**UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS**

Osmando Pereira Junior

**Novos operadores de fusão aplicados a descritores de
textura**

São Carlos

2019

Osmando Pereira Junior

**Novos operadores de fusão aplicados a descritores de
textura**

Tese apresentada à Escola de Engenharia de
São Carlos da Universidade de São Paulo,
para obtenção do título de Doutor em Ciên-
cias - Programa de Pós-Graduação em Enge-
nharia Elétrica.

Área de concentração: Processamento de Si-
nais e Instrumentação

Orientador: Prof. Dr. Adilson Gonzaga

**Trata-se da versão corrigida da tese. A versão original se encontra
disponível na EESC/USP que aloja o Programa de Pós-Graduação
de Engenharia Elétrica.**

**São Carlos
2019**

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da
EESC/USP com os dados inseridos pelo(a) autor(a).

P436n Pereira Junior, Osmando
Novos operadores de fusão aplicados a descritores
de textura / Osmando Pereira Junior; orientador Adilson
Gonzaga. São Carlos, 2019.

Tese (Doutorado) - Programa de Pós-Graduação em
Engenharia Elétrica e Área de Concentração em
Processamento de Sinais e Instrumentação -- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2019.

1. Operadores de fusão. 2. Descritores de textura.
3. Fusão de descritores. 4. Fusão de decisão. 5.
Extração de características. 6. Reconhecimento de
imagens. 7. Registro de imagens. 8. Distorção
arquitetural mamária. I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **OSMANDO PEREIRA JUNIOR.**

Título da tese: "Novos operadores de fusão aplicados a descritores de textura".

Data da defesa: 22/11/2019

Comissão Julgadora:

Prof. Associado **Adilson Gonzaga**
(Orientador)
(Escola de Engenharia de São Carlos/EESC)

Prof. Dr. **Ricardo José Ferrari**
(Universidade Federal de São Carlos/UFSCar)

Prof. Dr. **Aparecido Nilceu Marana**
(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP – Bauru)

Prof. Associado **Rudinei Goularte**
(Instituto de Ciências Matemáticas e de Computação/ICMC-USP)

Prof. Associado **Marcelo Andrade da Costa Vieira**
(Escola de Engenharia de São Carlos/EESC)

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Associado **Ivan Nunes da Silva**

Presidente da Comissão de Pós-Graduação:
Prof. Titular **Murilo Araujo Romero**

Resultado:

APROVADO

APROVADO

APROVADO

APROVADO

APROVADO

*Dedico este trabalho à minha família,
de ontem, de hoje e de amanhã.*

AGRADECIMENTOS

Sou devedor de amor e de gratidão para com todos!

Ao querido Pai eterno, agradeço pela existência e pelas incontáveis oportunidades de evolução.

A meus pais, Sônia e Osmando, agradeço pelo bilhete premiado de vinda ao planeta Terra, pelo amor infinito com o qual alimentaram minh'alma, desde minha tenra idade, pelo exemplo com o qual talharam minha personalidade e pelas gotas de suor derramadas em prol de que eu pudesse sempre me banhar nas límpidas águas do conhecimento.

A meus irmãos, Rafinha e Lucklu, por me inspirarem, por me ensinarem a arte da felicidade e por contribuírem com meu crescimento. Sou infinitamente grato.

À minha amada esposa, Camila, por estar ao meu lado durante o período de tormentas do processo de doutoramento, por participar da consagração de uma obra conjunta, intitulada *família nossa*, e por acalantar meu coração, trazendo cor e harmonia para minha vida.

À vovó Rosa e ao vovô José, pelo amor e por valorizarem a educação. À vovó Lola e ao vovô João, por ensinarem o valor da simplicidade e da honestidade.

Ao tio Libânio, por ser amigo e generoso, por ensinar o valor de cada segundo, por fazer o que ama e por amar o que faz.

À tia Rosa, agradeço pela inspiração, pelo primor, pelas cartas enviadas durante minha infância, incentivando-me a acreditar em meus sonhos, e pela doce acolhida no Ninho do Sol durante o meu doutoramento.

À minha madrinha, Sônia, por ensinar que a educação é feita com liberdade e com estímulo.

À tia Cleuza Helena Gonçalves, agradeço pelos telefonemas, pelas mensagens, pelo incentivo e pelo carinho.

Aos meus queridos primos, Guilherme, irmão da vida eterna, e Josiane, agradeço imensamente pelo carinho com que me receberam e por todo o apoio prestado durante o período em que estive pelas bandas de São Paulo.

A toda minha grande e amada família, de hoje e de ontem, avós, tios, primos, sobrinhos e amigos, pelo amor, pelos sorrisos e pelos abraços.

Ao mestre, prof. Dr. Adilson Gonzaga, por orientar-me nesta formação, por fazer-me pesquisador, pela paciência e presteza, pelas conversas agradáveis, pelo ensinamento inestimável e pela memorável biblioteca.

Ao prof. Dr. Marcelo Andrade da Costa Vieira, por contribuir com o enriquecimento deste trabalho, e aos professores Evandro Luís Linhari Rodrigues e Rudinei Goularte por se disporem a participar da banca de qualificação, ocasião em que teceram contribuições relevantes e valiosas para o desenvolvimento da pesquisa. Agradeço novamente aos professores Marcelo e Rudinei, assim como aos professores Aparecido Nilceu Marana e Ricardo José Ferrari, por participarem da banca de defesa da tese de doutorado, contribuindo significativamente para a melhoria do trabalho.

Sou grato aos amigos e aos colegas do LAVI pela companhia agradável, em especial à Carol e ao Helder, pela construção do conhecimento conjunto, pelas rodas de conversas esclarecedoras e por todo o apoio prestado durante o desenvolvimento do doutorado, principalmente nos instantes finais. Obrigado meus amigos.

Agradeço ainda aos queridos amigos da cidade de Araraquara, que contribuíram para o despertar da consciência. Sou muito feliz por esse reencontro! Assim como aos eternos amigos de Luz, que me amparam e me fortalecem durante todos os instantes de minha existência.

Por fim, agradeço a todos os servidores da EESC pelo apoio despendido durante estes poucos, mas intensos, anos em que estive na USP de São Carlos.

Muito obrigado, a todos!

“Lokāḥ Samastāḥ Sukhino Bhavantu”
“Que todos os seres sejam livres e felizes!”
(Autor desconhecido)

RESUMO

PEREIRA JR., O. **Novos operadores de fusão aplicados a descritores de textura**. 2019. 144p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2019.

A análise de texturas é fundamental em diversas aplicações de visão computacional e de reconhecimento de padrões. Diversos descritores de textura são propostos na literatura, com a finalidade de se representar adequadamente as imagens de textura, sendo a extração de características uma etapa essencial nesta tarefa. No entanto, o desempenho dos descritores está relacionado ao tipo de imagem em que são aplicados, não havendo um descritor que garanta o melhor resultado para todo conjunto de imagens. Procedimentos como a fusão da informação permitem obter resultados melhores que os obtidos com a aplicação das respectivas técnicas individualmente. A fim de contribuir para a melhor representação de texturas, esta tese propõe três novos descritores de textura baseados em fusão de características, *Completed Mean Local Mapped Pattern* (CMLMP), *Completed Median Local Mapped Pattern* (CMedianLMP) e *Completed Z with Tilted Z Local Mapped Pattern* (CZTZLMP), fundamentados na associação da metodologia de extração de informações complementares de uma textura pelas componentes Sinal, Magnitude e Centro, com uma regra específica de determinação das diferenças de nível de cinza de uma vizinhança. Propõe também dois novos operadores de fusão, $Soma^M$ e $Gram^M$ no espaço de fusão, cujas regras de combinação de características evidenciam as informações complementares entre os diferentes descritores combinados. Os métodos propostos foram aplicados a registro de imagens e a reconhecimento de distorção arquitetural mamária em mamografias digitais. Os respectivos desempenhos foram comparados aos de diversos descritores de textura e a operadores de fusão publicados na literatura. De acordo com os resultados obtidos, os descritores propostos apresentaram desempenho superior ao dos demais descritores com os quais foram comparados, o que incentiva a utilização dos mesmos em outras aplicações. Além disso, os operadores de fusão propostos permitiram resultados melhores que os obtidos com a aplicação dos demais operadores e resultaram em vetores de características de menor dimensão.

Palavras-chave: Operadores de fusão. Descritores de textura. Fusão de descritores. Fusão de características. Fusão de decisão. Extração de características. Reconhecimento de imagens. Registro de imagens. Distorção arquitetural mamária. Mamografia digital.

ABSTRACT

PEREIRA JR., O. **Novel fusion operators applied to texture descriptors**. 2019. 144p. Thesis (Doctoral Degree) - São Carlos School of Engineering, University of São Paulo, São Carlos, 2019.

Texture analysis is essential in computer vision and pattern recognition. Several texture descriptors are proposed in order to properly represent texture images. However, the descriptors performance is dependent on the images. It is a hard task to develop an ideal descriptor as there is not a descriptor that guarantees the best result for the whole set of images. Information fusion achieves better results than individual sources. In order to contribute to texture representation, we propose three novel fusion-based texture descriptors, namely: Completed Mean Local Mapped Pattern (CMLMP), Completed Median Local Mapped Pattern (CMedianLMP) and Completed Z with Tilted Z Local Mapped Pattern (CZTZLMP), based on Signal, Magnitude and Centre complementary information with a specific rule for determining gray level differences in a neighborhood. To combine features from texture descriptors, this work also proposes two fusion operators, $\text{Sum}^{\mathbb{M}}$ and $\text{Gram}^{\mathbb{M}}$. The proposed methods were applied to image registration and to the detection of breast architectural distortion in digital mammography, and then compared to several texture descriptors and fusion operators published in the literature. The experimental results show that the proposed fusion-based texture descriptors outperformed all other descriptors, which encourages their implementation in other applications. Moreover, the proposed fusion operators outperformed all other operators, and generated smaller feature vectors.

Keywords: Fusion operators. Texture descriptors. Feature fusion. Decision fusion. Feature extraction. Image recognition. Image registration. Breast architectural distortion. Digital mammography.

LISTA DE FIGURAS

Figura 1 – Evolução da percepção e da representação de textura (modificado de (LIU et al., 2019)).	35
Figura 2 – (a) Distâncias d entre os pares de <i>pixels</i> e (b) orientações $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ a se considerar para determinação das respectivas matrizes de coo- corrência de nível de cinza $\mathbf{GLCM}_{d,\theta}$ (modificado de (HARALICK; SHANMUGAM; DINSTEIN, 1973)).	37
Figura 3 – (a) Vizinhança 3×3 e (b) correspondente vizinhança circular de raio $R = 1$ e $N = 8$ em torno do pixel central $p(x, y)$	40
Figura 4 – Diferentes configurações de vizinhança circular para análise por multire- solução. (a) $R = 1$ e $N = 8$, (b) $R = 2$ e $N = 16$ e (c) $R = 3$ e $N = 24$ (modificado de (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002)).	41
Figura 5 – Comparação entre as regras de extração de características dos descritores LBP e CS-LBP, considerando uma vizinhança de raio $R = 1$ com oito <i>pi- xels</i> igualmente espaçados (modificado de (HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009)).	42
Figura 6 – Vizinhança 3×3 para extração da característica ZTZLBP para os <i>pixels</i> dispostos (a) no arranjo em Z e (b) no arranjo em Z invertido (TZ).	45
Figura 7 – Comportamento da função de mapeamento $f(g_{diff})$ para diferentes valores de β (modificado de (NEGRI et al., 2017)).	48
Figura 8 – Metodologias de determinação da intensidade de nível de cinza para comparação com os pixels da vizinhança na determinação dos códigos LMP e Median-LMP (modificado de (SOUZA; GONZAGA, 2019)).	52
Figura 9 – Vizinhança 3×3 em torno do <i>pixel</i> $p(x, y)$ (modificado de (SOUZA; GONZAGA, 2019)).	53
Figura 10 – Classificação das técnicas de fusão por níveis (adaptado de (DONG et al., 2009)).	56
Figura 11 – Exemplo de fusão por Soma de dois vetores de características.	60
Figura 12 – Exemplo de fusão por Concatenação de dois vetores de características.	61
Figura 13 – Exemplo de fusão por Combinação Conjunta de duas imagens de códigos.	62
Figura 14 – Vizinhança circular de raio $R = 1$ e $N = 8$ para o <i>pixel</i> p posicionado em (x, y)	67
Figura 15 – Vizinhança 3×3 e vizinhança circular de raio $R = 1$ e $N = 8$ para o <i>pixel</i> p posicionado em (x, y)	69
Figura 16 – Diferentes arranjos da vizinhança circular de raio $R = 1$ e $N = 8$ para extração das características CZTZLMP: (a) arranjo em Z e (b) arranjo em Z inclinado (TZ).	71

Figura 17 – K arranjos da vizinhança circular para extração das características CZTZLMP considerando (a) $R = 1$ e $N = 8$, (b) $R = 2$ e $N = 16$ e (c) $R = 3$ e $N = 24$	72
Figura 18 – Exemplo de fusão por Soma ^M de três vetores de características.	75
Figura 19 – Exemplo de fusão por Gram ^M de três vetores de características.	77
Figura 20 – Imagens do conjunto de treinamento: bark e boat (escala e rotação), bike e trees (imagem borrada), graf e wall (mudança de ponto de vista), leuven (mudança de iluminação), ubc (compressão JPEG).	81
Figura 21 – Imagens do conjunto de teste: asterix, belledone, bip e crolle (escala); east park, east south, ensigma e inria (escala e rotação).	82
Figura 22 – Imagens do conjunto de teste: lap e vangogh (escala); laptop e resid (escala e rotação); new york, zmars e zmonet (rotação).	83
Figura 23 – Mamografia digital com ocorrência de distorção arquitetural, sinalizada por radiologista com um círculo vermelho (extraído de (OLIVEIRA et al., 2019)).	104
Figura 24 – Extração das regiões de interesse em mamografia digital com distorção arquitetural (adaptado de (OLIVEIRA et al., 2019)).	104
Figura 25 – Exemplos de amostras extraídas de diferentes mamografias (a) com DA e (b) sem DA.	105
Figura 26 – Metodologia proposta de reconhecimento de distorção arquitetural. . .	106
Figura 27 – Método de validação cruzada por unidade, em que o banco de amostras consiste em todas as amostras de mamografia, com exceção das $N-1$ ROIs pertencentes à mesma mamografia que a amostra de consulta. . .	108
Figura 28 – Acurácia obtida com a aplicação do descritor (a) CMLMP_S, (b) CMLMP_M, e (c) CMLMP_C, com a variação de β e B	111
Figura 29 – Comportamento da função de mapeamento para o descritor CMLMP para β_S , β_M e β_C ótimos, com a variação das diferenças g_{diff} no intervalo (a) $[-100, 100]$ e (b) $[-2, 2]$	112
Figura 30 – Acurácia dos descritores individuais para diferentes valores de K , considerando $R = 1$ e $N = 8$	113
Figura 31 – Imagens de códigos obtidas com a aplicação dos descritores CMLMP, CLMP e CLBP sobre uma amostra com DA.	114
Figura 32 – Tempo médio de processamento dos descritores individuais (a) para geração dos vetores de características, em ms, e (b) para cálculo da distância euclidiana entre os vetores de características de duas amostras de mamografia digital, em μs	125
Figura 33 – Tempo médio de execução dos operadores de fusão.	126

Figura 34 – Tempo médio de processamento das melhores abordagens de fusão (a) para geração dos vetores de características, em ms, e (b) para cálculo da distância euclidiana entre os vetores de características de duas amostras de mamografia digital, em μs 127

LISTA DE TABELAS

Tabela 1 – Número de regiões de interesse detectadas pelo detector <i>Hessian-Affine</i> e de regiões correspondentes e não correspondentes, previamente conhecidas, por par de imagens do conjunto de treinamento.	89
Tabela 2 – Número de regiões de interesse detectadas pelo detector <i>Hessian-Affine</i> e de regiões correspondentes e não correspondentes, previamente conhecidas, por imagem do conjunto de teste.	90
Tabela 3 – AUC (%) para os descritores individuais aplicados sobre o conjunto de treinamento.	91
Tabela 4 – AUC (%) para as nove melhores abordagens de fusão aplicadas sobre o conjunto de treinamento.	91
Tabela 5 – Ganho (%) com a aplicação da abordagem de fusão $M-LMP \oplus (Z \oplus TZLBP)$ sobre o conjunto de treinamento.	92
Tabela 6 – AUC (%) para os melhores descritores individuais aplicados sobre o conjunto de teste.	93
Tabela 7 – AUC (%) para as três melhores abordagens de fusão da camada de características aplicadas sobre as imagens do conjunto de teste.	94
Tabela 8 – AUC (%) para as seis melhores abordagens de fusão de decisão aplicadas sobre as imagens do conjunto de teste.	94
Tabela 9 – Ganho (%) com a aplicação da abordagem de fusão $M-LMP \oplus CS-LBP$ sobre o conjunto de teste.	95
Tabela 10 – Tempo médio de processamento.	96
Tabela 11 – Diferentes abordagens para detecção de distorção arquitetural em mamografia	100
Tabela 12 – Parâmetros de cada descritor otimizados para a base de dados de mamografia digital	110
Tabela 13 – Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos descritores individuais sobre o banco de mamografias digitais (valores em %)	115
Tabela 14 – Ganho (%) com a aplicação do descritor individual proposto CMedianLMP_C sobre o banco de mamografia digital	117
Tabela 15 – Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($\parallel$), $Gram^M(\parallel_G)$ e $Soma^M(\parallel_+)$ sobre o descritor CMLMP na camada de características (valores em %)	118

Tabela 16 – Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($\parallel$), Gram ^M ($\parallel_G$) e Soma ^M ($\parallel_+$) sobre o descritor CMedianLMP na camada de características (valores em %)	119
Tabela 17 – Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($\parallel$), Gram ^M ($\parallel_G$) e Soma ^M ($\parallel_+$) sobre o descritor CZTZLMP na camada de características (valores em %)	120
Tabela 18 – Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($\parallel$), Gram ^M ($\parallel_G$) e Soma ^M ($\parallel_+$) sobre os descritores individuais CLMP, CLBP e WLD na camada de características (valores em %)	121
Tabela 19 – Ganho (%) com a aplicação da abordagem de fusão proposta CMedianLMP_ _{S_{3,24}} $\parallel_+$ CMedianLMP_ _{M_{2,16}} $\parallel_+$ CMedianLMP_ _C sobre o banco de mamografia digital	122
Tabela 20 – Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($\parallel$), Gram ^M ($\parallel_G$) e Soma ^M ($\parallel_+$) sobre os diferentes descritores na camada de características (valores em %)	123
Tabela 21 – Ganho (%) com a aplicação da abordagem de fusão proposta GLCM $\parallel_+$ CMedianLMP_ _{S_{3,24}} $\parallel_+$ CLMP_ _{M_{2,16}} $\parallel_+$ CMLMP_ _C sobre o banco de mamografia digital	124

LISTA DE ABREVIATURAS E SIGLAS

Acc	Acurácia
AUC	Área sobre a curva
CAD	Sistema computacional de auxílio ao diagnóstico
CC	Craniocaudal
CLBP	<i>Completed Local Binary Pattern</i>
CLMP	<i>Completed Local Mapped Pattern</i>
CMLMP	<i>Completed Mean Local Mapped Pattern</i>
CMedianLMP	<i>Completed Median Local Mapped Pattern</i>
CNN	Rede Neural Convolucional
CZTZLMP	<i>Completed Z with Tilted Z Local Mapped Pattern</i>
DA	Distorção arquitetural
E	Especificidade
FFDM	Mamografia digital de campo total
GLCM	Matriz de Coocorrência de Níveis de Cinza
LAVI	Laboratório de Visão Computacional
LBP	<i>Local Binary Pattern</i>
LMP	<i>Local Mapped Pattern</i>
M-LMP	<i>Mean Local Mapped Pattern</i>
MLO	Mediolateral-oblíqua
Median-LMP	<i>Median Local Mapped Pattern</i>
ROI	Região de interesse
S	Sensibilidade
SIFT	<i>Scale Invariant Feature Transform</i>
SVM	<i>Support Vector Machine</i>

TU	<i>Texture Unit</i>
WLD	<i>Weber Local Descriptor</i>
ZTZLBP	<i>Z with Tilted Z Local Binary Pattern</i>

LISTA DE SÍMBOLOS

β	Inclinação da curva sigmóide
B	Valor máximo do código de um determinado descritor na imagem de códigos
$\mathbb{M}$	Espaço de fusão
N_g	Quantidade de níveis de cinza da imagem de textura
$\mathbf{W}$	Matriz de pesos
$\oplus$	Operador de fusão Concatenação
$\ _+$	Operador de fusão Soma no espaço de fusão
$\ _G$	Operador de fusão Gram no espaço de fusão
$ $	Operador de fusão Combinação Conjunta

SUMÁRIO

1	CONTEXTO	29
1.1	Introdução	29
1.2	Motivação	30
1.3	Objetivos e Contribuições	31
1.4	Organização da tese	32
2	DESCRITORES DE TEXTURA	35
2.1	Matriz de Coocorrência dos Níveis de Cinza (GLCM)	36
2.2	<i>Scale Invariant Feature Transform (SIFT)</i>	38
2.3	<i>Local Binary Pattern (LBP)</i>	39
2.4	<i>Center-Symmetric Local Binary Pattern (CS-LBP)</i>	41
2.5	<i>Completed Local Binary Pattern (CLBP)</i>	43
2.6	<i>Z with Tilted Z Local Binary Pattern (ZTZLBP)</i>	45
2.7	<i>Local Mapped Pattern (LMP)</i>	47
2.8	<i>Center-Symmetric Local Mapped Pattern (CS-LMP)</i>	48
2.9	<i>Mean Local Mapped Pattern (M-LMP)</i>	49
2.10	<i>Completed Local Mapped Pattern (CLMP)</i>	50
2.11	<i>Median Local Mapped Pattern (Median-LMP)</i>	51
2.12	Considerações Finais	53
3	FUSÃO DA INFORMAÇÃO	55
3.1	Considerações Iniciais	55
3.2	Abordagens de Fusão aplicadas a Visão Computacional	57
3.3	Fusão de descritores de textura na etapa de Características	58
3.3.1	Soma de vetores de características	59
3.3.2	Concatenação de vetores de características	59
3.3.3	Combinação Conjunta de imagens de códigos	60
3.4	Fusão de descritores de textura na etapa de Decisão	61
3.4.1	Operador de fusão Adição	62
3.4.2	Operadores de fusão Fuzzy	63
3.4.3	Inferência Bayesiana	64
3.5	Considerações Finais	64
4	MÉTODOS PROPOSTOS	65
4.1	Descritores de textura baseados em fusão de características	65
4.1.1	<i>Completed Mean Local Mapped Pattern (CMLMP)</i>	66

4.1.2	<i>Completed Median Local Mapped Pattern (CMedianLMP)</i>	68
4.1.3	<i>Completed Z with Tilted Z Local Mapped Pattern (CZTZLMP)</i>	70
4.2	Operadores de Fusão	74
4.2.1	Espaço de Fusão de Características	74
4.2.2	Operador Soma no espaço de fusão	75
4.2.3	Operador Gram no espaço de fusão	76
4.3	Considerações Finais	78
5	FUSÃO DE DESCRITORES DE TEXTURA APLICADA A REGIS- TRO DE IMAGENS	79
5.1	Considerações iniciais	79
5.2	Materiais e Métodos	80
5.2.1	Definição dos Conjuntos de Imagens	80
5.2.2	Detecção das Regiões de Interesse	81
5.2.3	Determinação do <i>Ground Truth</i>	82
5.2.4	Aplicação dos Descritores de Textura	84
5.2.5	Métrica de Similaridade	85
5.2.6	Apresentação dos Resultados - Descritores Aplicados Individualmente	86
5.2.7	Fusão dos Descritores de Textura	87
5.2.7.1	Fusão de Descritores de Textura na Camada de Características	87
5.2.7.2	Fusão de Descritores de Textura na Camada de Decisão	88
5.3	Resultados	89
5.3.1	Resultados para o Conjunto de Treinamento	90
5.3.2	Validação dos Resultados: Aplicação Sobre o Conjunto de Teste	92
5.3.3	Análise do Custo Computacional	95
5.4	Considerações Finais	96
6	FUSÃO DE DESCRITORES DE TEXTURA APLICADA A RECO- NHECIMENTO DE DISTORÇÃO ARQUITETURAL EM MAMO- GRAFIA DIGITAL	99
6.1	Considerações Iniciais	99
6.2	Materiais e Métodos	103
6.2.1	Base de Mamografia Digital	103
6.2.2	Extração de Características e Fusão das Características	105
6.2.3	Classificação	108
6.3	Resultados	110
6.3.1	Otimização dos Parâmetros	110
6.3.2	Desempenho para Descritores Individuais	112
6.3.3	Desempenho para Descritores Baseados em Fusão de Características	117
6.3.4	Fusão de Diferentes Descritores de Textura	122

6.3.5	Análise do Custo Computacional	124
6.4	Considerações Finais	128
7	CONCLUSÕES	129
7.1	Principais Contribuições	132
7.2	Sugestões para Trabalhos Futuros	132
7.3	Publicações Oriundas do Trabalho	133
7.3.1	Premiações Oriundas do Trabalho	133
	 REFERÊNCIAS	 135

1 CONTEXTO

1.1 Introdução

A análise de textura é fundamental em diversas aplicações de visão computacional e de reconhecimento de padrões (LIU et al., 2019), sendo aplicada a reconhecimento de objetos (GU et al., 2016; WANG et al., 2011), registro de imagens (PAUL; PATI, 2017; DELLINGER et al., 2015; WANG; YOU; FU, 2015), classificação de texturas (VIEIRA; NEGRI; GONZAGA, 2018; GUO; ZHANG; ZHANG, 2010; ZHANG et al., 2007), predição de movimento em vídeo (BOONTHEP; CHAMNONGTHAI; PHENSADSA-ENG, 2018), detecção de objetos em imagens de vídeo (BERE, 2018), classificação de elementos na constituição de ligas de metais (WEBEL et al., 2018), reconhecimento de íris (SOUZA; GONZAGA, 2019), segmentação e classificação de imagens médicas (OWJI-MEHR; DANYALI; HELFROUSH, 2014), reconhecimento de expressões faciais (SINGH; MAURYA; MITTAL, 2012), sensoriamento remoto (CHEN et al., 2016), classificação de recifes de corais (MARY; DEJEY, 2018), dentre outras.

A textura se destaca dentre outras características inerentes de imagens, como tonalidade, cor e forma (KIM; YOO; SOHN, 2013; GUO; ZHAO; PIETIKÄINEN, 2012), por possuir informação de distribuição da intensidade dos níveis de cinza da imagem, assim como informação do valor das intensidades dos níveis de cinza (HARALICK; SHANMUGAM; DINSTEN, 1973). Diversos descritores de textura são propostos na literatura, com a finalidade de se representar adequadamente as imagens de textura, sendo a extração de características etapa essencial nesta tarefa (KIM; YOO; SOHN, 2013).

Cada descritor de textura aplica um método diferente de extração de características. Ojala, Pietikäinen e Harwood (1996) discorrem sobre a possibilidade de uma única medida de textura não possuir informação suficiente para descrever adequadamente o arranjo espacial e a distribuição de intensidade de nível de cinza de texturas locais, e afirmam que a análise de ocorrências conjuntas de duas ou mais características poderia ser mais discriminante para texturas.

Huang, Chan e Zhang (2002) afirmam que texturas naturais geralmente são constituídas por componentes estatística e estrutural, sendo assim, a utilização de um descritor de uma única abordagem não seria suficiente para descrever completamente imagens de textura.

Com a finalidade de se caracterizar melhor as diferentes informações das texturas, foram propostas diversas técnicas de extração de características complementares por um mesmo descritor (KIM; YOO; SOHN, 2013; CHEN et al., 2009). Guo, Zhang e Zhang (2010) propuseram o descritor *Completed Local Binary Pattern* (CLBP), o qual extrai três

diferentes informações da imagem de textura, sendo duas delas referentes à diferença local de intensidade de nível de cinza, e a terceira, considera uma informação global da imagem. Devido ao ganho de desempenho aportado pelo CLBP, algumas variações do método foram propostas, como o MS-CLBP (CHEN et al., 2016), que considera múltiplas representações da imagem original, cada uma em uma escala diferente, e extrai as informações complementares da imagem nas diversas escalas. Negri et al. (2017) propuseram o descritor *Color Intensity Local Mapped Pattern* (CILMP), composto por duas componentes, uma referente à informação de luminância e a outra, às características no espaço de cor RGB, sendo a abordagem referenciada por caracterizar a imagem com base em informações de textura da cor.

Diferentemente da abordagem de desenvolvimento de um único descritor que caracterize adequadamente a textura por meio da extração das informações complementares, também foram propostas abordagens baseadas na fusão das informações de diferentes descritores (WANG; PENG; ZHANG, 2017; CAMPOS; CORREIA; CALADO, 2015; SHIHAVUDDIN et al., 2013; NANNI; LUMINI, 2013; YUSHKEVICH et al., 2012). Na literatura, encontram-se referenciados diversos algoritmos e métodos de fusão da informação, principalmente de fusão de decisões, os quais proporcionam melhor desempenho a sistemas de processamento de imagens (GHASSEMIAN, 2016).

1.2 Motivação

Existem diversas aplicações em sistemas de visão computacional cuja solução representa um desafio. A determinação dos pares de *pixels* correspondentes, ou de regiões correspondentes, é uma delas (BAUMBERG, 2000; MARSHALL; DOUCETTE, 2011; LUTZ et al., 2015). Esta correspondência de imagens, também denominada registro de imagens, tem vasta aplicação em sistemas de reconstrução de informação tridimensional de uma cena, robótica móvel, rastreamento de objetos, sensoriamento remoto, inspeção industrial e controle de qualidade, biometria facial, vigilância e segurança, dentre outras (LUTZ et al., 2015; NANNI; LUMINI, 2013; MARSHALL; DOUCETTE, 2011; MA; CHAN; CANTERS, 2010; HOSSAIN et al., 2010; HUANG et al., 2008).

As imagens a serem analisadas podem ter sido capturadas em diferentes instantes de tempo, de diferentes pontos de vista e/ou por diferentes câmeras (LUTZ et al., 2015; HUANG et al., 2008), sendo assim, podem apresentar diversas transformações, como rotação, variação de iluminação, compressão, mudança de ponto de vista, e outras, sendo necessário que os descritores de características sejam invariantes a tais transformações. Dificilmente, um único descritor de características será suficientemente robusto a uma grande quantidade de transformações. Desse modo, a fusão de descritores pode contribuir com a melhoria do sistema de registro de imagens.

Outra aplicação de grande relevância consiste no reconhecimento de distorção

arquitetural mamária em imagens de mamografia digital. Dentre os indícios de formação de tumor na mama, a distorção arquitetural mamária é o que pode ser detectado em mamografias no estágio mais precoce da doença (BAHL et al., 2015; GAUR et al., 2013), sendo assim, a detecção de distorção arquitetural mamária caracteriza-se por ser uma tarefa essencial para o tratamento do câncer de mama (VERONESI et al., 2005; ELMORE et al., 2003). No entanto, devido às características da distorção arquitetural mamária, que é uma contração sutil no tecido mamário, sem borda definida e sem formação de massa (SULEIMAN et al., 2016; D'ORSI, 2013), o reconhecimento de distorção arquitetural mamária em mamografias não é evidente.

Diversas soluções baseadas em descritores de textura são propostas na literatura para essa tarefa (LIU et al., 2018; OLIVEIRA et al., 2018; CASTI et al., 2016), no entanto, ainda apresentam baixo desempenho, sendo necessária a investigação de outras abordagens. Alguns dos métodos que consideram a fusão de características extraídas por diferentes descritores permitem a obtenção de melhores resultados (LIU et al., 2018).

Para as duas aplicações citadas, mas também para diversas outras aplicações de visão computacional, há uma grande quantidade de descritores de textura publicados na literatura, sendo que o desempenho de cada descritor depende da aplicação em específico e/ou do conjunto de imagens utilizado. Entretanto, são poucas as publicações que apresentem propostas de operadores de fusão de características, sendo que os pesquisadores da área restringem-se ao uso de duas ou três abordagens de fusão consolidadas.

Com o intuito de contribuir para a melhor representação de texturas, aborda-se, neste trabalho, o uso de novos operadores de fusão de características aplicados a descritores de textura em sistemas de registro de imagens e de reconhecimento de distorção arquitetural mamária em mamografia digital.

1.3 Objetivos e Contribuições

Esta tese tem como objetivo geral melhorar a representação de imagens de textura por meio da fusão de descritores de textura. Especificamente, apresentam-se como objetivos principais: propor novos descritores de textura baseados em fusão de características que permitam melhor representação de texturas; propor novos operadores de fusão, cuja regra de fusão das características evidencie a complementaridade entre os diferentes descritores combinados; desenvolver uma metodologia de aplicação dos descritores e dos métodos de fusão a sistemas de registro de imagens e a sistemas de reconhecimento de distorção arquitetural em mamografias digitais.

A primeira contribuição da tese foi o desenvolvimento do operador de fusão Soma no espaço de fusão, Soma^M. O operador proposto combina os vetores de características dos diferentes descritores de modo que o vetor resultante apresente pequena dimensão e

alto desempenho.

A segunda contribuição foi o desenvolvimento de um segundo operador de fusão, denominado Gram no espaço de fusão, Gram^{M} , o qual combina os vetores de características dos diferentes descritores por meio da determinação da matriz de Gram dos vetores apropriadamente dispostos no espaço de fusão.

A terceira contribuição foi o desenvolvimento do descritor de textura baseado em fusão de características *Completed Mean Local Mapped Pattern* (CMLMP). O descritor proposto é baseado na metodologia de extração de características complementares da imagem, representadas por três componentes, proposta na formulação do descritor CLBP (GUO; ZHANG; ZHANG, 2010), associada à regra de extração de características do descritor M-LMP (FERRAZ et al., 2014).

A quarta contribuição foi o desenvolvimento de um segundo descritor de textura baseado em fusão de características, denominado *Completed Median Local Mapped Pattern* (CMedianLMP), baseado na metodologia de extração de características complementares da imagem associada à regra de extração de características do descritor Median-LMP (SOUZA; GONZAGA, 2019).

A quinta contribuição foi o desenvolvimento do descritor de textura baseado em fusão de características *Completed Z with Tilted Z Local Mapped Pattern* (CZTZLMP), o qual combina a metodologia de extração de características complementares com a regra de extração de características do descritor ZTZLBP (MARY; DEJEY, 2018).

A sexta e última contribuição desta tese foi o desenvolvimento e a implementação de duas metodologias de fusão de diferentes descritores de textura, uma aplicada a registro de imagens e a segunda, a reconhecimento de distorção arquitetural mamária em imagens de mamografia digital.

1.4 Organização da tese

Esta tese está organizada da seguinte maneira: neste primeiro Capítulo, foi introduzido o tema de pesquisa, sendo apresentadas as motivações e as contribuições do trabalho. No Capítulo 2 são apresentadas as principais características e a formulação matemática de descritores de textura de diferentes abordagens, assim como o procedimento de determinação dos respectivos vetores de características. No Capítulo 3 discorre-se sobre a fundamentação teórica da fusão da informação e apresenta-se a formulação matemática de alguns operadores de fusão de características e de fusão de decisão. Enquanto no Capítulo 4 apresentam-se as propostas de três novos descritores de textura baseados em fusão de características e de dois novos operadores de fusão de características, reserva-se ao Capítulo 5, a apresentação da metodologia e dos resultados obtidos com a fusão de descritores de textura aplicada a registro de imagens, sendo apresentados no Capítulo 6 a metodologia e

os resultados obtidos com a aplicação de fusão de descritores de textura em reconhecimento de distorção arquitetural em mamografia digital. A conclusão do trabalho, as principais contribuições e as sugestões de trabalhos futuros são apresentadas no Capítulo 7.

2 DESCRITORES DE TEXTURA

Como introduzido no capítulo anterior, a análise de texturas é fundamental em diversas aplicações de visão computacional e de reconhecimento de padrões (LIU et al., 2019). Diferentes metodologias foram propostas para caracterizar e classificar texturas em imagens. Uma breve evolução da percepção e da representação de textura é ilustrada na Figura 1. De modo geral, as metodologias descrevem a textura analisando um conjunto de medidas estatísticas sobre a distribuição espacial das intensidades de nível de cinza da textura (abordagem estatística), ou então analisando o arranjo espacial de elementos ou primitivas de texturas (abordagem estrutural ou geométrica), ou ainda analisando um conjunto de propriedades ou características da imagem no domínio da frequência (abordagem espectral ou baseada em processamento de sinais) (XIE; MIRMEHDI, 2008; TUCERYAN; JAIN, 1998; GOOL; DEWAELE; OOSTERLINCK, 1985; HARALICK, 1979).

Figura 1: Evolução da percepção e da representação de textura (modificado de (LIU et al., 2019)).

Neste Capítulo, discorre-se sobre descritores de textura, sendo apresentadas as principais características e a formulação matemática de descritores de diferentes abordagens. Com a aplicação de um descritor sobre uma determinada região da imagem, gera-se um vetor de características que passa a representar a região em específico. O tamanho do vetor de características retornado depende do descritor aplicado. Espera-se que cada região diferente tenha um vetor de características particular, também diferente, e que regiões iguais tenham os mesmos vetores de características, mesmo que as imagens tenham sofrido algum tipo de transformação, sendo então desejável que os descritores sejam invariantes a mudança de iluminação, rotação, escala, mudança de ponto de vista, transformações afins, dentre outras.

2.1 Matriz de Coocorrência dos Níveis de Cinza (GLCM)

De acordo com experimentos realizados por Julesz (1962), o sistema de visão humano é sensível a estatísticas de segunda ordem da imagem (GOOL; DEWAELE; OOSTERLINCK, 1985). Texturas que possuem estatísticas de primeira ordem diferentes são facilmente reconhecidas pelo sistema de visão humano, assim como texturas com mesma estatística de primeira ordem mas diferentes estatísticas de segunda ordem, sendo que estatísticas de ordem superior são pouco notadas pelo sistema humano. No entanto, caso as texturas tenham mesmas estatísticas de primeira e de segunda ordens, elas podem ser facilmente diferenciadas pelo sistema humano se possuírem elementos de texturas diferentes, situação para a qual Julesz (1981) introduziu o conceito de *textons*. São exemplos de elementos de texturas, ou *textons*: cantos, conectividade, granularidade, terminações, e outros.

Conforme apresentado por Julesz (1975), por Tuceryan e Jain (1998) e por Xie e Mirmehdi (2008), as estatísticas de primeira ordem relacionam-se à intensidade de nível de cinza da imagem, sendo o histograma de níveis de cinza, o contraste e a intensidade média de nível de cinza da imagem exemplos de estatísticas de primeira ordem, enquanto que a distribuição das intensidades de nível de cinza na imagem, a relação entre as intensidades de nível de cinza de dois ou mais *pixels* e a coocorrência de valores de nível de cinza entre *pixels* vizinhos definem as medidas de estatísticas de segunda ordem.

A Matriz de Coocorrência de Níveis de Cinza (GLCM), método proposto por Haralick (1971), Haralick, Shanmugam e Dinstein (1973), é uma das técnicas mais conhecidas de estatística de segunda ordem para análise de texturas (XIE; MIRMEHDI, 2008; GOOL; DEWAELE; OOSTERLINCK, 1985). A matriz $\mathbf{GLCM}_{d,\theta}$ representa a ocorrência conjunta de níveis de cinza, (i, j) , entre dois *pixels* posicionados a uma distância d , um em relação ao outro, medida em uma determinada orientação θ (HARALICK; SHANMUGAM; DINSTEIN, 1973; PACI; NANNI; SEVERI, 2013).

Para a determinação da matriz de coocorrência, cada *pixel* da imagem de textura é comparado a oito *pixels* vizinhos (com exceção dos *pixels* da borda da imagem, que têm um número menor de vizinhos), dispostos em uma vizinhança $J \times J$, em que $J = 2d + 1$, tal como ilustrado na Figura 2(a). Para distância unitária, $d = 1$, a vizinhança destacada em azul, hachurada na diagonal, é utilizada na determinação das ocorrências conjuntas de nível de cinza com relação ao *pixel* central $p(x, y)$; caso d seja igual a dois, a vizinhança a se considerar é a composta pelos *pixels* destacados em alaranjado, hachurados na horizontal.

Para cada distância, é possível determinar quatro matrizes de coocorrência, uma para cada orientação: $\mathbf{GLCM}_{d,0^\circ}$, $\mathbf{GLCM}_{d,45^\circ}$, $\mathbf{GLCM}_{d,90^\circ}$ e $\mathbf{GLCM}_{d,135^\circ}$. Conforme ilustrado na Figura 2(b), os *pixels* p_1 e p_5 da vizinhança são utilizados para determinação da ocorrência conjunta a $p(x, y)$ em caso de orientação igual a 0° ; para orientação igual

Figura 2: (a) Distâncias d entre os pares de *pixels* e (b) orientações $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ a se considerar para determinação das respectivas matrizes de coocorrência de nível de cinza $\mathbf{GLCM}_{d,\theta}$ (modificado de (HARALICK; SHANMUGAM; DINSTEIN, 1973)).

a 135° , utilizam-se os *pixels* p_2 e p_6 ; os *pixels* p_3 e p_7 são considerados para orientação θ igual a 90° ; e p_4 e p_8 , para orientação de 45° .

A Equação (2.1) representa a formulação geral da matriz de coocorrência de uma imagem de textura com intensidades de nível de cinza compreendidas no intervalo $[0, N_g - 1]$. O elemento $GLCM_{d,\theta}(i, j)$, $i, j \in [0, N_g - 1]$, corresponde à quantidade de *pixels* na imagem de textura, cuja intensidade de nível de cinza é igual a i , que possui vizinho com intensidade de nível de cinza igual a j distante d *pixels* na orientação θ (HARALICK, 1971).

$$\mathbf{GLCM}_{d,\theta} = \begin{bmatrix} \#(0,0) & \#(0,1) & \#(0,2) & \cdots & \#(0,N_g-1) \\ \#(1,0) & \#(1,1) & \#(1,2) & \cdots & \#(1,N_g-1) \\ \#(2,0) & \#(2,1) & \#(2,2) & \cdots & \#(2,N_g-1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \#(N_g-1,0) & \#(N_g-1,1) & \#(N_g-1,2) & \cdots & \#(N_g-1,N_g-1) \end{bmatrix} \quad (2.1)$$

O custo computacional da determinação da matriz $\mathbf{GLCM}_{d,\theta}$ é diretamente proporcional à dimensão $X \times Y$ da imagem de textura, motivo pelo qual em determinadas aplicações primeiro realiza-se a subamostragem da imagem, para posteriormente determinar a matriz de coocorrência sobre a imagem subamostrada. Ademais, a dimensão da matriz $\mathbf{GLCM}_{d,\theta}$ é igual a $N_g \times N_g$, em que N_g é a quantidade de níveis de cinza da imagem de textura. Em algumas aplicações, para evitar esparsidade na matriz $\mathbf{GLCM}_{d,\theta}$, é comum

quantizar as imagens de textura para níveis menores de intensidade de cinza (KASSNER; THORNHILL, 2010).

Haralick propôs a caracterização da textura por meio de quatorze medidas estatísticas determinadas sobre cada matriz de coocorrência $\mathbf{GLCM}_{d,\theta}$, sendo elas energia, correlação, contraste, homogeneidade, entropia, soma das variâncias, dentre outras. O vetor de características final, constituído por 28 elementos, é determinado pela média e pelo intervalo (*range*) de cada uma das quatorze medidas obtidas considerando as quatro diferentes orientações θ . A relação completa de todas as medidas estatísticas propostas por Haralick, assim como a formulação matemática, está disponível em (KASSNER; THORNHILL, 2010; HARALICK; SHANMUGAM; DINSTEIN, 1973).

Aplicado a três diferentes conjuntos de imagens, a implementação da Matriz de Coocorrência GLCM resultou em acurácia igual a 89% para classificação de imagens de minerais, 82% de imagens de fotografias aéreas, e de 83% para identificação de imagens de satélite (HARALICK; SHANMUGAM; DINSTEIN, 1973). Webel et al. (2018) aplicaram GLCM para classificação de diferentes microestruturas em ligas de aço por meio da análise de imagens de microscópio eletrônico de varredura (MEV). As imagens em estudo foram rotacionadas de 0° a 180° , com incremento de 1° . Uma matriz de coocorrência foi gerada para cada uma das 180 imagens rotacionadas, considerando orientação $\theta = 0$, para as quais foram extraídas quatro características: contraste, energia, correlação e homogeneidade. O vetor de características final foi formado pela média e pelo intervalo de cada característica calculados sobre as 180 imagens. Os autores informaram um aumento de 15% na acurácia, atingindo acurácia média de 90%.

2.2 Scale Invariant Feature Transform (SIFT)

SIFT - *Scale Invariant Feature Transform* (LOWE, 2004) é um descritor de regiões de interesse que gera vetores de características de 128 posições, apresenta invariância a rotação, escala e transformações afins, mas apresenta baixo desempenho para sistemas que possuem variação geométrica e variação de iluminação (ALITAPPEH; MAHMOUDI, 2013; NANNI; LUMINI, 2013).

Além de descrever as regiões de interesse, este descritor também detecta os pontos de interesse (*keypoints*). Para tal, aplica-se a função gaussiana a fim de encontrar pontos que são invariantes a escala, sendo ampliada a busca por meio da aplicação da diferença de gaussianas - DoG. Após a aplicação da diferença de gaussianas, armazenam-se apenas os pontos chave estáveis, desconsiderando-se os demais. A localização do gradiente (quantizada em matrizes de dimensão 4×4) e a respectiva orientação (quantizada em oito valores) são consideradas e ponderadas por um valor de peso definido, gerando o histograma 3D do descritor. A fim de reduzir os efeitos da variação da iluminação, o descritor é normalizado para possuir norma unitária; e com o objetivo de reduzir os efeitos de elevadas magnitudes

do gradiente, o descritor é truncado em 0,2 e normalizado novamente para obter norma unitária.

O descritor SIFT é utilizado na literatura em aplicações de reconhecimento de objetos (GU et al., 2016; WANG et al., 2011; LOWE, 2004), registro de imagens (HOSAIN et al., 2010; JUAN; GWUN, 2009; MIKOLAJCZYK et al., 2005; MIKOLAJCZYK; SCHMID, 2005; KE; SUKTHANKAR, 2004), classificação de texturas (ZHANG et al., 2007), predição de movimento em vídeo (BOONTHEP; CHAMNONGTHAI; PHENSADSAENG, 2018), dentre outras aplicações. Diversas variações do SIFT foram propostas. Os descritores OS-SIFT (XIANG; WANG; YOU, 2018), IAG-SIFT (PAUL; PATI, 2017), SAR-SIFT (DELLINGER et al., 2015) e AAG-SIFT (WANG; YOU; FU, 2015) foram propostos com o objetivo de melhorar o registro de imagens de radar (SAR). PCA-SIFT (KE; SUKTHANKAR, 2004) e BE-SIFT (ZHAO et al., 2015) resultam em menor custo computacional e/ou menor dimensionalidade do vetor de características. Uma versão do SIFT invariante a espelhamento de imagens, denominada *Flip invariant SIFT* (F-SIFT), foi proposta por Bere (2018) para aplicação em detecção de objetos em imagens de vídeo. Com o objetivo de aumentar a robustez à variação da iluminação, Alitappeh e Mahmoudi (2013) propuseram o descritor MGS-SIFT, que apresentou melhor desempenho que o descritor SIFT original.

Uma implementação em software do SIFT está disponível no sítio eletrônico do Grupo de Geometria Visual¹ do Departamento de Engenharia da Universidade de Oxford.

2.3 Local Binary Pattern (LBP)

Baseado no descritor *Texture Unit* (TU) (WANG; HE, 1990), o *Local Binary Pattern* (LBP) (OJALA; PIETIKÄINEN; HARWOOD, 1996) foi proposto como uma solução simples e eficiente para classificação de texturas.

Considerando uma vizinhança de dimensão 3×3 *pixels*, representada na Figura 3(a), a aplicação do LBP permitiu reduzir em 25,6 vezes o número de possíveis unidades de textura geradas, passando de 6.561, 3^8 possíveis códigos ou padrões TU, para 256, 2^8 possíveis códigos ou padrões LBP. A imagem original I , de intensidade de nível de cinza, cuja dimensão é $X \times Y$, é codificada *pixel a pixel*, com exceção dos *pixels* da borda da imagem. Como resultado, tem-se I_{LBP} , uma imagem de códigos LBP, cujo histograma, de 256 posições, representa o vetor de características gerado pelo descritor.

Uma versão multiresolução do LBP (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002) considera uma vizinhança circular simétrica de raio R ($R \in \mathbb{Z}^+$, $R > 0$) com N vizinhos ($N \in \mathbb{Z}^+$, $N > 1$) igualmente espaçados em torno do pixel central $p(x, y)$, conforme ilustrado na Figura 3(b). A operação de extração de características LBP é representada

¹ <http://www.robots.ox.ac.uk/~vgg/research/affine/descriptors.html>

Figura 3: (a) Vizinhança 3×3 e (b) correspondente vizinhança circular de raio $R = 1$ e $N = 8$ em torno do pixel central $p(x, y)$.

na Equação (2.2), em que $s(d)$ é a função degrau, ou função *Heaviside*, definida na Equação (2.3), $g_i|_{i=1}^N$ é a intensidade do nível de cinza do i -ésimo *pixel* da vizinhança circular, e $g(x, y)$ é a intensidade do nível de cinza do *pixel* central $p(x, y)$.

$$\text{LBP}_{R,N}(x, y) = \sum_{i=1}^N s(g_i - g(x, y)) \cdot 2^{i-1} \quad (2.2)$$

$$s(d) = \begin{cases} 1, & \text{se } d \geq 0, \\ 0, & \text{caso contrário.} \end{cases} \quad (2.3)$$

Considerando-se o *pixel* p posicionado em $(0, 0)$, as coordenadas (x_i, y_i) de cada vizinho $p_i|_{i=1}^N$ são definidas por $(-R \cdot \sin(2\pi(i-1)/N), R \cdot \cos(2\pi(i-1)/N))$. A intensidade de nível de cinza g_i dos vizinhos que não são localizados no centro de cada *pixel* é estimada por interpolação bilinear (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002).

A análise por multiresolução é caracterizada pela combinação da informação do descritor para diferentes configurações de raio R e quantidade de vizinhos N considerados na vizinhança circular (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002). A Figura 4 mostra três diferentes configurações de vizinhança circular a serem utilizadas na análise por multiresolução. Na primeira configuração, Figura 4(a), a vizinhança possui raio unitário e oito *pixels* vizinhos igualmente espaçados; a segunda configuração, Figura 4(b), consiste em uma vizinhança circular de raio igual a 2 com 16 *pixels* igualmente espaçados; e a terceira configuração, Figura 4(c), é caracterizada por uma vizinhança composta por 24 *pixels* igualmente espaçados e dispostos em uma circunferência de raio 3.

De acordo com Haralick (1979) e Haralick, Shanmugam e Dinstein (1973), a informação de textura é diretamente proporcional ao tamanho do raio de uma vizinhança,

Figura 4: Diferentes configurações de vizinhança circular para análise por multiresolução. (a) $R = 1$ e $N = 8$, (b) $R = 2$ e $N = 16$ e (c) $R = 3$ e $N = 24$ (modificado de (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002)).

sendo que, quanto maior o raio R da vizinhança, maior é a informação de textura contida na respectiva região da imagem. Desse modo, o uso de diferentes resoluções espaciais, dada pela variação do raio da vizinhança, e de diferentes resoluções angulares, dada pela variação da quantidade de vizinhos N na vizinhança circular simétrica, permite considerar as diferentes informações de textura contidas em uma região da imagem.

Desenvolvido para solucionar problemas de classificação de texturas (OJALA; PIETIKÄINEN; HARWOOD, 1996; OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002), o descritor LBP foi implementado para solução de diversos outros problemas, tais como detecção facial, reconhecimento facial, detecção de objetos em movimento, dentre outros (ZHAO; PIETIKÄINEN, 2007; AHONEN; HADID; PIETIKÄINEN, 2006; HEIKKILÄ; PIETIKÄINEN, 2006; HADID; PIETIKÄINEN; AHONEN, 2004). Um estudo comparativo de diversos descritores de textura baseados no LBP é apresentado em (LIU et al., 2017).

2.4 Center-Symmetric Local Binary Pattern (CS-LBP)

Com o objetivo de associar a boa performance do SIFT em descrever regiões de interesse à baixa complexidade computacional e à robustez do LBP a variação da iluminação, Heikkilä, Pietikäinen e Schmid (2009) propuseram um novo descritor, denominado *Center-Symmetric Local Binary Pattern*. Para tanto, mantiveram a metodologia do SIFT de geração do histograma 3D da região de interesse e substituíram a regra de localização do gradiente e das respectivas orientações por uma nova regra de extração de características baseada no LBP.

Diferentemente do SIFT, o descritor CS-LBP não detecta as regiões de interesse. Para tal, utilizam-se os detectores de regiões de interesse, tais como *Hessian-Affine*, *Harris-Affine*, *Hessian-Laplace* e *Harris-Laplace*, sendo apresentado por Ferraz et al. (2014),

Mikolajczyk et al. (2005) e Mikolajczyk e Schmid (2005) que o detector *Hessian-Affine* é o mais indicado para aplicações de correspondência de imagens. Antes de aplicar o descritor, as regiões detectadas são mapeadas para uma região circular de raio constante, para obter invariância a escala e a transformações afins, e são normalizadas para uma malha cartesiana de dimensão 41×41 *pixels*, com 16 células (4×4), sendo os valores dos *pixels* compreendidos no intervalo $[0, 1]$; a invariância à rotação é obtida rotacionando as regiões normalizadas na direção do gradiente dominante.

Gera-se um histograma de características para cada célula, sendo a característica de cada *pixel* da região determinada por meio da aplicação da Equação (2.4), em que g_i e $g_{i+(N/2)}$ correspondem aos valores de níveis de cinza do par de *pixels* simétricos em relação ao *pixel* central $p(x, y)$, com N *pixels* igualmente espaçados em um círculo de raio R , e T é o limiar de diferença de nível de cinza aplicado à função $s(d)$, conforme Equação (2.5).

$$\text{CS-LBP}_{R,N,T}(x, y) = \sum_{i=1}^{(N/2)} s(g_i - g_{i+(N/2)}) \cdot 2^{i-1} \quad (2.4)$$

$$s(d) = \begin{cases} 1, & \text{se } d > T, \\ 0, & \text{caso contrário.} \end{cases} \quad (2.5)$$

Enquanto para o LBP são gerados 256 diferentes códigos na imagem de características, a comparação de intensidade de nível de cinza entre os *pixels* simétricos na vizinhança circular de raio R permite ao CS-LBP gerar unicamente 16 códigos ou diferentes características, conforme mostrado na Figura 5.

Figura 5: Comparação entre as regras de extração de características dos descritores LBP e CS-LBP, considerando uma vizinhança de raio $R = 1$ com oito *pixels* igualmente espaçados (modificado de (HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009)).

Para evitar mudança abrupta no descritor quando da passagem da característica de uma célula para outra, é realizada uma interpolação bilinear para distribuir o peso das características em *bins* adjacentes do histograma. O histograma da região de interesse

é construído por meio da concatenação dos histogramas determinados para cada célula, formando um vetor de características de dimensão $M \times M \times 2^{N/2}$, em que M corresponde ao número de células por direção, sendo o histograma normalizado para obter norma unitária. Tal como para o descritor SIFT, limita-se o valor de cada *bin* em 0,2, com o objetivo de reduzir a influência dos elementos de alta repetibilidade, realizando novamente a normalização para obter norma unitária.

O CS-LBP, quando parametrizado com malha cartesiana igual a 4×4 ($M = 4$) e 8 vizinhos ($N = 8$), gera um vetor de características de 256 posições. Possui tolerância a mudança de iluminação, robustez em regiões finas da imagem (obtida pelo ajuste do parâmetro T) e menor custo computacional quando comparado ao SIFT (HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009).

2.5 Completed Local Binary Pattern (CLBP)

Ojala, Pietikäinen e Harwood (1996), ao apresentarem o descritor LBP, discorrem sobre a possibilidade de uma única medida de textura não possuir informação suficiente para descrever adequadamente o arranjo espacial e a distribuição de intensidade de nível de cinza de texturas locais, e afirmam que a análise de ocorrências conjuntas de duas ou mais características poderia ser mais discriminante para texturas.

Com o objetivo de proporcionar informação complementar ao LBP para classificação de texturas, Guo, Zhang e Zhang (2010) propuseram o descritor *Completed Local Binary Pattern*, composto por três operadores: Sinal (CLBP_S), Magnitude (CLBP_M) e Centro (CLBP_C). Os operadores Sinal e Magnitude são complementares entre si e derivam da diferença local de nível de cinza, enquanto que o operador Centro representa uma binarização, tendo como limiar uma informação global da imagem.

Cada um dos operadores é caracterizado por uma regra de extração de características e gera uma imagem de códigos específica, I_{CLBP_S} , I_{CLBP_M} e I_{CLBP_C} . O operador CLBP_S é equivalente ao descritor LBP original, cuja extração de características se dá pela Equação (2.2). O operador CLBP_M, definido pela Equação (2.6), relaciona a magnitude da diferença local com a média das magnitudes das diferenças locais de toda a imagem $X \times Y$, $\bar{D}$, cuja determinação se dá pela Equação (2.7).

$$CLBP_M_{R,N}(x, y) = \sum_{i=1}^N s(|g_i - g(x, y)| - \bar{D}) \cdot 2^{i-1} \quad (2.6)$$

$$\bar{D} = \frac{1}{(X - 2R) \cdot (Y - 2R)} \sum_{x=R+1}^{X-R} \sum_{y=R+1}^{Y-R} \left(\frac{1}{N} \sum_{i=1}^N |g_i - g(x, y)| \right) \quad (2.7)$$

A regra de extração de características do operador Centro aplicada ao *pixel* $p(x, y)$ é definida pela Equação (2.8), em que o código binário 1 é gerado caso $p(x, y)$ tenha

intensidade de nível de cinza maior ou igual à média das intensidades de nível de cinza de toda a imagem, $\bar{G}$; 0, caso contrário. $\bar{G}$ é definido pela Equação (2.9).

$$\text{CLBP_C}(x, y) = s(g(x, y) - \bar{G}) \quad (2.8)$$

$$\bar{G} = \frac{1}{X \cdot Y} \sum_{x=1}^X \sum_{y=1}^Y g(x, y) \quad (2.9)$$

O vetor de características de cada operador é igual ao histograma unidimensional das imagens de códigos. Considerando $N = 8$, os operadores Sinal e Magnitude geram vetores de características com 256 posições. Por se tratar de uma binarização, o número de elementos do vetor de características do operador Centro independe do número de *pixels* da vizinhança, sendo fixo e igual a 2.

O vetor de características CLBP final é obtido por meio da combinação dos operadores individuais. Os vetores de características CLBP_S e CLBP_M são concatenados, gerando o vetor CLBP_S_M. Do mesmo modo, mas usando a regra de fusão Combinação Conjunta, as imagens de códigos são combinadas gerando os vetores CLBP_S/M, CLBP_S/C, CLBP_M/C e CLBP_S/M/C. Uma terceira operação, denominada de híbrida, se caracteriza por concatenar os descritores individuais com o resultado das combinações conjuntas, resultando nas abordagens CLBP_S_M/C e CLBP_M_S/C. As regras de fusão Concatenação e Combinação Conjunta são detalhadas no Capítulo 3.

Com o objetivo de reduzir a dimensão dos vetores de características, Guo *et al.* (GUO; ZHANG; ZHANG, 2010) sugerem o uso da versão LBP uniforme e invariante a rotação, LBP^{riu2} . Diferentes configurações do descritor $CLBP^{riu2}$ apresentaram melhor acurácia que cinco outros descritores de textura, a saber LBP^{riu2} (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002), VAR (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002), LTP (TAN; TRIGGS, 2010), VZ_MR8 (VARMA; ZISSERMAN, 2005) e VZ_Joint (VARMA; ZISSERMAN, 2003), quando aplicado a bases de texturas com variação de iluminação e rotação.

O descritor CLBP foi aplicado na literatura em tarefas de reconhecimento de íris (SOUZA; GONZAGA, 2019), de classificação de texturas (VIEIRA; NEGRI; GONZAGA, 2018), de segmentação e classificação de imagens médicas (OWJIMEHR; DANYALI; HELFROUSH, 2014) e de reconhecimento de expressões faciais (SINGH; MAURYA; MITTAL, 2012).

Uma extensão do CLBP foi proposta por Chen *et al.* (2016). O descritor *Multi-Scale Completed Local Binary Pattern* (MS-CLBP), obtido pela concatenação das componentes Sinal e Magnitude, CLBP_S e CLBP_M, para diferentes resoluções de vizinhança, assim

como para diferentes escalas da imagem. Aplicado a sensoriamento remoto, obteve melhor desempenho que outras metodologias.

2.6 Z with Tilted Z Local Binary Pattern (ZTZLBP)

Na regra de extração de características do descritor LBP, a informação de nível de cinza do *pixel* central de uma determinada vizinhança é utilizada como limiar na determinação do código LBP da respectiva vizinhança. Para o descritor CS-LBP, a informação do *pixel* central da vizinhança não é considerada na determinação da respectiva característica, o que representa uma perda de informação complementar. O operador Centro do descritor CLBP, CLBP_C, gera um vetor de características baseado na informação de nível de cinza do *pixel* central de cada vizinhança da imagem, o qual contribui para melhor representação da textura quando combinado aos operadores Sinal e Magnitude, evidenciando a relevância da informação do *pixel* central na representação da textura.

Com o intuito de gerar um vetor de características de baixa dimensão e de alta discriminância para textura, Mary e Dejeu (2018) recentemente propuseram o descritor de textura ZTZLBP², caracterizado pela aplicação do LBP a dois arranjos de *pixels* não sobrepostos de uma mesma vizinhança 3×3 , arranjo em Z e arranjo em Z inclinado (TZ), em que o *pixel* central $p(x, y)$ é considerado tal como os *pixels* da vizinhança na determinação das características ZLBP e TZLBP. Os dois diferentes arranjos de *pixels* não sobrepostos são ilustrados na Figura 6.

Figura 6: Vizinhança 3×3 para extração da característica ZTZLBP para os *pixels* dispostos (a) no arranjo em Z e (b) no arranjo em Z invertido (TZ).

As regras de extração das características $ZLBP(x, y)$ e $TZLBP(x, y)$ são definidas pelas Equações (2.10) e (2.11), em que $s(d) = 1$, caso $d > 0$, e $s(d) = 0$, caso contrário; e

² Mary e Dejeu (2018) denominaram o descritor como $Z \oplus TZLBP$, no entanto, optou-se nesta tese referenciá-lo por ZTZLBP para evitar possíveis confusões entre o símbolo $\oplus$ do nome do descritor e o operador de fusão Concatenação.

$g_{diff_A}(i)$, em que A define o arranjo da vizinhança, $A = \{Z, TZ\}$, é a i -ésima diferença de intensidade de nível de cinza do respectivo arranjo.

$$ZLBP(x, y) = \sum_{i=1}^4 s(g_{diff_Z}(i)) \cdot 2^{i-1} \quad (2.10)$$

$$TZLBP(x, y) = \sum_{i=1}^4 s(g_{diff_TZ}(i)) \cdot 2^{i-1} \quad (2.11)$$

As diferenças de intensidade de nível de cinza, considerando o arranjo em Z , $g_{diff_Z}(i)|_{i=1}^4$, são definidas pela Equação (2.12) e, considerando o arranjo em TZ , $g_{diff_TZ}(i)|_{i=1}^4$, pela Equação (2.13).

$$\begin{aligned} g_{diff_Z}(1) &= g_1 - g_3 \\ g_{diff_Z}(2) &= g_3 - g(x, y) \\ g_{diff_Z}(3) &= g(x, y) - g_6 \\ g_{diff_Z}(4) &= g_6 - g_8 \end{aligned} \quad (2.12)$$

$$\begin{aligned} g_{diff_TZ}(1) &= g_4 - g_2 \\ g_{diff_TZ}(2) &= g_2 - g(x, y) \\ g_{diff_TZ}(3) &= g(x, y) - g_7 \\ g_{diff_TZ}(4) &= g_7 - g_5 \end{aligned} \quad (2.13)$$

As imagens de códigos I_{ZLBP} e I_{TZLBP} são geradas pela aplicação das regras de extração de características supracitadas sobre todos os *pixels* da imagem original I , com exceção dos *pixels* da borda da imagem. Os vetores de características ZLBP e TZLBP, com 16 posições cada, são iguais aos histogramas unidimensionais das respectivas imagens de códigos.

A composição do vetor de características final, ZTZLBP, se dá pela soma *bin a bin* dos vetores de características de cada arranjo, ZLBP e TZLBP, conforme definido pela Equação (2.14).

$$ZTZLBP = ZLBP + TZLBP \quad (2.14)$$

Os autores demonstraram que o descritor ZTZLBP permitiu melhor resultado que o LBP (OJALA; PIETIKÄINEN; HARWOOD, 1996), Uniform Pattern (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002), CS-LBP (HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009) e OC-LBP (ZHU; BICHOT; CHEN, 2013) quando aplicado a classificação de recifes de coral em vídeos e imagens submarinas.

2.7 Local Mapped Pattern (LMP)

Ferraz, Pereira-Jr e Gonzaga (2014b) apresentaram uma generalização do descritor LBP, a qual representa melhor as diferenças entre os níveis de cinza dos *pixels* da vizinhança e o nível de cinza do *pixel* central e caracteriza-se por ser uma solução mais flexível que o LBP, uma vez que o valor máximo do código da imagem de códigos é parametrizável, permitindo assim ajustar o tamanho do vetor de características. O novo descritor foi denominado *Local Mapped Pattern* (LMP).

A regra de extração de características do descritor LMP aplicado sobre o *pixel* $p(x, y)$ é apresentada na Equação (2.15), em que $\mathbf{W}$ é o vetor de pesos, sendo $W_i = 2^{(i-1)}$; B é o valor máximo do código da imagem de códigos LMP ($B \in \mathbb{Z}^+$, $B \geq 1$); e $f(g_{diff_i})$ é a função sigmóide.

$$\text{LMP}_{R,N}(x, y) = \left\lfloor \frac{\sum_{i=1}^N f(g_{diff_i}) W_i}{\sum_{i=1}^N W_i} (B - 1) + 0.5 \right\rfloor \quad (2.15)$$

A função sigmóide $f(g_{diff_i})$, cuja formulação matemática é apresentada na Equação (2.16), mapeia o valor das diferenças locais de nível de cinza para as respectivas contribuições na determinação do código LMP do *pixel* $p(x, y)$, sendo g_{diff_i} a diferença de nível de cinza entre o i -ésimo *pixel* da vizinhança e $p(x, y)$, dada por $g_{diff_i} = g_i - g(x, y)$. A inclinação da curva é definida pelo parâmetro β ($\beta \in \mathbb{R}$, $\beta > 0$).

$$f(g_{diff_i}) = \frac{1}{1 + e^{-\frac{g_{diff_i}}{\beta}}} \quad (2.16)$$

A performance do descritor LMP depende do parâmetro β , uma vez que a inclinação da curva sigmóide define a sensibilidade do método às nuances da vizinhança (NEGRI et al., 2017). A Figura 7 ilustra o comportamento da função de mapeamento para diferentes valores de β .

Observa-se que para valores pequenos de β , a curva sigmóide aproxima-se do formato da curva *Heaviside*, saturando a função de mapeamento em 0, caso g_{diff} seja negativo, ou em 1, caso contrário. À medida que β aumenta, aumenta-se também a quantidade de possíveis valores a serem retornados pela função de mapeamento, até que, para valores elevados de β , a função de mapeamento assemelha-se a uma reta inclinada, quase paralela ao eixo das abscissas, caso em que mapeia a diferença g_{diff} para valores próximos a 0,5.

Os parâmetros B e β devem ser otimizados para cada aplicação a fim de se obter melhor desempenho do descritor. O vetor de características LMP, cujo número de elementos é igual a B , equivale ao histograma 1D da imagem de códigos LMP.

Figura 7: Comportamento da função de mapeamento $f(g_{diff})$ para diferentes valores de β (modificado de (NEGRI et al., 2017)).

2.8 Center-Symmetric Local Mapped Pattern (CS-LMP)

O descritor CS-LMP, *Center-Symmetric Local Mapped Pattern* (FERRAZ; PEREIRA-JR; GONZAGA, 2014a), é uma extensão do LMP, baseado na metodologia do CS-LBP (HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009). Como representa melhor as nuances entre os níveis de cinza dos *pixels* da região, apresenta-se como uma solução mais flexível que o CS-LBP.

A aplicação do CS-LMP segue o mesmo procedimento apresentado para o CS-LBP: detecção das regiões de interesse usando um descritor de região, mapeamento das regiões detectadas para um círculo de raio constante seguida por normalização para uma malha cartesiana de dimensão 41×41 *pixels* com 16 células (4×4), os valores do nível de cinza dos *pixels* são compreendidos entre $[0,1]$; rotaciona-se as regiões normalizadas na direção do gradiente.

A característica de cada *pixel* da região, usada para compor o histograma de cada célula, é determinada pela aplicação da Equação (2.17), em que B é o valor máximo do código da imagem de códigos I_{CS-LMP} , $W(i)$ é o i -ésimo elemento da matriz de pesos $\mathbf{W}$, sendo $W(i) = 2^{i-1}$, e $f(g_{diff_i})$ é o mapeamento definido pela função sigmóide, apresentada na Equação (2.16).

$$CS-LMP_{R,N}(x, y) = \left\lfloor \frac{\sum_{i=1}^{(N/2)} f(g_{diff_i}) W_i}{\sum_{i=1}^{(N/2)} W_i} (B - 1) + 0.5 \right\rfloor \quad (2.17)$$

A diferença de intensidades de nível de cinza, g_{diff_i} , para o descritor CS-LMP, é definida pela Equação (2.18), em que g_i e $g_{i+(N/2)}$ são as intensidades de nível de cinza

do par de *pixels* simétricos ao *pixel* central $p(x, y)$, em uma vizinhança circular de raio R constituída por N *pixels* igualmente espaçados.

$$g_{diff_i} = g_i - g_{i+(N/2)} \quad (2.18)$$

Assim como para o CS-LBP, realiza-se uma interpolação bilinear para distribuir o peso das características em *bins* adjacentes do histograma. Os histogramas de cada célula são concatenados a fim de se determinar o histograma da região de interesse, formando um vetor de características de dimensão $M \times M \times B$, em que M corresponde ao número de células por direção. O histograma gerado é então normalizado para obter norma unitária, seguido pelo truncamento em 0,2 e nova normalização para norma unitária.

O descritor CS-LMP permite maior flexibilidade que o CS-LBP, uma vez que é possível ajustar o número de *bins* do histograma por meio do parâmetro B , obtendo vetores de características de menor dimensão e mantendo melhor desempenho que o CS-LBP. Utilizando-se a malha cartesiana de dimensão 4 x 4 (16 células) e $b = 15$, gera-se um vetor de características de 256 posições.

Os parâmetros R , N , B e β devem ser ajustados para cada aplicação a fim de obter melhor desempenho. Em aplicação de registro de imagens, Ferraz, Pereira-Jr e Gonzaga (2014a) demonstraram que o descritor CS-LMP supera o desempenho da respectiva variante do LBP (CS-LBP) (HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009).

2.9 Mean Local Mapped Pattern (M-LMP)

Uma variação do descritor LMP, denominado *Mean Local Mapped Pattern* (M-LMP), foi proposto por Ferraz et al. (2014) para classificação de objetos. A principal contribuição do descritor proposto está na maneira de se calcular as diferenças locais de intensidade de nível de cinza g_{diff} , sendo as intensidades de nível de cinza g_i dos N *pixels* da vizinhança circular de raio R comparadas à média das intensidades de nível de cinza da vizinhança, g_{mean} , e não a $g(x, y)$, como se dá para o descritor LMP original.

A extração da característica do *pixel* $p(x, y)$ segue a formulação apresentada na Equação (2.19), idêntica à regra de extração de características para o descritor LMP, sendo $W_i = 2^{i-1}$. A Equação (2.20) determina o cálculo da função de mapeamento das diferenças locais.

$$\text{M-LMP}_{R,N}(x, y) = \left\lceil \frac{\sum_{i=1}^N f(g_{diff_i}) W_i}{\sum_{i=1}^N W_i} (B - 1) + 0.5 \right\rceil \quad (2.19)$$

$$f(g_{diff_i}) = \frac{1}{1 + e^{-\frac{g_{diff_i}}{\beta}}} \quad (2.20)$$

A diferença local de intensidade de nível de cinza g_{diff_i} é determinada pela Equação (2.21), em que g_{mean} é calculado como apresentado pela Equação (2.22).

$$g_{diff_i} = g_i - g_{mean} \quad (2.21)$$

$$g_{mean} = \frac{1}{N} \sum_{i=1}^N g_i \quad (2.22)$$

Os códigos M-LMP gerados são armazenados na imagem de códigos I_{M-LMP} , cujo histograma 1D representa o respectivo vetor de características com comprimento igual a B . Os valores de β e B devem ser ajustados empiricamente. Os autores demonstraram que a aplicação do descritor M-LMP em registro de imagens e em reconhecimento de objetos resulta em melhor desempenho que os descritores CS-LMP (FERRAZ; PEREIRA-JR; GONZAGA, 2014a), CS-LBP (HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009) e SIFT (LOWE, 2004).

2.10 Completed Local Mapped Pattern (CLMP)

Tendo em vista que a metodologia LMP representa melhor as nuances nas regiões de uma determinada imagem e que o descritor CLBP apresenta melhor desempenho que o LBP devido à combinação das informações complementares dos operadores Sinal, Magnitude e Centro, Vieira, Negri e Gonzaga (2018) propuseram o descritor *Completed Local Mapped Pattern* (CLMP). Tal como para o CLBP, as diferenças locais de intensidade de nível de cinza são decompostas em duas componentes complementares, dando origem aos operadores Sinal (CLMP_S) e Magnitude (CLMP_M). Com o objetivo de agregar informação global de textura, considerou-se ainda o operador Centro (CLMP_C) idêntico ao CLBP_C.

A Equação (2.23) representa a regra de extração de características para os operadores Sinal e Magnitude aplicados ao *pixel* $p(x, y)$, o qual é rodeado por uma vizinhança circular de raio R e N *pixels*, tal como ilustrado na Figura 3. h define o operador, $h = \{S, M\}$; B_S e B_M representam os valores máximos dos códigos para os operadores Sinal e Magnitude, respectivamente; $f_h(g_{diff_{h,i}})$ é a função sigmóide para os respectivos operadores, representada na Equação (2.24), em que β_S e β_M são os parâmetros de inclinação da curva; e $\mathbf{W}$ é o vetor de pesos. Para prover invariância à rotação, os autores utilizaram vetor de pesos unitário, sendo $W_i = 1, \forall i$.

$$\text{CLMP_}h_{R,N}(x, y) = \left\lceil \frac{\sum_{i=1}^N f_h(g_{diff_{h,i}})W_i}{\sum_{i=1}^N W_i} (B_h - 1) + 0.5 \right\rceil \quad (2.23)$$

$$f_h(g_{diff_{h,i}}) = \frac{1}{1 + e^{-\frac{g_{diff_{h,i}}}{\beta_h}}} \quad (2.24)$$

Para obter melhor desempenho com o descritor CLMP, os parâmetros B_S , B_M , β_S e β_M devem ser ajustados para a aplicação em questão. Os autores utilizaram a técnica de otimização por enxame de partículas (PSO) para determinar os respectivos valores ótimos.

As diferenças locais de intensidade de nível de cinza para os operadores CLMP_S e CLMP_M são calculadas pelas Equações (2.25) e (2.26), respectivamente, em que $\bar{D}$ é a média das magnitudes das diferenças locais de toda a imagem, tal como determinado na Equação (2.7).

$$g_{diff_{S,i}} = g_i - g(x, y) \quad (2.25)$$

$$g_{diff_{M,i}} = |g_i - g(x, y)| - \bar{D} \quad (2.26)$$

Os códigos gerados para cada *pixel*, por cada operador, são armazenados nas imagens de códigos I_{CLMP_S} , I_{CLMP_M} e I_{CLMP_C} , sendo I_{CLMP_C} igual à respectiva imagem de códigos do descritor CLBP, I_{CLBP_C} . O vetor de características para cada operador é igual ao histograma 1D da respectiva imagem de códigos, e possuem comprimento igual a B_S , B_M e 2, respectivamente.

Tal como para o CLBP, para melhorar o desempenho do descritor CLMP, os operadores Sinal, Magnitude e Centro são combinados usando as regras de fusão Concatenação e Combinação Conjunta, as quais são descritas no Capítulo 3. [Vieira, Negri e Gonzaga \(2018\)](#) demonstraram que o descritor CLMP superou o desempenho do CLBP ([GUO; ZHANG; ZHANG, 2010](#)) em aplicação de classificação de texturas rotacionadas.

2.11 Median Local Mapped Pattern (Median-LMP)

[Souza e Gonzaga \(2019\)](#) analisam alguns descritores baseados no LMP e argumentam que, mesmo apresentando baixo custo computacional e alto desempenho, tais descritores são suscetíveis a queda de desempenho devido a ruídos anatômicos, a ruídos digitais e à variação da iluminação. Baseados no descritor *Median Binary Pattern* (MBP) ([HAFIANE; SEETHARAMAN; ZAVIDOVIQUE, 2007](#)), os autores propuseram um novo descritor, denominado *Median Local Mapped Pattern* (Median-LMP), que possui a propriedade de ser mais robusto a ruídos e à mudança de iluminação.

Tal como apresentado para o descritor M-LMP, a principal contribuição do descritor Median-LMP relaciona-se ao cálculo das diferenças locais de intensidade de nível de cinza g_{diff} , determinada pela Equação (2.27). Para a determinação do código Median-LMP, as

intensidades de nível de cinza g_i dos *pixels* da vizinhança $J \times J$ são comparadas à mediana das intensidades de nível de cinza da respectiva região $J \times J$, g_{median} .

$$g_{diff_i} = g_i - g_{median} \quad (2.27)$$

A Figura 8 ilustra a diferença de metodologia de determinação do valor de referência a ser utilizado na comparação com os *pixels* da vizinhança para determinação dos códigos LMP e Median-LMP.

A extração da característica do *pixel* $p(x, y)$ na vizinhança $J \times J$, representada na Figura 9, é dada pela Equação (2.28), sendo a função de mapeamento determinada pela Equação (2.20).

$$\text{Median-LMP}(x, y) = \left\lfloor \frac{\sum_{i=1}^N f(g_{diff_i}) W_i}{\sum_{i=1}^N W_i} (B - 1) + 0.5 \right\rfloor \quad (2.28)$$

100	50	25	100	50	25
10	100	50	25	50	30
100	10	255	30	15	20
100	100	15	20	50	25
10	50	30	10	50	30
100	15	20	100	15	20

(a) Imagem original com uma região local 3×3 destacada em vermelho.

100	50	25
10	255	30
100	15	20

(b) Vizinhança 3×3 afetada por ruído no *pixel* central.

100	50	25
10	255	30
100	15	20

(c) Intensidade de nível de cinza do *pixel* central, $g_c = 255$, para comparação com os *pixels* da vizinhança pelo LMP.

100	50	25
10	255	30
100	15	20

(d) Valor mediano das intensidades de nível de cinza dos *pixels* da vizinhança 3×3 , $g_{median} = 30$, para comparação com os *pixels* da vizinhança pelo Median-LMP.

Figura 8: Metodologias de determinação da intensidade de nível de cinza para comparação com os *pixels* da vizinhança na determinação dos códigos LMP e Median-LMP (modificado de (SOUZA; GONZAGA, 2019)).

g_1	g_2	g_3
g_4	$g(x,y)$	g_5
g_6	g_7	g_8

Figura 9: Vizinhança 3×3 em torno do *pixel* $p(x, y)$ (modificado de (SOUZA; GONZAGA, 2019)).

Os parâmetros β e B , que determinam a inclinação da curva sigmóide e o valor máximo do código na imagem de códigos Median-LMP, respectivamente, devem ser ajustados para obter melhor desempenho na aplicação em questão. Os autores utilizaram o algoritmo de Otimização por Enxame de Partículas (PSO) para determinar o valor de β , e utilizaram B fixo e igual a 256.

Foram utilizados dois vetores de pesos W diferentes, um unitário, $W1_i = 1, \forall i$, e um segundo de potência de base 2, cujos elementos são dados por $W2_i = 2^{i-1}$. Os experimentos realizados demonstraram que o descritor cuja extração de características considerou o vetor de pesos de potência de base 2, $W2$, apresentou melhor desempenho que aquele em que foi utilizado o vetor de pesos unitário, $W1$.

Em aplicações de reconhecimento de íris considerando variação da pupila e diferentes resoluções de amostragem da íris, o descritor Median-LMP superou o desempenho do método de Daugman (DAUGMAN, 2004) e de diversos descritores, tal como o MBP (HAFIANE; SEETHARAMAN; ZAVIDOVIQUE, 2007), LMP (FERRAZ; PEREIRA-JR; GONZAGA, 2014b), WLD (CHEN et al., 2009) e CLBP (OJALA; PIETIKÄINEN; HARWOOD, 1996).

2.12 Considerações Finais

Apresentaram-se, neste capítulo, as características e a formulação matemática de descritores de textura de diferentes abordagens. De acordo com os resultados apresentados em (SOUZA; GONZAGA, 2019; VIEIRA; NEGRI; GONZAGA, 2018; LIU et al., 2017; WANG; PENG; ZHANG, 2017; FERRAZ et al., 2014; HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009) não há um descritor que tenha o melhor desempenho para todos os conjuntos de imagens, sendo um grande desafio desenvolver um descritor que garanta os melhores resultados independente das características das imagens e da aplicação (NANNI; BRAHNAM; LUMINI, 2019).

De acordo com Ojala, Pietikäinen e Harwood (1996), é possível que uma única

medida de textura não possua informação suficiente para descrever adequadamente o arranjo espacial e a distribuição de intensidade de nível de cinza de texturas locais, sendo que a análise de ocorrências conjuntas de duas ou mais características poderia ser mais discriminante para texturas.

Huang, Chan e Zhang (2002) discorrem sobre descritores das abordagens estatística e estrutural, e afirmam que texturas naturais geralmente são constituídas por ambas componentes, estatística e estrutural, sendo assim, a utilização de um descritor de uma única abordagem não seria suficiente para descrever completamente imagens de textura.

Nesta tese, propõe-se a utilização de operadores de fusão para combinar diferentes descritores de textura a fim de melhorar a representação de imagens de textura em diferentes aplicações. A fundamentação teórica sobre fusão de descritores de textura é apresentada no capítulo a seguir.

3 FUSÃO DA INFORMAÇÃO

Como visto no capítulo anterior, há diferentes tipos de descritores de textura, os quais são vastamente aplicados na solução de diversos problemas, tais como correspondência de imagens, reconhecimento de objetos, reconhecimento de faces, classificação/identificação de objetos em vídeos, localização de robôs, dentre outras. No entanto, os descritores de textura possuem limitações para determinadas aplicações e/ou para determinados tipos de imagens, não havendo um descritor que apresente o melhor resultado dentre os demais para toda e qualquer aplicação e/ou para todo e qualquer conjunto de imagens (WANG; PENG; ZHANG, 2017; GUO; ZHAO; PIETIKÄINEN, 2012).

Neste capítulo, apresenta-se a teoria de fusão da informação e discorre-se sobre os diferentes níveis, assim como sobre as diferentes arquiteturas de fusão, enfatizando-se o fato de a fusão da informação permitir integrar as características dos diferentes métodos combinados, possibilitando, assim, obter resultados melhores que os obtidos pela aplicação de cada método individualmente.

3.1 Considerações Iniciais

Fusão de informação é uma técnica de processamento, associação e combinação de dados e/ou de informações provenientes de uma única fonte ou de múltiplas fontes com o objetivo de obter maior precisão, maior confiabilidade e melhor inferência do objeto em estudo ou melhor percepção do ambiente em questão (CASTANEDO, 2013; KHALEGHI et al., 2013; BOSTRÖM et al., 2007; HALL; LLINAS, 1997; WHITE, 1991), de modo que o resultado após a fusão seja representativo a ponto de uma única fonte ou sensor não poder fornecer, individualmente, resultado semelhante.

Com vasta aplicação em visão computacional, imagens médicas, automação industrial, aplicações militares e sensoriamento remoto (DONG et al., 2009; DURRANT-WHYTE; HENDERSON, 2008; LUO; YIH; SU, 2002), a fusão da informação pode ser implementada em quatro diferentes níveis ou camadas, sendo eles fusão de sinais, fusão de *pixels*, fusão de características e fusão de decisões (GHASSEMIAN, 2016; CASTANEDO, 2013; DONG et al., 2009; LUO; YIH; SU, 2002), tal como apresentado na Figura 10.

De acordo com Castanedo (2013) e Luo, Yih e Su (2002), a fusão de sinais pode ser utilizada em aplicações de tempo real, servindo como uma etapa adicional em processamento de sinais como um todo, manipulando diretamente os sinais advindos dos sensores; a fusão de *pixels* pode ser utilizada, por exemplo, para melhorar o desempenho de alguns procedimentos de pré-processamento de imagens, tal como a segmentação, sendo realizadas operações sobre elementos de imagens; os dois últimos níveis, fusão de características e

Figura 10: Classificação das técnicas de fusão por níveis (adaptado de (DONG et al., 2009)).

fusão de decisões, podem ser utilizados em sistemas de reconhecimento de objetos, ou de padrões, para melhorar a capacidade de reconhecimento do respectivo sistema, sendo que, na fusão de características, são aplicados operadores de fusão para combinar as características extraídas das imagens ou dos sinais, e na fusão de decisões, as operações de combinação são aplicadas sobre as informações de classes, símbolos e/ou decisões previamente tomadas pelo sistema.

Existem outros critérios de classificação de técnicas de fusão, sendo eles de acordo com: (I) a configuração dos dados de entrada ou das fontes de informação (sensores), classificando em dados ou fontes complementares, redundantes ou cooperativas (DURRANT-WHYTE, 1988 apud CASTANEDO, 2013); (II) o tipo de informação na entrada e na saída do elemento de fusão (CASTANEDO, 2013; GIGLI; BOSSÉ; LAMPROPOULOS, 2007; DASARATHY, 1997), classificando em dados na entrada e dados na saída (DAI-DAO), dados na entrada e características na saída (DAI-FEO), características na entrada e características na saída da fusão (FEI-FEO), características na entrada e decisão na saída (FEI-DEO) e informação de decisão na entrada e decisão na saída (DEI-DEO); (III) os diferentes níveis de dados de fusão considerados pelo *Joint Directors Laboratories - JDL* (CASTANEDO, 2013; WHITE, 1991), sendo considerados cinco diferentes níveis de dados, os quais podem ser agrupados em dois grupos de fusão, fusão de baixo nível e fusão de alto

nível; e (IV) o tipo de arquitetura (CASTANEDO, 2013; DURRANT-WHYTE; HENDERSON, 2008), sendo arquitetura centralizada, descentralizada, distribuída e hierárquica.

Os diferentes critérios de classificação da fusão são resultado do fato de a fusão da informação ser aplicada em diferentes áreas, sendo que os profissionais de cada área utilizam o critério que julgam mais pertinente. Essa multidisciplinaridade da fusão é consequência das vantagens que as respectivas técnicas aportam para os sistemas em que são aplicadas. Como vantagens do uso de fusão da informação, podem-se citar maior tolerância a falhas, maior cobertura espacial e temporal, maior robustez contra erros de medição (maior confiabilidade), redução da incerteza, a possibilidade de redução do custo do sistema (investimento em sensores mais baratos), aumento do desempenho do sistema, dentre outras.

3.2 Abordagens de Fusão aplicadas a Visão Computacional

Na literatura, encontram-se referenciados diversos algoritmos e métodos de fusão da informação (GHASSEMIAN, 2016). Como cada aplicação requer uma solução específica, não existe um método de fusão que seja apropriado para todas as situações, sendo necessário analisar cada problema individualmente (SHIHAVUDDIN et al., 2013).

Yushkevich et al. (2012) aplicaram fusão por votos ponderados a registro de imagens médicas, sendo que a técnica de fusão proporcionou correspondências mais consistentes quando comparado com outras abordagens.

Com o argumento de que diferentes descritores de textura apresentam informações complementares, Wang, Peng e Zhang (2017) propuseram um método estatístico de otimização dos pesos utilizados na combinação linear entre as informações extraídas por cinco diferentes descritores de textura a fim de solucionar um problema de detecção de fechamento de *loop*.

Campos, Correia e Calado (2015) aplicaram fusão na camada de decisão por meio dos operadores Adição, Produto e Voto Majoritário para melhorar a capacidade de localização visual de robôs, sendo os melhores resultados obtidos pelo operador Adição.

Para melhorar o desempenho de sistemas de classificação de recifes de corais, Shihavuddin et al. (2013) concatenaram os vetores de características dos descritores CLBP, GLCM, filtro de Gabor e *Opponent Angle and Hue Color Channel Histograms*, obtendo maior acurácia global que as demais abordagens testadas.

Em aplicação de classificação de texturas rotacionadas, Ojala, Pietikäinen e Mäenpää (2002) aplicaram a Combinação Conjunta sobre os descritores LBP e VAR para diferentes resoluções da vizinhança. Os resultados obtidos com a fusão dos dois descritores foram melhores que os alcançados com a aplicação de cada descritor individualmente.

Nanni e Lumini (2013) apresentaram um sistema de reconhecimento de objetos

usando *Bag of Features*. Na fase de classificação, utilizaram diferentes classificadores SVM, cada um treinado com um descritor de regiões específico. Realizaram a fusão da informação dos classificadores, caracterizando uma fusão na camada de decisão. Para tal, utilizaram dois operadores de fusão distintos, soma algébrica (Adição) das respostas dos classificadores e soma ponderada. Mostraram que os resultados obtidos com a fusão das respostas dos classificadores foram melhores que os obtidos com a aplicação de cada classificador individualmente.

Além das abordagens de fusão aplicadas sobre diferentes descritores, há ainda os descritores baseados em fusão de diferentes componentes. Como exemplo, têm-se alguns dos descritores citados no Capítulo 2, como CLBP, MS-CLBP, CLMP e WLD, cujo descritor final é obtido por meio da Concatenação e/ou da Combinação Conjunta das componentes Sinal, Magnitude e Centro para os três primeiros descritores; e excentricidade e orientação para o WLD. Ademais, [Negri et al. \(2017\)](#) propuseram um descritor de cor para classificação de texturas robusto à variação da iluminação, denominado *Color Intensity Local Mapped Pattern* (CILMP), cujo vetor de características consiste na Concatenação de duas componentes, uma de luminância e outra de características no espaço de cor RGB, sendo cada componente determinada para diferentes resoluções de vizinhança. A Concatenação de diferentes componentes foi também utilizada por [Kim, Yoo e Sohn \(2013\)](#) na formulação do descritor *Exact Order Based Descriptor* (EOD), em que as características globais (GEOF) e locais (LEOF) são concatenadas para gerar o vetor de características final do respectivo descritor de textura.

Diversos outros algoritmos e métodos de fusão são utilizados na literatura, sendo alguns deles: *Principal Component Analysis* - PCA, Intensidade - Matiz - Saturação (IHS), diferentes combinações aritméticas, como Transformada Brovey (cor normalizada), média ponderada, Redes Neurais Artificiais - RNA, Filtro de Kalman, Estimador de Máxima Verossimilhança - MLE, Inferência Bayesiana, Inferência Dempster-Shafer, Operadores de fusão fuzzy, Método de Monte Carlo, dentre outros ([GHASSEMIAN, 2016](#); [CASTANEDO, 2013](#); [KHALEGHI et al., 2013](#); [DONG et al., 2009](#); [DURRANT-WHYTE](#); [HENDERSON, 2008](#); [KITTLER et al., 1998](#)).

3.3 Fusão de descritores de textura na etapa de Características

Sejam I uma imagem de intensidade de níveis de cinza, de dimensão $X \times Y$, I_{DESC_1} , I_{DESC_2} , I_{DESC_3} e I_{DESC_N} as imagens de códigos geradas pela aplicação dos descritores $DESC_1$, $DESC_2$, $DESC_3$ e $DESC_N$ sobre a imagem I , respectivamente, e B_1 , B_2 , B_3 e B_N os valores máximos do código nas respectivas imagens de códigos. A representação 1D das imagens de códigos gera os vetores de características $\mathbf{v}_1 = [v_{11} \ v_{12} \ \cdots \ v_{1B_1}]$ para I_{DESC_1} , $\mathbf{v}_2 = [v_{21} \ v_{22} \ \cdots \ v_{2B_2}]$ para I_{DESC_2} , $\mathbf{v}_3 = [v_{31} \ v_{32} \ \cdots \ v_{3B_3}]$ para I_{DESC_3} e $\mathbf{v}_N = [v_{N1} \ v_{N2} \ \cdots \ v_{NB_N}]$ para I_{DESC_N} , com números de elementos iguais a B_1 , B_2 , B_3

e B_N , respectivamente.

Os vetores de características são a representação da imagem I em um determinado espaço de características. Cada descritor gera um vetor de características diferente, sendo assim, a mesma imagem I pode ser representada por diferentes pontos em um determinado espaço de características, cuja localização espacial depende do descritor utilizado. A fusão de vetores de características pode ser definida como a aplicação de regras de combinação sobre as características dos diferentes vetores, resultando em um novo vetor de características, $\mathbf{v}_F$, ou seja, uma nova representação da imagem I no espaço de características. O número de elementos de $\mathbf{v}_F$ é função do número de elementos dos vetores de características utilizados na fusão. Dois operadores de fusão de características comumente utilizados na literatura são a Soma e a Concatenação (WANG et al., 2017; WANG et al., 2011; GUO; ZHANG; ZHANG, 2010; WANG et al., 2009).

Outra operação comum de se realizar na etapa de características é a fusão de imagens de códigos, a qual diferencia-se da aplicada sobre os vetores de características por permitir considerar a informação de localização espacial das características nas imagens, o que pode ser ainda mais discriminativo. O operador Combinação Conjunta demonstrou bons resultados quando aplicado aos descritores LBP (OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002), CLBP (GUO; ZHANG; ZHANG, 2010) e CLMP (VIEIRA; NEGRI; GONZAGA, 2018).

3.3.1 Soma de vetores de características

O operador Soma (+) (WANG et al., 2011; WANG et al., 2009) aplicado a dois ou mais vetores de características, $\mathbf{v}_{Soma} = \mathbf{v}_1 + \mathbf{v}_2 + \dots + \mathbf{v}_N$, realiza a soma *bin a bin* dos elementos dos respectivos vetores de características, tal como definido na Equação (3.1), com a restrição de que os vetores combinados devem possuir o mesmo número de elementos, $B_1 = B_2 = \dots = B_N$. Como a soma é realizada *bin a bin*, o número de elementos do vetor $\mathbf{v}_{Soma}$ resultante é igual ao número de elementos dos vetores de características combinados, $B_{Soma} = B_1$.

$$\begin{aligned} \mathbf{v}_{Soma} &= \mathbf{v}_1 + \dots + \mathbf{v}_N \\ &= \begin{bmatrix} v_{11} + \dots + v_{N1} & v_{12} + \dots + v_{N2} & \dots & v_{1B_1} + \dots + v_{NB_1} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^N v_{i1} & \sum_{i=1}^N v_{i2} & \dots & \sum_{i=1}^N v_{iB_1} \end{bmatrix} \end{aligned} \quad (3.1)$$

A fusão por Soma de dois vetores de características, $\mathbf{v}_1$ e $\mathbf{v}_2$, é ilustrada na Figura 11.

3.3.2 Concatenação de vetores de características

O operador Concatenação ($\oplus$) (WANG et al., 2017; WANG et al., 2011) aplicado a dois ou mais vetores de características, $\mathbf{v}_{Conc} = \mathbf{v}_1 \oplus \mathbf{v}_2 \oplus \dots \oplus \mathbf{v}_N$, realiza a justaposição

Figura 11: Exemplo de fusão por Soma de dois vetores de características.

dos elementos dos respectivos vetores de características, tal como definido na Equação (3.2). O número de elementos do vetor $\mathbf{v}_{Conc}$ é igual à soma do número de elementos dos vetores de características concatenados, $B_{Conc} = B_1 + B_2 + \dots + B_N$.

$$\begin{aligned} \mathbf{v}_{Conc} &= \mathbf{v}_1 \oplus \mathbf{v}_2 \oplus \dots \oplus \mathbf{v}_N \\ &= \begin{bmatrix} v_{11} & \dots & v_{1B_1} & v_{21} & \dots & v_{2B_2} & \dots & v_{N1} & \dots & v_{NB_N} \end{bmatrix} \end{aligned} \quad (3.2)$$

A fusão por Concatenação de dois vetores de características, $\mathbf{v}_1$ e $\mathbf{v}_3$, é ilustrada na Figura 12.

3.3.3 Combinação Conjunta de imagens de códigos

O operador Combinação Conjunta ($| \cdot |$) (NEGRI et al., 2018; WANG et al., 2017; OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002) aplicado a duas ou mais imagens de códigos, $\mathbf{v}_{Comb} = I_{DESC_1} | I_{DESC_2} | \dots | I_{DESC_N}$, realiza o histograma conjunto N -dimensional ($\mathbf{H}_{Comb}$) das respectivas imagens de códigos, seguido pela representação 1D do histograma N -dimensional gerado. O número de elementos do vetor $\mathbf{v}_{Comb}$ é igual ao produto entre os números de elementos dos vetores de características combinados, $B_{Comb} = B_1 \times B_2 \times \dots \times B_N$.

Figura 12: Exemplo de fusão por Concatenação de dois vetores de características.

A fusão por Combinação Conjunta de duas imagens de códigos, I_{Desc_1} e I_{Desc_3} , é ilustrada na Figura 13, sendo o histograma conjunto entre duas imagens de códigos definido na Equação (3.3), em que $\#(i, j)$ representa o número de elementos da imagem I_{DESC_1} que possuem valor de código igual a i e que ocupam as mesmas posições dos elementos da imagem I_{DESC_3} com valor de código igual a j .

$$\mathbf{H}_{Comb} = \begin{bmatrix} \#(0,0) & \#(0,1) & \#(0,2) & \cdots & \#(0, B_3-1) \\ \#(1,0) & \#(1,1) & \#(1,2) & \cdots & \#(1, B_3-1) \\ \#(2,0) & \#(2,1) & \#(2,2) & \cdots & \#(2, B_3-1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \#(B_1-1,0) & \#(B_1-1,1) & \#(B_1-1,2) & \cdots & \#(B_1-1, B_3-1) \end{bmatrix} \quad (3.3)$$

3.4 Fusão de descritores de textura na etapa de Decisão

De acordo com Gigli, Bossé e Lampropoulos (2007), a fusão na camada de decisões apresenta menor custo computacional que fusão na camada de características, além de ser mais robusta à inclusão e/ou retirada de fontes de informação. No entanto, possui a

Figura 13: Exemplo de fusão por Combinação Conjunta de duas imagens de códigos.

desvantagem de não permitir a recuperação de informações relevantes que por ventura tenham sido ignoradas ou perdidas na camada de características.

Alguns dos operadores de fusão de decisão comumente utilizados são Soma (ou Adição), Soma Ponderada, Voto Majoritário, Produto, Mínimo, Máximo, regras Fuzzy e regras Bayesianas.

3.4.1 Operador de fusão Adição

O operador Adição (Add)¹ aplicado a duas informações de decisão, d_1 e d_2 , é definido na Equação (3.4) como a soma algébrica das respectivas informações de decisão.

$$d_{add} = d_1 + d_2 \quad (3.4)$$

¹ Adotou-se neste trabalho a denominação Adição para este operador a fim de diferenciá-lo do operador de fusão Soma, aplicado a vetores de características.

De acordo com Kittler et al. (1998), a fusão por soma das informações de decisões é mais resistente a erros de estimação quando comparada a outros operadores, como Produto, Mínimo, Máximo, Mediana e Voto Majoritário.

3.4.2 Operadores de fusão Fuzzy

Apresenta-se, em (KHALEGHI et al., 2013), um modelo de operadores de fusão *fuzzy* cujo uso é recomendado para combinar dados parciais ou dados vagos.

Seja um conjunto fuzzy $F \subseteq \mathbf{X}$ definido pela função de pertinência $\mu_F(x)$, como apresentado na Equação (3.5).

$$\mu_F(x) \in [0, 1] \quad \forall x \in \mathbf{X} \quad (3.5)$$

Os quatro operadores fuzzy de fusão são apresentados nas equações (3.6), para o operador de Interseção; (3.7), para o Produto de dois conjuntos fuzzy; (3.8), para o operador de União entre dois conjuntos fuzzy; e (3.9), para o operador Soma Algébrica.

$$\mu_1^\cap(x) = \min(\mu_{F_1}(x), \mu_{F_2}(x)) \quad \forall x \in \mathbf{X} \quad (3.6)$$

$$\mu_2^\cap(x) = \mu_{F_1}(x) \cdot \mu_{F_2}(x) \quad \forall x \in \mathbf{X} \quad (3.7)$$

$$\mu_1^\cup(x) = \max(\mu_{F_1}(x), \mu_{F_2}(x)) \quad \forall x \in \mathbf{X} \quad (3.8)$$

$$\mu_2^\cup(x) = \mu_{F_1}(x) + \mu_{F_2}(x) - (\mu_{F_1}(x) \cdot \mu_{F_2}(x)) \quad \forall x \in \mathbf{X} \quad (3.9)$$

Os operadores de fusão Interseção e União de dois conjuntos *fuzzy*, semelhante ao modo como foram definidos nesta seção, são também aplicados a dados que não sejam *fuzzy*, caso em que são referenciados como operadores Mínimo e Máximo, respectivamente, como apresentado em (CAMPOS; CORREIA; CALADO, 2015; KITTLER et al., 1998). O mesmo ocorre para o operador Produto, cuja denominação é mantida.

Conforme apresentado por Wang, Peng e Zhang (2017), os operadores de fusão na camada de decisão comumente realizam combinação linear entre as informações de decisão dos descritores individuais. Desse modo, o operador de fusão Soma Algébrica pode também ser aplicado a dados que não sejam *fuzzy*, sendo compreendido como uma combinação linear entre os operadores Adição e Produto. Esse operador é referenciado nesta tese pela denominação operador de fusão de decisão Combinação Linear.

3.4.3 Inferência Bayesiana

De acordo com [Castanedo \(2013\)](#), o uso de inferência Bayesiana em fusão se dá na camada de decisões. A fusão baseada na inferência Bayesiana combina evidências com teoria probabilística, e fundamentam-se nas regras do Teorema de Bayes, conforme Equação (3.10) ([CAMPOS; CORREIA; CALADO, 2015](#); [GIGLI; BOSSÉ; LAMPROPOULOS, 2007](#)).

$$P(Y|x) = \frac{P(x|Y) \cdot P(Y)}{P(x)} , \quad (3.10)$$

em que a probabilidade *a posteriori*, $P(Y|x)$, representa a probabilidade de a hipótese Y ser verdadeira, uma vez que ocorreu o evento x, sendo dada pelo produto entre a probabilidade *a priori* da hipótese $P(Y)$ e a probabilidade de ocorrência do evento x sabendo que a hipótese Y é verdadeira. A probabilidade $P(x)$ é uma constante de normalização.

Para utilizar a inferência Bayesiana, é necessário conhecer previamente os valores das probabilidades condicional $P(x|Y)$ e *a priori* $P(Y)$. Os operadores de fusão Adição e Produto são comumente utilizados para combinar as probabilidades *a posteriori* oriundas de diferentes fontes de informação.

3.5 Considerações Finais

Discorreu-se, neste capítulo, sobre fusão da informação, sendo apresentadas algumas abordagens de fusão aplicadas a processamento de imagens. Notou-se uma pluralidade de metodologias, todas direcionadas a melhorar o desempenho de diversos sistemas por meio da combinação da informação de diferentes fontes, seja pela fusão de elementos de imagens, de características extraídas das imagens ou de informações de decisão. Foram apresentados alguns operadores de fusão de características e de decisão.

A fim de contribuir com o desenvolvimento de pesquisas voltadas a processamento de imagens e a fusão da informação, apresentam-se no capítulo a seguir a proposta de dois novos operadores de fusão de características. Ademais, são propostos três novos descritores de textura baseados em fusão de características.

4 MÉTODOS PROPOSTOS

As regras de aplicação de diferentes descritores de textura foram apresentadas no Capítulo 2. Como o desempenho dos descritores de textura é associado ao tipo de imagem e à aplicação específica, diversas abordagens combinam diferentes descritores de textura a fim de se obter melhores resultados. A teoria sobre fusão da informação foi apresentada no Capítulo 3, assim como a metodologia de aplicação dos operadores de fusão de características Soma, Concatenação e Combinação Conjunta, e dos operadores de fusão de decisão Adição, Mínimo (Interseção *fuzzy*), Máximo (União *fuzzy*), Produto e Combinação (Soma Algébrica *fuzzy*).

Com o objetivo de obter melhores desempenhos em aplicações de visão computacional e de reconhecimento de imagens, apresentam-se, neste capítulo, a proposta de três novos descritores de textura baseados em fusão de características e de dois novos operadores de fusão.

4.1 Descritores de textura baseados em fusão de características

Propõe-se a nomenclatura descritores de textura baseado em fusão de características aos descritores que apresentam, cada um, nas respectivas metodologias de aplicação, diferentes regras de extração de características, sendo cada regra referente a uma componente individual, e cujo vetor de características final seja constituído pela fusão das diferentes componentes individuais. Os descritores CLBP (GUO; ZHANG; ZHANG, 2010), CLMP (VIEIRA; NEGRI; GONZAGA, 2018), CILMP (NEGRI et al., 2017), WLD (CHEN et al., 2009) e EOD (KIM; YOO; SOHN, 2013) são exemplos de descritores de textura baseados em fusão de características.

Em concordância com o que foi exposto nos capítulos anteriores, sabe-se que:

- S1. a fusão de informações complementares de diferentes descritores possibilita a obtenção de resultados superiores ao melhor resultado dentre os obtidos pela aplicação dos descritores individuais;
- S2. o descritor *Completed Local Binary Pattern* apresenta uma metodologia de extração de informações complementares de uma imagem, representando-as em diferentes componentes, as quais são combinadas para gerar o vetor de características final;
- S3. os descritores de textura baseados no LMP apresentaram melhor desempenho que os respectivos descritores baseados no LBP para diferentes aplicações;
- S4. os descritores *Mean Local Mapped Pattern* (M-LMP) e *Median Local Mapped Pattern* (Median-LMP) possibilitaram a obtenção de melhores resultados que o LMP, e que

outras variações do LMP, ao proporem uma modificação no cálculo das diferenças locais de intensidade de nível de cinza;

- S5. o descritor *Z with Tilted Z Local Binary Pattern* (ZTZLBP), ao extrair duas características por *pixel* da imagem, resultou em melhor desempenho que diversos outros descritores.

Dispondo dessas asserções, propõem-se três novos descritores de textura baseados em fusão de características, *Completed Mean Local Mapped Pattern* (CMLMP), *Completed Median Local Mapped Pattern* (CMedianLMP) e *Completed Z with Tilted Z Local Mapped Pattern* (CZTZLMP), os quais são apresentados a seguir.

4.1.1 *Completed Mean Local Mapped Pattern* (CMLMP)

O descritor *Completed Mean Local Mapped Pattern* baseia-se na extração de informações complementares de uma textura por diferentes componentes, tal como realizado pelos descritores CLBP (GUO; ZHANG; ZHANG, 2010) e CLMP (VIEIRA; NEGRI; GONZAGA, 2018), utilizando o método de extração de características do descritor MLMP (FERRAZ et al., 2014). Como é composto por três diferentes componentes, Sinal (CMLMP_S), Magnitude (CMLMP_M) e Centro (CMLMP_C), é possível representar a textura por diferentes vetores de características, cada um resultado de uma combinação específica das respectivas componentes.

A componente Sinal considera as diferenças locais de intensidade de nível de cinza em uma vizinhança circular para cada *pixel* da imagem; a componente Magnitude relaciona as diferenças locais com uma informação global da imagem de textura; e a componente Centro compara a informação de nível de cinza de cada *pixel* com uma informação global da imagem. Cada componente é caracterizada por uma regra específica de extração de características e gera uma imagem de códigos, I_{CMLMP_S} , I_{CMLMP_M} e I_{CMLMP_C} , composta por um código para cada *pixel* da imagem de textura, com exceção dos *pixels* da borda da imagem.

Seja p um *pixel* da imagem de textura de dimensão $X \times Y$, posicionado em (x, y) , de intensidade de nível de cinza $g(x, y)$ e rodeado por uma vizinhança circular de raio R ($R > 0$) com N vizinhos igualmente espaçados, como ilustrado na Figura 14. Ao se considerar p posicionado na coordenada $(0,0)$, as coordenadas dos *pixels* da vizinhança $p_i \mid_{i=1}^N$ são iguais a $(-R\sin(2\pi(i-1)/N), R\cos(2\pi(i-1)/N))$, e a intensidade de nível de cinza de p_i é igual a g_i . A intensidade de nível de cinza dos vizinhos cuja localização não coincida com o centro de cada *pixel* pode ser estimada por interpolação bilinear (VIEIRA; NEGRI; GONZAGA, 2018; OJALA; PIETIKÄINEN; MÄENPÄÄ, 2002).

As regras de extração de características para as componentes CMLMP_S, CMLMP_M e CMLMP_C aplicadas sobre $p(x, y)$ são apresentadas na Equação (4.1), em

Figura 14: Vizinhança circular de raio $R = 1$ e $N = 8$ para o *pixel* p posicionado em (x, y) .

que h define a componente, $h = \{S, M, C\}$; W é um vetor de pesos para o qual $W_i = 2^{(i-1)}$; B_S , B_M e B_C , com $B \in \mathbb{Z}^+$, são os valores máximos dos códigos nas respectivas imagens de códigos; e f_{h_i} é a função sigmóide de mapeamento, específica para cada componente, definida pela Equação (4.2).

$$\text{CMLMP_}h(x, y) = \left\lfloor \frac{\sum_{i=1}^N f_{h_i} W_i}{\sum_{i=1}^N W_i} (B_h - 1) + 0.5 \right\rfloor \quad (4.1)$$

$$f_{h_i} = \frac{1}{1 + e^{-\frac{g_{diff_{h,i}}}{\beta_h}}} \quad (4.2)$$

A função sigmóide f mapeia o valor da diferença de intensidade de nível de cinza g_{diff} para um padrão na imagem de códigos. A inclinação da curva sigmóide é definida pelo parâmetro (β) , sendo β um número real positivo. Tal como apresentado no Capítulo 2 para outros descritores baseados no LMP, o desempenho do descritor CMLMP é afetado pelo valor de β . Desse modo, faz-se necessário otimizar os valores de β_S , β_M e β_C para cada aplicação e/ou conjunto de imagens. Recomenda-se, também, otimizar os valores de B_S , B_M e B_C .

As diferenças de intensidade de nível de cinza para as componentes Sinal, Magnitude e Centro, $g_{diff_{S,i}}$, $g_{diff_{M,i}}$ e $g_{diff_{C,i}}$, são definidas nas Equações (4.3), (4.4) e (4.5), respectivamente.

$$g_{diff_{S,i}} = g_i - g_{mean} \quad (4.3)$$

$$g_{diff_{M,i}} = |g_i - g_{mean}| - G \quad (4.4)$$

$$g_{diff_{C,i}} = g(x, y) - G \quad (4.5)$$

em que g_{mean} é a média dos valores de intensidade de nível de cinza dos N vizinhos, determinada pela Equação (4.6), e G é o valor médio de intensidade de nível de cinza de toda a imagem $X \times Y$, definido conforme Equação (4.7).

$$g_{mean} = \frac{1}{N} \sum_{i=1}^N g_i \quad (4.6)$$

$$G = \frac{1}{X \cdot Y} \sum_{x=1}^X \sum_{y=1}^Y g_{(x,y)} \quad (4.7)$$

Uma vez determinadas as imagens de códigos, geram-se diferentes vetores de características para representar a imagem de textura, sendo consideradas abordagem individual e abordagens de fusão. Para a abordagem individual, cada componente é considerada um descritor de textura, sendo que a representação 1D das imagens de códigos I_{CMLMP_S} , I_{CMLMP_M} e I_{CMLMP_C} gera os respectivos vetores de características, cujo número de elementos é B_S , B_M e B_C , respectivamente.

Nas abordagens de fusão, diferentes operadores de fusão são aplicados sobre os vetores de características, assim como sobre as imagens de características das componentes Sinal, Magnitude e Centro. Quatro diferentes abordagens são geradas para cada operador de fusão e para cada configuração de vizinhança:

- $CMLMP_S_{R,N}(Op_{Fusao})CMLMP_M_{R,N}$
- $CMLMP_S_{R,N}(Op_{Fusao})CMLMP_C$
- $CMLMP_M_{R,N}(Op_{Fusao})CMLMP_C$
- $CMLMP_S_{R,N}(Op_{Fusao})CMLMP_M_{R,N}(Op_{Fusao})CMLMP_C$

O termo (Op_{Fusao}) representa o operador de fusão a ser utilizado. O número de elementos do vetor de características gerado para cada abordagem é função de B_S , B_M e B_C e depende do operador de fusão utilizado.

4.1.2 Completed Median Local Mapped Pattern (CMedianLMP)

O descritor *Completed Median Local Mapped Pattern* (CMedianLMP) diferencia-se do CMLMP pela regra de extração de características, sendo aplicado o método de extração de características do descritor Median-LMP (SOUZA; GONZAGA, 2019). É também composto por três diferentes componentes, CMedianLMP_S, CMedianLMP_M e CMedianLMP_C, sendo possível representar a textura por diferentes vetores de características, resultados da combinação das diferentes componentes.

Cada componente é caracterizada por uma regra específica de extração das características e gera uma imagem de códigos, $I_{CMedianLMP_S}$, $I_{CMedianLMP_M}$ e $I_{CMedianLMP_C}$, composta por um código para cada *pixel* da imagem de textura, com exceção dos *pixels* da borda da imagem. O método de extração das características para as diferentes componentes do descritor CMedianLMP, considera a vizinhança circular apresentada para o CMLMP, mas também a vizinhança quadrada 3×3 centrada na coordenada (x, y) , tal como ilustrado na Figura 15.

Figura 15: Vizinhança 3×3 e vizinhança circular de raio $R = 1$ e $N = 8$ para o *pixel* p posicionado em (x, y) .

As regras de extração de características para as componentes CMedianLMP_S, CMedianLMP_M e CMedianLMP_C aplicadas sobre $p(x, y)$ são apresentadas na Equação (4.8), em que $h = \{S, M, C\}$; W é um vetor de pesos para o qual $W_i = 2^{(i-1)}$; B_S , B_M e B_C , com $B \in \mathbb{Z}^+$, são os valores máximos dos códigos nas respectivas imagens de códigos; e f_{h_i} é a função sigmóide de mapeamento, específica para cada componente, definida pela Equação (4.2). Recomenda-se otimizar os valores dos parâmetros β_S , β_M , β_C , B_S , B_M e B_C a fim de se obter melhor desempenho na aplicação específica.

$$CMedianLMP_h(x, y) = \left\lceil \frac{\sum_{i=1}^N f_{h_i} W_i}{\sum_{i=1}^N W_i} (B_h - 1) + 0.5 \right\rceil \quad (4.8)$$

As diferenças de intensidade de nível de cinza para as componentes Sinal, Magnitude e Centro, $g_{diff_{S,i}}$, $g_{diff_{M,i}}$ e $g_{diff_{C,i}}$, são definidas nas Equações (4.9), (4.10) e (4.11), respectivamente.

$$g_{diff_{S,i}} = g_i - g_{median} \quad (4.9)$$

$$g_{diff_{M,i}} = |g_i - g_{median}| - G_{median} \quad (4.10)$$

$$g_{diff_{C,i}} = g_{median} - G_{median} \quad (4.11)$$

em que g_i é a intensidade de nível de cinza do i -ésimo *pixel* da vizinhança circular, g_{median} é a mediana dos valores de intensidade de nível de cinza dos *pixels* da região 3×3 centrada na coordenada (x, y) , determinada pela Equação (4.12), e G_{median} é o valor médio de intensidade de nível de cinza da imagem $X \times Y$ filtrada pelo filtro da mediana, considerando janela quadrada de dimensão 3×3 , definido conforme Equação (4.13).

$$g_{median} = \text{mediana} \left(g(x-1, y-1), g(x-1, y), g(x-1, y+1), g(x, y-1), \right. \\ \left. g(x, y), g(x, y+1), g(x+1, y-1), g(x+1, y), g(x+1, y+1) \right) \quad (4.12)$$

$$G_{median} = \frac{1}{(X-2) \cdot (Y-2)} \sum_{x=2}^{X-1} \sum_{y=2}^{Y-1} g_{median} \quad (4.13)$$

Assim como apresentado para o descritor CMLMP, geram-se diferentes vetores de características para representar a imagem de textura, sendo consideradas abordagem individual e abordagens de fusão. Para a abordagem individual, cada componente é considerada um descritor de textura, sendo que a representação 1D das imagens de códigos $I_{CMedianLMP_S}$, $I_{CMedianLMP_M}$ e $I_{CMedianLMP_C}$ gera os respectivos vetores de características, cujo número de elementos é B_S , B_M e B_C , respectivamente.

Nas abordagens de fusão, diferentes operadores de fusão são aplicados sobre os vetores de características, assim como sobre as imagens de características das componentes Sinal, Magnitude e Centro. Quatro diferentes abordagens são geradas para cada operador de fusão e para cada configuração de vizinhança:

- CMedianLMP_ $S_{R,N}$ (Op_{Fusao})CMedianLMP_ $M_{R,N}$
- CMedianLMP_ $S_{R,N}$ (Op_{Fusao})CMedianLMP_ C
- CMedianLMP_ $M_{R,N}$ (Op_{Fusao})CMedianLMP_ C
- CMedianLMP_ $S_{R,N}$ (Op_{Fusao})CMedianLMP_ $M_{R,N}$ (Op_{Fusao})CMedianLMP_ C

O termo (Op_{Fusao}) representa o operador de fusão a ser utilizado. O número de elementos do vetor de características gerado para cada abordagem é função de B_S , B_M e B_C e depende do operador de fusão utilizado.

4.1.3 Completed Z with Tilted Z Local Mapped Pattern (CZTZLMP)

O descritor de texturas *Completed Z with Tilted Z Local Mapped Pattern* (CZTZLMP) apresenta a mesma metodologia de geração de componentes complementares descrita para os descritores CMLMP e CMedianLMP, sendo utilizada regra de extração de características

baseada no descritor ZTZLBP (MARY; DEJEY, 2018). As componentes CZTZLMP_S, CZTZLMP_M e CZTZLMP_C são caracterizadas por uma regra específica de extração de características e geram as imagens de códigos $I_{CZTZLMP_S}$, $I_{CZTZLMP_M}$ e $I_{CZTZLMP_C}$, compostas por um código para cada *pixel* da imagem de textura, com exceção dos *pixels* da borda da imagem. A componente CZTZLMP_C possui a mesma regra de extração de características apresentada para CMLMP_C, sendo assim, $I_{CZTZLMP_C} = I_{CMLMP_C}$.

O método de extração das características para as componentes Sinal e Magnitude do descritor CZTZLMP considera K arranjos de *pixels* não sobrepostos da vizinhança circular de raio R com N vizinhos igualmente espaçados em torno do *pixel* $p(x, y)$, sendo um arranjo em Z e $K - 1$ arranjos em Z inclinado ($TZ_i \mid_{i=1}^{K-1}$). O número de arranjos K é função do número de vizinhos em torno de $p(x, y)$, sendo $K = \frac{N}{4}$. Mostram-se, na Figura 16, os arranjos em Z e em Z inclinado considerando vizinhança circular de raio unitário composta por oito vizinhos.

Figura 16: Diferentes arranjos da vizinhança circular de raio $R = 1$ e $N = 8$ para extração das características CZTZLMP: (a) arranjo em Z e (b) arranjo em Z inclinado (TZ).

A fim de ilustrar o agrupamento dos *pixels* vizinhos para diferentes configurações de vizinhança, mostram-se, na Figura 17, os K arranjos da vizinhança circular para extração das características CZTZLMP, sendo $K = 2$ para $R = 1$ e $N = 8$, $K = 4$ para $R = 2$ e $N = 16$, e $K = 6$ para $R = 3$ e $N = 24$.

As regras de extração das características para as componentes CZTZLMP_S e CZTZLMP_M aplicadas sobre $p(x, y)$ são apresentadas na Equação (4.14), em que h indica a componente, sendo $h = \{S, M\}$; W é o vetor de pesos para o qual $W_i = 2^{(i-1)}$; B_S e B_M , com $B \in \mathbb{Z}^+$, são os valores máximos dos códigos nas respectivas imagens de códigos; e $f_{h_{k,i}}$ é a função sigmóide de mapeamento da i -ésima diferença de intensidade de nível de cinza entre os *pixels* do k -ésimo arranjo da vizinhança circular, específica para cada componente e definida pela Equação (4.15).

$$CZTZLMP_h_{R,N}(x, y) = \frac{1}{K} \sum_{k=1}^K \left[\frac{\sum_{i=1}^4 f_{h_{k,i}} W_i}{\sum_{i=1}^4 W_i} (B_h - 1) + 0.5 \right] \quad (4.14)$$

Figura 17: K arranjos da vizinhança circular para extração das características CZTZLMP considerando (a) $R = 1$ e $N = 8$, (b) $R = 2$ e $N = 16$ e (c) $R = 3$ e $N = 24$.

$$f_{h_{k,i}} = \frac{1}{1 + e^{-\frac{gdiff_{h_{k,i}}}{\beta_h}}} \quad (4.15)$$

As diferenças de intensidade de nível de cinza para a componente Sinal considerando N *pixels* igualmente espaçados em torno do *pixel* central $p(x, y)$, cuja intensidade de nível de cinza é $g(x, y)$, são definidas na Equação (4.16), em que o índice a_k , definido na Equação (4.17), indica o *pixel* inicial e o índice m_k , definido na Equação (4.18), o *pixel* final do k -ésimo arranjo da vizinhança.

$$\begin{aligned} gdiff_{S_{k,1}} &= g_{a_k} - g_{a_k + \frac{N}{4}} \\ gdiff_{S_{k,2}} &= g_{a_k + \frac{N}{4}} - g(x, y) \\ gdiff_{S_{k,3}} &= g(x, y) - g_{a_k - \frac{N}{4}} \\ gdiff_{S_{k,4}} &= g_{a_k - \frac{N}{4}} - g_{m_k} \end{aligned} \quad (4.16)$$

$$a_k = \left\lfloor \frac{5}{8}N - k + 2 \right\rfloor \quad (4.17)$$

$$m_k = \begin{cases} \left\lfloor \frac{1}{8}N - k + 2 \right\rfloor, & k < \left\lfloor \frac{1}{8}N + 2 \right\rfloor \\ \left\lfloor \frac{9}{8}N - k + 2 \right\rfloor, & k \geq \left\lfloor \frac{1}{8}N + 2 \right\rfloor \end{cases} \quad (4.18)$$

Para a componente Magnitude, as diferenças de intensidade de nível de cinza entre os *pixels* dispostos no k -ésimo arranjo, $gdiff_{M_{k,i}}$, são definidas na Equação (4.19). Elas relacionam a magnitude das respectivas diferenças de intensidade de nível de cinza obtidas para a componente Sinal com a informação global G , sendo G o valor médio de intensidade

de nível de cinza de toda a imagem de textura de dimensão $X \times Y$, definido conforme Equação (4.20).

$$g_{diff_{M_{k,i}}} = |g_{diff_{S_{k,i}}}| - G \quad (4.19)$$

$$G = \frac{1}{X \cdot Y} \sum_{x=1}^X \sum_{y=1}^Y g_{(x,y)} \quad (4.20)$$

Como mencionado, a regra de extração das características para a componente Centro, CZTZLMP_C, é igual à apresentada para CMLMP_C, sendo definida pela Equação (4.21).

$$CZTZLMP_C = \left\lfloor \frac{1}{1 + e^{-\frac{g(x,y) - G}{\beta_C}}} (B_C - 1) + 0.5 \right\rfloor \quad (4.21)$$

Para se obter melhor desempenho na aplicação específica, recomenda-se otimizar os valores dos parâmetros β_S , β_M , β_C , B_S , B_M e B_C .

Da mesma maneira como definido para os descritores CMLMP e CMedianLMP, uma vez determinadas as imagens de códigos $I_{CZTZLMP_S}$, $I_{CZTZLMP_M}$ e $I_{CZTZLMP_C}$, os vetores de características para cada componente individual são gerados pela representação 1D das respectivas imagens de códigos, cujo número de elementos é B_S , B_M e B_C , respectivamente.

Nas abordagens de fusão, diferentes operadores de fusão são aplicados sobre os vetores de características, assim como sobre as imagens de características das componentes Sinal, Magnitude e Centro. Quatro diferentes abordagens são geradas para cada operador de fusão e para cada configuração de vizinhança, sendo eles:

- CZTZLMP_ $S_{R,N}$ (Op_{Fusao})CZTZLMP_ $M_{R,N}$
- CZTZLMP_ $S_{R,N}$ (Op_{Fusao})CZTZLMP_C
- CZTZLMP_ $M_{R,N}$ (Op_{Fusao})CZTZLMP_C
- CZTZLMP_ $S_{R,N}$ (Op_{Fusao})CZTZLMP_ $M_{R,N}$ (Op_{Fusao})CZTZLMP_C

O termo (Op_{Fusao}) representa o operador de fusão a ser utilizado. O número de elementos do vetor de características gerado para cada abordagem é função de B_S , B_M e B_C e depende do operador de fusão utilizado.

4.2 Operadores de Fusão

4.2.1 Espaço de Fusão de Características

Sejam $\mathbf{v}_1 = [v_{11} \ \cdots \ v_{1B_1}]$, $\mathbf{v}_2 = [v_{21} \ \cdots \ v_{2B_2}]$, $\mathbf{v}_3 = [v_{31} \ \cdots \ v_{3B_3}]$ e $\mathbf{v}_N = [v_{N1} \ \cdots \ v_{NB_N}]$ vetores de características com número de elementos igual a B_1 , B_2 , B_3 e B_N , respectivamente.

$\mathbb{M} = \mathbf{v}_1 \| \mathbf{v}_2 \| \cdots \| \mathbf{v}_N$ é a matriz de características composta pelos N vetores de características, definida na Equação (4.22), cuja dimensão é $N \times B_{max}$, em que $B_{max} = \max(B_1, B_2, \cdots, B_N)$. O operador $\|$ indica que cada vetor de características $\mathbf{v}_i$ é disposto como uma linha da matriz de características. Caso os vetores de características sejam de dimensões diferentes, adota-se preenchimento por zero. Os elementos de $\mathbb{M}$ são representados por m_{ij} , $\mathbb{M} = (m_{ij})$, em que $i \in \{1, \cdots, N\}$ e $j \in \{1, \cdots, B_{max}\}$; e a j -ésima coluna da matriz de características é representada por $\mathbb{M}_{kj}$.

$$\begin{aligned} \mathbb{M} &= \mathbf{v}_1 \| \mathbf{v}_2 \| \cdots \| \mathbf{v}_N \\ &= \begin{bmatrix} \mathbf{v}_1 & \mathbf{0}_{1 \times (B_{max} - B_1)} \\ \mathbf{v}_2 & \mathbf{0}_{1 \times (B_{max} - B_2)} \\ \vdots & \vdots \\ \mathbf{v}_N & \mathbf{0}_{1 \times (B_{max} - B_N)} \end{bmatrix} \\ &= [\mathbb{M}_{k1} \ \mathbb{M}_{k2} \ \cdots \ \mathbb{M}_{kB_{max}}] \end{aligned} \quad (4.22)$$

A matriz de características composta pelos vetores $\mathbf{v}_1$, $\mathbf{v}_2$ e $\mathbf{v}_3$, com $B_1 = B_2$ e $B_3 > B_1$, é definida na Equação(4.23).

$$\begin{aligned} \mathbb{M} &= \mathbf{v}_1 \| \mathbf{v}_2 \| \mathbf{v}_3 \\ &= \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1B_1} & 0 & 0 & \cdots & 0 \\ v_{21} & v_{22} & \cdots & v_{2B_1} & 0 & 0 & \cdots & 0 \\ v_{31} & v_{32} & \cdots & v_{3B_1} & v_{3B_1+1} & v_{3B_1+2} & \cdots & v_{3B_3} \end{bmatrix} \\ &= \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1B_3} \\ m_{21} & m_{22} & \cdots & m_{2B_3} \\ m_{31} & m_{32} & \cdots & m_{3B_3} \end{bmatrix} \\ &= [\mathbb{M}_{k1} \ \mathbb{M}_{k2} \ \cdots \ \mathbb{M}_{kB_3}] \end{aligned} \quad (4.23)$$

Propõem-se dois novos operadores de fusão de vetores de características, os quais são aplicados aos vetores de características dispostos na matriz $\mathbb{M}$. Como $\mathbb{M}$ representa a disposição dos vetores de características no formato adequado para serem combinados, define-se $\mathbb{M}$ como sendo o espaço de fusão de características.

Os operadores de fusão propostos são denominados operador Soma no espaço de fusão ($\|_+$) e operador Gram no espaço de fusão ($\|_G$).

4.2.2 Operador Soma no espaço de fusão

O operador Soma $^{\mathbb{M}}$ ($\|_+$), lê-se “Soma no espaço de fusão”, aplicado a dois ou mais vetores de características, $\mathbf{v}_{Soma_{\mathbb{M}}} = \mathbf{v}_1 \|_+ \mathbf{v}_2 \|_+ \cdots \|_+ \mathbf{v}_N$, realiza a soma algébrica dos elementos em $\mathbb{M}_{kj}$, tal como definido na equação (4.24). O número de elementos do vetor $\mathbf{v}_{Soma_{\mathbb{M}}}$ é igual ao máximo número de elementos dentre os vetores de características combinados, $B_{Soma_{\mathbb{M}}} = B_{max}$.

$$\begin{aligned} \mathbf{v}_{Soma_{\mathbb{M}}} &= \mathbf{v}_1 \|_+ \mathbf{v}_2 \|_+ \cdots \|_+ \mathbf{v}_N \\ &= \left[\sum_{k=1}^N m_{k1} \quad \sum_{k=1}^N m_{k2} \quad \cdots \quad \sum_{k=1}^N m_{kB_{max}} \right] \end{aligned} \quad (4.24)$$

A fusão por Soma $^{\mathbb{M}}$ de três vetores de características, $\mathbf{v}_1$, $\mathbf{v}_2$ e $\mathbf{v}_3$, é ilustrada na Figura 18.

Figura 18: Exemplo de fusão por Soma $^{\mathbb{M}}$ de três vetores de características.

4.2.3 Operador Gram no espaço de fusão

O operador $\text{Gram}^{\mathbb{M}} (\|_G)$, lê-se “Gram no espaço de fusão”, aplicado a dois ou mais vetores de características, $\mathbf{v}_{\text{Gram}^{\mathbb{M}}} = \mathbf{v}_1 \|_G \mathbf{v}_2 \|_G \cdots \|_G \mathbf{v}_N$, determina a matriz de Gram (PARK et al., 2018; GATYS; ECKER; BETHGE, 2015) das características dispostas no espaço de fusão, $\mathbf{G}_{\mathbb{M}} = \mathbb{M}^t \cdot \mathbb{M}$, definida na Equação (4.25), seguido pela representação $1D$ da matriz gerada.

$$\begin{aligned}
 \mathbf{G}_{\mathbb{M}} &= \mathbb{M}^t \cdot \mathbb{M} \\
 &= \begin{bmatrix} \mathbb{M}_{k1}^t \\ \mathbb{M}_{k2}^t \\ \vdots \\ \mathbb{M}_{kB_{max}}^t \end{bmatrix} \cdot \begin{bmatrix} \mathbb{M}_{k1} & \mathbb{M}_{k2} & \cdots & \mathbb{M}_{kB_{max}} \end{bmatrix} \\
 &= \begin{bmatrix} \mathbb{M}_{k1}^t \cdot \mathbb{M}_{k1} & \mathbb{M}_{k1}^t \cdot \mathbb{M}_{k2} & \cdots & \mathbb{M}_{k1}^t \cdot \mathbb{M}_{kB_{max}} \\ \mathbb{M}_{k2}^t \cdot \mathbb{M}_{k1} & \mathbb{M}_{k2}^t \cdot \mathbb{M}_{k2} & \cdots & \mathbb{M}_{k2}^t \cdot \mathbb{M}_{kB_{max}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{M}_{kB_{max}}^t \cdot \mathbb{M}_{k1} & \mathbb{M}_{kB_{max}}^t \cdot \mathbb{M}_{k2} & \cdots & \mathbb{M}_{kB_{max}}^t \cdot \mathbb{M}_{kB_{max}} \end{bmatrix} \quad (4.25) \\
 &= \begin{bmatrix} \sum_{k=1}^N m_{k1}^2 & \sum_{k=1}^N (m_{k1} \cdot m_{k2}) & \cdots & \sum_{k=1}^N (m_{k1} \cdot m_{kB_{max}}) \\ \sum_{k=1}^N (m_{k1} \cdot m_{k2}) & \sum_{k=1}^N m_{k2}^2 & \cdots & \sum_{k=1}^N (m_{k2} \cdot m_{kB_{max}}) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^N (m_{k1} \cdot m_{kB_{max}}) & \sum_{k=1}^N (m_{k2} \cdot m_{kB_{max}}) & \cdots & \sum_{k=1}^N m_{kB_{max}}^2 \end{bmatrix} \\
 &= \begin{bmatrix} G_{\mathbb{M}11} & G_{\mathbb{M}12} & \cdots & G_{\mathbb{M}1B_{max}} \\ G_{\mathbb{M}21} & G_{\mathbb{M}22} & \cdots & G_{\mathbb{M}2B_{max}} \\ \vdots & \vdots & \ddots & \vdots \\ G_{\mathbb{M}B_{max}1} & G_{\mathbb{M}B_{max}2} & \cdots & G_{\mathbb{M}B_{max}B_{max}} \end{bmatrix}
 \end{aligned}$$

Cada elemento $G_{\mathbb{M}ij}$ da matriz de Gram $\mathbf{G}_{\mathbb{M}}$, definido na Equação (4.26), é resultado do produto escalar entre os elementos $\mathbb{M}_{ki}$ e $\mathbb{M}_{kj}$, $i, j \in [1, B_{max}]$.

$$\begin{aligned}
 G_{\mathbb{M}ij} &= \langle \mathbb{M}_{ki}, \mathbb{M}_{kj} \rangle \\
 &= \sum_{k=1}^N (m_{ki} \cdot m_{kj}) \quad (4.26)
 \end{aligned}$$

Como a matriz $\mathbf{G}_{\mathbb{M}}$ é simétrica, o vetor de características combinado por operador de fusão $\text{Gram}^{\mathbb{M}}$, $\mathbf{v}_{\text{Gram}^{\mathbb{M}}}$, definido na Equação (4.27), é o resultado da justaposição dos elementos da diagonal principal da matriz $\mathbf{G}_{\mathbb{M}}$ e dos elementos acima da diagonal principal,

$G_{M_{ij}}$, $j \geq i$, como ilustrado na Figura 19.

$$\begin{aligned} \mathbf{v}_{GRAM_M} &= \mathbf{v}_1 \parallel_G \mathbf{v}_2 \parallel_G \cdots \parallel_G \mathbf{v}_N \\ &= \begin{bmatrix} G_{M_{11}} & \cdots & G_{M_{1B_{max}}} & G_{M_{22}} & \cdots & G_{M_{2B_{max}}} & \cdots & G_{M_{B_{max}B_{max}}} \end{bmatrix} \end{aligned} \quad (4.27)$$

Figura 19: Exemplo de fusão por $Gram^M$ de três vetores de características.

O número de elementos do vetor $\mathbf{v}_{Gram_M}$, B_{Gram_M} , é definido pela Equação (4.28).

$$B_{Gram_M} = \frac{B_{max}^2 + B_{max}}{2} \quad (4.28)$$

4.3 Considerações Finais

Neste capítulo, foram propostos três novos descritores de textura baseados em fusão de características, *Completed Mean Local Mapped Pattern* (CMLMP), *Completed Median Local Mapped Pattern* (CMedianLMP) e *Completed Z with Tilted Z Local Mapped Pattern* (CZTZLMP), todos os três baseados na metodologia de extração de características complementares da imagem, representada pelas componentes Sinal, Magnitude e Centro. Diferenciam-se, uns dos outros, pela regra de extração das características, sendo baseados na regra dos descritores de textura M-LMP, Median-LMP e ZTZLBP, respectivamente. Cada um dos descritores gera diferentes representações de uma mesma imagem (diferentes vetores de características) por meio da aplicação de operadores de fusão sobre as diferentes componentes.

Apresentou-se, também, a proposta de dois novos operadores de fusão de características, sendo formalizado um novo espaço de representação de vetores de características, denominado espaço de fusão de características. Os operadores de fusão propostos foram denominados Soma no espaço de fusão e operador Gram no espaço de fusão.

As metodologias propostas foram aplicadas a registro de imagens e a reconhecimento de distorção arquitetural em mamografia digital. Os resultados obtidos foram comparados aos de diversos outros descritores de textura e operadores de fusão publicados na literatura. Os desempenhos obtidos nas aplicações citadas são apresentados nos Capítulos 5 e 6, respectivamente.

5 FUSÃO DE DESCRITORES DE TEXTURA APLICADA A REGISTRO DE IMAGENS

Nos capítulos 2 e 3, discorreu-se sobre a importância de se descrever adequadamente as imagens de textura, a fim de se determinar uma característica, ou um vetor de características, que permita diferenciar a respectiva imagem das demais, e sobre a possibilidade da fusão, ou combinação, de diferentes métodos gerar melhores resultados que os obtidos com a aplicação de cada método individualmente.

Apresentaram-se as características e a formulação matemática de alguns descritores de textura utilizados na literatura, sendo observado que, para determinadas aplicações, ou para determinados conjuntos de imagens, alguns descritores resultam em melhor desempenho que outros, sendo desconhecido um descritor que apresente o melhor resultado dentre os demais para toda e qualquer aplicação e/ou para todo e qualquer conjunto de imagens, fato este que impulsiona diversos pesquisadores a tentarem desenvolver descritores de melhor desempenho.

Foram descritos, também, diferentes topologias de fusão, as quais consideram a aplicação de operadores de fusão sobre diversas etapas do processo de classificação de imagens, assim como algumas técnicas utilizadas e as vantagens de aplicação, evidenciando-se que a fusão da informação permite integrar as características complementares dos diferentes métodos combinados, melhorando o resultado do processo em que é utilizada.

Tendo como base a relevância da descrição adequada de imagens de textura e o ganho de desempenho proporcionado pela fusão da informação, apresentam-se, neste capítulo, a metodologia e os resultados obtidos para a fusão de descritores de textura aplicada a registro de imagens.

5.1 Considerações iniciais

A determinação dos pares de *pixels* correspondentes, ou de regiões correspondentes, em diferentes imagens, representa um desafio em sistemas de visão computacional (LUTZ et al., 2015; MARSHALL; DOUCETTE, 2011; BAUMBERG, 2000). Esta correspondência de imagens, também denominada registro de imagens, tem vasta aplicação em sistemas de reconstrução de informação tridimensional de uma cena, robótica móvel, rastreamento de objetos, sensoriamento remoto, inspeção industrial e controle de qualidade, biometria, vigilância e segurança, dentre outras (LUTZ et al., 2015; NANNI; LUMINI, 2013; MARSHALL; DOUCETTE, 2011; MA; CHAN; CANTERS, 2010; HOSSAIN et al., 2010; HUANG et al., 2008).

As imagens a serem analisadas podem ter sido capturadas em diferentes instantes de tempo, de diferentes pontos de vista e/ou por diferentes câmeras (LUTZ et al., 2015;

HUANG et al., 2008). Para realizar a correspondência dessas imagens, são definidas métricas de similaridade, sendo comumente utilizados descritores de textura, os quais consideram tanto a informação de intensidade de nível de cinza dos *pixels* de uma imagem como a de localização (PATIL et al., 2013; MA; CHAN; CANTERS, 2010; HOSSAIN et al., 2010; ROS; LAURENT, 2006; BAUMBERG, 2000).

Como a fusão da informação combina as características complementares dos diferentes métodos considerados, resultando em melhor desempenho que o obtido com a aplicação de cada método individualmente, enuncia-se a hipótese de que a fusão de descritores de textura aplicada a registro de imagens resulta em desempenho melhor que o obtido com a aplicação dos descritores individuais.

5.2 Materiais e Métodos

Com o propósito de verificar a validade da hipótese de que a fusão de descritores de textura melhora o desempenho do processo de determinação de pares correspondentes, desenvolveu-se um procedimento para aplicação de diferentes operadores de fusão na camada de características e na camada de decisão. Na camada de características, os operadores de fusão Soma e Concatenação foram aplicados sobre os vetores de características dos descritores de textura SIFT, CS-LBP, CS-LMP, M-LMP e ZTZLBP, enquanto que na camada de decisão, aplicaram-se os operadores de fusão Mínimo, Adição, Produto e Combinação sobre a informação de distância euclidiana entre os vetores de características dos pares de regiões de interesse para diferentes imagens.

5.2.1 Definição dos Conjuntos de Imagens

Os diferentes operadores de fusão são aplicados sobre um conjunto de imagens de treinamento e a validação do desempenho se dá pela aplicação destes sobre um conjunto de imagens de teste. Ambos conjuntos de imagens adotados¹ são vastamente utilizados pela literatura (YUN-XI; FANG, 2016; FERRAZ; PEREIRA-JR; GONZAGA, 2014a; FERRAZ et al., 2014; KIM; YOO; SOHN, 2013; JUAN; GWUN, 2009; HEIKKILÄ; PIETIKÄINEN; SCHMID, 2006; MIKOLAJCZYK et al., 2005; MIKOLAJCZYK; SCHMID, 2005).

O conjunto de treinamento, mostrado na Figura 20, é composto por oito pares de imagens, os quais apresentam diferentes tipos de transformações, sendo elas: escala e rotação (pares de imagens (a), (b) e (e), (f)), borramento (pares de imagens (c), (d) e (k), (l)), mudança de ponto de vista (pares de imagens (g), (h) e (o), (p)), mudança de iluminação (par de imagens (i), (j)) e compressão JPEG (par de imagens (m), (n)).

Nas Figuras 21 e 22, apresentam-se os quinze pares de imagens que compõem o conjunto de teste, os quais possuem três tipos de transformações: somente escala (pares de

¹ Disponíveis em <http://www.robots.ox.ac.uk/~vgg/research/affine/> e <http://lear.inrialpes.fr/people/mikolajczyk/>.

Figura 20: Imagens do conjunto de treinamento: bark e boat (escala e rotação), bike e trees (imagem borrada), graf e wall (mudança de ponto de vista), leuven (mudança de iluminação), ubc (compressão JPEG).

imagens 21(a),(b); 21(c),(d); 21(e),(f); 21(g),(h); 22(a),(b); e 22(i),(j)), escala e rotação (pares de imagens 21(i),(j); 21(k),(l); 21(m),(n); 21(o),(p); 22(c),(d); e 22(g),(h)), e somente rotação (pares de imagens 22(e),(f); 22(k),(l); e 22(m),(n)).

5.2.2 Detecção das Regiões de Interesse

Uma vez definidos os conjuntos de imagens, deve-se proceder à detecção das regiões de interesse. Para tal, aplicou-se o detector de regiões de interesse *Hessian Affine*, com limiar igual a 500, disponibilizado por Mikolajczyk ² sobre as imagens dos conjuntos de treinamento e de teste, obtendo $s_{1_{Img}}$ regiões de interesse na pose 1 da imagem Img e $s_{2_{Img}}$ regiões de interesse na pose 2 da respectiva imagem.

Utilizou-se o detector de regiões *Hessian Affine*, pois, de acordo com Mikolajczyk et al. (2005), não existe um detector de regiões de interesse que possua o melhor desempenho

² Disponível em <<http://www.robots.ox.ac.uk/~vgg/research/affine/detectors.html>>.

Figura 21: Imagens do conjunto de teste: asterix, belledone, bip e crolle (escala); east park, east south, ensigma e inria (escala e rotação).

para todo tipo de transformação e para todo tipo de imagem, no entanto, demonstrou-se que os detectores MSER e *Hessian Affine* obtiveram os melhores resultados para os experimentos realizados, sendo que os detectores *Hessian Affine* e *Harris Affine* são os que retornam um maior número de regiões de interesse, sendo indicados para correspondência de imagens; além disso, Ferraz et al. (2014) e Mikolajczyk e Schmid (2005) mostraram que o detector *Hessian Affine* permite melhores resultados que o *Harris Affine*.

5.2.3 Determinação do *Ground Truth*

Cada par de regiões de interesse é um possível par correspondente, podendo ser par correspondente, caso seja uma correspondência correta (quando a i -ésima região da pose 1 da imagem corresponda à j -ésima região da pose 2 da respectiva imagem), ou um par não correspondente, caso seja uma falsa correspondência. O conjunto de pares correspondentes (corretas correspondências previamente conhecidas ou *ground truth*) foi

Figura 22: Imagens do conjunto de teste: lap e vangogh (escala); laptop e resid (escala e rotação); new york, zmars e zmonet (rotação).

gerado considerando o critério de erro de sobreposição³, utilizado por [Heikkilä, Pietikäinen e Schmid \(2009\)](#) e por [Mikolajczyk e Schmid \(2002\)](#).

O erro de sobreposição, determinado pela Equação (5.1), é definido pela razão entre a interseção e a união das duas regiões de interesse, A e B, quantificando o grau de similaridade, ou de correspondência, entre as duas regiões, sendo A, para o trabalho proposto, a i-ésima região da pose 1 de uma determinada imagem, e B, a j-ésima região da pose 2 da respectiva imagem. [Ferraz, Pereira-Jr e Gonzaga \(2013\)](#) afirmam que o erro de sobreposição representa a informação do “quanto uma região A da imagem 1 se sobrepõe a uma região B da imagem 2, sendo a imagem 2 projetada sobre a imagem 1 por uma

³ Traduzido do inglês *overlap error*

matriz de homografia H'' .

$$\epsilon_S = 1 - \frac{A \cap H^T B H}{A \cup H^T B H} \quad (5.1)$$

Sendo assim, para se determinar o erro de sobreposição, é necessário conhecer a matriz de homografia H entre as poses 1 e 2 de uma determinada imagem. No caso em estudo, a matriz de homografia, para cada par de imagens utilizado, foi disponibilizada juntamente com o banco de imagens considerado, tanto para o conjunto de treinamento como para o conjunto de teste.

De acordo com Ferraz et al. (2014) e com Heikkilä, Pietikäinen e Schmid (2009), consideram-se correspondências corretas todas as possíveis correspondências cujo erro de sobreposição for menor que 50% ($\epsilon_S < 0,5$).

5.2.4 Aplicação dos Descritores de Textura

Conhecidas as regiões de interesse de cada par de imagens (poses 1 e 2), aplicam-se sobre elas os descritores SIFT, CS-LBP, CS-LMP, M-LMP e ZTZLBP. O descritor SIFT foi aplicado usando o software específico disponibilizado por Lowe (2004)⁴; os descritores CS-LBP, CS-LMP e M-LMP foram aplicados conforme especificado no Capítulo 2, sendo adotado $R = 2$, $N = 8$ e $B = 16$ para os descritores CS-LMP e M-LMP, $\beta_{CS-LMP} = 0,6$ e $\beta_{M-LMP} = 1,2$. Conforme determinado em (FERRAZ; PEREIRA-JR; GONZAGA, 2014b), esses são os valores R , N , B e β que proporcionaram melhor desempenho para os descritores em questão. Os parâmetros do descritor CS-LBP foram ajustados em $R = 2$, $N = 8$ e $T = 0,01$, que, conforme experimentos relatados por Heikkilä, Pietikäinen e Schmid (2006), permitiram obter melhor resultado para a maioria das imagens testadas.

Para a aplicação do descritor ZTZLBP, combinou-se o algoritmo original apresentado por Mary e Dejeu (2018), descrito no Capítulo 2, com a metodologia do descritor SIFT (LOWE, 2004), tal como adotado para os demais descritores. Considerou-se, ainda, as componentes ZLBP e TZLBP independentes uma da outra, sendo cada uma delas utilizada como um descritor individual. Os descritores ZLBP e TZLBP foram aplicados às regiões de interesse de dimensão 41×41 *pixels*, as quais foram divididas em uma malha cartesiana de dimensão 4×4 . Para a extração de características de cada *pixel*, adotou-se vizinhança circular de raio R igual a 2 com oito vizinhos simetricamente espaçados. Gerou-se, assim, para cada região de interesse, os histogramas 3D ZLBP e TZLBP com 256 bins cada um. Consideraram-se dois diferentes descritores baseados no ZTZLBP, os quais foram gerados pela combinação das componentes individuais, ZLBP e TZLBP, por meio de duas abordagens: operador de fusão Soma e operador de fusão Concatenação. Ao aplicar o operador de fusão Soma sobre os vetores de características ZLBP e TZLBP, gerou-se o

⁴ Disponível em <<http://www.robots.ox.ac.uk/~vgg/research/affine/descriptors.html>>

descritor Z+TZLBP, com 256 elementos, o qual corresponde ao descritor ZTZLBP original, diferindo unicamente por usar uma vizinhança circular em torno do *pixel* central em detrimento a uma região quadrada de dimensão $J \times J$. Com a aplicação do operador de fusão Concatenação sobre os histogramas 3D individuais, gerou-se o descritor $Z \oplus \text{TZLBP}$, com 512 elementos. Para as duas abordagens consideradas, os vetores de características finais são normalizados para obter norma unitária, truncados em 0,2 e renormalizados para norma unitária.

Com a aplicação dos descritores, são gerados os vetores de características da k -ésima região da n -ésima pose da imagem Img para o descritor $Desc$, $\mathbf{v}_{Img_{n,k_Desc}}$, em que k varia de 1 a $s_{1_{Img}}$, n é igual a 1, para pose 1, e 2, para pose 2, e $Desc = \{\text{SIFT}, \text{CS-LBP}, \text{CS-LMP}, \text{M-LMP}, \text{Z+TZLBP}, \text{Z} \oplus \text{TZLBP}\}$. Os vetores de características para o descritor SIFT, $\mathbf{v}_{Img_{n,k_SIFT}}$, possuem 128 posições; enquanto os vetores de características dos descritores CS-LBP, CS-LMP, M-LMP e Z+TZLBP, $\mathbf{v}_{Img_{n,k_CS-LBP}}$, $\mathbf{v}_{Img_{n,k_CS-LMP}}$, $\mathbf{v}_{Img_{n,k_M-LMP}}$ e $\mathbf{v}_{Img_{n,k_Z+TZLBP}}$, respectivamente, possuem 256 posições cada; e, como mencionado, os vetores de características do descritor $Z \oplus \text{TZLBP}$, $\mathbf{v}_{Img_{n,k_Z \oplus \text{TZLBP}}}$, possuem 512 posições.

Com o intuito de atribuir invariância a rotação aos descritores $Z \oplus \text{TZLBP}$, Z+TZLBP, M-LMP, CS-LMP, and CS-LBP, foram consideradas K direções do gradiente quando da formação do histograma de gradientes (LOWE, 2004). Antes da construção de cada descritor, as regiões de interesse normalizadas foram rotacionadas nas K direções do gradiente, sendo gerados K vetores de características para cada região de interesse e para cada descritor⁵, dos quais, foram mantidos unicamente o par de vetores de características para o qual a distância foi mínima. Este procedimento garante melhor desempenho que o obtido pela comparação das regiões de interesse na direção do máximo gradiente.

5.2.5 Métrica de Similaridade

Dispondo dos vetores de características de cada região de interesse das imagens, para os diferentes descritores, determina-se a distância euclidiana entre os vetores de características para cada par de regiões, por meio da Equação (5.2).

$$\mathbf{d}_{Img_{i,j_Desc}} = \|\mathbf{v}_{Img_{1,i_Desc}} - \mathbf{v}_{Img_{2,j_Desc}}\| \quad (5.2)$$

em que $\mathbf{d}_{Img_{i,j_Desc}}$ é a norma $L2$ ou norma euclidiana entre os vetores de características da i -ésima região da pose 1, $\mathbf{v}_{Img_{1,i_Desc}}$, e da j -ésima região da pose 2, $\mathbf{v}_{Img_{2,j_Desc}}$, da imagem Img para o descritor $Desc$.

Cada par de regiões de interesse é, assim, representada por seis valores de distância euclidiana: $\mathbf{d}_{Img_{i,j_SIFT}}$, $\mathbf{d}_{Img_{i,j_CS-LBP}}$, $\mathbf{d}_{Img_{i,j_CS-LMP}}$, $\mathbf{d}_{Img_{i,j_M-LMP}}$, $\mathbf{d}_{Img_{i,j_Z+TZLBP}}$

⁵ <https://github.com/balasanjeevi/utills-matlab>

e $\mathbf{d}_{\text{Im}g_{i,j} - Z \oplus TZLMP}$, sendo que cada par de imagens possui L pares de regiões de interesse, em que $L = s_{1_{\text{Im}g}} \cdot s_{2_{\text{Im}g}}$.

5.2.6 Apresentação dos Resultados - Descritores Aplicados Individualmente

Os resultados da aplicação individual de cada descritor são apresentados no formato de área sobre a curva *recall* x *precision* (AUC), tendo como base a informação de corretas correspondências retornadas e sendo utilizada, como métrica de similaridade, a distância euclidiana entre os vetores de características das regiões de interesse para cada imagem.

Sejam Verdadeiros Positivos (VP), o número de pares de regiões previamente conhecidos como correspondentes que são corretamente classificados como correspondentes; Falsos Negativos (FN), o número de pares de regiões previamente conhecidos como correspondentes e que são erroneamente classificados como não correspondentes; Verdadeiros Negativos (VN), o número de pares de regiões de interesse previamente conhecidos como não correspondentes e que são corretamente classificados como não correspondentes; e Falsos Positivos (FP), o número de pares de regiões previamente conhecidos como não correspondentes e que são erroneamente classificados como correspondentes.

As curvas *recall* x *precision* (BAEZA-YATES; RIBEIRO-NETO, 1999 apud FER-RAZ et al., 2014) apresentam a evolução da precisão da informação de distância euclidiana à medida que se aumenta o número de regiões de interesse retornadas, sendo cada ponto da curva um par (*recall*, *precision*), em que *recall* é a razão entre o número de correspondências corretas retornadas, VP , e o número total de correspondências corretas previamente conhecidas, $VP + FN$ (número de elementos que constituem o *ground truth*), dado pela Equação (5.3),

$$\text{recall} = \frac{VP}{VP + FN} , \quad (5.3)$$

e *precision* é a razão entre o número de correspondências corretas retornadas, VP , e o número de todas as possíveis correspondências retornadas, $VP + FP$ (correspondências corretas e falsas correspondências), conforme Equação (5.4).

$$\text{precision} = \frac{VP}{VP + FP} \quad (5.4)$$

Para gerar as curvas *recall* x *precision* para cada descritor, os pares de regiões possíveis correspondentes são ordenados em ordem crescente de distância euclidiana, iniciando a leitura pelo par de menor distância. Caso o respectivo par de regiões retornado for correspondente, VP é acrescido de uma unidade, aumentando o *recall* e fazendo com que a precisão atual seja maior ou igual à precisão anterior, caso contrário, apenas FP é acrescido de uma unidade, mantendo *recall* inalterado e reduzindo a precisão. Quanto mais próximo o ponto de inflexão da curva *recall* x *precision* estiver do ponto (1,1) no

gráfico, melhor é o desempenho do descritor, ou seja, quanto maior for a área sobre a curva $recall \times precision$, melhor é o desempenho do descritor, sendo $AUC \in [0, 1]$.

A AUC (FERRAZ et al., 2014; HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009) é calculada pela integral trapezoidal discreta, dada pela Equação (5.5). Neste trabalho de doutorado, a referida métrica de desempenho foi calculada usando a função *trapz()* do *MATLAB*[®], considerando 500 pontos sobre a curva $recall \times precision$, com passo igual a 0,002.

$$AUC \approx \frac{1}{2} \sum_{i=1}^{N-1} (precision_i + precision_{i-1}) \cdot (recall_i - recall_{i-1}) \quad , \quad (5.5)$$

em que N é o número de pontos ($recall$, $precision$) plotados na respectiva curva.

5.2.7 Fusão dos Descritores de Textura

Com o objetivo de melhorar o desempenho de sistemas de registro de imagens, realizou-se a fusão de seis descritores de textura, $Z \oplus TZLBP$, $Z + TZLBP$, SIFT, M-LMP, CS-LMP, e CS-LBP, tomados dois a dois. Para tal, aplicaram-se dois diferentes operadores de fusão na camada de características e quatro operadores de fusão na camada de decisão.

5.2.7.1 Fusão de Descritores de Textura na Camada de Características

Os operadores de fusão Soma e Concatenação são aplicados aos vetores de características de cada região de interesse das imagens, para os diferentes descritores de textura, tomados dois a dois.

Sejam $\mathbf{v}_{Img_{n,k_Desc1}}$ o vetor de características da k -ésima região da n -ésima pose da imagem Img para o descritor $Desc1$, cujo comprimento é B_1 , e $\mathbf{v}_{Img_{n,k_Desc2}}$ o vetor de características da mesma região e da mesma pose da imagem Img , mas extraído para o descritor $Desc2$, com comprimento B_2 .

O vetor de características $\mathbf{v}_{Img_{n,k_Desc1+Desc2}}$, dado pela Equação (5.6), representa o vetor de características da k -ésima região da n -ésima pose da imagem Img para a fusão por Soma dos descritores $Desc1$ e $Desc2$, cujo comprimento B_{Soma} é igual a B_1 , que é igual a B_2 .

$$\mathbf{v}_{Img_{n,k_Desc1+Desc2}} = \mathbf{v}_{Img_{n,k_Desc1}} + \mathbf{v}_{Img_{n,k_Desc2}} \quad (5.6)$$

Ao utilizar o operador Concatenação sobre os vetores de características dos descritores individuais, $Desc1$ e $Desc2$, gerou-se $\mathbf{v}_{Img_{n,k_Desc1 \oplus Desc2}}$, dado pela Equação (5.7),

cujo comprimento B_{Conc} é igual a $B_1 + B_2$.

$$\mathbf{V}_{Img_{n,k_Desc1 \oplus Desc2}} = \mathbf{V}_{Img_{n,k_Desc1}} \oplus \mathbf{V}_{Img_{n,k_Desc2}} \quad (5.7)$$

Foram gerados quatorze diferentes vetores combinados por Concatenação ao aplicar o respectivo operador de fusão sobre os conjuntos de descritores $D_1 = [Z \oplus TZLBP, M-LMP, CS-LMP, CS-LBP, SIFT]$ e $D_2 = [Z+TZLBP, M-LMP, CS-LMP, CS-LBP, SIFT]$, e seis diferentes vetores combinados pelo operador de fusão Soma aplicado ao conjunto de descritores $D_3 = [Z+TZLBP, M-LMP, CS-LMP, CS-LBP]$, tomados dois a dois.

Por simplicidade, quando da apresentação dos resultados, a Concatenação entre os vetores de características para os descritores $Desc1$ e $Desc2$ é representada por $Desc1 \oplus Desc2$, e a Soma por $Desc1 + Desc2$.

5.2.7.2 Fusão de Descritores de Textura na Camada de Decisão

Para realizar a fusão dos descritores na camada de decisão, foram aplicados quatro diferentes operadores de fusão, Mínimo, Adição, Produto e Combinação, sobre o conjunto de distância euclidiana dos pares de regiões possíveis correspondentes para os diferentes descritores, tomados dois a dois, sendo o operador Mínimo dado pela equação (5.8).

$$\min_{Img_{i,j_}(Desc1, Desc2)} = \min(d_{Img_{i,j_}Desc1}, d_{Img_{i,j_}Desc2}), \quad (5.8)$$

em que $\min_{Img_{i,j_}(Desc1, Desc2)}$ representa a mínima distância euclidiana entre os vetores de características da i -ésima região da pose 1 da imagem Img e da j -ésima região da pose 2 da respectiva imagem, dentre as distâncias retornadas pelos descritores $Desc1$ e $Desc2$.

A Adição das distâncias euclidianas entre os vetores de características para os descritores $Desc1$ e $Desc2$, representada por $add_{Img_{i,j_}(Desc1, Desc2)}$, segue a formulação apresentada na Equação (5.9).

$$add_{Img_{i,j_}(Desc1, Desc2)} = d_{Img_{i,j_}Desc1} + d_{Img_{i,j_}Desc2} \quad (5.9)$$

O operador Produto é aplicado tal como apresentado pela Equação (5.10),

$$prod_{Img_{i,j_}(Desc1, Desc2)} = d_{Img_{i,j_}Desc1} \cdot d_{Img_{i,j_}Desc2}, \quad (5.10)$$

sendo $prod_{Img_{i,j_}(Desc1, Desc2)}$ o produto das distâncias euclidianas entre os vetores de características da i -ésima região da pose 1 da imagem Img e da j -ésima região da pose 2 da respectiva imagem para os descritores $Desc1$ e $Desc2$.

A combinação das distâncias euclidianas entre os vetores de características da i -ésima região da pose 1 da imagem Img e da j -ésima região da pose 2 da respectiva

imagem para os descritores $Desc1$ e $Desc2$, representada por $comb_{Img_{i,j}}(Desc1, Desc2)$, é dada por (5.11).

$$comb_{Img_{i,j}}(Desc1, Desc2) = d_{Img_{i,j}}_{Desc1} + d_{Img_{i,j}}_{Desc2} - (d_{Img_{i,j}}_{Desc1} \cdot d_{Img_{i,j}}_{Desc2}) \quad (5.11)$$

Os quatro operadores de fusão de decisão são aplicados sobre as distâncias euclidianas entre os vetores de características determinados pelos conjuntos de descritores D_1 e D_2 , apresentados na seção anterior, gerando quatorze diferentes distâncias combinadas para cada operador de fusão de decisão.

Por simplicidade, quando da apresentação dos resultados, os operadores de fusão de decisão Mínimo, Adição, Produto e Combinação, serão representados por $min(Desc1, Desc2)$, $add(Desc1, Desc2)$, $prod(Desc1, Desc2)$ e $comb(Desc1, Desc2)$, respectivamente.

As setenta e seis abordagens de fusão, vinte da camada de características e cinquenta e seis da camada de decisão (quatorze para cada operador de fusão de decisão), são aplicadas sobre todas as imagens do conjunto de treinamento, sendo aplicadas sobre as imagens do conjunto de teste apenas as abordagens que obtiverem os melhores resultados sobre o conjunto de treinamento. Os resultados são apresentados no formato de área sobre a curva *recall x precision* (AUC).

5.3 Resultados

Como descrito nas seções anteriores, para realizar o registro de imagens, primeiro detectou-se as regiões de interesse de cada par de imagens. Para tal, aplicou-se o detector de regiões *Hessian-Affine*.

Tabela 1: Número de regiões de interesse detectadas pelo detector *Hessian-Affine* e de regiões correspondentes e não correspondentes, previamente conhecidas, por par de imagens do conjunto de treinamento.

Imagens	Regiões Detectadas		Número de Regiões	
	Pose 1	Pose 2	Correspondentes	Não correspondentes
bark	493	510	225	251.205
bike	1025	765	2.402	781.723
boat	3146	1433	3.529	4.504.689
graf	2454	2434	3.480	5.969.556
leuven	1125	664	1.818	745.182
trees	3685	4217	7.668	15.531.977
ubc	1570	1615	4.713	2.530.837
wall	1375	1449	1.353	1.991.022

O número de regiões de interesse detectadas em cada imagem e o de regiões correspondentes e não correspondentes, previamente conhecido, é apresentado na Tabela 1, para o conjunto de treinamento, e na Tabela 2, para o conjunto de teste. Nota-se que,

Tabela 2: Número de regiões de interesse detectadas pelo detector *Hessian-Affine* e de regiões correspondentes e não correspondentes, previamente conhecidas, por imagem do conjunto de teste.

Imagens	Regiões Detectadas		Número de Regiões	
	Pose 1	Pose 2	Correspondentes	Não correspondentes
asterix	339	545	1.202	183.553
belledone	833	221	50	184.043
bip	940	1810	2.143	1.669.257
crolle	385	145	177	55.648
east park	1680	1578	2.957	2.648.083
east south	2608	1611	3.321	4.198.167
ensigma	1377	1037	2.062	1.425.887
inria	1006	621	459	624.267
lap	782	605	1.823	471.287
laptop	742	613	1.286	453.560
new york	888	843	2.039	746.545
resid	1319	1171	2.685	1.541.864
vangogh	448	488	1.152	217.472
zmars	2335	2305	6.715	5.375.460
zmonet	538	586	1.844	313.424

para determinadas imagens, a aplicação do detector de regiões resultou em um número de regiões detectadas na pose 1 muito próximo ao obtido na pose 2, como é o caso das imagens *bark*, *graf*, *ubc* e *wall*, para o conjunto de treinamento, e *east park*, *laptop*, *new york*, *vangogh*, *zmars* e *zmonet*, para o conjunto de teste. Para as outras imagens, há discrepância significativa entre os respectivos valores.

Cada região da pose 1 é comparada com todas as regiões da pose 2 a fim de se determinar o par correspondente, sendo que, caso exista correspondência, apenas uma região da pose 2 é correspondente à região da pose 1. Nota-se uma alta densidade de regiões não correspondentes, sendo o total de regiões não correspondentes da ordem de mil vezes maior que o total de regiões correspondentes para os dois conjuntos de imagens.

5.3.1 Resultados para o Conjunto de Treinamento

Os valores de AUC para os descritores individuais aplicados sobre o conjunto de treinamento são apresentados na Tabela 3, em que os melhores resultados encontram-se destacados. Para as imagens *bark*, *boat* e *graf*, os melhores resultados foram obtidos com a aplicação do descritor M-LMP. O descritor CS-LMP resultou em melhor AUC para a imagem *ubc*. Para as quatro outras imagens (*bike*, *leuven*, *trees* e *wall*), os melhores resultados foram obtidos com a aplicação do descritor CS-LBP. O descritor $Z \oplus TZLBP$ superou o desempenho do $Z + TZLBP$ e do SIFT para todas as imagens do conjunto de treinamento, no entanto apresentou menor AUC que os descritores M-LMP, CS-LMP e CS-LBP. Nota-se ainda uma queda de desempenho expressiva para todos os descritores, em especial para M-LMP, CS-LMP e CS-LBP, quando aplicados às imagens *boat*, *graf* e *trees*; para estas imagens, o percentual retornado de falsas correspondências supera significativamente o de correspondências corretas. Os descritores M-LMP, CS-LMP e CS-LBP, apresentam desempenho médio semelhante uns aos outros e superam em 20% o

obtido pelo descritor SIFT, que apresentou o menor desempenho. O melhor desempenho médio, $\overline{AUC}$ igual a 65,5%, foi obtido com a aplicação do descritor CS-LBP.

Tabela 3: AUC (%) para os descritores individuais aplicados sobre o conjunto de treinamento.

Descritor	B	bark	bike	boat	graf	leuven	trees	ubc	wall	$\overline{AUC}$
M-LMP	256	83,1	77,6	56,4	46,8	70,5	32,7	77,3	74,8	64,9
CS-LMP	256	82,9	79,4	55,7	43,5	74,5	33,4	77,5	75,7	65,3
CS-LBP	256	82,1	80,0	55,5	43,8	75,5	33,9	77,4	76,0	65,5
SIFT	128	64,3	39,6	44,0	24,7	49,4	16,4	60,0	55,5	44,2
Z+TZLBP	256	70,7	69,9	41,4	35,2	69,5	25,7	70,5	64,5	55,9
Z $\oplus$ TZLBP	512	75,2	73,9	49,4	39,0	73,3	30,8	74,6	72,4	61,1

As células destacadas em cinza escuro representam o melhor resultado obtido.

Tabela 4: AUC (%) para as nove melhores abordagens de fusão aplicadas sobre o conjunto de treinamento.

Abordagem de Fusão	bark	bike	boat	graf	leuven	trees	ubc	wall	$\overline{AUC}$
M-LMP $\oplus$ CS-LBP	84,2	81,2	58,6	50,7	75,9	36,1	78,9	77,4	67,9
M-LMP $\oplus$ (Z $\oplus$ TZLBP)	84,5	80,8	58,5	51,9	76,5	36,6	79,0	78,9	68,3
CS-LMP $\oplus$ (Z $\oplus$ TZLBP)	83,7	80,2	57,7	46,9	76,7	36,1	78,8	78,8	67,4
add(M-LMP,Z $\oplus$ TZLBP)	83,9	80,1	58,0	50,7	76,3	36,8	79,0	78,8	67,9
prod(M-LMP,CS-LBP)	83,8	81,4	58,5	49,7	76,3	36,2	78,7	77,5	67,8
prod(M-LMP,Z $\oplus$ TZLBP)	83,7	80,1	57,9	50,5	76,7	36,7	79,0	78,8	67,9
comb(M-LMP,CS-LBP)	84,3	81,3	58,5	50,0	76,1	36,2	78,9	77,7	67,9
comb(M-LMP,Z $\oplus$ TZLBP)	84,0	80,6	58,1	51,1	76,6	36,8	79,1	78,7	68,1
comb(CS-LMP,Z $\oplus$ TZLBP)	83,6	80,2	57,2	47,4	77,0	36,4	78,8	78,7	67,4

A célula destacada em cinza escuro representa o melhor resultado obtido.

Com a aplicação dos operadores de fusão sobre os descritores de textura (camada de características) e sobre as distâncias euclidianas (camada de decisão), geraram-se setenta e seis diferentes combinações. Destas, nove superaram o desempenho dos descritores individuais em todas as imagens do conjunto de treinamento. Apresentam-se, na Tabela 4, os valores de AUC para as nove melhores combinações, sendo três delas obtidas pela aplicação do operador Concatenação sobre os descritores de características, M-LMP $\oplus$ CS-LBP, M-LMP $\oplus$ (Z $\oplus$ TZLBP) e CS-LMP $\oplus$ (Z $\oplus$ TZLBP), e as seis outras melhores, obtidas com a aplicação dos operadores de fusão de decisão Adição, add(M-LMP,Z $\oplus$ TZLBP), Produto, prod(M-LMP,CS-LBP) e prod(M-LMP,Z $\oplus$ TZLBP), e Combinação, comb(M-LMP,CS-LBP), comb(M-LMP,Z $\oplus$ TZLBP) e comb(CS-LMP,Z $\oplus$ TZLBP). O melhor desempenho dentre as combinações de vetores de características foi obtido com a abordagem M-LMP $\oplus$ (Z $\oplus$ TZLBP), cujo valor médio de AUC, igual a 68,3%, é semelhante ao melhor desempenho dentre as abordagens de fusão por decisão, obtido por comb(M-LMP,Z $\oplus$ TZLBP).

Apresenta-se, na Tabela 5, o ganho obtido com a aplicação da abordagem de fusão M-LMP $\oplus$ (Z $\oplus$ TZLBP) quando comparada ao desempenho dos descritores individuais que a compõem, M-LMP e Z $\oplus$ TZLBP, e ao melhor desempenho individual para cada imagem do conjunto de treinamento. Para a imagem *bark*, a abordagem de fusão apresentou AUC igual

Tabela 5: Ganho (%) com a aplicação da abordagem de fusão $M-LMP \oplus (Z \oplus TZLBP)$ sobre o conjunto de treinamento.

Abordagem	bark	bike	boat	graf	leuven	trees	ubc	wall	$\overline{AUC}$
M-LMP $\oplus$ (Z $\oplus$ TZLBP)	84,5	80,8	58,5	51,9	76,5	36,6	79,0	78,9	68,3
M-LMP $\oplus$ CS-LBP	$\uparrow 0,3$	$\downarrow 0,4$	$\downarrow 0,1$	$\uparrow 1,2$	$\uparrow 0,6$	$\uparrow 0,5$	$\uparrow 0,1$	$\uparrow 1,5$	$\uparrow 0,4$
M-LMP	$\uparrow 1,4$	$\uparrow 3,2$	$\uparrow 2,1$	$\uparrow 5,1$	$\uparrow 6,0$	$\uparrow 3,9$	$\uparrow 1,7$	$\uparrow 4,1$	$\uparrow 3,4$
Z $\oplus$ TZLBP	$\uparrow 9,3$	$\uparrow 6,9$	$\uparrow 9,1$	$\uparrow 12,9$	$\uparrow 3,2$	$\uparrow 5,8$	$\uparrow 4,4$	$\uparrow 6,5$	$\uparrow 7,2$
Melhor resultado individual	$\uparrow 1,4$	$\uparrow 0,8$	$\uparrow 2,1$	$\uparrow 5,1$	$\uparrow 1,0$	$\uparrow 2,7$	$\uparrow 1,5$	$\uparrow 2,9$	$\uparrow 2,8$

a 84,5%, representando um acréscimo de 1,4% e de 9,3% em relação à aplicação individual dos descritores M-LMP e Z $\oplus$ TZLBP, respectivamente. O maior ganho se deu com a aplicação da abordagem de fusão sobre a imagem *graf*, cujo desempenho representa um acréscimo de 5,1% e de 12,9% sobre o desempenho dos respectivos descritores individuais. Ao comparar o desempenho médio, a abordagem M-LMP $\oplus$ (Z $\oplus$ TZLBP) apresentou um acréscimo de 3,4% em relação ao M-LMP, de 7,2% em relação ao Z $\oplus$ TZLBP e de 2,8% em relação ao melhor resultado individual. Quando comparada à abordagem M-LMP $\oplus$ CS-LBP, a abordagem de fusão M-LMP $\oplus$ (Z $\oplus$ TZLBP) apresenta AUC superior em seis das oito imagens do conjunto de treinamento, com acréscimo de 0,1% na imagem *ubc* a 1,5% na imagem *wall*, no entanto, apresenta AUC menor nas imagens *bike* e *boat*, com decréscimo de 0,4% e 0,1%, respectivamente. O desempenho médio da abordagem M-LMP $\oplus$ (Z $\oplus$ TZLBP) apresenta um acréscimo de 0,4% em relação à $\overline{AUC}$ da abordagem M-LMP $\oplus$ CS-LBP.

O descritor Z $\oplus$ TZLBP, quando aplicado individualmente, apresentou desempenho inferior ao M-LMP, CS-LMP e CS-LBP, no entanto, quando combinado ao descritor M-LMP, seja por Concatenação, seja pelo operador de fusão de decisão Combinação, resultou no melhor desempenho obtido dentre as demais abordagens de fusão. Acredita-se, assim, que o descritor Z $\oplus$ TZLBP apresenta informação complementar à do descritor M-LMP. Como discutido anteriormente, este é o objetivo de se utilizar fusão da informação: combinar as informações complementares provenientes de diferentes fontes, no caso em estudo, de diferentes descritores de textura.

5.3.2 Validação dos Resultados: Aplicação Sobre o Conjunto de Teste

Aplicou-se sobre o conjunto de teste somente os descritores individuais que resultaram em bom desempenho sobre o conjunto de treinamento. Para a escolha, definiu-se limiar igual a 60%. Os descritores M-LMP, CS-LMP, CS-LBP e Z $\oplus$ TZLBP apresentaram $\overline{AUC} > 60,0\%$. O desempenho destes quatro descritores individuais aplicados sobre o conjunto de teste é discriminado na Tabela 6, em que os melhores resultados são mostrados em destaque.

O descritor CS-LMP apresentou o melhor resultado para sete das quinze imagens do conjunto de teste, assim como o maior valor médio de AUC, $\overline{AUC} = 73,3\%$. A aplicação individual do descritor M-LMP, assim como do CS-LBP, gerou a melhor AUC para três

Tabela 6: AUC (%) para os melhores descritores individuais aplicados sobre o conjunto de teste.

Imagem	M-LMP	CS-LMP	CS-LBP	$Z \oplus \text{TZLBP}$
asterix	70,9	74,6	72,8	71,6
belledone	69,7	73,8	75,3	61,3
bip	71,3	72,5	72,0	68,2
crolle	86,1	87,6	86,7	83,3
east park	56,1	57,4	58,5	59,7
east south	64,6	66,0	65,5	59,9
ensigma	45,2	45,9	47,9	48,8
inria	67,8	66,8	61,9	55,7
lap	80,4	83,8	82,1	78,4
laptop	77,3	79,4	77,5	72,1
new york	86,6	86,3	85,8	83,6
resid	66,4	70,8	71,4	67,4
vangogh	77,3	80,0	80,4	76,8
zmars	78,4	78,3	77,9	75,7
zmonet	76,8	77,0	76,4	70,5
$\overline{AUC}$	71,7	73,3	72,8	68,9

As células destacadas em cinza escuro representam o melhor resultado obtido.

imagens do conjunto de teste, e o desempenho do descritor $Z \oplus \text{TZLBP}$ superou o dos demais descritores em duas das imagens do conjunto de teste. Tal como observado no conjunto de treinamento, todos os descritores individuais apresentaram queda expressiva de desempenho quando aplicados a duas imagens do conjunto de teste, sendo elas *east park* e *ensigma*.

Para validar os resultados obtidos no conjunto de treinamento, aplicaram-se, sobre as imagens do conjunto de teste, as nove melhores abordagens de fusão apresentadas na seção anterior. Os resultados obtidos para os operadores de fusão aplicados sobre os vetores de características, $\text{M-LMP} \oplus \text{CS-LBP}$, $\text{M-LMP} \oplus (Z \oplus \text{TZLBP})$ e $\text{CS-LMP} \oplus (Z \oplus \text{TZLBP})$, considerando métrica de desempenho AUC, são discriminados na Tabela 7, e os obtidos para os seis operadores de fusão de decisão, $\text{add}(\text{M-LMP}, Z \oplus \text{TZLBP})$, $\text{prod}(\text{M-LMP}, \text{CS-LBP})$, $\text{prod}(\text{M-LMP}, Z \oplus \text{TZLBP})$, $\text{comb}(\text{M-LMP}, \text{CS-LBP})$, $\text{comb}(\text{M-LMP}, Z \oplus \text{TZLBP})$ e $\text{comb}(\text{CS-LMP}, Z \oplus \text{TZLBP})$, são apresentados na Tabela 8.

Nota-se que as combinações $\text{M-LMP} \oplus (Z \oplus \text{TZLBP})$ e $\text{comb}(\text{M-LMP}, Z \oplus \text{TZLBP})$ superaram os desempenhos dos melhores descritores individuais para doze das quinze imagens do conjunto de teste (80% das imagens), enquanto as demais combinações apresentaram resultado melhor que o melhor dentre os descritores individuais em ao menos dez das quinze imagens do conjunto de teste (66.7% das imagens). Ao se considerar o valor médio de AUC, todas as abordagens de fusão superaram o melhor resultado dentre os obtidos pelos descritores individuais, sendo que a abordagem $\text{M-LMP} \oplus \text{CS-LBP}$ apresentou o melhor resultado, $\overline{AUC} = 74,6\%$, representando um acréscimo de 1,3% ao maior valor obtido dentre os descritores individuais, o qual foi atingido com a aplicação do descritor

Tabela 7: AUC (%) para as três melhores abordagens de fusão da camada de características aplicadas sobre as imagens do conjunto de teste.

Imagem	M-LMP $\oplus$ CS-LBP	M-LMP $\oplus$ (Z $\oplus$ TZLBP)	CS-LMP $\oplus$ (Z $\oplus$ TZLBP)
asterix	74,4	73,4	74,7
belledone	74,4	72,9	73,6
bip	73,6	73,2	72,8
crolle	88,3	88,3	88,2
east park	61,1	63,5	62,0
east south	67,2	66,5	66,6
ensigma	48,5	50,1	49,4
inria	69,0	68,3	65,0
lap	84,2	84,8	84,6
laptop	80,7	81,4	80,7
new york	87,2	87,2	86,6
resid	72,3	72,0	72,5
vangogh	80,1	79,4	79,9
zmars	79,3	79,4	78,9
zmonet	78,4	77,6	77,3
$\overline{AUC}$	74,6	74,5	74,2

As células destacadas em cinza claro representam os casos para os quais a abordagem de fusão supera o melhor desempenho dentre os descritores individuais.

A célula destacada em cinza escuro representa o melhor resultado $\overline{AUC}$ obtido dentre as abordagens de fusão.

Tabela 8: AUC (%) para as seis melhores abordagens de fusão de decisão aplicadas sobre as imagens do conjunto de teste.

Imagem	add(n1,n4)*	prod(n1,n3)*	prod(n1,n4)*	comb(n1,n3)*	comb(n1,n4)*	comb(n2,n4)*
asterix	73,1	73,4	73,7	73,5	73,5	74,7
belledone	73,0	74,8	73,0	74,7	73,1	73,8
bip	72,9	73,5	72,9	73,6	73,1	72,9
crolle	87,9	88,4	87,7	88,4	87,9	88,3
east park	62,6	60,5	62,7	60,7	62,7	62,2
east south	66,2	66,9	66,2	66,9	66,2	66,4
ensigma	50,0	48,4	49,8	48,3	49,6	49,3
inria	68,2	68,6	67,8	69,3	69,0	65,6
lap	83,9	83,8	84,1	83,9	84,3	84,5
laptop	80,5	80,4	80,8	80,4	81,0	80,4
new york	86,9	87,3	86,9	87,3	87,0	86,4
resid	71,7	72,4	71,9	72,5	71,9	72,4
vangogh	78,8	80,2	79,2	80,4	79,3	79,9
zmars	79,2	79,2	79,1	79,3	79,3	78,9
zmonet	76,8	78,1	76,6	78,3	77,3	77,0
$\overline{AUC}$	74,1	73,8	74,2	74,5	74,3	74,2

* n1 = M-LMP, n2 = CS-LMP, n3 = CS-LBP, n4 = Z $\oplus$ TZLBP.

As células destacadas em cinza claro representam os casos para os quais a abordagem de fusão supera o melhor desempenho dentre os descritores individuais. A célula destacada em cinza escuro representa o melhor resultado $\overline{AUC}$ obtido dentre as abordagens de fusão.

CS-LMP.

O ganho obtido com a aplicação da abordagem de fusão M-LMP $\oplus$ CS-LBP quando comparada ao desempenho dos descritores individuais que a compõem, M-LMP e CS-LBP,

Tabela 9: Ganho (%) com a aplicação da abordagem de fusão M-LMP $\oplus$ CS-LBP sobre o conjunto de teste.

Imagem	M-LMP $\oplus$ CS-LBP	M-LMP $\oplus$ (Z $\oplus$ TZLBP)	M-LMP	CS-LBP	>DESC _{ind}
asterix	74,4	$\uparrow 1,0$	$\uparrow 3,5$	$\uparrow 1,6$	$\downarrow 0,2$
belledone	74,4	$\uparrow 1,5$	$\uparrow 4,7$	$\downarrow 0,9$	$\downarrow 0,9$
bip	73,6	$\uparrow 0,4$	$\uparrow 2,3$	$\uparrow 1,6$	$\uparrow 1,1$
crolle	88,3	0	$\uparrow 2,2$	$\uparrow 1,6$	$\uparrow 0,7$
east park	61,1	$\downarrow 2,4$	$\uparrow 5,0$	$\uparrow 2,6$	$\uparrow 1,4$
east south	67,2	$\uparrow 0,7$	$\uparrow 2,6$	$\uparrow 1,7$	$\uparrow 1,2$
ensigma	48,5	$\downarrow 1,6$	$\uparrow 3,3$	$\uparrow 0,6$	$\downarrow 0,3$
inria	69,0	$\uparrow 0,7$	$\uparrow 1,2$	$\uparrow 7,1$	$\uparrow 1,2$
lap	84,2	$\downarrow 0,6$	$\uparrow 3,8$	$\uparrow 2,1$	$\uparrow 0,4$
laptop	80,7	$\downarrow 0,7$	$\uparrow 3,4$	$\uparrow 3,2$	$\uparrow 1,3$
new york	87,2	0	$\uparrow 0,6$	$\uparrow 1,4$	$\uparrow 0,6$
resid	72,3	$\uparrow 0,3$	$\uparrow 5,9$	$\uparrow 0,9$	$\uparrow 0,9$
vangogh	80,1	$\uparrow 0,7$	$\uparrow 2,8$	$\downarrow 0,3$	$\downarrow 0,3$
zmars	79,3	$\downarrow 0,1$	$\uparrow 0,9$	$\uparrow 1,4$	$\uparrow 0,9$
zmonet	78,4	$\uparrow 0,8$	$\uparrow 1,6$	$\uparrow 2,0$	$\uparrow 1,4$
$\overline{AUC}$	74,6	$\uparrow 0,1$	$\uparrow 2,9$	$\uparrow 1,8$	$\uparrow 1,3$

e ao melhor desempenho individual ($>DESC_{ind}$) para cada imagem do conjunto de teste encontra-se detalhado na Tabela 9. Nota-se que a abordagem M-LMP $\oplus$ CS-LBP supera o descritor individual M-LMP em todas as imagens, com acréscimo de 0,6% na imagem *new york* a 5,9% na imagem *resid*. Ademais, apresenta AUC superior à obtida com a aplicação individual do descritor CS-LBP em treze das quinze imagens do conjunto de teste, com acréscimo de 0,6% na imagem *ensigma* a 7,1% na imagem *inria*, e AUC inferior nas imagens *belledone* e *vangogh*, com decréscimo de 0,9% e 0,3%, respectivamente. Ao comparar o desempenho médio, a abordagem M-LMP $\oplus$ CS-LBP apresenta um acréscimo de 2,9% em relação ao M-LMP, de 1,8% em relação ao CS-LBP e de 1,3% em relação ao melhor resultado individual. Quando comparada à abordagem M-LMP $\oplus$ (Z $\oplus$ TZLBP), a abordagem de fusão M-LMP $\oplus$ CS-LBP apresenta AUC superior em oito das quinze imagens do conjunto de teste, com acréscimo de 0,3% na imagem *resid* a 1,5% na imagem *belledone*, no entanto, apresenta AUC menor em cinco imagens, com decréscimo de 0,1% na imagem *zmars* a 2,4% na imagem *east park*. As duas abordagens de fusão apresentaram o mesmo valor de AUC quando aplicadas às imagens *crolle* e *new york*. O desempenho médio da abordagem M-LMP $\oplus$ CS-LBP apresenta um acréscimo de 0,1% em relação à $\overline{AUC}$ da abordagem M-LMP $\oplus$ (Z $\oplus$ TZLBP).

5.3.3 Análise do Custo Computacional

Apresenta-se, na Tabela 10, o tempo médio de processamento para os descritores individuais M-LMP, CS-LMP, CS-LBP e Z $\oplus$ TZLBP, assim como para as nove melhores abordagens de fusão, em que t_{feat} refere-se ao tempo de geração do vetor de características e t_{dist} é o tempo de determinação da distância euclidiana entre os pares de vetores de características para os descritores em questão e para as respectivas abordagens de fusão. Enquanto t_{feat} é função da dimensão da imagem de textura e da complexidade da regra de extração de características do descritor, t_{dist} é diretamente proporcional a B , sendo B

o comprimento do vetor de características.

Tabela 10: Tempo médio de processamento.

Descritor	B	t_{feat} (ms)	t_{dist} (μ s)
M-LMP	256	73,8	0,78
CS-LMP	256	73,8	0,78
CS-LBP	256	64,6	0,77
$Z \oplus$ TZLBP	512	66,8	1,20
M-LMP $\oplus$ CS-LBP	512	138,4	1,14
M-LMP $\oplus$ ($Z \oplus$ TZLBP)	768	140,6	1,66
CS-LMP $\oplus$ ($Z \oplus$ TZLBP)	768	140,6	1,64
add(M-LMP, $Z \oplus$ TZLBP)	-	140,6	1,98
prod(M-LMP, CS-LBP)	-	138,4	1,55
prod(M-LMP, $Z \oplus$ TZLBP)	-	140,6	1,98
comb(M-LMP, CS-LBP)	-	138,4	1,55
comb(M-LMP, $Z \oplus$ TZLBP)	-	140,6	1,98
comb(CS-LMP, $Z \oplus$ TZLBP)	-	140,6	1,98

Nota-se que a geração do vetor de características representa um custo computacional significativamente mais expressivo que a determinação da distância euclidiana entre dois vetores de características, sendo t_{feat} medido em milissegundos e t_{dist} , em microsegundos. Para as diferentes abordagens de fusão, t_{feat} foi determinado pela soma dos respectivos valores de tempo de geração dos descritores individuais sobre os quais o operador de fusão foi aplicado. Do mesmo modo, t_{dist} , para as abordagens de fusão de decisão, foi determinado pela soma dos respectivos tempos para os descritores individuais sobre os quais o operador de fusão de decisão foi aplicado. Em uma visão geral, o tempo médio de processamento para as abordagens de fusão é aproximadamente duas vezes maior que o respectivo tempo para os descritores individuais.

5.4 Considerações Finais

Com o objetivo de melhorar o desempenho de sistemas de registro de imagens, aplicaram-se seis diferentes operadores de fusão, dois na camada de características (Soma e Concatenação), e quatro na camada de decisão (Mínimo, Adição, Produto e Combinação), sobre seis diferentes descritores de textura, $Z \oplus$ TZLBP, $Z +$ TZLBP, M-LMP, CS-LMP, CS-LBP e SIFT. Utilizaram-se dois conjuntos de imagens com diferentes tipos de transformações, sendo o conjunto de treinamento composto por oito pares de imagens e o de teste, por 15 pares de imagens.

A aplicação dos descritores individuais sobre o conjunto de treinamento corroborou com a afirmação de que não há um descritor que apresente o melhor resultado para todo tipo de imagem, uma vez que o descritor M-LMP apresentou o melhor desempenho para três das oito imagens do conjunto de treinamento, o descritor CS-LMP, para uma imagem,

e o descritor CS-LBP superou o desempenho dos demais para as outras quatro imagens do conjunto de treinamento. Nove das setenta e seis abordagens de fusão superaram o melhor desempenho dentre os obtidos pelos descritores individuais para todas as imagens do conjunto de treinamento.

Comportamento semelhante foi observado para o conjunto de teste. Não houve um descritor individual que apresentasse o melhor desempenho para todas as imagens do conjunto, sendo que duas das nove melhores abordagens de fusão superaram o maior desempenho dos descritores individuais para doze das quinze imagens do conjunto de treinamento, e as demais sete abordagens de fusão superaram o maior desempenho dos descritores individuais em ao menos dez imagens do conjunto de teste; ademais, ao se considerar o valor médio de AUC, todas as abordagens de fusão superaram o melhor resultado dentre os obtidos pelos descritores individuais.

Esses resultados são promissores e demonstram que a fusão de descritores de textura contribui para o aumento do desempenho de sistemas de registro de imagens quando comparado à aplicação de descritores de textura individuais. Com o intuito de expandir tal análise a outras aplicações, apresentam-se, no capítulo a seguir, a metodologia e os resultados referentes à fusão de descritores de textura aplicada a reconhecimento de distorção arquitetural em mamografia digital.

6 FUSÃO DE DESCRITORES DE TEXTURA APLICADA A RECONHECIMENTO DE DISTORÇÃO ARQUITETURAL EM MAMOGRAFIA DIGITAL

Apresentou-se, no capítulo anterior, a fusão de descritores de textura aplicada a registro de imagens, sendo aplicados operadores de fusão nas camadas de características e de decisão. Os resultados obtidos foram promissores, pois nove abordagens de fusão possibilitaram melhor desempenho global que os descritores individuais. Com o objetivo de continuar a investigação acerca de fusão de descritores de textura, apresentam-se, neste capítulo, a metodologia e os resultados obtidos com a aplicação de fusão a reconhecimento de distorção arquitetural em mamografia digital, sendo implementados os descritores de textura baseados em fusão de características, CMLMP, CMedianLMP e CZTZLMP, e os operadores Soma^M e Gram^M, propostos neste trabalho de doutorado, cujo desempenho é comparado a outros métodos publicados na literatura.

6.1 Considerações Iniciais

Distorção arquitetural mamária (DA) é o indicador de possível formação de câncer de mama mais prematuro que pode ser detectado em imagens de mamografia e sua ocorrência está normalmente associada à formação de tumores malignos (BAHL et al., 2015). Outros sinais em mamografias indicadores de câncer de mama são microcalcificações e nódulos. Enquanto estes indicam a possível presença de um tumor em determinado estágio de evolução, a distorção arquitetural indica que, no local de sua ocorrência, um tumor pode estar começando a ser formado (BAHL et al., 2015; GAUR et al., 2013).

De acordo com a Organização Mundial de Saúde (OMS), o câncer de mama é uma das principais doenças que causam a morte prematura de mulheres (WHO, 2019). No entanto, a detecção do tumor em estágio inicial de formação pode aumentar em até 30% a taxa de sucesso do tratamento, que leva à cura da doença, além de permitir outras opções de tratamento, com menor efeito colateral aos pacientes. Desse modo, a detecção de distorção arquitetural evidencia-se como uma tarefa essencial para o tratamento do câncer de mama (VERONESI et al., 2005; ELMORE et al., 2003), sendo a mamografia digital o exame de triagem e/ou rastreamento mais comum utilizado em todo o mundo para a detecção precoce desta doença (GLYNN et al., 2011; KARELLAS; VEDANTHAM, 2008).

Entretanto, pelo fato da distorção arquitetural mamária ser caracterizada por uma contração sutil no tecido mamário, sem borda definida e sem formação de massa (SULEIMAN et al., 2016; D'ORSI, 2013), detectá-la em exames de mamografia é uma tarefa de difícil realização, altamente associada à experiência dos radiologistas (BAHL et al., 2015). Não obstante, são comuns casos em que os radiologistas não conseguem identificar

a distorção arquitetural mamária (JASIONOWSKA; PRZELASKOWSKI, 2010). Em estudos de análise retrospectiva de exames de mamografia, para casos falsos negativos, constatou-se uma significativa incidência de distorção arquitetural mamária (RAY et al., 2015; GAUR et al., 2013).

Para auxiliar na detecção de DA, sistemas computacionais de auxílio ao diagnóstico (CAD) são utilizados na prática médica, servindo como uma segunda opinião para os radiologistas (YANG et al., 2007; DEAN; ILVENTO, 2006). No entanto, tais sistemas apresentam baixo desempenho, principalmente se comparados ao desempenho de detecção de microcalcificações e nódulos. Diversas técnicas de reconhecimento de padrões e aprendizagem de máquina vêm sendo investigadas para melhorar o desempenho de sistemas CAD aplicados a reconhecimento de DA em mamografias, sendo descritores de textura comumente utilizados (OLIVEIRA et al., 2018; RANGAYYAN; BANIK; DESAUTELS, 2010).

Mesmo sendo a mamografia digital de campo total (FFDM) o exame de triagem mais utilizado em todo o mundo para auxiliar no diagnóstico de doenças da mama, a maioria dos trabalhos publicados na literatura aborda detecção de DA em mamografias digitalizadas. Isso se dá principalmente pela dificuldade em se obter bancos de mamografias digitais com distorção arquitetural detectada por radiologistas e documentadas com laudo médico, e por haver dois bancos de mamografias digitalizadas disponíveis e de fácil acesso na internet, sendo eles DDSM (HEATH et al., 2000) e mini-MIAS (SUCKLING et al., 1994), os quais são vastamente utilizados por diversos pesquisadores, como mostram os dados apresentados na Tabela 11.

Tabela 11: Diferentes abordagens para detecção de distorção arquitetural em mamografia

Abordagem	Banco de Dados	#Imagens com DA	#Imagens sem DA	Desempenho*
Nemoto et al. (2009)	FCR - privado	25	-	S = 80,0%
Jasionowska et al. (2010)	DDSM	33	258**	S = 68,0%
Minavathi et al. (2011)	mini-MIAS	19	152	S = 94,7%, E = 89,6%
	DDSM	23	97	
Rangayyan et al. (2013)	digitalizado - privado	106	52	S = 80,0%, AUC = 75,0%
Mohammadi et al. (2013)	DDSM	150**	150**	Acc = 91,2%
	DDSM	60**	25	Acc = 92,9%, S = 93,3%
Kamra et al. (2016)	mini-MIAS	13	36	Acc = 95,3%, S = 85,7%
	digitalizado - privado	23**	50**	Acc = 88,0%, S = 87,5%
Casti et al. (2016)	DDSM	122	785	S = 81,0%
Oliveira et al. (2017)	FFDM - privado	100	100**	Acc = 90,0%, AUC = 97,3%
	DDSM	63	-	S = 90,0%
Liu et al. (2018)	mini-MIAS	19	-	
Oliveira et al. (2019)	FFDM - privado	201	201	Acc = 75,3%, AUC = 82,9%

*Acurácia (Acc), Sensibilidade (S), Especificidade (E) e Área sobre a curva ROC (AUC)

**Valores que se referem ao número de ROIs, e não ao número de mamografias

Nemoto et al. (2009) propuseram uma metodologia para detecção de DA, para a qual extraíram nove características das regiões suspeitas, das quais seis foram extraídas da imagem de intensidade de nível de cinza (sendo quatro medidas de contraste, energia e *skewness*), duas calculadas sobre a imagem filtrada por filtro de Laplace (*skewness*

e momento de segunda ordem), e a nona característica refere-se à distância entre a região suspeita e a pele. As diferentes características foram concatenadas em um único vetor. Aplicada a 25 mamografias computadorizadas, a metodologia proposta resultou em sensibilidade S igual a 80,0%.

Jasionowska e Przelaskowski (2010) extraíram treze características das regiões supeitas, das quais cinco referem-se ao descritor GLCM (constraste, correlação, energia, entropia e homogeneidade), outras seis, a medidas estatísticas (variância, desvio padrão, *kurtosis*, *skewness*, energia e entropia), e duas são calculadas usando aproximação por espectro angular. As diferentes características foram concatenadas em um único vetor. Ao utilizarem classificador SVM, obtiveram sensibilidade S igual 68,0%.

Na metodologia proposta por Minavathi, Murali e Dinesh (2011), aplicada a mamografias digitalizadas, foram utilizados dois classificadores, SVM e Perceptron Multicamadas (PMC), sendo o melhor resultado obtido por SVM com *kernel* Gaussiano. Outra abordagem que utiliza classificador PMC foi proposta por Rangayyan et al. (2013), na qual extraíram treze características referentes à espicularidade e à dispersão angular das espículas na região suspeita, obtendo sensibilidade igual a 80,0% e AUC igual a 75,0%.

Mohammadi et al. (2013) aplicaram o descritor local de textura *Monogenic Binary Coding* (MBC) para realizar o reconhecimento de distorção arquitetural em ROIs extraídas de mamografias digitalizadas, obtendo acurácia igual a 91,2%.

Kamra et al. (2016) apresentaram uma abordagem para a qual extraiu 284 características das regiões suspeitas, das quais 280 referem-se ao descritor GLCM com diferentes valores de distância entre os *pixels*, 2 são fractais e as demais características foram extraídas do espectro de Fourier. As diferentes características foram concatenadas em um único vetor. Utilizando classificador SVM, a abordagem foi aplicada a três diferentes bancos de mamografia digitalizada, sendo dois de domínio público (DDSM e mini-MIAS) e um privado, este último composto por 23 ROIs com DA e 50 ROIs sem DA, para o qual obteve-se acurácia igual a 88,0%, sensibilidade de 87,5%, especificidade igual a 88,2% e AUC de 93,0%

Casti et al. (2016) propuseram uma metodologia de detecção de anomalias em mamografia digitalizada, microcalcificações, nódulos e DA, para a qual foi extraído um conjunto de sete características (*área*, *compactness*, *elongation*, distância, direcionalidade, direcionalidade ponderada e dispersão). Utilizando o classificador *Fisher Linear Discriminant Analysis*, obtiveram sensibilidade igual 81.0% referente à detecção de DA.

Liu et al. (2018) utilizaram uma modificação do método *Twin Bounded Support Vector Machine* (TBSVM) para detectar regiões com distorção arquitetural em mamografia digitalizada. Propuseram a extração de 241 características, sendo 228 referentes ao descritor GLCM, 11 referentes a medidas de padrões espiculados e as demais referem-se a duas

medidas de entropia, Renyi e Tsallis. As diferentes características foram concatenadas em um único vetor, sendo posteriormente aplicado um processo de redução de dimensionalidade. A sensibilidade obtida foi igual a 90,0%.

Recentemente, pesquisadores do Laboratório de Visão Computacional (LAVI) da USP de São Carlos publicaram quatro trabalhos referentes ao reconhecimento de distorção arquitetural em mamografia, para os quais compararam-se os resultados da aplicação de diferentes descritores, sendo eles LBP, CS-LBP, LMP, GLCM, filtro de Gabor e uma abordagem de rede neural convolucional (CNN), *Autoencoder*, cuja classificação se deu com redes neurais PMC. No primeiro deles (OLIVEIRA et al., 2016), os autores compararam o desempenho dos descritores LBP, LMP e GLCM (contraste, correlação, energia e homogeneidade) aplicados a mamografias digitalizadas do banco de dados DDSM, sendo o melhor resultado obtido com a aplicação do LMP, cujo valor de sensibilidade foi igual a 97,0% e de acurácia, 95,0%. No segundo trabalho (OLIVEIRA et al., 2017), foi apresentado um banco de mamografia digital gerado pelos autores, composto por 100 mamografias com DA. Foram extraídas duas regiões de interesse de 256×256 pixels de cada mamografia, uma com a DA situada no centro da respectiva ROI e a segunda, de uma região saudável da mama, sem a presença de DA. Foram aplicados os mesmos descritores do trabalho anterior, LBP, LMP e GLCM (contraste, correlação, energia e homogeneidade). Os melhores resultados foram obtidos com a aplicação do descritor LMP, acurácia igual a 90,0% e AUC igual a 97,3%, e com a aplicação da medida estatística energia do descritor GLCM, para a qual obteve-se acurácia de 96,0% e AUC igual a 94,6%.

No terceiro trabalho (OLIVEIRA et al., 2018), os autores apresentaram uma metodologia de detecção e reconhecimento de distorção arquitetural em mamografia digital, cuja finalidade era reduzir o número de falsos positivos retornados pelo sistema CAD, mantendo todos os verdadeiros positivos. Para tal, extraíram diversas características, sendo as principais valor médio, intervalo e variância sobre a imagem de magnitude de Gabor, contraste sobre a imagem de intensidade de níveis de cinza, *skewness* sobre a imagem de códigos LBP, contraste sobre a imagem de códigos CS-LBP e correlação sobre a imagem de códigos LMP. O melhor resultado foi obtido com a aplicação do descritor LMP. No quarto trabalho (OLIVEIRA et al., 2019), os autores apresentaram uma nova versão do banco de mamografia digital, expandida para 350 mamografias digitais, das quais 175 contêm DA, e as demais 175, são imagens sem DA. Foram extraídas 50 ROIs de cada mamografia, totalizando 17500 regiões de interesse. No processo de recorte das ROIs nas mamografias com DA, garantiu-se que o centro da DA estaria contida nas respectivas ROIs. Para o reconhecimento da distorção arquitetural, utilizaram-se os descritores LBP, GLCM (apenas a medida estatística de energia) e o *Autoencoder*. Os melhores resultados obtidos para cada descritor foram $\text{Acc} = 75,3\%$ e $\text{AUC} = 82,9\%$ para o *Autoencoder*, $\text{Acc} = 74,7\%$ e $\text{AUC} = 80,2\%$ para GLCM e $\text{Acc} = 64,7\%$ e $\text{AUC} = 70,1\%$ para o LBP.

Estimulado pelos resultados apresentados nos trabalhos supracitados, em que cada descritor individual apresentou um comportamento particular para cada banco de mamografia utilizado, assim como pelo fato de alguns autores aplicarem o operador Concatenação a diferentes vetores de características, enuncia-se a hipótese de que a fusão de descritores de textura aplicada a reconhecimento de distorção arquitetural em mamografia digital resulta em desempenho melhor que o obtido com a aplicação dos descritores individuais.

6.2 Materiais e Métodos

Com o propósito de verificar a validade da hipótese de que a fusão de descritores de textura melhora o desempenho do processo de reconhecimento de distorção arquitetural em mamografia digital, desenvolveu-se um procedimento para aplicação de diferentes operadores de fusão na camada de características e na camada de decisão, semelhante ao apresentado no Capítulo 5. Primeiramente, aplicaram-se os descritores de textura baseados em fusão de características, propostos neste trabalho de doutorado, sobre as regiões de interesse das mamografias, sendo eles CMLMP, CMedianLMP e CZTZLMP. O desempenho da aplicação individual dos descritores é comparado ao obtido com a aplicação de outros descritores publicados na literatura, a saber CLMP, CLBP, ZTZLBP, WLD, GLCM e filtro de Gabor. Em seguida, realizou-se a análise com a aplicação dos operadores de fusão propostos, Soma^M e Gram^M, sobre os vetores de características dos descritores de textura, cujos desempenhos foram comparados ao obtido com a aplicação dos operadores Concatenação e Combinação Conjunta.

6.2.1 Base de Mamografia Digital

Oliveira et al. (2019) organizaram um banco de mamografia digital composto por 350 mamografias, tanto vista mediolateral oblíqua (MLO) como vista craniocaudal (CC), das quais 175 apresentam ocorrência de distorção arquitetural e as demais 175 não possuem DA. Todas as mamografias possuem laudo médico e as distorções arquiteturais foram contornadas manualmente por um radiologista experiente. As mamografias foram obtidas no Instituto de Radiologia (InRAD) da Universidade de São Paulo (USP) com aprovação pelo Comitê de Ética em Pesquisa¹.

Apresenta-se, na Figura 23, uma das mamografias do banco de dados, pose mediolateral oblíqua, com ocorrência de distorção arquitetural sinalizada pelo radiologista e reforçada com um círculo vermelho concêntrico ao centro da DA.

Uma vez sinalizada a DA, foram extraídas, de cada mamografia, 50 regiões de interesse (ROI) de dimensão 256×256 , sendo garantido que o centro da distorção arquitetural esteja contido em cada ROI. O processo de extração das diferentes amostras para as

¹ CAAE #56699016:7:0000:0065

Figura 23: Mamografia digital com ocorrência de distorção arquitetural, sinalizada por radiologista com um círculo vermelho (extraído de (OLIVEIRA et al., 2019)).

Figura 24: Extração das regiões de interesse em mamografia digital com distorção arquitetural (adaptado de (OLIVEIRA et al., 2019)).

mamografias com DA é ilustrado na Figura 24, sendo que os organizadores do banco de mamografia digital definiram um limiar de distância entre o centro da DA e a borda da região de interesse para evitar que o centro da DA fique muito próximo das bordas. Para as mamografias sem DA, as 50 amostras de cada imagem foram extraídas aleatoriamente, considerando-se apenas busca dentro da região mamária. O banco de mamografia digital

Figura 25: Exemplos de amostras extraídas de diferentes mamografias (a) com DA e (b) sem DA.

consiste, assim, em 17.500 amostras de mamografias digitais (OLIVEIRA et al., 2019). Algumas amostras extraídas com DA e sem DA são apresentadas na Figura 25.

Utilizou-se, para os experimentos realizados nesta tese, um subconjunto do banco de mamografia digital descrito acima, composto por 10 amostras de cada mamografia digital MLO. São 88 mamografias com DA e 88 sem DA, totalizando 1.760 amostras.

6.2.2 Extração de Características e Fusão das Características

Apresenta-se na Figura 26 o diagrama em blocos da metodologia proposta de reconhecimento de distorção arquitetural em mamografia digital, sendo o problema modelado como classificação de amostras em duas classes: com distorção arquitetural, classe DA, e sem distorção arquitetural, classe não DA. Aplicaram-se os descritores propostos *Completed Mean Local Mapped Pattern* (CMLMP), *Completed Median Local Mapped Pattern* (CMedianLMP) e *Completed Z with Tilted Z Local Mapped Pattern* (CZTZLMP) sobre as

Figura 26: Metodologia proposta de reconhecimento de distorção arquitetural.

amostras extraídas das mamografias digitais, considerando três diferentes configurações de vizinhança, sendo elas $R = 1$ e $N = 8$, $R = 2$ e $N = 16$, e $R = 3$ e $N = 24$. Os parâmetros B_S , B_M , B_C , β_S , β_M e β_C , referentes ao valor máximo do código gerado e à inclinação da curva sigmóide das respectivas componentes Sinal, Magnitude e Centro dos descritores, foram otimizados por algoritmo genético, utilizando Acurácia como função objetivo e 20% das imagens do banco de dados.

Também foram aplicados outros seis descritores publicados na literatura: CLMP (VIEIRA; NEGRI; GONZAGA, 2018), CLBP (CHEN et al., 2016), ZTZLBP (MARY; DEJEY, 2018), WLD (CHEN et al., 2009), GLCM (HARALICK; SHANMUGAM; DINSTEN, 1973) e filtro de Gabor (RANGAYYAN; AYRES, 2007; RANGAYYAN; AYRES, 2006).

Para os dois primeiros, CLMP e CLBP, foram utilizadas as mesmas configurações de vizinhança apresentadas para os descritores propostos. Os parâmetros B_S , B_M , β_S e β_M para o CLMP também foram otimizados por algoritmo genético. Utilizou-se a versão do CLBP uniforme e invariante a rotação com duas transições, $CLBP^{riu2}$. Para o WLD, foram utilizadas vizinhança quadrada 3×3 , oito orientações e oito excentricidades. As componentes do WLD, WLD_ϵ e WLD_ϕ , foram aplicadas individualmente. Para o descritor GLCM, foram consideradas as métricas energia, correlação, entropia e soma das variâncias para $d = 2$ e $d = 3$ na orientação 90° , e as métricas contraste e homogeneidade para $d = 2$ e $d = 3$ nas orientações $\theta = 90^\circ$ e $\theta = 135^\circ$, resultando em vetor de características GLCM com dezesseis elementos.

Para o descritor filtro de Gabor, calcularam-se treze diferentes características sobre as imagens de magnitude e de orientação, assim como para a imagem de intensidade de níveis de cinza, tal como apresentado por Rangayyan et al. (2013), sendo as características IDS, índice de divergência das espículas, e RWD, diferença ponderada radial, calculadas sobre as imagens de magnitude e de intensidade; RWD angular sobre a imagem de orientação; AWD, diferença ponderada angular, e AWDES, diferença ponderada angular sobre a entropia das espículas, calculadas sobre as imagens de magnitude, de orientação e de intensidade; e as medidas estatísticas média e variância foram calculadas sobre a imagem de magnitude.

Ao aplicar os descritores sobre as amostras de mamografia, geraram-se as imagens de códigos I_{DESC} , com exceção para os descritores GLCM e filtro de Gabor, cuja aplicação gera diretamente os vetores de características GLCM e Gabor. A representação 1D das imagens de códigos, determinada pelo histograma 1D de I_{DESC} , gerou os vetores de características individuais $ZTZLBP_{1,8}$, WLD_ϵ , WLD_ϕ , $DESC_S_{R,N}$, $DESC_M_{R,N}$ e $DESC_C_{R,N}$, em que $DESC = \{CMLMP, CMedianLMP, CZTZLMP, CLMP, CLBP\}$.

Combinaram-se as informações complementares entre diferentes descritores com a aplicação dos operadores de fusão Soma^M e Gram^M, propostos neste trabalho, assim como do operador Concatenação. Estes três operadores foram aplicados, primeiramente, sobre os vetores de características das diferentes componentes de um mesmo descritor, considerando resolução única (mesma configuração de vizinhança) e multiresolução (diferentes configurações de vizinhança). Posteriormente, os operadores de fusão foram aplicados sobre os vetores de características de diferentes descritores, tomados dois a dois, três a três e quatro a quatro. Para evitar predominância de determinado vetor de características sobre outro devido a diferença de grandeza entre as características, realizou-se a normalização por norma Z_s dos vetores de características antes da fusão (ABDEL-NASSER et al., 2016).

Para se combinar as imagens de códigos, aplicou-se o operador de fusão Combinação Conjunta. Primeiramente, sobre as imagens de códigos I_{DESC} geradas pelas diferentes componentes de um mesmo descritor, tanto para resolução única como para multiresolução.

Posteriormente, aplicou-se o operador Combinação Conjunta sobre as imagens de códigos geradas por diferentes descritores, tomados dois a dois e três a três.

Utilizou-se ainda a fusão Híbrida, que caracteriza-se pela Concatenação do vetor de características de um determinado descritor (DESC1) com o vetor de características resultante da Combinação Conjunta entre duas imagens de códigos geradas por dois outros descritores (DESC2|DESC3). Foram consideradas as configurações resolução única e multiresolução para um mesmo descritor, assim como fusão entre diferentes descritores.

6.2.3 Classificação

Dispondo dos vetores de características individuais e combinados, seguiu-se para a etapa de classificação das amostras de mamografia em DA e em não DA, para a qual utilizou-se o classificador K vizinhos mais próximos (K-NN) com o método de validação cruzada por unidade². Como mostrado na Figura 27, durante o processo de busca pelos vizinhos mais próximos, a amostra de consulta, a qual deseja-se classificar, é comparada a todas as amostras do banco de mamografias, com exceção das N-1 ROIs pertencentes à mesma mamografia que a amostra de consulta.

Figura 27: Método de validação cruzada por unidade, em que o banco de amostras consiste em todas as amostras de mamografia, com exceção das N-1 ROIs pertencentes à mesma mamografia que a amostra de consulta.

Determina-se a similaridade entre duas amostras pelo cálculo da distância Euclidiana entre o vetor de características da amostra de consulta (**A**) e o vetor de características da

² Traduzido do inglês *leave-one-out cross validation method*.

i-ésima amostra do banco de mamografias ($\mathbf{M}_i$), determinada pela Equação (6.1), em que B é o número de elementos dos vetores de características.

$$d(\mathbf{A}, \mathbf{M}_i) = \sqrt{\sum_{b=1}^B (A_b - M_{i_b})^2} \quad (6.1)$$

A amostra de consulta é classificada como sendo da mesma classe à qual pertence a maioria dos vizinhos dentre os K vizinhos mais próximos. Ou seja, se $\frac{K}{2} + 1$ das amostras dentre os K vizinhos mais próximos forem DA, a amostra de consulta é classificada como sendo DA, caso contrário, como não DA. Inicialmente, foi utilizado $K = 1$ (classificador 1-NN); posteriormente, otimizou-se o valor de K a fim de melhorar o desempenho da classificação.

A performance do sistema proposto de reconhecimento de distorção arquitetural é medida pelas métricas de desempenho Sensibilidade (S), Acurácia (Acc) e Especificidade (E). A Sensibilidade, ou *recall*, tal como apresentado no Capítulo 5, é definida pela Equação (6.2), sendo o percentual de amostras DA corretamente classificadas como DA; a Acurácia, Equação (6.3), é a razão entre todas as amostras corretamente classificadas (DA e não DA) e o número total de amostras no banco de mamografias; e a Especificidade é o percentual de amostras não DA corretamente classificadas como não DA, sendo definida na Equação (6.4).

$$S = \frac{VP}{VP + FN} \quad (6.2)$$

$$Acc = \frac{VP + VN}{VP + FN + VN + FP} \quad (6.3)$$

$$E = \frac{VN}{VN + FP} \quad (6.4)$$

Na modelagem do problema, considerou-se verdadeiro positivo (VP) como sendo o número de amostras DA corretamente classificadas como DA; falso negativo (FN) é o número de amostras DA erroneamente classificadas como não DA; verdadeiro negativo (VN) é o número de amostras não DA corretamente classificadas como não DA; e falso positivo (FP), refere-se ao número de amostras não DA erroneamente classificadas como DA.

A ocorrência de FN é o pior caso nesta aplicação, sendo o custo de FN maior que o de FP. Deste modo, um determinado descritor (individual ou combinado) pode ser considerado mais apropriado que outros para uso em reconhecimento de distorção arquitetural ao resultar no maior valor de Sensibilidade com alta Acurácia. Ainda assim, é importante observar o valor da Especificidade, para garantir que sejam desconsideradas as abordagens

que permitam classificar bem as amostras DA, mas que classifiquem erroneamente grande número de amostras não DA.

6.3 Resultados

6.3.1 Otimização dos Parâmetros

Como apresentado nas seções anteriores, os descritores de textura baseados em fusão de características CMLMP, CMedianLMP, CZTZLMP e CLMP, são paramétricos, sendo necessário otimizar os valores de B e β para cada componente. Apresentam-se, na Tabela 12, os valores otimizados dos respectivos parâmetros para cada componente dos descritores, considerando diferentes configurações de vizinhança. Utilizou-se otimização por algoritmos genéticos (NEGRI et al., 2017) configurado como descrito a seguir: $\beta \in [10^{-2}, 40]$, $B \in [2, 256]$, tamanho da população igual a 50 indivíduos e número máximo de gerações igual a 40, sendo usado 20% das amostras do banco de imagens.

Tabela 12: Parâmetros de cada descritor otimizados para a base de dados de mamografia digital

Descritor	β_S	β_M	β_C	B_S	B_M	B_C
CMLMP _{1,8} *	0,1516	20,4247	26,2579	41	6	8
CMLMP _{2,16} *	0,1403	14,3807	-	26	18	-
CMLMP _{3,24} *	0,0406	13,9464	-	34	24	-
CMedianLMP _{1,8} *	0,0332	8,3250	26,7356	40	7	19
CMedianLMP _{2,16} *	0,2390	14,5275	-	83	38	-
CMedianLMP _{3,24} *	0,3654	13,3607	-	62	30	-
CZTZLMP _{1,8} *	0,4103	5,5100	26,2579	9	25	8
CZTZLMP _{2,16} *	2,2125	3,8102	-	15	48	-
CZTZLMP _{3,24} *	4,4987	6,8023	-	32	19	-
CLMP _{1,8}	3,9095	39,8446	-	31	2	-
CLMP _{2,16}	1,3648	2,2203	-	32	20	-
CLMP _{3,24}	5,8183	2,9051	-	7	4	-

* Métodos propostos.

A escolha destes parâmetros interfere significativamente no desempenho da classificação, como ilustrado na Figura 28, em que pode-se visualizar o comportamento do desempenho do sistema, em termos de Acurácia, quando da aplicação do descritor CMLMP, com a variação de β no intervalo $(0, 6]$ e de B no intervalo $(0, 64]$. Salienta-se que os parâmetros utilizados para gerar essas curvas estão em um intervalo menor que o utilizado na etapa de otimização com a finalidade de permitir melhor visualização do comportamento das curvas, sendo que os valores ótimos não estão necessariamente compreendidos nos respectivos gráficos. Nota-se que pequenas variações nos valores de β_S e/ou de B_S interferem abruptamente no valor de Acurácia para o descritor CMLMP_S. Por outro lado, as componentes Magnitude, CMLMP_M, e Centro, CMLMP_C, apresentam uma superfície mais suave para o intervalo considerado.

Figura 28: Acurácia obtida com a aplicação do descritor (a) CMLMP_S, (b) CMLMP_M, e (c) CMLMP_C, com a variação de β e B .

A influência dos parâmetros otimizados β_S , β_M e β_C sobre a função de mapeamento para o descritor CMLMP é mostrada na Figura 29. Inicialmente, ao analisar um grande intervalo para os valores das diferenças entre os níveis de cinza em uma determinada vizinhança, g_{diff} , a função de mapeamento sigmóide, para o CMLMP_S, apresenta comportamento semelhante ao da função degrau (Heaviside); tal comportamento é consequência do pequeno valor atribuído a β_S ($\beta_S = 0,1516$). Conquanto mapeie $f_S(g_{diff})$ para 0, caso $g_{diff} \leq -1$, ou para 1, caso $g_{diff} \geq 1$, representa um mapeamento suave para g_{diff} compreendido no intervalo $(-1, 1)$. As funções de mapeamento para CMLMP_C e CMLMP_M apresentam curvas similares entre si, no entanto, como $\beta_C > \beta_M$, $f_C(g_{diff})$ retorna valores mais próximos a 0,5 que os retornados por $f_M(g_{diff})$ para o mesmo valor de g_{diff} . Ressalta-se que não somente o comportamento da função de mapeamento é diferente para cada componente do descritor, mas também o ponto de operação das diferenças de intensidade, uma vez que as regras de determinação de g_{diff} para as componentes Sinal,

Figura 29: Comportamento da função de mapeamento para o descritor CMLMP para β_S , β_M e β_C ótimos, com a variação das diferenças g_{diff} no intervalo (a) $[-100, 100]$ e (b) $[-2, 2]$.

Magnitude e Centro para um mesmo descritor são diferentes entre si, desse modo, para uma mesma vizinhança, cada componente terá um ponto diferente de atuação na função de mapeamento, uma vez que $g_{diff_S} \neq g_{diff_M} \neq g_{diff_C}$.

6.3.2 Desempenho para Descritores Individuais

Após a otimização dos parâmetros, os descritores individuais foram aplicados ao conjunto de amostras de mamografias para reconhecimento de distorção arquitetural, considerando vizinhança circular com $R = 1$ e $N = 8$. Inicialmente, utilizou-se o classificador Primeiro Vizinho Mais Próximo (K-NN com $K = 1$), para o qual foram obtidos resultados insatisfatórios, apresentando baixo valor de Acurácia. Desse modo, variou-se o valor de K no intervalo $[1, 500]$ a fim de se determinar o número de vizinhos mais próximos a ser considerado na etapa de classificação. Apresenta-se, na Figura 30, o comportamento dos descritores individuais para os diferentes valores de K considerados.

Nota-se que o desempenho da maioria dos descritores aumenta com o aumento de K no intervalo de 1 a 20, sendo que a partir de 20 vizinhos mais próximos, não há aumento significativo do desempenho dos descritores, com exceção do CLBP_C, que apresenta um ganho de desempenho ao passar de $K = 20$ para $K = 50$, seguida por nova queda de desempenho para $K > 50$. Os descritores CMedianLMP_M e Gabor, apresentam desempenho máximo em $K = 20$, seguido por redução significativa do valor de Acurácia quando da passagem para $K = 50$, sendo que o descritor Gabor apresenta queda abrupta, de aproximadamente 75% em $K = 20$ para aproximadamente 55% em $K = 200$. A partir da análise do comportamento da Acurácia em função do número de vizinhos mais próximos para diferentes descritores, optou-se pelo uso de $K = 20$.

Figura 30: Acurácia dos descritores individuais para diferentes valores de K , considerando $R = 1$ e $N = 8$.

Com a aplicação dos descritores de textura sobre as amostras do banco de mamografias digitais, diversas imagens de códigos foram geradas, uma para cada descritor, com exceção do GLCM e do filtro de Gabor. Apresentam-se, na Figura 31, as imagens de códigos obtidas com a aplicação dos descritores $\text{CMLMP}_{1,8}$, $\text{CLMP}_{1,8}$ e $\text{CLBP}_{1,8}^{\text{riu}2}$ sobre uma amostra com distorção arquitetural. Nota-se que as imagens de códigos para as componentes Sinal possuem mais informação de textura que as respectivas imagens de códigos para a componentes Magnitude, as quais retêm maior informação espacial. As imagens de códigos da componente Centro são resultado de uma transformação de intensidade de nível de cinza no domínio do espaço aplicada sobre a imagem original, sendo as imagens de códigos obtidas para os descritores CLMP_C e CLBP_C resultado de uma binarização aplicada sobre a imagem original com limiar igual ao valor médio das intensidades de cinza de toda a imagem, e a imagem de códigos para o CMLMP_C resulta de uma transformação de intensidade de nível de cinza aplicada com função de mapeamento sigmóide ($\beta_C = 26.2579$), seguida por uma quantização em 3 bits, com valores de códigos compreendidos no intervalo $[0,7]$ ($B_C = 8$). Esta figura ilustra a

Figura 31: Imagens de códigos obtidas com a aplicação dos descritores CMLMP, CLMP e CLBP sobre uma amostra com DA.

complementaridade de informação entre as diferentes componentes de um mesmo descritor, assim como entre diferentes descritores.

Os resultados de Sensibilidade, Acurácia e Especificidade obtidos com a aplicação dos descritores individuais para diferentes configurações de vizinhança são discriminados na Tabela 13. O maior valor de sensibilidade foi obtido para o descritor CMedianLMP_M_{1,8}, $S = 84,7\%$, no entanto, foi retornado um baixo valor de Especificidade, $E = 55,3\%$, significando que o descritor classifica bem as amostras que são da classe DA, mas erra bastante quando da classificação de amostras da classe não DA. O segundo maior valor de Sensibilidade foi obtido com o descritor GLCM, $S = 81,4\%$, apresentando Acurácia igual a $76,9\%$ e Especificidade de $72,4\%$. Os descritores CMedianLMP_C e CMLMP_C

Tabela 13: Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos descritores individuais sobre o banco de mamografias digitais (valores em %)

Descritor	B	S	Acc	E
CMLMP_ $S_{1,8}$ *	41	75,8	77,4	79,1
CMLMP_ $M_{1,8}$ *	6	74,1	67,2	60,3
CMLMP_C, CZTZLMP_C*	8	78,3	77,7	77,1
CMLMP_ $S_{2,16}$ *	26	76,5	76,3	76,0
CMLMP_ $M_{2,16}$ *	18	76,3	70,8	65,3
CMLMP_ $S_{3,24}$ *	34	79,7	77,0	74,3
CMLMP_ $M_{3,24}$ *	24	73,6	69,7	65,7
CMedianLMP_ $S_{1,8}$ *	40	75,7	76,9	78,1
CMedianLMP_ $M_{1,8}$ *	7	84,7	70,0	55,3
CMedianLMP_C*	19	80,6	78,8	77,1
CMedianLMP_ $S_{2,16}$ *	83	75,3	74,5	73,8
CMedianLMP_ $M_{2,16}$ *	38	72,1	69,3	66,6
CMedianLMP_ $S_{3,24}$ *	62	76,6	74,2	71,7
CMedianLMP_ $M_{3,24}$ *	30	70,8	69,1	67,4
CZTZLMP_ $S_{1,8}$ *	9	67,1	72,4	77,7
CZTZLMP_ $M_{1,8}$ *	25	75,2	72,3	69,3
CZTZLMP_ $S_{2,16}$ *	15	71,7	74,4	77,2
CZTZLMP_ $M_{2,16}$ *	48	72,3	69,7	67,1
CZTZLMP_ $S_{3,24}$ *	32	73,0	76,4	79,9
CZTZLMP_ $M_{3,24}$ *	19	69,8	70,5	71,3
CLMP_ $S_{1,8}$	31	72,6	75,3	78,0
CLMP_ $M_{1,8}$	2	73,6	75,8	78,0
CLMP_C, CLBP_C	2	50,9	54,3	57,6
CLMP_ $S_{2,16}$	32	73,2	73,9	74,6
CLMP_ $M_{2,16}$	20	75,7	74,2	72,6
CLMP_ $S_{3,24}$	7	70,5	74,9	79,4
CLMP_ $M_{3,24}$	4	72,5	72,4	72,3
CLBP_ $S_{1,8}^{riu2}$	10	64,2	71,8	79,3
CLBP_ $M_{1,8}^{riu2}$	10	72,6	74,2	75,8
CLBP_ $S_{2,16}^{riu2}$	18	70,0	71,3	72,5
CLBP_ $M_{2,16}^{riu2}$	18	74,8	74,3	73,9
CLBP_ $S_{3,24}^{riu2}$	26	72,7	70,9	69,1
CLBP_ $M_{3,24}^{riu2}$	26	71,3	72,4	73,6
ZTZLBP $_{1,8}$	16	71,0	73,4	75,8
WLD $_{\epsilon}$	8	75,8	75,3	74,5
WLD $_{\Phi}$	8	56,5	65,2	73,9
GLCM	16	81,4	76,9	72,4
Gabor	15	70,5	73,5	76,5

As células destacadas em cinza escuro e em cinza claro representam o melhor e o segundo melhor resultado obtido.

* Métodos propostos.

retornaram os maiores valores de Acurácia, iguais a 78,8% e 77,7%, respectivamente.

Dentre as diferentes configurações do descritor CMLMP, nota-se que as componen-

tes Sinal apresentam melhor desempenho (S, Acc e E) que as componentes Magnitude, sendo o maior valor de Sensibilidade obtido para CMLMP_S_{3,24}. Algumas componentes apresentam um equilíbrio entre os três valores retornados, como observado para o descritor CMLMP_S_{2,16} e para CMLMP_C, o que significa que apresentam desempenho semelhante na classificação de amostras DA e de amostras não DA. Outros descritores têm comportamento diferente entre as duas classes, como o caso geral das três componentes de Magnitude abordadas para o CMLMP, que apresentam Sensibilidade aproximadamente 15% superior aos valores de Especificidade.

O descritor CMedianLMP apresenta comportamento semelhante ao observado para o CMLMP: as componentes Sinal apresentam desempenho superior às componentes Magnitude, com exceção para CMedian_LMP_{1,8}, como já mencionado; e as componentes Magnitude também apresentam Sensibilidade notoriamente maior que Especificidade. Apresentando menor desempenho que os outros dois descritores propostos, as componentes Sinal do CZTZLMP apresentam maior Especificidade que Sensibilidade, sendo o maior valor de Especificidade dentre todos os descritores obtido com a aplicação do CZTZLMP_S_{3,24}, igual a 79,9%, e Sensibilidade igual a 73,0%.

Os descritores propostos, quando aplicados individualmente, superaram o desempenho dos descritores CLMP, CLBP, ZTZLBP, WLD e Gabor. O pior desempenho foi obtido com o descritor CLBP_C (igual ao CLMP_C). Tal comportamento poderia ser associado à dimensão do descritor, uma vez que apresenta vetor de características com apenas dois elementos. No entanto, o descritor CLMP_M_{1,8}, de mesma dimensão, apresenta desempenho aproximadamente 20% superior. É provável que o baixo desempenho do CLBP_C seja pelo fato de a imagem de códigos por ele gerada ser uma binarização da imagem original, como já discutido, sendo possível que essa informação seja pouco discriminante.

Apresenta-se, na Tabela 14, o ganho obtido com a aplicação do descritor individual CMedianLMP_C, o qual retornou o melhor desempenho dentre os descritores propostos, quando comparado ao desempenho dos descritores individuais da literatura. O descritor proposto supera em Sensibilidade e em Acurácia todas as componentes dos descritores CLMP, CLBP, ZTZLBP, WLD e Gabor, com acréscimo de 4,8% em S e de 3,5% em Acc em relação ao WLD_e a 29,7% em S e 24,5% em Acc em relação ao CLMP_C. No entanto, apresenta menor Sensibilidade que o descritor GLCM, com decréscimo de 0,8% em S, mas acréscimo de 1,9% em Acc. Ao analisar a métrica de desempenho Especificidade, nota-se que o descritor proposto supera quatorze dos dezoito descritores individuais da literatura, mas apresenta Especificidade inferior aos demais quatro descritores: CLMP_S_{1,8} e CLMP_M_{1,8}, com decréscimo de 0,9% em E, mas acréscimo de 8,0% e 7,0% em S, respectivamente, e de 3,5% e 3,0% em Acc, respectivamente; CLMP_S_{3,24}, com decréscimo de 2,3% em E, mas acréscimo de 10,1% em S e de 3,9% em Acc; e CLBP_S_{1,8}^{riu2}, com decréscimo de 2,2% em E, mas acréscimo de 16,4% em S e de 7,0% em Acc.

Tabela 14: Ganho (%) com a aplicação do descritor individual proposto CMedianLMP_C sobre o banco de mamografia digital

Descritor	B	S	Acc	E
CMedianLMP_C*	19	80,6	78,8	77,1
CLMP_S _{1,8}	31	↑8,00	↑3,50	↓0,90
CLMP_M _{1,8}	2	↑7,00	↑3,00	↓0,90
CLMP_C, CLBP_C	2	↑29,7	↑24,5	↑19,5
CLMP_S _{2,16}	32	↑7,40	↑4,90	↑2,50
CLMP_M _{2,16}	20	↑4,90	↑4,60	↑4,50
CLMP_S _{3,24}	7	↑10,1	↑3,90	↓2,30
CLMP_M _{3,24}	4	↑8,10	↑6,40	↑4,80
CLBP_S _{1,8} ^{riu2}	10	↑16,4	↑7,00	↓2,20
CLBP_M _{1,8} ^{riu2}	10	↑8,00	↑4,60	↑1,30
CLBP_S _{2,16} ^{riu2}	18	↑10,6	↑7,50	↑4,60
CLBP_M _{2,16} ^{riu2}	18	↑5,80	↑4,50	↑3,20
CLBP_S _{3,24} ^{riu2}	26	↑7,90	↑7,90	↑8,00
CLBP_M _{3,24} ^{riu2}	26	↑9,30	↑6,40	↑3,50
ZTZLBP _{1,8}	16	↑9,60	↑5,40	↑1,30
WLD _ε	8	↑4,80	↑3,50	↑2,60
WLD _Φ	8	↑24,1	↑13,6	↑3,20
GLCM	16	↓0,80	↑1,90	↑4,70
Gabor	15	↑10,1	↑5,30	↑0,60

*Método proposto.

Nenhum dos descritores, aplicados individualmente, apresentou o maior resultado para as três métricas de desempenho analisadas. A fim de obter melhores resultados para o processo de reconhecimento de DA, aplicaram-se os operadores propostos Soma^M e Gram^M inicialmente aos descritores baseados em fusão de características, sendo combinadas as diferentes componentes para um mesmo descritor. Os resultados são comparados aos obtidos com a aplicação dos operadores Concatenação e Combinação Conjunta, os quais são apresentados na próxima seção.

6.3.3 Desempenho para Descritores Baseados em Fusão de Características

Os quatro operadores de fusão na camada de características abordados nesta tese foram aplicados sobre as componentes dos descritores baseados em fusão de características CMLMP, CMedianLMP, CZTZLMP, CLMP, CLBP e WLD, descritor a descritor, considerando as diferentes configurações de vizinhança. Foram geradas 162 diferentes combinações para cada descritor, com exceção do WDL, para o qual geraram-se apenas quatro combinações, tendo em vista que possui somente duas componentes. Os resultados foram compilados, sendo apresentados apenas os mais relevantes, os quais possibilitaram o melhor desempenho para cada abordagem de fusão considerada.

São apresentados, na Tabela 15, os melhores resultados obtidos com a fusão das

Tabela 15: Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($|$), Gram^M ($||_G$) e Soma^M ($||_+$) sobre o descritor CMLMP na camada de características (valores em %)

Descritor	B	S	Acc	E
CMLMP_ $S_{3,24}$	34	79,7	77,0	74,3
CMLMP_C	8	78,3	77,7	77,1
CMLMP_ $S_{1,8}$	41	75,8	77,4	79,1
CMLMP_ $S_{3,24} \oplus$ CMLMP_ $M_{3,24} \oplus$ CMLMP_C	66	84,0	81,0	78,0
CMLMP_ $S_{3,24} \oplus$ CMLMP_ $M_{2,16} \oplus$ CMLMP_C	60	84,1	81,3	78,4
CMLMP_ $S_{3,24} $ CMLMP_C	272	85,8	79,4	73,1
CMLMP_ $S_{3,24} $ CMLMP_ $M_{1,8} $ CMLMP_C	1.632	84,5	78,8	73,1
CMLMP_ $M_{3,24} \oplus$ (CMLMP_ $S_{3,24} $ CMLMP_C)	296	85,7	79,4	73,1
CMLMP_ $M_{1,8} \oplus$ (CMLMP_ $S_{3,24} $ CMLMP_C)	278	85,8	79,4	73,1
CMLMP_ $S_{3,24} _G$ CMLMP_ $M_{3,24}$	595	80,2	77,1	74,0
CMLMP_ $S_{3,24} _G$ CMLMP_ $M_{2,16}$	595	79,9	76,4	73,0
CMLMP_ $S_{3,24} _+$ CMLMP_ $M_{3,24} _+$ CMLMP_C	34	82,4	79,5	76,7
CMLMP_ $S_{3,24} _+$ CMLMP_ $M_{1,8} _+$ CMLMP_C	34	82,6	80,3	78,1

As células destacadas em cinza escuro e em cinza claro representam o melhor e o segundo melhor resultado obtido.

componentes do descritor CMLMP, considerando resolução única e multiresolução, sendo as três primeiras linhas da tabela preenchidas com os melhores desempenhos individuais. Para os operadores Concatenação e Soma^M, a fusão por multiresolução resultou em melhor desempenho que a fusão por resolução única, sendo que o operador Concatenação apresenta incremento de 1,5% na Sensibilidade e de 1% na Acurácia, quando comparado ao Soma^M, apresentando valores semelhantes para Especificidade. Para os operadores Combinação Conjunta e Gram^M, a fusão por resolução única gerou melhores resultados, sendo que o operador Combinação Conjunta apresenta Sensibilidade 5,6% superior ao Gram^M, Acurácia 2,3% superior e Especificidade 0,9% menor. Nota-se ainda que ambos operadores implicam em aumento da dimensão dos vetores de características, sendo que o vetor de características para a abordagem por multiresolução do operador Combinação Conjunta apresenta alta dimensão ($B = 1.632$). A fusão híbrida não agregou valor ao sistema, uma vez que reproduziu o resultado obtido pelo descritor CMLMP_ $S_{3,24}|$ CMLMP_C. Todas as abordagens de fusão apresentadas resultaram em maior Sensibilidade que o melhor resultado individual; somente os descritores combinados por Gram^M apresentaram menor Acurácia que as abordagens individuais; e nenhum dos descritores combinados obtiveram Especificidade superior à da componente CMLMP_ $S_{1,8}$. Nota-se ainda que a componente individual CMLMP_ $S_{3,24}$ que resultou em maior valor de Sensibilidade dentre as demais componentes do CMLMP, integra todas as abordagens de fusão cujos resultados foram relevantes, e que a componente CMLMP_C, com maior valor de Acurácia e segundo maior valor de Sensibilidade dentre as demais componentes individuais para o descritor em questão, compõe 8 das 10 abordagens relevantes.

Tabela 16: Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($|$), Gram^M ($\|_G$) e Soma^M ($\|_+$) sobre o descritor CMedianLMP na camada de características (valores em %)

Descritor	B	S	Acc	E
CMedianLMP_M _{1,8}	7	84,7	70,0	55,3
CMedianLMP_C	19	80,6	78,8	77,1
CMedianLMP_S _{1,8}	40	75,7	76,9	78,1
CMedianLMP_S _{3,24} $\oplus$ CMedianLMP_C	81	85,8	80,7	75,6
CMedianLMP_S _{3,24} $ $ CMedianLMP_C	1178	86,0	78,6	71,3
CMedianLMP_M _{1,8} $\oplus$ $\oplus$ (CMedianLMP_S _{3,24} $ $ CMedianLMP_C)	1185	86,0	78,6	71,3
CMedianLMP_M _{1,8} $\ _G$ CMedianLMP_C	190	80,0	78,5	76,9
CMedianLMP_S _{3,24} $\ _+$ CMedianLMP_M _{2,16} $\ _+$ $\ _+$ CMedianLMP_C	62	86,9	81,0	75,0

As células destacadas em cinza escuro e em cinza claro representam o melhor e o segundo melhor resultado obtido.

Os resultados dos operadores de fusão aplicados ao descritor CMedianLMP são apresentados na Tabela 16. O operador de fusão Soma^M resultou nos melhores resultados de Sensibilidade e de Acurácia, representando um incremento de 0,9% e 0,3% em relação aos segundos maiores valores, obtidos pelos operadores Combinação Conjunta e Concatenação, respectivamente. Ademais, o operador Soma^M apresenta o vetor de características de menor dimensão (B = 62) dentre as abordagens de fusão analisadas, contrastando com os gerados pela Combinação Conjunta e pela fusão Híbrida, iguais a 1.178 e 1.185 características, respectivamente. Tal como observado para o descritor CMLMP, nenhum operador de fusão implicou em maior Especificidade que a obtida pelos descritores individuais. O operador Gram^M foi o único que apresentou Sensibilidade menor que a obtida pelos descritores individuais. Novamente, o operador Híbrido não contribuiu para melhoria dos resultados, uma vez que apenas reproduziu os resultados obtidos com o descritor Combinação Conjunta. Nota-se ainda que a componente individual de maior Acurácia, CMedianLMP_C, integra todas as abordagens de fusão relevantes, e que a componente de maior Sensibilidade, CMedianLMP_M_{1,8}, compõe duas das cinco abordagens relevantes apresentadas. A componente CMedianLMP_S_{3,24}, que não apresenta os melhores resultados individuais, integra quatro das abordagens de fusão relevantes, inclusive a de maior Sensibilidade e Acurácia, CMedianLMP_S_{3,24} $\|_+$ CMedianLMP_M_{2,16} $\|_+$ CMedianLMP_C, que apresenta Sensibilidade e Acurácia 2,2% superiores aos melhores valores obtidos pelas componentes individuais do CMedianLMP, e Especificidade 3,1% inferior.

De acordo com os dados apresentados na Tabela 17, referentes à aplicação dos diferentes operadores de fusão sobre as componentes do descritor CZTZLMP, sendo considerado tanto resolução única como multiresolução, todas as abordagens de fusão resultaram em maior Sensibilidade e Acurácia que o melhor resultado apresentado pelos

Tabela 17: Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($|$), Gram^M ($\parallel_G$) e Soma^M ($\parallel_+$) sobre o descritor CZTZLMP na camada de características (valores em %)

Descritor	B	S	Acc	E
CZTZLMP_C	8	78,3	77,7	77,1
CZTZLMP_S _{3,24}	32	73,0	76,4	79,9
CZTZLMP_S _{3,24} $\oplus$ CZTZLMP_C	40	80,2	78,5	76,7
CZTZLMP_S _{2,16} CZTZLMP_M _{2,16} CZTZLMP_C	5760	82,8	78,0	73,2
CZTZLMP_S _{3,24} $\oplus$ $\oplus$ (CZTZLMP_M _{3,24} CZTZLMP_C)	184	82,0	78,0	74,0
CZTZLMP_S _{1,8} $\parallel_G$ CZTZLMP_M _{1,8} $\parallel_G$ $\parallel_G$ CZTZLMP_S _{2,16} $\parallel_G$ CZTZLMP_M _{2,16} $\parallel_G$ $\parallel_G$ CZTZLMP_S _{3,24} $\parallel_G$ CZTZLMP_M _{3,24} $\parallel_G$ $\parallel_G$ CZTZLMP_C	1176	79,0	78,7	78,4
CZTZLMP_S _{3,24} $\parallel_+$ CZTZLMP_C	32	79,9	78,2	76,5

As células destacadas em cinza escuro e em cinza claro representam o melhor e o segundo melhor resultado obtido.

descritores individuais. No entanto, nenhum descritor combinado apresentou melhor Especificidade que a obtida com a aplicação do descritor individual CZTZLMP_S_{3,24}. O maior valor de Sensibilidade foi atingido para o operador Combinação Conjunta, cujo desempenho representa um incremento de 3,8% em relação ao operador Gram^M e 2,9% em relação ao operador Soma^M. No entanto, o vetor de características obtido com o respectivo operador possui 5.760 elementos. O operador Híbrido resultou em um vetor de características de 184 elementos, com o qual obteve-se Sensibilidade 0,8% inferior, Acurácia de mesmo valor e Especificidade 0,8% maior que a melhor abordagem obtida com o operador Combinação Conjunta. A maior Acurácia foi obtida com o operador Gram^M aplicado à fusão das sete componentes CZTZLMP, obtendo vetor de características de 1.176 elementos, sendo que a aplicação do operador Concatenação permitiu Acurácia muito próxima com um vetor de características de dimensão reduzida, B = 40, Sensibilidade 1,2% maior, mas Especificidade 1,7% menor. Observa-se que a componente CZTZLMP_S_{3,24}, que retornou o maior valor de Especificidade dentre todos os descritores individuais, compõe quatro das cinco abordagens de fusão relevantes; e a componente CZTZLMP_C, que apresenta os maiores resultados de Sensibilidade e de Acurácia dentre as componentes individuais do descritor CZTZLMP, integra as cinco abordagens de fusão relevantes.

A combinação das diferentes componentes do descritor CLMP, apresentada na Tabela 18, permitiu incrementar a Sensibilidade em 3,2%, com o custo da redução da Acurácia em 0,8% e da Especificidade em 6,8%. A Combinação Conjunta resultou nos melhores desempenhos para o descritor CLMP, com um vetor de características de baixa dimensão, B = 80, sendo que a fusão Híbrida não apresentou contribuições. Tanto o CLMP quanto o CLBP obtiveram os melhores resultados dentre as abordagens de fusão combinando as componentes de Magnitude entre si para diferentes vizinhanças. Este

Tabela 18: Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($|$), Gram^M ($\parallel_G$) e Soma^M ($\parallel_+$) sobre os descritores individuais CLMP, CLBP e WLD na camada de características (valores em %)

Descritor	B	S	Acc	E
CLMP_M _{2,16}	20	75,7	74,2	72,6
CLMP_M _{1,8}	2	73,6	75,8	78,0
CLMP_S _{3,24}	7	70,5	74,9	79,4
CLBP_M ^{riu2} _{2,16}	18	74,8	74,3	73,9
WLD _ε	8	75,8	75,3	74,5
CLMP_M _{2,16} $\oplus$ CLMP_M _{3,24}	24	75,8	74,2	72,6
CLMP_M _{2,16} $ $ CLMP_M _{3,24}	80	78,9	75,0	71,1
CLMP_M _{1,8} $\oplus$ (CLMP_M _{2,16} $ $ CLMP_M _{3,24})	82	78,9	75,0	71,1
CLMP_M _{2,16} $\parallel_G$ CLMP_C	210	77,5	74,3	71,0
CLMP_M _{2,16} $\parallel_+$ CLMP_M _{3,24}	20	76,0	74,3	72,5
CLBP_M ^{riu2} _{1,8} $\oplus$ CLBP_M ^{riu2} _{2,16}	28	73,8	74,9	76,1
CLBP_M ^{riu2} _{1,8} $ $ CLBP_M ^{riu2} _{2,16}	180	74,8	74,5	74,3
CLBP_M ^{riu2} _{3,24} $\oplus$ (CLBP_M ^{riu2} _{1,8} $ $ CLBP_M ^{riu2} _{2,16})	206	74,8	74,5	74,3
CLBP_M ^{riu2} _{1,8} $\parallel_G$ CLBP_M ^{riu2} _{2,16} $\parallel_G$ CLBP_M ^{riu2} _{3,24}	351	72,3	72,7	73,1
CLBP_M ^{riu2} _{1,8} $\parallel_+$ CLBP_M ^{riu2} _{2,16}	18	73,8	73,4	73,1
WLD _ε $\oplus$ WLD _Φ	16	71,5	76,1	80,8
WLD _ε $ $ WLD _Φ	64	73,5	77,1	80,7
WLD _ε $\parallel_G$ WLD _Φ	36	61,6	71,4	81,1
WLD _ε $\parallel_+$ WLD _Φ	8	66,0	72,9	79,8

As células destacadas em cinza escuro e em cinza claro representam o melhor e o segundo melhor resultado obtido.

comportamento é diferente do que foi observado para os descritores CMLMP, CMedianLMP e CZTLMP, que obtiveram os melhores resultados com a combinação das diferentes componentes: Sinal e Magnitude, Sinal e Centro ou Magnitude e Centro. As diferentes abordagens de fusão geradas para o CLBP permitiram aumentar a Especificidade em 2,2% e a Acurácia em 0,6%, no entanto, não proporcionou aumento de Sensibilidade. Em sua formulação original, o descritor final do WLD é constituído por meio da Combinação Conjunta das duas componentes, sendo esta a abordagem que permitiu melhor Acurácia ao descritor. O maior valor de Especificidade foi obtido com a aplicação do operador Gram^M sobre as componentes do WLD.

Ao comparar estes resultados com os obtidos nas tabelas anteriores, percebe-se que os descritores de textura baseados em fusão de características, CMLMP e CMedianLMP, propostos neste trabalho, apresentaram os melhores resultados de Sensibilidade e de Acurácia dentre os descritores individuais e dentre os descritores combinados, considerando fusão das diferentes componentes de um mesmo descritor. As abordagens de fusão para o descritor CZTZLMP, também proposto, apresentaram melhor resultado que a maioria dos demais descritores considerados, apresentando menor Acurácia somente para o descritor

Tabela 19: Ganho (%) com a aplicação da abordagem de fusão proposta CMedianLMP_ $S_{3,24}$ || $+$ CMedianLMP_ $M_{2,16}$ || $+$ CMedianLMP_C sobre o banco de mamografia digital

Descritor	B	S	Acc	E
CMedianLMP_ $S_{3,24}$ $+$ CMedianLMP_ $M_{2,16}$ $+$ CMedianLMP_C*	62	86.9	81.0	75.0
CMLMP_ $S_{3,24}$ CMLMP_C*	272	$\uparrow 1,10$	$\uparrow 1,6$	$\uparrow 1,9$
CZTZLMP_ $S_{2,16}$ CZTZLMP_ $M_{2,16}$ CZTZLMP_C*	5760	$\uparrow 4,10$	$\uparrow 3,0$	$\uparrow 1,8$
CLMP_ $M_{2,16}$ CLMP_ $M_{3,24}$	80	$\uparrow 8,00$	$\uparrow 6,0$	$\uparrow 3,9$
CLBP_ $M_{1,8}^{riu2}$ CLBP_ $M_{2,16}^{riu2}$	180	$\uparrow 12,1$	$\uparrow 6,5$	$\uparrow 0,7$
WLD $_{\epsilon}$ WLD $_{\Phi}$	64	$\uparrow 13,4$	$\uparrow 3,9$	$\downarrow 5,7$
CMedianLMP_C*	19	$\uparrow 6,30$	$\uparrow 2,2$	$\downarrow 2,1$
GLCM	16	$\uparrow 5,50$	$\uparrow 4,1$	$\uparrow 2,6$

*Métodos propostos.

CMedianLMP_C. A abordagem de fusão CMedianLMP_ $S_{3,24}$ || $+$ CMedianLMP_ $M_{2,16}$ || $+$ CMedianLMP_C apresentou o melhor desempenho dentre os demais descritores, individuais e combinados, cujo ganho de aplicação é discriminado na Tabela 19. Quando comparada à melhor abordagem de fusão dos descritores CLMP, CLBP e WLD, apresenta acréscimo de 8,0%, 12,1% e 13,4% em Sensibilidade, respectivamente, e de 6,0%, 6,5% e 3,9% em Acurácia, respectivamente. Em relação aos descritores individuais de maior desempenho, CMedianLMP_C e GLCM, apresenta incremento de 6,3% e 5,5% em S, respectivamente, e de 2,2% e 4,1% em Acc, respectivamente.

6.3.4 Fusão de Diferentes Descritores de Textura

Apresentam-se, na Tabela 20, os resultados obtidos com a fusão entre diferentes descritores de textura combinados dois a dois, três a três e quatro a quatro. Nota-se que as abordagens de fusão obtidas com a aplicação dos operadores de fusão propostos, Soma^M e Gram^M, permitiram os melhores resultados para a maioria dos casos. Ademais, é notável a contribuição que o descritor GLCM aportou ao sistema de reconhecimento de distorção arquitetural quando combinado com os descritores baseados em fusão de características, integrando vinte e nove das trinta abordagens de fusão de vetores de características apresentadas na tabela. Outra constatação relevante refere-se à dimensão dos vetores de características retornadas pelo operador de fusão Combinação Conjunta, sendo que uma das abordagens para descritores tomados dois a dois, CMedianLMP_ $S_{2,16}$ | CLMP_ $S_{2,16}$, apresenta $B = 2.656$, e as três abordagens para descritores tomados três a três apresentam B igual a 18.848, 9.424 e 3.968, o que pode inviabilizar a implementação do respectivo operador.

O maior resultado de Sensibilidade, $S = 88,4\%$, foi obtido com a aplicação do operador Soma^M sobre os descritores GLCM, CMedianLMP_ $S_{3,24}$, CLMP_ $M_{2,16}$ e CMLMP_C, sendo a maior Sensibilidade global (maior desempenho dentre todos os obtidos para os

Tabela 20: Sensibilidade(S), Especificidade(E) e Acurácia(Acc) obtida com a aplicação dos operadores de fusão Concatenação ($\oplus$), Combinação Conjunta ($|$), Gram^M ($\parallel_G$) e Soma^M ($\parallel_+$) sobre os diferentes descritores na camada de características (valores em %)

Descritor	B	S	Acc	E
CMedianLMP_S _{2,16} $\oplus$ GLCM	99	87,5	80,6	73,8
CMLMP_S _{2,16} $\oplus$ GLCM	42	87,3	81,3	75,2
CMedianLMP_S _{3,24} $\oplus$ GLCM	78	87,0	80,9	74,8
CMLMP_S _{3,24} CMedianLMP_C	646	85,6	79,4	73,2
CMedianLMP_S _{3,24} CMLMP_C	496	85,3	79,1	72,8
CMedianLMP_S _{2,16} CLMP_S _{2,16}	2.656	84,8	73,8	62,7
CMLMP_S _{2,16} $\parallel_G$ GLCM	351	88,0	81,6	75,2
CMLMP_S _{3,24} $\parallel_G$ GLCM	595	86,9	81,3	75,7
CMLMP_S _{1,8} $\parallel_G$ GLCM	861	86,7	80,9	75,0
CMedianLMP_S _{3,24} $\parallel_+$ GLCM	62	87,3	80,7	74,1
CMedianLMP_S _{2,16} $\parallel_+$ GLCM	83	87,0	80,7	74,3
CMLMP_S _{3,24} $\parallel_+$ GLCM	34	86,7	81,0	75,2
CMLMP_S _{3,24} $\oplus$ GLCM $\oplus$ CMLMP_C	58	87,0	80,7	74,4
CMedianLMP_S _{3,24} $\oplus$ GLCM $\oplus$ CMedianLMP_C	97	87,0	81,0	75,0
CMLMP_S _{2,16} $\oplus$ GLCM $\oplus$ CMLMP_C	50	86,8	81,5	76,3
CMedianLMP_S _{3,24} ZTZLBP _{1,8} CMedianLMP_C	18.848	87,4	78,8	70,2
CMedianLMP_S _{3,24} WLD _{ϵ} CMedianLMP_C	9.424	87,4	79,1	70,9
CMedianLMP_S _{3,24} WLD _{Φ} CMLMP_C	3.968	86,8	79,3	71,7
GLCM $\oplus$ (CMLMP_S _{3,24} CMedianLMP_M _{1,8})	254	86,4	80,7	75,1
GLCM $\oplus$ (CMLMP_S _{3,24} CMLMP_C)	288	85,8	79,4	73,1
GLCM $\oplus$ (CMLMP_S _{3,24} CMedianLMP_C)	662	85,6	79,4	73,2
CMLMP_S _{2,16} $\parallel_G$ GLCM $\parallel_G$ CMLMP_C	351	87,7	81,4	75,0
CMLMP_S _{3,24} $\parallel_G$ GLCM $\parallel_G$ CMLMP_C	595	87,0	81,2	75,3
CMedianLMP_S _{1,8} $\parallel_G$ GLCM $\parallel_G$ CMLMP_C	820	86,3	81,1	75,9
CMedianLMP_S _{3,24} $\parallel_+$ GLCM $\parallel_+$ CMedianLMP_C	62	87,2	80,3	73,4
CMLMP_S _{3,24} $\parallel_+$ GLCM $\parallel_+$ CMLMP_C	34	86,7	80,3	74,0
CMedianLMP_S _{3,24} $\parallel_+$ CLMP_M _{2,16} $\parallel_+$ CMedianLMP_C	62	86,7	80,6	74,5
GLCM $\oplus$ CMLMP_S _{3,24} $\oplus$ CMedianLMP_M _{2,16} $\oplus$ CMLMP_C	96	87,4	80,6	73,8
GLCM $\oplus$ CMedianLMP_S _{3,24} $\oplus$ WLD _{ϵ} $\oplus$ CMedianLMP_C	105	87,4	81,4	75,3
GLCM $\oplus$ CMedianLMP_S _{3,24} $\oplus$ CLMP_M _{2,16} $\oplus$ CMedianLMP_C	117	87,3	81,1	74,9
GLCM $\parallel_G$ CMLMP_S _{2,16} $\parallel_G$ CLBP_M _{3,24} ^{riu2} $\parallel_G$ CMLMP_C	351	88,0	81,4	74,8
GLCM $\parallel_G$ CMLMP_S _{2,16} $\parallel_G$ WLD _{Φ} $\parallel_G$ CMLMP_C	351	87,7	81,7	75,7
GLCM $\parallel_G$ CMLMP_S _{2,16} $\parallel_G$ ZTZLBP _{1,8} $\parallel_G$ CMLMP_C	351	87,7	81,5	75,2
GLCM $\parallel_+$ CMedianLMP_S _{3,24} $\parallel_+$ CLMP_M _{2,16} $\parallel_+$ CMLMP_C	62	88,4	81,1	73,8
GLCM $\parallel_+$ CMedianLMP_S _{3,24} $\parallel_+$ CLMP_M _{2,16} $\parallel_+$ CMedianLMP_C	62	88,1	80,7	73,4
GLCM $\parallel_+$ CMedianLMP_S _{3,24} $\parallel_+$ CMLMP_M _{2,16} $\parallel_+$ CMedianLMP_C	62	87,5	80,3	73,2

As células destacadas em cinza escuro e em cinza claro representam o melhor e o segundo melhor resultado obtido.

descritores individuais e para as abordagens de fusão), representando um incremento de 7,8% e de 7,0% no desempenho dos descritores individuais CMedianLMP_C e GLCM, respectivamente, tal como discriminado na Tabela 21. A respectiva abordagem de fusão resultou em Acurácia igual a 81,1%, representando um incremento de 2,3% em relação ao CMedianLMP_C, melhor desempenho individual, e de 4,2% em relação ao GLCM; e Especificidade igual a 73,8%, com decréscimo de 3,3% quando comparada ao desempenho do descritor individual CMedianLMP_C, mas incremento de 1,4% em relação ao GLCM.

O operador Gram^M aplicado sobre os descritores GLCM, CMLMP_S_{2,16}, WLD _{Φ} e CMLMP_C, resultou na maior Acurácia global, Acc = 81,7%, incremento de 2,9% comparado ao melhor desempenho individual, apresentando Sensibilidade igual a 87,7% e Especificidade igual a 75,7%. Ao considerar fusão entre dois descritores, a abordagem CMLMP_S_{2,16} $\parallel_G$ GLCM resultou em Acc = 81,6%, com S = 88,0% e E = 75,2%, resultados

Tabela 21: Ganho (%) com a aplicação da abordagem de fusão proposta GLCM $\parallel_+$ CMedianLMP_S_{3,24} $\parallel_+$ CLMP_M_{2,16} $\parallel_+$ CMLMP_C sobre o banco de mamografia digital

Descritor	B	S	Acc	E
GLCM $\parallel_+$ CMedianLMP_S _{3,24} $\parallel_+$ CLMP_M _{2,16} $\parallel_+$ CMLMP_C*	62	88,4	81,1	73,8
CMLMP_S _{2,16} $\parallel_G$ GLCM	351	↑0,4	↓0,5	↓1,4
CMedianLMP_S _{3,24} $\parallel_+$ CMedianLMP_M _{2,16} $\parallel_+$ $\parallel_+$ CMedianLMP_C*	62	↑1,5	↑0,1	↓1,2
CMedianLMP_C*	19	↑7,8	↑2,3	↓3,3
GLCM	16	↑7,0	↑4,2	↑1,4

*Métodos propostos.

bem semelhantes à primeira abordagem, mas com a vantagem de combinar apenas dois descritores.

Nenhuma das abordagem de fusão apresentou Especificidade superior ao melhor desempenho individual. O maior valor de Especificidade foi obtido com a Concatenação dos vetores de características dos descritores CMLMP_S_{2,16}, GLCM e CMLMP_C, sendo $E = 76,3\%$, com $S = 86,8\%$ e $Acc = 81,5\%$. Quando comparado à maior Sensibilidade global, representa decréscimo de $1,6\%$ em S, incremento de $0,4\%$ em Acc e de $2,5\%$ em E, ao passo que, comparado com o melhor desempenho de Especificidade obtido pela aplicação individual de descritores, representa um incremento de $13,8\%$ na Sensibilidade, $5,1\%$ na Acurácia e um decréscimo de $3,6\%$ na Especificidade.

6.3.5 Análise do Custo Computacional

O tempo médio de processamento para os descritores individuais é apresentado na Figura 32, em que t_{desc} , medido em ms, refere-se ao tempo de geração dos vetores de características e t_{dist} , em μs , é o tempo para se calcular a distância euclidiana entre os vetores de características de duas amostras de mamografia digital. O tempo de geração dos vetores de características é função da complexidade das regras de extração das características e da quantidade de *pixels* da vizinhança. Nota-se que os menores valores de t_{desc} foram obtidos com a implementação da componente Centro para os diferentes descritores, CLBP_C (0,9 ms), CLMP_C, CMLMP_C (1,3 ms), CZTZLMP_C e CMedianLMP_C (3,6 ms), cuja regra de extração de características é mais simples que a das componentes Sinal e Magnitude dos respectivos descritores, e que os maiores valores foram retornados pelos descritores GLCM (78,1 ms), WLD (99,8 ms) e Gabor, sendo este último o de maior custo computacional dentre todos os descritores individuais, $t_{desc}(\text{Gabor})$ igual a 51,3 segundos.

A geração das diferentes componentes do descritor CLBP se deu em aproximadamente metade do tempo transcorrido para as respectivas componentes dos descritores CLMP, CMLMP, CMedianLMP e CZTZLMP, contudo, apresentam baixo desempenho: os valores de Sensibilidade e de Acurácia retornados pelo CLBP são menores que os retornados pelos demais descritores. Ao comparar o tempo de geração dos vetores de

(a)

(b)

Figura 32: Tempo médio de processamento dos descritores individuais (a) para geração dos vetores de características, em ms, e (b) para cálculo da distância euclidiana entre os vetores de características de duas amostras de mamografia digital, em μs .

características destes quatro descritores, percebe-se que as componentes Sinal e Magnitude de mesma configuração, igual número de *pixels* na vizinhança, apresentaram valores semelhantes: $t_{desc}(\text{CMLMP_S}_{1,8}) \approx t_{desc}(\text{CMLMP_M}_{1,8})$, $t_{desc}(\text{CMedianLMP_S}_{2,16}) \approx t_{desc}(\text{CMedianLMP_M}_{2,16})$ e assim consecutivamente; ademais, observa-se similaridade entre as respectivas componentes de diferentes descritores e mesma configuração: $t_{desc}(\text{CMLMP_S}_{1,8}) \approx t_{desc}(\text{CMedianLMP_S}_{1,8})$ e $t_{desc}(\text{CMLMP_S}_{1,8}) \approx t_{desc}(\text{CLMP_S}_{1,8})$, $t_{desc}(\text{CMLMP_M}_{2,16}) \approx t_{desc}(\text{CMedianLMP_M}_{2,16})$ e

$t_{desc}(\text{CMLMP_M}_{2,16}) \approx t_{desc}(\text{CLMP_M}_{2,16})$, e assim consecutivamente. Nota-se ainda que o tempo de geração de vetores de características aumenta gradativamente em função do aumento do número de *pixels* na vizinhança.

O tempo para se calcular a distância euclidiana entre os vetores de características, t_{dist} , é função do número de elementos dos respectivos vetores e independe de t_{desc} . O menor valor médio de t_{dist} , $0,86 \mu s$, foi obtido para os descritores CLBP_C e CLMP_S_{1,8}, ambos apresentando comprimento B igual a 2. O descritor CMedianLMP_S_{2,16}, cujo vetor de características possui 83 elementos, apresentou, dentre os descritores individuais, o maior valor médio de t_{dist} , igual a $1,49 \mu s$.

Nota-se que o tempo de geração dos vetores de características é significativamente maior que o tempo decorrido durante o cálculo da distância euclidiana entre dois vetores de características. Desse modo, o custo computacional de uma determinada abordagem de fusão é mais sensível ao descritor de textura a ser combinado que ao número de elementos do vetor de características do respectivo descritor. Outro fator que pode impactar em aumento do custo computacional de uma abordagem de fusão é o operador de fusão utilizado. Apresenta-se, na Figura 33, o tempo médio de execução dos operadores Concatenação, Soma^M, Gram^M e Combinação Conjunta ao se combinar descritores de textura tomados dois a dois, três a três e quatro a quatro.

O operador Concatenação apresentou o menor tempo médio de execução, t_{fus} igual a $1,1 \mu s$ ao combinar dois descritores; o segundo menor tempo médio de execução foi em decorrência da aplicação do operador Soma^M ($2,3 \mu s$); e o terceiro menor, do operador Gram^M ($20,5 \mu s$). Nota-se que o aumento do número de descritores combinados tem pouca influência sobre o tempo médio de execução destes três operadores, apresentando maior

Figura 33: Tempo médio de execução dos operadores de fusão.

O tempo médio de processamento das abordagens de fusão que apresentaram os melhores desempenhos de Sensibilidade e de Acurácia é mostrado na Figura 34. A combinação dos descritores GLCM, CMedianLMP_ $S_{3,24}$, CLMP_ $M_{2,16}$ e CMLMP_C pelo operador Soma^M, que resultou no maior desempenho global ($S = 88,4\%$), apresentou t_{desc} igual a 161,9 ms, o terceiro maior tempo médio de geração dentre as abordagens consideradas, e t_{dist} igual a 1,61 μs , um dos menores tempos de cálculo da distância euclidiana dentre as abordagens consideradas. O valor retornado de t_{desc} é consequência, principalmente, da geração do vetor de características do descritor GLCM (78,1 ms), o terceiro mais dispendioso dentre os descritores individuais, e do descritor CMedianLMP_ $S_{3,24}$ (51,6 ms), enquanto que o baixo valor de t_{dist} é resultado da propriedade do operador de fusão Soma^M de gerar vetores de características combinados de menor comprimento dentre as abordagens equivalentes geradas pelos demais operadores de fusão analisados.

Com a combinação de apenas dois descritores, a abordagem de fusão CMLMP_ $S_{2,16}$ || $_G$ GLCM que possui desempenho próximo ao melhor resultado global ($S = 88,0\%$), apresentou t_{desc} igual a 105,5 ms, 56,4 ms menor que a solução de maior desempenho, e t_{dist} igual a 3,73 μs .

6.4 Considerações Finais

A utilização de descritores de textura em detecção e reconhecimento de distorção arquitetural mamária em mamografias digitais é recorrente e tem como finalidade contribuir para o desenvolvimento de sistemas de auxílio ao diagnóstico prematuro de câncer de mama. No entanto, o desempenho das técnicas utilizadas ainda é baixo, principalmente quando comparado a soluções implementadas para detecção de outros indicadores de câncer, mas em estágio mais evoluído da doença, como microcalcificações e nódulos.

Com o objetivo de contribuir para a melhoria do desempenho de sistemas de reconhecimento de distorção arquitetural em mamografias digitais, apresentaram-se, neste capítulo, a metodologia e os resultados da aplicação de três descritores de textura baseados em fusão, propostos neste trabalho, CMLMP, CMedianLMP e CZTZLMP, assim como de quatro operadores de fusão sobre nove diferentes descritores de textura, sendo que dois dos operadores aplicados, Soma^M e Gram^M, foram propostos neste trabalho.

A aplicação dos descritores propostos resultou nos maiores valores de Sensibilidade e de Acurácia dentre os obtidos pelos descritores individuais e pela fusão das diferentes componentes de um mesmo descritor. Ao se considerar fusão entre diferentes descritores, os melhores resultados foram obtidos com a aplicação dos operadores de fusão propostos.

Esses resultados condizem com os obtidos na aplicação de fusão de descritores a registro de imagens. Discorre-se, no capítulo a seguir, sobre as conclusões referentes às duas implementações.

7 CONCLUSÕES

A utilização de descritores de textura é bastante recorrente em aplicações de visão computacional e de reconhecimento de padrões. No entanto, dificilmente encontra-se um descritor que apresente o melhor desempenho para diversos conjuntos de imagens, assim como para diversas aplicações, sendo a utilização destes limitada a aplicações específicas e a um conjunto restrito de imagens. Neste sentido, foi proposta nesta tese a utilização de operadores de fusão sobre diversos descritores de texturas que, aplicados a registro de imagens com diferentes tipos de transformações, permitiram a obtenção de desempenho superior ao dos descritores individuais.

Foram também propostos descritores baseados em fusão, os quais apresentaram melhor desempenho ao serem aplicados a reconhecimento de distorção arquitetural em mamografia digital quando comparados com diversos descritores frequentemente utilizados neste tipo de aplicação. Como diferentes descritores apresentam informações complementares entre si, propôs-se ainda operadores de fusão que, ao combinar os diferentes descritores, permitiram a obtenção dos melhores resultados globais.

A extração de informações complementares de uma textura por diferentes componentes, metodologia empregada na formulação do descritor CLBP e que motivou a proposta de outros descritores, permite uma melhor representação das imagens de textura, contribuindo com o aperfeiçoamento da resposta de diversas aplicações de reconhecimento de imagens. Os descritores propostos foram baseados nesta metodologia associada a métodos de extração de características mais robustos a ruído, como o empregado na formulação dos descritores M-LMP e Median-LMP, o que lhes conferiu melhor capacidade em discriminar as texturas.

Pelo fato das texturas serem caracterizadas por possuírem informação de distribuição da intensidade dos níveis de cinza da imagem, assim como informação do valor das intensidades dos níveis de cinza, é importante que o descritor represente adequadamente as diferenças de intensidade dos *pixels* em uma vizinhança na imagem. O descritor LMP, e os diversos descritores nele baseados, realizam esta melhor representação por empregarem a função sigmóide no mapeamento das diferenças de nível de cinza. Os descritores propostos são baseados na metodologia do LMP, mais precisamente do CLMP, apresentando melhor capacidade de representar as nuances nos *pixels* de uma mesma vizinhança.

O ganho de desempenho pela implementação da metodologia do LMP é evidenciada quando da aplicação dos descritores CLBP_C e CMLMP_C sobre amostras de mamografias. O CLBP_C aplica uma binarização sobre a imagem original com limiar igual ao valor médio das intensidades de cinza de toda a imagem, obtendo Sensibilidade igual a

50,9%, Acurácia igual a 54,3% e Especificidade igual a 57,6%. O descritor CMLMP_C proposto realiza uma transformação de intensidade de nível de cinza aplicada com função de mapeamento sigmóide, seguida por uma quantização em 3 bits, com valores de códigos compreendidos no intervalo $[0,7]$, obtendo Sensibilidade de 75,8%, Acurácia de 77,4% e Especificidade igual a 79,1%, representando um acréscimo de mais de 20% no desempenho do sistema de reconhecimento.

A maioria dos trabalhos publicados na literatura, mesmo os mais recentes, que tentam prover soluções para o sistema de detecção e de reconhecimento de distorção arquitetural, utiliza banco de mamografias digitalizadas, ainda que a mamografia digital seja o exame de triagem mais utilizado para auxiliar no diagnóstico de doenças da mama. Os trabalhos que utilizam mamografia digital, dispõem de poucas imagens de distorção arquitetural. O banco de dados utilizado nesta tese consiste em 176 mamografias digitais, das quais 88 possuem distorção arquitetural. Todas as mamografias têm laudo médico e a ocorrência de distorção arquitetural, com respectiva localização, foi identificada por um radiologista experiente. Sendo assim, os resultados presentes nesta tese contribuem com a literatura, podendo servir como referência para outros trabalhos.

Os resultados globais obtidos, quando da aplicação dos descritores e dos operadores de fusão propostos no reconhecimento de distorção arquitetural em mamografias digitais, destacam-se dos demais métodos implementados. Quando comparado aos resultados publicados na literatura, possuem maior desempenho que técnicas de aprendizagem profunda, a exemplo do Autoencoder. Esses resultados são indício da relevância dos métodos propostos.

Por mais que não se possa garantir quando a fusão de informações proverá melhores resultados que os obtidos pelas fontes individuais, demonstrou-se nesta tese que a fusão dos descritores de textura apresenta desempenho superior ao melhor resultado dos descritores individuais nas duas aplicações consideradas. Não houve um descritor individual que apresentasse o melhor desempenho dentre os demais descritores para todas as imagens quando implementados em um sistema de registro de imagens. Com a aplicação dos operadores de fusão Soma e Concatenação sobre os vetores de características dos descritores SIFT, CS-LBP, CS-LMP, M-LMP, Z+TZLBP e $Z \oplus TZLBP$, obteve-se a melhor AUC global para os dois conjuntos de imagens utilizados. Resultado semelhante foi obtido ao aplicar os operadores de fusão de decisão Mínimo, Adição, Produto e Combinação às informações de decisão.

Observou-se que o desempenho da abordagem de fusão depende dos descritores combinados, assim como dos operadores de fusão utilizados. Ao combinar as diferentes componentes do descritor CMLMP, o melhor resultado foi obtido com a Combinação Conjunta das componentes CMLMP_ $S_{3,24}$ e CMLMP_C ($S = 85,8\%$, $Acc = 79,4\%$ e $E = 78,4\%$). Para o CMedianLMP, o melhor resultado foi obtido ao utilizar o operador Soma^M sobre

as componentes CMedianLMP_ $S_{3,24}$, CMedianLMP_ $M_{2,16}$ e CMedianLMP_ C , obtendo $S = 86,9\%$, $Acc = 81,0\%$ e $E = 75,0\%$. Enquanto que o melhor resultado quando da fusão sobre o descritor CZTZLMP foi obtido ao realizar a Combinação Conjunta das componentes CZTZLMP_ $S_{2,16}$, CZTZLMP_ $M_{2,16}$ e CZTZLMP_ C ($S = 82,8\%$, $Acc = 78,0\%$ e $E = 73,2\%$).

O tempo de processamento é uma limitante em diversas aplicações de processamento de imagens. Para reduzir o custo computacional da aplicação, é desejável que os descritores gerem vetores de características de baixa dimensão. O operador Combinação Conjunta resulta em vetores de características de dimensão muito elevada. Desde que os descritores a serem combinados apresentem baixa dimensão, o operador Concatenação resultará em vetores de baixa dimensão, no entanto, não apresenta os melhores desempenhos. O operador de fusão Soma mantém o número de elementos dos vetores de características a serem combinados, no entanto, tem a restrição de combinar unicamente vetores de características de mesmo número de elementos. Uma das contribuições desse trabalho foi a proposta de dois operadores de fusão de características, Soma e Gram no espaço de fusão. A aplicação do operador Soma^M permite combinar vetores de características de diferentes dimensões, e o número de elementos do vetor de características resultante é igual ao número de elementos do descritor individual de maior dimensão. O aumento de dimensão causado pelo operador Gram^M é limitado, sendo função do número de elementos do descritor individual de maior dimensão, não aumentando com o aumento do número de descritores a serem combinados.

Os melhores desempenhos de Sensibilidade e de Acurácia para a aplicação de reconhecimento de distorção arquitetural foram obtidos com a aplicação dos descritores propostos e dos operadores de fusão propostos. Ao se considerar a aplicação de operadores de fusão sobre as componentes de um mesmo descritor, o melhor resultado foi obtido com a aplicação do operador de fusão Soma^M sobre as componentes CMedianLMP_ $S_{3,24}$, CMedianLMP_ $M_{2,16}$ e CMedianLMP_ C . Ao realizar a fusão entre diferentes descritores, o melhor resultado foi obtido com a aplicação do operador Soma^M sobre os descritores GLCM, CMedianLMP_ $S_{3,24}$, CLMP_ $M_{2,16}$ e CMLMP_ C , sendo obtido Sensibilidade igual a 88,4%, Acurácia igual a 81,1% e Especificidade igual a 73,8%.

De acordo com os resultados obtidos, conclui-se que a fusão de descritores de textura contribui com o aumento do desempenho de sistemas de registro de imagens e de reconhecimento de distorção arquitetural em mamografia digital quando comparada à aplicação de descritores de textura individuais. Os descritores propostos apresentaram desempenho superior ao dos demais descritores com os quais foram comparados, o que incentiva a implementação dos mesmos em outras aplicações. Os operadores de fusão propostos permitiram resultados melhores que os obtidos com a aplicação dos demais operadores, gerando vetores de características de menor dimensão. Sendo assim, os objetivos

propostos nesta tese de doutorado foram atingidos.

7.1 Principais Contribuições

As principais contribuições desta tese podem ser enumeradas como a seguir:

- Proposta de dois novos operadores de fusão de vetores de características, Soma no espaço de fusão ($\text{Soma}^{\mathbb{M}}$) e Gram no espaço de fusão ($\text{Gram}^{\mathbb{M}}$);
- Proposta de três novos descritores de textura baseados em fusão de características, *Completed Mean Local Mapped Pattern* (CMLMP), *Completed Median Local Mapped Pattern* (CMedianLMP) e *Completed Z with Tilted Z Local Mapped Pattern* (CZTZLMP);
- Implementação de uma metodologia de fusão de diferentes descritores de textura aplicados a reconhecimento de distorção arquitetural em mamografias digitais;
- Implementação de uma metodologia de fusão de diferentes descritores de textura aplicados a registro de imagens.

7.2 Sugestões para Trabalhos Futuros

De acordo com o trabalho realizado e com os resultados obtidos, sugerem-se as seguintes propostas de trabalhos futuros:

- A aplicação de operadores de fusão de características e de decisões sobre descritores de textura contribuiu com o aumento do desempenho de sistemas de registro de imagens quando comparada à aplicação de descritores de textura individuais. Em reconhecimento de distorção arquitetural em mamografias digitais, foram implementados apenas operadores de fusão de características. Sugere-se investigar o desempenho de operadores de fusão de decisão aplicados a descritores de textura para reconhecimento de distorção arquitetural em mamografias digitais;
- Sugere-se ainda a aplicação dos operadores de fusão propostos para combinar descritores de textura convencionais com Redes Neurais Convolucionais (CNN);
- Avaliar o desempenho da metodologia de fusão aplicada, substituindo o classificador K-NN por outros classificadores, tal como SVM ou Redes Neurais Perceptron Multicamadas;
- Neste trabalho, os parâmetros β e B foram otimizados conjuntamente, mas para cada descritor individualmente. Sugere-se realizar a otimização conjunta dos parâmetros considerando a abordagem de fusão.

- Investigar o desempenho das abordagens de fusão propostas aplicadas a outros problemas de visão computacional, tal como reconhecimento da região ocular.

7.3 Publicações Oriundas do Trabalho

- [1] Pereira Jr, O.; Ferraz, C. T.; Gonzaga, A. Image correspondence using a fusion of local region descriptors. In: **XIV Workshop de Visão Computacional (WVC 2018)**, 2018.
- [2] Pereira Jr, O; Oliveira H. C. R.; Ferraz, C. T.; Saito, J. H.; Vieira, M. A. C.; Gonzaga, A. A novel fusion-based texture descriptor to improve the detection of architectural distortion in digital mammography. **Journal of Digital Imaging**. 2019 (Submetido)

7.3.1 Premiações Oriundas do Trabalho

- Prêmio de 3º lugar concedido ao trabalho “Image correspondence using a fusion of local region descriptors”, apresentado em seção técnica no XIV Workshop de Visão Computacional - WVC 2018.

REFERÊNCIAS

- ABDEL-NASSER, M.; MELENDEZ, J.; MORENO, A.; DOMENEC, P. The impact of pixel resolution, integration scale, preprocessing, and feature normalization on texture analysis for mass classification in mammograms. **International Journal of Optics**, v. 2016, 2016.
- AHONEN, T.; HADID, A.; PIETIKÄINEN, M. Face description with local binary patterns: Application to face recognition. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 28, n. 12, p. 2037 – 2041, December 2006.
- ALITAPPEH, R. J.; MAHMOUDI, F. MGS-SIFT: A new illumination invariant feature based on sift descriptor. **International Journal of Computer Theory and Engineering**, v. 5, n. 1, p. 99 – 103, February 2013.
- BAEZA-YATES, R. A.; RIBEIRO-NETO, B. **Modern Information Retrieval**. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1999.
- BAHL, M.; BAKER, J. A.; KINSEY, E. N.; GHATE, S. V. Architectural Distortion on Mammography: Correlation With Pathologic Outcomes and Predictors of Malignancy. **American Journal of Roentgenology**, American Roentgen Ray Society, v. 205, n. 6, p. 1339–1345, 2015.
- BAUMBERG, A. Reliable feature matching across widely separated views. In: **Proceedings of IEEE Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2000. p. 774 – 781.
- BERE, S. S. Duplicate Video and Object Detection by Video Key Frame Using F-SIFT. In: **International Conference on Computing Communication Control and Automation (ICCUBEA)**. [S.l.: s.n.], 2018.
- BOONTHEP, N.; CHAMNONGTHAI, K.; PHENSADSAENG, P. A FPGA-Based SIFT Architecture for Motion Estimation in Video Coding. In: **2018 Global Wireless Summit (GWS)**. [S.l.: s.n.], 2018. p. 383 – 388.
- BOSTRÖM, H.; ANDLER, S. F.; BROHEDE, M.; JOHANSSON, R.; KARLSSON, E.; LAERE, J. V.; NIKLASSON, L.; NILSSON, M.; PERSSON, A.; ZIEMKE, T. **On the definition of information fusion as a field of research**. [S.l.], 2007.
- CAMPOS, F. M.; CORREIA, L.; CALADO, J. M. F. Robot visual localization through local feature fusion: An evaluation of multiple classifiers combination approaches. **Journal of Intelligent and Robotic Systems**, v. 77, n. 2, p. 377 – 390, February 2015.
- CASTANEDO, F. A review of data fusion techniques. **The Scientific World Journal**, v. 2013, p. 19, 2013.
- CASTI, P.; MENCATTINI, A.; SALMERI, M.; ANCONA, A.; MANGERI, F.; PEPE, M.; RANGAYYAN, R. Contour-independent detection and classification of mammographic lesions. **Biomedical Signal Processing and Control**, v. 25, p. 165–177, 2016.

- CHEN, C.; ZHANG, B.; SU, H.; LI, W.; WANG, L. Land-use scene classification using multi-scale completed local binary patterns. **Signal, Image and Video Processing (SIVIP)**, v. 10, n. 4, p. 745 – 752, 2016.
- CHEN, J.; SHAN, S.; HE, C.; ZHAO, G.; PIETIKÄINEN, M.; CHEN, X.; GAO, W. WLD: A robust local image descriptor. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 32, n. 9, p. 1705 – 1720, 2009.
- DASARATHY, B. V. Sensor fusion potential exploitation-innovative architectures and illustrative applications. **Proceedings of the IEEE**, v. 85, n. 1, p. 24–38, Jan 1997.
- DAUGMAN, J. How iris recognition works. **IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY**, v. 14, n. 1, p. 21 – 30, January 2004.
- DEAN, J. C.; ILVENTO, C. C. Improved Cancer Detection Using Computer-Aided Detection with Diagnostic and Screening Mammography - prospective study of 104 cancers. **American Journal of Roentgenology**, v. 187, p. 20–28, 2006.
- DELLINGER, F.; DELON, J.; GOUSSEAU, Y.; MICHEL, J.; TUPIN, F. SAR-SIFT: A SIFT-Like Algorithm for SAR Images. **IEEE Transactions on Geoscience and Remote Sensing**, v. 53, n. 1, p. 453 – 466, January 2015.
- DONG, J.; ZHUANG, D.; HUANG, Y.; FU, J. Advances in multi-sensor data fusion: Algorithms and applications. **Sensors**, v. 9, p. 7771 – 7784, September 2009.
- D’ORSI, C. J. (Ed.). **ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System**. [S.l.]: American College of Radiology, 2013.
- DURRANT-WHYTE, H.; HENDERSON, T. C. Multisensor data fusion. In: SICILIANO, B.; KHATIB, O. (Ed.). **Springer Handbook of Robotics**. [S.l.]: Springer, 2008. cap. 25.
- DURRANT-WHYTE, H. F. Sensor models and multisensor integration. **Int. J. Rob. Res.**, v. 7, n. 6, p. 97 – 113, Dec. 1988.
- ELMORE, J. G.; NAKANO, C. Y.; KOEPESELL, T. D.; DESNICK, L. M.; D’ORSI, C. J.; RANSOHOFF, D. F. International variation in screening mammography interpretations in community-based programs. **J. Natl. Cancer Inst.**, v. 95, n. 18, p. 1384–1393, 2003.
- FERRAZ, C. T.; PEREIRA-JR, O.; GONZAGA, A. A new region of interest descriptor for image correspondence. In: **IX Workshop de Visão Computacional (WVC 2013)**. [S.l.: s.n.], 2013.
- _____. Feature description based on center-symmetric local mapped patterns. In: **Proceedings of the 29th Annual ACM Symposium on Applied Computing (SAC’14)**. [S.l.: s.n.], 2014. p. 39 – 44.
- _____. Feature description based on mean local mapped pattern. In: **X Workshop de Visão Computacional (WVC 2014)**. [S.l.: s.n.], 2014. p. 32 – 37.
- FERRAZ, C. T.; PEREIRA-JR, O.; ROSA, M. V.; GONZAGA, A. Object recognition based on bag of features and a new local pattern descriptor. **International Journal of Pattern Recognition and Artificial Intelligence**, v. 28, n. 8, p. 14550101 – 145501032, 2014.

- GATYS, L. A.; ECKER, A. S.; BETHGE, M. A neural algorithm of artistic style. **arXiv**, August 2015.
- GAUR, S.; DIALANI, V.; SLANETZ, P. J.; EISENBERG, R. L. Architectural Distortion of the Breast. **American Journal of Roentgenology**, v. 201, p. 662–670, 2013.
- GHASSEMIAN, H. A review of remote sensing image fusion methods. **Information Fusion**, v. 32, p. 75 – 89, November 2016.
- GIGLI, G.; BOSSÉ, E.; LAMPROPOULOS, G. A. An optimized architecture for classification combining data fusion and data-mining. **Information Fusion**, v. 8, n. 4, p. 366 – 378, 2007.
- GLYNN, C. G.; FARRIA, D. M.; MONSEES, B. S.; SALCMAN, J. T.; WIELE, K. N.; HILDEBOLT, C. F. Effect of transition to digital mammography on clinical outcomes. **Radiology**, v. 260, n. 3, p. 664–670, 2011.
- GOOL, L.; DEWAELE, P.; OOSTERLINCK, A. Texture analysis anno 1983. **COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING**, v. 29, n. 3, p. 336 – 357, March 1985.
- GU, Q.; YANG, J.; CUI, G.; KONG, L.; ZHENG, H.; KLETTE, R. Multi-scale vehicle logo recognition by directional dense SIFT flow parsing. In: **International Conference on Image Processing (ICIP)**. [S.l.: s.n.], 2016. p. 3827 – 3831.
- GUO, Y.; ZHAO, G.; PIETIKÄINEN, M. Discriminative features for texture description. **Pattern Recognition**, v. 45, n. 10, p. 3834 – 3843, February 2012.
- GUO, Z.; ZHANG, L.; ZHANG, D. A completed modeling of local binary pattern operator for texture classification. **IEEE Transactions on Image Processing**, v. 19, n. 6, p. 1657–1663, 2010.
- HADID, A.; PIETIKÄINEN, M.; AHONEN, T. A discriminative feature space for detecting and recognizing faces. In: **IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR'04)**. [S.l.: s.n.], 2004.
- HAFIANE, A.; SEETHARAMAN, G.; ZAVIDOVIQUE, B. Median binary pattern for textures classification. In: **International Conference Image Analysis and Recognition**. [S.l.: s.n.], 2007. p. 387 – 398.
- HALL, D. L.; LLINAS, J. An introduction to multisensor data fusion. In: **Proceedings of the IEEE**. [S.l.: s.n.], 1997. v. 85, n. 1, p. 6 – 23.
- HARALICK, R. M. On a texture-context feature extraction algorithm for remotely sensed imagery. In: **IEEE Conference on Decision and Control**. [S.l.: s.n.], 1971.
- _____. Statistical and structural approaches to texture. **Proceedings of the IEEE**, v. 67, n. 5, p. 786 – 804, May 1979.
- HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, SMC-3, n. 6, p. 610 – 621, November 1973.

- HEATH, M.; BOWYER, K.; KOPANS, D.; MOORE, R.; KEGELMEYER, W. P. The digital database for screening mammography. In: **Proceedings of the 5th international workshop on digital mammography**. [S.l.: s.n.], 2000. p. 212–218.
- HEIKKILÄ, M.; PIETIKÄINEN, M. A texture-based method for modeling the background and detecting moving objects. **IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE**, v. 28, n. 4, p. 657 – 662, April 2006.
- HEIKKILÄ, M.; PIETIKÄINEN, M.; SCHMID, C. Description of interest regions with center-symmetric local binary patterns. In: _____. **Computer Vision, Graphics and Image Processing: Proceedings of the 5th Indian Conference (ICVGIP 2006)**. [S.l.]: Springer Berlin Heidelberg, 2006. p. 58 – 69.
- _____. Description of interest regions with local binary patterns. **Pattern Recognition**, v. 42, p. 425 – 436, 2009.
- HOSSAIN, T.; TENG, S. W.; LU, G.; LACKMANN, M. An enhancement to sift-based techniques for image registration. In: **International Conference on Digital Image Computing: Techniques and Applications**. [S.l.: s.n.], 2010.
- HUANG, X.; WANG, B.; CHENG, M.; HUANG, S.; JU, Y. Image registration and data fusion for different phases of contrast enhanced liver ct data. In: **The 2nd International Conference on Bioinformatics and Biomedical Engineering (ICBBE)**. [S.l.: s.n.], 2008.
- HUANG, Y.; CHAN, K. L.; ZHANG, Z. Texture classification by multi-model feature integration using Bayesian networks. **Pattern Recognition Letters**, v. 24, n. 1 - 3, p. 393 – 401, 2002.
- JASIONOWSKA, M.; PRZELASKOWSKI, A. A two-step method for detection of architectural distortions in mammograms. **Information Technology in Biomedicine**, v. 69, p. 73–84, 2010.
- JUAN, L.; GWUN, O. A comparison of SIFT, PCA-SIFT and SURF. **International Journal of Image Processing**, v. 3, n. 4, p. 143 – 152, August 2009.
- JULESZ, B. Visual pattern discrimination. **IRE Transactions on Information Theory**, v. 8, n. 2, p. 84 – 92, February 1962.
- _____. Experiments in the visual perception of texture. **Scientific American**, v. 232, n. 4, p. 34 – 43, April 1975.
- _____. Textons, the elements of texture perception, and their interactions. **Nature**, v. 290, p. 91 – 97, March 1981.
- KAMRA, A.; JAIN, V. K.; SINGH, S.; MITTAL, S. Characterization of architectural distortion in mammograms based on texture analysis using support vector machine classifier with clinical evaluation. **Journal of Digital Imaging**, v. 29, n. 1, p. 104–114, 2016.
- KARELLAS, A.; VEDANTHAM, S. Breast cancer imaging: a perspective for the next decade. **Medical Physics**, v. 35, n. 11, p. 4878–4897, Nov 2008.

- KASSNER, A.; THORNHILL, R. Texture analysis: A review of neurologic mr imaging applications. **American Journal of Neuroradiology**, v. 31, n. 5, p. 809 – 816, May 2010.
- KE, Y.; SUKTHANKAR, R. PCA-SIFT: A more distinctive representation for local image descriptors. In: **Conference on Computer Vision and Pattern Recognition - CVPR**. [S.l.: s.n.], 2004.
- KHALEGHI, B.; KHAMIS, A.; KARRAY, F. O.; RAZAVI, S. N. Multisensor data fusion: A review of the state-of-the-art. **Information Fusion**, v. 14, n. 1, p. 28 – 44, 2013.
- KIM, B.; YOO, H.; SOHN, K. Exact order based feature descriptor for illumination robust image matching. **Pattern Recognition**, v. 46, p. 3268 – 3278, May 2013.
- KITTLER, J.; HATEF, M.; DUIN, R. P.; MATAS, J. On combining classifiers. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 20, n. 3, p. 226 – 239, March 1998.
- LIU, L.; CHEN, J.; FIEGUTH, P.; ZHAO, G.; CHELLAPPA, R.; PIETIKÄINEN, M. From BoW to CNN: two decades of texture representation for texture classification. **Int J Comput Vis**, v. 127, n. 1, p. 74 – 109, January 2019.
- LIU, L.; FIEGUTH, P.; GUO, Y.; WANG, X.; PIETIKÄINEN, M. Local binary features for texture classification: Taxonomy and experimental study. **Pattern Recognition**, v. 62, p. 135 – 160, 2017.
- LIU, X.; ZHAI, L.; ZHU, T.; LIU, J.; ZHANG, K.; HU, W. Multiple TBSVM-RFE for the detection of architectural distortion in mammographic images. **Multimedia Tools and Applications**, v. 77, n. 12, p. 15773 – 15802, 2018.
- LOWE, D. G. Distinctive image features from scale-invariant keypoints. **International Journal of Computer Vision**, v. 60, n. 2, p. 91 – 110, 2004.
- LUO, R. C.; YIH, C.-C.; SU, K. L. Multisensor fusion and integration: approaches, applications, and future research directions. **IEEE Sensors Journal**, v. 2, n. 2, p. 107 – 119, 2002.
- LUTZ, A.; GRACE, K.; MESSER, N.; EZEKIEL, S.; BLASCH, E.; ALFORD, M.; BUBALO, A.; CORNACCHIA, M. Bandelet transformation based image registration. In: **Applied Imagery Pattern Recognition Workshop (AIPR)**. [S.l.: s.n.], 2015.
- MA, J.; CHAN, J. C.-W.; CANTERS, F. Fully automatic subpixel image registration of multiangle chris/proba data. **IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING**, v. 48, n. 7, p. 2829 – 2839, July 2010.
- MARSHALL, J.; DOUCETTE, P. Image registration and a metric to assess the transformation function. In: **Applied Imagery Pattern Recognition Workshop (AIPR)**. [S.l.: s.n.], 2011.
- MARY, N. A. B.; DEJEY, D. Classification of coral reef submarine images and videos using a novel Z and Tilted Z local binary pattern ($Z \oplus TZLBP$). **Wireless Pers Commun**, v. 98, n. 3, p. 2427 – 2459, February 2018.

MIKOLAJCZYK, K.; SCHMID, C. An affine invariant interest point detector. In: **Proceedings of the 7th European Conference on Computer Vision-Part I**. [S.l.: s.n.], 2002. p. 128 – 142.

_____. A performance evaluation of local descriptors. **IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE**, v. 27, n. 10, p. 1615 – 1630, October 2005.

MIKOLAJCZYK, K.; TUYTELAARS, T.; SCHMID, C.; ZISSERMAN, A.; MATAS, J.; SCHAFFALITZKY, F.; KADIR, T.; GOOL, L. V. A comparison of affine region detectors. **International Journal of Computer Vision**, v. 65, n. 1, p. 43 – 72, November 2005.

MINAVATHI; MURALI, S.; DINESH, M. S. Model based approach for Detection of Architectural Distortions and Spiculated Masses in Mammograms. **International Journal on Computer Science and Engineering**, v. 3, n. 11, p. 3534–3546, 2011.

Mohammadi, E.; Fatemizadeh, E.; Sheikhzadeh, H.; Khoubani, S. A textural approach for recognizing architectural distortion in mammograms. In: **2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP)**. [S.l.: s.n.], 2013. p. 136–140.

NANNI, L.; BRAHNAM, S.; LUMINI, A. Texture descriptors for representing feature vectors. **Expert Systems with Applications**, v. 122, p. 163 – 172, 2019.

NANNI, L.; LUMINI, A. Heterogeneous bag-of-features for object/scene recognition. **Applied Soft Computing**, v. 13, p. 2171 – 2178, 2013.

NEGRI, T.; ZHOU, F.; OBRADOVIC, Z.; GONZAGA, A. A robust descriptor for color texture classification under varying illumination. In: **International Conference on Computer Vision, Theory and Applications**. [S.l.: s.n.], 2017. p. 378–388.

NEGRI, T. T.; ZHOU, F.; OBRADOVIC, Z.; GONZAGA, A. Extended color local mapped pattern for color texture classification under varying illumination. **Journal of Electronic Imaging**, v. 26, n. 1, p. 011008–1 – 011008–12, February 2018.

NEMOTO, M.; HONMURA, S.; SHIMIZU, A.; FURUKAWA, D.; KOBATAKE, H.; NAWANO, S. A pilot study of architectural distortion detection in mammograms based on characteristics of line shadows. **International Journal of Computer Assisted Radiology and Surgery**, v. 4, n. 1, p. 27–36, 2009.

OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern Recognition**, v. 29, n. 1, p. 51 – 59, 1996. ISSN 0031-3203.

OJALA, T.; PIETIKÄINEN, M.; MÄENPÄÄ, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 24, n. 7, p. 971 – 987, 2002.

OLIVEIRA, H. C.; MELO, C. F. E.; CATANI, J. H.; BARROS, N.; VIEIRA, M. A. C. Exploratory learning with convolutional autoencoder for discrimination of architectural distortion in digital mammography. In: HAHN, H. K.; MORI, K. (Ed.). **Proc. of SPIE Medical Imaging 2019: Computer-Aided Diagnosis**. [S.l.]: SPIE, 2019. p. 87.

- OLIVEIRA, H. C. R.; MENCATTINI, A.; CASTI, P.; MARTINELLI, E.; Di Natale, C.; CATANI, J. H.; BARROS, N.; MELO, C. F.; GONZAGA, A.; VIEIRA, M. A. C. Reduction of false-positives in a CAD scheme for automated detection of architectural distortion in digital mammography. In: MORI, K.; PETRICK, N. (Ed.). **Medical Imaging 2018: Computer-Aided Diagnosis**. [S.l.]: SPIE, 2018. v. 10575, p. 96.
- OLIVEIRA, H. C. R.; MORAES, D. R.; RECHE, G. A.; BORGES, L. R.; CATANI, J. H.; BARROS, N. de; MELO, C. F. E.; GONZAGA, A.; VIEIRA, M. A. C. A new texture descriptor based on local micro-pattern for detection of architectural distortion in mammographic images. In: ARMATO, S. G.; PETRICK, N. A. (Ed.). **Proc. SPIE Medical Imaging 2017: Computer-Aided Diagnosis**. [S.l.: s.n.], 2017. v. 10134, p. 101342U.
- OLIVEIRA, H. C. R. de; MORAES, D. R.; GONZAGA, A.; VIEIRA, M. A. da C. Detection of architectural distortion in digitized screen-film mammograms using texture descriptors. In: **Workshop de Visão Computacional**. [S.l.: s.n.], 2016.
- OWJIMEHR, M.; DANYALI, H.; HELFROUSH, M. S. Fully automatic segmentation and classification of liver ultrasound images using completed lbp texture features. In: **The Iranian Conference on Electrical Engineering**. [S.l.: s.n.], 2014. p. 1956 – 1960.
- PACI, M.; NANNI, L.; SEVERI, S. An ensemble of classifiers based on different texture descriptors for texture classification. **Journal of King Saud University**, v. 25, p. 235 – 244, 2013.
- PARK, E.; CUI, X.; KIM, W.; KIM, H. End-to-end fingerprints liveness detection using convolutional networks with gram module. **arXiv**, March 2018.
- PATIL, S.; NADAR, J. S.; GADA, J.; MOTGHARE, S.; NAIR, S. S. Comparison of various stereo vision cost aggregation methods. **International Journal of Engineering and Innovative Technology**, v. 2, n. 8, p. 222 – 226, February 2013.
- PAUL, S.; PATI, U. C. SAR Image Registration Using An Improved Anisotropic Gaussian-SIFT Algorithm. In: **IEEE India Council International Conference (INDICON)**. [S.l.: s.n.], 2017.
- RANGAYYAN, R. M.; AYRES, F. J. Gabor filters and phase portraits for the detection of architectural distortion in mammograms. **Medical and Biological Engineering and Computing**, v. 44, n. 10, p. 883 – 894, October 2006.
- _____. Design and performance analysis of oriented feature detectors. **Journal of Electronic Imaging**, v. 16, n. 2, p. 023007, April 2007.
- RANGAYYAN, R. M.; BANIK, S.; CHAKRABORTY, J.; MUKHOPADHYAY, S.; DESAUTELS, J. E. L. Measures of divergence of oriented patterns for the detection of architectural distortion in prior mammograms. **International Journal of Computer Assisted Radiology and Surgery**, v. 8, n. 4, p. 527–545, 2013.
- RANGAYYAN, R. M.; BANIK, S.; DESAUTELS, J. E. L. Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer. **Journal of Digital Imaging**, v. 23, n. 5, p. 611–631, 2010.

- RAY, K. M.; TURNER, E.; SICKLES, E. A.; JOE, B. N. Suspicious Findings at Digital Breast Tomosynthesis Occult to Conventional Digital Mammography: Imaging Features and Pathology Findings. **The Breast Journal**, p. 1–5, 2015.
- ROS, J.; LAURENT, C. Description of local singularities for image registration. In: **The 18th International Conference on Pattern Recognition (ICPR'06)**. [S.l.: s.n.], 2006.
- SHIHAVUDDIN, A. S. M.; GRACIAS, N.; GARCIA, R.; GLEASON, A. C. R.; GINTERT, B. Image-based coral reef classification and thematic mapping. **Remote Sensing**, v. 5, n. 4, p. 1809 – 1841, April 2013.
- SINGH, S.; MAURYA, R.; MITTAL, A. Application of complete local binary pattern method for facial expression recognition. In: **International Conference on Intelligent Human Computer Interaction**. [S.l.: s.n.], 2012.
- SOUZA, J. M. de; GONZAGA, A. Human iris feature extraction under pupil size variation using local texture descriptors. **Multimedia Tools and Applications**, p. 1 – 28, 2019.
- SUCKLING, J.; PARKER, J.; DANCE, D.; ASTLEY, S.; HUTT, I.; BOGGIS, C.; RICKETTS, I.; STAMATAKIS, E.; CERNEAZ, N.; KOK, S. et al. The mammographic image analysis society digital mammogram database. In: **Excerpta Medica. International Congress Series**. [S.l.: s.n.], 1994. v. 1069, p. 375–378.
- SULEIMAN, W. I.; MCENTEE, M. F.; LEWIS, S. J.; RAWASHDEH, M. A.; GEORGIAN-SMITH, D.; HEARD, R.; TAPIA, K.; BRENNAN, P. C. In the Digital Era, Architectural Distortion Remains a Challenging Radiological Task. **Clinical Radiology**, The Royal College of Radiologists, v. 71, n. 1, p. e35–e40, 2016.
- TAN, X.; TRIGGS, B. Enhanced local texture feature sets for face recognition under difficult lighting conditions. **IEEE Transactions on Image Processing**, v. 19, n. 6, p. 1635 – 1650, June 2010.
- TUCERYAN, M.; JAIN, A. K. Texture analysis. In: _____. **The Handbook of Pattern Recognition and Computer Vision**. [S.l.: World Scientific Publishing Co., 1998. p. 207 – 248.
- VARMA, M.; ZISSERMAN, A. Texture classification: are filter banks necessary? In: **IEEE Computer Society Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2003.
- _____. A statistical approach to texture classification from single images. **Int J Comput Vision**, v. 62, n. 1-2, p. 61 – 81, April 2005.
- VERONESI, U.; BOYLE, P.; GOLDBIRSCH, A.; ORECCHIA, R.; VIALE, G. Breast Cancer. **The Lancet**, v. 365, p. 1727–1741, 2005.
- VIEIRA, R. T.; NEGRI, T. T.; GONZAGA, A. Improving the classification of rotated images by adding the signal and magnitude information to a local texture descriptor. **Multimedia Tools and Applications**, v. 77, n. 23, p. 31041 – 31066, Dec. 2018.
- WANG, F.; YOU, H.; FU, X. Adapted Anisotropic Gaussian SIFT Matching Strategy for SAR Registration. **IEEE Geoscience and Remote Sensing Letters**, v. 12, n. 1, p. 160 – 164, January 2015.

- WANG, H.; YANG, K.; GAO, F.; LI, J. Normalization Methods of SIFT Vector for Object Recognition. In: **International Symposium on Distributed Computing and Applications to Business, Engineering and Science**. [S.l.: s.n.], 2011. p. 175 – 178.
- WANG, K.; BICHOT, C.-E.; LI, Y.; LI, B. Local binary circumferential and radial derivative pattern for texture classification. **Pattern Recognition**, v. 67, p. 213 – 229, July 2017.
- WANG, L.; HE, D.-C. Texture classification using texture spectrum. **Pattern Recognition**, v. 23, n. 8, p. 905 – 910, 1990.
- WANG, X.; PENG, G.; ZHANG, H. Combining multiple image descriptions for loop closure detection. **J. Intell. Robot. Syst.**, p. 1 – 21, 2017.
- WANG, Z.; HAN, Q.; NIU, X.; BUSCH, C. Feature-level fusion of iris and face for personal identification. In: YU, W.; HE, H.; ZHANG, N. (Ed.). **Advances in Neural Networks – ISNN 2009**. [S.l.: s.n.], 2009. v. 5553, p. 356 – 364.
- WANG, Z.; WANG, E.; WANG, S.; DING, Q. Multimodal biometric system using face-iris fusion feature. **Journal of Computers**, v. 6, n. 5, p. 931 – 938, May 2011.
- WEBEL, J.; GOLLA, J.; BRITZ, D.; MÜCKLICH, F. A new analysis approach based on haralick texture features for the characterization of microstructure on the example of low-alloy steels. **Materials Characterization**, v. 144, p. 584 – 596, October 2018.
- WHITE, F. E. **DATA FUSION LEXICON**. [S.l.], 1991.
[Http://www.dtic.mil/dtic/tr/fulltext/u2/a529661.pdf](http://www.dtic.mil/dtic/tr/fulltext/u2/a529661.pdf).
- WHO. World health organization - breast cancer. 2019. Disponível em: [<http://www.who.int/>](http://www.who.int/). Acesso em: apr. 5th of 2019.
- XIANG, Y.; WANG, F.; YOU, H. OS-SIFT: A Robust SIFT-Like Algorithm for High-Resolution Optical-to-SAR Image Registration in Suburban Areas. **IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING**, v. 56, n. 6, p. 3078 – 3090, June 2018.
- XIE, X.; MIRMEHDI, M. A galaxy of texture features. In: _____. **Handbook of Texture Analysis**. [S.l.]: Imperial College Press, 2008. p. 375 – 406.
- YANG, S. K.; MOON, W. K.; CHO, N.; PARK, J. S.; CHA, J. H.; KIM, S. M.; KIM, S. J.; IM, J.-G. Screening mammography-detected cancers: Sensitivity of a computer-aided detection system applied to full-field digital mammograms. **Radiology**, v. 244, n. 1, p. 104–111, 2007.
- YUN-XI, X.; FANG, C. Local image descriptor of feature combination and rotation invariant space division combination. **Acta Automatica Sinica**, v. 42, n. 4, p. 617 – 630, 2016.
- YUSHKEVICH, P. A.; WANG, H.; PLUTA, J.; AVANTS, B. B. From label fusion to correspondence fusion: a new approach to unbiased groupwise registration. In: **Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.** [S.l.: s.n.], 2012. p. 956 – 963.

ZHANG, J.; MARSZALEK, M.; LAZEBNIK, S.; SCHMID, C. Local features and kernels for classification of texture and object categories: A comprehensive study. **International Journal of Computer Vision**, v. 73, n. 2, p. 213 – 238, 2007.

ZHAO, G.; PIETIKÄINEN, M. Dynamic texture recognition using local binary patterns with an application to facial expressions. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 29, n. 6, p. 915 – 928, April 2007.

ZHAO, J.; LIU, H.; FENG, Y.; YUAN, S.; CAI, W. BE-SIFT: A More Brief and Efficient SIFT Image Matching Algorithm for Computer Vision. In: ARMATO, S. G.; PETRICK, N. A. (Ed.). **International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing**. [S.l.: s.n.], 2015. p. 568 – 574.

ZHU, C.; BICHOT, C.-E.; CHEN, L. Image region description using orthogonal combination of local binary patterns enhanced with color information. **Pattern Recognition**, v. 46, n. 7, p. 1949 – 1963, July 2013.