esse padrão, com tonalidades próximas tanto para o motivo central quanto para o fundo, que é sempre claro. As imagens da categoria *cavalo* compõem-se de cavalos baios ou castanhos quase sempre centralizados, sobre um fundo verde ou verde amarelado. As imagens da categoria *flor*, que

apresentaram desempenho um pouco pior, são compostas, em sua grande maioria, por uma flor de coloração exuberante sobre um fundo bem mais escuro. O que faz com que o desempenho seja pior é o fato da coloração das flores ser mais variada. As demais categorias apresentaram um grau de confusão muito elevado.

4.4.2 O descritor de estrutura de cores (CSD)

A Tabela 4.37 mostra a matriz de confusão para a base de imagens coloridas SIMPLIcity utilizando a técnica do descritor de estrutura de cores. O valor da ANMRR encontrada foi de 0,5107.

Tabela 4.37 – Matriz de confusão para SIMPLIcity, utilizando a técnica do descritor de estrutura de cores.

	<i>índigenas</i>	<i>praia</i>	<i>construção</i>	<i>ônibus</i>	<i>dinossauro</i>	<i>elefante</i>	<i>flor</i>	<i>cavalo</i>	<i>montanha</i>	<i>comida</i>
<i>índigenas</i>	4065	645	810	1617	129	1321	523	634	334	2262
<i>praia</i>	657	3045	1235	898	12	1085	288	223	2247	214
<i>construção</i>	1144	1411	3561	1792	37	1303	379	211	1718	475
<i>ônibus</i>	765	489	926	3695		164	48	45	503	354
<i>dinossauro</i>	267	391	156	3	9381	931	297	114	383	663
<i>elefante</i>	1083	1063	1017	265	202	2911	262	1134	997	962
<i>flor</i>	125	211	207	111	2	179	6418	702	219	574
<i>cavalo</i>	245	75	38	14		576	614	6360	89	189
<i>montanha</i>	565	2446	1733	1210	67	1044	607	365	3339	161
<i>comida</i>	984	124	217	295	70	386	464	112	71	4046

O desempenho do descritor CSD foi melhor do que o desempenho do CLD. Ele é calculado pela quantização do espaço de cores HMMD seguido pela contabilização, para cada *bin*, do número de elementos de dimensão 8 x 8 *pixels* dos quais a cor faz parte. O resultado disso é um histograma que mede a concentração espacial de cada cor na imagem.

Novamente, os menores graus de confusão são apresentados pelas imagens das categorias *dinossauro*, *flor* e *cavalo*. Houve uma inversão na ordem das categorias *flor* e *cavalo*, sendo que o desempenho de ambas foi bem melhor do que o apresentado pela técnica CLD. As demais categorias apresentaram diminuição da confusão. Duas categorias

apresentaram um nível de melhora menor: *praia* e *elefante*. As categorias com maior nível de confusão são *praia*, *construção*, *ônibus*, *elefante* e *montanha*. A explicação provável está na variação de cores para as imagens de cada uma dessas categorias. Analisando a matriz de confusão, pode-se verificar os pares *praia/montanha* e *construção/montanha*. As classes *praia* e *montanha* contêm muitas imagens com céu azulado, cor acinzentada das rochas e areia e verde de matas e relvas. A confusão do par *construção/montanha* é menor e deve-se, possivelmente, à presença das tonalidades azuladas do céu e acinzentada das rochas e das construções.

4.4.3 O descritor de cores dominantes (DCD)

A Tabela 4.38 mostra a matriz de confusão para a base de imagens coloridas SIMPLIcity utilizando a técnica do descritor de cores dominantes. O valor da ANMRR encontrada foi de 0,6409.

Tabela 4.38 – Matriz de confusão para SIMPLIcity, utilizando a técnica do descritor de cores dominantes.

	<i>índigenas</i>	<i>praia</i>	<i>construção</i>	<i>ônibus</i>	<i>dinossauro</i>	<i>elefante</i>	<i>flor</i>	<i>cavalo</i>	<i>montanha</i>	<i>comida</i>
<i>índigenas</i>	3226	837	1613	1213	152	1416	1064	1344	754	1908
<i>praia</i>	427	2544	889	651	267	828	2985	1646	1800	476
<i>construção</i>	1732	1804	2470	1790	116	1544	790	619	1435	952
<i>ônibus</i>	1117	1183	1673	3231	31	665	652	283	2154	883
<i>dinossauro</i>	58	157	51	30	8568	478	33	193	105	109
<i>elefante</i>	1108	1159	957	416	445	2663	234	1315	892	751
<i>flor</i>	432	146	561	447	13	161	2713	284	135	908
<i>cavalo</i>	233	151	105	48	91	717	237	3237	148	289
<i>montanha</i>	469	1597	951	1464	74	796	174	276	2164	296
<i>comida</i>	1098	322	630	610	143	632	1018	703	313	3328

O desempenho do descritor DCD foi o pior de todos, um pouco atrás do CLD. Esse descritor é composto pelo percentual de incidência das cores dominantes da imagem. O número máximo de cores dominantes é oito, porém, para as imagens testadas, o número de cores do descritor variou de três a cinco. Nessa técnica não é feita uma quantização

grossa do espaço de cores utilizado, o RGB. Os melhores desempenhos (confusão mais baixa) foram observados nas categorias *dinossauros* (bastante superior) e *indígenas*, *ônibus*, *cavalo* e *comida*. Para essas classes, a ocorrência de cores semelhantes nas imagens da mesma classe, é grande. As demais categorias apresentam grande variação interna de tonalidades, o que explica os níveis mais elevados de confusão. No entanto, com exceção da categoria *dinossauro*, a ocorrência de confusão nas outras classes é muito grande.

4.4.4 O descritor escalável de cores (SCD)

A Tabela 4.39 mostra a matriz de confusão para a base de imagens coloridas SIMPLIcity utilizando a técnica do descritor escalável de cores. O valor da ANMRR encontrada foi de 0,49323.

Tabela 4.39 – Matriz de confusão para SIMPLIcity, utilizando a técnica do descritor escalável de cores.

	<i>indígenas</i>	<i>praia</i>	<i>construção</i>	<i>ônibus</i>	<i>dinossauro</i>	<i>elefante</i>	<i>flor</i>	<i>cavalo</i>	<i>montanha</i>	<i>comida</i>
<i>indígenas</i>	5554	979	1260	1228	96	1649	891	719	646	2104
<i>praia</i>	363	2530	1123	494	30	555	120	124	1585	124
<i>construção</i>	805	1934	4106	626	172	1625	477	82	1406	343
<i>ônibus</i>	666	1190	660	5478	52	629	1126	261	2160	703
<i>dinossauro</i>	32	25	82	5	8244	343		3	132	237
<i>elefante</i>	771	956	1295	367	699	2843	196	704	869	617
<i>flor</i>	96	83	206	172		93	4921	472	111	324
<i>cavalo</i>	190	137	57	79	5	595	854	6823	167	178
<i>montanha</i>	158	1704	770	793	137	552	102	108	2570	71
<i>comida</i>	1265	362	341	658	465	1016	1213	604	254	5199

Como pode ser observado na Tabela 4.39, o desempenho do descritor escalável de cores medido pela ANMRR foi o melhor de todos. O descritor SCD é, na verdade, um histograma de cores expressas no espaço de cores HSV cujas componentes são quantizadas uniformemente. O número de intervalos de quantização é maior na componente H, como era de se esperar. Novamente a categoria de melhor desempenho foi *dinossauro*. Em

seguida, em um segundo patamar estão as categorias *cavalo*, *indígenas*, *ônibus*, *comida* e *flor*, nessa ordem. As piores categorias foram *construção*, *elefante*, *montanha* e *praia*, da melhor para a pior. As categorias *praia* e *montanha* constituem um par de confusão observável. A presença de céu azul, ou tonalidades de verde nas imagens dessas categorias explicam esse par.

4.4.5 O conjunto de histogramas de cores e da direção de variação da magnitude do gradiente

A Tabela 4.40 mostra a matriz de confusão para a base de imagens coloridas SIMPLIcity utilizando a técnica dos histogramas HDCIG + HUCUB + HBCNB. O valor da ANMRR encontrada foi de 0,50741.

Tabela 4.40 – Matriz de confusão para SIMPLIcity, utilizando a técnica dos histogramas HDCIG + HUCUB + HBCNB.

	<i>indígenas</i>	<i>praia</i>	<i>construção</i>	<i>ônibus</i>	<i>dinossauro</i>	<i>elefante</i>	<i>flor</i>	<i>cavalo</i>	<i>montanha</i>	<i>comida</i>
<i>indígenas</i>	5073	846	1969	1365	138	1949	1283	461	600	2761
<i>praia</i>	302	3182	1159	523	6	400	87	164	2224	119
<i>construção</i>	1111	1558	3302	722	167	1411	385	136	1357	376
<i>ônibus</i>	380	638	465	4748	6	475	522	110	1257	583
<i>dinossauro</i>	58		33		8762	112			19	108
<i>elefante</i>	1240	906	1469	766	541	2991	277	1264	907	973
<i>flor</i>	32	18	103	35		36	4534	306	66	214
<i>cavalo</i>	257	290	163	68		1161	1374	6988	166	312
<i>montanha</i>	194	2255	932	962	37	526	82	91	3125	84
<i>comida</i>	1253	207	305	711	243	839	1356	380	179	4370

Como pode ser observado na Tabela 4.40, os resultados da técnica dos histogramas HDCIG + HUCUB + HBCNB são muito próximos aos resultados alcançados pelo descritor SCD. As semelhanças ocorrem tanto na medida ANMRR quanto nas confusões das classes, essas em nível um pouco menor.

A métrica definida para o cálculo das distâncias para a técnica dos histogramas HDCIG + HUCUB + HBCNB é uma combinação ponderada de dois componentes. O primeiro mede as distâncias L1 dos histogramas de cor (HUCUB e HBCNB) e as combina equilibradamente. O segundo, que pode ser considerado um componente de textura, mede as distâncias L1 do histograma da direção de variação da magnitude do gradiente (HDCIG). Os autores dessa técnica, Nezamabadi-Pour e Kabir (NEZAMABADI-POUR e KABIR, 2004), testaram-na com pesos iguais para as componentes de cor e de textura. Para avaliar o impacto de pesos diferentes no valor da ANMRR, foram realizadas medidas da ANMRR para as seguintes 11 combinações diferentes de pesos: (0; 1), (0,1; 0,9), (0,2; 0,8),..., (0,9; 0,1) e (1; 0). O primeiro valor do par indica o peso da componente de cor e o segundo indica o peso da componente de textura. O resultado desse teste é mostrado no gráfico da Figura 4.13.

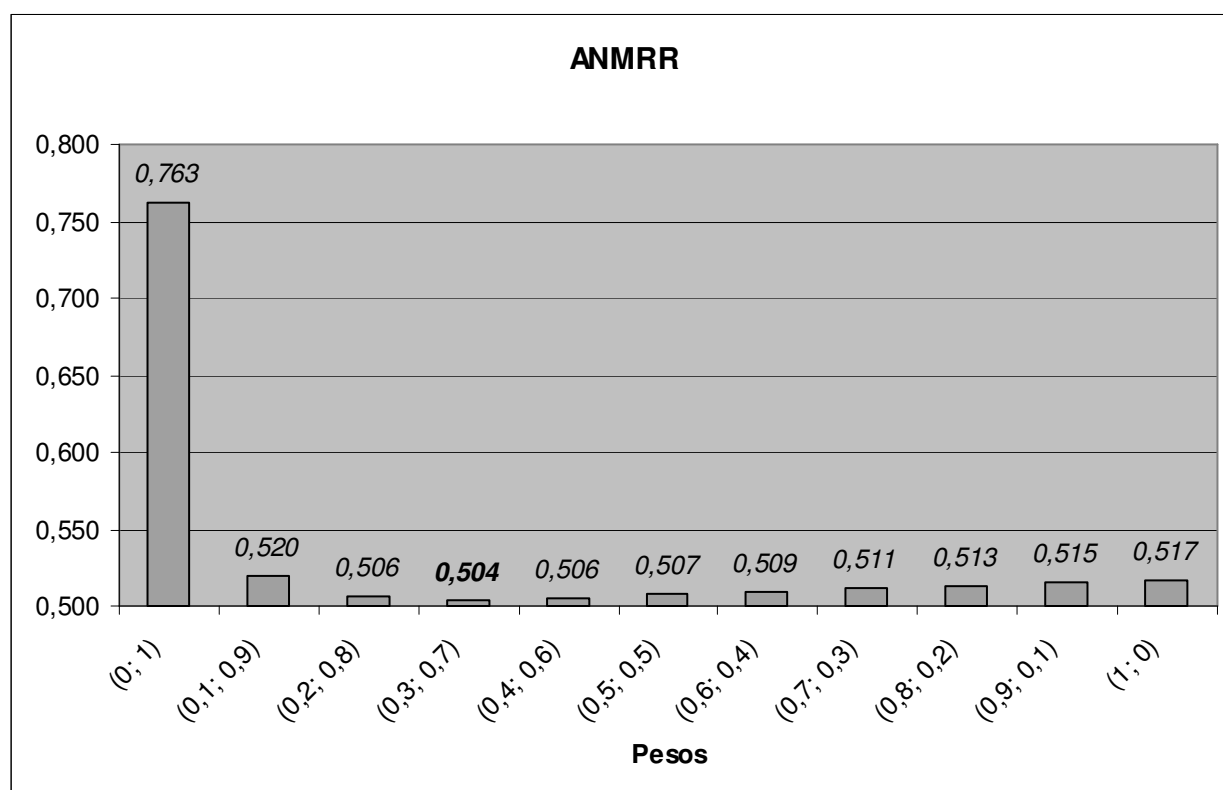


Figura 4.13 – Gráfico dos valores da ANMRR em função dos valores dos pesos das componentes de cor e de textura na métrica da técnica dos histogramas HDCIG + HUCUB + HBCNB.

Observando-se o gráfico da Figura 4.13, pode-se concluir que o valor da ANMRR é mínimo para o par de pesos (0,3; 0,7), embora os pesos utilizados por Nezamabadi-Pour e Kabir (NEZAMABADI-POUR e KABIR, 2004), (0,5; 0,5) resultam em um valor de ANMRR bastante próximo do valor ótimo. Mais importante do que isso é o comportamento do gráfico para os valores extremos. No caso em que o peso da componente de cor é zero, correspondente ao par de pesos (0; 1), o valor da ANMRR atinge um pico, enquanto que para o outro extremo, quando o peso da componente de textura é zero, a ANMRR apresenta um valor bem mais baixo. Isso permite concluir que a componente de textura não desempenha papel tão importante quanto a componente de cor, no cálculo da distância.

4.4.6 Comentários sobre os testes de semelhança de imagens coloridas

Os testes de semelhança de imagens coloridas aplicados nas imagens da base de busca SIMPLIcity mostraram que a melhor técnica testada foi a do descritor escalável de cores (SCD).

As técnicas do conjunto de histogramas da direção de variação da magnitude do gradiente (HDCIG) e do descritor de estrutura de cores (CSD) apresentaram desempenho ligeiramente inferior, sendo que o desempenho da técnica HDCIG foi melhor do que o desempenho da técnica CSD.

Apresentando desempenhos um pouco piores do que os métodos SCD, CSD e HDCIG, as técnicas do descritor de leiante de cores (CLD) e do descritor de cores dominantes (DCD) obtiveram valores de ANMRR aproximadamente iguais.

O fato do SCD e do HDCIG apresentarem desempenhos melhores já era esperado. O número de características dessas técnicas é bastante superior do que os demais. Raciocínio análogo, porém inverso, é válido para o CLD e para o DCD. A técnica SCD determina 256 coeficientes por imagem. A técnica HDCIG determina um total que supera os 200 coeficientes, se for utilizada a assinatura para o HBCNB, ou 1517 coeficientes se for utilizado o histograma para o HBCNB. A técnica CSD implementada utilizou 184 coeficientes. Por outro lado, a técnica CLD utiliza apenas 12 coeficientes e a técnica DCD

utiliza um número variável, que depende do número de cores dominantes da imagem, mas que não ultrapassa 56 para o caso máximo. Um caso típico utiliza 28 coeficientes.

No entanto, nos testes realizados, nem todos os recursos do DCD foram utilizados. A coerência espacial da cor, por exemplo, não foi considerada, apesar de se constituir em uma informação importante que, se utilizada pode melhorar o desempenho desse método. Essa coerência espacial indica como os *pixels* de cada cor dominante estão distribuídos pela imagem: se estão concentrados em blocos, ou se estão dispersos. Essa informação é semelhante àquela medida pelo descritor de estrutura de cores (CSD), que teve desempenho melhor do que o do DCD testado.

A Tabela 4.41 apresenta o resumo dos resultados obtidos para a base de busca SIMPLIcity. As técnicas estão classificadas em ordem decrescente de desempenho. A técnica que apresentou melhor desempenho está em destaque nessa tabela.

Tabela 4.41 – Comparação dos resultados obtidos nos testes realizados nas imagens coloridas da base de busca SIMPLIcity.

<i>Técnica</i>	<i>ANMRR</i>
<i>SCD</i>	0,493
<i>HDCIG</i>	0,507
<i>CSD</i>	0,511
<i>CLD</i>	0,635
<i>DCD</i>	0,641

A Figura 4.14 mostra um gráfico comparativo dos resultados obtidos nos testes realizados nas imagens coloridas da base de busca SIMPLIcity.

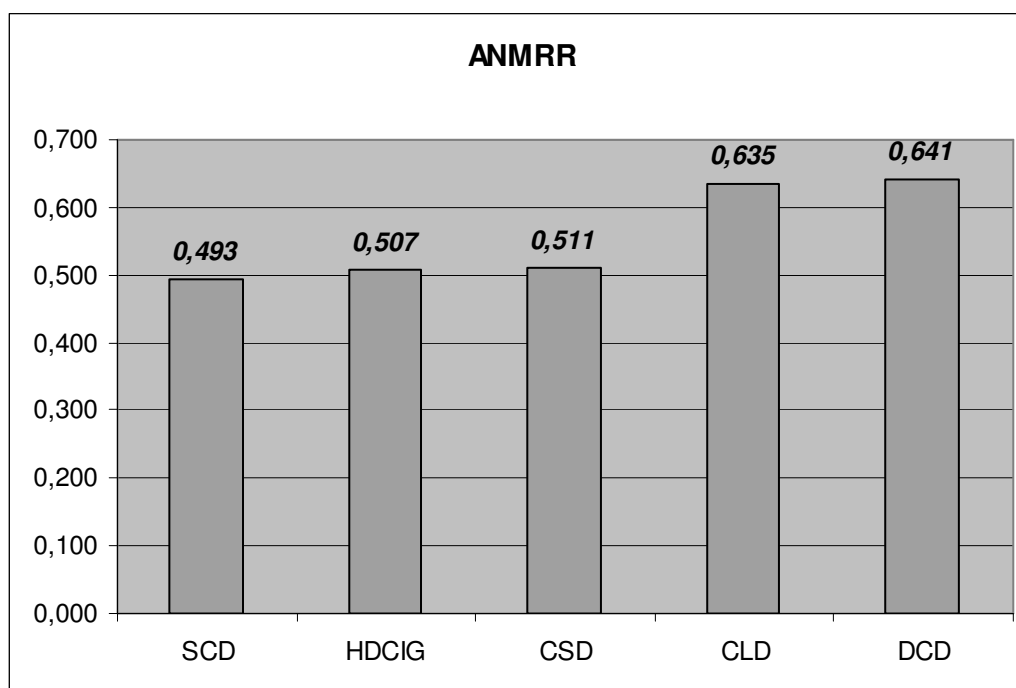


Figura 4.14 – Gráfico comparativo dos resultados obtidos nos testes realizados nas imagens coloridas da base de busca SIMPLIcity.

O gráfico da Figura 4.14 mostra que os resultados dos testes quantitativos realizados nas imagens da base de imagens coloridas SIMPLIcity definiram dois patamares de desempenho. Os melhores resultados foram apresentados por métodos que utilizam histogramas de cores: SCD e HDCIG. O método do descritor de estrutura de cores (CSD) também compôs o patamar de melhor desempenho. Esse método é um histograma de cor modificado que leva em consideração a distribuição mais concentrada ou mais espalhada das cores pela imagem.

Localizados em um patamar de desempenhos menos satisfatórios, surgiram as técnicas CLD E DCD. Apesar do desempenho mais fraco, o DCD inspira possibilidades de melhoria, já que uma de suas características, a coerência espacial, não foi utilizada no cálculo da distância. O melhor desempenho da técnica CSD, que como a coerência espacial do descritor de cores dominantes, também leva em consideração a distribuição das cores pela imagem, sugere essa possibilidade de aumento do desempenho da técnica DCD.

Para finalizar essa análise, a Tabela 4.42 mostra a comparação entre as matrizes de confusão dos métodos testados nas imagens da base SIMPLIcity. Os elementos dessa tabela foram convertidos para valores percentuais para facilitar a análise dos dados.

Tabela 4.42 – Comparação dos valores percentuais das matrizes de confusão para as técnicas testadas nas imagens da base SIMPLicity.

		<i>indígenas</i>	<i>praia</i>	<i>construção</i>	<i>ônibus</i>	<i>dinossauro</i>	<i>elefante</i>	<i>flor</i>	<i>cavalo</i>	<i>montanha</i>	<i>comida</i>
<i>SCD</i>	<i>indígenas</i>	56	10	13	12	1	17	9	7	7	21
	<i>praia</i>	4	26	11	5	0	6	1	1	16	1
	<i>construção</i>	8	20	41	6	2	16	5	1	14	3
	<i>ônibus</i>	7	12	7	55	1	6	11	3	22	7
	<i>dinossauro</i>	0	0	1	0	83	3	0	0	1	2
	<i>elefante</i>	8	10	13	4	7	29	2	7	9	6
	<i>flor</i>	1	1	2	2	0	1	50	5	1	3
	<i>cavalo</i>	2	1	1	1	0	6	9	69	2	2
	<i>montanha</i>	2	17	8	8	1	6	1	1	26	1
	<i>comida</i>	13	4	3	7	5	10	12	6	3	53
<i>HDCIG</i>	<i>indígenas</i>	51	9	20	14	1	20	13	5	6	28
	<i>praia</i>	3	32	12	5	0	4	1	2	22	1
	<i>construção</i>	11	16	33	7	2	14	4	1	14	4
	<i>ônibus</i>	4	6	5	48	0	5	5	1	13	6
	<i>dinossauro</i>	1	0	0	0	89	1	0	0	0	1
	<i>elefante</i>	13	9	15	8	5	30	3	13	9	10
	<i>flor</i>	0	0	1	0	0	0	46	3	1	2
	<i>cavalo</i>	3	3	2	1	0	12	14	71	2	3
	<i>montanha</i>	2	23	9	10	0	5	1	1	32	1
	<i>comida</i>	13	2	3	7	2	8	14	4	2	44
<i>CSD</i>	<i>indígenas</i>	41	7	8	16	1	13	5	6	3	23
	<i>praia</i>	7	31	12	9	0	11	3	2	23	2
	<i>construção</i>	12	14	36	18	0	13	4	2	17	5
	<i>ônibus</i>	8	5	9	37	0	2	0	0	5	4
	<i>dinossauro</i>	3	4	2	0	95	9	3	1	4	7
	<i>elefante</i>	11	11	10	3	2	29	3	11	10	10
	<i>flor</i>	1	2	2	1	0	2	65	7	2	8
	<i>cavalo</i>	2	1	0	0	0	6	6	64	1	2
	<i>montanha</i>	6	25	18	12	1	11	6	4	34	2
	<i>comida</i>	10	1	2	3	1	4	5	1	1	41
<i>CLD</i>	<i>indígenas</i>	25	8	13	17	1	12	18	12	7	21
	<i>praia</i>	5	26	15	10	1	7	1	4	18	3
	<i>construção</i>	11	18	18	14	0	10	7	8	20	8
	<i>ônibus</i>	12	11	12	21	1	9	5	3	13	7
	<i>dinossauro</i>	4	2	3	3	93	17	0	2	1	5
	<i>elefante</i>	11	10	10	13	4	24	1	6	9	8
	<i>flor</i>	8	1	4	3	0	1	42	3	1	13
	<i>cavalo</i>	11	7	8	4	0	9	8	57	4	13
	<i>montanha</i>	3	15	14	11	0	7	1	1	27	1
	<i>comida</i>	12	2	4	5	1	5	17	6	1	22
<i>DCD</i>	<i>indígenas</i>	33	8	16	12	2	14	11	14	8	19
	<i>praia</i>	4	26	9	7	3	8	30	17	18	5
	<i>construção</i>	17	18	25	18	1	16	8	6	14	10
	<i>ônibus</i>	11	12	17	33	0	7	7	3	22	9
	<i>dinossauro</i>	1	2	1	0	87	5	0	2	1	1
	<i>elefante</i>	11	12	10	4	4	27	2	13	9	8
	<i>flor</i>	4	1	6	5	0	2	27	3	1	9
	<i>cavalo</i>	2	2	1	0	1	7	2	33	1	3
	<i>montanha</i>	5	16	10	15	1	8	2	3	22	3
	<i>comida</i>	11	3	6	6	1	6	10	7	3	34

4.5 Os Testes Qualitativos da Base de Imagens Coloridas VisTex

Os testes qualitativos nas imagens da base de busca VisTex foram executados com a finalidade de avaliar visualmente o comportamento das técnicas do descritor escalável de cores (SCD), do conjunto de histogramas da direção de variação da magnitude do gradiente (HDCIG), do descritor de estrutura de cores (CSD), do descritor de leiaute de cores (CLD) e do descritor de cores dominantes (DCD).

Por ser uma base composta por um número mais reduzido de imagens, quando comparada com a base SIMPLIcity, ela foi útil para executar os testes de implementação desses métodos e também para possibilitar uma análise preliminar dos seus desempenhos. Além disso, as imagens da base Vistex apresentam imagens e texturas coloridas de maior variedade e com maior complexidade. Isso permitiu verificar melhor o comportamento dos métodos testados, principalmente do HDCIG quando as texturas da imagem são tão importantes quanto as tonalidades de cor que elas apresentam.

Os resultados qualitativos seguem com bastante proximidade os resultados quantitativos apresentados na seção 4.4 deste capítulo. A única observação adicional refere-se à percepção, bastante subjetiva, é verdade, de que a técnica HDCIG apresentou um resultado melhor.

As Figuras 4.15 a 4.17 mostram alguns exemplos de busca por semelhança utilizando a técnica HDCIG nas imagens coloridas da base VisTex. Cada linha corresponde a uma busca. Em cada linha, a imagem da esquerda é a imagem de busca e as sete demais imagens imediatamente à direita são as imagens mais semelhantes retornadas, classificadas em ordem decrescente de similaridade. A segunda imagem de cada linha é a mais semelhante e a oitava imagem é a menos semelhante da linha. A orientação da figura foi modificada para melhor visualização na página.

Figura 4.15 – Primeiro exemplo de busca por semelhança na base Vistex utilizando a técnica HDCIG.

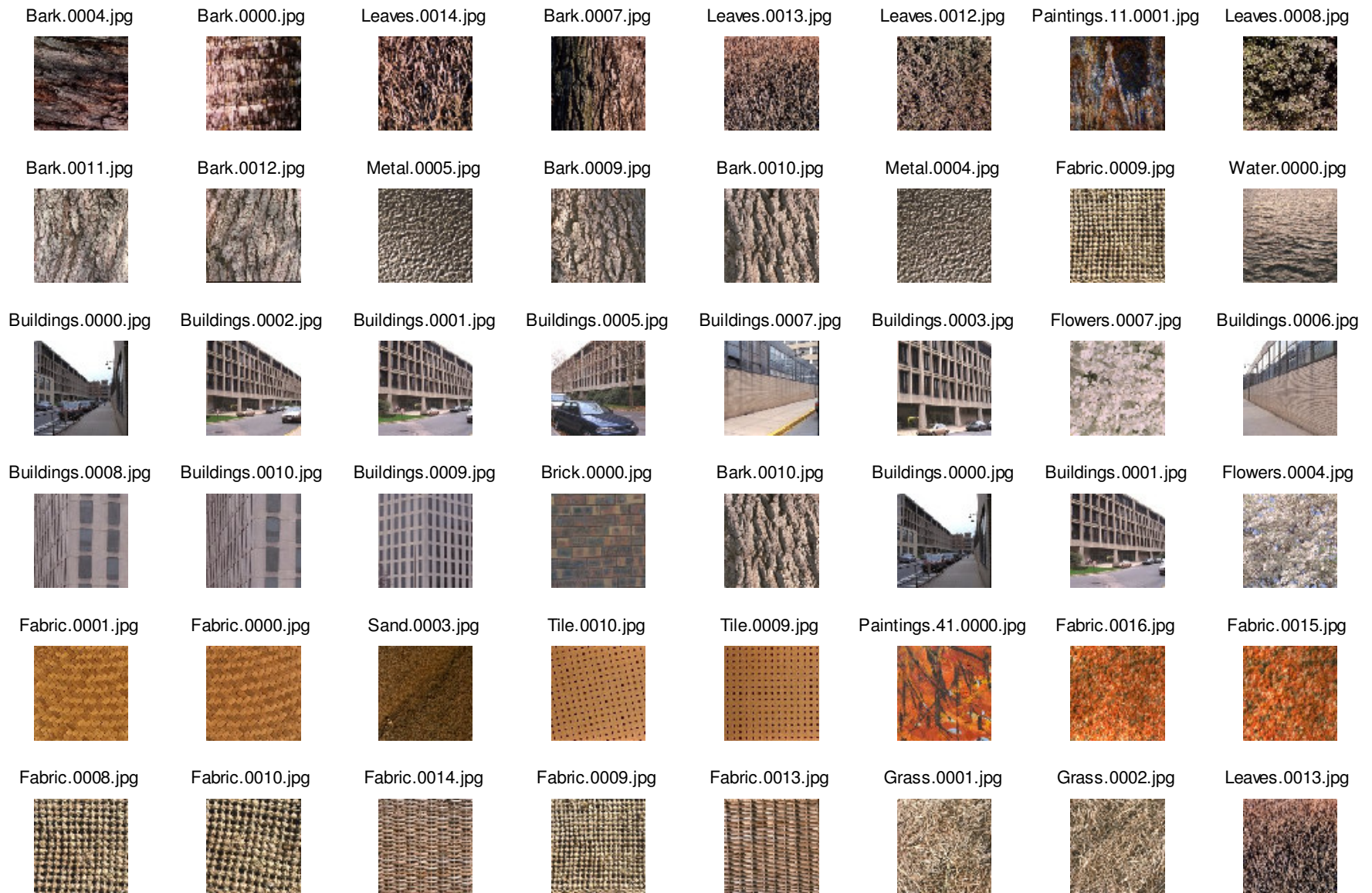


Figura 4.16 – Segundo exemplo de busca por semelhança na base VISTex utilizando a técnica HDCIG.

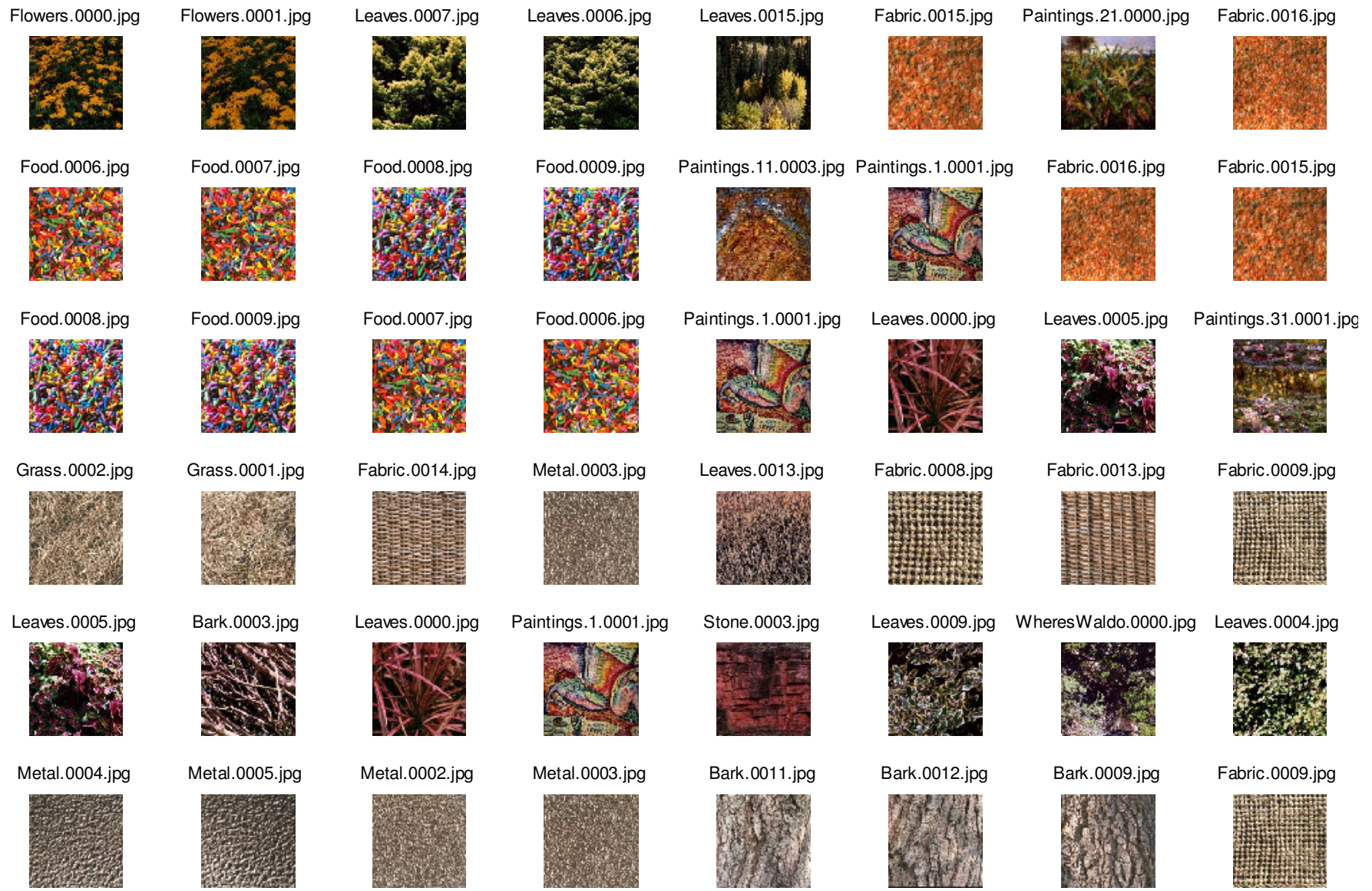
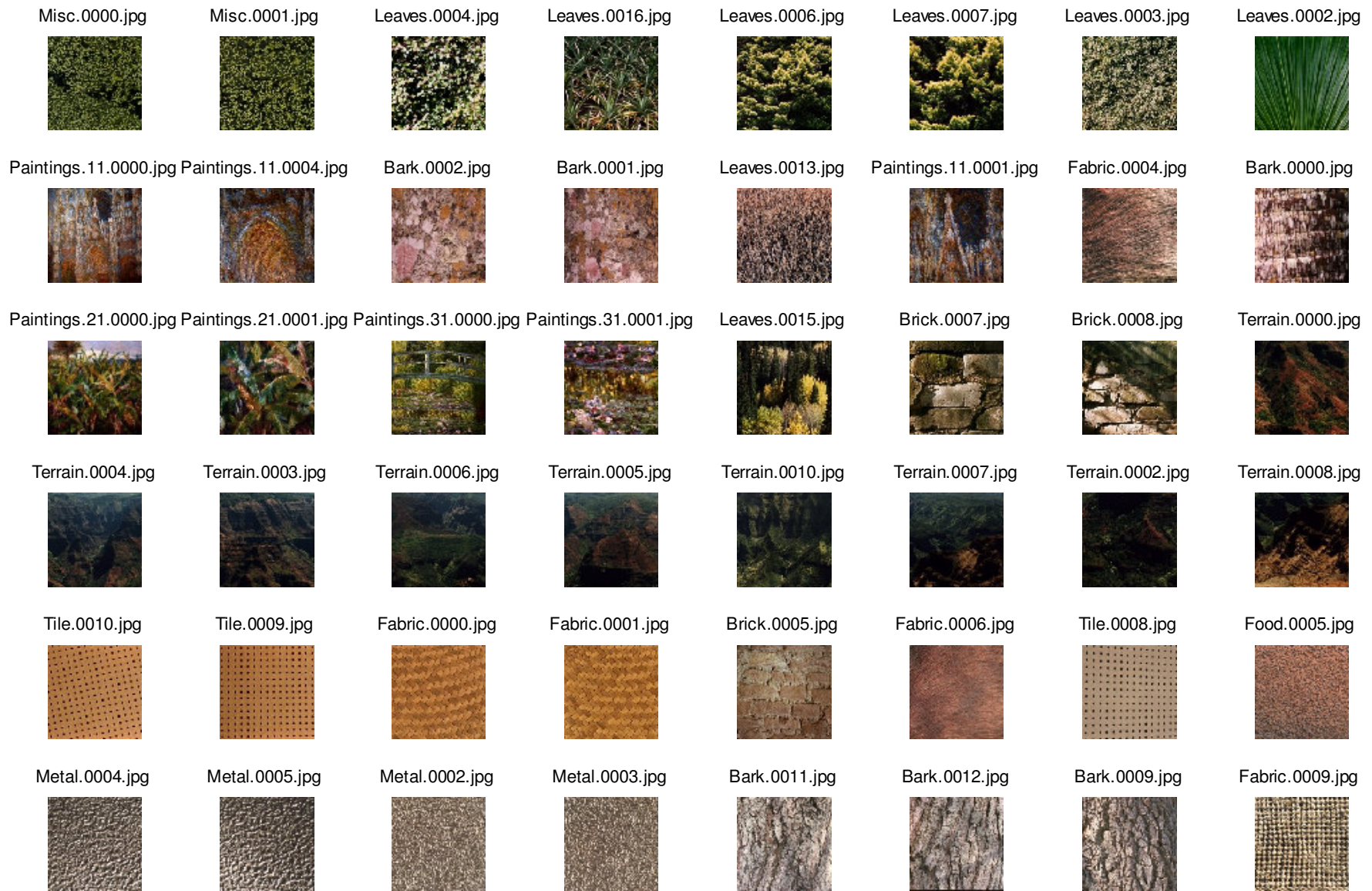


Figura 4.17 – Terceiro exemplo de busca por semelhança na base Vistex utilizando a técnica HDCIG.



4.6 Conclusões

O objetivo deste capítulo foi apresentar os testes realizados para avaliar o desempenho das principais técnicas de busca de imagens por semelhança que foram descritas no Capítulo 3 deste trabalho.

Foram apresentadas as características das imagens e texturas componentes das bases de busca, as ferramentas utilizadas, os tipos de testes realizados e o equipamento utilizado para a execução.

O procedimento dos testes foi padrão, constituído basicamente das seguintes etapas: cálculo dos vetores de características para cada imagem, cálculo das distâncias entre cada imagem e cada uma das demais, classificação em ordem decrescente das distâncias das imagens retornadas para cada imagem de busca e cálculo do desempenho da técnica.

O desempenho foi medido pela utilização da ANMRR, taxa média de retorno modificada e normalizada, proposta e utilizada nos testes do padrão MPEG-7. Essa métrica permite a comparação entre as diversas técnicas utilizadas e, por ser normalizada, possibilita uma noção comparativa real dos resultados dos testes realizados.

Três tipos de testes foram realizados: testes quantitativos sobre texturas utilizando técnicas invariantes e não invariantes à rotação, testes quantitativos sobre imagens coloridas, e sobre essas imagens, testes qualitativos que tiveram como objetivo orientar a implementação dos diversos algoritmos.

Nos testes quantitativos, os resultados obtidos foram apresentados por meio de tabelas comparativas de desempenho e de matrizes de confusão. Estas matrizes são um bom instrumento para a visualização das inversões de classe que aparecem nas imagens recuperadas. Para ilustrar os testes qualitativos foram apresentados os resultados das buscas por meio das listas resumidas de imagens retornadas, ao lado das imagens de busca correspondentes.

Os resultados dos testes quantitativos realizados nas imagens da base Brodatz 512 x 512 *pixels* mostraram que as técnicas que mais se destacaram foram a PBLIRU16, a

MCNC e a MedVar, nesta ordem, seguidas pela HDCIG. O desempenho da PBLIRU16 foi excelente, apresentando poucas inversões de classe. O desempenho da MCNC foi muito bom, também apresentando baixo nível de confusão. O desempenho da técnica MedVar foi surpreendente, mas foi devido, principalmente, às características específicas das texturas utilizadas no teste. O desempenho do HDCIG não foi dos melhores, mas sugere a possibilidade de melhoria pela utilização de uma componente de textura mais adequada.

Os piores desempenhos para a base Brodatz 512 x 512 *pixels* foram apresentados pelo HistGrad e pelo HTD. O resultado do descritor de texturas homogêneas não surpreendeu negativamente porque esse descritor não é invariante à rotação e, portanto, é incompatível com a base de testes utilizada. Surpreendentemente ruim foi o resultado do HistGrad, que superou as piores expectativas. A proposta da técnica HistGrad é interessante, seu mecanismo é lógico e faz sentido, ao menos teoricamente, mas os resultados obtidos apresentaram uma discrepância muito grande. Esse desacordo chegou a levantar a suspeita sobre a incorreção da implementação que, no entanto, foi dissipada pela análise dos gráficos dos histogramas para os casos mais discrepantes.

Os resultados dos testes quantitativos para a base de texturas Brodatz 128 x 128 *pixels* permitiram uma visão mais realista dos desempenhos das técnicas testadas. As técnicas que mais se destacaram foram a PBLIRU16 e MCNC. MedVar e HDCIG foram colocadas em um patamar inferior de resultados. A técnica HistGrad teve seu desempenho ruim confirmado. A surpresa positiva desses testes foi o resultado obtido com a combinação da PBLIRU15, MCNC e MedVar. O desempenho dessa combinação posicionou-a isolada no patamar de melhor desempenho.

Os testes das técnicas HDT e EHD com as imagens da base Brodatz 128 x 128 *pixels* foram destacados dos demais por se tratar de técnicas não invariantes à rotação. A técnica HTD apresentou o melhor desempenho, que foi bastante superior ao observado para a técnica EHD. Isso indica que o banco de filtros de Gabor, utilizado na HTD apresenta uma aplicabilidade razoável na análise de texturas.

Os testes quantitativos realizados nas imagens da base de imagens coloridas SIMPLIcity não trouxeram nenhuma surpresa. Os melhores resultados foram apresentados

por métodos que utilizam histogramas de cores: SCD e HDCIG. Um terceiro método compôs o patamar de melhor desempenho: o descritor de estrutura de cores (CSD). Esse método é um histograma de cor modificado que leva em consideração a distribuição mais concentrada ou mais espalhada das cores pela imagem.

Localizados em um patamar de desempenhos menos satisfatórios, surgiram as técnicas CLD E DCD. Apesar do desempenho mais fraco, o DCD inspira possibilidades de melhoria, já que uma de suas características, a coerência espacial, não foi utilizada no cálculo da distância. O melhor desempenho da técnica CSD sugere essa possibilidade de aumento do desempenho da técnica DCD.

Os testes qualitativos sugeriram as técnicas HDCIG, HSC e CSD como as mais apropriadas. Os resultados visuais para as sete imagens mais semelhantes nas buscas apontam para a técnica HDCIG como sendo a mais promissora para aperfeiçoamentos. Cabe ressaltar que os testes qualitativos trazem uma carga subjetiva muito grande para que sejam considerados para outros objetivos, que não apenas os de sugerir os caminhos mais apropriados.

Este capítulo apresentou os testes realizados, resultados obtidos e conclusões importantes sobre esses testes.

O próximo capítulo apresenta as conclusões e as contribuições desta dissertação, e as sugestões para trabalhos futuros que possibilitem prosseguimento aos estudos realizados.

5 CONCLUSÕES, CONTRIBUIÇÕES E TRABALHOS FUTUROS

5.1 Introdução

A recuperação de imagens semelhantes em bancos de dados é um campo de pesquisa amplo, complexo e que apresenta grande demanda por aplicativos que apresentem bons resultados. Entre várias características para a determinação da semelhança entre imagens digitais, a textura e a cor se destacam por permitir uma análise preliminar mais rápida que reduz o universo de busca.

Várias técnicas que analisam textura e cor foram descritas neste trabalho e foram testadas algumas delas que utilizam histogramas. Foram elas: a média e a variância, a matriz de co-ocorrência de níveis de cinza (MCNC), os padrões binários locais invariantes à rotação e uniformes na vizinhança 16 (PBLIRU16), os histogramas combinados de cor e textura (HDCIG + HUCUB + HBCNB), o histograma de direção e magnitude do gradiente, o descritor de texturas homogêneas (HTD), o descritor de histograma de bordas (EHD), o descritor escalável de cores (SCD), o descritor de estrutura de cores (CSD), o descritor de leiaute de cores (CLD) e o descritor de cores dominantes (DCD).

Os resultados desses testes foram apresentados, o desempenho foi medido pela ANMRR e as matrizes de confusão foram calculadas e analisadas. Finalmente foi realizada a comparação entre as técnicas testadas.

5.2 Conclusões

Este trabalho apresentou um conjunto significativo de técnicas de determinação de semelhança entre imagens digitais utilizando a comparação de histogramas como ferramenta fundamental. Algumas das principais técnicas propostas na literatura foram implementadas e testadas quanto a seu desempenho e os resultados dos testes foram mostrados de forma comparativa.

Dois grupos de características visuais foram avaliados: as texturas e as cores. As características texturais obtidas da análise da distribuição de níveis de cinza mostraram-se eficientes quanto aos objetivos propostos. Os testes utilizando as características relacionadas às cores também se revelaram promissores, porém, sua aplicação, testes e análise mais profunda dos resultados requerem recursos maiores, não disponíveis quando da realização deste trabalho.

Das técnicas testadas sobre texturas em níveis de cinza, a utilização de padrões binários locais uniformes e invariantes à rotação, PBLIRU16, combinada com as características das matrizes de co-ocorrência de níveis de cinza, MCNC, e com a média e variância apresentou o melhor desempenho, destacando-se dos demais.

Quanto às técnicas testadas nas imagens coloridas, três delas se destacaram: descritor escalável de cores (SCD), descritor de estrutura de cores (CSD) e a combinação dos histogramas de textura e de cores (HDCIG + HUCUB + HBCNB). Os resultados obtidos pela aplicação dos descritores escaláveis de cor (SCD), dos descritores de estrutura de cores (CSD) e da combinação do histograma de direção de variação da magnitude do gradiente com o histograma de blocos de cor uniforme e do histograma de blocos de duas cores (HDCIG + HUCUB + HBCNB) foram bons tanto nos testes de avaliação quantitativa quanto na avaliação qualitativa. Os descritores de leiaute de cores (CLD) e de cores dominantes (DCD) apresentaram um desempenho quantitativo pior. Com relação ao descritor de cores dominantes (DCD) deve-se enfatizar que a característica de coerência de cor, nele presente, não foi utilizada no cálculo da semelhança.

5.3 Contribuições

Apesar de sua característica eminentemente introdutória, esta dissertação trouxe algumas contribuições ao campo de estudo da determinação de semelhança entre imagens digitais.

Em primeiro lugar está a reunião, em um mesmo estudo comparativo, de técnicas diversas que se baseiam nas características texturais, ou nas características de cor, ou na combinação de ambas.

Outra contribuição importante foi a de apresentar as matrizes de confusão para cada teste realizado.

Importante também foi a proposição e teste da combinação entre as técnicas PBLIRU16, MCNC e da média e da variância. Essa combinação apresentou resultados muito superiores àqueles obtidos pela aplicação das técnicas isoladamente. Como essa técnica não foi testada completamente em seus detalhes, sugere-se o prosseguimento desses estudos para a otimização dos resultados.

A realização dos testes da técnica HDCIG + HUCUB + HBCNB nas imagens da base de texturas Brodatz 128 x 128 *pixels* foi esclarecedora porque mostrou que o desempenho da componente de textura HDCIG pode ser melhorado.

Os testes que compararam o desempenho da técnica HDCIG + HUCUB + HBCNB em função dos pesos relativos das componentes de textura (HDCIG) e cor (HUCUB + HBCNB) também contribuíram para mostrar que essa técnica comporta o prosseguimento dos estudos visando seu aperfeiçoamento.

Outra contribuição deste trabalho foi a determinação do comportamento do desempenho da técnica HistGrad em função do número de coeficientes da FFT utilizados no vetor de características. Apesar do desempenho ruim apresentado por essa técnica, os valores obtidos diferiram dos valores sugeridos por seus autores Fountain e Tan (FOUNTAIN e TAN, 1998).

5.4 Trabalhos Futuros

Alguns estudos podem ser sugeridos para a verificação da melhoria de desempenho das técnicas apresentadas neste trabalho. No campo das texturas, algum progresso pode ser obtido na combinação dos PBLIRU16 com as MCNC e das médias e das variâncias. Será de grande utilidade um estudo que determine as principais características das MCNC, da média e da variância que, combinadas com os PBLIRU16, obtenha os melhores resultados com o menor tamanho de vetor de características possível.

Outro trabalho que pode ser realizado é o prosseguimento das pesquisas iniciadas por Tamura e outros (TAMURA *et alii*, 1978) na busca de características texturais

que correspondam a padrões visuais. Uma abertura a essa continuidade foi deixada no padrão MPEG-7 e algumas propostas já foram realizadas por Manjunath e outros (MANJUNATH *et alii*, 1996). No entanto, esse caminho, permanece ainda aberto a estudos futuros.

No campo das imagens coloridas, o desenvolvimento de uma métrica que inclua a coerência de cor na técnica DCD visando à melhoria de desempenho é uma boa possibilidade. Outra possibilidade de projetos futuros é a pesquisa para a substituição do histograma HDCIG na técnica combinada HDCIG + HUCUB + HBCNB. Nessa técnica, o papel dos histogramas de cor sobressai em relação ao histograma de textura HDCIG. Por isso, é possível que sua substituição por outra característica textural seguida pela necessária alteração e ajuste da métrica de comparação possa permitir melhorias importantes do desempenho, que já é bastante adequado.

Os progressos no campo da determinação de semelhança entre imagens visando à busca em bancos de dados foram notáveis. Muito já foi realizado. Entretanto, muitas técnicas ainda podem ser desenvolvidas para melhorar os resultados nesse campo.

REFERÊNCIAS BIBLIOGRÁFICAS

- AMADASUN, M., KING, R. **Textural features corresponding to textural properties**, *IEEE Transactions on Systems, Man and Cybernetics* SMC-19:5 (1989) 1264-1274.
- BARALDI, A., PARMIGGIANI, F. **An investigation of the textural characteristics associated with Gray level cooccurrence matrix statistical parameters**, *IEEE Transactions on Geoscience and Remote Sensing* 33:2 (1995) 293-304.
- BRODATZ, P. **Texture — a photographic album for artists and designers**. New York : Reinhold, 1968.
- CHA, S., SRIHARI, S. N. **On measuring the distance between histograms**, *Pattern Recognition* 35 (2002) 1355-1370.
- CHANG, S., SIKORA, T., PURI, A. **Overview of the MPEG-7 standard**, *IEEE Transactions on Circuits and Systems for Video Technology* 11:6 (2001) 703-715.
- CHEN, K., CHEN, S. **Color texture segmentation using feature distributions**, *Pattern Recognition Letters* 23 (2002) 755-771.
- CHENG, Y., CHEN, S. **Image classification using color, texture and regions**, *Image Vision and Computing* 21 (2003) 759-776.
- CHU, A., SEHGAL, C. M., GREENLEAF, J.F. **Use of gray level distribution of run lengths for texture analysis**, *Pattern Recognition Letters* 11 (1990) 415-420.
- CONNERS, R. W., HARLOW, C. A. **A theoretical comparison of texture algorithms**, *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-2:2 (1980) 204-222.
- DASARATHY, B. V., HOLDER, E. B. **Image characterizations based on joint gray level-run length distributions**, *Pattern Recognition* 12 (1991) 497-502.
- DIMAI, A. **Rotation invariant texture description using general moment invariants and Gabor filters**, *Proceedings of the 11th Scandinavian Conference on Image analysis* (1999) 391-398, vol I, June.
- DENG, Y., MANJUNATH, B. S., KENNEY C., MOORE, M. S., SHIN H. **An efficient color representation for image retrieval**, *IEEE transactions on Image Processing* 10:1 (2001) 140-147.
- DUDA, Richard O., HART, Peter E., STORK, David G. **Pattern classification**. 2 ed. New York : John Willey & Sons, 2000.
- FOUNTAIN, S. R., TAN, T. N. **Efficient rotation invariant texture features for content-based image retrieval**, *Pattern Recognition* 31:11 (1998) 1725-1732.
- GALLOWAY, M. M. **Texture analysis using gray level run lengths**, *Computer Graphics and Image Processing* 4 (1975) 172-179.
- GONDRA, D. R., HEISTERKAMP, D. R. **Content-based image retrieval with the normalized information distance**, *Computer Vision and Image Understanding* (2008), doi:10.1016/j.cviu.2007.11.001.
- GONZALEZ, R. C., WOODS, R. E. **Processamento de imagens digitais**. São Paulo : Editora Edgard Blücher, 2000.

- HAFNER, J., SAWHNEY, H. S., EQUITZ, W., FLICKNER, M., NIBLACK, W. **Efficient color histogram indexing for quadratic form distance functions**, *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-17:7 (1995) 729-736.
- HARALICK, R., SHANMUGAM, K., DINSTEIN, I. **Textural features for image classification**, *IEEE Transactions on Systems, Man and Cybernetics* SMC-3:6 (1973) 610-621.
- HARALICK, R. **Statistical and structural approaches to texture**, *Proceedings of the IEEE* 67:5 (1979) 786-804.
- HARWOOD, D., OJALA, T., PIETIKÄINEN, M., KELMAN, S., DAVIS, L. **Texture classification by center-symmetric auto-correlation, using Kullback discrimination of distributions**, *Pattern Recognition Letters* 16 (1995) 1-10.
- HAYES, Jr., K. C., SHAH, A. N., ROSENFELD, A. **Texture coarseness: further experiments**, *IEEE Transactions on Systems, Man and Cybernetics* SMC-4 (1974) 467-472.
- HE, D., WANG, L. **Texture unit, texture spectrum and texture analysis**, *IEEE Transactions on Geoscience and Remote Sensing* 28:4 (1990) 509-512.
- HE, D., WANG, L. **Texture features based on texture spectrum**, *Pattern Recognition* 24:5 (1991) 391-399.
- ILONEN, J., KÄMÄRÄINEN, K., KÄLVIÄINEN, H. **Efficient computation Gabor features**, *Research Report 100*, Lappeenranta University of Technology, Lappeenranta, Finlandia, 2005.
- KÄMÄRÄINEN, J., KYRKI, V., IILONEN, J., KÄLVIÄINEN, H. **Improving similarity measures of histograms using smoothing projections**, *Pattern Recognition Letters* 24 (2003) 2009-2019.
- KNUTH, D. E. **The art of computer programming**. 2 ed. Reading : Addison-Wesley, 1998. vol 3.
- LAWS, K. I. **Rapid texture identification**, *Proceedings of SPIE Conference on Image Processing for Missile Guidance* (1980) 376-380.
- LLOYD, S. P. **Least square quantization in PCM**, *IEEE Transactions on Information Theory* IT-28:2 (1982) 129-137.
- MANJUNATH, B. S., MA, W. Y. **Texture features for browsing and retrieval of image data**, *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-18:8 (1996) 837-842.
- MANJUNATH, B. S., OHM, J., VASUDEVAN, V. V., YAMADA, A. **Color and texture descriptors**, *IEEE Transactions on Circuits and Systems for Video Technology* 11:6 (2001) 703-715.
- MEHTRE, B. M., KANKANHALLI, M. S., NARASIMHALU, A. D., MAN, G. C. **Color matching for image retrieval**, *Pattern Recognition Letters* 16 (1995) 325-331.
- NEZAMABADI-POUR, H., KABIR, E. **Image retrieval using histograms of uni-color and bi-color blocks and directional changes in intensity gradient**, *Pattern Recognition Letters* 25 (2004) 1547-1557.
- OHM, J., BUNJAMIN, F., LIEBSCH, W., MAKAI, B., MÜLLER, K., SMOLIC, A., ZIER, D. **A set of visual feature descriptors and their combination in a low-level description scheme**, *Signal Processin: Image Communication* 16 (2000) 157-179.

- OJALA, T., PIETIKÄINEN, M., HARWOOD, D. **A comparative study of texture measures with classification based on feature distributions**, *Pattern Recognition* 29 (1996) 51-59.
- OJALA, T., PIETIKÄINEN, M., MÄENPÄÄ, T. **Multiresolution gray-scale and rotation invariant classification with local binary patterns**, *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-24:7* (2002) 971-987.
- OJALA, T., VALKEALAHTI, K., OJA, E., PIETIKÄINEN, M. **Texture discrimination with multidimensional distributions of signed gray-level differences**, *Pattern Recognition* 34 (2001) 727-739.
- PARKER, J. R. **Algorithms for image processing and computer vision**. New York : John Willey & Sons, 1997.
- PEDRINI, H., SCHWARTZ, W. R. **Análise de imagens digitais**. São Paulo : Thomson Learning, 2008.
- PIETIKÄINEN, M., OJALA, T., XU, Z. **Rotation-invariant texture classification using feature distributions**, *Pattern Recognition* 33 (2000) 43-52.
- PO, L., WONG, K. **A new palette histogram similarity measure for MPEG-7 dominant color descriptor**, *Proceedings of the IEEE International Conference on ImageProcessing* (2004) 1533-1536.
- POOLE, D. **Álgebra linear**. São Paulo : Pioneira Thomson Learning, 2004.
- RANDEM, T., HUSØY, J. H. **Filtering for texture classification: a comparative study**, *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-21:4* (1999) 291-310.
- REED, T. R., WECHSLER, H. **Segmentation of textured images and gestalt organization usin spatial/spatial-frequency representations**, *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-12:1* (1990) 1-12.
- SALOMON, D. **Data compression: the complete reference**. 3 ed. New York : Springer-Verlag, 2004.
- SERATOSA, F., SANFELIU, A. **Signatures versus histograms: distances and algorithms**, *Pattern Recognition* 39 (2006) 921-934.
- SIMPLIcity – **Image database**. James Z. Wang Research Group. College of Information Sciences and Technology. The Pennsylvania State University. Acessível em <http://wang.ist.psu.edu/~jwang/test1.tar>. 01/01/2008.
- SIPI-USC - **Texture database**. Signal and Image Processing Institute. University of Southern California, SIPI-USC,. Acessível em <http://sipi.usc.edu/database/database.cgi?volume=textures>. 01/01/2008.
- SIKORA, T. **The MPEG-7 visual standard for content description – an overview**, *IEEE Transactions on Circuits and Systems for Video Technology* 11:6 (2001) 696-702.
- TAMURA, H., MORI, S., YAMAWAKI, T. **Textural features corresponding to visual perception**, *IEEE Transactions on Systems, Man and Cybernetics SMC-8:6* (1978) 460-473.
- TEXTURAS BRODATZ - **Brodatz Textures**. Página de Trygve Randen. University of Stavanger. Norway. Acessível em <http://www.ux.uis.no/~tranden/brodatz.html>. 01/01/2008.

- UNSER, M. **Sum and difference histograms for texture classification**, *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8:1 (1986) 118-125.
- VisTex – **Vision Texture database**. MediaLab. MIT. Acessível em <http://vismod.media.mit.edu/pub/VisTex/>. 01/01/2008.
- WU, P., MANJUNATH, B. S., NEWSAM, S., SHIN, H. D. **A texture descriptor for browsing and similarity retrieval**, *Signal Processing: Image Communication* 16 (2000) 33-43.
- YANG, N., HANG, W., KUO, C., LI, T. **A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval**, *Journal of Visual Communication and Image Representation* 19 (2008) 92-105.