

SIMONE BELLO KAMINSKI AIRES

**MODELO HIERÁRQUICO PERCEPTIVO
PARA RECUPERAÇÃO DE MARCAS
REGISTRADAS VISUAIS**

Tese apresentada ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de Doutor em Informática.

CURITIBA

2015

SIMONE BELLO KAMINSKI AIRES

**MODELO HIERÁRQUICO PERCEPTIVO
PARA RECUPERAÇÃO DE MARCAS
REGISTRADAS VISUAIS**

Tese apresentada ao Programa de Pós-Graduação em
Informática da Pontifícia Universidade Católica do
Paraná como requisito parcial para obtenção do título
de Doutor em Informática.

Área de Concentração: *Ciência da Computação*

Orientadora: Profa. Dra. Cinthia Obladen de
Almendra Freitas

Co-orientador: Prof. Dr. Luiz Eduardo Soares de
Oliveira

CURITIBA

2015

Dedico este trabalho ao
meu marido e minha filha,
João Paulo e Isabella.
Aos meu pais, Luiz e Marilda,
e a toda minha família
com grande amor...

"Nas grandes batalhas da vida, o primeiro passo para a vitória é o desejo de vencer!"

Mahatma Gandhi

AGRADECIMENTOS

Agradeço a Deus porque tive saúde e sabedoria para vencer as dificuldades durante a realização deste trabalho, sem Ele nada seria possível.

À minha orientadora Profa. Dra Cinthia Obladen de Almendra Freitas, pela orientação, esclarecimento, paciência, incentivo constante e preciosos ensinamentos ao longo do desenvolvimento deste trabalho .

Ao prof. Dr. Luiz Eduardo Soares de Oliveira pelas orientações, importantes contribuições e valiosas discussões que permitiram a realização deste trabalho.

Aos meus pais Luiz e Marilda pelo amor e confiança a mim dedicados.

Ao meu querido marido João Paulo pelo amor, pela força, pelo incentivo e infinita compreensão.

À minha filha Isabella, pelos incentivos e carinhos que me mostram o que realmente importa na vida.

Ao meu colega Juliano, que esteve sempre disposto à esclarecer dúvidas e auxiliar em algumas dificuldades em matemática.

À UTFPR pelo apoio recebido durante a realização deste trabalho.

Ao PPGIa pela oportunidade e suporte oferecidos ao desenvolvimento deste trabalho.

À CAPES pelo apoio financeiro.

A todos que direta ou indiretamente colaboraram na execução deste trabalho.

SUMÁRIO

CAPÍTULO 1	16
INTRODUÇÃO	16
1.1. DESCRIÇÃO DO PROBLEMA	18
1.2. OBJETIVOS.....	20
1.2.1 <i>Geral</i>	20
1.2.2 <i>Específicos</i>	21
1.3. JUSTIFICATIVA.....	22
1.4. INEDITISMO DO TRABALHO	22
1.5. MOTIVAÇÃO	22
1.6. CONTRIBUIÇÕES	23
1.7. HIPÓTESE DE PESQUISA	24
1.8. METODOLOGIA CIENTÍFICA	24
1.9. ESTRUTURA DO TRABALHO	25
CAPÍTULO 2	26
FUNDAMENTAÇÃO TEÓRICA	26
2.1. CONSIDERAÇÕES INICIAIS	26
2.2. MARCAS REGISTRADAS (MR)	26
2.2.1 <i>Marcas Registradas no Brasil</i>	28
2.3. DESCRITORES DA FORMA	31
2.3.1 <i>Descritores Baseados em Contorno</i>	33
2.3.2 <i>Descritores Baseados em Região</i>	40
2.3.3 <i>Descritores da Forma aplicados em sistemas de recuperação de MRV</i>	50
2.4. PRINCÍPIOS DA GESTALT.....	52
2.5. CLASSIFICADORES	60
2.5.1 <i>Redes Neurais Artificiais</i>	62
2.5.2 <i>Máquinas de Vetor Suporte (SVM)</i>	64
2.5.3 <i>Múltiplos Classificadores</i>	66
2.6. MECANISMOS DE ZONEAMENTO	69
2.6.1 <i>Conceitos de Mecanismos de Zoneamento</i>	70
2.6.2 <i>Zoneamento Perceptivo Assimétrico</i>	76
2.7. MEDIDAS DE SIMILARIDADE.....	77
2.7.1 <i>Distância Euclidiana</i>	80
2.7.2 <i>Distância de Mahalanobis</i>	80
2.7.3 <i>Distância entre Histogramas</i>	81
2.8. MEDIDAS DE DESEMPENHO NA RECUPERAÇÃO	82
2.9. CONSIDERAÇÕES	83
CAPÍTULO 3	84
TRABALHOS RELACIONADOS	84
3.1. CONSIDERAÇÕES	98
CAPÍTULO 4	99
MÉTODO PROPOSTO	99
4.1. CONSIDERAÇÕES INICIAIS.....	99
4.2. PREMISSAS METODOLÓGICAS	100

4.3 BASE DE DADOS	100
4.4. MODELO HIERÁRQUICO PERCEPTIVO	102
4.5 PRÉ-PROCESSAMENTO.....	105
4.6 EXTRAÇÃO DE CARACTERÍSTICAS	106
4.6.1. <i>Primitiva com base na percepção humana das formas</i>	107
4.7. CLASSIFICAÇÃO.....	109
4.7.1 <i>Redes Neurais Artificiais (RNA)</i>	110
4.7.2 <i>Máquinas de Vetor Suporte (SVM)</i>	113
4.8. MECANISMOS DE ZONEAMENTO	115
4.9. MEDIDAS DE SIMILARIDADE.....	118
4.9.1 <i>Distância Mínima entre pontos-chave</i>	119
4.9.2 <i>Distância Mínima entre zoneamentos perceptivos</i>	121
4.10. CONSIDERAÇÕES FINAIS	124
CAPÍTULO 5.....	125
EXPERIMENTOS.....	125
5.1. CONSIDERAÇÕES INICIAIS	125
5.2. EXPERIMENTOS	126
5.2.1 <i>Experimentos utilizando Deficiências de Concavidade e Convexidade</i>	126
5.2.2 <i>Experimentos utilizando Cadeias de Freeman (FCC)</i>	131
5.2.3 <i>Comparativo entre extratores de contorno</i>	137
5.2.4 <i>Experimentos Scale Invariant Feature Transform (SIFT)</i>	138
5.2.5 <i>Experimentos Momentos Invariantes de Hu</i>	146
5.3 COMPARATIVO GERAL DE TODOS OS EXTRATORES (<i>BEST OF BEST</i>)	147
5.4 RESULTADOS RNAs	151
5.5 RESULTADOS SVM	154
5.6 COMPARATIVO DOS RESULTADOS OBTIDOS COM OUTROS TRABALHOS APRESENTADOS NA LITERATURA.....	157
5.7. CONSIDERAÇÕES FINAIS	160
CAPÍTULO 6.....	161
CONCLUSÃO.....	161
REFERÊNCIAS	166
APÊNDICE A	181
APÊNDICE B.....	188
APÊNDICE C	190

LISTA DE FIGURAS

Figura 1.1: Sistema CBIR básico	18
Figura 1.2: Comparação das trademarks Guess e Gucci	19
Figura 2.1: Exemplos de marcas registradas	27
Figura 2.2: Classificação de Vienna	28
Figura 2.3: (a) Marcas de Produtos e serviços (b) Marcas de Certificação Marcas Coletivas	(c) 29
Figura 2.4: Exemplos de similaridade da forma baseada em contorno e região.	31
Figura 2.5: Descritores da forma	32
Figura 2.6: Exemplos de formas onde se aplicam descritores baseados em contorno ...	34
Figura 2.7: (a) Cadeia de Freeman 4 direções; (b) Cadeia de Freeman 8 direções.	35
Figura 2.8: Exemplo parcial codificação Cadeias de Freeman	36
Figura 2.9: Rotulação dos pixels do fundo de duas imagens contidas na base de dados	38
Figura 2.10: Verificação de Escape e Laço Fechado.....	39
Figura 2.11: Exemplos de forma em que aplicam-se descritores baseados em região...	40
Figura 2.12: Diferenças Gaussianas DoG.....	44
Figura 2.13: Máximo e Mínimo da DoG.....	45
Figura 2.14: Histograma de orientação dos pontos-chave.....	48
Figura 2.15: Construção descritor do ponto-chave 2x2.....	49
Figura 2.16: Princípio da Proximidade.....	53
Figura 2.17: Princípio da Proximidade em MRVs	53
Figura 2.18: Princípio da Similaridade.....	54
Figura 2.19: Princípio da Similaridade em MRVs	54
Figura 2.20: Princípio da Continuidade.....	55
Figura 2.21: Princípio da Continuidade em MRVs	55
Figura 2.22: Princípio do Fechamento em MRVs.....	56
Figura 2.23: Princípio Figura-Fundo MRV Carrefour	57
Figura 2.24: Princípio Figura-Fundo MRVs	57
Figura 2.25: Princípio da Pregnância em MRVs.....	58
Figura 2.26: Princípio da Segregação.....	59
Figura 2.27: Movimentos oculares durante a percepção visual.	60
Figura 2.28: Espaço de características.....	61

Figura 2.29: Exemplo de classificação supervisionada.....	62
Figura 2.30: Modelo de um neurônio	63
Figura 2.31: Distâncias de hiperplanos e vetores de suporte.....	65
Figura 2.32: Hiperplanos: separação superior e inferior do espaço binário	66
Figura 2.33: Fases do Sistema de Múltiplos Classificadores	67
Figura 2.34: Combinação de estilo em MCS.....	68
Figura 2.35: Partições onde 1,2,3,4,5 e 6 são o número de posições das partes	71
Figura 2.36: (a) anéis circulares (b) anéis de invólucro convexo.....	71
Figura 2.37: Caractere com zoneamento 2x2 e 3x3	72
Figura 2.38: Zoneamento linear - espectrograma sinal de áudio.....	73
Figura 2.39: Particionamento em camadas.....	73
Figura 2.40: Metodologia auto-adaptativa	74
Figura 2.41: Zoneamento por Diagramas de Voronoi.....	75
Figura 2.42: Exemplo subdivisão malha 4 x 4	75
Figura 2.43: Mecanismo de Zoneamento: Z = 4, 5H, 5V e 7 regiões	76
Figura 2.44: Esquema de seleção dinâmica de mecanismos de zoneamento	77
Figura 2.45: (a)Dados do objeto (b) Matriz de distância entre objetos i,j.....	78
Figura 2.46: Cálculo da distância de Mahalanobis.....	80
Figura 3.1: MRVs (a) Mundo real, (b) Base de dados do Reino Unido.....	84
Figura 3.2: Correspondência por Deformable Template.....	85
Figura 3.3: Agrupamento em famílias.....	86
Figura 3.4: Sucessos na recuperação do ARTISAN.....	87
Figura 3.5: Sistema para recuperação de MRVs baseado na forma	87
Figura 3.6: Exemplos de elementos da Gestalt extraídos das MRVs.....	89
Figura 3.7: Resultados de consultas com sucesso	89
Figura 3.8: Exemplos de consultas complexas.....	90
Figura 3.9: Múltiplo padrão de agrupamento para uma imagem de MRV	90
Figura 3.10: Alguns agrupamentos com sucesso	91
Figura 3.11: Exemplo de parte do questionário aplicado aos 10 participantes	92
Figura 3.12: Exemplos localização MRV utilizando SIFT	93
Figura 4.1: Exemplos de MRVs contidas na base de dados.....	101
Figura 4.2: Modelo Hierárquico Perceptivo	103
Figura 4.3: Par de imagens similares.....	107
Figura 4.4: Princípio da Gestalt: figura-fundo.....	109

Figura 4.5: RNAs Múltiplos Classificadores.....	112
Figura 4.6: Construção SVM Multiclasses - Treinamento	114
Figura 4.7: Construção SVM Multiclasses - Teste.....	115
Figura 4.8: Zoneamento perceptivo em quatro regiões.	116
Figura 4.9: Zoneamento perceptivo em sete regiões.	116
Figura 4.10: Zoneamento perceptivo em cinco regiões horizontal.	117
Figura 4.11: Zoneamento perceptivo em cinco regiões vertical.....	117
Figura 4.12: Zoneamento perceptivo - MRV da base de dados	119
Figura 4.13: Correspondência entre pontos-chave.	120
Figura 4.14: Vetor de características - Zoneamento perceptivo.....	121
Figura 4.15: Correspondência entre Zoneamentos Perceptivos	122
Figura 4.16: Cálculo da média por zona z_i	123
Figura 5.1: Top-24 MRVs mais similares 1279931.bmp - $Z=0$	130
Figura 5.2: Top-24 MRVs mais similares 1138293_.bmp - $Z=5v$	131
Figura 5.3: FCC Top-24 MRVs mais similares 914.bmp - $Z=0$	135
Figura 5.4: FCC Top-24 MRVs mais similares 1190540.bmp - $Z=4$	136
Figura 5.5: SIFT Top-24 MRVs mais similares 1575268_.bmp - Ponto-a-ponto ($Z=0$)	140
Figura 5.6: SIFT Top-24 MRVs mais similares 967049.bmp - $Z=7$	144
Figura 5.7: Momentos de Hu Top-24 MRVs mais similares 1439229_.bmp - $Z=0$	147
Figura 6.1: Outdoor coca-cola	164

LISTA DE TABELAS

Tabela 2.1: Rótulos do Vetor de Características	39
Tabela 2.2: Extração de primitivas	50
Tabela 2.3: Medidas de Similaridade	79
Tabela 3.1: Trabalhos relacionados à recuperação de MRVs	96
Tabela 4.1: Conjunto verdade.....	102
Tabela 4.2: Primitivas e Princípios da Gestalt.....	108
Tabela 4.3: Composição dos conjuntos de treinamento, validação e teste – RNA	110
Tabela 4.4: Configurações RNAs.....	111
Tabela 4.5: Composição conjuntos treinamento e teste - SVM	113
Tabela 5.1: Deficiências de Concavidade e Convexidade.....	127
Tabela 5.2: Cadeias de Freeman.....	133
Tabela 5.3: Comparativos entre os resultados obtidos com os extratores de contornos	137
Tabela 5.4: SIFT ponto-a-ponto ($Z=0$).....	139
Tabela 5.5: SIFT zoneamentos perceptivos.....	142
Tabela 5.6: Comparativo R_n para todos os zoneamentos	145
Tabela 5.7: Resultados Momentos de Hu.....	146
Tabela 5.8: Comparativo Best extratores - R_n ($Z=0$).....	148
Tabela 5.9: Best de cada extrator R_n	149
Tabela 5.10: Diferenças R_n para $Z=0$ e $Z=4, 5h, 5v$ e 7	150
Tabela 5.11: Best de cada extrator P_n	152
Tabela 5.12: Voto 4 Redes Classe Modular	153
Tabela 5.13: Matriz de confusão Teste - RNAs	154
Tabela 5.14: SVM Multiclasses	155
Tabela 5.15: Matriz de confusão Teste - predict	156
Tabela 5.16: Comparativo resultados por MRV.....	157
Tabela 5.17: Comparativo resultados por MRV.....	158
Tabela 5.18: Comparativo resultados Média Geral	159

Tabela A.1: Distância de Histogramas - Concavidade e Convexidade	182
Tabela A.2: Distância de Histogramas - Cadeias de Freeman	183
Tabela A.3: Distância de Histogramas - SIFT.....	184
Tabela A.4: Comparativo DE e DH - Deficiências de Concavidade e Convexidade - Rn Top-100	185
Tabela A.5: Comparativo DE e DH - Cadeias de Freeman - Rn Top-100.....	186
Tabela A.6: Comparativo DE e DH - SIFT- Rn Top-100	187
Tabela B.1: SIFT - Rn Top-100.....	188
Tabela B.2: Conc.Conv. com Cadeias de Freeman - Top-100	189
Tabela C.1: Combinação 4 extratores Z=0 (único vetor normalizado) - Top-100.....	190
Tabela C.2: Combinação Concavidade e Convexidade e Cadeias de Freeman Top-100	191
Tabela C.3: Combinação Concavidade/Convexidade e SIFT Top-100	192
Tabela C.4: Combinação Concavidade/Convexidade e M. Hu Z=0 Top-100	193

LISTA DE QUADROS

Quadro 2.1: Algoritmo Distância Mínima entre histogramas	82
Quadro 2.2: Exemplo algoritmo Distância $H(A)$ e $H(B)$	82

LISTA DE ABREVIATURAS

ABNT	Associação Brasileira de Normas Técnicas
ART	Transformada Angular Radial
ARTISAN	<i>Automatic Retrieval of Trademarks Images by Shape Analysis</i>
CAPTCHA	<i>Completely Automated Public Turing test to tell Computers and Humans Apart</i>
CBIR	<i>Content Based Image Retrieval</i>
CDS	<i>Context- Dependent Similarity</i>
CSS	Curvatura Espaço Escala
CSV	<i>Component Similarity Vector</i>
DoG	<i>Difference of Gaussian</i>
DTC	Transformada Discreta de Cosseno
FC	Função de Custo
FCC	<i>Freeman Chain Code</i>
INMETRO	Instituto Nacional de Metrologia, Normalização e Qualidade Industrial
INPI	Instituto Nacional da Propriedade Intelectual
IPO	<i>Intellectual Property Office</i>
LDA	Análise de Discriminante Linear
MCS	<i>Multiple Classifier Systems</i>
MD	Análise Discriminante por Mahalanobis
MEC	<i>Minimum Enclosing Circle</i>
MLP	<i>Multiple Layer Perceptron</i>
MLP-NN	<i>Multiple Layer Perceptron Neural Network</i>
MOEAs	Algoritmos Evolucionários Multi-Objetivo
MPEG-7	<i>Moving Picture Experts Group</i>
MRV	Marca Registrada Visual
NSLO	Norte-Sul-Leste-Oeste
P_n	<i>Normalized Precision</i>
R_n	<i>Normalized Recall</i>
RNA	Redes Neurais Artificiais
SIFT	Scale Invariant Feature Transform
SVD	<i>Singular Value Decomposition</i>
SVM	<i>Support Vector Machine</i>
TAE	Teoria de Aprendizagem Estatística
WBG	<i>Maximum Weighted Bipartite Graph</i>
WIPO	<i>World Intellectual Property Organization</i>
WWF	<i>World Wide Fund for Nature</i>

Resumo

Uma Marca Registrada Visual (MRV) permite diferenciar um produto ou serviço de uma empresa frente a seus concorrentes, tornando-o especial e único visando obter visibilidade no mercado. O registro de uma MRV e sua avaliação para distinção, têm se tornado um trabalho complexo e entediante para os escritórios que realizam esta atividade, visto que há milhões de MRVs já registradas e muitos outros pedidos são apresentados periodicamente. Estes escritórios têm de que garantir que todas as novas MRVs sejam distintas o suficiente para evitar confusão com aquelas já registradas. Desta forma, o presente trabalho se justifica diante da complexidade da tarefa no registro de uma nova MRV, sendo necessário estabelecer uma abordagem que possa ser utilizada ou estendida, abrangendo a maior variedade de imagens de MRVs. Neste estudo um Modelo Hierárquico Perceptivo em dois níveis é proposto e testado, sendo que tal modelo busca considerar aspectos da percepção humana das formas por meio da aplicação de mecanismos de zoneamento perceptivo. O 1º. Nível realiza a seleção dinâmica de extratores de características, utilizando dois extratores de contorno (Deficiências de Concavidade e Convexidade e Cadeias de Freeman) e dois extratores por região (Momentos Invariantes de Hu e SIFT – *Scale Invariant Feature Transform*). A seleção dinâmica dos extratores é realizada utilizando SVM. No 2º. Nível são aplicados os mecanismos de zoneamento perceptivo de acordo com o extrator de características selecionado dinamicamente no 1º. Nível. O cálculo das similaridades é realizado utilizando Distância Euclidiana, Distância mínima entre pontos-chave e Distância Mínima entre zoneamentos perceptivos. Um *ranking* com as Top-100 imagens mais similares é obtido. Com o objetivo de avaliar o desempenho na recuperação das MRVs, foram utilizadas duas medidas: R_n (*recall*) e P_n (*precision*). A similaridade calculada por meio da Distância Mínima entre zoneamentos perceptivos, proposta para o extrator SIFT, é uma abordagem que obteve resultados satisfatórios; uma vez que melhora as taxas obtidas para aproximadamente 50% das MRVs consultadas. Assim, os resultados para o Modelo proposto foram de $R_n = 0.93$ e $P_n = 0.75$, sendo superiores aos encontrados na literatura, demonstrando que o modelo apresentado é uma importante ferramenta na recuperação de MRVs. Além disso, o conceito Figura-Fundo do princípio da Gestalt, ainda não explorado na recuperação de MRVs, é implementado no trabalho por meio das Deficiências de Concavidade e Convexidade, tendo como resultado $R_n = 0.83$ e $P_n = 0.61$, demonstrando ser possível incorporar o conceito Figura-Fundo nos sistemas computacionais.

Palavras-Chave: Marcas Registradas Visuais, Gestalt, Figura-Fundo, Zoneamento, Modelo Hierárquico Perceptivo, Extração de Primitivas.

Abstract

A trademark allows distinguishing a company's product or service from its competitors, making it special and unique to obtain visibility in the global market. A trademark registration and its evaluation for distinction, has become a complex and tedious work for the offices that perform this activity, since millions of trademarks are already registered and many other requests are made periodically. Each patent office must ensure that all new trademarks are sufficiently distinct to avoid confusion with existing brands. In this way, the present work is justified by the complexity of the task in registering a new trademark, being necessary to establish an approach that can be used or extended, including the greatest variety of trademarks. In this study, a Perceptual Hierarchical Model on two levels is proposed and tested. It incorporates aspects of human perception in this process by means of a non-symmetrical perceptual zoning mechanism. The 1st level performs the dynamic selection of features extractors where we used two contour-based extractors (Concavity and Convexity Deficiencies and Freeman Chain Code) and two region-based extractors (Invariant Moments of Hu and SIFT - Scale Invariant Feature Transform). The dynamic selection of extractors is carried out through an SVM. In the 2nd level, the perceptual zoning mechanisms in accordance with the features extractor, selected dynamically in the first level, are applied. The calculation of similarities was performed using Euclidean Distance, Minimum Distance between keypoints and Minimum Distance between perceptual zoning mechanisms. A ranking of the Top-100 most similar images is provided. In order to evaluate the performance in the trademarks retrieval, two measures were used: Rn (recall) and Pn (precision). The similarity calculated using the Minimum Distance between perceptual zoning mechanisms proposed to SIFT extractor, is an approach that achieved satisfactory results; it improves the rates obtained for around 50% of the trademarks queried. Thus, the results for the proposed model were $R_n = 0.93$ and $P_n = 0.75$, being higher than those found in the literature, demonstrating that the model is an important tool for trademark retrieval. In addition, the Gestalt principle called Figure-Ground, unexplored in trademark retrieval systems, implemented in this work by means of the Concavity and Convexity Deficiencies, achieved results for R_n and $P_n = 0.83$ and 0.61 , respectively, demonstrating that it is possible to incorporate this concept in computer systems.

Keywords: Trademarks, Gestalt, Figure-Ground, Zoning, Perceptual Hierarchical Model, Feature Extraction.

Capítulo 1

Introdução

Atualmente tecnologias permitem gerar informações na forma visual, e estas representam uma enorme quantidade de dados armazenados na Web ou em um domínio específico. O aumento dos bancos de dados multimídia tem motivado pesquisas para o desenvolvimento de ferramentas que buscam extrair automaticamente conhecimento de grandes volumes de conteúdos digitais.

Por intermédio destas ferramentas automatizadas que utilizam motores de busca avançada, torna-se possível a recuperação de informações usando, não somente cadeias de caracteres como também a de um objeto de interesse. Neste caso, as pesquisas seriam baseadas em conteúdo a fim de permitir aos usuários recuperar informações visuais relativas à moda, publicidade, medicina, arquitetura, entre outros, aumentando assim, a importância de sistemas de recuperação com base em imagens digitais.

Neste mundo digital, as empresas tentam estabelecer o nome de sua marca no mercado para o reconhecimento mundial. Para isso, elas utilizam elementos visuais (cor, textura e forma) e/ou não visuais (som e aroma), permitindo o fortalecimento dos seus produtos. Estes elementos são as Marcas Registradas Visuais (do inglês, *trademarks*), as quais desempenham um papel importante no fortalecimento da identidade da empresa, visto que toda companhia ou organização quer uma marca distinta para obter visibilidade no mercado.

Uma Marca Registrada (MR) pode consistir de textos, imagens e, ocasionalmente de outras mídias, como aroma e som [EAK98]. Marcas Registradas compostas somente por textos podem ser gerenciadas por meio de índices alfabéticos impressos ou *software* de reconhecimento de textos. Mesmo assim o registro deste tipo de marca registrada pode ser complexo, diante da necessidade de identificar as MRs que

possuam, por exemplo, som associado, sendo que tais sons podem ser semelhantes quando falados em voz alta, bem como, podem conter grafias semelhantes [EAK98].

O registro da MRV e sua avaliação para distinção, têm se tornado um trabalho muito entediante para os escritórios que realizam esta atividade. Milhões de MRVs já estão registradas e milhões de pedidos são apresentados para registro [EAK98]. Cada escritório de patentes deve garantir que todas as novas MRVs sejam distintas o suficiente para evitar confusão com aquelas já registradas. Para isso, precisam manter detalhes de todas as MRVs existentes, de maneira que possam ser rapidamente acessadas para identificar possíveis conflitos no momento do seu registro.

Os Sistemas de Recuperação de Imagens Baseado em Conteúdo (*Content Based Image Retrieval* - CBIR) [YOS99] [DAT08] [PEN09] [SNO10] [SIL11] têm sido utilizados para a recuperação automática de imagens, a partir de um banco de dados utilizando-se, por exemplo, cor, textura e características da forma. Esses sistemas buscam superar as dificuldades encontradas nos sistemas de recuperação de imagem baseados em texto [SIL11].

Os sistemas CBIR utilizam muitos métodos da área de processamento de imagens e visão computacional, sendo considerados por alguns pesquisadores como um subconjunto destas áreas [KUM11]. Nos sistemas CBIR, a pesquisa é feita analisando o conteúdo das imagens. Explica-se que o termo "conteúdo" pode se referir às cores, formas, texturas ou qualquer outra informação que pode ser derivada da própria imagem.

Um sistema CBIR típico permite aos usuários formular consultas por meio da apresentação de uma imagem exemplo (*query*). Em seguida, o sistema identifica as imagens armazenadas no banco de dados, cujas características sejam similares aos da imagem de consulta e exibe miniaturas das imagens semelhantes na tela. A Figura 1.1 apresenta um sistema CBIR básico, sendo que este sistema funciona em duas etapas: na primeira etapa são extraídas as características a partir da imagem de consulta; na segunda etapa, as características extraídas são verificadas ou "correspondidas" com as características das imagens do banco de dados.

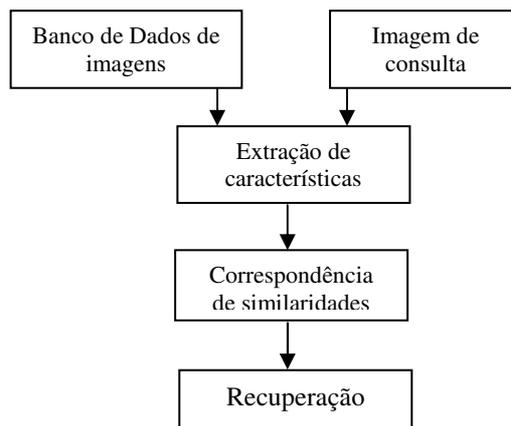


Figura 1.1: Sistema CBIR básico

Fonte: Adaptado de [KUM11]

Existem diferentes técnicas e abordagens disponíveis para a verificação, distinção e recuperação de MRVs em grandes bases de dados. No entanto, o desenvolvimento de sistemas de recuperação automática tem desafiado os pesquisadores devido ao alto grau de dificuldade em reproduzir a capacidade humana de percepção da forma.

Este estudo situa-se no contexto de recuperação de MRVs. Um fator determinante para um bom desempenho na recuperação é a seleção do conjunto de características a serem extraídas das MRVs. O método de extração deve identificar as características particulares de cada uma, salientando suas similaridades em relação às outras MRVs.

O presente trabalho visa contribuir para os sistemas CBIR voltados à recuperação de MRVs, buscando incorporar aspectos da percepção humana das formas neste processo por meio de um Modelo Hierárquico Perceptivo. Neste Modelo Hierárquico Perceptivo lança-se mão de um mecanismo de zoneamento perceptivo não-simétrico baseado nos princípios da Gestalt já utilizado com sucesso em reconhecimento de caracteres manuscritos [FRE07].

1.1. Descrição do problema

A fim de manter a integridade e a visibilidade de suas MRVs, as empresas realizam buscas constantes na Internet e nos meios de comunicação em geral (revistas, jornais, vídeos, entre outros) para verificar a existência de MRVs similares à sua MRV,

as quais podem estar sendo usadas sem autorização. No entanto, a mão de obra e os custos associados a este esforço aumentam a cada ano [ABE09].

Segundo informações do canal de notícias Fashion Forward¹ no ano de 2009, um processo foi instaurado em Nova York por solicitação da marca italiana *Gucci* contra a marca americana *Guess*. A alegação da *Gucci* é que a *Guess* imita quatro características específicas da marca italiana, conforme apresentado na Figura 1.2. Pode-se observar a listra verde e vermelha; a estampa dos G's conectados, e o G sozinho ou isolado dentro de uma moldura quadrada.



Figura 1.2: Comparação das trademarks Guess e Gucci

Fonte: [FFW12]

A acusação da *Gucci* afirma que a *Guess* está tentando arruinar o *design* aplicado nos produtos comercializados por eles, utilizando elementos similares de maneira a confundir os clientes. Para o advogado da *Guess*, apesar de existirem similaridades quanto ao tamanho do nome das marcas (cinco letras) e suas MRVs, jamais um cliente da *Guess* confundiria seus produtos com o da *Gucci*. Entretanto, fica a indagação: "Realmente é parecido ou não?".

Em face desse tipo de disputa judicial, diversas ferramentas e técnicas de processamento de imagem podem ser usadas para tentar auxiliar na resolução de diferentes problemas relacionados à imagem, texto, gráficos e cores, pois uma MRV é, na verdade, uma combinação desses elementos.

¹ <http://ffw.com.br/noticias/moda/gucci-x-guess-similaridade-copia-inspiracao-ou-coincidencia/>

No entanto, a recuperação de MRVs ainda é um problema aberto e de difícil solução, conforme apresentado na revisão bibliográfica no Capítulo 3. Dessa forma, alguns problemas importantes na recuperação de MRVs devem ser considerados [EAK98] [JAI98] [JIA06]:

- 1) É necessário estabelecer uma abordagem, que possa ser utilizada ou estendida para abranger uma maior variedade de imagens de MRVs.
- 2) Realizar buscas de dados similares em grandes volumes de dados, com uma boa precisão e tempo de resposta razoável.
- 3) Encontrar descritores que traduzam as propriedades visuais utilizadas para descrição das imagens que sejam invariantes em termos de rotação, escala e translação;
- 4) A descrição semântica das imagens de MRVs deve encontrar resultados visualmente similares;
- 5) O desenvolvimento de características da forma que sejam capazes de capturar a percepção humana na similaridade das formas;
- 6) Uso de descritores que sejam capazes de modelar as características geométricas abstratas encontradas nas MRVs;
- 7) Utilizar medidas de similaridade da forma, que garanta bom desempenho de recuperação, e que os itens recuperados sejam coerentes com a percepção visual humana.

Pode-se observar pelos problemas destacados a complexidade em desenvolver sistemas para a recuperação de MRVs, visto que muitos aspectos estão relacionados à percepção visual humana ou mesmo à diferença entre a visualização humana das formas e a representação possível nos sistemas computacionais.

1.2. Objetivos

1.2.1 Geral

O objetivo principal deste trabalho é utilizar um Modelo Hierárquico aplicando conceitos da percepção humana das formas para extração de características e recuperação de MRVs em bases de dados digitais. Este modelo proposto nominamos Modelo Hierárquico Perceptivo.

1.2.2 Específicos

Para este estudo foram definidos alguns objetivos específicos, a saber:

- entender MRVs;
- desenvolver um modelo de recuperação hierárquico, ou seja, em níveis;
- testar e determinar um conjunto de extratores de características para compor o Modelo Hierárquico;
- utilizar classificação dinâmica dos extratores de características;
- aplicar mecanismos de zoneamento não-simétricos;
- aplicar medidas de similaridade, propondo um cálculo de similaridade baseada em regiões perceptivas;
- aplicar conceitos da Gestalt e da percepção humana da forma na recuperação de MRVs;
- implementar um cenário de teste como prova de conceito para avaliar e validar o método proposto, analisando os resultados obtidos.

Visto que o presente trabalho apresenta um Modelo Hierárquico Perceptivo, resumem-se a seguir os níveis que compõem tal modelo:

- Nível 1: extração de primitivas utilizando quatro extratores, dois baseados em contorno e dois baseados em região. Os extratores baseados em contorno da forma são: Codificação de Freeman [FRE78] e Concavidade e Convexidade (*Convex Deficiencies*) [PAR97]. Extratores baseados em região: Momentos Invariantes de Hu [HU_62] e SIFT (*Scale Invariant Feature Transform*) [LOW04]. Uma seleção dinâmica para determinar o melhor extrator (*Best*) é realizada por meio de um classificador.
- Nível 2: aplicam-se os mecanismos de zoneamento perceptivo [FRE07] ao melhor extrator de primitivas determinado no Nível 1. Obtém-se o melhor zoneamento (*Best-Zoning*) a ser utilizado na recuperação das MRVs e, finalmente, são aplicadas medidas de similaridade para estabelecer um *ranking* considerando as Top-100 imagens similares.

Deve-se destacar que o mecanismo de zoneamento é um método de análise de informações locais em um dado padrão particionado. A análise de informações locais

permite evidenciar elementos-chave, contribuindo para o processo de recuperação. Esta discussão encontra-se apresentada na Seção 2.6 do Capítulo 2.

1.3. Justificativa

O presente estudo justifica-se diante da complexidade na tarefa de recuperação de MRVs, tendo em vista a descrição do problema presente na Seção 1.1.

Com objetivo de dar continuidade a aspectos da percepção humana e mecanismos de zoneamento já estudados por [FRE02] [FRE07] [COR02] [AIR05] [HIR11], um modelo hierárquico perceptivo é proposto, buscando auxiliar com soluções para o problema de recuperação de MRVs já estudados por [EAK98] [HUN04] [CER06] [JIA06] [ABE06] [BAG07] [ANU13] [SAH13] [SHA14].

Este trabalho está inserido no contexto da recuperação de imagens de MRVs, com a perspectiva de espelhar-se na percepção humana da forma, uma vez que acredita-se que a área de percepção humana das formas pode trazer complementariedade à etapa de extração de características. O estudo investiga a utilização de regiões perceptivas aplicadas ao processo de recuperação de MRVs.

1.4. Ineditismo do Trabalho

Este trabalho propõe um Modelo Hierárquico Perceptivo, resultante da utilização de 4 (quatro) extratores de características, associado a mecanismos de zoneamentos perceptivos, para o problema de recuperação de MRVs

Além disso, não foram identificadas outras pesquisas envolvendo zoneamentos perceptivos aplicados à recuperação de MRVs, bem como medidas de similaridade que utilizam esses zoneamentos para tal tarefa. Um conjunto de MRVs para consulta, definidas e classificadas por especialistas do escritório de patentes do Reino Unido (IPO), possibilitou a realização de testes como prova de conceito [IPO12].

1.5. Motivação

A principal motivação deste trabalho é a definição de um modelo suportado por uma solução computacional para o problema de recuperação de MRVs utilizando percepção humana das formas.

Deve-se ressaltar que a recuperação de MRVs é um problema extremamente desafiador, uma vez que o desenvolvimento de soluções computacionais além de produzirem resultados estatisticamente comprobatórios (ou seja, taxas de recuperação compatíveis com as encontradas na literatura) oferece resultados que contribuem com a área de CBIR, especificamente na recuperação de MRVs.

A percepção humana das formas baseia-se na hipótese de que, o ser humano percebe o todo e não as partes [FRE02] [GOM04]. Porém, mesmo as pessoas possuindo diferentes percepções visuais, existem elementos invariantes e relevantes para identificar objetos ou pessoas. Essas informações invariantes e os elementos relevantes devem ser incorporados ao processo de recuperação de MRVs.

O estudo desses aspectos foi objeto desta pesquisa, uma vez que os especialistas, assim como o modelo proposto, utilizam elementos da percepção humana para realizar o processo de recuperação de MRVs [FRE07].

1.6. Contribuições

Conforme destacado, a recuperação de MRVs é uma tarefa de alta complexidade, em virtude das diversidades da forma e elementos abstratos que uma MRV pode apresentar. Com isso, os sistemas de reconhecimento precisam contar com mecanismos para garantir a eficiência na recuperação.

Assim, este trabalho complementa as abordagens propostas por [EAK98] [HUN04] [JIA06] [ABE06] [HIR09], aplicando um modelo hierárquico que incorpora regiões perceptivas advindas do estudo do reconhecimento humano das formas.

Trata-se de uma proposta para ampliar os estudos a cerca da teoria da *Gestalt* [EAK98] [AIR05] [FRE07] [JAI06] [ABE06] [HIR09] [HIR11], aplicada ao reconhecimento de MRVs, sendo que esta Teoria sugere que os seres humanos avaliam formas complexas globalmente, mas em um segundo momento, são os detalhes que fazem a diferença [FRE02].

Até a presente data, os trabalhos que utilizam a Teoria da Gestalt na recuperação de MRVs, aplicam a extração de características evidenciando as forças de percepção (saliências visuais): proximidade, fechamento, continuidade, paralelismo, similaridade, simetria [EAK98] [ABE06] [JIA06] [HIR09]. No entanto, o princípio da Gestalt Figura-Fundo ainda não foi explorado. Adicionalmente ao Modelo Hierárquico proposto, complementa-se a abordagem proposta por [ABE06] [HIR09], na qual são utilizadas as

forças de percepção, explorando-se as Deficiências de Concavidade e Convexidade. Pode-se então, tal qual no princípio Figura-Fundo, explorar em 1o. plano os pixels pretos contidos na imagem, e em 2º. plano aplicar-se um procedimento de rotulação do fundo da imagem (*background*) considerando-se os pixels brancos. Tais elementos encontram-se apresentados e discutidos na Seção 2.3.1.2.

Além disso, com base na revisão bibliográfica, ainda não foram apresentados trabalhos que se baseiam em princípios mais simples da Gestalt, que tratam o fato do ser humano fazer uma análise global e posteriormente uma análise local para evidenciar os detalhes que passam a ser um elemento-chave na recuperação das MRVs. Assim, o presente trabalho tem como principal contribuição a proposta de um modelo hierárquico que contemple aspectos da percepção da forma, por meio de extratores de características associados a um mecanismo perceptivo de zoneamento.

1.7. Hipótese de Pesquisa

A hipótese básica desta pesquisa, é que é possível desenvolver um Modelo Hierárquico composto por uma combinação de extratores de características e mecanismos de zoneamento, produzindo resultados superiores aos apresentados por métodos os quais não utilizam zoneamentos perceptivos e não selecionam dinamicamente seus extratores.

1.8. Metodologia Científica

Este trabalho consiste de uma pesquisa exploratória de natureza aplicada, pois objetiva estudar a percepção humana e gerar conhecimento para a solução de um problema específico da área de recuperação de Marcas Registradas Visuais (MRV).

Como método científico adotou-se o método dedutivo, uma vez que com base em um conhecimento técnico e científico já formalmente conhecido é possível desenvolver e avaliar uma solução computacional que ofereça suporte de maneira consistente com tais conhecimentos (e/ou premissas).

Trata-se de uma pesquisa quantitativa, uma vez que a abordagem adotada para análise do método proposto ocorre por meio dos resultados mensuráveis obtidos com os experimentos realizados.

Em relação aos procedimentos técnicos, foram realizados levantamentos bibliográficos que fundamentaram o desenvolvimento do modelo proposto. Atividades experimentais foram realizadas subsequentemente neste método para prova de conceito e análise do método propriamente dito.

Inicialmente foram realizados estudos com os objetivos de entender a tarefa de recuperação de MRVs. Também foram realizados estudos sobre Sistemas de Recuperação de Imagens por Conteúdo (CBIR - *Content-based Image Retrieval*) e métodos aplicados à recuperação de MRVs. Todos esses elementos, permitiram definir metas a serem atingidas nesta pesquisa, principalmente com relação às taxas de recuperação que poderiam e deveriam ser obtidas pelo método proposto.

Como segunda etapa desta pesquisa, o modelo proposto foi definido e os extratores de características foram selecionados, estudados e implementados. Em seguida, foram aplicados os mecanismos de zoneamento perceptivo, e procede-se uma análise dos resultados obtidos pelos extratores de modo individual e combinados. Definindo-se também o método de classificação dinâmica do extrator.

Na etapa final, foram implementados métodos para cálculos de similaridade. Por meio de experimentos, definiu-se os melhores extratores e zoneamentos a serem utilizados buscando alcançar as melhores taxas de recuperação.

1.9. Estrutura do Trabalho

Este trabalho está organizado em seis capítulos. No Capítulo 2 apresenta-se a fundamentação teórica. O Capítulo 3 apresenta os trabalhos relacionados. O Capítulo 4 descreve o método proposto para a recuperação de MRVs. Os experimentos realizados são apresentados no Capítulo 5, bem como a análise dos resultados. No Capítulo 6 são apresentadas as conclusões e também as propostas de trabalhos futuros.

Capítulo 2

Fundamentação Teórica

2.1. Considerações Iniciais

Neste Capítulo apresenta-se a fundamentação teórica necessária para tratar o problema de recuperação de Marcas Registradas Visuais (MRVs). São apresentados os Descritores da Forma, os princípios da Teoria da Gestalt, mecanismos de zoneamento (regiões perceptivas), classificadores, medidas de similaridade e as medidas de desempenho na recuperação.

2.2. Marcas Registradas (MR)

Uma marca registrada (do inglês *Trademark*) pode ser um conjunto de palavras ou símbolos utilizado para identificar um produto ou serviço. As empresas utilizam as MRVs para distinguir-se de outras empresas.

Imaginando situações cotidianas, pode-se considerar como “marca”, uma lembrança forte deixada em nossas mentes. Estas “marcas” podem ser bens, serviços, lugares, pessoas, experiências, que surgem em nossas mentes quando se entra em contato com algo que permita relembrar [VER10].

As equipes de *marketing* das empresas buscam criar “marcas” de seus produtos nas mentes dos consumidores, logo não é uma tarefa simples fazer com que os consumidores se lembrem de uma determinada MRV ao consumir um produto [DUH12]. Pois estas MRVs conseguem trazer lembranças aos consumidores dos

produtos que representam. Porém, elas estão em constante avaliação de experiências positivas ou negativas em relação a cada produto a ela associado [VER10].

Para Veras [VER10], quanto mais positivas forem estas experiências, maior é o vínculo criado com a MRV, e a tendência é consumir mais os produtos associados a ela. Porém, deve-se também considerar que o inverso é verdadeiro. Quanto maior o nível de relação do consumidor com a MRV pressupõe-se que esta faz parte de sua vida e que não será substituída sem um bom motivo. A Figura 2.1, apresenta diversas imagens de MRVs conhecidas mundialmente.



Figura 2.1: Exemplos de marcas registradas

Uma MRV precisa ser única, pois esta mantém a visibilidade do produto ou serviço ao qual ela representa. Havendo MRVs similares, o consumidor ficará confuso, o que leva a problemas de lealdade e identidade visual. Com o intuito de evitar esse problema, as MRVs devem ser devidamente registradas em escritórios de marcas e patentes [EAK98] [VER10].

Como enfatizado na Seção 1.1, existem diversos problemas quando o registro de uma nova MRV é solicitado. A permissão para uso desta MRV só é concedida depois de realizada uma busca na base de dados verificando a existência de possíveis MRVs similares. A Lei que regulamenta o registro de novas MRVs no Brasil (Lei no. 9279) é apresentada na Seção 2.2.1.

Eakins et al. [EAK98] apresentam que alguns métodos de recuperação de MRVs utilizam o sistema Vienna para a classificação através de códigos, desenvolvido pela *World Intellectual Property Organization* (WIPO)². Entretanto, existem dois

² <http://www.wipo.int/portal/en/index.html>

inconvenientes inerentes a qualquer sistema de classificação manual: 1) A classificação manual é demorada e potencialmente propensa a erros; 2) Sistemas de classificação por códigos nem sempre são úteis para a recuperação, particularmente para imagens abstratas (por exemplo, as MRVs).

O propósito da Classificação de Vienna é facilitar as buscas antecipadas de MRVs e auxiliar no trabalho de reclassificação quando documentos são trocados em nível internacional [WIP12]. Os países que fazem parte do acordo de *Vienna* não precisam elaborar sua própria classificação ou manter uma já existente. A classificação constitui um sistema hierárquico que procede do genérico para o específico (*top-down*), dividindo todos os elementos figurativos em categorias, divisões e seções, conforme apresentado na Figura 2.2.

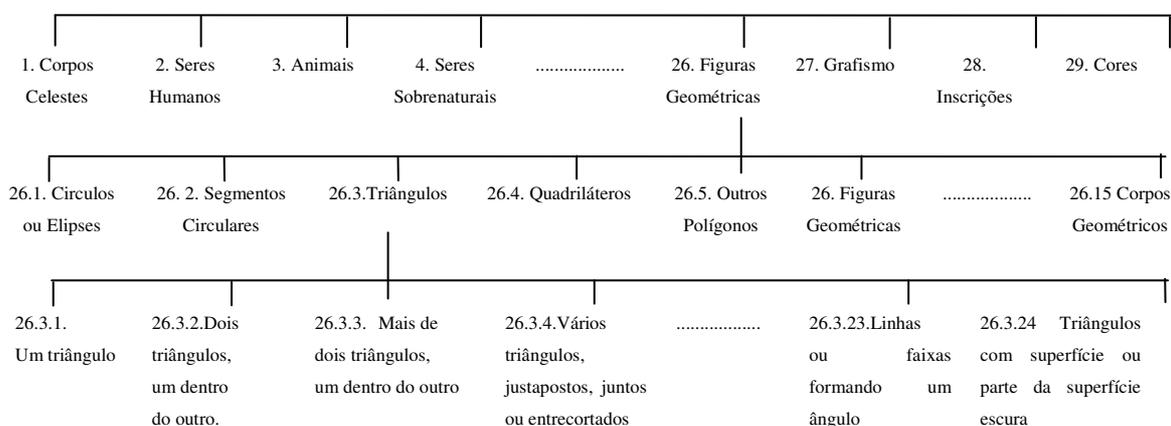


Figura 2.2: Classificação de Vienna

Fonte: Adaptado de [EAK98] [WIP12]

Observa-se na Figura 2.2, que no primeiro nível existem 29 classificações, sendo que a classificação 26 corresponde a Figuras Geométricas e possui 15 classificações, e a classificação 26.3, que corresponde a triângulos tem 24 classificações. Portanto, percebe-se que realizar uma classificação de uma MRV utilizando a Classificação de Vienna não é uma tarefa trivial.

2.2.1. Marcas Registradas no Brasil

A Lei nº 9.279, de 14 de maio de 1996, regulamenta direitos e obrigações relativos à propriedade industrial³. No Título III é apresentada a regulamentação

³ http://www.planalto.gov.br/ccivil_03/leis/19279.htm

referente ao registro de Marcas. Conforme descrito na Lei, tem-se como Marcas sinais distintivos visualmente e que não sejam proibidos legalmente. São considerados três tipos de Marcas [BRA96]:

Marca de produto ou serviço: são utilizadas para distinguir produto ou serviço, em diferentes áreas de atuação no mercado. Na Figura 2.3-a são apresentadas algumas marcas brasileiras, que apresentam formatos bem distintos.

Marca de certificação: usada para atestar a conformidade de um produto ou serviço, baseada em uma norma ou especificação técnica. Por exemplo, o selo da Associação Brasileira de Normas Técnicas (ABNT) ou o selo Instituto Nacional de Metrologia, Normalização e Qualidade Industrial (INMETRO). Tais marcas podem ser visualizadas na Figura 2.3-b.

Marca coletiva: usada para identificar produtos ou serviços prestados por uma entidade que representa um grupo, por exemplo, marcas de cooperativas. A Figura 2.3-c apresenta a Marca da Cooperativa Agroindustrial COAMO e do Sistema Cooperativo de Crédito SICREDI.



Figura 2.3: (a) Marcas de Produtos e serviços (b) Marcas de Certificação
(c) Marcas Coletivas

Conforme consta no artigo 128 da Lei 9279/1996, o registro de Marcas pode ser requerido por pessoa física ou jurídica de direito público ou privado. Porém, as pessoas de direito privado, só podem requerer registro de marca da atividade que exercem efetiva e licitamente [BRA96].

O artigo 130 da mesma lei destaca que a propriedade da marca é adquirida por meio do registro emitido pelo Instituto Nacional de Propriedade Industrial (INPI),

garantindo a utilização exclusiva em todo território nacional. O titular da marca não poderá impedir, conforme o artigo 132, que setores do comércio ou de distribuição utilizem os símbolos que representem sua própria identificação em conjunto com a marca do produto. Ainda, não poderá impedir que a marca seja citada em discursos ou em publicações (obras científicas, literárias, entre outras), desde que a utilização apresente conotação comercial ou que prejudique a imagem da marca.

O registro da marca tem o prazo de 10 (dez) anos, contados a partir da data em que o registro foi concedido, podendo ser prorrogado por períodos iguais e sucessivos, conforme consta no Capítulo V artigo 133.

De acordo com o Capítulo VII, após o registro ser protocolado, reclamações à concessão da MRV podem ser realizadas em um prazo de 60 (sessenta) dias. As novas Marcas são enviadas para exame, somente após o prazo de oposições ter sido esgotado. Concluído o exame, será emitida decisão, deferindo ou não o pedido de registro.

Após aprovação do registro, é emitido um certificado, no qual constará, conforme artigo 164 do Capítulo X, a marca, o número e data do registro, nome, nacionalidade e endereço do titular, os produtos ou serviços, as características do registro e a prioridade estrangeira.

O Capítulo XI, destaca que o registro de uma MRV pode ser cancelado, no caso do titular de uma marca registrada em outro país, signatário da Convenção da União de Paris⁴, vir a reivindicar por intermédio de uma ação judicial, conforme prevê a Convenção. Tal processo de nulidade pode ser iniciado por qualquer pessoa com legítimo interesse, no prazo de 180 (cento e oitenta) dias contados da data da expedição do certificado de registro.

Como mencionado anteriormente, no Brasil o órgão que regulamenta as novas MRVs é o INPI⁵. Quando o usuário faz o pedido de uma nova MRV, é necessário associar os produtos ou serviços que aquela MRV visa proteger. Para tal associação, o INPI utiliza a Classificação Internacional de Produtos e Serviços de Nice, que possui uma lista de 45 classes com informações sobre os diversos tipos de produtos e serviços e o que pertence a cada classe. Para realizar o exame dos elementos figurativos da MRV - avaliar a similaridade entre as MRVs já registradas - é utilizada a Classificação de

⁴ Convenção da União de Paris para a Proteção da Propriedade Industrial
http://www.planalto.gov.br/ccivil_03/decreto/1990-1994/anexo/and1263-94.pdf

⁵ <http://www.inpi.gov.br/portal/>

Vienna. Em 2014, o INPI recebeu o depósito de 156.989 pedidos de renovações e novos registros de MRV. Destes pedidos, 85.810⁶ novos registros foram concedidos.

Na busca ou recuperação de MRVs, o julgamento por similaridade pode ser baseado em um número de critérios, incluindo a forma global, a forma dos componentes da imagem, e a configuração espacial dos componentes. No entanto, é complexo para os sistemas de classificação manual refletir esta ampla gama de critérios [EAK98].

2.3. Descritores da Forma

Para Bober [BOB01], a noção de forma do objeto apesar de intuitiva pode ter diversos significados. Os objetos do mundo real são 3-D, mas imagens e vídeos geralmente contém projeções 2-D. No caso das projeções 2-D, existem duas noções de similaridade: 1) baseada em região e 2) baseada em contorno, conforme apresentadas na Figura 2.4.

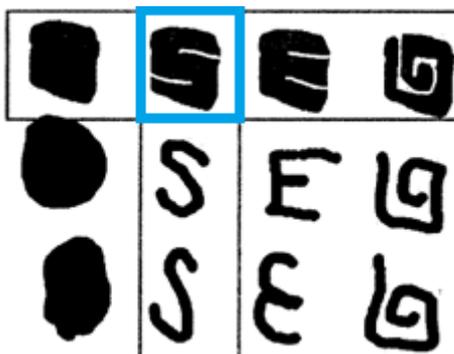


Figura 2.4: Exemplos de similaridade da forma baseada em contorno e região.

Fonte: [BOB01]

Observa-se na Figura 2.4 que as imagens da primeira linha possuem distribuição espacial similar dos pixels, conforme os critérios baseados em região. No entanto, possuem claramente contornos distintos. Em relação à similaridade baseada em contorno, esta é observada nos objetos mostrados em cada coluna. Considerando uma consulta com o objeto da primeira linha e segunda coluna (destaque em cinza claro), as imagens consideradas similares seriam as da primeira linha (quando a consulta é baseada em região) ou da segunda coluna (quando a consulta é baseada em contorno). A extração de características por região considera a distribuição dos pixels na imagem.

⁶ <http://www.inpi.gov.br/portal/artigo/estatisticas>

Dessa forma, analisando a Figura 2.4 coluna 1 e linha 1, as imagens similares seriam as contidas na primeira coluna e na primeira linha. No entanto, percebe-se que os contornos distintos contidos nas imagens da primeira linha são ignorados pelo extrator de região, causando confusão com as imagens da primeira coluna.

Os descritores de formas são obtidos por modelos matemáticos que representam a forma do objeto como um todo ou de uma região específica. Os descritores ideais devem possuir invariância à translação, rotação, escala e ponto de início. Métodos considerados eficientes aliam estas características associadas à robustez em relação a ruído na imagem [BOB01]. Na Figura 2.5 são apresentados alguns descritores da forma, sendo importante observar que os métodos apresentados nesta figura não são os únicos existentes.

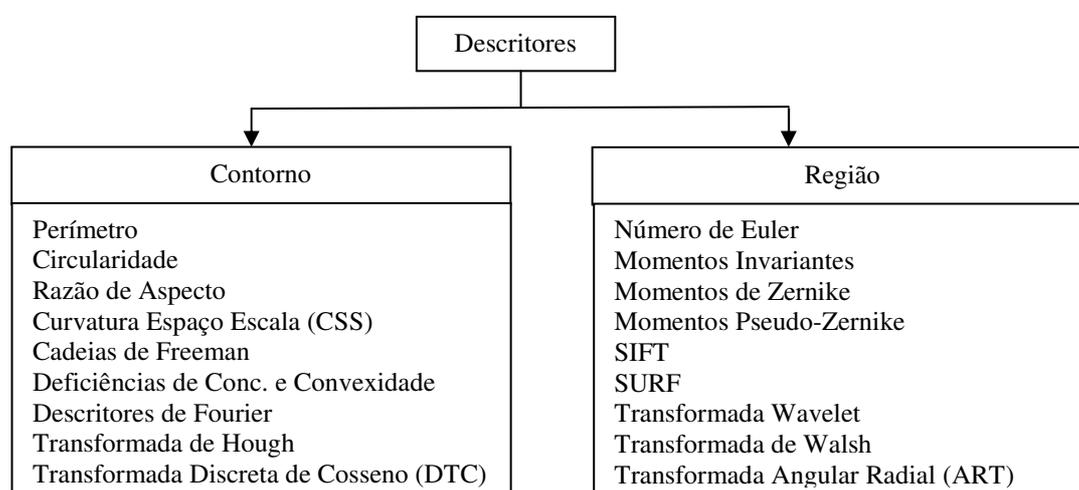


Figura 2.5: Descritores da forma

A abordagem baseada em contorno é muito utilizada na literatura, conforme apresentado na revisão teórica do Capítulo 3. Entretanto, esta abordagem pode sofrer efeitos negativos na presença de ruídos.

Por outro lado, a abordagem baseada em região pode ser mais robusta na recuperação por considerar a informação de toda a forma ao invés de considerar apenas a informação do contorno, podendo ser utilizada em aplicações de propósito geral como, por exemplo, reconhecimento de faces, análise de imagens médicas, detecção de objetos, entre outras [ZHA04].

Dentro destas abordagens, pode-se ainda classificar os descritores como sendo globais ou estruturais. Os descritores globais consideram o contorno ou a região da forma como um todo, sem realizar divisões em subpartes. Os descritores estruturais

dividem o contorno ou a região em segmentos e possibilitam a comparação parcial de formas, conseguindo, portanto, suportar oclusão de formas. Entretanto, eles são mais complexos computacionalmente, especialmente para o cálculo de similaridades [PEN09].

Para Kim e Kim [KIM00], os descritores de forma baseados em contorno podem não ser adequados para formas complexas que consistem em várias regiões disjuntas tais como MRVs ou logotipos, emblemas, *cliparts* e caracteres [FRE07] ou outras formas extraídas de cenários naturais. Por exemplo, o contorno pode ser alterado drasticamente se houver uma pequena abertura ou um objeto tocar em objetos vizinhos.

No entanto, descritores de forma baseados em regiões, como Momentos Invariantes, são mais confiáveis para as formas que têm contornos complexos, porque elas dependem não só dos pixels do contorno, mas também de todos os pixels que constituem as formas. Deve-se considerar que Momentos Invariantes são sensíveis a ruídos [KIM00].

Hung et al. [HUN04] argumentam que alguns autores indicam que as MRVs tanto são similares no que se refere ao contorno quanto em relação à região. Assim, a medida de similaridade deve considerar o contorno do objeto de uma MRV, bem como a região do interior do objeto. Para Gonzales e Woods [GON08] é uma prática comum o uso combinado dos descritores de contorno com os descritores de região.

2.3.1. Descritores Baseados em Contorno

Para Erpen [ERP04] estes métodos representam as propriedades da forma pelo seu esboço (contorno). Os descritores desta categoria são utilizados, segundo a especificação MPEG-7 (*Moving Picture Experts Group*), em aplicações nas quais espera-se que exista uma alta variabilidade da forma devido a deformações no objeto. A Figura 2.6, apresenta algumas imagens nas quais os descritores baseados em contorno são aplicáveis.

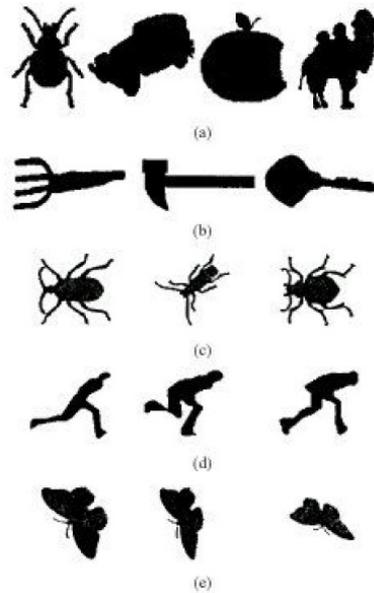


Figura 2.6: Exemplos de formas onde se aplicam descritores baseados em contorno

Fonte: [ERP04]

Observando a Figura 2.6, pode-se notar que:

- Os objetos contidos na Figura 2.6 (a) são distintos, pois as características da forma estão contidas no contorno;
- Os objetos apresentados na Figura 2.6 (b) possuem uma distribuição dos pixels similares por região, mas diferentes propriedades de contorno;
- A Figura 2.6 (c) apresenta formas que são similares para a percepção humana. Porém, existe uma variabilidade significativa das imagens, sendo que estas devem ser consideradas similares;
- Na Figura 2.6 (d) deformações não-rígidas são apresentadas;
- E distorções no contorno ocorridas por transformações perspectivas, comuns em imagens e vídeos são apresentadas na Figura 2.6 (e).

Um descritor de contorno eficiente deve ser capaz de distinguir dentro de cada grupo as imagens da Figura 2.6 (a) e (b) e, ainda considerar similares as imagens da Figura 2.6 (c) (d) e (e).

Na literatura pesquisada são apresentados diversos descritores da forma por contorno [CON07] [GON08], tal como destacado na Figura 2.5. Nos experimentos realizados neste trabalho, foram utilizados dois descritores da forma por contorno:

Cadeias de Freeman e Concavidades/Convexidades (*Convex Deficiencies*) [PAR97], apresentados nas seções a seguir.

2.3.1.1 Cadeias de Freeman

O código de cadeia de Freeman resulta em uma representação fiel do contorno de imagens utilizando apenas um caractere por ponto, que indica a direção como identificação do próximo pixel do contorno, conforme Figura 2.7.

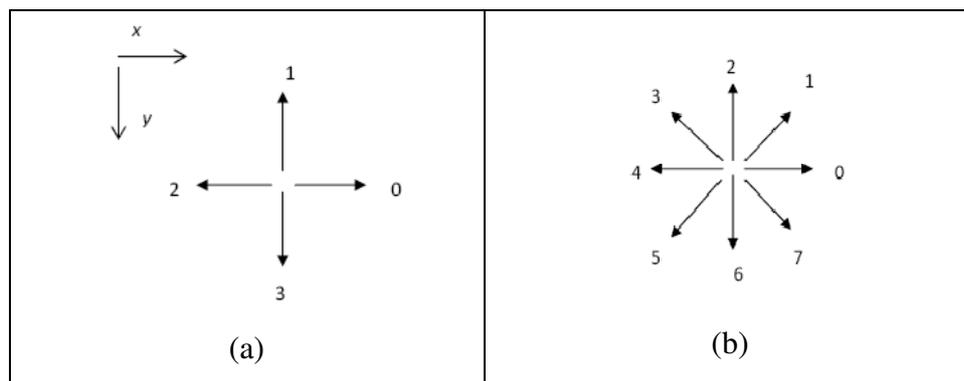


Figura 2.7: (a) Cadeia de Freeman 4 direções; (b) Cadeia de Freeman 8 direções.

As Cadeias de Freeman são uma das representações da forma usadas para extrair um contorno através de uma sequência conectada de segmentos de reta de comprimento e direção específicos [JUS09].

Um código de cadeia é gerado pela sequência de contornos de um objeto seguindo a direção dos ponteiros do relógio, e atribui uma direção aos segmentos ligando cada par de pixels.

Inicialmente, deve-se determinar o pixel inicial no contorno do objeto. A seguir faz-se uma verificação dos pixels adjacentes a uma das oito direções possíveis em torno do pixel atual, com o objetivo de encontrar o próximo pixel. A busca ao próximo pixel deve ser no sentido horário, iniciando na posição 0 (zero) e terminando na posição 1 (um), conforme Figura 2.7-b. Ao analisar cada um dos oito pixels adjacentes, será encontrado pelo menos um que também é um pixel pertencente ao contorno. O código determinado para o par de pixels verificado (atual e novo) segue conforme indicado na Figura 2.8. Exemplificando, na Figura 2.8, podemos verificar o pixel inicial, e ao lado direito dele temos outro pixel, neste caso a codificação para este par será 0 (zero). Esta regra também vale para o par segundo e terceiro pixel, o quarto pixel está posicionado

abaixo e à direita, para este par (terceiro e quarto pixel) a codificação será 7 (sete). A regra segue até que seja alcançado o pixel da posição inicial.

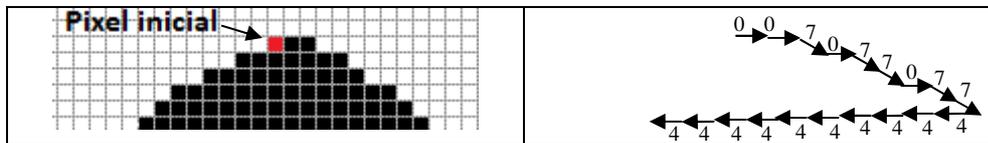


Figura 2.8: Exemplo parcial codificação Cadeias de Freeman

Para o exemplo apresentado na Figura 2.8, a codificação final será 007077077444444444444444444411010101.

Segundo Jusoh e Zain [JUS09] este método pode apresentar problemas por duas razões principais:

1. A cadeia resultante de códigos tende a ser bastante longa;
2. Qualquer pequena perturbação ao longo do contorno devido a ruído ou segmentação imperfeita, causa alterações na codificação que pode não estar relacionada à forma do contorno.

De acordo com Jusoh e Zain [JUS09] os códigos de cadeia oferecem uma compressão sem perdas e a preservação de todas as informações topológicas e morfológicas que trazem benefício em termos de rapidez e eficácia para a análise de padrões de linha.

Neste trabalho, foram realizados experimentos utilizando a análise do contorno das imagens por meio da codificação das cadeias de Freeman [FRE74].

2.3.1.2 Concavidades e Convexidades

Para evidenciar as propriedades topológicas e geométricas das formas, pode-se realizar uma análise das partes côncavas e convexas das MRVs [PAR97]. Para Freitas et al. [FRE01] existem dois tipos de métodos que permitem extrair as concavidades, a saber:

- métodos com base na rotulação das áreas brancas das imagens ou do fundo das imagens (*background*),
- métodos com base no tratamento do contorno das formas.

O método composto pela rotulação dos pixels do fundo da imagem das MRVs foi utilizado no presente estudo de maneira a extrair para cada imagem um vetor de características (*feature vector*).

O vetor de características é composto baseando-se na rotulação de cada pixel do fundo da imagem, mapeando 24 situações distintas. Dois tipos de verificação são considerados, a saber:

- Verificação dos 4 vizinhos, norte-sul-leste-oeste (NSLO) para cada pixel branco na MRV
- Verificação de falsos laços com a possibilidade de escape em 4 direções de sondagem para os pixels determinados como fechados (laço) para as 4 direções principais (NSLO).

A Tabela 2.1 apresenta a rotulação do vetor de características da MRV, contendo 24 rótulos com as verificações de direção, sondagem de regiões de escape e laço fechado.

É importante ressaltar que para utilizar a rotulação conforme Tabela 2.1, é necessário que as etapas de pré-processamento (limiarização [OTS79], e *bouding box*) sejam aplicadas nas imagens de MRVs. Após o pré-processamento, o algoritmo de rotulação avalia todos os pixels contidos na imagem. Se o pixel for preto seu rótulo será ".", senão verificam-se as direções NSLO. Considera-se caminho "fechado" se um pixel preto for encontrado na direção que está sendo realizada a verificação, e se não existirem pixels pretos até a região da borda considera-se como caminho "aberto". No entanto, encontrando um pixel preto em uma direção específica verifica-se se este pixel encontra-se em um laço fechado (caso as posição NSLO encontrem pixels pretos), se isso ocorrer seu rótulo será "8". Ainda, verifica-se a possibilidade de haver alguma área de "escape", nas regiões acima ou abaixo, à direita ou à esquerda, conforme Figuras 2.9 e 2.10.

A Figura 2.9 exemplifica a rotulação dos pixels do fundo da imagem de duas MRVs, podendo-se verificar facilmente os pixels pretos (1o. plano) representados pelo símbolo (.) e pixels brancos rotulados (2o. plano) conforme a Tabela 2.1. Na Figura 2.10, são apresentadas regiões de escapes e laços fechados.

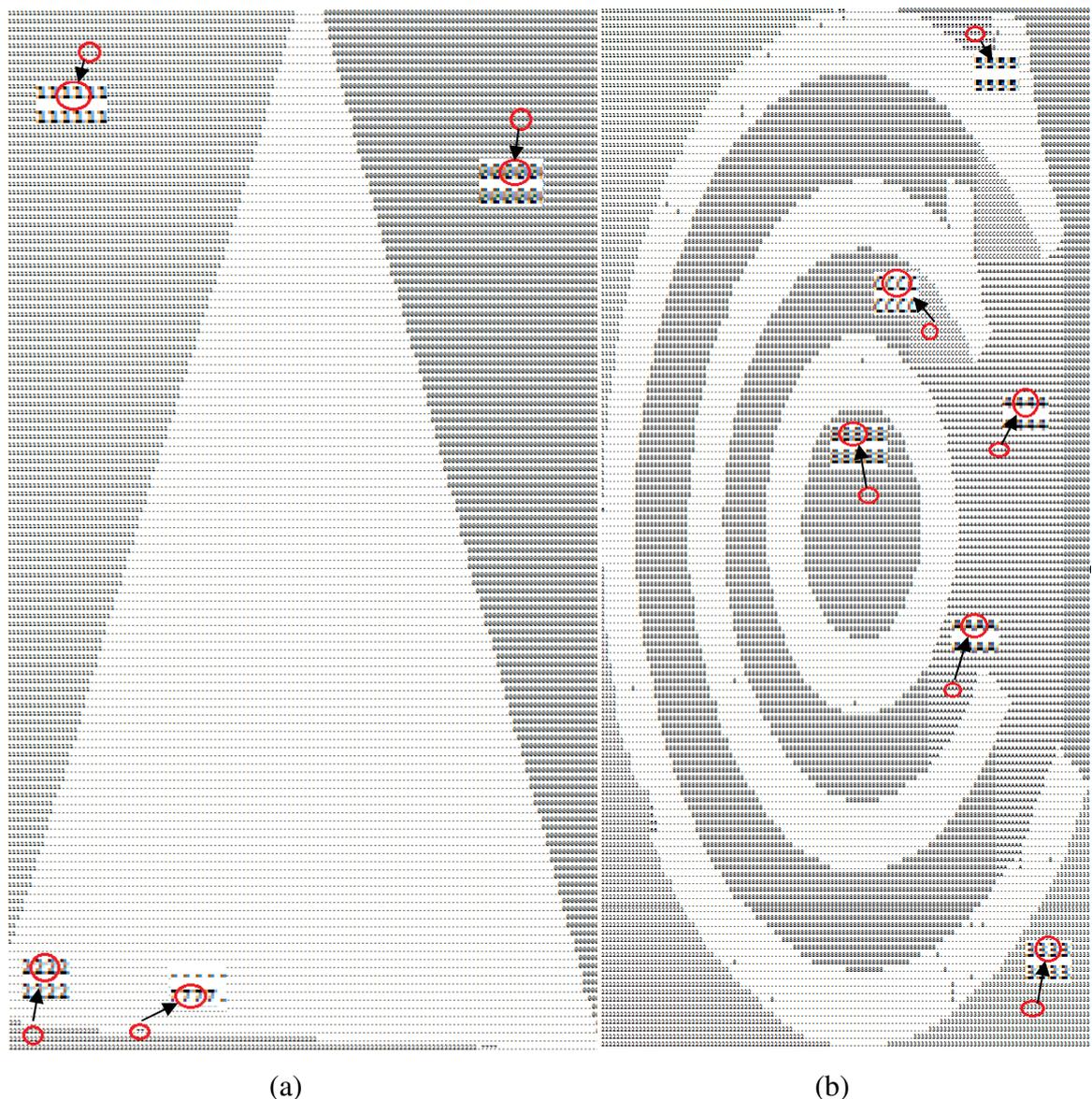


Figura 2.9: Rotulação dos pixels do fundo de duas imagens contidas na base de dados

Observa-se na Figura 2.9-a, que são encontrados rótulos 0, 1, 2 e 7. Conforme a Tabela 2.1, tais rótulos representam respectivamente: pixel com abertura acima e à direita (0), pixel com abertura acima e à esquerda (1), pixel com abertura abaixo e à esquerda (2) e pixel com abertura abaixo (7). Para a Figura 2.9-b os rótulos em destaque são 5, C, 4, 8, A e 3. Novamente, de acordo com a Tabela 2.1, tem-se que os rótulos representam respectivamente: pixel com abertura acima (5), pixel com escape para direita e abaixo (C), pixel com abertura à direita (4), pixel contido em laço fechado (8), pixel com escape à direita e acima (A) e pixel com abertura à direita e abaixo.

A Figura 2.10, apresenta a verificação de regiões de escapes e laços fechados, os pixels em destaque são os mesmos apresentados em destaque na Figura 2.9-b.

Tabela 2.1: Rótulos do Vetor de Características

Rótulo	Tipo	Busca
0	Aberto à direita e acima	(NSLO)
1	Aberto à esquerda e acima	(NSLO)
2	Aberto à esquerda e abaixo	(NSLO)
3	Aberto à direita e abaixo	(NSLO)
4	Aberto à direita	(NSLO)
5	Aberto acima	(NSLO)
6	Aberto à esquerda	(NSLO)
7	Aberto abaixo	(NSLO)
8	Laço fechado	(NSLO)
A	Escape para direita e acima	Sondagem
B	Escape para esquerda e acima	Sondagem
C	Escape para direita e abaixo	Sondagem
D	Escape para esquerda e abaixo	Sondagem
E	Escape acima e direita	Sondagem
F	Escape acima e esquerda	Sondagem
G	Escapa abaixo e direita	Sondagem
H	Escape abaixo e esquerda	Sondagem
J	Aberto esquerda/direita	(NSLO)
K	Aberto acima/abaixo	(NSLO)
L	Fechado abaixo	(NSLO)
M	Fechado acima	(NSLO)
N	Fechado direita	(NSLO)
O	Fechado esquerda	(NSLO)
.	Pixel preto	

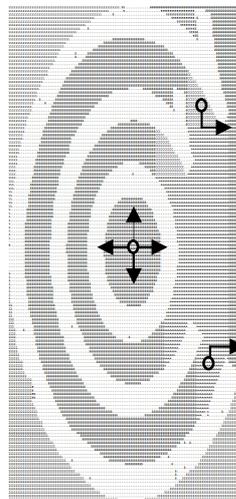


Figura 2.10: Verificação de Escape e Laço Fechado

Observa-se na Figura 2.10, que dois pixels encontraram região de escape à direita. No entanto, o pixel no centro da Figura 2.10 está contido em um laço fechado (rotulação = 8).

2.3.2. Descritores Baseados em Região

O descritores baseados em região podem descrever objetos complexos que consistem em várias regiões desconexas, como também objetos simples com ou sem furos/laços.

Na Figura 2.11, observa-se que as imagens são similares em seus próprios conjuntos, e dissimilares em relação aos demais.

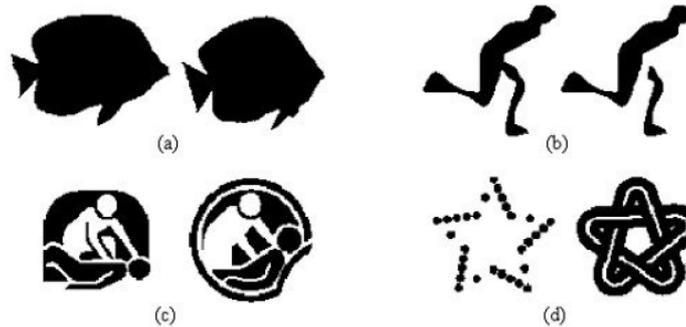


Figura 2.11: Exemplos de forma em que aplicam-se descritores baseados em região

Fonte: Adaptada de [ERP04]

Na literatura são apresentados diversos descritores da forma por região [CON07] [GON08], tal como destacado na Figura 2.5. Para os experimentos realizados neste trabalho foram utilizados 2 (dois) descritores da forma por região: Momentos Invariantes de Hu e SIFT (*Scale Invariant Feature Transform*), apresentados nas seções a seguir.

2.3.2.1 Momentos Invariantes de Hu

A extração de Momentos é uma técnica robusta para decomposição de uma imagem em um conjunto finito de características invariantes quanto à escala, posição e orientação. Em termos práticos, a utilização de Momentos para o reconhecimento de imagens requer a seleção de um subconjunto de valores de momento que contenha informações suficientes para caracterizar unicamente cada imagem [ROD08].

Os momentos de imagem mais utilizados são os momentos regulares definidos a partir da Equação 2.1:

$$m_{pq} = \sum_{x=1}^{nx} \sum_{y=1}^{ny} x^p y^q f(x, y) \quad (2.1)$$

Na Equação 2.1, m_{pq} representam o momento de ordem $(p+q)$ da função intensidade $f(x,y)$ [CHA11b]. Onde n_x e n_y representam respectivamente a largura e a altura da imagem digital. Considerando uma imagem binária esta terá valores resultantes para a função $f(x,y)$ iguais a 0 ou 1.

Os momentos regulares definem medidas importantes sobre os objetos de interesse, e úteis na identificação, por exemplo, os momentos regulares de ordem 0 e 1 são usados para o cálculo do baricentro ou centro de massa do objeto, conforme a Equação 2.2:

$$x_c = \frac{m_{10}}{m_{00}} \quad y_c = \frac{m_{01}}{m_{00}} \quad (2.2)$$

Com a informação obtida dos baricentros obtemos os momentos centrais, conforme apresentados na Equação 2.3:

$$\mu_{pq} = \sum_1^{n_x} \sum_1^{n_y} (x - x_c)^p (y - y_c)^q f(x,y) \quad (2.3)$$

Os momentos centrais normalizados representados por n_{pq} , são definidos na Equação 2.4 [CHA11b]:

$$n_{pq} = \frac{\mu_{pq}}{m_{00}^{\left(\frac{p+q}{2}+1\right)}} \quad (2.4)$$

Uma combinação de momentos que são invariantes à escala, rotação e translação é apresentada por Hu [HU_62] [CHA11b]. Estes momentos são representados por sete equações chamadas de momentos de Hu ou momentos invariantes, que são apresentados nas Equações 2.5, 2.6, 2.7, 2.8, 2.9, 2.10 e 2.11:

$$I_1 = n_{20} + n_{02} \quad (2.5)$$

$$I_2 = (n_{20} - n_{02})^2 + (2n_{11})^2 \quad (2.6)$$

$$I_3 = (n_{30} - 3n_{12})^2 + (3n_{21} - n_{03})^2 \quad (2.7)$$

$$I_4 = (n_{30} + n_{12})^2 + (n_{21} + n_{03})^2 \quad (2.8)$$

$$I_5 = (n_{30} - 3n_{12})(n_{30} + n_{12})[(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2] \quad (2.9)$$

$$+ (3n_{21} - n_{03})(n_{21} + n_{03})[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2]$$

$$I_6 = (n_{20} - n_{02}) [(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] \quad (2.10)$$

$$+ 4n_{11}(n_{30} + n_{12})(n_{21} + n_{03})$$

$$I_7 = (3n_{21} - n_{03})(n_{30} + n_{12})[(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2]$$

$$+ (n_{30} - 3n_{12})(n_{21} + n_{03})[3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] \quad (2.11)$$

Os Momentos Invariantes de Hu destacam-se por serem bem sucedidos em muitas aplicações, além de incorporarem invariâncias à escala, rotação e translação [ROD08]. Os trabalhos desenvolvidos por [JAI98][CHA99] utilizaram Momentos Invariantes de Hu para extrair informações da forma e auxiliar na recuperação de MRV.

2.3.2.2 Scale Invariant Feature Transform (SIFT)

SIFT é uma técnica de visão computacional, desenvolvida por David Lowe [LOW04] que permite detectar e extrair descritores locais, sendo um descritor invariante à escala, rotação, translação e parcialmente invariante à iluminação [GON10].

O algoritmo SIFT possui duas etapas distintas, detecção e descrição. A detecção é realizada por meio do cálculo da diferença Gaussiana e a descrição utiliza histogramas de gradientes orientados para verificar a vizinhança dos pontos de interesse (*keypoints*). SIFT é composto de quatro etapas principais, sendo que as duas primeiras fazem parte do detector e as duas últimas descrevem a formação dos descritores. Tais etapas são descritas a seguir.

a) Detecção de extremos

Esta etapa busca pontos que são invariantes à mudança de escala, detectando pontos com a câmera próxima ou distante do objeto de interesse. Isto é alcançado verificando características estáveis em diferentes escalas, utilizando uma função Gaussiana, chamada de função de espaço-escala [GON10].

Uma imagem $I(x,y)$, passa a ser definida por $L(x, y, \sigma)$, no espaço-escala. Esta função é produzida pela convolução de uma função Gaussiana, $G(x, y, \sigma)$, com a imagem $I(x,y)$, conforme Equações 2.12 e 2.13:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x,y) \quad (2.12)$$

Sendo:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-((x^2 + y^2)/2\sigma^2)} \quad (2.13)$$

O filtro é variável à escala através do parâmetro σ . A eficiência da busca por pontos-chave é aumentada com a aplicação da função DoG (*Difference of Gaussian*), conforme Equação 2.14:

$$DoG = G(x, y, k\sigma) - G(x, y, \sigma) \quad (2.14)$$

Para detectar pontos-chave estáveis no espaço-escala, é proposto por Lowe [LOW04] usar a convolução de uma imagem utilizando filtro DoG apresentado nas Equações 2.15 e 2.16:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x,y) \quad (2.15)$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (2.16)$$

Subtraindo as imagens borradas, calcula-se a função DoG, utilizando um filtro Gaussiano em escalas σ e $k\sigma$. Essa função Gaussiana obtém amostras das imagens nas quais os ruídos são minimizados e características são realçadas. Com a variação de σ é possível encontrar características em diferentes escalas [GON10].

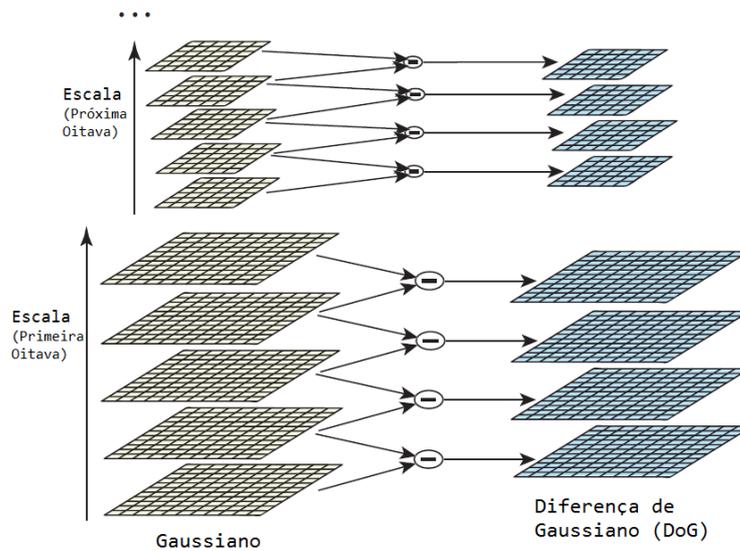


Figura 2.12: Diferenças Gaussianas DoG

Fonte: Adaptada de [LOW04]

Na Figura 2.12, pode-se observar que para cada oitava no espaço-escala, a imagem original é repetidamente convoluída com a Gaussiana para produzir um conjunto de imagens no espaço-escala apresentados à esquerda. Imagens em escalas adjacentes são subtraídas para produzir imagens da Diferença-Gaussiana (DoG) à direita. Uma vez processada a oitava, a resolução da imagem é reduzida (*downsample*) tornando cada segundo pixel da imagem no centro da oitava, gerando uma nova oitava e assim retornando ao passo inicial [LOW04].

Deste ponto em diante, é realizada a detecção de extremos em cada intervalo de cada oitava da pirâmide DoG. Um extremo define-se como qualquer valor no DoG maior que todos os seus vizinhos no espaço-escala.

Os extremos são determinados pelos valores de máximo ou mínimo locais para cada $D(x, y, \sigma)$. Estes podem ser obtidos comparando-se a intensidade de cada ponto com a intensidade de seus oito vizinhos na sua escala, com os nove vizinhos na escala superior e inferior [GON10], conforme Figura 2.13.

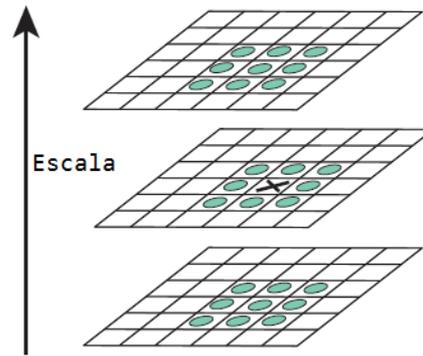


Figura 2.13: Máximo e Mínimo da DoG

Fonte: Adaptada de [LOW04]

A próxima etapa verifica a localização dos pontos-chave, descartando possíveis pontos instáveis.

b) Precisão na localização de pontos-chave

Os pontos detectados como extremos são candidatos a ponto-chave. Uma vez que um ponto-chave é encontrado, é necessário realizar um detalhamento de informações como: localização, escala e taxa de curvatura. Estas informações permitem descartar pontos com baixo contraste (sensíveis a ruído) ou mal localizados [LOW99] [GON10].

Brow e Lowe [BRO02], propõem uma melhoria no método inicial proposto por [LOW99]. Este método aprimorado consiste em ajustar funções quadráticas 3D do ponto de amostragem local para determinar a localização interpolada do máximo. Isto é realizado utilizando a expansão de Taylor da função DoG aplicada à imagem, $D(x, y, \sigma)$ deslocada de modo que a origem da expansão esteja localizada no ponto de amostragem. Conforme apresentado nas Equações 2.17 e 2.18.

$$D(x) = D + \frac{\partial D}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \dots \quad (2.17)$$

Sendo D e suas derivadas avaliadas no ponto de amostragem e $x = (x, y, \sigma)^T$ que representa o deslocamento deste ponto. A localização do extremo, \hat{x} , é determinada calculando a derivada de $D(\hat{x})$ em relação à \hat{x} e igualando a zero (Equação 2.18).

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial \bar{x}} \quad (2.18)$$

O valor da função $D(\hat{x})$, é utilizada para rejeitar extremos com baixo contraste. Isso pode ser obtido, substituindo Equação 2.18 na Equação 2.17 e obtendo a Equação 2.19:

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (2.19)$$

Para Brown e Lowe [BRO02] é aconselhável rejeitar valores para $|D(\hat{x})|$ inferiores a um limiar. Estabelece-se o valor 0.03 desde que os pixels em nível de cinza estejam normalizados entre 0 e 1.

Para a estabilidade, não é suficiente rejeitar somente pontos com baixo contraste [BEL06]. A função DoG possui resposta "forte" junto a bordas, mesmo que a localização ao longo da borda seja mal determinada e desta forma, instável em relação a ruído. Pontos em bordas poderiam ser pontos de interesse, porém não é desejável [LOW04]. Para a eliminação destes pontos-chave, utiliza-se a matriz Hessiana 2×2 , H , calculada na localização e escala dos pontos-chave na função D , conforme Equação 2.20.

$$H(x, y) = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (2.20)$$

Sendo:

- D_{xy} é a derivada de $D(x, y, \sigma)$ na localização e escala em relação a x e y ;
- D_{xx} é a derivada segunda em relação a x ;
- D_{yy} é a derivada segunda em relação a y .

Estima-se as derivadas através das diferenças entre pontos vizinhos à localização e escala, conforme Equações 2.21, 2.22 e 2.23:

$$D_{xx} = D(x + 1, y, \sigma) - 2D(x, y, \sigma) + D(x - 1, y, \sigma) \quad (2.21)$$

$$D_{yy} = D(x, y + 1, \sigma) - 2D(x, y, \sigma) + D(x, y - 1, \sigma) \quad (2.22)$$

$$D_{xy} = \left(\frac{D(x - 1, y + 1, \sigma) - D(x + 1, y + 1, \sigma)}{+ D(x + 1, y - 1, \sigma) - D(x - 1, y - 1, \sigma)} \right) / 4 \quad (2.23)$$

O autovalor da matriz H com a maior magnitude determina α , e β o de menor. Calcula-se a soma dos autovalores pelo traço de H e o produto pelo seu determinante. Conforme Equações 2.24 e 2.25.

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (2.24)$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (2.25)$$

Nos casos onde o determinante é negativo, as curvaturas possuem sinais diferentes, e o ponto é descartado, não sendo considerado um extremo [LOW04]. Sendo r a razão entre o autovalor de maior magnitude e o de menor, assim $\alpha = r\beta$ (Equação 2.26).

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r\beta+\beta)^2}{r\beta^2} = \frac{(r+1)^2}{r} \quad (2.26)$$

A Equação 2.26 depende somente da razão entre os autovalores. O valor de $(r+1)/r$ é mínimo em que os dois autovalores são iguais e cresce com r . Para verificar se razão de curvatura está abaixo de um limiar, r , tem-se a Equação 2.27.

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (2.27)$$

Lowe [LOW04] propõe o uso de $r = 10$, assim eliminam-se pontos-chaves que não são estáveis, apesar de estarem próximos de extremidades.

c) Orientação dos Descritores

Atribuindo uma orientação consistente para cada ponto-chave, com base nas propriedades locais da imagem, o descritor do ponto-chave pode ser representado em relação e esta orientação e, portanto alcançar invariância à rotação [LOW04].

Para cada amostragem da imagem na escala, $L(x, y, \alpha)$, calcula-se a magnitude $m(x, y)$ e orientação $\theta(x, y)$ do gradiente utilizando a diferença de pixels, conforme Equações 2.28 e 2.29.

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (2.28)$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y))) \quad (2.29)$$

Um histograma das orientações dos pixels vizinhos ao ponto-chave é construído. O histograma possui 36 regiões, considerando-se todas as orientações possíveis (0 a 2π), conforme Figura 2.14.

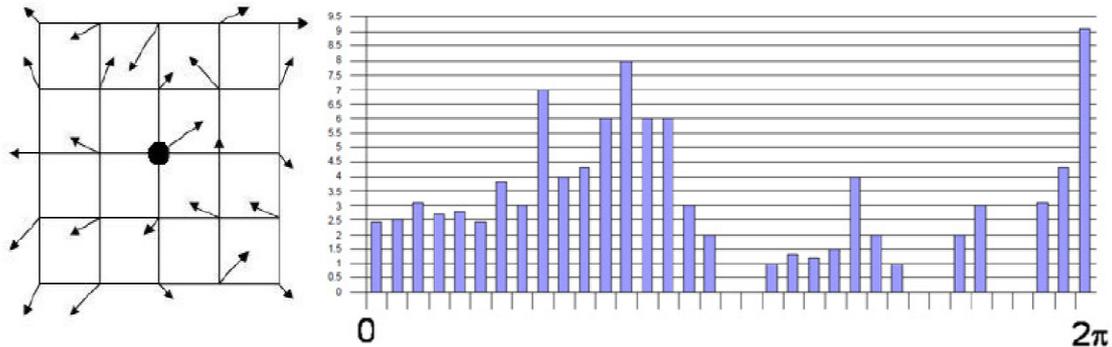


Figura 2.14: Histograma de orientação dos pontos-chave

Fonte: Adaptada de [GON10]

Cada ponto vizinho ao ponto-chave é adicionado ao histograma considerando os pesos da magnitude $m(x,y)$ e da janela Gaussiana circular com α' igual a 1,5 vezes maior que a escala do ponto-chave [LOW04]. Essa janela é definida pela Equação 2.30:

$$g(\Delta x, \Delta y, \alpha') = \frac{1}{2\pi\sigma'^2} e^{-(\Delta x^2 + \Delta y^2)/2\sigma'^2} \quad (2.30)$$

Sendo Δx e Δy as distâncias entre cada ponto verificado e o ponto-chave. Os pesos calculados para cada ponto da vizinhança (x,y) são atualizados conforme Equação 2.31:

$$h'_\theta = h_\theta + \alpha m(x,y) \cdot g(\Delta x, \Delta y, \alpha') \quad (2.31)$$

$$\alpha = \begin{cases} \frac{d}{i}, & d < i \\ 0, & d > i \end{cases}$$

Sendo h'_θ a atualização de h_θ , e d é a distância absoluta em graus entre a orientação do ponto e o θ discretizado, e i é o intervalo em graus entre os θ 's discretizados.

Os picos no histograma de orientações correspondem às direções dominantes dos gradientes locais. Também são considerados os picos que correspondem ao mínimo de 80% do valor deste máximo. Então, um mesmo ponto-chave poderá ter mais de uma orientação associada.

O pico do histograma é utilizado para definir a sua orientação. Havendo múltiplos picos de elevada amplitude, o ponto-chave receberá múltiplas orientações, tornando-se mais estável para futura identificação. Uma parábola é utilizada para interpolar os três valores do histograma mais próximos ao pico, de forma a obter uma melhor exatidão na sua orientação. Ressalta-se que cada ponto-chave possui quatro dimensões: posição x e y ; magnitude e orientação.

d) Descritor Local

O descritor do ponto-chave é criado computando-se as magnitudes e orientações dos gradientes ao redor da localização do ponto-chave. A Figura 2.15 descreve este procedimento, no qual os gradientes são representados pelas pequenas setas em cada amostra da localização. São definidas $n \times n$ regiões de amostragem com $k \times k$ pixels cada ao redor da localização do ponto-chave.

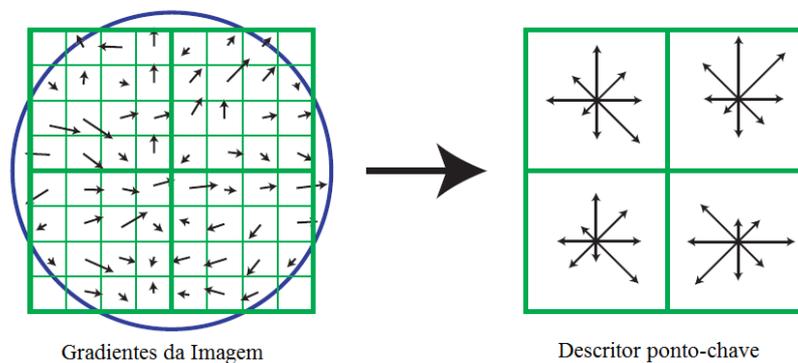


Figura 2.15: Construção descritor do ponto-chave 2x2

Fonte: Adaptada de [LOW04]

Para dar peso à magnitude do gradiente em cada ponto vizinho ao ponto-chave, utiliza-se uma função Gaussiana, com uma janela de suavização de escala σ igual à metade da largura da janela do descritor. Isso evita mudanças súbitas do descritor a pequenas mudanças na posição da janela, e diminui a ênfase nos gradientes longe do centro do descritor, que são mais afetados por erros.

Após a suavização dos gradientes, o descritor consiste em um vetor com valores do histograma. No exemplo da Figura 2.15, o histograma tem oito valores de orientação,

cada um criado ao longo em uma janela de 4x4 pixels. O vetor de características resultante possui 128 elementos com uma janela de 16x16 pixels.

Para que o descritor tenha invariância à iluminação, este é normalizado. Pois duas imagens de um mesmo objeto podem possuir variações de luminosidade que modifiquem os descritores obtidos.

Para cada ponto-chave um descritor é construído, e cada imagem é representada por diversos descritores conforme a quantidade de pontos-chave identificados. Obtém-se um conjunto de descritores robustos, que podem ser utilizados para realizar correspondência entre imagens contidas em bases de dados.

Os detalhes aqui apresentados não pretendem esgotar o estudo dos descritores SIFT e mais informações podem ser encontradas em [LOW04].

2.3.3. Descritores da Forma Aplicados em Sistemas de Recuperação de MRV

A seguir são apresentados trabalhos que buscam auxiliar na recuperação de MRVs. A Tabela 2.2 apresenta um resumo dos trabalhos relacionados à extração de primitivas no contexto de recuperação de MRVs.

Tabela 2.2: Extração de primitivas

(continua)	
Autor	Descritores da Forma
[JAI98]	Histograma de detecção de bordas. Momentos invariantes.
[EAK98]	<ul style="list-style-type: none"> • <i>área relativa</i> • <i>right-angleness</i> • <i>directedness</i> • <i>straightness</i> • nitidez • relação de aspecto • circularidade • transparência • complexidade
[ALW99]	<ul style="list-style-type: none"> • co-linearismo • co-curvilinearismo • paralelismo • proximidade de ponto final (<i>end-point proximity</i>)
[CHA99]	Descritor de Fourier Momentos invariantes de Hu, No. de Euler, Excentricidade e Circularidade
[KIM00]	Momentos de Zernike
[HUN04]	Assinatura do contorno ART (Transformada Angular Radial)

(conclusão)

Autor	Descritores da Forma
[HUS06]	Triangularidade, retangularidade e elipticidade. Área, relação de aspecto (<i>aspect ratio</i>), circularidade e convexidade. Descritor de Fourier
[JIA06]	Momentos de Zernike Transformada de Hough (segmentos de linhas, arcos, linhas paralelas, círculos concêntricos e polígonos)
[ABE06]	Circularidade, diâmetro equivalente, taxa de ocupação e <i>fillet</i> gradiente. Cálculo de curvatura. Relação entre a Área e a Menos distância entre 2 componentes da imagem.
[BAG07]	<i>Scale Invariant Feature Transform (SIFT)</i>
[ANU13]	Momentos de Zernike, <i>Edge Gradient Co-occurrence Matrix</i>
[SAH13]	<i>Scale Invariant Feature Transform (SIFT)</i>
[SHA14]	Momentos invariantes, Transformada Decomposição Valores Singular (SVD - <i>Singular Value Decomposition</i>), Transformada Discreta de Cosseno 2D (DCT - <i>Discrete Cosine Transform</i>)

Observa-se na Tabela 2.2 que diversos autores utilizam as duas abordagens dos descritores da forma, fazendo uma composição dos descritores de contorno e região.

Os sistemas de recuperação de MRVs utilizam descritores da forma para extrair as características das imagens de MRVs. A extração de características codifica o conteúdo de uma imagem em um vetor de características, enquanto uma função calcula a distância e permite definir a similaridade entre dois vetores de características e, conseqüentemente, entre duas imagens.

Neste trabalho, os experimentos realizados utilizaram quatro tipos de extratores de características diferentes, extraídas para o mesmo conjunto de imagens. Foram utilizados: Momentos Invariantes de Hu (descritor por região); SIFT (descritor por região); Cadeia de Freeman (descritor de contorno); Concavidade e Convexidade [PAR97] [FRE07](descritor de contorno).

Durante o processo de reconhecimento de formas, a visão humana suporta diferentes distorções e fenômenos associados à forma. Assim, descritores de forma devem ser capazes de reconhecer imagens semelhantes mesmo que existam objetos rotacionados, transladados, escalonados, com ruído, distorcidos ou parcialmente encobertos [ZHA04].

Para Eakins et al. [EAK98] é fundamental que modelos de percepção humana sejam incorporadas nos sistemas de recuperação de MRVs. Para tal funcionalidade ser utilizada, uma alternativa é incorporar os princípios da Gestalt a exemplo de alguns autores [EAK98][JAI98] [JIA06][ABE06][HIR09] quando o escopo são MRVs, tendo

sido estes princípios aplicados com sucesso em reconhecimento de palavras e caracteres manuscritos [FRE04] [FRE07] [FRE07a] [FRE08] [OLI09] [HIR11].

2.4. Princípios da Gestalt

O poder das informações visuais de ativar nossas emoções vem sendo observado desde a Antiguidade: "A mente desperta mais lentamente por meio do ouvido do que do olho", disse Horácio em sua Arte Poética ao comparar as apresentações realizadas no palco em relação às narrativas verbais [GOM12].

Atualmente mais de seis mil idiomas são utilizadas no mundo todo, conforme informações do compêndio Ethnologue⁷, considerado o maior inventário de línguas. No entanto a linguagem visual é universal.

A Gestalt é uma Escola de Psicologia Experimental alemã, que atuou no campo da teoria da forma, com contribuições relevantes aos estudos da percepção, linguagem, aprendizagem, memória, inteligência, entre outros. Tais estudos foram realizados por um grupo de pesquisadores alemães, Max Wertheimer (1880-1943), Wolfgang Köhler (1887-1967) e Kurt Koffka (1886-1941).

Segundo a Gestalt, a primeira sensação ao observar algo já é da forma, pois vemos relações entre os elementos e não partes isoladas. A sensação é global. Portanto, enxerga-se o todo e não partes dele [FRE02] [GOM04].

Os princípios da Gestalt regem há tempos a criação de MRVs ou imagens comerciais utilizadas em publicidade e propaganda. Esses princípios ajudam as pessoas a assimilarem informações e a entenderem as mensagens que são transmitidas. Sendo uma solução sutil e criativa para a concepção de uma marca visual (MRV) tornando-a simples e memorável.

Apresentam-se a seguir os princípios da Gestalt com os quais pode-se detalhar estas forças de percepção [GOM04][GOM12]:

Princípio da Proximidade: este princípio descreve que elementos próximos tendem a serem vistos como um só grupo, e por conseguinte constituem uma unidade.

Uma das estratégias que nosso cérebro usa é formar grupos de objetos que possuam uma característica em comum, seja por forma, cor, tamanho, textura, brilho,

⁷ <http://www.ethnologue.com/>

peso, direção, entre outros. Observando a Figura 2.16, é possível notam-se 16 círculos ou 4 grupos distintos contendo círculos?



Figura 2.16: Princípio da Proximidade

O fenômeno observado na Figura 2.16 dá-se pelo agrupamento de formas, através da proximidade dos objetos. Este princípio é claramente visto na MRV da Unilever, apresentada na Figura 2.17-a, pois não percebemos os 24 elementos que compõem a MRV separadamente, mas tendemos a agrupá-los pela forma que estão organizados e pela distância um dos outros, formando a letra “U”. Outro exemplo, na Figura 2.17-b a MRV do restaurante Four Season apresenta elementos que são facilmente agrupados considerando os tamanhos dos objetos.



(a)



(b)

Figura 2.17: Princípio da Proximidade em MRVs

Princípio da Similaridade: este princípio define objetos que compartilham características visuais parecidas tendem a se agrupar. Isso pode acontecer com a forma, cor, textura, direção e até mesmo a sensação de massa das unidades em questão. Os agrupamentos por similaridade, podem ser observados na Figura 2.18.

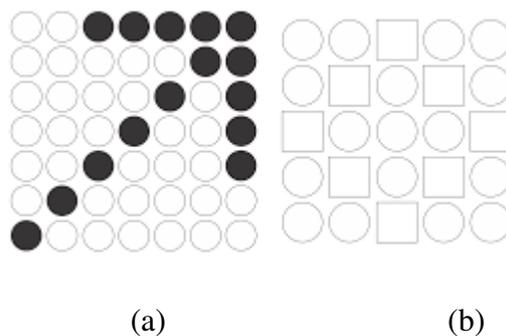


Figura 2.18: Princípio da Similaridade

No caso da Figura 2.18-a observa-se a similaridade pela cor e preenchimento, enquanto a Figura 2.18-b, a organização visual dá-se pela similaridade da forma (círculos e quadrados) .

O princípio da similaridade pode ser observado nas MRVs contidas na Figura 2.19.



Figura 2.19: Princípio da Similaridade em MRVs

Na Figura 2.19-a, observa-se a similaridade pela cor laranja que compõe a raposa. Na Figura 2.19-b nota-se a similaridade pelos elementos com o mesmo peso e cor, também observam-se blocos distintos que delimitam os edifícios, tal sensação é obtida pelo Princípio da Proximidade.

Princípio da Continuidade: este princípio demonstra a tendência dos elementos acompanharem uns aos outros, ou a continuidade de um movimento em uma direção já estabelecida [GOM04]. Têm-se a sensação visual de que as partes se sucedem de modo coerente, sem quebras ou interrupções.

A forma da Figura 2.20-a é então percebida como duas linhas que se cruzam ao invés de quatro linhas que se tocam ou dois (ou mesmo quatro) ângulos opostos. Na Figura 2.20-b, temos a sensação que a espiral está em movimento e não tem fim.



Figura 2.20: Princípio da Continuidade

A Lei da Continuidade apresenta que os pontos conectados por linhas retas ou curvas são vistos de modo a sugestionar que seguem uma determinada direção. Ou seja, ao invés de ver linhas e ângulos separados, as linhas são vistas como estando juntas. A Figura 2.21, apresenta o Princípio da Continuidade aplicado em MRVs.

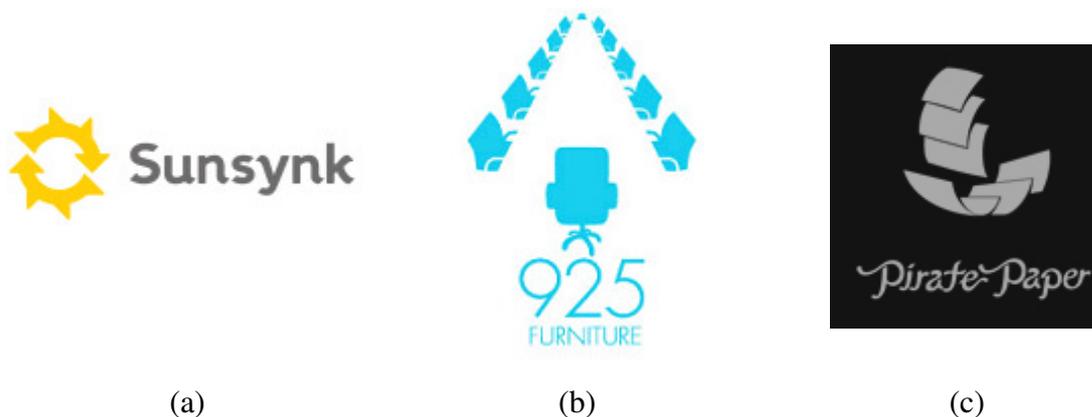


Figura 2.21: Princípio da Continuidade em MRVs

Quando vários elementos apontam para a mesma direção, nosso cérebro tende a acompanhar a direção para a qual os objetos apontam, este fenômeno é facilmente observado na MRVs da Sunsynk, contida na Figura 2.21-a. Quando observamos a Figura 2.21-b, temos a sensação de que as cadeiras continuam infinitamente, nas mesmas posições e distâncias. Na MRV da Pirate Paper (Figura 2.21-c) conseguimos ver o barco formado pelas folhas de papel que flutuam horizontalmente e verticalmente no fundo preto.

Princípio do Fechamento: a sensação de fechamento visual da forma descreve a tendência humana de interpretar o estímulo visual como completo, mesmo quando algumas das informações de contorno estão ausentes. Nosso cérebro completa as

imagens para que elas façam sentido, em que as forças de organização da forma dirigem-se espontaneamente para uma ordem espacial que tende para a formação de unidades .

A Figura 2.22-a e Figura 2.22-b apresentam dois exemplos de MRVs bem conhecidas, o famoso panda da WWF e o cavalheiro caminhando da marca de uísque Johnnie Walker. Ambos utilizam a teoria do fechamento, pois nosso cérebro tende a completar essas imagens abstratas formando imagens das quais já conhecemos (um urso panda e uma pessoa).



Figura 2.22: Princípio do Fechamento em MRVs

A formação de unidades, ocasionada pelo princípio do fechamento, é importante, uma vez que a Gestalt parte da teoria que a primeira sensação da forma é global e unificada. Não se vê partes isoladas, mas relações entre elas [GOM04]. Na Figura 2.22-c observa-se a MRV da Fórmula 1, o segundo elemento em vermelho trata-se de uma imagem abstrata, porém unindo ambos os elementos que compõem a imagem forma-se um elemento conhecido "F1". Para Freitas [FRE02], comprova-se que existe a tendência psicológica de se unir intervalos e estabelecer ligações entre os elementos.

Princípio de Figura-Fundo: baseia-se que a atenção visual do observador alterna-se entre o padrão que emerge como fundo (*background*) ou figura. É fácil lembrar da MRV do Carrefour com aquele "C" utilizando o princípio Figura-Fundo, elemento que passa despercebido por muita gente. Normalmente, estes elementos causam surpresa, mas são de fácil identificação, como mostrado na Figura 2.23-a. Mais

surpreendente ainda, é o fato de que esta MRV foi "adaptada" caracterizando uso indevido da marca, conforme Figura 2.23-b.

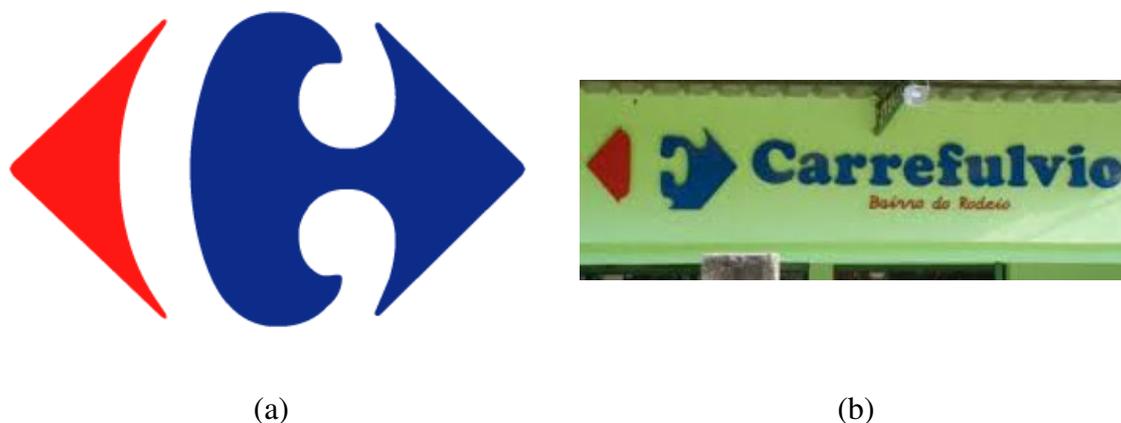


Figura 2.23: Princípio Figura-Fundo MRV Carrefour

Quando observamos atentamente a MRV da rede Carrefour, a soma dos elementos abstratos ajuda a perceber a letra C. Ao analisarmos os elementos que compõem esta MRV, e unindo ambos os elementos (vermelho e azul) estes lembram uma flor-de-lis, símbolo da aristocracia francesa.

O Princípio da Figura-Fundo baseia-se na atenção visual do observador. Tendemos a organizar as percepções no objeto como figura, e no plano contra o qual ele se destaca sendo o fundo (*background*). Na Figura 2.24-a, podemos observar três MRVs que utilizam este conceito.



Figura 2.24: Princípio Figura-Fundo MRVs

Na Figura 2.24-a, observa-se facilmente dois cálices pretos lado a lado no fundo branco, se alternar o objeto de interesse pode-se ver uma casa. A Figura 2.24-b apresenta uma caneta em um fundo branco, ou alternando o ponto de interesse pode-se

observar uma colher com um fundo preto. A MRV do Bar Safari, apresenta uma girafa em um fundo laranja ou duas garrafas em um fundo preto.

Princípio da Pregnância: é considerado a Lei básica da Percepção Visual da Gestalt [GOM04]. Este princípio é geral e abrange os demais princípios [FRE02]. A pregnância diz que todas as formas tendem a serem vistas em seu caráter mais simples tanto quanto permitem as condições dadas. Por exemplo, um homem e um barco a vela podem ser vistos como uma linha reta e um triângulo - quanto mais simples a figura mais fácil ela será assimilada - maior será seu grau de pregnância e melhor será sua relação com o cérebro. Naturalmente, quanto mais confusa a organização visual da forma do objeto, menor será seu grau de pregnância. A MRV das Olimpíadas é um exemplo do Princípio da Pregnância, em que os círculos estão unidos de forma harmônica e parecem dar suporte uns aos outros (Figura 2.25-a).



Figura 2.25: Princípio da Pregnância em MRVs

Na Figura 2.25-b, pode-se observar a MRV da Microsoft conhecida mundialmente. Tal MRV é simples e marcante, suas cores e formas possuem simplicidade e "boa" relação com o cérebro, sendo um excelente exemplo do Princípio da Pregnância.

Princípio da Segregação: este princípio baseia-se na capacidade de separar, identificar e evidenciar unidades da forma, tanto em sua totalidade quanto em partes deste todo [GOM04]. A Segregação poder ser feita através de diversos elementos, como: pontos, linhas, volumes, cores, brilho, textura, entre outros. Na Figura 2.26-a pode-se observar a MRV da Macintosh, o cérebro separa as imagens dividindo em 2 partes facilmente percebidas: 1) um rosto visto de frente e 2) um rosto visto de perfil.



(a)



(b)

Figura 2.26: Princípio da Segregação

Pode-se ainda definir níveis de segregação, onde identificam-se unidades principais de um todo mais complexo. Na Figura 2.26-b, pode-se observar um cenário tendo como elementos principais: o carro, o mar e a estrada. Ainda, no carro pode-se segregar outros elementos como: rodas, faróis, portas, pára-brisa, entre outros. A percepção das unidades visíveis podem se esgotar ou define-se um limite que é considerado suficiente para uma dada leitura visual.

Sabe-se, que os olhos estão equipados para localizar pequenos detalhes, segregando elementos que são essenciais [FRE02] [GOM04].

No processo de percepção de um objeto, os olhos movimentam-se sucessivamente e se prendem em partes mais relevantes do objeto. Na Figura 2.27, uma câmera é utilizada para descrever os movimentos oculares na percepção e reconhecimento de faces. Tal experimento foi realizado por pesquisadores do Instituto de Neurociência da *Rostov State University* na Rússia.

Para Freitas [FRE02] "a percepção visual e o reconhecimento são processos comportamentais e não podem ser completamente entendidos pelos limitados sistemas computacionais. Estes sistemas devem considerar aspectos comportamentais e cognitivos contidos no processo de percepção".

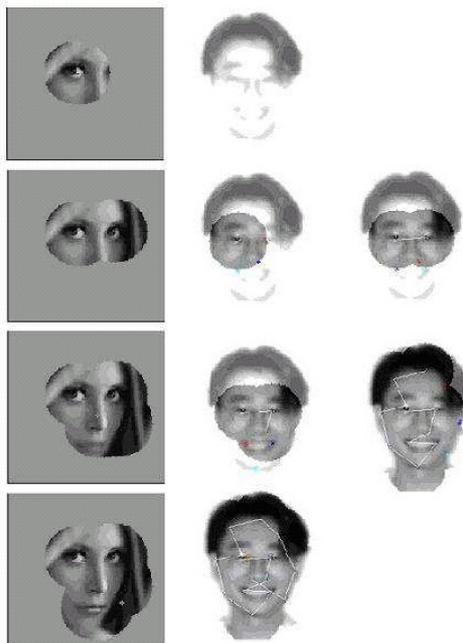


Figura 2.27: Movimentos oculares durante a percepção visual.

Fonte: <http://www.rybak-et-al.net/vnc.html>

Ao projetar Sistemas CBIR deve-se observar que a percepção depende de cooperação e interação entre os processos globais e locais, sendo importante considerar que em uma primeira observação o ser humano processa informações globais, e em seguida processa informações locais para melhorar o reconhecimento, seja de um caractere, de um objeto ou de uma MRV [FRE07].

Pode-se observar na Figura 2.27 que a percepção visual ocorre analisando partes da imagem e evidenciando características locais. Desta forma, os sistemas computacionais podem utilizar os mecanismos de zoneamento, apresentados na Seção 2.6, para auxiliar e aprimorar o processo de reconhecimento.

2.5. Classificadores

Em reconhecimento de padrões, a classificação é utilizada para associar um determinado objeto a uma ou mais classes. Essa associação ocorre conforme as características de cada objeto [BOR12]. A análise digital de imagens possui ferramentas para extração de características de objetos, a fim de obter medidas que possam ser usadas no momento da classificação.

As características extraídas dos objetos formam um espaço multidimensional, chamado espaço de características, onde cada objeto é representado por um vetor de características, sendo que cada objeto é um ponto nesse espaço [GOM07], um exemplo de espaço é apresentado na Figura 2.28.

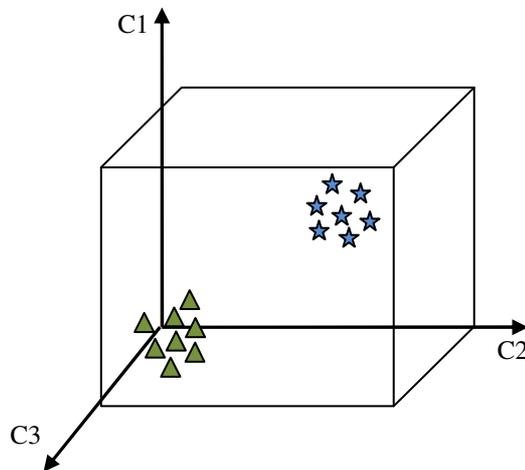


Figura 2.28: Espaço de características

Observa-se na Figura 2.28, que o exemplo de espaço é formado por três características (C_1 , C_2 , C_3), identificando-se claramente dois grupos distintos de objetos.

As técnicas de classificação podem ser divididas em dois grupos: classificação supervisionada e classificação não-supervisionada:

- **Classificação supervisionada:** constitui as classes conforme a base de conhecimento formada por exemplos dos padrões conhecidos. Pode-se então, a partir da base de conhecimento, atribuir um objeto desconhecido x como pertencendo a uma determinada classe C_1 .
- **Classificação não-supervisionada (*clustering*):** não há informações sobre as classes que cada padrão pertence. Agrupam-se as classes em função de sua disposição no espaço de características, inserindo objetos similares na mesma classe.

A Figura 2.29, apresenta o processo de classificação supervisionada. Primeiramente, os dados contidos na base de dados, são preparados em um conjunto de exemplos para que possam ser processados. O conjunto de exemplos é composto por valores característicos de cada exemplo, e pela classe a qual o exemplo pertence. A figura apresenta o formato padrão de exemplos pertencentes a uma base de dados T com m exemplos e n atributos.

Após a extração de características, esse conjunto de exemplos será submetido ao classificador para a etapa de treinamento. O objetivo do treinamento é encontrar uma função que mapeie cada exemplo T_i à sua classe y_i correspondente [BOR12].

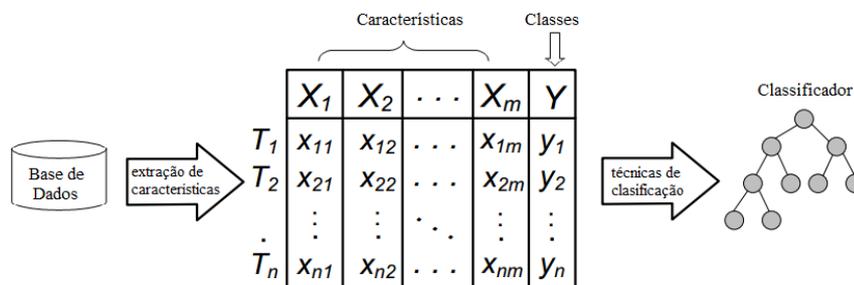


Figura 2.29: Exemplo de classificação supervisionada

Fonte: Adaptado de [REZ05]

Após a etapa de treinamento obtém-se um classificador que seja capaz de prever corretamente a classe de novos exemplos, que não haviam sido usados no treinamento [REZ05].

Neste trabalho, os experimentos realizados utilizaram classificação supervisionada. Nas seções seguintes são apresentados 2 (dois) classificadores: Redes Neurais Artificiais (RNA) e Máquina de Vetores de Suporte (SVM).

2.5.1 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNA) são técnicas computacionais que utilizam modelos matemáticos inspirados nas estruturas biológicas do cérebro humano. Os primeiros trabalhos utilizam uma rede desenvolvida por Rosenblatt [ROS56], denominada Perceptron. A rede Perceptron tem a limitação de resolver somente problemas linearmente separáveis, infelizmente tal característica não está presente na maioria dos problemas reais. Em meados dos anos de 1980, um modelo expandido do algoritmo Perceptron foi apresentado, o modelo *Multi Layer Perceptron* (MLP). Tal modelo impulsionou os estudos nessa área de pesquisa e diversas RNA foram desenvolvidas e estão disponíveis para uso em diferentes tipos de aplicação, como por exemplo em reconhecimento de padrões e simulação [NIE04] [RUM86].

Uma definição clássica para Redes Neurais Artificiais é feita por Haykin (2001):

" Uma Rede Neural é um processador maciçamente paralelamente distribuído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso".

O aprendizado de uma RNA é realizado pelo uso de um algoritmo de aprendizagem que define como os pesos, e eventualmente suas conexões, podem ser alterados durante o treinamento [BOR12] [NIE04].

Uma RNA é constituída de um conjunto de elementos de processamento, chamados neurônios, que interagem entre si. A Figura 2.30 apresenta o modelo de um neurônio.

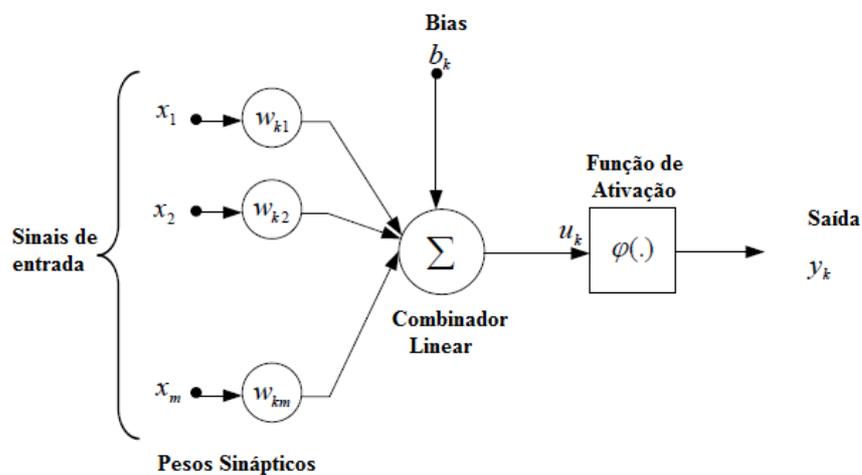


Figura 2.30: Modelo de um neurônio

Fonte: Adaptado de [HAY01]

Pode-se observar na Figura 2.30, que x_j representa as entradas do neurônio e w_{kj} os pesos sinápticos associados a cada entrada. Um combinador linear soma os sinais de entrada, ponderado pela respectiva sinapse do neurônio. Uma função de ativação restringe a amplitude da saída de um neurônio, que geralmente é limitada ao intervalo fechado de $[0,1]$ ou $[-1,1]$ [BOR12].

Matematicamente, um neurônio pode ser representado conforme Equação 2.32:

$$e_T = \sum_{i=0}^D x_i \cdot w_i \quad (2.32)$$

As RNA são capazes de realizar três tipos de aprendizagem [NIE04]:

- Aprendizagem Supervisionada: essa aprendizagem realiza um mapeamento das variáveis de entrada com um conjunto de variáveis de

saída. A rede compara o valor de saída desejado fazendo correções no modelo até atingir um erro aceitável;

- Aprendizagem Não-Supervisionada (*clustering*): nesta aprendizagem, tem-se um conjunto de dados de entrada e deseja agrupar esses dados em conjuntos compostos por elementos similares entre si, sendo que elementos contidos em conjuntos diferentes tenham características distintas;
- Aprendizagem com reforço: esta aprendizagem é semelhante a aprendizagem supervisionada, no entanto não se sabe o valor de saída correto para cada exemplo de entrada. Identifica-se somente se a saída está correta ou não.

Os trabalhos realizados por [ALW99] [SHA14], utilizam redes neurais artificiais em seus experimentos para a recuperação de MRVs.

Neste estudo foi utilizada aprendizagem supervisionada, e as RNAs foram construídas com arquitetura MLP com *backpropagation*. Maiores detalhes sobre a construção das RNA são descritos no Capítulo 4.

2.5.2 Máquinas de Vetores de Suporte (SVM)

As máquinas de vetores suporte (*Support Vector Machine* - SVM), desenvolvidas por Vapnik no ano de 1971 e também apresentado em [VAP99], são um procedimento construtivo universal de aprendizagem com base na teoria de aprendizagem estatística (TAE). Isto significa, que a SVM pode ser utilizada para o aprendizado de várias representações como: as redes neurais; funções de base radial e funções polinomiais.

A formulação mais simples de uma SVM é aquela que trata problemas linearmente separáveis. De forma a colocar o maior número de pontos da mesma classe do mesmo lado, enquanto maximiza a distância de cada classe a esse hiperplano, conforme Figura 2.31-a e 2.31-b. O hiperplano gerado pela SVM é determinado por um subconjunto dos pontos das duas classes, chamado vetores de suporte (Figura 2.31-c) [CHA06].

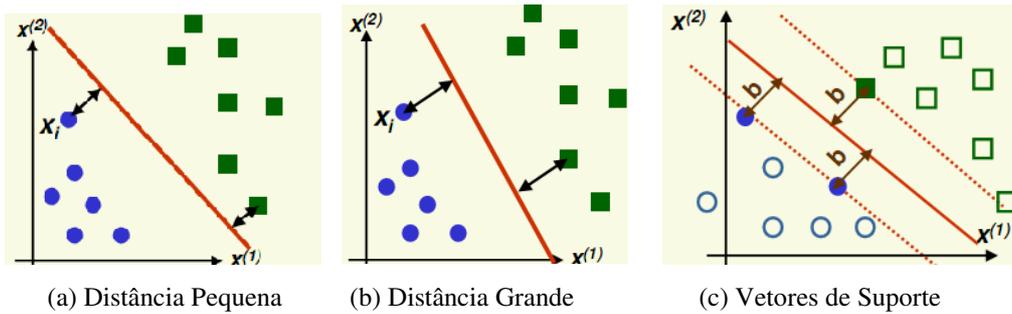


Figura 2.31: Distâncias de hiperplanos e vetores de suporte

Fonte: Adaptado de [OLI09b]

A classificação através de uma SVM usualmente envolve dados de treinamento e teste, que consiste de alguns exemplos contidos na base de dados. Cada exemplo no conjunto de treinamento contém um valor objetivo (classe) e diversos atributos (características). O objetivo da SVM é produzir um modelo que funcione de forma adequada com exemplos não conhecidos - exemplos que não foram utilizados durante o treinamento - adquirindo desta forma a capacidade de prever as saídas para novas entradas [RUF11].

Na classificação binária, o problema consiste em encontrar uma função paramétrica, linear ou não, para um hiperplano de separação dos pontos em dois conjuntos. Seja T um conjunto de treinamento com n dados $x_i \in X$ e seus respectivos rótulos $y_i \in Y$, em que X constitui os dados e $Y = \{+1, -1\}$. Onde, T é linearmente separável, uma vez que seja possível separar as classes $+1$ e -1 por um hiperplano linear. Obtêm-se três hiperplanos: hiperplano de separação (H_0), hiperplano superior (H_1) e hiperplano inferior (H_2), conforme Equações 2.33, 2.34 e 2.35 [SCA05]:

$$y = w^t x + b = 0 \quad (2.33)$$

$$y = w^t x + b = +1 \quad (2.34)$$

$$y = w^t x + b = -1 \quad (2.35)$$

Os pontos que definem os hiperplanos H_1 e H_2 são vetores de suporte. Sendo a distância entre esses hiperplanos calculada conforme Equação 2.36, com o objetivo de encontrar os parâmetros w que maximizem essas distâncias e definir a função objetivo apresentada na Equação 2.37:

$$D(H_1, H_2) = \frac{2}{\|w\|} = \frac{2}{w^t w}, \quad (2.36)$$

$$\text{Minimizar } \frac{1}{2} w^t w \quad (2.37)$$

Para não haver pontos entre os hiperplanos H_1 e H_2 , aplica-se a Equação 2.38:

$$y(w^t x - b) \geq +1 \quad (2.38)$$

Os hiperplanos definidos conforme as equações podem então ser visualizados na Figura 2.32.

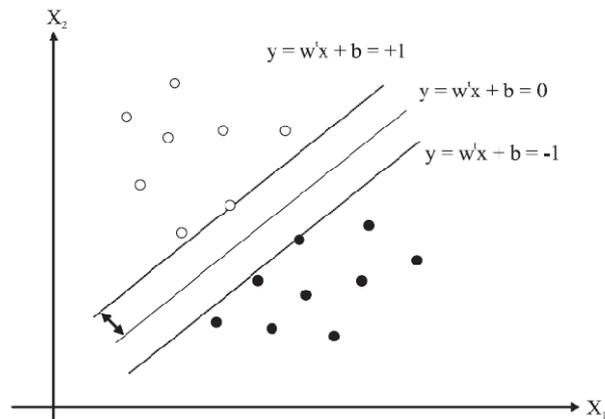


Figura 2.32: Hiperplanos: separação superior e inferior do espaço binário

Fonte: Adaptado de [SCA05]

Este modelo é estendido para tratar problemas linearmente não-separáveis. Neste caso, introduz-se a variável de folga ($\xi_i \geq 0, i = 1, \dots, N$) de forma a criar uma penalidade na função objetivo e uma folga nas restrições [SCA05], conforme apresentado na Equação 2.39:

$$\text{Min } \frac{1}{2} w^t w + C \left(\sum_{i=1}^N \xi_i \right) \quad (2.39)$$

Os trabalhos realizados por Qi [QI_09] e Haitao et al. [HAI10], utilizam SVM em seus experimentos para a recuperação de MRVs.

2.5.3 Múltiplos Classificadores

Muitos métodos de extração de características e técnicas de classificação foram propostos nas últimas décadas para auxiliar na recuperação de MRVs. Porém, isoladamente nenhum conseguiu uma solução completa para o problema.

A etapa de classificação é fundamental em sistemas de reconhecimento e recuperação de imagens. Nos últimos anos um grande número de pesquisas são dedicados a Sistemas de Múltiplos Classificadores (MCS - *Multiple Classifier Systems*) [DU_12] [NAB12] [CAV13] [BRI14], surgindo novas soluções para cada uma das etapas que compõem esse sistema. Brito et. al. [BRI14] apresentam três fases possíveis para um Sistema de Múltiplos Classificadores: a) Geração, b) Seleção e c) Integração; ilustradas na Figura 2.33.



Figura 2.33: Fases do Sistema de Múltiplos Classificadores

Fonte: Adaptado de [BRI14]

Na Figura 2.33 a primeira fase constitui a geração de um *pool* de classificadores que irão compor o sistema, na segunda fase somente um classificador ou um subconjunto de classificadores é selecionado, na última fase a integração dos resultados determina a decisão final. No entanto, Brito et al. [BRI14] ressaltam que esta representação não é única, uma vez que as fases de seleção e integração podem ser facultativas. Sendo possível, por exemplo, que todos os classificadores sejam utilizados para a tomada de decisão, sem precisar passar pela fase de seleção. Ainda, em sistemas em que a fase de seleção indica apenas um classificador, não é necessária a fase de integração.

Para Cavalin et al. [CAV13], o *pool* de classificadores que compõe o MCS deve conter membros que são complementares e diversificados, assim a combinação entre os classificadores superará o melhor membro do conjunto [BRO05] [SAN06].

Peijun Du et al. [DU_12] apresentam três estilos para o MCS:

- 1) Algoritmos baseados em manipulação de amostras de treinamento. Os métodos mais populares são *Bagging* e *Boosting* [BRE96] [FRE96].
- 2) Combinação por concatenação: o resultado de classificação gerado por um classificador é utilizado como entrada do próximo classificador. Os resultados obtidos através de cada classificador são igualmente transferidos

para o próximo classificador, o resultado final é obtido através do último classificador na cadeia de classificadores (Figura 2.34-a).

- 3) Combinação em paralelo: vários classificadores são projetados independentemente sem qualquer interação mútua, e seus resultados são combinados de acordo com uma determinada estratégia [XU_92] [RAH99] (Figura 2.34-b). Se a abordagem é bem projetada, o MCS pode obter resultados precisos [RAN06]. Algumas das abordagens de conjunto mais populares e bem sucedidas são Voto Majoritário, Integral Fuzzy, Teoria da Evidência Dempster-Shafer, entre outros [DU_12].

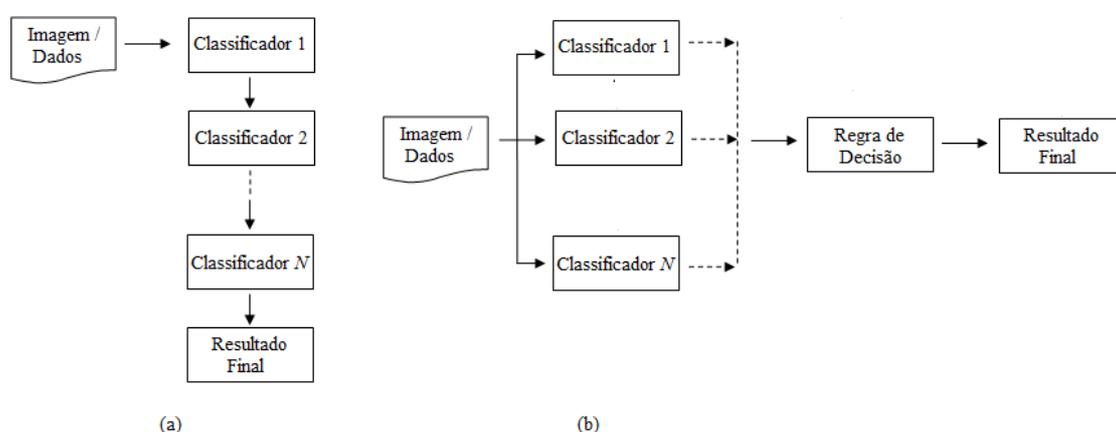


Figura 2.34: Combinação de estilo em MCS

Peijun Du et. al. acrescentam que além da combinação por concatenação e combinação paralela, a combinação hierárquica que combina os dois estilos citados também é utilizada [DU_12].

Para Brito et al [BRI14], pesquisas recentes indicam que para a fase da geração do *pool* de classificadores (fase 1), a direção mais promissora é gerar um conjunto de classificadores precisos e diversificados. A diversidade das saídas dos classificadores é um requisito vital para o sucesso do conjunto [DU_12].

Na segunda fase (Seleção), apresentada por Brito et. al. [BRI14], um classificador ou um subconjunto de classificadores é selecionado. A seleção de classificadores pode ser estática ou dinâmica. A preferência pela seleção dinâmica, se dá pelo fato de selecionar classificadores mais precisos para cada padrão desconhecido [BRI14].

A fase da Seleção de classificadores, apresentada por Brito et al. [BRI14] na Figura 2.33, não é utilizada no Modelo Hierárquico Perceptivo apresentado neste

trabalho, porém mais informações sobre a fase de Seleção podem ser obtidas nos estudos realizados por [GIA01] [KO_08] [CAV13] [DU_12] [NAB12] [BRI14].

A terceira fase do MCS (Integração), consiste na integração das saídas dos classificadores selecionados para a tomada da decisão final.

Para Xu et al. [XU_92], existem três diferentes tipos de saída dos classificadores:

- 1) Nível de Medidas: cada classificador gera um valor numérico para cada classe, indicando a probabilidade do padrão de entrada pertencer a esta classe;
- 2) Nível de *Rank*: cada classificador gera um *ranking* ordenado, sendo que o valor mais alto representa a primeira escolha.
- 3) Nível Abstrato: cada classificador gera apenas um rótulo, que é a classe escolhida para o padrão de entrada.

Jain et al. [JAI00] consideram que o Nível de Medidas transmite informações mais relevantes, enquanto o Nível Abstrato contém a menor quantidade de informações sobre a decisão a ser adotada.

Algumas regras para combinação dos classificadores, podem ser encontradas nos trabalhos apresentados por Kittel et al. [KIT98] e por Jain et al. [JAI00].

2.6. Mecanismos de Zoneamento

Nas últimas décadas, muitos estudos utilizaram os mecanismos de zoneamento ou métodos de decomposição regional para fazer uma análise do padrão localmente. Estes mecanismos auxiliam a solução de problemas em diversas áreas de pesquisa, por exemplo: dígitos [RAD03] [HEG12] [IMP10] ; caracteres [FRE07] [MUR11] [HIR11] [PIR12] [SAF13]; gêneros musicais [COS12], entre outros.

Os mecanismos de zoneamento podem ser utilizados para descobrir potenciais candidatos quando ocorre confusão em uma determinada parte dos objetos analisados [FRE07].

Durante o processo de percepção, os seres humano concentram-se nas partes significativas do objeto a ser reconhecido, buscando uma interpretação eficaz e eficiente. No entanto, sabe-se realmente quais são as partes mais significativas? Onde

as partes significativas estão localizadas no objeto a ser reconhecido? A seguir são apresentados os conceitos de mecanismos de zoneamento e as abordagens baseadas no zoneamento perceptivo considerando uma estratégia assimétrica.

Maiores informações podem ser obtidas em artigos que apresentam uma revisão das últimas décadas de pesquisa na área de zoneamento [PIR12][IMP12].

2.6.1 Conceitos de Mecanismos de Zoneamento

Zoneamento é uma maneira simples para obter informação local, sendo utilizado para extrair informações topológicas dos padrões [LEC00]. O objetivo do zoneamento é a obtenção de características locais, em vez de características globais.

Sendo I a imagem de um objeto a ser reconhecido, e Z_M o método de zoneamento considerado como uma partição de I em M sub-imagens (sendo M um valor inteiro e maior que um). Os zoneamentos são denominados como $Z_M = \{z_1, z_2, \dots, z_M\}$, onde cada partição do zoneamento provê informações relacionadas com a parte específica do padrão a ser reconhecido [IMP2012], e ainda permite identificar a posição em que as características do padrão são detectadas [FRE07].

O *design* do zoneamento, é a maneira pela qual a partição da imagem I em M sub-imagens é definida, sendo considerado de duas formas diferentes: estática e dinâmica.

- Estática: são abordagens tradicionais, onde o *design* utiliza informação *a priori* da distribuição das características nas classes do padrão. Em geral topologias estáticas são projetadas utilizando grades regulares que se sobrepõem ao padrão da imagem [IMP2012], determinando partições uniformes de forma fixa ou simétrica.
- Dinâmica: são projetadas de acordo com resultados de procedimentos otimizados [IMP2012]. O zoneamento não é uniforme, sendo que as partições podem ser definidas de acordo com a densidade do padrão da imagem [FRE07].

Para Freitas et al. [FRE07] “dependendo do domínio da aplicação ou da experiência do pesquisador, o zoneamento pode ser realizado com base nas motivações intuitivas ou com base na maneira mais fácil, isto é, zoneamento fixo ou simétrico”.

Um dos estudos mais tradicionais apresentados, utilizando mecanismos de zoneamento, foi apresentado por Suen et al. [SUE94].

Suen et al.[SUE94] e Li et al. [LI_95] aplicam mecanismos de zoneamento em seus experimentos com caracteres manuscritos. Estes analisam quatro diferentes configurações onde o caractere é dividido em Z partes, sendo $Z = 2, 4$ e 6 , conforme Figura 2.35.

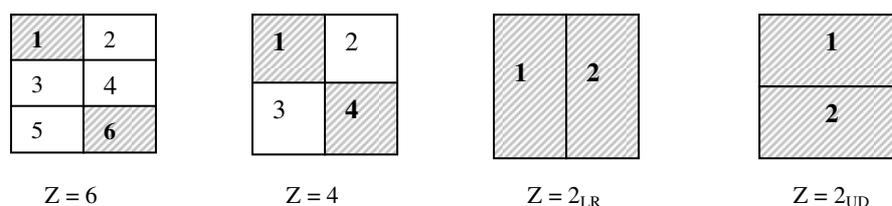


Figura 2.35: Partições onde 1,2,3,4,5 e 6 são o número de posições das partes

Fonte: Adaptado de [SUE94]

Suen et al. [SUE94] observam que existem partes do padrão chamadas de cruciais, onde o reconhecimento é perfeito, ou seja, 100%. Encontrar partes distintas de várias partições de um padrão é um passo preliminar em direção a uma robusta análise de padrões e reconhecimento de caracteres.

Roy et. al [ROY09] apresenta uma nova topologia para reconhecimento de caracteres baseada na forma. Sua abordagem utiliza anéis circulares e anéis de invólucros convexos, conforme ilustrado na Figura 2.36. Os anéis circulares são definidos com círculos concêntricos, cujo centro é o círculo mínimo de inclusão do caractere (MEC - *minimum enclosing circle*). Da mesma forma, os anéis de invólucro convexo são construídos a partir do invólucro convexo da forma do caractere.

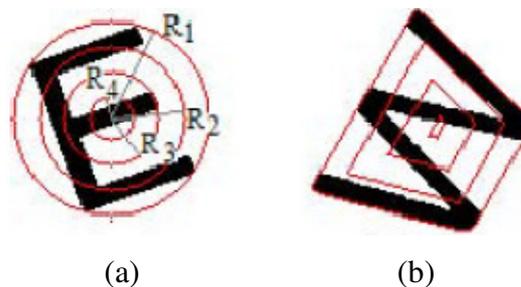


Figura 2.36: (a) anéis circulares (b) anéis de invólucro convexo

Fonte: [ROY09]

No trabalho de Murthy et al. [MUR11], uma técnica é apresentada onde a localização dos pixels pretos do caractere é considerada para contribuir como uma característica única no padrão. Seus experimentos são executados na base de dados Devanagari para dígitos e caracteres manuscritos, usando características com e sem a função de localização de pixels pretos. A Figura 2.37, ilustra os zoneamentos aplicados em um caractere.

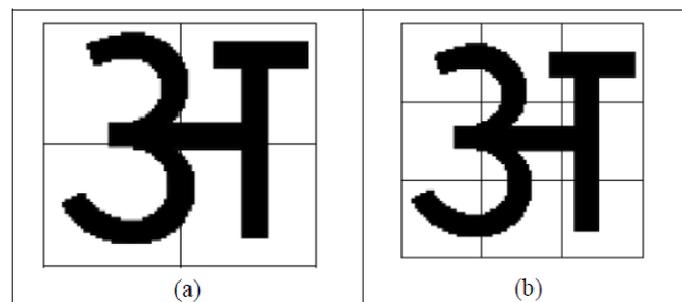


Figura 2.37: Caractere com zoneamento 2x2 e 3x3

No trabalho apresentado por Costa et al. [COS12] uma abordagem alternativa para classificação de gêneros musicais que utiliza a conversão de sinais de áudio em espectrogramas, e a extração de características de textura para estas imagens de tempo-frequência (*time-frequency*) são usadas para modelar gêneros musicais em um sistema de classificação. Costa et al. [COS12] consideram que mesmo não havendo uma relação direta entre dimensões musicais e a extração de características em imagens, espera-se que a extração de características baseada na imagem possa capturar diferentes informações das abordagens que trabalham diretamente com áudio. Com o objetivo de extrair informações locais do padrão, um zoneamento linear é utilizado. No zoneamento aplicado a imagem do espectrograma é dividida em 10 (dez) zonas lineares de igual tamanho, conforme ilustrado na Figura 2.38.

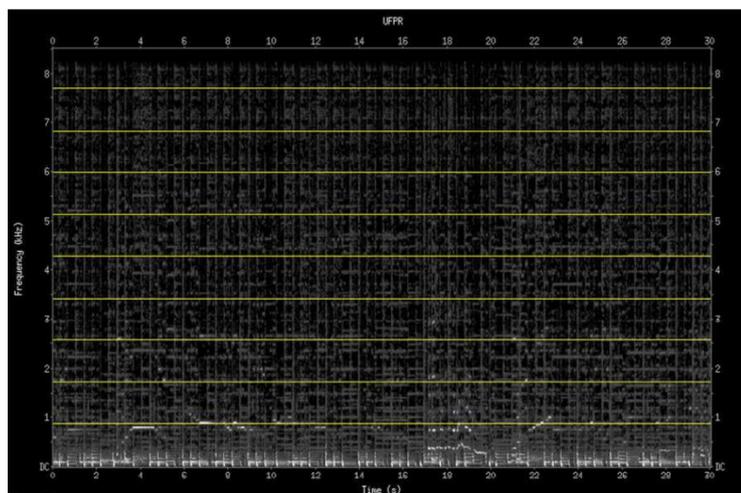


Figura 2.38: Zoneamento linear - espectrograma sinal de áudio

No estudo apresentado por Chang et al. [CHAN11a], um descritor da forma baseado em contorno é proposto, este descritor baseia-se na distância de pontos do contorno da forma. Inicialmente são computadas as distâncias para cada ponto do contorno em relação ao centro geométrico ou centro de massa. Um esquema de particionamento de camadas é desenvolvido para dividir os pontos da fronteira em vários grupos baseados em suas distâncias. Chang et. al. [CHAN11a] defendem que o esquema não somente descreve as características globais de uma imagem, como também extrai o conteúdo do interior da forma. A Figura 2.39, apresenta o particionamento em camadas proposto por [CHAN11a]. A Figura 2.39-a apresenta o particionamento em três camadas, a Figura 2.39-b ilustra a camada externa, na Figura 2.39-c observa-se a camada do meio e a Figura 2.39-d apresenta a camada interna.

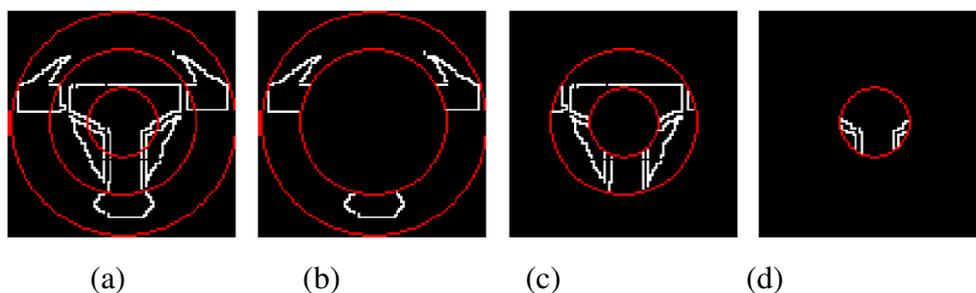


Figura 2.39: Particionamento em camadas

Fonte: [CHAN11a]

Mecanismos de zoneamento estáticos e simétricos, são utilizados por diversos autores. Blumenstein et al. [BLU03] apresentaram um estudo em que o caractere é

dividido em seis partições de igual tamanho. Morita et al. [MOR04] utilizaram a mesma estratégia para dígitos manuscritos. Xiang et al. [XIA04] realizaram a extração de características dividindo a imagem de entrada de caracteres de placas de carro em $n \times m$ ($n = 4$, $m = 4$) zonas. Koerich [KOE05] faz uma análise da imagem de entrada, dividindo o caractere manuscrito em 3×2 partições. Sharma e Gupta [SHA10], utilizam partições regulares de tamanhos 3×2 , 3×3 e 4×4 , para extrair a densidade dos pixels da imagem padrão. Rajashekararadhya e Ranjan [RAJ09] utilizam um zoneamento em partições 5×5 , para cada partição a distância média do centróide do caractere em relação aos pixels de cada linha/coluna são considerados como características.

Mecanismos de zoneamento dinâmico são mais complexos e são utilizados para encontrar o melhor zoneamento a ser aplicado no padrão a ser reconhecido.

Radtke et al. [RAD03] apresentaram uma abordagem dinâmica para definir o zoneamento no reconhecimento *off-line* de dígitos manuscritos, usando Algoritmos Evolucionários Multi-Objetivo (MOEAs). A ideia é uma metodologia auto-adaptativa para definir a estratégia de zoneamento, com m zonas não sobrepostas e uma taxa de erro aceitável, sem necessidade de intervenção humana durante a fase de pesquisa. A Figura 2.40, ilustra os zoneamentos encontrados com a metodologia proposta, bem como o número de zonas e as taxas de erro.

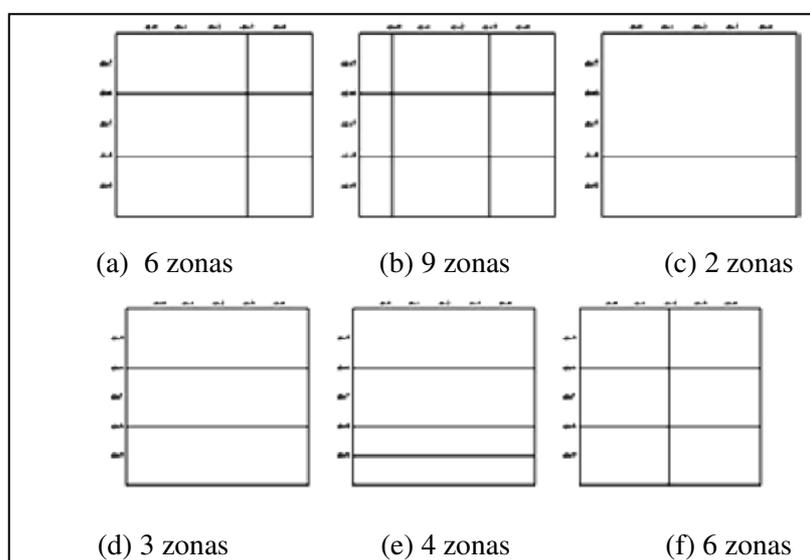


Figura 2.40: Metodologia auto-adaptativa

Fonte: [RAD03]

Impedovo et al. [IMP06] [IMP11] definem uma topologia de zoneamento ótima. O *design* do zoneamento é considerado um problema de otimização, e o zoneamento

ideal é encontrado como sendo o zoneamento para o qual a Função de Custo (FC) associada à classificação é mínimo. Diagramas de Voronoi [AUR91] [OKA92] [BER00] são considerados para descrição do zoneamento. Estes diagramas são compostos por um conjunto distinto de pontos (chamados Pontos de Voronoi), onde cada ponto que compõe este conjunto delimita uma região, que representa uma zona em um método de zoneamento qualquer. Algoritmos genéticos são utilizados para encontrar o zoneamento ótimo. A Figura 2.41 ilustra os zoneamento obtidos por meio do Diagrama de Voronoi, nesta figura observa-se o conjunto de pontos bem como as regiões delimitadas por cada ponto.

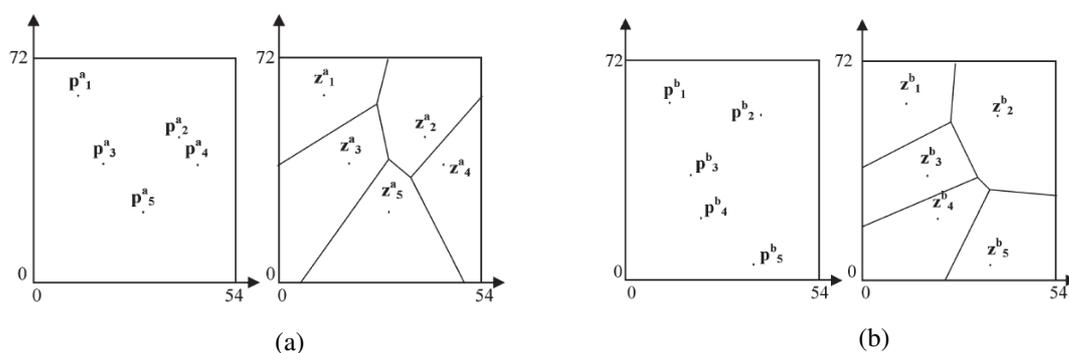


Figura 2.41: Zoneamento por Diagramas de Voronoi.

Fonte: [IMP06]

Para a recuperação de MRVs, poucos trabalhos utilizam mecanismos de zoneamento para auxiliar nesta tarefa. Conforme Kumar [KUM11], um dos sistemas tradicionais na área de recuperação de MRV, foi desenvolvido por Kato et al. [KAT92].

Kato et al. [KAT92] em seu trabalho realizam a extração de características através da subdivisão de uma imagem de MRV em uma malha. A imagem é fisicamente subdividida em uma malha de 4 x 4 ou 8 x 8 partições, conforme apresentado na Figura 2.42.



Figura 2.42: Exemplo subdivisão malha 4 x 4

Fonte: [KAT92]

Para Abe et al. [ABE06] o resultado da extração de características para as partições agrupadas ou contornos que se cruzam em várias partes do zoneamento,

apresentado, apresentado por Kato et al [KAT92] é problemática. O fato da divisão física da imagem antes do processo de extração de primitivas afeta os resultados, sendo necessário aplicar o mecanismo de zoneamento após a extração de primitivas [FRE07].

2.6.2 Zoneamento Perceptivo Assimétrico

No trabalho apresentado por Freitas et.al. [FRE07] é realizada uma análise das partes significativas de caracteres manuscritos usando a matriz de confusão obtida no processo de reconhecimento. A idéia consiste em olhar para a relação entre as regiões e as confusões, permitindo compreender quais as partes do caractere estão causando as confusões. A abordagem não usa qualquer algoritmo de busca complexo e extenso para projetar o zoneamento. Utiliza-se a informação perceptiva contida nas matrizes de confusão para propor uma estratégia perceptiva assimétrica, conforme apresentado na Figura 2.43.

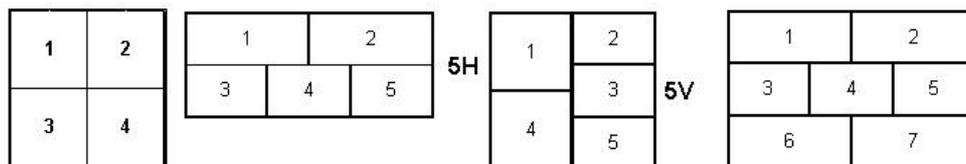


Figura 2.43: Mecanismo de Zoneamento: $Z = 4$, 5H, 5V e 7 regiões

Fonte: Adaptado de [FRE07]

Hirabara et al [HIR11] apresenta um método composto pela seleção dinâmica de zoneamento aplicado em dois níveis distintos visando o reconhecimento de caracteres manuscritos. No primeiro nível, uma Rede Neural Convencional e um esquema de zoneamento tradicional baseado em quatro zonas iguais ($Z = 4$) é usado para prever os três melhores resultados de reconhecimento de um padrão desconhecido. A ideia é prever o melhor esquema de zoneamento a partir das opções apresentadas em [FRE07], como mostrado na Figura 2.44. O segundo nível usa o zoneamento selecionado para extrair as características e realimentar uma rede neural MLP-NN (*Multi-Layer Perceptron Neural Network*), utilizando uma arquitetura classe-modular, que decide o resultado final do reconhecimento.

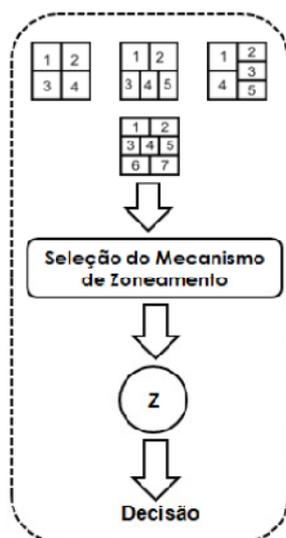


Figura 2.44: Esquema de seleção dinâmica de mecanismos de zoneamento

Fonte: Adaptado de [HIR11]

Para o presente estudo o mecanismo de zoneamento tem por base as áreas propostas por [SUE94] [AIR05] [FRE07]. As características da percepção humanas foram utilizadas durante os estudos a fim de contribuir com a definição das regiões perceptivas. Observa-se localmente os detalhes através do uso dos modelos de zoneamentos perceptivos proposto por Freitas et al. [FRE07]. No Capítulo 4 são apresentados mais detalhes sobre o mecanismo de zoneamento aplicado as MRVs.

2.7. Medidas de Similaridade

Kumar [KUM11] apresenta como ponto relevante na recuperação de imagens de MRVs, a capacidade do sistema CBIR calcular a similaridade (ou distância) entre qualquer par de imagens. Pois em um sistema, qualquer imagem pode ser consultada e qualquer imagem pode ser um membro do banco de dados.

Por meio de medidas de similaridade, pode-se medir o quanto dois padrões, que representam duas imagens, são parecidos.

A similaridade é determinada por um valor normalmente o inverso de um valor de distância, uma vez que quanto mais similares são duas imagens, menor é a distância entre elas dentro de um espaço de representação [PEN09].

Linden [LIN09] apresenta que os relacionamentos relevantes entre os objetos podem ser descritos por uma matriz contendo uma medida de dissimilaridade ou de proximidade entre cada par de objetos.

Cada elemento da matriz M_{ij} consiste em um valor numérico que demonstra quão próximos os objetos i e j são. Algumas métricas calculam a similaridade, outras calculam a dissimilaridade, mas em essência elas são idênticas [LIN09]. Todos os coeficientes de similaridade são funções $d : M \times M \Rightarrow \mathbb{R}$, sendo que M representa o conjunto de objetos foco do trabalho. Basicamente, estas funções permitem realizar a transformação da matriz de dados apresentada na Figura 2.45-a, em uma matriz de distância apresentada na Figura 2.45-b.

$$M = \begin{bmatrix} x_{11} & \dots & x_{1p} \\ \dots & \dots & \dots \\ x_{i1} & \dots & x_{ip} \\ \dots & \dots & \dots \\ x_{n1} & \dots & x_{np} \end{bmatrix} \quad (a)$$

$$d = \begin{bmatrix} 0 & & & & & \\ d_{21} & 0 & & & & \\ d_{31} & d_{32} & 0 & & & \\ \vdots & \vdots & \vdots & & & \\ d_{n1} & d_{n2} & \dots & \dots & 0 & \end{bmatrix} \quad (b)$$

Figura 2.45: (a)Dados do objeto (b) Matriz de distância entre objetos i,j .

A Figura 2.45 $d(i,j)$ representa a distância entre os elementos i e j . As funções de similaridade devem obedecer alguns critérios básicos [LIN09]:

- $d_{i,j} \geq 0$, sendo $i,j \in M$;
- $d_{i,j} = d_{j,i}$, sendo $i, j \in M$. Esta regra significa que a distância entre dois elementos não varia, independente do ponto ao qual ela é medida. Podemos observar na Figura 2.22-b que a matriz de distâncias é triangular inferior, sendo esta simétrica e os valores acima da diagonal estão implícitos;
- $d_{ij} + d_{ji} \geq d_{ik}$, sendo $i, j, k \in M$. Conhecida como desigualdade triangular, especifica basicamente que a menor distância entre dois pontos é uma reta.

Na literatura pesquisada são apresentadas diversas medidas de similaridade, tais como: Distância Euclidiana, Distância de Mahalanobis, Distância de Minkowski, Distância de Manhattan, Distância Bottleneck, Distância de Hausdorff, Distância de Fréchet, Distância entre histogramas [CHA02], Métrica de similaridade [LI_04], entre outras. Essas distâncias são aplicadas de acordo com o contexto dos problemas apresentados [JAI98] [ZHA04] [ABE06] [PEN09].

A Tabela 2.3 apresenta algumas medidas de similaridade utilizadas em sistemas de recuperação de MRVs. Pode-se observar que a Distância Euclidiana é uma medida utilizada por diversos autores para determinar o *ranking* de imagens similares à imagem de consulta.

Autores como [ALW99] [HUS06] definem novas medidas ou a combinação de diversas medidas, buscando melhorar o nível de correspondência nas consultas a base de dados.

Tabela 2.3: Medidas de Similaridade

Autor	Medida de Similaridade
[JAI98]	Template Deformável
[EAK98]	Distância Euclidiana
[CHA99]	Distância Euclidiana multiplicada pelo peso x atribuído aos vetores de características
[KIM00]	Soma das diferenças absolutas ponderadas de cada momento de Zernike
[HUN04]	Distância Euclidiana
[HUS06]	Vetor de Componentes Similares – CSV (<i>Component Similarity Vector</i>)
[JIA06]	Gráfico bipartido com máxima ponderada (WBG)
[ABE06]	Distância de Mahalanobis
[BAG07]	Dois vizinhos mais próximos (<i>Nearest neighbors</i>), associa- um limiar para eliminar falsos positivos.
[ANU13]	Calculada em dois estágios. Primeiro estágio utiliza somente descritor global (D_g) através da distância euclidiana. No segundo estágio utiliza-se o descritor local (D_l), valor final é calculado com os resultados dos dois descritores multiplicados por um peso w .
[SHA14]	Interação de votos de 3 redes neurais.

Nos experimentos realizados neste trabalho, foram utilizadas as Distâncias Euclidiana, Distância de Mahalanobis e Distância entre Histogramas. Com o objetivo de determinar o grau de similaridade entre a MRV de consulta e as MRVs armazenadas na base de dados.

2.7.1 Distância Euclidiana

A distância Euclidiana corresponde à distância geométrica entre dois objetos no plano multidimensional. Calcula-se a distância entre dois vetores de características percorrendo os vetores apenas uma vez, conforme apresentado na Equação 2.40:

$$d_{x,y} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.40)$$

Sendo x_i e y_i , os vetores de características dos objetos I e O respectivamente, e n é a dimensão do vetor de características destes objetos.

2.7.2 Distância de Mahalanobis

A distância de Mahalanobis é calculada considerando o espalhamento dos dados. Tal distância é calculada sempre entre um ponto e um conjunto de amostras, como mostrado na Figura 2.46.

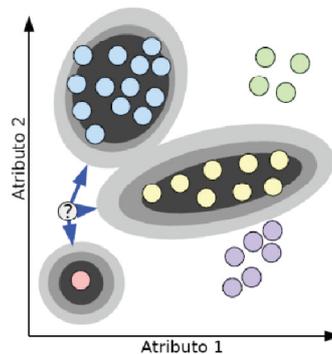


Figura 2.46: Cálculo da distância de Mahalanobis

Fonte: Adaptado de [KOR06]

A distância de Mahalanobis difere da distância Euclidiana pois considera a correlação entre os conjuntos de dados. A fórmula para distância de Mahalanobis entre dois vetores da mesma distribuição que possuam uma matriz de covariância Σ é dada pela Equação 2.41:

$$D(x, m_j) = (x - m_j) \sum_j^{-1} (x - m_j)^T \quad (2.41)$$

Sendo x o vetor de características de um objeto I , e m_j o vetor de características da média do conjunto de objetos.

2.7.3 Distância entre Histogramas

A distância entre dois histogramas proposta por [CHA02] é aplicada na seleção de características, indexação e recuperação de imagem, classificação de padrões e *clustering*. A distância pode ser expressa em termos de valores dos elementos medidos. Dados 2 conjuntos de n elementos A e B , considera-se o problema de encontrar a diferença mínima de pares entre dois conjuntos. O problema visa determinar a melhor atribuição um-para-um entre dois conjuntos, de tal modo que a soma de todas as diferenças entre dois elementos individuais de um par é minimizado. Dados n elementos $a_i \in A$ e n elementos $b_j \in B$, define-se a atribuição da Diferença Mínima entre os pares conforme Equação 2.42:

$$D(A, B) = \min_{A, B} \left(\sum_{i, j=0}^{n-1} d(a_i, b_j) \right) \quad (2.42)$$

Onde D e d são designados como: D_{nom} e d_{nom} , D_{ord} e d_{ord} , e D_{mod} e d_{mod} . Sendo as medidas nominal, ordinal e modular respectivamente, apresentadas nas Equações 2.43 a 2.45:

$$nominal: d_{nom}(x, x') = \begin{cases} 0 & \text{if } x = x' \\ 1 & \text{else} \end{cases} \quad (2.43)$$

$$ordinal: d_{ord}(x, x') = |x - x'| \quad (2.44)$$

$$modulo: d_{mod}(x, x') = \begin{cases} |x - x'| & \text{if } |x - x'| \leq \frac{b}{2} \\ b - |x - x'| & \text{else} \end{cases} \quad (2.45)$$

O menor valor calculado para a distância dos histogramas A e B ($D(A, B)$) é considerado o histograma mais similar. Como $H(A)$ é a menor representação de A , definimos a distância medida entre histogramas, $D(H(A), H(B)) = D(A, B)$, dada na Equação 2.60. Além disso, usa-se $D(A, B)$ como uma forma abreviada da distância entre dois histogramas, $D(H(A), H(B))$.

No Quadro 2.1 apresenta-se o algoritmo para encontrar a distância mínima entre dois histogramas conforme [CHA02].

```

prefixsum = 0
h_dist = 0
for i=0 to b -1
    prefixsum += A[i] - B[j]
    h_dist += |prefixsum|
return(h_dist)

```

Quadro 2.1: Algoritmo Distância Mínima entre histogramas

No Quadro 2.2. é exemplificado o funcionamento do algoritmo contido no Quadro 2.1.

4	1	1	0	0	0	3	1	...	(1)
2	1	1	0	0	0	3	3	...	(2)
2	0	0	0	0	0	0	-2	...	(3)
2	2	2	2	2	2	2	0	⇒	14 (4)

Quadro 2.2: Exemplo algoritmo Distância H(A) e H(B)

As linhas (1) e (2) representam os histogramas (*A*) e (*B*), respectivamente. A linha (3) é a diferença entre os elementos (1) e (2) em cada nível. A linha (4) é o *prefix* soma acumulada dos elementos da linha (3). Nota-se que o último elemento em *prefix* soma é sempre zero se ambos os histogramas forem de mesmo tamanho. No passo final é adicionado o valor absoluto de cada elemento no *prefix* soma, neste caso é igual a 14. Pode-se encontrar mais informações sobre a distância de histograma em [CHA02].

2.8 Medidas de Desempenho na Recuperação

São introduzidas por [SAL71] as medidas *Normalized Recall* (R_n) e *Normalized Precision* (P_n), para a avaliação do seu sistema de recuperação bibliográfico inteligente. Essas medidas podem ser usadas para avaliar qualquer sistema de recuperação que gera a saída em ordem de classificação. Isto inclui a maioria dos sistemas de recuperação [EAK98]. Ambas medidas classificam o desempenho de um sistema de recuperação em resposta a uma consulta, sendo de 0 (pior caso) para 1 (recuperação perfeita). Qualquer uma destas medidas é suficiente para expressar o desempenho de recuperação. A única

diferença entre elas é que R_n dá um maior peso para o sucesso na recuperação dos primeiros itens, enquanto P_n dá peso igual a todas as recuperações [EAK98] [ALW98]. As medidas R_n e P_n , são apresentadas nas Equações 2.46 e 2.47:

$$R_n = 1 - \frac{\sum_{i=1}^n R_i - \sum_{i=1}^n i}{n(N-n)} \quad (2.46)$$

$$P_n = 1 - \frac{\sum_{i=1}^n (\log R_i - \sum_{i=1}^n (\log i))}{\log \left(\frac{N!}{(N-n)! n!} \right)} \quad (2.47)$$

Sendo R_i é a posição de classificação em que documento relevante i é realmente recuperado, n é o número total de documentos relevantes, e N é o tamanho total da coleção documento.

Neste trabalho as Medidas de Desempenho na Recuperação foram utilizadas para avaliar o desempenho na recuperação de MRVs, conforme os experimentos apresentados no Capítulo 5.

2.9. Considerações

Neste Capítulo, foi apresentada a fundamentação teórica necessária para a realização do presente trabalho. Um breve resumo sobre a lei que regulamenta o registro de uma MRV no Brasil foi apresentado. As empresas muitas vezes utilizam os Princípios da Gestalt como ferramenta para construção de sua MRV. Tendo como objetivo que sua MRV tenha uma "boa" relação com o cérebro de seus consumidores, e que esta seja marcante. Com o objetivo de analisar o padrão localmente, mecanismo de zoneamentos foram apresentados, observa-se que estes mecanismos são aplicados em diversas áreas de conhecimento. Assim neste trabalho este zoneamentos são utilizados para a recuperação de MRVs. O processo de classificação é uma etapa importante na tarefa de reconhecimento e/ou recuperação, de tal modo que a Seção 2.5 apresenta os classificadores utilizados nos experimentos. Para calcular a similaridade entre as MRVs aplicam-se as medidas de similaridades, que são apresentadas na Seção 2.7. O desempenho obtido pelo sistema é avaliado através das medidas de desempenho na recuperação, conforme Seção 2.8. No Capítulo 3 serão apresentados os trabalhos relacionados.

Capítulo 3

Trabalhos Relacionados

Atualmente uma grande quantidade de dados visuais é produzida através da internet em *sites* como Youtube, Facebook e outras redes sociais online que contribuem para a explosão de dados digitais. Neste cenário, deve-se considerar que as imagens podem ser geradas em qualquer lugar, com qualquer tipo de iluminação e fundo. Dada a relevância econômica das MRVs, as empresas necessitam de sistemas inteligentes para encontrar evidências de uso malicioso, indevido ou não autorizado de suas MRVs [SAH13].

Esta pesquisa foi conduzida utilizando a base de dados do escritório de patentes de MRVs comerciais do Reino Unido. Esta base de dados, contém imagens em níveis de cinza e não tem fundo (*background*) complexo. No entanto, diversas pesquisas utilizam imagens do mundo real, e questões como iluminação e fundo tornam o processo de recuperação da MRV ainda mais difícil. A Figura 3.1-a apresenta uma imagem do mundo real e Figura 3.1-b apresenta uma imagem da base de dados utilizada nos experimentos.



(a)

(b)

Figura 3.1: MRVs (a) Mundo real, (b) Base de dados do Reino Unido

Com o objetivo de compreender as pesquisas realizadas na recuperação de MRVs, apresenta-se o estado da arte relacionado aos sistemas Sistemas de Recuperação de Imagens Baseada em Conteúdo (CBIR - *Content Based Image Retrieval*) [YOS99] [DAT08] [PEN09] [SNO10] [SIL11].

No trabalho apresentado por Jain et al. [JAI98], um sistema baseado na informação das características da forma é proposto, visando complementar os sistemas tradicionais baseados em texto. O sistema propõe uma hierarquia em duas fases. Na primeira fase chamada de “estágio de poda rápida”, a extração de características é feita por meio de histogramas de direção das bordas e momentos invariantes. Na segunda fase uma técnica baseada em *Deformable Templates* é utilizada, nesta o mapa de bordas da MRV a ser consultado é deformado para fazer correspondência e verificar a similaridade, com o mapa de bordas das Top-N imagens de MRVs a serem recuperadas. A Figura 3.2-a apresenta a posição inicial da *template* de uma cabeça de touro sobreposta ao mapa de borda da MRV da cabeça de um touro gerado pela Transformada de Hough. Na Figura 3.2-b apresenta-se a correspondência final.

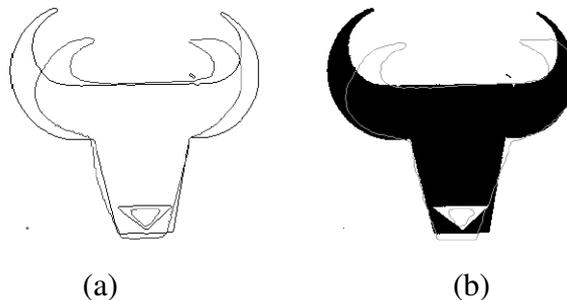


Figura 3.2: Correspondência por Deformable Template.

Fonte: Adaptado de [JAI98]

Eakins et al. [EAK98], em seus estudos realizam uma análise de cada imagem visando caracterizar os componentes-chave da forma, e utilizam agrupamento de regiões da imagem em famílias para espelhar-se na percepção humana, conforme apresentado na Figura 3.3. Os problemas de recuperação ocorridos na *United Kingdom Patent Office*, motivaram o desenvolvimento do projeto ARTISAN (*Automatic Retrieval of Trademark Images by Shape Analysis*).

Na pesquisa de Eakins [EAK98], os examinadores de MRVs foram observados e durante esse processo, nota-se que eles deveriam identificar e recordar as características mais

distintivas da MRV de consulta. Com a intenção de “imitar” o julgamento dos examinadores, o Sistema ARTISAN busca extrair evidências da psicologia da Gestalt, que sugere que os humanos não percebem imagens complexas como a soma de suas partes. Em vez disso, eles percebem elementos da imagem apresentando próximos uns aos outros, similaridades de tamanho e forma, continuidade de linhas, ou fechamento. Os módulos que compõem o Sistema Artisan são: 1) Extração de regiões de bordas para imagens *bitmap* e aproximação por linha reta e segmentos de arco-circular.; 2) reprocessamento da representação de bordas para remover anomalias causadas pelos ruídos na imagem original; 3) Agrupamento de regiões de bordas em famílias; 4) Construção de "invólucros" para famílias de bordas próximas e 5) Extração e armazenagem de características globais da forma. Na Figura 3.3 é apresentado o processamento de uma MRV conforme os módulos que compõe o sistema.

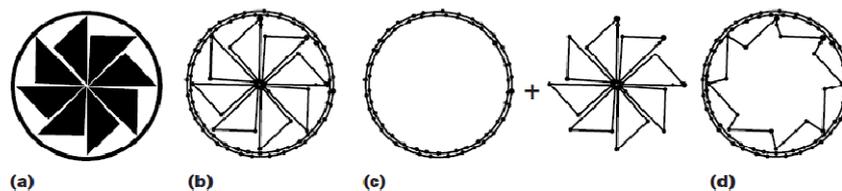


Figura 3.3: Agrupamento em famílias

Fonte: Adaptado de [EAK98]

A Figura 3.3 apresenta um processamento típico de uma MRV pelo sistema ARTISAN. A Figura 3.3-a representa a MRV original. Inicialmente é realizada a extração de contornos, conforme Figura 3.3-b. A seguir, um grupo de contornos é separado em famílias com base na similaridade perceptiva. O primeiro grupo contém dois círculos exteriores considerando uma família, e o segundo grupo contém oito triângulos considerando outra família, conforme Figura 3.3-c. Finalmente, a Figura 3.3-d apresenta a família de contorno construída.

A Figura 3.4, apresenta imagens de MRVs recuperadas pelo sistema ARTISAN [EAK98]. Pode-se observar que duas imagens são consideradas como sucesso na recuperação. No entanto, o sistema apresenta uma imagem considerada como parcialmente recuperada.



Figura 3.4: Sucessos na recuperação do ARTISAN.

Fonte: Adaptado de [EAK98]

Chan e King [CHA99] em seus estudos, propõem agrupar diferentes características da forma sugerindo um conjunto de fatores para a recuperação de imagens de MRVs. Cinco características da forma são usadas para capturar o contorno e as partes internas de uma MRV. Os Descritores de Fourier são usados para capturar as aproximações de contornos. Momentos invariantes, número de Euler, excentricidade e circularidade são utilizados para extrair informação das partes internas, conforme apresentado na Figura 3.5. Havendo mais de um componente em uma MRV, a imagem será conectada por meio de um operador morfológico de fechamento. Para integrar as características da forma, um método de aprendizado supervisionado utilizando um algoritmo genético é proposto, visando encontrar fatores de peso nas funções de dissimilaridade por meio da integração das cinco características da forma.

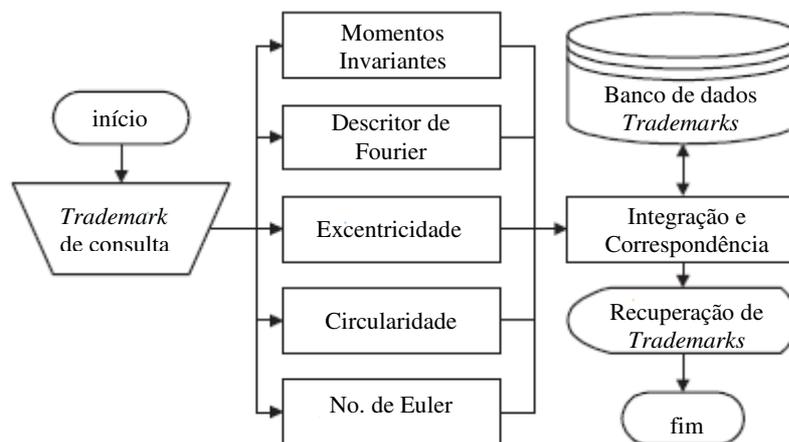


Figura 3.5: Sistema para recuperação de MRVs baseado na forma

Fonte: Adaptado de [CHA99]

Kim [KIM00] utiliza Momentos de Zernike como um descritor da forma. Em seus estudos apresenta que estes Momentos de Zernike têm poder discriminante, invariância de

rotação e imunidade a ruídos. Nos experimentos realizados, Momentos de Zernike foram efetivos na recuperação de MRVs em grandes bases de dados.

Hung et al. [HUN04] apresentam um sistema de recuperação de MRVs em duas fases que integram contorno e atributos de região. Na primeira fase, o contorno da MRV a ser consultado é extraído e utilizado para rejeitar correspondências improváveis de MRVs do banco de dados. Na segunda etapa, as características da região obtida pela ART (*Transformada Angular Radial*) é então usada para pesquisar as MRVs mais similares com a MRV consultada. A medida de similaridade para a MRV considera o contorno, bem como a região do interior da MRV.

No trabalho apresentado por Cerri et al. [CER06], um sistema baseado em contexto é proposto. Três diferentes famílias de descritores da forma, baseado em funções de tamanho, são utilizados: **Família 1) Distância dos pontos**: aplica-se um conjunto de funções de medida parametrizadas utilizando o posicionamento dos pontos na imagem; **Família 2) Projeções**: definem-se funções de medidas para extrair características topológicas da MRV e **Família 3) Jumps**: auxiliam na correspondência de segmentos de linhas. A dissimilaridade é definida com base nas três funções de tamanho, que combinadas determinam um *score*. A base de dados utilizada nos experimentos é da *UK Patent Office* com mais de 10 mil imagens de MRVs comerciais, um conjunto verdade contendo 24 MRVs que foram utilizados para testes.

Hussain e Eakins [HUS06] apresentam um método para agrupamento visual de imagens multi-componentes, tais como MRVs, usando as propriedades topológicas de mapas auto-organizáveis. O método consiste em duas etapas: primeira etapa consiste na construção de um mapa 2D baseado em características extraídas a partir de componentes da imagem e, na segunda etapa um vetor de componentes similares é construído a partir de uma imagem de consulta, que é usado para gerar um mapa 2D das imagens recuperadas. Hussain e Eakins argumentam que na maioria dos sistemas de recuperação de MRVs os resultados são apresentados como uma lista ordenada, de acordo com alguma medida de similaridade. No entanto, estas listas 1D podem dificultar a visão de como as imagens similares não-adjacentes são relacionadas. Então, uma maneira de resolver isso é colocar imagens em uma superfície 2D onde seu posicionamento pode refletir a distância mútua entre as imagens, em um espaço de característica. Este tipo de visualização pode fornecer pistas visuais, como por qual motivo em particular as MRVs se aglomerem em torno da MRV de consulta, e por que outras são

colocadas mais distantes, facilitando a visualização do usuário, realçando as estruturas e padrões na imagem.

Jiang et al. [JIA06], apresentam uma abordagem usando a seleção adaptativa de características visuais para recuperação de MRVs. Em seu trabalho, é investigada a aplicação dos princípios da Gestalt em base de dados de MRVs, uma vez que as MRVs são imagens geralmente ricas em características geométricas abstratas. Para seus experimentos foram utilizadas cinco saliências visuais: a simetria, a continuidade, a proximidade, o paralelismo e encerramento, derivados dos princípios da Gestalt. A Figura 3.6 apresenta exemplos da extração de elementos da Gestalt em imagens de MRVs

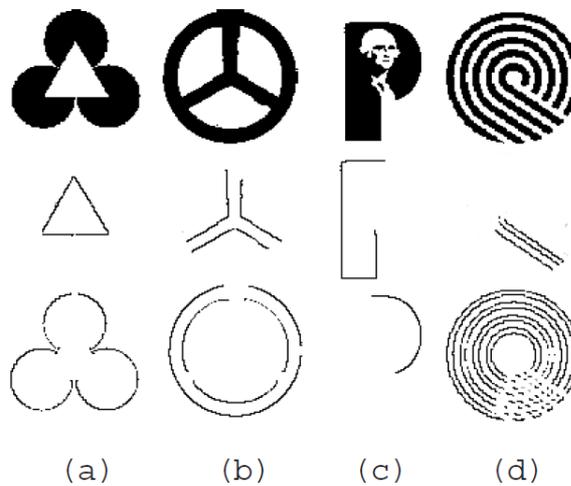


Figura 3.6: Exemplos de elementos da Gestalt extraídos das MRVs

Fonte: Adaptado de [JIA06]

A Figura 3.7, apresenta alguns sucessos na recuperação de MRVs, apresentado em [JIA06].

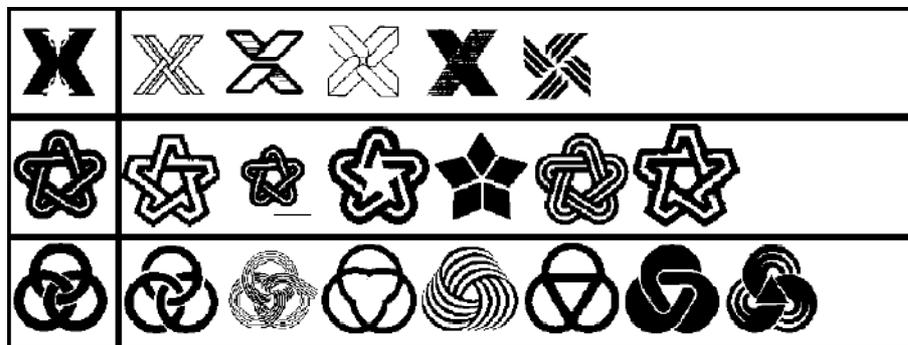


Figura 3.7: Resultados de consultas com sucesso

Fonte: Adaptado de [JIA06]

Um exemplo de problema na recuperação de MRV é apresentado na Figura 3.8, sendo considerada por [JIA06] como uma consulta “difícil”, pois MRVs relevantes não foram recuperadas com sucesso.

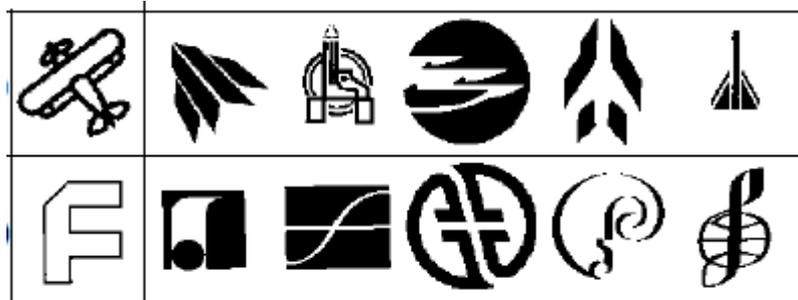


Figura 3.8: Exemplos de consultas complexas

Fonte: Adaptado de [JIA06]

Abe et al. [ABE06] apresentam em seu trabalho um método de reconhecimento de áreas de agrupamento em imagens de MRVs, modelando características para medir o grau de atração entre pares de componentes da imagem. As características extraídas são: proximidade, similaridade da forma e fechamento entre dois componentes. Inicialmente, as características que representam os fatores de agrupamento na psicologia da Gestalt são encontradas. A seguir, uma análise discriminante é realizada utilizando a distância de Mahalanobis para determinar se a combinação de elementos-alvo são agrupados ou não agrupados usando tais características. Este processo é executado para todas as combinações de todos os elementos presentes na MRV. A seguir o múltiplo agrupamento de regiões é combinado para se tornar o padrão de saída, conforme apresentado na Figura 3.9. Os resultados dos agrupamentos obtidos nos experimentos foram comparados com agrupamentos realizados por humanos, conforme apresentado na Figura 3.10.



Figura 3.9: Múltiplo padrão de agrupamento para uma imagem de MRV

Fonte: Adaptado de [ABE06]

Consulta	Padrão de agrupamentos obtidos pelos participantes	Padrão de agrupamento pelo método proposto

Figura 3.10: Alguns agrupamentos com sucesso

Fonte: Adaptado de [ABE06]

Befane et al. [BEF12] apresentam uma extensão do trabalho realizado por [ABE06], com o objetivo de reconhecer agrupamentos de área, baseado no princípio da Gestalt da "boa continuidade". Em seus estudos, o autor considera que ao reconhecer áreas com "boa continuidade" os fatores físicos sugeridos pela psicologia da Gestalt são: 1) proximidade entre os pares de componentes dentro o conjunto de componentes em uma imagem e 2) a similaridade da forma entre esses componentes. O método proposto mede a proximidade entre cada par de componente e "julga" se cada par pode ser parte do agrupamento de área usando Máquina de Vetores Suporte (SVM). Um exemplo de agrupamento é apresentado na Figura 3.11. O valor de proximidade entre os pares (T_i) é obtido pela SVM. Importante observar que o conjunto de treinamento para a SVM foi criado a partir da percepção humana. Para isso, os dados de treinamento foram coletados com base em um questionário aplicado em dez participantes sobre as 75 MRVs utilizadas nos experimentos, conforme Figura 3.11. A similaridade da forma é calculada utilizando Circularidade e Diâmetro equivalente, o conjunto de treinamento da SVM é composto, tal e qual na verificação de proximidade entre os pares, com base nos resultados obtidos no questionário aplicado nos dez participantes (Figura 3.11). Para ambos os casos, proximidade entre os pares e similaridade, foram realizados experimentos utilizando três métodos: Análise de Discriminante Linear (LDA), Análise Discriminante por Mahalanobis (MD) e Máquina de Vetor Suporte Linear (SVM). Com base nos resultados obtidos, o autor determina que SVM foi a máquina discriminante ideal para o

método proposto, pois obteve melhores resultados tanto para proximidade quanto para similaridade da forma.

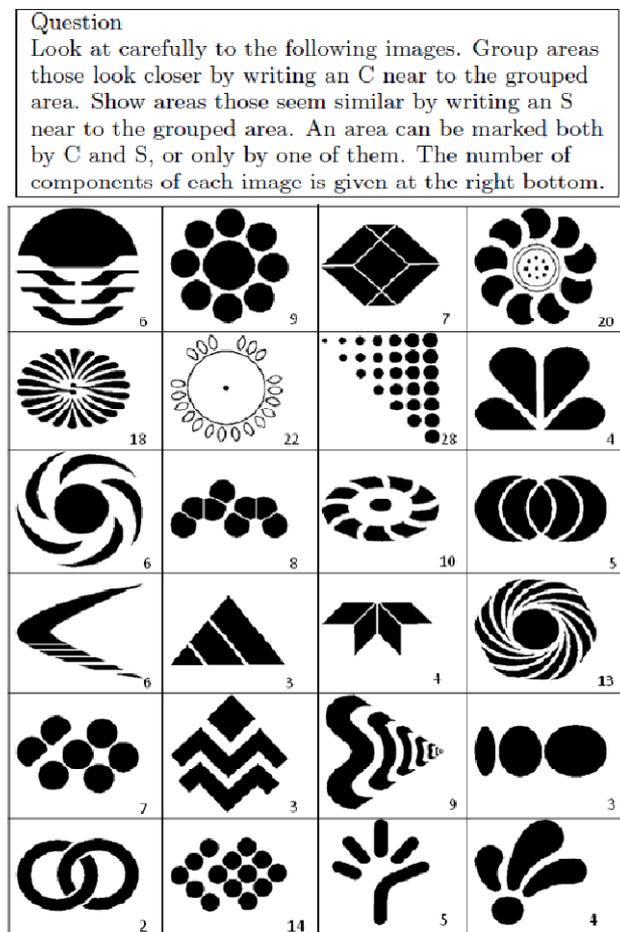


Figura 3.11: Exemplo de parte do questionário aplicado aos 10 participantes

Fonte: [BEF12]

No trabalho de Shaaban [SHA14], é apresentado um sistema para recuperação de MRV baseado na integração de vários classificadores. As imagens de MRVs são convertidas de RGB para níveis de cinza, e também são redimensionadas para um tamanho de 60x60 pixels. Tal sistema utiliza três extratores de características: Momentos Invariantes; Decomposição em Valores Singulares (SVD - *Singular Value Decomposition*) e Transformada Discreta do Cosseno 2D (DCT - *2D Discrete Cosine Transform*). O próximo estágio do sistema é dividido em treinamento e recuperação. Três classificadores neurais são construídos com base nas três características extraídas, os pesos de cada classificador são obtidos e a decisão final é obtida pela integração dos classificadores.

Bagdanov et al. [BAG07] descrevem um sistema para detectar MRV em vídeos de esportes. As empresas (patrocinadores) gastam milhões de euros com marketing esportivo, logo desejam saber se sua MRV teve um bom nível de visibilidade nestes eventos, que justifiquem os altos valores cobrados para esse tipo de propaganda. Este trabalho geralmente é realizado manualmente por anotadores humanos que assistem tais eventos esportivos e anotam todas as vezes que a MRV do patrocinador aparece no vídeo. Tal anotação é muito trabalhosa, uma vez que é necessário assistir ao vídeo mais de uma vez. O autor apresenta que as MRV em vídeos são caracterizadas por: Deformações de perspectiva, Borrão do movimento e Oclusão. O sistema proposto obtém as características da MRV através de uma "bolsa" de pontos SIFT. Devido à grande quantidade de dados, as características SIFT obtidas em cada quadro (*frame*) do vídeo são armazenadas em uma base de dados para posterior recuperação. Cada MRV contida na base de dados de consulta é comparada com a base de dados que contém as características de cada quadro (*frame*) do vídeo. Para cada ponto SIFT da MRV consultada, buscam-se os dois vizinhos mais próximos nos quadros do vídeo e computa-se o *score* de correspondência/similaridade. Uma lista de quadros (*frames*) candidatos é retornada. Os quadros (*frames*) são agrupados temporariamente para definir os intervalos onde a MRV possivelmente está presente. Dois exemplos da localização de uma MRV utilizando SIFT são apresentados na Figura 3.12.

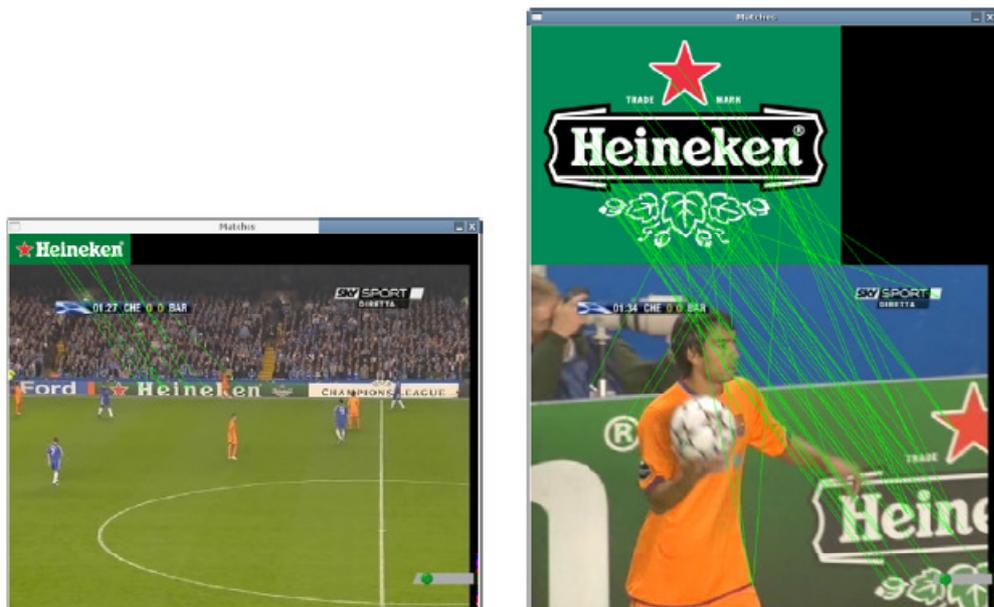


Figura 3.12: Exemplos localização MRV utilizando SIFT

Fonte: [BAG07]

Sahib et al. [SAH13] apresentam uma solução para detecção e reconhecimento de MRVs baseada na definição de uma "Similaridade Dependente do Contexto" (CDS - *Context-Dependent Similarity*). As MRVs de referência e as de teste, são representadas como uma constelação de características locais (pontos de interesse SIFT, regiões, entre outros) e a correspondência/similaridade é feita através da minimização de uma função de energia misturada. A base de dados utilizada nos experimentos foi a MICC-Logos, que possui 13 classes de MRVs, cada classe contém de 15 a 87 MRVs do mundo-real obtidas da internet.

No trabalho apresentado por Anuar et al. [ANU13], um sistema para recuperação de MRVs é proposto, integrando descritor global e local. Como descritor global são empregados Coeficientes de Momentos de Zernike e o descritor local é obtido pela Matriz de Co-ocorrência dos gradientes de borda derivada de informações de contornos. Para o autor as informações de contornos são consideradas importantes na percepção humana das similaridades visuais. Os experimentos são realizados utilizando duas base de dados do padrão MPEG-7, base de dados da forma contendo 1400 imagens e base de dados de MRVs contendo 3260 imagens. Os cálculos de similaridade são apresentados utilizando quatro distâncias: Manhattan, Euclidiana, Manhattan Normalizada e Euclidiana Normalizada.

Jain et al. [JAI98] realizaram uma comparação entre os resultados obtidos pelo sistema automatizados e apresentados por seres humanos. Foi solicitado a cinco indivíduos para recuperar imagens da base de dados usando as mesmas imagens de consulta avaliadas pelo sistema. Os autores concluíram que para todas as consultas, a recuperação obtida pelos cinco avaliadores foram consistentes pelos seguintes motivos: 1) humanos podem facilmente decidir os tons de cinza do primeiro plano dos objetos; 2) humanos tendem a abstrair a imagem de consulta para algumas informações já conhecidas anteriormente. Por exemplo, na consulta da imagem da cabeça de um touro, os seres humanos tendem a recuperar as imagens que se assemelhavam a cabeça de um touro (Figura 3.2), o sistema por sua vez não entende o conceito da cabeça de um touro e recupera a maioria das imagens que são similares a um triângulo. Para Jain et al. [JAI98] estas constatações demonstram a necessidade de desenvolver extratores de características de maneira que os extratores sejam capazes de incorporar a percepção humana auxiliando o desenvolvimento de sistemas de recuperação de imagens.

Conforme proposto e aplicado no presente trabalho, o princípio da Figura-Fundo está sendo utilizado de modo a auxiliar o sistema computacional no que diz respeito ao primeiro

motivo apontado por Jain et al. [JAI98]. As características extraídas, pelo extrator de concavidade e convexidade (Seção 2.3.1.2), permitem determinar o primeiro e o segundo plano nas imagens. Desta maneira, busca-se no sistema um desempenho similar ao processo de percepção visual humana. Assim, o conceito de Figura-Fundo torna-se uma importante ferramenta perceptiva para auxiliar na recuperação de MRVs.

O sistema ARTISAN [EAK98], é apresentado por diversos autores [HUS06] [JIA06] [CER06] [ABE06] [HIR09] [BEF12] como um dos mais importantes estudos na área de recuperação de MRVs. Kumar [KUM11] considera o sistema ARTISAN como inovador, pois este sistema incorpora princípios derivados da psicologia da Gestalt para a recuperação de MRVs, estas muitas vezes consistem de desenhos geométricos abstratos.

A Tabela 3.1, apresenta alguns trabalhos realizados na recuperação de MRVs. Estão relacionados os autores, os descritores da forma, as medidas de similaridade, a base de dados e a quantidade de imagens (tamanho), bem como as taxas de recuperação. Mesmo sabendo que comparar resultados não é uma tarefa trivial, a Tabela 3.1 tem a finalidade de demonstrar de maneira resumida os trabalhos já realizados e que estão relacionados ao tema da Tese, ou seja, recuperação de MRVs no contexto de sistema CBIR.

Tabela 3.1: Trabalhos relacionados à recuperação de MRVs

Autor	Extração de Primitivas	Medidas de Similaridade	Base		Resultados Média Geral
			Nome	Tamanho	
[JAI98] ⁸	Histograma de detecção de bordas. Momentos invariantes de Hu.	<i>Deformable Template</i>	-	1100	R _F = 0,71 P _F = 0,18 R _U = 0,40 P _U = 0,10
[EAK98]	<i>Relative área, right-angleness, directedness, straightness, sharpness, aspect ratio, circularity, transparency, complexity.</i>	Distância Euclidiana	UK Patent Office	10745	R _n = 0,90 P _n = 0,63 L _n = 0,56
[ALW98] ⁹	<i>Co-linearism, co-curvilinearism, parallelism, and end-point proximity.</i>	Combinação de múltiplas medidas de similaridade.	UK Patent Office	1000	R _n = 0,97 P _n = 0,86 L _n = 0,87
[CHA99] ¹⁰	Descritor de Fourier. Momentos invariantes de Hu, No. de Euler, Excentricidade e Circularidade.	Distância euclidiana x peso atribuído aos vetores de características (Algoritmos genéticos).	-	1360	Top-3 = 50% Top-5 = 70% Top-20 = 100%
[KIM00]	Momentos de Zernike.	Soma das diferenças absolutas ponderadas de cada Momento de Zernike	<i>Benchmark trademark database in MPEG-7</i>	6000	CE-1 = 87,22% CE-2 = 75.64%
[HUN04]	Assinatura do contorno. ART (Transformada Angular Radial).	Distância Euclidiana	<i>Benchmark trademark database in MPEG-7</i>	3000	ARR = 0,94 ANMRR=0,101
[HUS06]	<i>Triangularity, rectangularity, and ellipticity. Área, aspect ratio, circularity and convexity.</i> Descritor de Fourier.	Vetor de componentes similares – CSV	UK Patent Office	10745	R _{avg} = 0,55 P _{avg} = 0,09
[JIA06]	Momentos de Zernike. Transformada de Hough (segmentos de linhas, arcos, linhas paralelas, círculos concêntricos e polígonos)	<i>Maximum weighted bipartite graph</i> (WBG)	<i>Benchmark trademark database in MPEG-7</i>	3000	R _n = 0,87 P _n = 0,66 L _n = 0,61
[ABE06]	Circularidade, diâmetro equivalente, taxa de ocupação e <i>fillet</i> gradiente. Cálculo de curvatura. Relação entre a Área e a Menor distância entre 2 componentes da imagem.	Distância de Mahalanobis	<i>Japan Patent Office</i>	1873	R _n = 0,83 P _n = 0,75 L _n = -----

⁸ Média apresentada para a recuperação para 3 *trademarks* de consulta. Sendo que R_F representa as imagens com preenchimento e R_U imagens não preenchidas.

⁹ Média apresentada na recuperação de 6 *trademarks* de consulta.

¹⁰ Porcentagens apresentadas para 10 *trademarks* de consulta. O *ranking* foi estabelecido conforme análise de 5 voluntários.

[CER06]	Size Functions ** Também utiliza Top-1% da base ** Apresenta que algumas imagens não tiveram bons resultados em função da má digitalização e então o descritor não obteve sucesso.	São combinadas as três famílias de funções de tamanho para determinar o <i>score</i> de similaridade.	<i>UK Patent Office</i>	10745	R _n = 0.81 P _n = 0.56 L _n = 0.40
[BAG07] ¹¹	SIFT	Fórmula baseada nos dois vizinhos mais próximos de cada ponto-chave.	-		R _n = 85% P _n = 80%
[BEF12] ¹²	Proximidade entre pares. Circularidade e Diâmetro equivalente	SVM	-	75	R _{n1} = 100% P _{n1} = 100% F-value ₁ = 100% R _{n2} = 100% P _{n2} = 100% F-value ₂ = 100% R _{n3} = 91.11% P _{n3} = 80.39% F-value ₃ = 85.41% R _{n4} = 72.97% P _{n4} = 87.09% F-value ₄ = 79.40%
[SAH13] ¹³	SIFT	Minimização de uma função de energia misturada.	MICC-Logos	13 classe, cada classe composta de 15 a 87 MRVs.	FAR = 0.18 FRR = 0.12
[SHA14] ¹⁴	Momentos Invariantes; Decomposição em Valores Singular (SVD - Singular Value Decomposition) Transformada Discreta de Cosseno 2D (DCT - 2D Discrete Cosine Transform).	Integração de 3 RNAs	- imagens obtidas pela internet	200	93.33%

¹¹ O autor apresenta os resultados em porcentagem. As fórmulas para cálculo de R_n e P_n são diferentes as utilizadas neste trabalho.

¹² O autor apresenta os resultados por Grupos (1, 2, 3 e 4). Os valores para R_n, P_n e F-values, foram obtidos pelo classificador SVM conforme o grupo avaliado.

¹³ O autor apresenta resultados para taxas de Falsa Aceitação (FA) e Falsa Rejeição (FR).

¹⁴ Não há informação como este cálculo foi realizado. O valor 93.3% é apresentado como taxa de recuperação.

3.1. Considerações

Com base na fundamentação teórica e nos trabalhos relacionados, pode-se verificar que a tarefa de recuperação de MRVs não é uma tarefa simples. Destaca-se que é necessário incorporar a percepção humana neste processo e, para isso utilizam-se os princípios da Gestalt. Neste trabalho um Modelo Hierárquico Perceptivo é apresentado, e o princípio da Gestalt denominado Figura-Fundo é explorado por meio do uso do extrator de Deficiências de Concavidade e Convexidade. Com o objetivo de evidenciar informações locais, ressaltando os detalhes não percebidos quando se olha globalmente a imagem da MRV, um mecanismo de zoneamento perceptivo não-simétrico é aplicado. O Capítulo 4, apresenta detalhadamente o Modelo Hierárquico Perceptivo.

Capítulo 4

Método Proposto

4.1. Considerações Iniciais

Neste capítulo são apresentados os elementos necessários para o desenvolvimento deste trabalho. A Seção 4.2 apresenta as Premissas Metodológicas e a base de dados utilizada nos experimentos é apresentada na Seção 4.3. Na Seção 4.4 tem-se uma visão geral do método proposto, descrevendo-se cada uma das etapas que o compõem. Nas seções seguintes são abordados o protocolo adotado para a realização dos experimentos de prova de conceito e a validação do método proposto.

A concepção do método proposto tem por base um Modelo Hierárquico Perceptivo, o qual permite recuperar MRVs similares dentro do contexto de sistemas CBIR, sendo composto pelas seguintes etapas:

- Pré-processamento: limiarização, extração de contornos e construção de uma “caixa” (*bouding box*) para delimitação das MRVs;
- Extração de primitivas considerando os seguintes extratores: Cadeias de Freeman, Concavidade/ Convexidade, Momentos de Hu e SIFT;
- Seleção dinâmica de extrator de características;
- Aplicação de mecanismos de zoneamento perceptivo não-simétrico;
- Cálculo de similaridade entre as MRVs;
- Determinação do *ranking* de imagens de retorno a partir de uma imagem de consulta (*query*).

Nas seções a seguir, são descritas detalhadamente as etapas que compõem o modelo proposto, permitindo testar e validar o processo de recuperação de MRVs.

4.2 Premissas Metodológicas

O problema central deste trabalho é a recuperação de MRVs em uma base de dados de instituições responsáveis pelo registro de marcas e, portanto, torna-se necessário estar atento às similaridades existentes entre marcas distintas. Dada uma MRV deve-se apontar as MRVs mais similares de modo a evitar problemas de plágio ou confusão entre marcas distintas, tal qual descrito no Capítulo 2. Neste contexto, é necessário extrair informações que possam identificar a MRV sob análise no universo de MRVs já existentes e registradas. Para obter informações da MRV, quatro extratores de características foram utilizados: Deficiência de Concavidade e Convexidade [PAR97], Cadeias de Freeman [FRE74], SIFT (*Scale Invariant Feature Transform*) [LOW04] e Momentos Invariantes de Hu [HU_62]; conforme apresentado na Seção 2.3. Tais extratores formam um conjunto de classes de extratores, os quais são submetidos a um processo de classificação por meio de SVM (Máquinas de Vetores Suporte). Dessa forma, existem classes de extratores, sendo o objetivo encontrar o extrator que melhor represente a MRV sob consulta.

Para auxiliar na solução do problema de recuperação de MRVs utilizam-se os mecanismos de zoneamento perceptivo propostos por Freitas et al. [FRE07] com o intuito de destacar padrões importantes em determinadas regiões das MRVs, conforme apresentado na Seção 2.6. Finalmente, o método proposto apresenta o conjunto contendo as Top-100 (o que representa 1% da base de dados) imagem de MRVs mais similares a MRV consultada.

4.3 Base de Dados

A base de dados utilizada nos experimentos, pertence ao Escritório de Propriedade Intelectual do Reino Unido (IPO – *Intellectual Property Office*) [IPO12]. Para que se pudesse trabalhar com esta base de dados foi necessário solicitar autorização de uso, ressaltando que somente após a concessão desta pode-se realizar os experimentos.

A base de dados possui 10.745 imagens de MRVs em níveis de cinza. O formato original é .tif, não existindo uma padronização no tamanho das imagens. A Figura 4.1 ilustra algumas imagens contidas na base de MRVs. Cabe ressaltar que não existem

classes pré-definidas - as imagens não se encontram rotuladas - sendo que as imagens são somente nominadas por meio de numeração específica.

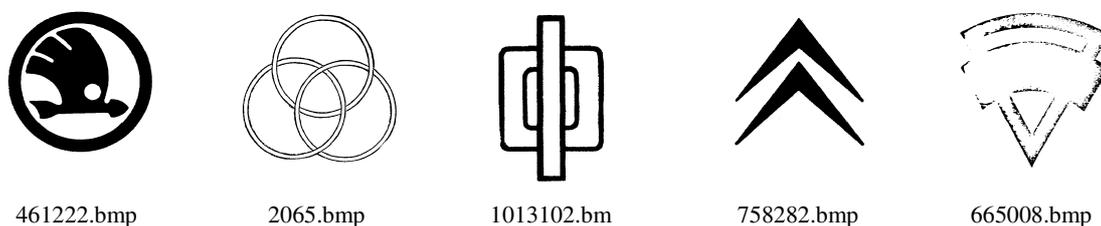


Figura 4.1: Exemplos de MRVs contidas na base de dados

Na Figura 4.1, pode-se observar a diversidade de imagens contidas na base de dados. Observam-se também imagens que possuem "baixa" qualidade como por exemplo a imagem 665008.bmp, fato que deteriora o processo de extração de primitivas e, portanto, a recuperação propriamente dita de MRVs.

Um conjunto adicional à base de MRVs, denominado "conjunto verdade", contém 20 MRVs selecionadas pelos especialistas do escritório de patentes [EAK98]. Este conjunto é proposto para realizar testes de desempenho, conforme mencionado e utilizado nos experimentos realizados por Eakins et al. [EAK98] e Cerri et al. [CER06].

No "conjunto verdade" cada uma das 20 imagens de MRVs está associada às imagens mais similares da base de dados de acordo com a avaliação dos especialistas. A Tabela 4.1 apresenta a composição de cada conjunto verdade para as 20 imagens de teste pré-determinadas.

Na Tabela 4.1, o "conjunto verdade" possui 287 imagens de MRVs associadas às 20 (vinte) MRVs de consulta, perfazendo um total de 307 imagens de MRVs. O mesmo "conjunto verdade" é utilizado nos trabalhos apresentados por Eakins et. al. [EAK98] e Cerri et. al. [CER06] para validação de seus experimentos.

Para os experimentos realizados neste trabalho, considerou-se o mesmo "conjunto verdade" contraposto ao conjunto completo de 10.151 imagens de MRVs contidas nesta base de dados. O desempenho de recuperação é calculado pelas medidas *Normalized Recall* (R_n) e *Normalized Precision* (P_n), conforme apresentado na Seção 2.8.

Tabela 4.1: Conjunto verdade

	MRV de consulta		No. de imagens similares
1	1037814.bmp		25
2	1055261.bmp		15
3	1138103.bmp		11
4	1138293_.bmp		09
5	1190540.bmp		09
6	1259886.bmp		17
7	1267206.bmp		10
8	1279931.bmp		19
9	1289047.bmp		24
10	1376861.bmp		10
11	1439229_.bmp		09
12	1486213_.bmp		15
13	1575268_.bmp		12
14	2010916_.bmp		12
15	2016658_.bmp		16
16	2018809_.bmp		11
17	2042822_.bmp		20
18	665322.bmp		09
19	914.bmp		22
20	967049.bmp		12
	TOTAL		287
	TOTAL GERAL		287 + 20 = 307

4.4. Modelo Hierárquico Perceptivo

No presente trabalho, um modelo hierárquico perceptivo é apresentado composto por dois níveis: 1) extração de características e 2) mecanismos de zoneamento perceptivos. A etapa de extração de características ocorre no Nível 1 e os mecanismos de zoneamento são aplicados no Nível 2, como ilustrado na Figura 4.2.

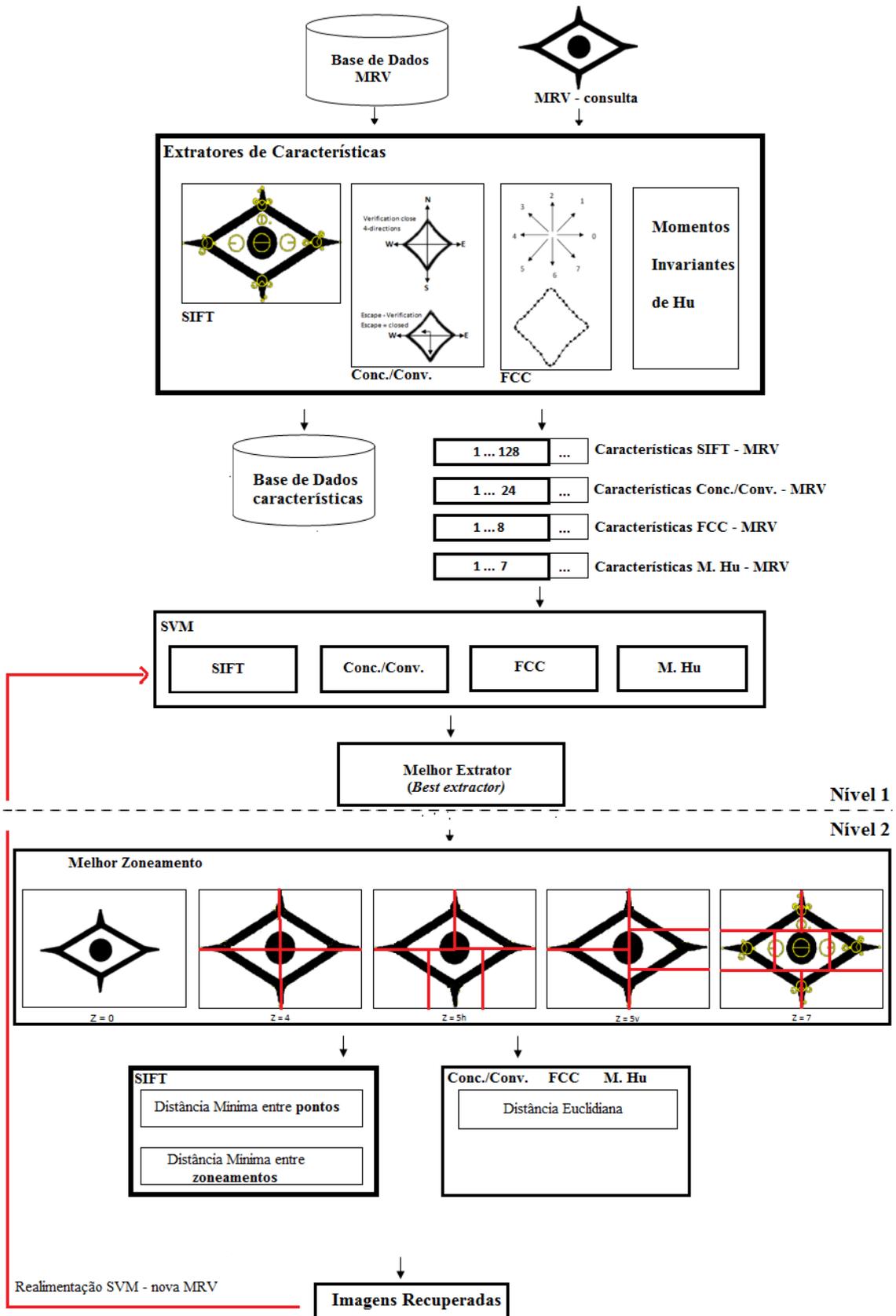


Figura 4.2: Modelo Hierárquico Perceptivo

A Figura 4.2 exibe uma visão geral do Modelo Hierárquico Perceptivo utilizado nos experimentos. No Nível 1 do modelo, um sistema de classificação é utilizado com o objetivo de encontrar o melhor extrator de características para representar a imagem a ser consultada. Tal classificador é construído conforme apresentado na Seção 4.7.2. No Nível 2 define-se o melhor zoneamento perceptivo a ser aplicado. Os níveis deste modelo são descritos a seguir:

1. **Nível 1:** São extraídas as características utilizando-se quatro classes de extratores:
 - a. dois extratores de contorno: Cadeias de Freeman e Deficiências de Concavidade e Convexidade;
 - b. dois extratores por região: SIFT e Momentos Invariantes de Hu.

Após o processo de extração das características da MRV de consulta faz-se a seleção dinâmica do extrator por meio do classificador SVM. Cada extrator representa uma classe, tal qual explicado anteriormente, sendo o objetivo neste nível que o SVM indique qual o melhor extrator para representação da MRV sob consulta.

2. **Nível 2:** Conhecendo o melhor extrator, os zoneamentos perceptivos propostos por [FRE07] (Seção 2.6.2) são aplicados na MRV de consulta. Na sequência, calcula-se a correspondência entre as MRVs contidas na base de dados (Seção 4.3) e a MRV de consulta. Caso o método SIFT seja definido como melhor extrator no Nível 1, dois cálculos de similaridade podem ser aplicados: distância mínima entre pontos-chave (Seção 4.9.1) e distância mínima entre zoneamentos perceptivos (Seção 4.9.2) [FRE07]. Para os demais extratores aplica-se a Distância Euclidiana entre os vetores de características.

Destaca-se que como resultado final, são retornadas as Top-100 MRVs mais similares, o que representa 1% da base de dados. Assim, por meio do cálculo do *normalized recall* (R_n), para todos os zoneamentos perceptivos propostos, o melhor zoneamento é definido e, portanto, pode ser utilizado na definição das Top-100 MRVs que serão retornadas ao usuário. A medida de recuperação *normalized recall* (R_n) foi escolhida por considerar um maior peso para o sucesso na recuperação dos primeiros itens.

4.5 Pré-Processamento

Na etapa de pré-processamento diversas operações são realizadas para o tratamento da imagem digitalizada, tais como realce de contraste, minimização de ruídos, segmentação e limiarização, de modo que as características da imagem possam ser realçadas e problemas possam ser minimizados. No presente estudo foi aplicada a etapa de pré-processamento, composta pelas técnicas descritas a seguir:

- Limiarização (*thresholding*): esta tarefa consiste em converter a imagem original em níveis de cinza em uma imagem binária. Na imagem original, as variações de nível de cinza ocorrem em um intervalo compreendido do preto (valor mínimo) ao branco (valor máximo). O extremo maior corresponde ao fundo branco, enquanto que o extremo menor representa a imagem que está em primeiro plano. A tarefa de limiarização consiste em determinar o valor para um limiar de escala de cinza e todos os valores abaixo deste limiar escolhido, associá-los ao preto. No presente trabalho utilizou-se o método de OTSU [OTS79];
- Extração de contornos (bordas): o contorno pode ser obtido por processos morfológicos de dilatação e erosão, utilizando um elemento estruturante em cruz [FAC96]. O elemento estruturante em cruz permite obter contorno da espessura de um pixel [SOL05]. O processo de dilatação tem como objetivo aumentar o número de pixels nas bordas da imagem, enquanto que o processo da erosão é inverso, ou seja, reduzindo o número de pixels nas bordas da imagem. Após aplicar o processo de erosão e dilatação, faz-se a sobreposição das duas imagens resultantes (erodida e dilatada) e realiza-se a subtração dos pixels, obtendo a imagem do contorno da MRV. Neste trabalho, os contornos foram extraídos pelo Matlab, por meio da função *edge*;
- Caixa (*Bounding Box*): este pré-processamento busca extrair da imagem os limites acima, abaixo, esquerda e direita criando ao redor da MRV uma “caixa”. Este procedimento visa auxiliar o mecanismo de zoneamento (*zoning*), de acordo com as regiões perceptivas a serem aplicadas.

As etapas de pré-processamento descritas anteriormente foram aplicadas nas imagens contidas na base de dados conforme o extrator de características. Para Cadeias de Freeman, as imagens da base de dados passaram pelas etapas de limiarização e extração de contornos. No caso do extrator de Deficiências de Concavidade e Convexidade foram aplicadas limiarização e delimitação por caixa (*bouding box*). Para os extratores Momentos de Hu e SIFT, somente foram aplicadas a etapa de limiarização.

4.6 Extração de Características

Atualmente, apesar dos inúmeros esforços na área de Recuperação de Imagens Baseada em Conteúdo (CBIR) de modo a encontrar o melhor conjunto de características e também o melhor método de combinar tais primitivas, esta área ainda permanece um desafio.

Os sistemas de recuperação de MRVs utilizam descritores da forma para extrair as características das imagens de MRVs [SIL11]. A extração de características codifica o conteúdo de uma imagem, por exemplo, por meio de um vetor de primitivas, enquanto uma função que calcula a distância define a similaridade entre dois vetores e, conseqüentemente, entre duas imagens.

Neste trabalho, o método avaliou experimentalmente 4 (quatro) classes de extratores de primitivas diferentes, conforme apresentado na Figura 4.2.

Foram definidos 2 (dois) extratores de contornos: Deficiências de Concavidade e Convexidade; e 2 (dois) extratores por região: Momentos Invariantes de Hu e SIFT (*Scale Invariant Feature Transform*) - conforme apresentado na Seção 2.3.

De acordo com os experimentos apresentados no Capítulo 5, os extratores obtiveram resultados individuais superiores aos resultados obtidos por meio da combinação entre eles. Desta forma, tornou-se necessário definir dinamicamente qual o melhor extrator a ser aplicado a cada MRV de consulta. Para esta tarefa utilizou-se o classificador SVM para as classes de extratores. Maiores informações sobre o classificador SVM encontram-se na Seção 4.7.2.

Nos trabalhos relacionados, apresentados no Capítulo 3, alguns se destacam por lançar mão dos conhecimentos da área de percepção humana por meio dos princípios da Gestalt [EAK98] [ABE06] [JIA06]. Neste trabalho, estudou-se e aplicou-se o princípio da Gestalt denominado **Figura-Fundo** na recuperação de MRVs, o qual baseia-se na hipótese do cérebro organizar os elementos em perspectivas, definindo o que é figura ou

fundo a cada cena observada. Analisando tal conceito pode-se associar o extrator que explora informações de Concavidade e Convexidade para evidenciar as perspectivas de Primeiro Plano (figura) e Segundo Plano (fundo) nas MRVs sob consulta. Mais detalhes sobre estes estudos e a aplicação deste princípio segue apresentado na Seção 4.6.1.

4.6.1. Primitiva com base na percepção humana das formas

Conforme apresentado por Eakins et. al. [EAK98], os seres humanos percebem elementos da imagem apresentados próximos uns aos outros, considerando assimilaridades de tamanho e forma, continuidade de linhas, ou fechamento.

A Figura 4.3, apresenta um par de imagens semelhantes em relação à triangularidade. A Figura 4.3-b é um triângulo, entretanto na Figura 4.3-a somente a parte esquerda é um triângulo e as demais partes são elementos trapezoidais [ABE06]. No entanto, o ser humano percebe a totalidade da imagem, por meio do princípio do fechamento o cérebro despreza as falhas ou linhas de separação e observa o todo, o triângulo [ABE06].



Figura 4.3: Par de imagens similares

Fonte: Adaptado de [ABE06]

Para Abe et. al. [ABE06] é necessário considerar não apenas as características de cada componente da imagem, mas a atração entre eles. Para isso, alguns princípios da Gestalt como Proximidade e Continuidade podem ser utilizados. Os métodos aplicados para explorar tais conceitos, podem ser visualizados na Tabela 4.2. Neste trabalho, busca-se explorar o Princípio Figura-Fundo por meio de informações de concavidade e convexidade, uma vez que é possível alternar a perspectiva entre o primeiro plano (figura) e o segundo plano (fundo) das imagens. A Tabela 4.2 apresenta um conjunto de primitivas para extrair informações das imagens de MRVs, sendo estes agrupados de acordo com os princípios da Gestalt.

O conjunto apresentado na Tabela 4.2 contém primitivas baseadas na forma, tais como: triangularidade, circularidade, arcos e linhas; visto que estas primitivas têm uma conexão com a percepção humana das formas, uma vez que a forma do objeto observado é importante no processo de percepção da imagem propriamente dita.

Tabela 4.2: Primitivas e Princípios da Gestalt

Princípio da Gestalt	Primitivas
Proximidade	Linhas, círculos e arcos (concêntricos). Distâncias entre áreas distintas [JIA06].
Similaridade	Relação de aspecto/Proporcionalidade, taxa de ocupação. Circularidade. Triangularidade, retangularidade e elipticidade [ABE06] [HUS06].
Continuidade	Linhas e círculos (colineares e/ou paralelos) próximos o bastante que podem ser agrupados/unidos [JIA06].
Fechamento	Cálculo de curvatura [ALW99].
Figura-Fundo	Deficiências de concavidade e convexidade [AIR14].

Algumas das características contidas na Tabela 4.2, foram estudadas pelos pesquisadores e apresentadas no Capítulo 3 sobre trabalhos relacionados [EAK98][JIA06][ABE06]. Com o objetivo de complementar o conjunto de características, foram realizados experimentos observando o conceito da Gestalt chamado de Figura-Fundo. Tal princípio baseia-se no fato de que a atenção visual do observador alterna entre o padrão que emerge como fundo ou figura, conforme Figura 4.4.

As Figuras 4.4-a e 4.4-b apresentam duas imagens clássicas do Princípio Figura-Fundo, sobre as quais pode-se verificar tais imagens rotuladas utilizando informações de Concavidade e Convexidade nas Figuras 4.4-c e 4.4-d. É interessante observar, que mesmo a imagem estando rotulada, a alternância entre Figura e Fundo é facilmente verificada, é possível enxergar em cada uma das imagens rotuladas duas figuras tal qual a imagem original.

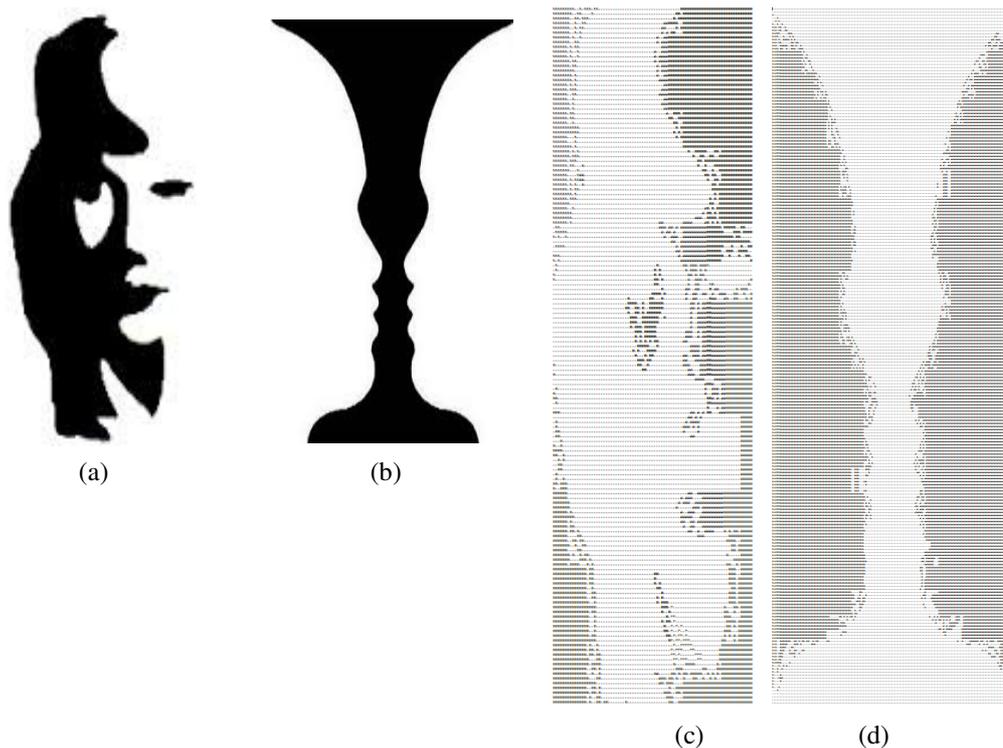


Figura 4.4: Princípio da Gestalt: figura-fundo.

Por meio da rotulação aplicada à imagem, o padrão a ser considerado como primeiro ou segundo plano dependerá do “foco de atenção” que o sistema dará aos pixels pretos e brancos contidos na imagem. Podendo apontar o “foco de atenção” em apenas um deles.

Conforme resultados apresentados na Seção 5.2.1, a utilização de concavidade e convexidade obteve bons resultados individuais, o que fortaleceu o uso deste extrator para complementar os conjuntos de primitivas propostos por [ABE06][JIA06].

4.7. Classificação

Após a etapa de extração de características, torna-se necessário definir o melhor extrator a ser utilizado para encontrar as imagens mais similares à MRV de consulta. Nos experimentos realizados, apresentados no Apêndice C, a combinação de extratores não melhorou os resultados individuais para a MRV consultada. Assim, tornou-se necessário definir uma estratégia que fosse capaz de selecionar dinamicamente o melhor extrator (*Best Extractor*) a ser utilizado. Para tal tarefa, foram realizados experimentos utilizando dois classificadores: Redes Neurais Artificiais (RNA) e Máquinas de Vetores

Suporte (SVM), conforme apresentado na Seção 2.5. Detalhes sobre a construção desses classificadores estão contidos nas Seções 4.7.1 e 4.7.2.

Ambos classificadores foram treinados e testados com base no "conjunto verdade", contido na base de dados da UK *Patent Office* [IPO12]. Como explicado anteriormente, o "conjunto verdade" é composto por 20 imagens de consulta, sendo que cada imagem de consulta está associada às imagens mais similares conforme avaliação dos peritos do escritório de patentes do Reino Unido [EAK98], como apresentado na Tabela 4.1.

4.7.1 Redes Neurais Artificiais (RNA)

Redes Neurais Artificiais com aprendizagem supervisionada possuem um conjunto de variáveis de entrada e um conjunto de saída esperado. A RNA compara o valor de saída com o valor desejado, fazendo correções ao modelo de forma a encontrar um erro aceitável. Após a etapa de treinamento, pode-se apresentar um novo conjunto de entrada desconhecido à RNA e sua tarefa é classificar este novo padrão corretamente. Conceitos de RNAs foram apresentados no Capítulo 2.

Com base nos resultados individuais apresentados na Tabela 5.11 da Seção 5.3, foram construídos os conjuntos de treinamento e validação das RNAs. O número de MRVs para cada conjunto é apresentado na Tabela 4.3.

Tabela 4.3: Composição dos conjuntos de treinamento, validação e teste – RNA

RNAs	Conjuntos	No. de imagens por classe de extrator				Total
		SIFT	CC	FCC	Hu	
SIFT	Treinamento	99	20	20	20	145
	Validação	42	10	10	3	65
CC	Treinamento	62	20	20	6	108
	Validação	27	9	9	3	48
FCC	Treinamento	34	11	11	6	62
	Validação	14	4	4	3	25
Hu	Treinamento	6	2	2	2	12
	Validação	3	1	1	1	6
Total						287
	Teste					20
Conjunto Verdade						307

Na Tabela 4.3 o total de imagens de MRVs utilizadas para compor os conjuntos é de 307, sendo o total de imagens contidas no "conjunto verdade" que compõe a base de dados (Seção 4.3). Das 307 imagens de MRVs, 20 são utilizadas para compor o conjunto de teste e as 287 imagens restantes compõem os conjuntos de treinamento e

validação. A distribuição destas imagens para cada RNA construída consta da Tabela 4.3.

Os conjuntos de treinamento e validação foram construídos considerando o total de imagem pertencente a cada classe de extrator, sem considerar as 20 MRV de teste. A divisão dos conjuntos de treinamento e validação foram respectivamente 70% e 30% das imagens de cada classe. Por exemplo, para o extrator SIFT, 141 imagens de MRVs pertencem a esta classe então 70% representa 99 imagens e 30% representa 42 imagens (Tabela 4.3). Adicionalmente aos conjuntos de treinamento e validação, foram inseridas imagens pertencentes as outras três classes de extratores de maneira a realizar o treinamento da RNA apresentando exemplos de MRV que não pertencem a classe que está sendo treinada.

O treinamento das redes neurais foi realizado usando uma rede MLP (*Multi-Layer Perceptron*) com camada escondida, sendo que o número de neurônios na camada escondida e o número de épocas são exibidos na Tabela 4.4. O algoritmo de aprendizagem utilizado é *Back Propagation* padrão, com parâmetro de aprendizagem 0,2. Os pesos foram inicializados aleatoriamente com valores entre -1 e 1. Para casos onde uma característica não foi encontrada, assume-se o valor de 0,0001. Este valor justifica-se pelo fato de as redes neurais serem aversivas a zeros, pois as ligações entre os neurônios são basicamente multiplicações e o uso desses valores retarda o processo de aprendizagem. Assim, assume-se um valor pequeno, próximo de zero, que após a normalização dos vetores de características ainda seja menor que qualquer outro valor encontrado nos vetores.

Tabela 4.4: Configurações RNAs

RNA	No. de entradas	No. Neurônio na camada escondida	No. de épocas
Conc./Conv.	24	12	2000
FCC	8	4	14000
M. Hu	7	4	1200
SIFT	128	64	800

Observa-se na Tabela 4.4, que o número de neurônios na camada escondida é a metade do número de neurônios contidos na camada de entrada. O número de épocas de treinamento é variável conforme o extrator utilizado em cada RNA, tais valores foram obtidos observando a curva de aprendizagem apresentada pelo JNNS no decorrer do treinamento.

A Figura 4.5 apresenta uma visão geral da construção das RNAs sob a premissa de se constituir Múltiplos Classificadores. Deste modo, cada RNA treinada é especialista em reconhecer uma classe de extrator. Cada RNA tem como saída valores 1 e 0; sendo que o valor 1 representa imagem "reconhecida" e o valor 0 representa imagem "não reconhecida" para a classe de extrator avaliada. Dada uma MRV de consulta, são extraídas informações utilizando os quatro extratores descritos anteriormente - Concavidade/Convexidade, Cadeias de Freeman (FCC), Momentos de Hu e SIFT. Os vetores de características são enviados, normalizados, para as respectivas RNAs (Figura 4.5).

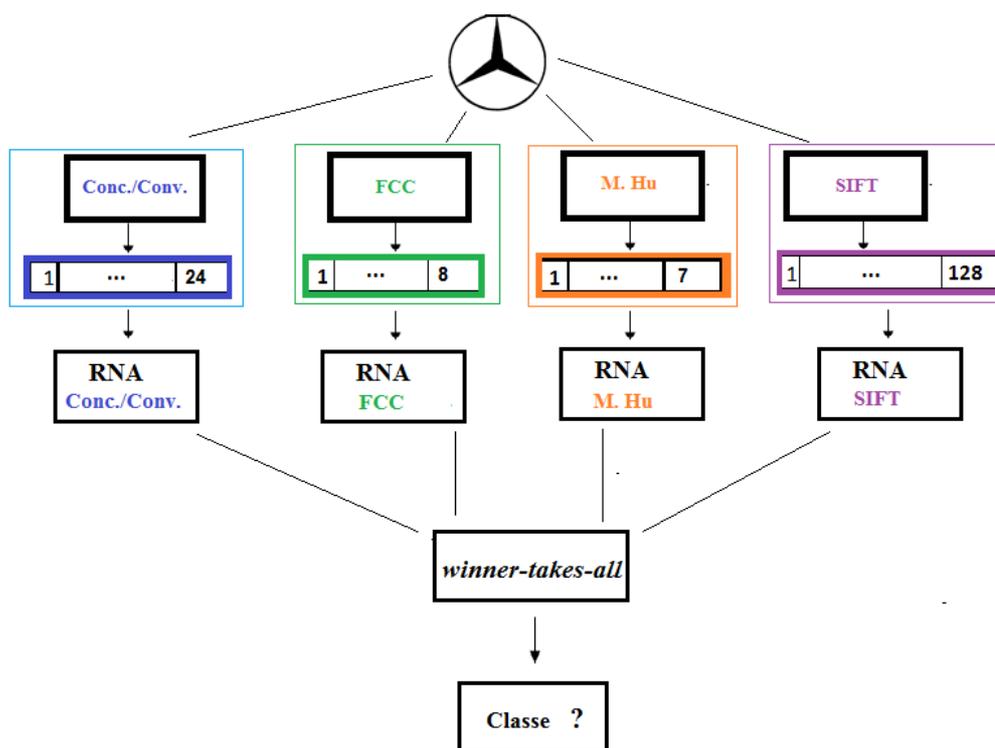


Figura 4.5: RNAs Múltiplos Classificadores.

Na Figura 4.5, uma MRV de consulta é enviada para as 4 RNAs, cada RNA é especializada e pode reconhecer uma classe de extrator. Após o processamento, obtém-se um valor que representa o "voto" de cada uma das redes. A combinação do "voto" de cada RNA foi realizada utilizando-se Voto Majoritário, conforme Seção 2.5.3 e ilustrado na Figura 4.5. O maior valor é considerado como a classe vencedora (*Best extractor*).

Foram executados experimentos com uma única RNA com o objetivo de classificar as quatro classes de extratores, porém estes não obtiveram melhores resultados que a RNA Múltiplos classificadores.

4.7.2 Máquinas de Vetor Suporte (SVM)

O classificador SVM utilizado neste trabalho foi utilizado por meio da ferramenta LIBSVM amplamente utilizado e disponível em [CHA11c] [CHA15].

Para construção do classificador SVM, é necessário definir conjuntos de treinamento e teste. Estes conjuntos foram criados utilizando informações das 307 imagens de MRVs contidas no "conjunto verdade" (Tabela 4.1). O conjunto de treinamento da SVM, que possui 287 imagens foi construído com base nos resultados individuais apresentados na Tabela 5.11 da Seção 5.3, tal qual a metodologia utilizada para a construção dos conjuntos das RNAs. As saída para cada classe apresentada para a SVM estão distribuídas conforme apresentado na Tabela 4.5.

Tabela 4.5: Composição conjuntos treinamento e teste - SVM

Classes	Treinamento	Teste
Conc./Conv.	89	20
FCC	48	20
M. de Hu	9	20
SIFT	141	20
TOTAL	287	20

Diferentemente das RNAs, somente um classificador SVM foi construída para o problema de classificação. A composição do vetor de características é realizada pela união dos vetores de características dos quatro extratores e a classe a qual cada MRV pertence, conforme apresentado na Figura 4.6.

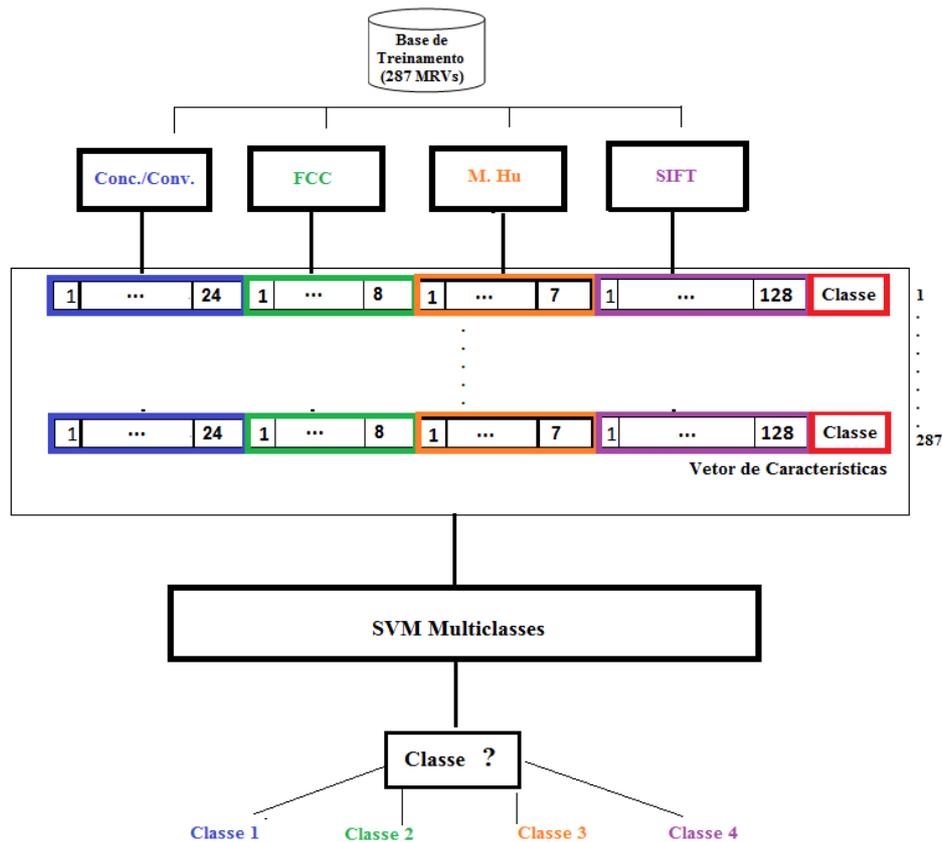


Figura 4.6: Construção SVM Multiclasses - Treinamento

A Figura 4.6 apresenta a composição do conjunto de treinamento da SVM Multiclasses utilizada nos experimentos, sendo treinada e testada para as quatro classes que representam os quatro extratores aplicados ao Modelo. A classe 1 representa o extrator Concavidade/Convexidade (vetor de dimensão 24), a classe 2 representa Cadeias de Freeman - FCC (vetor de dimensão 8), a classe 3 representa Momentos Invariantes de Hu (vetor de dimensão 7) e a classe 4 representa o extrator SIFT (vetor de dimensão 128). O vetor de características possui uma dimensão total igual a 168.

As 20 imagens de consulta foram inseridas somente no arquivo de teste. A composição do vetor de características para o conjunto de teste (Tabela 4.1) é igual ao procedimento adotado para formar o conjunto de treinamento (Figura 4.6). Como resultado final, a SVM indicará o melhor extrator a ser utilizado para a MRV de consulta.

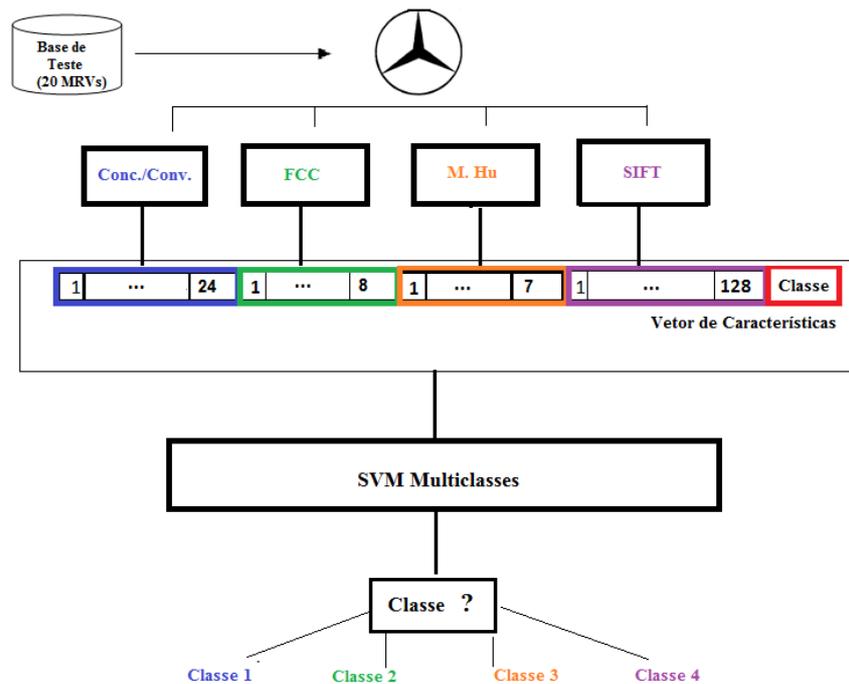


Figura 4.7: Construção SVM Multiclasses - Teste

Na Figura 4.7, uma MRV contida no conjunto de teste é enviada para classificação. Inicialmente são extraídas as características utilizando-se os 4 extratores e formando-se o vetor de características da MRV. A seguir, as informações contidas no vetor de características são enviadas para a SVM Multiclasses. Ao final, a classe de extrator que melhor representa a MRV é determinada. Os resultados obtidos e a matriz de confusão são apresentadas na Seção 5.5.

4.8. Mecanismos de Zoneamento

Conforme apresentado na Seção 2.6, os mecanismos de zoneamento permitem uma análise de informações locais em um dado padrão particionado, sendo que a análise de informações locais permite evidenciar elementos-chave no processo de recuperação.

A ideia apresentada por [FRE07] é aplicável com sucesso em reconhecimento de caracteres e consiste em analisar a relação entre as regiões e as confusões, permitindo compreender quais regiões dos caracteres estão causando as confusões. Neste trabalho, busca-se utilizar as regiões consideradas como confusões, para evidenciar as similaridades entre as imagens de MRVs. Dessa forma, pode-se utilizar as similaridades

em cada região para evidenciar características locais, importantes na recuperação de MRVs similares.

As Figuras 4.8 a 4.11 apresentam imagens de MRVs conforme o zoneamento proposto em [FRE07], sendo que tais zoneamentos tem o objetivo de criar um centro de atenção em informações locais contidas na MRVs.

A Figura 4.8, apresenta um zoneamento simétrico em quatro regiões. Este zoneamento divide a imagem de MRV em quatro regiões de mesmo tamanho.

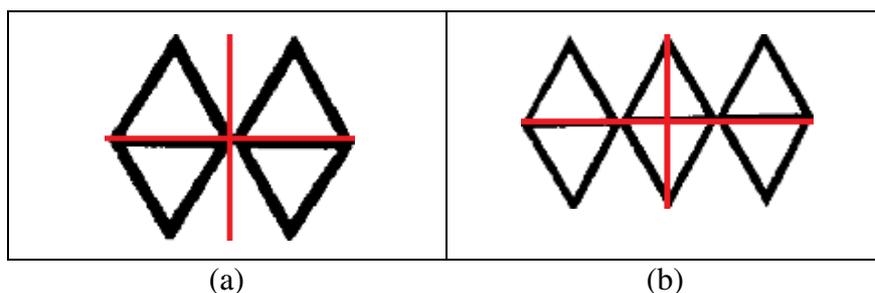


Figura 4.8: Zoneamento perceptivo em quatro regiões.

A Figura 4.9 apresenta um zoneamento em sete regiões. O objetivo deste zoneamento é evidenciar o centro de atenção nas informações contidas no centro da MRV. Esta informação é facilmente observada, quando o foco de atenção está nas três divisões realizadas no meio da MRV.

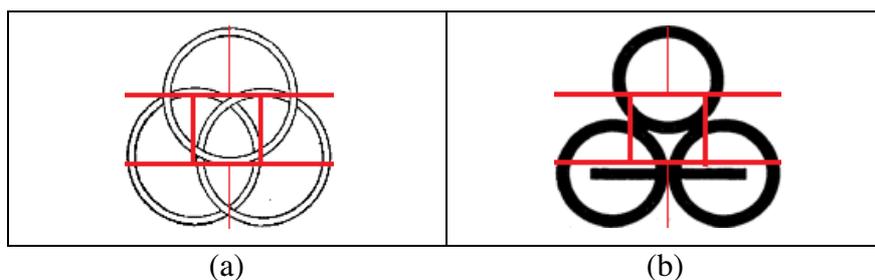


Figura 4.9: Zoneamento perceptivo em sete regiões.

Realizando uma análise global na Figura 4.10-a e Figura 4.10-b, temos um triângulo. Entretanto, as informações contidas na parte inferior da MRV evidenciam suas diferenças, este fato pode ser observado quando aplicado o zoneamento em cinco regiões horizontal.

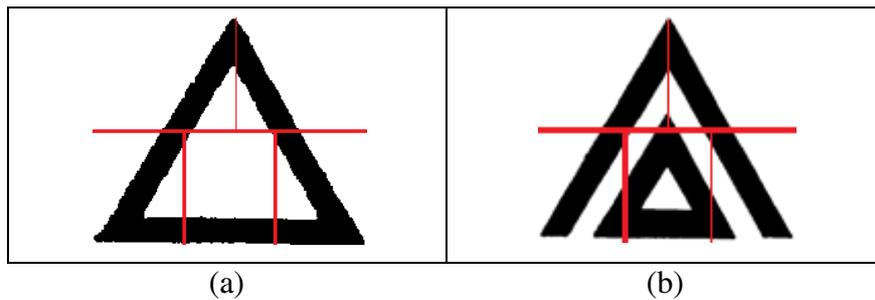


Figura 4.10: Zoneamento perceptivo em cinco regiões horizontal.

A Figura 4.11, apresenta um zoneamento em cinco regiões vertical, com o objetivo de ressaltar o centro de atenção para as informações contidas na parte direita da MRV.

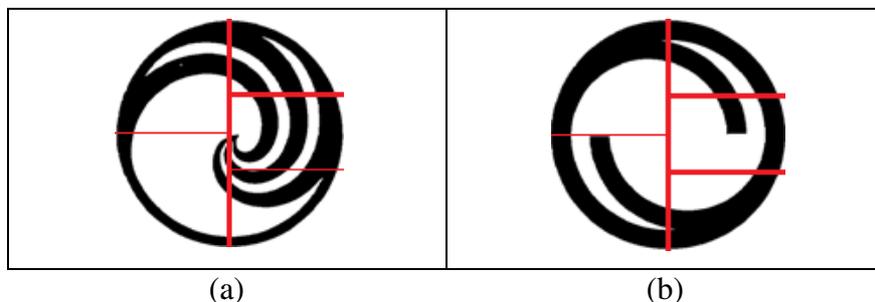


Figura 4.11: Zoneamento perceptivo em cinco regiões vertical.

Após a definição do melhor extrator, os zoneamentos perceptivos são aplicados às MRVs. Os resultados obtidos nos experimentos são apresentados e discutidos no Capítulo 5.

É importante observar que, para o extrator Momentos Invariantes de Hu não é possível aplicar os zoneamentos perceptivos, uma vez que é necessário conhecer a localização dos pontos/pixels na imagem para determinar as regiões e isso não pode ser determinado por meio deste tipo de extrator visto que este não trata a forma propriamente dita, mas a da distribuição dos pontos que formam a imagem.

Os zoneamentos perceptivos são utilizados neste trabalho não somente para avaliar o padrão particionado, mas também para realizar cálculo de similaridade considerando as regiões perceptivas, conforme apresentado na Seção 4.9.2.

Para determinar o melhor zoneamento (*Best Zoning*) a ser aplicado na MRV de consulta são calculadas as medidas de similaridade apresentadas na Seção 4.9.

4.9. Medidas de Similaridade

Uma parte importante na recuperação de uma MRV é a existência de um método eficiente de correspondência (*matching*) por meio da determinação de similaridades entre uma forma consultada e as formas existentes na base de dados. Assim, a correspondência define uma pontuação de similaridade/dissimilaridade entre a imagem de consulta e as imagens contidas no banco de dados. O resultado da similaridade é um *ranking* das imagens mais similares à imagem de consulta.

O julgamento de similaridade entre a MRV de consulta e as MRVs da base de dados é baseado na distância entre os vetores de características extraído das imagens. Pode-se observar na Tabela 2.3 que a Distância Euclidiana é uma das medidas utilizadas nos trabalhos de recuperação de MRVs.

Nos experimentos deste trabalho foram utilizadas Distância Euclidiana, Distância de Mahalanobis e Distância entre histogramas visando estabelecer o grau de similaridade entre as imagens de MRVs (ver Seção 2.7); sendo proposto para o extrator SIFT uma medida de similaridade baseada em regiões perceptivas, conforme Seção 4.9.2. As medidas de similaridade calculadas foram baseadas nas distâncias de seus vetores de características, com ou sem zoneamento. No entanto, a Distância Euclidiana obteve melhores resultados, conforme apresentado e discutido no Capítulo 5. Os demais resultados são apresentados nos Apêndices A e B.

Com o objetivo de complementar o cálculo de similaridade é apresentado neste trabalho um cálculo baseado nos zoneamentos perceptivos [FRE07]. Este cálculo somente pode ser aplicado em extratores de características invariantes à rotação e translação e que possuem informação de localização dos pontos/pixel considerados "estratégicos" nas imagens ou que definem o "foco de atenção". Dessa forma, sabe-se exatamente em qual região perceptiva cada ponto considerado relevante está localizado.

Os extratores de características que não possuem invariância a rotação e translação, obterão informações distintas para cada região na imagem avaliada inviabilizando a comparação direta destas regiões.

As informações obtidas pelo extrator SIFT são interessantes, pois é possível conhecer a localização de cada ponto-chave na imagem. Assim, pode-se determinar em qual zoneamento perceptivo o ponto-chave está posicionado, sendo fácil encontrar o número de pontos-chave por zoneamento e, então, determinar em qual região perceptiva

é concentrada o maior número de informações relevantes. A Figura 4.12 apresenta uma MRV com o zoneamento perceptivo aplicado.

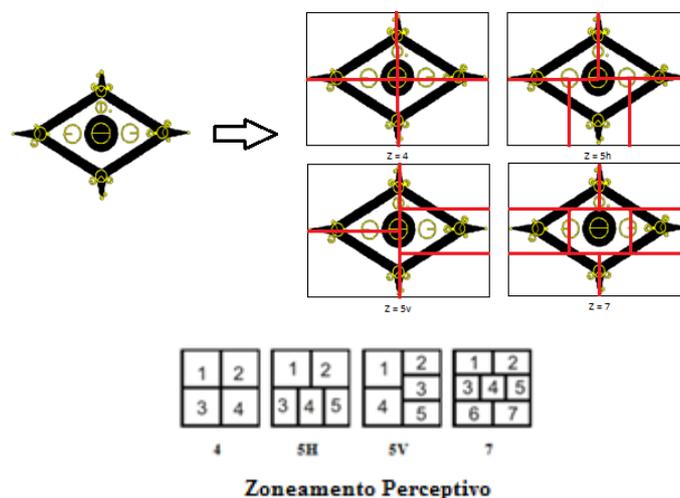


Figura 4.12: Zoneamento perceptivo - MRV da base de dados

Observa-se na Figura 4.12, uma MRV conforme os zoneamentos perceptivos propostos por [FRE07], cada zoneamento busca evidenciar uma região específica na imagem da MRV (ver Seção 4.8).

Como apresentado anteriormente, para o extrator SIFT utiliza-se para cálculo de similaridade a Distância Mínima entre pontos-chave e a Distância Mínima entre Zoneamentos Perceptivos, conforme Seções 4.9.1 e 4.9. 2.

4.9.1 Distância Mínima entre Pontos-Chave

Para calcular a similaridade entre duas imagens, a distância mínima entre o ponto-chave de imagem de consulta (I_a) e os pontos-chave das imagens contidas na base de dados (I_b) é computada. Para cada descritor em I_a , o descritor mais próximo em I_b (aplicando Distância Euclidiana) é encontrado. A Figura 4.13 mostra a correspondência dos pontos-chave utilizando a imagem de consulta I_a com duas imagens da base de dados. A distância mínima entre cada par de pontos-chave é usada para determinar a pontuação final.

Tal cálculo é similar ao proposto pela equipe de desenvolvedores da plataforma VLFEAT¹⁵, por meio da função `vl_ubcmatch()`¹⁶. A plataforma VLFEAT, é

¹⁵ <http://www.vlfeat.org>

¹⁶ http://www.vlfeat.org/matlab/vl_ubcmatch.html

desenvolvida por um time de pesquisadores de Visão Computacional, criado por Andrea Vedaldi e Brian Fulkerson no ano de 2007. Essa equipe desenvolve uma coleção de algoritmos de Visão Computacional *open source* disponíveis para a comunidade no site <http://www.vlfeat.org>. A função `vl_ubcmatch()`, é disponibilizada na plataforma VLFEAT, sendo que tal função foi desenvolvida conforme sugerido por Lowe et. al. [LOW04].

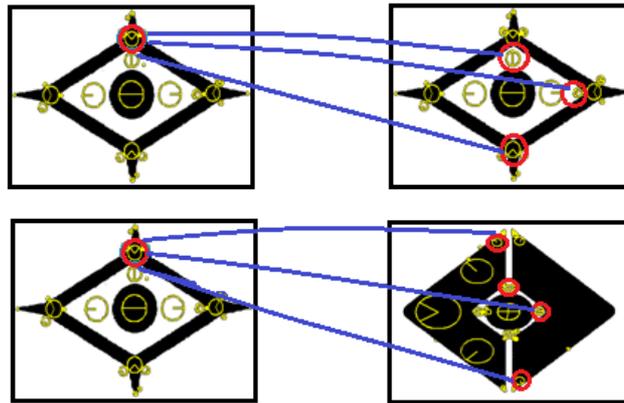


Figura 4.13: Correspondência entre pontos-chave.

Na Figura 4.13 pode-se observar que os pontos-chave estão localizados nos pontos com destaque amarelo. Nestes é possível determinar exatamente a posição de cada ponto-chave, e assim encontrar os pontos mais similares e quais são suas coordenadas (x,y) nas imagens.

Sendo a imagem I composta de n pontos-chave, $I = \{P_1, P_2, \dots, P_n\}$, onde P_i é formado por 128 características SIFT, sendo $P = \{x_1, x_2, \dots, x_n\}$.

O valor da similaridade entre cada ponto $P = \{x_1, x_2, \dots, x_n\}$ da imagem I_1 e cada ponto $Q = \{y_1, y_2, \dots, y_n\}$ da imagem I_2 é calculado pela Distância Euclidiana, conforme Equação 4.1.

$$d(P, Q) = \sqrt{\sum_{i=1}^{128} (x_i - y_i)^2} \quad (4.1)$$

A distância final é calculada pela soma das menores distâncias entre os pontos-chave das MRVs, conforme Equação 4.2.

$$d(I_1, I_2) = \sum_{i=1}^n \min_{j=1..n} d(P_i, Q_j) \quad (4.2)$$

Os resultados obtidos, utilizando a distância mínima entre pontos-chave, foram superiores ao cálculo da Distância Euclidiana considerando somente os vetores de características das imagens. Pode-se verificar tais resultados no Apêndice A na Tabela A.6.

4.9.2 Distância Mínima entre Zoneamentos Perceptivos

Neste trabalho é proposto observar os resultados obtidos pelo extrator SIFT usando zoneamento perceptivo. O vetor de características para cada imagem é construído considerando a localização dos pontos-chave de acordo com a sua localização no zoneamento.

A ideia é calcular a similaridade, fazendo uma comparação entre as regiões perceptivas. O extrator SIFT utiliza uma abordagem local, propõem-se realizar uma observação global do padrão dos pontos-chave que compõem a região, tendo como objetivo encontrar regiões perceptivas mais próximas, não somente os pontos-chave mais próximos.

Assim, calcula-se a média dos descritores de pontos-chave que pertencem a mesma região de zoneamento (região perceptiva). Para $Z = 4$ é gerado um vetor de dimensão $128 \times 4 = 512$ descritores. Para $Z = 5H$ e $Z = 5V$ são gerados vetores com dimensão de $128 \times 5 = 640$ descritores e para $Z = 7$ o vetor apresenta dimensão de $128 \times 7 = 896$ descritores. Um exemplo destes vetores de características aplicados ao zoneamento $Z = 4$ e $Z = 5V$ está representado na Figura 4.14.

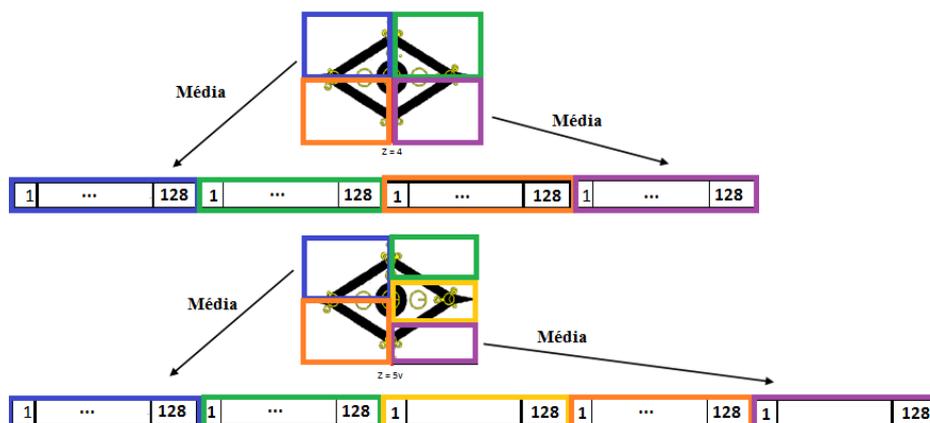


Figura 4.14: Vetor de características - Zoneamento perceptivo.

Considerando o que foi explicado anteriormente, calcula-se a similaridade com base em cada região do zoneamento proposto. A Figura 4.15 apresenta a similaridade por regiões, considerando um zoneamento $Z = 4$.

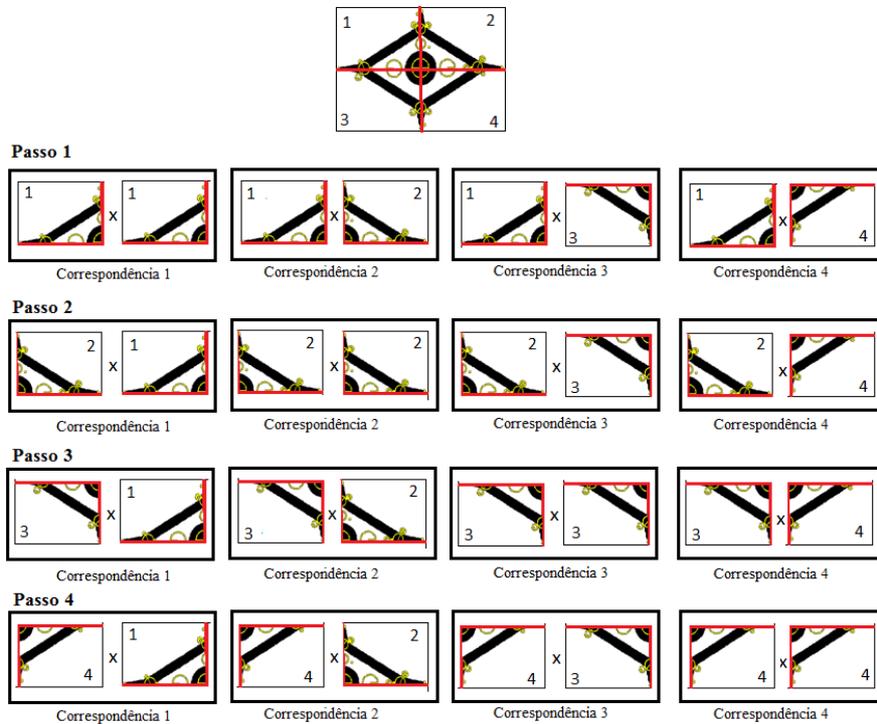


Figura 4.15: Correspondência entre Zoneamentos Perceptivos

Na Figura 4.15, observa-se também a comparação entre as quatro regiões. No passo 1, a região 1 é comparada com as regiões 1, 2, 3 e 4. No passo 2, a região 2 é comparada com regiões 1, 2, 3 e 4, e assim sucessivamente. A distância mínima é então calculada pela Distância Euclidiana. O número de regiões depende do zoneamento perceptivo que será aplicado. A seguir, buscam-se as imagens mais similares em toda a base de dados.

Sendo a imagem I particionada em n zonas, tem-se $I = \{Z_1, Z_2, \dots, Z_n\}$, onde cada zona Z_i é composta por m pontos-chave, $Z = \{p_1, p_2, \dots, p_m\}$. Cada ponto-chave P_i é formado por 128 características SIFT, $P_i = \{x_1^i, x_2^i, \dots, x_{128}^i\}$.

Para obter um único vetor de características para cada zona Z_i que compõe a imagem I , calcula-se a média entre os pontos daquela região conforme ilustrado na Figura 4.16.

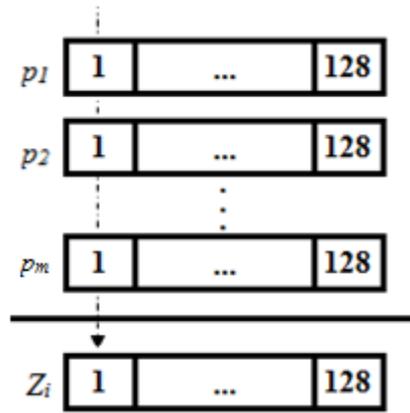


Figura 4.16: Cálculo da média por zona z_i

Sendo $Z_i = \{x_{m1}^i, x_{m2}^i, \dots, x_{m128}^i\}$ e x_m^i a média das características entre os pontos $\{p_1, p_2, \dots, p_m\}$ que compõem cada zona Z_i da imagem I .

A distância final entre as imagens $I_1 = \{K_1, K_2, \dots, K_n\}$ e $I_2 = \{W_1, W_2, \dots, W_n\}$ é calculada pela soma das menores distâncias entre os zoneamentos das MRVs, conforme Equação 4.3.

$$d(I_1, I_2) = \sum_{i=1}^n \min_{j=1..n} d(K_i, W_j) \quad (4.3)$$

Sendo que, $d(K_i, W_j)$ denota a Distância Euclidiana conforme Equação 4.1.

É importante ressaltar que neste trabalho utilizamos quatro tipos de zoneamentos, $Z=4, 5h, 5v$ e 7 .

Exemplificando na Equação 4.3, supondo um zoneamento em 4 regiões ($Z=4$), tem-se a Equação 4.4.

$$d(I_1, I_2) = \sum_{i=1}^4 \min_{j=1..4} d(K_i, W_j) \quad (4.4)$$

Desta forma, é possível determinar as regiões mais similares uma vez que elas possuem a menor distância. Os resultados dos experimentos são apresentados no Capítulo 5.

4.10. Considerações Finais

Neste capítulo, foi descrito o sistema proposto para a recuperação de MRVs utilizando regiões perceptivas. O modelo hierárquico perceptivo é composto das etapas de: pré-processamento, extração de características, classificação, seleção dinâmica de extratores (*best-extractor*), zoneamento perceptivo, seleção do *best-zoning* e medidas de similaridade. Uma primitiva baseada nos princípios da Gestalt é apresentado, com o objetivo de complementar as abordagens já existentes. Um cálculo de similaridade, utilizando os zoneamentos perceptivos é proposto. No próximo capítulo são apresentados os experimentos realizados e os resultados obtidos.

Capítulo 5

Experimentos

5.1. Considerações Iniciais

Para validar os resultados obtidos com o Modelo Hierárquico Perceptivo apresentado no Capítulo 4, foi necessária a realização de experimentos com o objetivo de provar conceitos, atentando-se ao fato de que diferentes variáveis afetam os resultados das pesquisas, pois as MRVs muitas vezes possuem características geométricas abstratas. Outro problema a ser considerado é a diferença de escala, rotação e translação entre as imagens que formam a base de dados. Ainda, o segundo plano das imagens (*background*) pode aumentar a complexidade da tarefa de recuperação de MRVs, uma vez que MRVs no "mundo-real" podem ter o segundo plano (*background*) complexo, por exemplo: fotos e vídeos em redes sociais, CAPTCHA (*Completely Automated Public Turing test to tell Computers and Humans Apart*), vídeos esportivos, filmes, entre outros.

Assim, nos experimentos foram utilizadas MRVs contidas na base de dados do escritório de patentes do Reino Unido, como já mencionado anteriormente, e estas, não possuem segundo plano (*background*) complexo. Deve-se considerar também o tamanho da base de dados, a abordagem de classificação adotada e as características das bases de dados utilizada para treinamento e teste do protocolo experimental. A seguir são apresentados os resultados obtidos nos experimentos realizados, além disso são analisados tais resultados a fim de permitir o entendimento dos ganhos do sistema proposto e as dificuldades encontradas.

5.2. Experimentos

5.2.1 Experimentos utilizando Deficiências de Concavidade e Convexidade

A extração de concavidade e convexidade evidencia as propriedades topológicas e geométricas das formas, analisando partes côncavas e convexas das MRVs. Neste trabalho, utiliza-se também a extração de concavidade e convexidade para evidenciar características de primeiro e segundo planos, com base no princípio da Gestalt denominado **Figura-Fundo**.

Para os experimentos utilizando Concavidade e Convexidade aplicam-se nas imagens de MRV, contidas na base de dados, as etapas de pré-processamento: limiarização e *bouding box*. Tais etapas de pré-processamento foram apresentadas na Seção 4.5.

A Tabela 5.1 apresenta os resultados obtidos para cinco experimentos, utilizando o extrator de Concavidade e Convexidade. O primeiro experimento não aplica nenhum tipo de zoneamento ($Z=0$), enquanto que os demais utilizam os zoneamentos perceptivos ($Z=4, 5h, 5v$ e 7). As taxas apresentadas adotam as medidas de recuperação *Normalized recall* (R_n) e *Normalized precision* (P_n).

Avaliando os resultados apresentados na Tabela 5.1, sem aplicar zoneamento perceptivo ($Z=0$), observa-se que as MRVs 1279931.bmp (), 1439229_.bmp (), 2042822_.bmp () e 967049.bmp () obtiveram melhores resultados para R_n . No entanto, observa-se que aplicando os zoneamentos, os resultados obtidos para 16 (dezesesseis) das 20 (vinte) MRVs consultadas foram melhores que os resultados obtidos sem zoneamento. Em uma análise considerando $Z=0$, a MRV 1138293_.bmp () obteve o melhor resultado sendo $R_n = 0,955971$, enquanto que a MRV 2018809_.bmp () obteve o menor resultado para $R_n = 0,475659$. Tais imagens são muito distintas entre si, sendo importante observar o preenchimento de pixels pretos. A MRV 2018809_.bmp () possui um primeiro plano mais "denso" e poucas informações de segundo plano são utilizadas para compor o vetor de características, o que influencia negativamente os resultados.

Tabela 5.1: Deficiências de Concavidade e Convexidade

MRV		Z=0		Z=4		Z=5h		Z=5v		Z=7		Best
		R _n	P _n									
1037814.bmp		0,913979	0,782846	0,917314	0,762984	0,923221	0,786305	0,917584	0,783546	0,923100	0,774441	0,923221
1055261.bmp		0,897219	0,799174	0,911045	0,822548	0,909472	0,818118	0,894974	0,786815	0,865497	0,766886	0,911045
1138103.bmp		0,737104	0,435358	0,812021	0,520611	0,811405	0,527783	0,814224	0,528480	0,851087	0,562724	0,851087
1138293_.bmp		0,955971	0,859226	0,991243	0,899630	0,990405	0,897357	0,991243	0,899616	0,990652	0,897995	0,991243
1190540.bmp		0,851790	0,691153	0,880515	0,642869	0,921053	0,724056	0,885327	0,643001	0,905818	0,659493	0,921053
1259886.bmp		0,848089	0,542946	0,916820	0,697936	0,918092	0,701132	0,901978	0,686491	0,866385	0,607781	0,918092
1267206.bmp		0,677049	0,497747	0,729084	0,529420	0,718567	0,469719	0,738139	0,498974	0,730832	0,519419	0,738139
1279931.bmp		0,953005	0,721983	0,925494	0,694398	0,929644	0,704843	0,926351	0,679819	0,931582	0,703108	0,953005
1289047.bmp		0,747940	0,550679	0,825443	0,503299	0,840423	0,509318	0,827632	0,527055	0,813889	0,528199	0,840423
1376861.bmp		0,794684	0,463138	0,905872	0,588321	0,913152	0,599321	0,915806	0,596344	0,911718	0,590614	0,915806
1439229_.bmp		0,677129	0,532709	0,506026	0,503031	0,508710	0,503475	0,509959	0,503592	0,528769	0,506607	0,677129
1486213_.bmp		0,680433	0,333413	0,762888	0,409133	0,744234	0,366673	0,770837	0,418284	0,734799	0,380268	0,770837
1575268_.bmp		0,654612	0,511416	0,674113	0,509097	0,707316	0,535079	0,678999	0,510881	0,661593	0,511883	0,707316
2010916_.bmp		0,816665	0,494341	0,863391	0,556050	0,871914	0,571378	0,867837	0,559927	0,816451	0,551183	0,871914
2016658_.bmp		0,659382	0,284570	0,705319	0,339452	0,724799	0,417170	0,733135	0,419806	0,756254	0,440499	0,756254
2018809_.bmp		0,475659	0,214888	0,547983	0,301919	0,544648	0,264543	0,530832	0,234504	0,498198	0,217590	0,547983
2042822_.bmp		0,758863	0,318822	0,706144	0,276825	0,685935	0,259610	0,715997	0,284798	0,722987	0,288364	0,758863
665322.bmp		0,832462	0,624937	0,863169	0,696605	0,837925	0,694613	0,856050	0,694654	0,863386	0,773361	0,863386
914.bmp		0,753039	0,527269	0,842490	0,571237	0,816295	0,560898	0,834093	0,576369	0,863311	0,588797	0,863311
967049.bmp		0,885958	0,871627	0,865320	0,867350	0,855350	0,838470	0,865791	0,867424	0,864364	0,831706	0,885958
Média		0,778552	0,552912	0,807585	0,584636	0,808628	0,587493	0,808839	0,585019	0,805034	0,585046	0,833303
Desvio Padrão		0,117453	0,184702	0,125187	0,173364	0,126273	0,177843	0,127017	0,175998	0,129089	0,176185	0,109458

Para o zoneamento em quatro regiões ($Z=4$), Tabela 5.1, três MRVs das 20 MRVs consultadas obtiveram melhores resultados comparados aos demais zoneamentos testados, sendo: 1055261.bmp (), 1138293_.bmp () e 2018809_.bmp (). Comparando os resultados individuais aplicando zoneamento em quatro regiões ($Z=4$) com os resultados sem zoneamento ($Z=0$), observa-se uma melhora significativa em R_n para 17 MRVs das 20 MRVs consultadas. Tal desempenho pode ser facilmente verificado em função da média geral, sendo $Z=0$ $R_n = 0,778552$ e para $Z=4$ $R_n = 0,807585$. Particularmente, a MRV 1138293_.bmp () melhorou significativamente seu resultado atingindo $R_n = 0,991243$ para $Z=4$, se comparado ao resultado obtido para $Z=0$ sendo $R_n = 0,955971$.

Observa-se ainda nos resultados apresentados na Tabela 5.1, que a MRV 1190540.bmp ( ) conseguiu melhorar seus resultados quando aplicado o zoneamento em cinco regiões horizontal ($Z=5h$). O valor obtido para $Z=4$ foi $R_n = 0,880515$ e $P_n = 0,642869$, atingindo para o zoneamento $Z=5h$ o valor para $R_n = 0,921053$ e $P_n = 0,724056$. A Média Geral obteve uma melhora pouco significativa, sendo $Z=4$ $R_n = 0,807585$ e $P_n = 0,587493$ e para $Z=5h$ $R_n = 0,808628$ e $P_n = 0,587493$.

No entanto, para a MRV 914.bmp () o zoneamento com o "foco de atenção" na parte inferior da MRV ($Z=5h$) não obtém melhores resultados, sendo que para $Z=4$ os valores são $R_n = 0,842490$ e $P_n = 0,571237$, e estes pioram aplicando $Z=5h$ sendo $R_n = 0,816295$ e $P_n = 0,560898$. Deve-se considerar que a MRV está em um *bounding box*, e a região inferior desta MRV contém muitos pixels pretos (1o. plano) e as características de fundo (2o. plano) não são suficientes para representar e permitir diferenciação do referido padrão. Deve-se ressaltar que das 20 MRVs consultadas seis MRVs obtiveram melhores resultados aplicando o zoneamento de cinco regiões horizontal ($Z=5h$). Dos zoneamentos aplicados no extrator de Concavidade e Convexidade, $Z=5h$ obteve o maior número de melhores resultados em relação aos demais zoneamentos testados.

Observando os resultados para o zoneamento em cinco regiões vertical ($Z=5v$), apresentado na Tabela 5.1, verifica-se que as imagens 1267206.bmp (), 1376861.bmp () e 1486213.bmp (), obtiveram resultados melhores em relação aos demais zoneamentos apresentados. Isso implica que a região direita da MRV, para estas imagens, obteve informações importantes que possibilitaram melhora nos

resultados apresentados. A MRV 1138293_.bmp () atingiu resultado igual ao obtido para o zoneamento em quatro regiões ($Z=4$), $R_n = 0,991243$, sendo este resultado o melhor valor obtido para a MRV em questão.

Na Tabela 5.1, as MRVs 1138103.bmp (), 2016658.bmp (), 665322.bmp () e 914.bmp () conseguem melhorar os resultados anteriormente apresentados, aplicando o zoneamento em sete regiões ($Z=7$). A MRV 1138103.bmp () obteve para o zoneamento $Z=5$ $R_n = 0,814224$ e $P_n = 0,528480$, atingindo com $Z=7$ $R_n = 0,851087$ e $P_n = 0,562724$. A imagem 2016658.bmp () obteve para $Z=5$ $R_n = 0,733135$ e $P_n = 0,419806$, obtendo com $Z=7$ valores para $R_n = 0,756254$ e $P_n = 0,440499$. E a MRV 914.bmp () atingiu para $Z=4$ valor de $R_n = 0,842490$ e $P_n = 0,571237$, melhorando os resultados para $Z=7$ sendo $R_n = 0,863311$ e $P_n = 0,588797$.

Comparando todos os zoneamentos aplicados para o extrator de Concavidade e Convexidade, ressalta-se a escolha do melhor zoneamento perceptivo (*Best*). A medida de recuperação utilizada para determinar o melhor zoneamento (*Best-Zoning*) a ser aplicado na MRV é *Normalized Recall* (R_n). Tal medida foi escolhida uma vez que R_n dá um maior peso para o sucesso na recuperação dos primeiros itens (Seção 2.8). Observa-se que obter o melhor zoneamento (*Best-Zoning*) para cada MRV aumenta a taxa atingida por R_n na média final.

Considerando a medida de recuperação R_n a MRV 2018809_.bmp () não obteve bons resultados utilizando o extrator de concavidade e convexidade, sendo o pior resultado entre as MRV consultadas. A MRV 1138293_.bmp () obteve o melhor resultado, atingindo entre as TOP-100 imagens mais similares resultados muito próximos ao conjunto verdade proposto pelos especialistas.

A Média Geral para R_n obtida pelos experimentos sem zoneamento foi inferior aos resultados obtidos usando zoneamento perceptivo. Isso deve-se ao fato de observar o padrão particionado, enfatizando regiões de interesse nas imagens de MRVs. A Média Geral sem zoneamento ($Z=0$) foi de $R_n = 0,778552$ e a Média Geral do melhor zoneamento perceptivo ($Z=5$) foi de $R_n = 0,808839$, sendo esta média muito próxima as médias dos demais zoneamentos aplicados.

Analisando os resultados para R_n as MRVs 1289047.bmp () e 2016658_.bmp () obtiveram melhoras significativas utilizando zoneamentos perceptivos em relação as MRVs sem zoneamento aplicado. No entanto, algumas MRVs como 1279931.bmp () e 2042822_.bmp () obtiveram melhores resultados sem aplicar o zoneamento perceptivo ($Z=0$).

A Figura 5.1 apresenta os resultados parciais obtidos pela MRV 1279931.bmp () sem os zoneamentos perceptivos ($Z=0$), sendo importante lembrar que os valores para R_n e P_n são calculados para Top-100 imagens retornadas. Entretanto, as MRVs apresentadas na Figura 5.1 são as Top-24 imagens consideradas mais similares. Se todas as imagens fossem exibidas pelo sistema em uma única tela as MRVs seriam muito pequenas dificultando a avaliação, por isso a apresentação feita pelo sistema é realizada em conjuntos de 24 MRVs.



Figura 5.1: Top-24 MRVs mais similares 1279931.bmp - $Z=0$

Observa-se nas MRVs similares apresentadas na Figura 5.1 que a região superior da MRV obtém informações importantes para a recuperação. No entanto, nenhum dos zoneamentos propostos tem o "foco de atenção" para essa região. Isso explica que o melhor resultado obtido para essa MRV foi utilizando informações da imagem toda.

A Figura 5.2 apresenta as MRVs mais similares a MRV 1138293_.bmp (). Tal MRV obteve bons resultados aplicando Concavidade/Convexidade para o zoneamento $5v$, sendo $R_n = 0,991243$ e $P_n = 0,899616$.

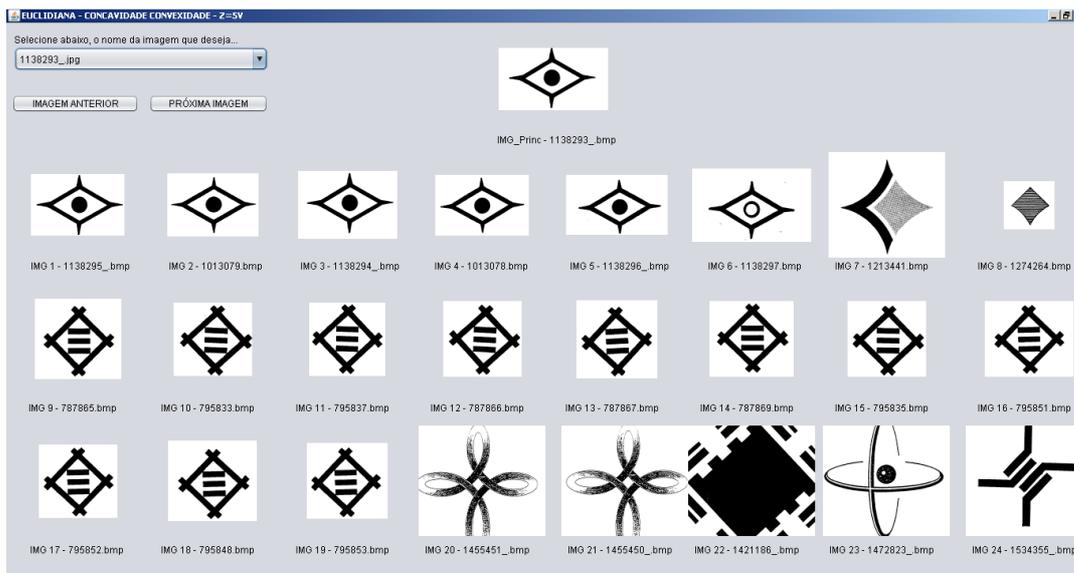


Figura 5.2: Top-24 MRVs mais similares 1138293_.bmp - Z=5v

Avaliando as MRVs similares apresentadas na Figura 5.2, é possível verificar que o extrator de características Concavidade/Convexidade desempenhou seu papel de forma satisfatória, uma vez que as imagens possuem regiões abertas acima/direita, acima/esquerda, abaixo/direita, abaixo/esquerda, e a região central possui laço fechado. Além disso, esta MRV obteve melhor taxa para R_n com zoneamento $Z=4$, dada a simetria particular dos elementos que compõem a MRV.

5.2.2 Experimentos utilizando Cadeias de Freeman (FCC)

O método de Cadeias de Freeman codifica informações de contorno da forma. Estas informações são relevantes no processo de reconhecimento/percepção de objetos, conforme apresentado neste trabalho com MRVs.

Para realizar os experimentos utilizando o extrator denominado Cadeias de Freeman aplicam-se nas imagens de MRV, contidas na base de dados, as etapas de pré-processamento: limiarização e *bouding box*.

A Tabela 5.2 apresenta os resultados obtidos para cinco experimentos utilizando o extrator de Cadeias de Freeman. O experimento inicial foi realizado sem os zoneamentos perceptivos, ou seja $Z=0$, os demais experimentos usam os zoneamentos ($Z=4$, $5h$, $5v$ e 7). As taxas apresentadas utilizam as medidas de recuperação *Normalized recall* (R_n) e *Normalized precision* (P_n).

Na Tabela 5.2, são apresentados os resultados dos experimentos utilizando Cadeias de Freeman. Considerando os experimentos realizados observa-se que sete

MRVs obtiveram melhores resultados sem utilizar os zoneamentos perceptivos ($Z=0$), para as outras 13 (treze) MRVs os melhores resultados foram obtidos por intermédio dos zoneamentos. Ainda, a MRV 914.bmp () obteve o melhor resultado entre as 20 MRVs consultadas, sendo $R_n = 0,968185$ e $P_n = 0,897077$. Este resultado é superior às taxas conseguidas para essa MRV utilizando o extrator de Concavidade/Convexidade ($R_n = 0,863311$). A MRV 1259886.bmp () obteve o menor resultado, para $Z=0$, sendo $R_n = 0,385766$ e $P_n = 0,168797$. Esta MRV (1259886.bmp - ) obteve melhores resultados com o extrator de Concavidade/Convexidade ($R_n = 0,918092$).

Tabela 5.2: Cadeias de Freeman

MRV		Z=0		Z=4		Z=5h		Z=5v		Z=7		Best
		R _n	P _n	R _n								
1037814.bmp		0,786716	0,671470	0,735814	0,550394	0,726180	0,553618	0,704129	0,555218	0,712802	0,625581	0,786716
1055261.bmp		0,792322	0,670208	0,796275	0,644120	0,769863	0,639253	0,838733	0,707400	0,851924	0,701442	0,851924
1138103.bmp		0,789920	0,509501	0,805323	0,372756	0,791383	0,354443	0,745381	0,382513	0,744222	0,384973	0,805323
1138293_.bmp		0,795050	0,755605	0,795346	0,766077	0,738261	0,533840	0,795030	0,769398	0,566887	0,244804	0,795346
1190540.bmp		0,891746	0,526413	0,978533	0,769044	0,978661	0,769147	0,940430	0,620914	0,849088	0,546287	0,978661
1259886.bmp		0,385766	0,168797	0,474768	0,216487	0,493127	0,220446	0,514331	0,221858	0,453709	0,215912	0,514331
1267206.bmp		0,885593	0,496955	0,899220	0,553407	0,884902	0,527754	0,872736	0,507216	0,860588	0,500349	0,899220
1279931.bmp		0,576236	0,402842	0,697555	0,479903	0,662669	0,330343	0,704563	0,464139	0,674190	0,292359	0,704563
1289047.bmp		0,725937	0,493850	0,842939	0,693950	0,850733	0,694277	0,792996	0,543084	0,762702	0,482991	0,850733
1376861.bmp		0,902968	0,769857	0,545114	0,267459	0,579254	0,274922	0,547149	0,252840	0,587879	0,272650	0,902968
1439229_.bmp		0,466881	0,259210	0,543472	0,332173	0,523040	0,329099	0,540327	0,336161	0,585114	0,354505	0,585114
1486213_.bmp		0,766058	0,408752	0,752942	0,430114	0,761248	0,410166	0,785705	0,478207	0,785280	0,438379	0,785705
1575268_.bmp		0,855858	0,490704	0,753183	0,501514	0,711034	0,393347	0,777441	0,532522	0,773753	0,490768	0,855858
2010916_.bmp		0,788071	0,499863	0,867253	0,429345	0,905291	0,450636	0,911028	0,447578	0,924310	0,457094	0,924310
2016658_.bmp		0,852163	0,388191	0,824992	0,419987	0,806435	0,406151	0,802116	0,419789	0,770505	0,405107	0,852163
2018809_.bmp		0,616720	0,302656	0,518155	0,278520	0,522629	0,279879	0,543948	0,282964	0,590748	0,292621	0,61672
2042822_.bmp		0,452776	0,163271	0,592982	0,234223	0,579641	0,214502	0,678367	0,252812	0,629554	0,246097	0,678367
665322.bmp		0,892959	0,683507	0,697752	0,645010	0,699073	0,645223	0,724366	0,590543	0,756129	0,550308	0,892959
914.bmp		0,968185	0,897077	0,881323	0,796558	0,890291	0,798146	0,863989	0,703193	0,800398	0,672825	0,968185
967049.bmp		0,849690	0,811136	0,782888	0,385342	0,742386	0,343343	0,823368	0,440893	0,855509	0,814165	0,855509
Média		0,752081	0,518493	0,739291	0,488319	0,730805	0,458427	0,745307	0,475462	0,726765	0,449461	0,805234
Desvio Padrão		0,165368	0,208886	0,139602	0,183753	0,138098	0,178143	0,126652	0,158182	0,123532	0,167965	0,126512

Avaliando os resultados obtidos para o zoneamento em quatro regiões, apresentados na Tabela 5.2, observa-se que quatro MRVs obtiveram melhores resultados em relação aos demais zoneamentos, sendo: 1138103.bmp () , 1138293_.bmp () , 1190540.bmp () e 1267206.bmp () . A imagem 1190540.bmp () obteve melhor resultado aplicando-se o zoneamento em quatro regiões, atingindo $R_n = 0,978533$ e $P_n = 0,769044$. Comparando com o resultado obtido por esta MRV utilizando FCC não aplicando zoneamento ($Z=0$) tem-se $R_n = 0,891746$ e $P_n = 0,526413$. Uma melhora significativa é obtida pela MRV 2010916_.bmp () para o zoneamento em quatro regiões ($Z=4$), tal MRV obteve para FCC sem zoneamento o valor para $R_n = 0,788071$ e $P_n = 0,499863$, e FCC com zoneamento em quatro regiões ($Z=4$) $R_n = 0,867253$ e $P_n = 0,429345$.

Na Tabela 5.2, considerando o zoneamento em cinco regiões horizontais ($Z=5h$) somente uma MRV obteve o melhor resultado entre os experimentos, sendo 1289047.bmp () obtendo $R_n = 0,850733$. Tal MRV também obteve o melhor resultado com este zoneamento ($Z=5h$) nos experimentos utilizando Concavidade/Convexidade apresentados na Tabela 5.1.

Para os experimentos com cinco regiões verticais ($Z=5v$), apresentados na Tabela 5.2, quatro das 20 MRVs obtiveram melhores resultados comparados aos demais experimentos, sendo: 1259886.bmp () , 1279931.bmp () , 1486213_.bmp () e 2042822_.bmp () . A MRV 2042822_.bmp () obteve para FCC associado ao zoneamento $Z=5v$ valor para $R_n = 0,678367$ e $P_n = 0,252812$, comparando com FCC associado ao zoneamento em sete regiões ($Z=7$) obteve $R_n = 0,629554$ e $P_n = 0,246097$. Para a imagem 1279931.bmp () , houve uma pequena melhora nos resultados sendo $R_n = 0,704563$ e $P_n = 0,464139$, em relação ao zoneamento $Z=4$ que obteve para $R_n = 0,697555$ e $P_n = 0,479903$. É importante analisar que a MRV 1486213_.bmp () também obteve o melhor resultado em Concavidade/Convexidade para o mesmo zoneamento $Z=5v$, isso define que o "foco de atenção" no lado esquerdo da MRV possui informações relevantes no processo de recuperação.

Observa-se na Tabela 5.2, para o zoneamento em sete regiões ($Z=7$), que a MRV 1055261.bmp () consegue resultados superiores aos demais zoneamentos, sendo $R_n = 0,851924$ e $P_n = 0,701442$, comparando com FCC associado $Z=5v$ que obteve $R_n = 0,838733$ e $P_n = 0,707400$. Importante observar que o zoneamento $Z=7$ não obteve

sucesso com a MRV 1138293_.bmp (), atingindo valores para $R_n = 0,566887$ e $P_n = 0,244804$. O melhor valor obtido para 1138293_.bmp (), foi com o zoneamento $Z=4$ sendo $R_n = 0,795346$ e $P_n = 0,766077$. Nota-se que quatro das 20 MRVs consultadas obtiveram melhores resultados com o zoneamento em 7 regiões ($Z=7$).

Realizando um comparativo da medida de recuperação *Normalized recall* (R_n) em relação aos zoneamentos perceptivos aplicados, é possível definir o melhor zoneamento (*Best*) a ser aplicado em cada uma das MRVs consultadas.

Considerando a medida de recuperação R_n as MRVs 1190540.bmp () e 914.bmp (), obtiveram bons resultados. Os conjuntos retornados das TOP-100 mais similares são muito próximos ao conjunto verdade proposto pelos especialistas. Tais resultados também são superiores aos obtidos utilizando o extrator de concavidade e convexidade.

A definição do melhor zoneamento para cada MRV consultada (*Best*), permitiu aumentar a Média Geral, atingindo $R_n = 0,805234$. Superando o valor conseguido para FCC sem zoneamento ($Z=0$) onde $R_n = 0,752081$. Das 20 MRVs consultadas oito MRVs obtiveram melhor resultado sem zoneamento. Considerando este fato, é importante observar que os demais zoneamentos não podem ser ignorados, pois a escolha do melhor zoneamento representa um ganho significativo na Média Geral.

A Figura 5.3 apresenta as MRVs mais similares a MRV 914.bmp () utilizando o extrator Cadeias de Freeman sem zoneamentos perceptivos ($Z=0$).

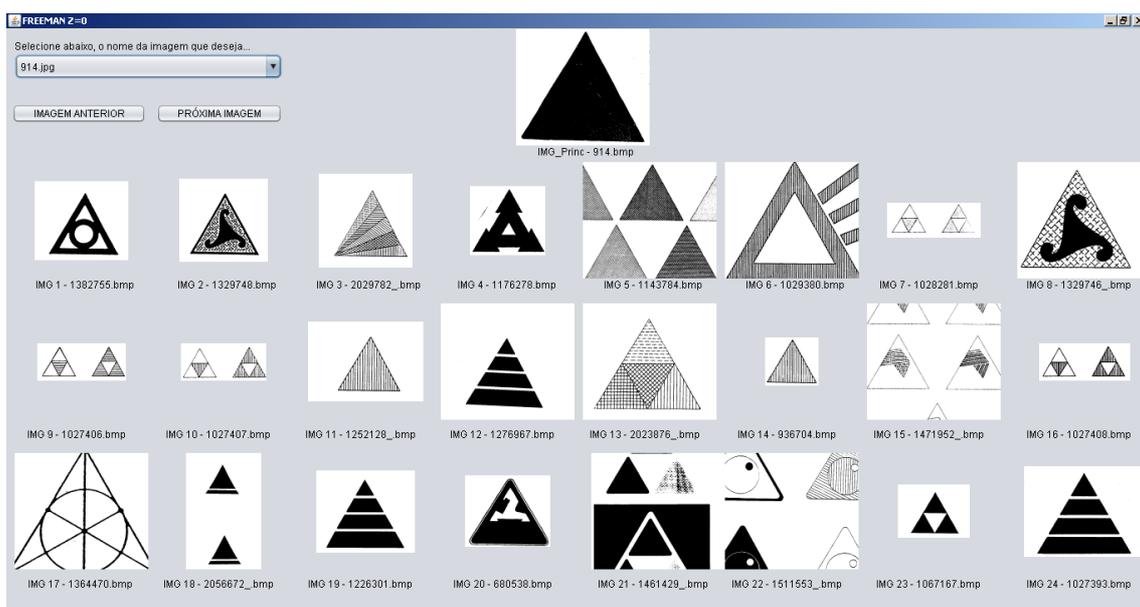


Figura 5.3: FCC Top-24 MRVs mais similares 914.bmp - $Z=0$

Observa-se na Figura 5.3 as MRVs similares a 914.bmp, o descritor de contornos obteve bons resultados para a MRVs consultada. O triângulo contido na MRV 914.bmp pode ser visualizado nas demais MRVs retornadas.

Na Figura 5.4 são apresentadas MRVs similares a MRV 1190540.bmp, por meio do extrator FCC associado ao zoneamento em quatro regiões perceptivas ($Z=4$).

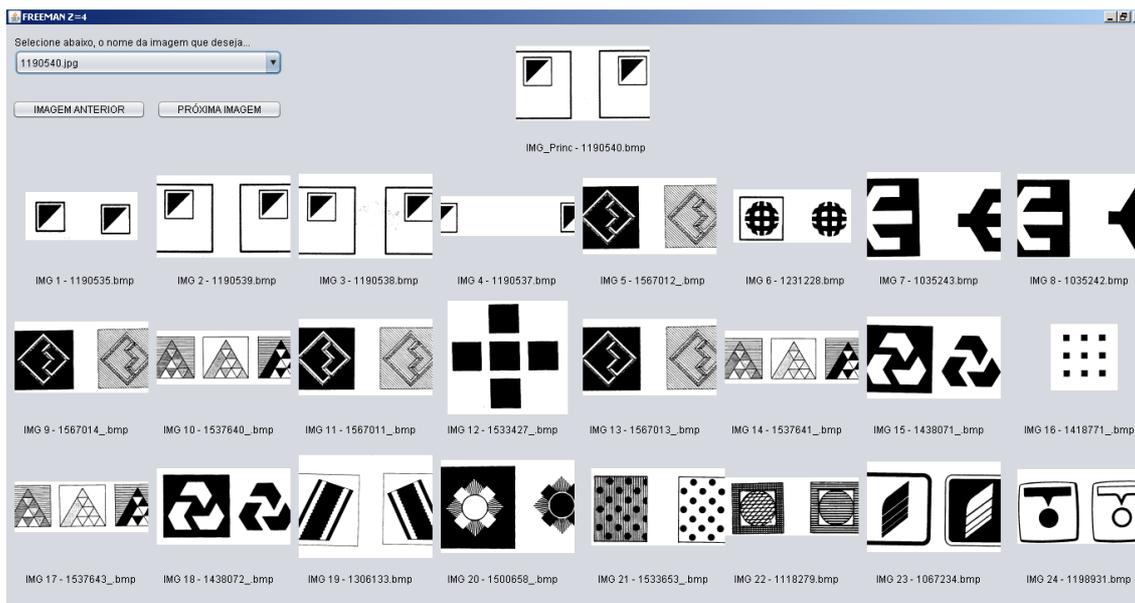


Figura 5.4: FCC Top-24 MRVs mais similares 1190540.bmp - $Z=4$

Observa-se na Figura 5.4 que as imagens similares obtêm bons resultados nas posições iniciais. No entanto, como esperado considerando a utilização de um extrator de contorno, predomina a forma da moldura quadrada contida nas demais MRVs retornadas. As variações ocorrem em relação às informações contidas "dentro" da moldura quadrada. Tais informações não são recuperadas avaliando somente o contorno da MRV, sendo necessário utilizar outro extrator de características para resolver este problema. O zoneamento perceptivo é realizado após a extração das Cadeias de Freeman, procedimento adotado em todos os experimentos. Para trabalhos futuros, é possível realizar experimentos aplicando o zoneamento antes da extração das Cadeias de Freeman, a fim de obter informações da parte interna da MRV.

5.2.3 Comparativo entre extratores de contorno

Para Roberta Berns [BER02], as informações contidas no contorno de um objeto têm papel importante na percepção da forma. Tais informações de contornos são importantes em diversas áreas cognitivas. Por exemplo, no processo de alfabetização, a criança precisa distinguir formas similares F e E ou M e N, bem como agrupar objetos semelhantes e reconhecer objetos em diferentes perspectivas (um losango e um quadrado). Desde a infância as informações de contornos são importantes no processo de reconhecimento, tais informações devem ser agregadas aos sistemas automatizados de recuperação de imagens com o objetivo de incorporar mecanismos que modelem a percepção humana neste processo.

Wang e Hong [WAN12] argumentam que as informações sobre os contornos da forma são importantes no processo de recuperação de MRVs. Assim, a Tabela 5.5 apresenta resultados comparando os dois extratores de contorno utilizados nos experimentos.

Para a construção da Tabela 5.3 foram selecionadas MRVs que possuem altas taxas de recuperação, usando como limiar valor de $R_n = 0,89$.

Tabela 5.3: Comparativos entre os resultados obtidos com os extratores de contornos

MRV		Conc./Conv.		Best Zoning	FCC		Best Zoning
		R_n	P_n		R_n	P_n	
1190540.bmp		0,921053	0,724056	5h	0,978661	0,769147	5h
1267206.bmp		0,738139	0,498974	5v	0,899220	0,553407	4
1376861.bmp		0,915806	0,596344	5v	0,902968	0,769857	0
2010916_.bmp		0,871914	0,571378	5h	0,924310	0,457094	7
914.bmp		0,863311	0,588797	7	0,968185	0,897077	0
665322.bmp		0,863386	0,773361	7	0,892959	0,683507	0
1055261.bmp		0,911045	0,822548	4	0,851924	0,701442	7
1138293_.bmp		0,991243	0,899630	4/5v	0,795346	0,766077	4
1279931.bmp		0,953005	0,721983	0	0,704563	0,464139	5v
1259886.bmp		0,918092	0,701132	5h	0,514331	0,221858	5v

Observa-se na Tabela 5.3 que as MRVs que possuem contorno bem definido, sem falhas ou interrupções, obtiveram melhores resultados utilizando FCC. No entanto, o extrator de concavidade e convexidade conseguem melhores resultados com as MRVs que apresentam tais problemas no contorno. Por exemplo, a imagem 1190540.bmp

( ) possui $R_n = 0,978661$ para FCC e para Conc./Conv. $R_n = 0,921053$, sendo uma diferença significativa na taxa de recuperação. Já a imagem 1279931.bmp () , que possui "falhas" na parte superior da MRV, obteve resultado muito superior através da extração de concavidade e convexidade obtendo $R_n=0,953005$, sendo que FCC atingiu para esta MRV $R_n = 0,704563$.

Algumas imagens não possuem formas geométricas completas (quadrados, círculos, entre outros), no entanto a mente ignora informações contraditórias, preenchendo as informações incompletas criando formas que sejam familiares [GOM04]. A extração de concavidade e convexidade, utilizada como o Princípio Figura-Fundo, contribui para avaliar informações importantes na MRV nas quais falhas, ruídos ou problemas de interrupções estão presentes nas imagem causando deformações no contorno ou informações incompletas. Este extrator alcança além de informações do contorno, características internas da forma Isso explica os bons resultados obtidos em MRVs que possuem tal deficiência. Em contraponto, o extrator das cadeias de Freeman obtém bons resultados em imagem com contornos menos complexos.

A maior média geral obtida pelo extrato de concavidade e convexidade ($Z=5v$) é $R_n = 0,808839$, sendo esta superior à maior média geral atingida pelo extrator FCC sem zoneamento ($Z=0$) $R_n = 0,752081$. Ainda, selecionando o melhor (*best*) zoneamento para cada MRV consultada, têm-se para Concavidade/Convexidade $R_n = 0,833303$ e para FCC $R_n = 0,805234$, conforme Tabelas 5.1 e 5.2 .

5.2.4 Experimentos com Scale Invariant Feature Transform (SIFT)

O método SIFT foi desenvolvido para extrair e descrever pontos-chave em imagens com o objetivo de reconhecimento de objetos [LOW04]. Tal método é apresentado na Seção 2.3.2.2.

Após a etapa de extração de características, torna-se necessário encontrar as imagens mais similares à imagem consultada. Em muitos extratores, um vetor de características descreve uma imagem e, com base neste vetor, faz-se os cálculos de similaridade. No caso do extrator SIFT, existe um vetor de características para cada ponto-chave encontrado na imagem, sendo que cada MRV consultada apresenta diversos pontos-chave. Assim, é necessário definir uma metodologia para utilizar estes diversos vetores de características que descrevem uma única MRV. Na Seção 4.9 foram

apresentadas duas metodologias para executar essa tarefa, os resultados alcançados são apresentados a seguir.

5.2.4.1 Experimentos com SIFT - Distância Mínima entre Pontos-Chave

Verifica-se a distância mínima entre todos os pontos-chave contidos na imagem da MRV, conforme a Seção 4.9.1. Para este experimento, não são utilizados os zoneamentos perceptivos, considera-se a imagem toda. A Tabela 5.4 apresenta os resultados obtidos com esse método.

Tabela 5.4: SIFT ponto-a-ponto ($Z=0$)

MRV		R_n	P_n
1037814.bmp		0,891590	0,711048
1055261.bmp		0,933973	0,777527
1138103.bmp		0,995948	0,936223
1138293_.bmp		0,886293	0,773267
1190540.bmp		0,993955	0,910710
1259886.bmp		0,759615	0,475968
1267206.bmp		0,599372	0,420676
1279931.bmp		0,925177	0,626926
1289047.bmp		0,814588	0,566206
1376861.bmp		0,418783	0,200150
1439229_.bmp		0,668146	0,534463
1486213_.bmp		0,942933	0,672818
1575268_.bmp		0,981653	0,944844
2010916_.bmp		0,960244	0,778580
2016658_.bmp		0,654860	0,359836
2018809_.bmp		0,870800	0,412429
2042822_.bmp		0,576280	0,202454
665322.bmp		0,451238	0,374748
914.bmp		0,606266	0,284138
967049.bmp		0,964240	0,862979
Média		0,794798	0,591300
Desvio Padrão		0,187940	0,243030

Observa-se na Tabela 5.4 que as MRVs 1138103.bmp (⊗), 1190540.bmp (☐ ☐) e 1575268_.bmp (⊠), obtiveram ótimos resultados comparados ao conjunto verdade proposto pelos especialistas. Sendo respectivamente, $R_n = 0,995948$, $R_n = 0,993955$ e $R_n = 0,981653$. No entanto, as MRVs 1376861.bmp (⊠) e 665322.bmp (⚠) não atingiram bons resultados, obtendo $R_n = 0,418783$ e $R_n = 0,451238$, respectivamente. Tais resultados ruins podem ter ocorrido em razão do extrator SIFT possui invariância a rotação. Senão que a observação do objeto em diferentes perspectivas pode ser uma desvantagem do método aplicado (por exemplo, a comparação de um losângulo e um quadrado).

A Média Geral obtida foi de $R_n = 0,794798$. Pode-se observar a alta variabilidade nos resultados das MRVs consultadas, através do Desvio Padrão.

A Figura 5.5 apresenta as MRVs mais similares à MRV 1575268_.bmp, resultante dos experimentos utilizando o extrator SIFT aplicando a distância mínima entre os pontos-chave ($Z=0$).

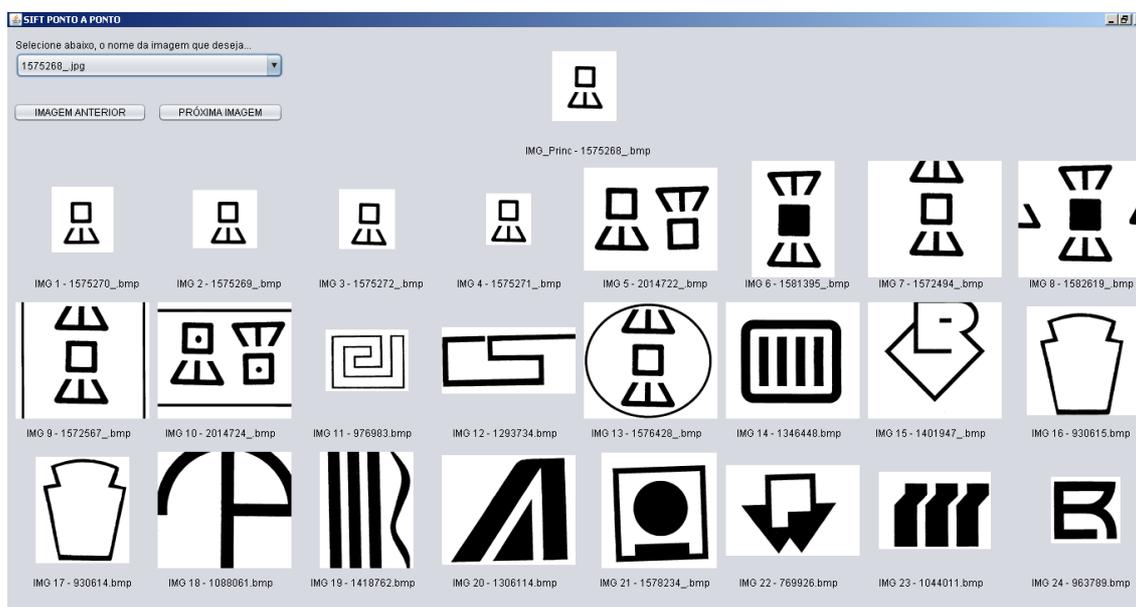


Figura 5.5: SIFT Top-24 MRVs mais similares 1575268_.bmp - Ponto-a-ponto ($Z=0$)

Analisando as MRVs similares apresentadas na Figura 5.5, nota-se que os resultados iniciais são próximos a MRV consultada. Porém, algumas MRVs dadas como similares são bem distintas da MRV consultada, os pontos-chave determinam as características que são extraídas, portanto a localização destes pontos torna-se crucial no processo de recuperação. Observa-se que algumas MRVs foram recuperadas com base

na região inferior da MRV consultada onde existem "ranhuras", e outras são recuperadas pela região superior onde existe um quadrado. Para resolver tal problema é necessário em trabalhos futuros estabelecer um método para rejeitar falsos positivos. É importante ressaltar que os cálculos para R_n consideram as Top-100 MRVs retornadas, na tela apresentada na Figura 5.5 estão apenas as Top-24 iniciais.

5.2.4.2 Experimentos SIFT - Distância Mínima entre Zoneamentos Perceptivos

Neste trabalho, é apresentado o cálculo de similaridade através da distância mínima entre os zoneamentos perceptivos, conforme apresentado na Seção 4.9.2, uma vez que a similaridade é calculada considerando a região perceptiva e não somente o ponto-chave. A Tabela 5.5 apresenta os resultados obtidos com os experimentos calculando a similaridade por região perceptiva.

Observa-se nos resultados dos experimentos, apresentados na Tabela 5.5, que para nove MRVs as taxas de R_n são superiores as obtidas somente considerando os pontos-chave (Tabela 5.4).

Para os experimentos aplicando o zoneamento em quatro regiões ($Z=4$), apresentados na Tabela 5.5, observa-se que as MRVs 1267206.bmp () , 2016658.bmp () e 2042822_.bmp () , atingiram melhores resultados comparados aos demais zoneamentos perceptivos utilizados. A Média Geral da distância mínima entre zoneamentos perceptivos é inferior aos resultados sem zoneamento (ponto-a-ponto), sendo respectivamente $R_n = 0,759230$ e $R_n = 0,794798$. Porém, a variabilidade entre os resultados é reduzida conforme calculado pelo Desvio Padrão = 0,160777, valor menor que o experimento sem zoneamento obtendo Desvio Padrão = 0,18794.

Conforme apresentado na Tabela 5.5 as MRVs 1279931.bmp () , 1575268_.bmp () e 2010916_.bmp () , obtiveram melhores resultados em cinco regiões horizontais ($Z=5h$) comparados aos demais zoneamentos utilizados. Observa-se que as MRVs 1575268_.bmp () e 2010916_.bmp () também obtiveram os melhores resultados para Concavidade/Convexidade com o zoneamento perceptivo $Z=5h$, esses resultados demonstram que as informações contidas na parte inferior da MRV são importantes para o sucesso na recuperação destas MRVs. Um comparativo é apresentado na Seção 5.2.4.3.

Tabela 5.5: SIFT zoneamentos perceptivos

MRV		Z=4		Z=5h		Z=5v		Z=7	
		R _n	P _n						
1037814.bmp		0,909962	0,629839	0,909204	0,611144	0,899676	0,668898	0,916212	0,705711
1055261.bmp		0,877828	1,000000	0,860642	1,000000	0,872455	1,000000	0,893679	1,000000
1138103.bmp		0,957504	0,882161	0,960167	0,880483	0,966075	0,864116	0,918128	0,662318
1138293_.bmp		0,870944	0,818994	0,857950	0,816322	0,832928	0,811755	0,872477	0,819197
1190540.bmp		0,985079	0,872139	0,981047	0,866358	0,981638	0,866384	0,990107	0,936076
1259886.bmp		0,743989	0,457253	0,739375	0,437543	0,812843	0,480499	0,804781	0,479696
1267206.bmp		0,674480	0,337479	0,649571	0,401276	0,673346	0,371220	0,463179	0,210027
1279931.bmp		0,836064	0,569257	0,878197	0,615778	0,823353	0,540232	0,690402	0,335721
1289047.bmp		0,912972	1,000000	0,862003	1,000000	0,930775	1,000000	0,738282	1,000000
1376861.bmp		0,447579	0,204803	0,401035	0,195311	0,456602	0,204654	0,369934	0,185026
1439229_.bmp		0,588335	0,462159	0,563049	0,444008	0,593933	0,518938	0,591424	0,522259
1486213_.bmp		0,623172	1,000000	0,612872	1,000000	0,556577	1,000000	0,640213	1,000000
1575268_.bmp		0,806383	0,545988	0,877158	0,742980	0,802701	0,602623	0,833917	0,739717
2010916_.bmp		0,940297	0,632514	0,955740	0,644590	0,951294	0,632075	0,954368	0,664281
2016658_.bmp		0,618315	0,308690	0,498113	0,242492	0,611458	0,325255	0,553638	0,256524
2018809_.bmp		0,479433	0,216234	0,573713	0,243308	0,610095	0,270800	0,711269	0,346493
2042822_.bmp		0,767 372	1,000000	0,761277	1,000000	0,754782	1,000000	0,754901	1,000000
665322.bmp		0,630152	0,354889	0,628651	0,356117	0,643211	0,362615	0,596343	0,282715
914.bmp		0,657706	1,000000	0,612795	1,000000	0,698974	1,000000	0,721093	1,000000
967049.bmp		0,857029	0,560394	0,955329	0,889750	0,893111	0,591994	0,977619	0,909217
Média		0,759230	0,642640	0,756894	0,669373	0,768291	0,655603	0,749598	0,652749
Desvio Padrão		0,160777	0,28371	0,17613	0,287954	0,153938	0,271724	0,17546	0,302379

O distância mínima entre zoneamento perceptivo em cinco regiões verticais ($Z=5v$), obteve o melhor resultado utilizando o extrator SIFT para as MRVs: 1138103.bmp (⊗), 1259886.bmp (◀), 1289047.bmp (//), 1376861.bmp (⊠), 1439229_.bmp (▽) e 665322.bmp (▲). A MRV 1376861.bmp (⊠), obteve o melhor resultado também para Concavidade/Convexidade aplicando o zoneamento em cinco regiões verticais (Tabela 5.1). Ainda, a MRV 1259886.bmp (◀), obteve para o extrator Cadeias de Freeman o melhor resultado também utilizando $Z=5v$. Das 20 MRVs consultadas seis obtiveram melhor resultado com este zoneamento ($Z=5v$), atingindo Média Geral com $R_n = 0,768291$, sendo a maior taxa em relação aos demais zoneamentos utilizados.

As MRVs 1037814.bmp (●), 1055261.bmp (Ⓢ), 1138293_.bmp (◊), 1190540.bmp (▣ ▣), 1486213.bmp (//), 2018809_.bmp (⊕), 967049.bmp (▲) e 914.bmp (▲) conseguiram o melhor resultado comparando aos demais zoneamentos perceptivos aplicados utilizando zoneamento em sete regiões ($Z=7$). Tal zoneamento evidencia características contidas no meio da MRV, sendo que oito das 20 MRVs obtiveram melhor resultado utilizando $Z=7$, se comparados aos demais zoneamentos perceptivos. No entanto, alguns resultados ruins, por exemplo, para as MRVs 1376861.bmp ($R_n = 0,369934$) e 1267206.bmp ($R_n = 0,463179$) prejudicaram a média final obtendo $R_n = 0,749598$.

O zoneamento em sete regiões também obteve os melhores resultados com o extrator das Cadeias de Freeman para as MRVs 1055261.bmp (Ⓢ) e 967049.bmp (▲), e com o extrator de Concavidade/Convexidade para a MRV 914.bmp (▲). As informações contidas na região central (meio) da MRV auxiliaram na recuperação destas MRVs.

A Figura 5.6 apresenta as 24 MRVs mais similares à MRV 967049, utilizando SIFT associado a distância mínima entre os zoneamentos em sete regiões ($Z=7$). Um comparativo das distâncias mínimas de MRVs sem zoneamento e com os zoneamentos perceptivos é realizado na Tabela 5.6.

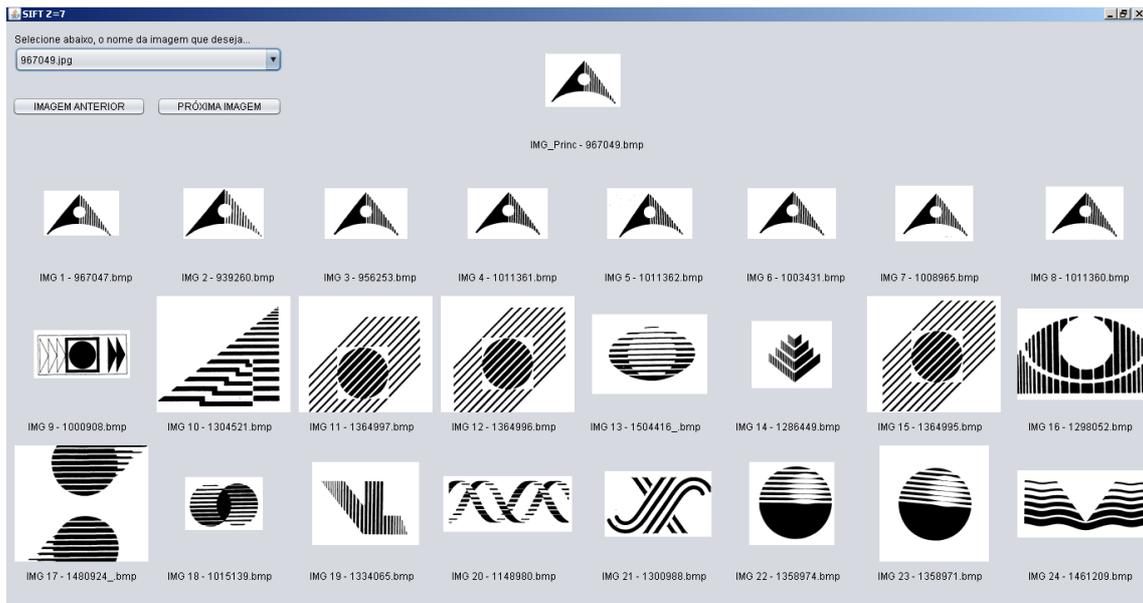


Figura 5.6: SIFT Top-24 MRVs mais similares 967049.bmp - Z=7

Observa-se nos resultados iniciais, na Figura 5.6, que as MRVs apresentadas são iguais a MRV consultada. Algumas MRVs recuperadas mantêm o padrão, onde uma parte da imagem é "sólida" (pixels pretos) e na outra extremidade aparecem as "ranhuras". Outras MRVs concentram seus pontos-chave nas "ranhuras existentes", não consideram informações referentes à forma do objeto, o que é facilmente observado quando se utiliza um extrator de contornos.

5.2.4.3 Comparativo Distância Mínima entre pontos-chave e Distância Mínima entre Zoneamentos Perceptivos

A Tabela 5.6 apresenta um comparativo entre os resultados obtidos nos experimentos utilizando as duas abordagens propostas para calcular a similaridade, utilizando características extraídas pelo SIFT.

Tabela 5.6: Comparativo R_n para todos os zoneamentos

MRV		Z=0	Z=4	Z=5h	Z=5v	Z=7	Best	Z ¹⁷
1037814.bmp		0,891590	0,909962	0,909204	0,899676	0,916212	0,916212	7
1055261.bmp		0,933973	0,877828	0,860642	0,872455	0,893679	0,933973	0
1138103.bmp		0,995948	0,957504	0,960167	0,966075	0,918128	0,995948	0
1138293_.bmp		0,886293	0,870944	0,857950	0,832928	0,872477	0,886293	0
1190540.bmp		0,993955	0,985079	0,981047	0,981638	0,990107	0,993955	0
1259886.bmp		0,759615	0,743989	0,739375	0,812843	0,804781	0,812843	5v
1267206.bmp		0,599372	0,674480	0,649571	0,673346	0,463179	0,67448	4
1279931.bmp		0,925177	0,836064	0,878197	0,823353	0,690402	0,925177	0
1289047.bmp		0,814588	0,912972	0,862003	0,930775	0,738282	0,930775	5v
1376861.bmp		0,418783	0,447579	0,401035	0,456602	0,369934	0,456602	5v
1439229_.bmp		0,668146	0,588335	0,563049	0,593933	0,591424	0,668146	0
1486213_.bmp		0,942933	0,623172	0,612872	0,556577	0,640213	0,942933	0
1575268_.bmp		0,981653	0,806383	0,877158	0,802701	0,833917	0,981653	0
2010916_.bmp		0,960244	0,940297	0,955740	0,951294	0,954368	0,960244	0
2016658_.bmp		0,654860	0,618315	0,498113	0,611458	0,553638	0,65486	0
2018809_.bmp		0,870800	0,479433	0,573713	0,610095	0,711269	0,870800	0
2042822_.bmp		0,576280	0,767372	0,761277	0,754782	0,754901	0,767372	4
665322.bmp		0,451238	0,630152	0,628651	0,643211	0,596343	0,643211	5v
914.bmp		0,606266	0,657706	0,612795	0,698974	0,721093	0,721093	7
967049.bmp		0,964240	0,857029	0,955329	0,893111	0,977619	0,977619	7
Média		0,794798	0,759230	0,756894	0,768291	0,749598	0,835709	
Desvio Padrão		0,18794	0,18794	0,160777	0,17613	0,153938	0,17546	

Observa-se na Tabela 5.6, que 11 (onze) MRVs possuem melhor resultado avaliando a imagem toda - aplicando a distância mínima entre os pontos-chave (Z=0). Porém, deve-se ressaltar que nove MRVs obtiveram melhores resultados fazendo o cálculo da distância entre as regiões perceptivas. A maior média foi obtida pela distância entre os pontos-chave $R_n = 0,794798$. É importante considerar que os resultados obtidos pelas distância entre os zoneamentos perceptivos, selecionando o melhor zoneamento para cada MRV consultada, contribuem para melhora a Média Geral *Best* obtendo $R_n = 0,835709$.

¹⁷ Zoneamento Perceptivo

5.2.5 Experimentos Momentos Invariantes de Hu

Nesta Seção são apresentados os resultados obtidos com os experimentos utilizando Momentos Invariantes de Hu. Conforme mencionado anteriormente, tal extrator não retorna informações que permitam realizar o zoneamento perceptivo, o zoneamento depende da informação do posicionamento dos pixels na imagem. A Tabela 5.7, apresenta os resultados dos experimentos.

Tabela 5.7: Resultados Momentos de Hu

MRV		R _n	P _n
1037814.bmp		0,913139	0,525324
1055261.bmp		0,806506	0,351165
1138103.bmp		0,644352	0,270830
1138293_.bmp		0,774164	0,444298
1190540.bmp		0,500296	0,403793
1259886.bmp		0,703590	0,299476
1267206.bmp		0,587655	0,264234
1279931.bmp		0,629465	0,260122
1289047.bmp		0,303496	0,130680
1376861.bmp		0,739788	0,328769
1439229_.bmp		0,872499	0,521309
1486213_.bmp		0,570504	0,235497
1575268_.bmp		0,807586	0,438651
2010916_.bmp		0,580407	0,329455
2016658_.bmp		0,482813	0,186160
2018809_.bmp		0,440264	0,199723
2042822_.bmp		0,588977	0,218254
665322.bmp		0,592318	0,329596
914.bmp		0,522645	0,212052
967049.bmp		0,818171	0,332914
Média		0,643982	0,314115
Desvio Padrão		0,15823	0,109265

Observa-se na Tabela 5.7 que a imagem 1439229_.bmp () alcança resultados muito superiores aos obtidos pelos demais extratores de características, sendo $R_n = 0,872499$ e $P_n = 0,521309$. O melhor resultado para essa MRV foi $R_n = 0,677129$ e $P_n = 0,532709$ com o extrator de Concavidade/Convexidade, sem zoneamento ($Z=0$). A

MRV 1037814.bmp () obteve $R_n = 0,913139$, sendo este resultado comparável com o resultado apresentado pelo extrator de Concavidade/Convexidade com zoneamento cinco regiões horizontal ($Z=5h$) sendo $R_n = 0,923221$. As 18 (dezoito) MRVs restantes obtiveram resultados inferiores, comparadas as taxas alcançadas pelos demais extratores. É importante ressaltar que nos experimentos utilizando Momentos de Hu executados por Cerri et al. [CER06], os valores obtidos foram $R_n = 0,65$ e para $P_n = 0,32$. Tais resultados, considerando os arredondamentos, são iguais aos obtidos nos experimentos realizados neste trabalho ($R_n = 0,643982$ e $P_n = 0,314115$).

A Figura 5.7 apresenta as MRVs mais similares à MRV 1439229_.bmp. Para o extrator de Momentos de Hu não foram aplicados os zoneamentos perceptivos, os experimentos consideram informações da imagem toda.

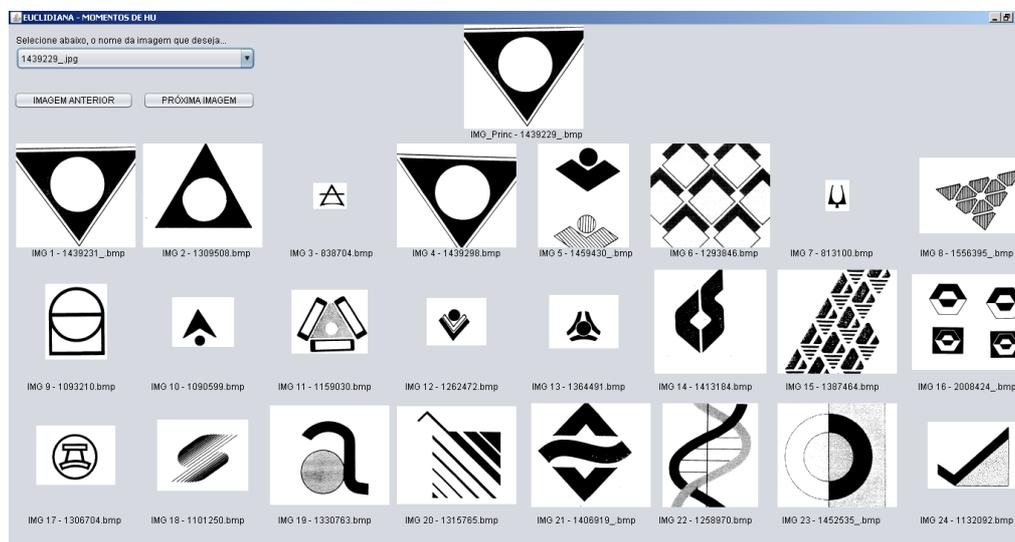


Figura 5.7: Momentos de Hu Top-24 MRVs mais similares 1439229_.bmp - $Z=0$

Observa-se nas MRVs similares retornadas, apresentadas na Figura 5.7, que as primeiras MRVs são muito próximas à MRV consultada. Algumas MRVs são similares ao triângulo invertido contido na MRV consultada, outras, porém, não obtêm essa informação da forma. Observa-se que a invariância a rotação pode ser uma desvantagem para o extrator de características. Momentos de Hu considera a distribuição de pixels nas imagens, isso explica algumas MRVs retornadas nos resultados obtidos.

5.3 Comparativo Geral de Todos os Extratores (*Best of Best*)

Nesta Seção é realizado um comparativo entre os resultados de todos os extratores utilizados nos experimentos. São apresentados os melhores resultados,

destacando o melhor valor de R_n para cada MRV consultada, bem como o melhor zoneamento perceptivo.

A Tabela 5.8¹⁸, apresenta um comparativo entre os melhores resultados obtidos para cada extrator de características, sem zoneamento perceptivo ($Z=0$). Este comparativo apresenta os valores obtidos para R_n considerando cada uma das 20 MRVs consultadas.

Tabela 5.8: Comparativo Best extratores - R_n ($Z=0$)

MRV		Conc./Conv.	FCC	SIFT	M. Hu	Best	Extrator
1037814.bmp		0,91	0,79	0,89	0,91	0,91	Conc./Conv.
1055261.bmp		0,90	0,79	0,93	0,81	0,93	SIFT
1138103.bmp		0,74	0,79	1,00	0,64	1,00	SIFT
1138293_.bmp		0,96	0,80	0,89	0,77	0,96	Conc./Conv.
1190540.bmp		0,85	0,89	0,99	0,50	0,99	SIFT
1259886.bmp		0,85	0,39	0,76	0,70	0,85	Conc./Conv.
1267206.bmp		0,68	0,89	0,60	0,59	0,89	FCC
1279931.bmp		0,95	0,58	0,93	0,63	0,95	Conc./Conv.
1289047.bmp		0,75	0,73	0,81	0,30	0,81	SIFT
1376861.bmp		0,79	0,90	0,42	0,74	0,90	FCC
1439229_.bmp		0,68	0,47	0,67	0,87	0,87	M. Hu
1486213_.bmp		0,68	0,77	0,94	0,57	0,94	SIFT
1575268_.bmp		0,65	0,86	0,98	0,81	0,98	SIFT
2010916_.bmp		0,82	0,79	0,96	0,58	0,96	SIFT
2016658_.bmp		0,66	0,85	0,65	0,48	0,85	FCC
2018809_.bmp		0,48	0,62	0,87	0,44	0,87	SIFT
2042822_.bmp		0,76	0,45	0,58	0,59	0,76	Conc./Conv.
665322.bmp		0,83	0,89	0,45	0,59	0,89	FCC
914.bmp		0,75	0,97	0,61	0,52	0,97	FCC
967049.bmp		0,89	0,85	0,96	0,82	0,96	SIFT
Média		0,78	0,75	0,79	0,64	0,91	
Desvio Padrão		0,12	0,17	0,19	0,16	0,06	

Com base nos resultados da Tabela 5.8, o melhor extrator sem aplicar zoneamento perceptivo foi SIFT (*Scale Invariant Feature Transform*) atingindo Média Geral $R_n = 0,79$. Sendo importante observar que SIFT obteve a maior variação nas taxas de R_n entre as MRVs consultadas, através do Desvio Padrão = 0,19. No entanto,

¹⁸ A Tabela 5.10 apresenta os valores arredondados para duas casas decimais. Tal arredondamento foi realizado para facilitar as comparações futuras com outros trabalhos na literatura.

selecionando o melhor extrator para cada MRV, a Média Geral obtida foi $R_n = 0,91$, sendo superior aos valores encontrados para R_n na literatura [EAK98] [CER06].

Na Tabela 5.9 são apresentados os resultados obtidos para R_n evidenciando o melhor resultado de cada extrator (*Best of Best*) e consideram-se os resultados dos extratores associados aos zoneamentos perceptivos. Assim, é possível determinar qual o melhor extrator e também o melhor zoneamento perceptivo para cada MRV consultada.

É importante observar que as taxas de R_n melhoram em relação aos experimentos realizados sem zoneamento ($Z=0$), apresentados na Tabela 5.8.

Tabela 5.9: *Best* de cada extrator R_n

MRV		Conc/Conv	FCC	SIFT	Hu	Best	Extrator	Zoneamento
1037814.bmp		0,923	0,79	0,916	0,91	0,92	Conc./Conv.	5h
1055261.bmp		0,91	0,85	0,93	0,81	0,93	SIFT	0
1138103.bmp		0,85	0,81	1,00	0,64	1,00	SIFT	0
1138293_.bmp		0,99	0,80	0,89	0,77	0,99	Conc./Conv.	4/5v
1190540.bmp		0,92	0,98	0,99	0,50	0,99	SIFT	0 / 7
1259886.bmp		0,92	0,51	0,81	0,70	0,92	Conc./Conv.	5h
1267206.bmp		0,74	0,90	0,67	0,59	0,90	FCC	4
1279931.bmp		0,95	0,70	0,93	0,63	0,95	Conc./Conv.	0
1289047.bmp		0,84	0,85	0,93	0,30	0,93	SIFT	5v
1376861.bmp		0,92	0,90	0,46	0,74	0,92	Conc./Conv.	5v
1439229_.bmp		0,68	0,59	0,67	0,87	0,87	M. Hu	0
1486213_.bmp		0,77	0,79	0,94	0,57	0,94	SIFT	0
1575268_.bmp		0,71	0,86	0,98	0,81	0,98	SIFT	0
2010916_.bmp		0,87	0,92	0,96	0,58	0,96	SIFT	0 / 5h
2016658_.bmp		0,76	0,85	0,65	0,48	0,85	FCC	0
2018809_.bmp		0,55	0,62	0,87	0,44	0,87	SIFT	0
2042822_.bmp		0,76	0,68	0,77	0,59	0,77	SIFT	4
665322.bmp		0,86	0,89	0,64	0,59	0,89	FCC	0
914.bmp		0,86	0,97	0,72	0,52	0,97	FCC	0
967049.bmp		0,89	0,86	0,98	0,82	0,98	SIFT	7
Média		0,83	0,81	0,84	0,64	0,93		
Desvio Padrão		0,11	0,13	0,15	0,16	0,06		

Comparando os resultados individuais das MRVS, na Tabela 5.9, é possível determinar o melhor extrator (*Best extractor*) e o melhor zoneamento perceptivo (*Best Zoning*) para cada MRV. Observa-se que os zoneamentos melhoram as taxas obtidas

para R_n , dez das 20 MRVs atingiram valores superiores utilizando os zoneamentos. Sendo que duas, entre as dez MRVs obtiveram valores iguais utilizando ou não os zoneamentos perceptivos. Verifica-se que, selecionando o melhor extrator e o melhor zoneamento perceptivo houve reflexo no aumento da Média Geral.

O melhor extrator associado aos zoneamentos perceptivos foi SIFT com Média Geral $R_n = 0,84$. Porém, selecionando o melhor extrator e considerando os zoneamentos perceptivos para cada MRV consultada, a Média Geral Final (*Best of Best*) aumenta consideravelmente obtendo $R_n = 0,93$, sendo esta média geral superior às taxas encontradas na literatura para a mesma base de dados, com o mesmo conjunto verdade das 20 MRVs de consulta.

Deve-se considerar que, para oito MRVs houve um aumento na taxa de R_n quando aplicados os zoneamentos, para a imagem 1289047.bmp () a taxa obtida pelo extrator SIFT utilizando a distância mínima entre os pontos-chave foi de $R_n = 0,81$. No entanto, após aplicar SIFT com a distância mínima entre os zoneamentos perceptivos a taxa atingida foi $R_n = 0,93$, representando uma diferença de 0,12. A Tabela 5.10, apresenta o aumento nos resultados das MRVs com o uso do zoneamento perceptivo.

Tabela 5.10: Diferenças R_n para $Z=0$ e $Z=4, 5h, 5v$ e 7

MRV		Extrator sem zoneamento	Best	Extrator com zoneamentos		Best	Ganho do Sistema R_n
		(Z=0)	R_n	Extrator	Zoneamentos	R_n	
1037814.bmp		Conc./Conv.	0,91	Conc./Conv.	5h	0,92	+0,01
1138293_.bmp		Conc./Conv.	0,96	Conc./Conv.	4/5v	0,99	+0,03
1259886.bmp		Conc./Conv.	0,85	Conc./Conv.	5h	0,92	+0,07
1267206.bmp		FCC	0,89	FCC	4	0,90	+0,01
1289047.bmp		SIFT	0,81	SIFT	5v	0,93	+0,12
1376861.bmp		FCC	0,90	Conc./Conv.	5v	0,92	+0,02
2042822_.bmp		Conc./Conv.	0,76	SIFT	4	0,77	+0,01
967049.bmp		SIFT	0,96	SIFT	7	0,98	+0,02
Média <i>Best</i>			0,91	0,91	0,93	0,93	+0,02
Des. Pad. <i>Best</i>			0,06	0,06	0,06	0,06	

Na Tabela 5.10 observa-se que seis MRVs mantiveram o mesmo extrator de características, e suas taxas R_n aumentaram após o uso do zoneamento perceptivo. Entretanto, duas MRVs obtiveram alteração no melhor extrator quando considerado o zoneamento perceptivo, sendo elas 1376861.bmp () e 2042822_.bmp ()

A Tabela 5.9, é parâmetro para a construção dos conjuntos de treinamento e validação das RNAs e da SVM, sendo que estes conjuntos são formados de acordo com o conjunto verdade associado a cada MRV, como explicado na Seção 4.7.

É importante destacar que as RNAs e a SVM foram treinadas com as características sem zoneamento ($Z=0$). O zoneamento perceptivo foi considerado para definir os melhores resultados obtidos pelas MRVs (Tabela 5.9), estes resultados são parâmetro para definir a qual classe de extrator cada MRVs pertence, de modo a compor os conjuntos de treinamento, validação e teste.

Os valores obtidos para *Normalized Precision* (P_n) são apresentados na Tabela 5.11. Cabe destacar que os valores foram considerados conforme os resultados apresentados na Tabela 5.9, em que são definidos o melhor extrator e o melhor zoneamento com base na medida de recuperação *Normalized Recall* (R_n). Neste contexto, os valores obtidos por R_n definem o melhor extrator e o melhor zoneamento a ser aplicado pelo modelo Hierárquico Perceptivo.

Nas Seções 5.4 e 5.5 são apresentados os resultados dos experimentos realizados utilizando Redes Neurais Artificiais (RNAs) e Máquinas de Vetores Suporte (SVM). Tais experimentos buscam definir dinamicamente o melhor extrator a ser aplicado nas MRVs de consulta.

5.4 Resultados RNA

Foram construídas quatro RNA, cada uma especialista em um dos quatro extratores de características utilizados no modelo. O objetivo é que as RNAs apresentem um "voto" ou *score* para cada MRV consultada. Por intermédio da regra do Voto Majoritário determina-se a classe de extrator que a MRV pertence, ou seja, o melhor extrator a ser utilizado para a MRV em questão. Determinar o melhor extrator a ser utilizado é importante para maximizar os resultados obtidos.

As entradas das RNA utilizam o vetor de características sem zoneamento ($Z=0$). Consequentemente, a MRV enviada para consulta também possui informações sem o zoneamento perceptivo. A definição do zoneamento ocorre somente no segundo nível, após determinar o melhor extrator a ser utilizado.

Tabela 5.11: Best de cada extrator P_n

MRV		Extrator	Zoneamento	Best P_n
1037814.bmp		Conc./Conv.	5h	0,79
1055261.bmp		SIFT	0	0,78
1138103.bmp		SIFT	0	0,94
1138293_.bmp		Conc./Conv.	4	0,90
1190540.bmp		SIFT	0	0,91
1259886.bmp		Conc./Conv.	5h	0,70
1267206.bmp		FCC	4	0,55
1279931.bmp		Conc./Conv.	0	0,72
1289047.bmp		SIFT	5v	1,00
1376861.bmp		Conc./Conv.	5v	0,60
1439229_.bmp		M. Hu	0	0,52
1486213_.bmp		SIFT	0	0,67
1575268_.bmp		SIFT	0	0,94
2010916_.bmp		SIFT	0	0,78
2016658_.bmp		FCC	0	0,39
2018809_.bmp		SIFT	0	0,41
2042822_.bmp		SIFT	4	1,00
665322.bmp		FCC	0	0,68
914.bmp		FCC	0	0,90
967049.bmp		SIFT	7	0,91
Média				0,75
Desvio Padrão				0,19

Foram treinadas diversas configurações de RNAs para cada classe de extrator. No entanto, as RNAs que obtiveram melhores resultados foram: SIFT com 800 épocas, Concavidade e Convexidade com duas mil épocas, Cadeias de Freeman com 14 mil épocas e Momentos Invariantes de Hu com 1,2 mil épocas. O número de neurônios da camada escondida é a metade do número de neurônios contidos na camada de entrada. A saída de cada rede é única.

A Tabela 5.12, apresenta os votos de cada RNA, em relação as 20 (vinte) MRVs consultadas.

Tabela 5.12: Voto 4 Redes Classe Modular

MRV		VOTOS - RNA Classe Modular				1o	2o.	<i>Best</i> ¹⁹
		SIFT (800)	Conc (2000)	FCC (14000)	M. Hu (1200)			
1037814.bmp		0,95236	0,97905	0,00008	0,00867	Conc/Conv	SIFT	Conc/Conv
1055261.bmp		0,99999	0,27133	0,28678	0,63788	SIFT	M. Hu	SIFT
1138103.bmp		1,00000	0,87204	0,80897	0,65175	SIFT	Conc/Conv	SIFT
1138293_.bmp		0,00922	0,99783	0,00023	0,88685	Conc/Conv	M. Hu	Conc/Conv
1190540.bmp		0,99999	0,01674	0,95978	0,86756	SIFT	FCC	SIFT
1259886.bmp		0,99123	0,99895	0,01870	0,88385	Conc/Conv	SIFT	Conc/Conv
1267206.bmp		0,99732	0,83992	0,96964	0,26522	SIFT	FCC	FCC
1279931.bmp		0,30128	0,97460	0,99949	0,64149	FCC	Conc/Conv	Conc/Conv
1289047.bmp		0,99998	0,00314	0,00376	0,88932	SIFT	M. Hu	SIFT
1376861.bmp		0,76713	0,99727	0,00002	0,90666	Conc/Conv	M. Hu	Conc/Conv
1439229_.bmp		0,98466	0,99996	0,08600	0,93371	Conc/Conv	SIFT	M. Hu
1486213_.bmp		0,99993	0,00013	0,00020	0,83963	SIFT	M. Hu	SIFT
1575268_.bmp		0,99999	0,52399	0,98127	0,95444	SIFT	FCC	SIFT
2010916_.bmp		1,00000	0,33248	0,73473	0,91601	SIFT	M. Hu	SIFT
2016658_.bmp		0,90784	0,21290	0,97164	0,94157	FCC	M. Hu	FCC
2018809_.bmp		1,00000	0,85700	0,78291	0,00666	SIFT	Conc/Conv	SIFT
2042822_.bmp		0,99736	1,00000	0,99787	0,96080	Conc/Conv	FCC	SIFT
665322.bmp		0,99975	0,04270	0,91989	0,69355	SIFT	FCC	FCC
914.bmp		0,98913	0,05405	0,97614	0,36632	SIFT	FCC	FCC
967049.bmp		0,99813	0,10478	0,43212	0,91617	SIFT	M. Hu	SIFT

Observa-se na Tabela 5.12, que as imagens 1439229_.bmp () e 2042822_.bmp () não obtiveram os resultados esperados nas RNAs, porque não foram classificadas corretamente. Quatro MRVs obtiveram o resultado esperado com o 2o. melhor voto. As 14 MRVs restantes conseguiram ser classificadas corretamente nas suas classes de extratores.

Deve-se considerar que a MRV 2042822_.bmp () não foi classificada corretamente pois foram utilizadas as características sem zoneamento ($Z=0$) para treinar as RNA. Conforme discutido na Tabela 5.10, esta MRV altera o melhor extrator quando aplicado o zoneamento perceptivo. Este fato prejudicou o resultado obtido pela RNA para esta MRV.

¹⁹ O valor para *Best* foi retirado da Tabela 5.11, onde são apresentados os resultados para definir o melhor extrator a ser utilizado.

A MRV 1439229_.bmp (), que deveria ser classificada pertencendo à classe de extrator Momentos de Hu, não obteve sucesso. Tal fato deve-se à quantidade de exemplos desta MRV contidos no conjunto verdade (apenas nove MRVs). É importante observar que somente essa MRV obteve melhor resultado com Momentos de Hu, resultando em um conjunto pequeno de MRVs. Infelizmente, o número de MRVs para o conjunto de treinamento, validação e testes foram insuficientes para resolver tal tarefa.

Os resultados obtidos pelas RNAs não foram suficientes para resolver completamente a tarefa para determinar o melhor extrator a ser utilizado pela MRV de consulta. No entanto, das 20 MRVs de consulta 14 MRVs conseguiram ser classificadas corretamente. Das 6 MRVs que não obtiveram êxito, quatro MRVs alcançaram sua classe de extrator considerando o segundo voto e duas não atingiram o objetivo esperado. A Tabela 5.13 apresenta a matriz de confusão resultante deste experimento.

Tabela 5.13: Matriz de confusão Teste - RNAs

	CONC	FCC	Hu	SIFT
CONC	5	1	0	0
FCC	0	1	0	3
Hu	1	0	0	0
SIFT	1	0	0	9
				14 acertos
				6 Erros
				Acurácia =70 %

Em face dos resultados obtidos pelas RNAs, foram realizados experimentos utilizando Máquinas de Vetores Suporte (SVM). Os resultados obtidos pela SVM são apresentados e discutidos a seguir.

5.5 Resultados SVM

As Máquinas de Vetores Suporte (SVM) foram desenvolvidas com o objetivo de realizar tarefas de classificação, sendo utilizadas com sucesso em aplicações de reconhecimento de padrões tais como: identificação de autoria [AMA14], recuperação de MRVs [QI_9] [HAI10] entre outras aplicações [RUF11] [LIU14].

Com o objetivo de avaliar o desempenho da SVM na classificação de extratores, comparando com os resultados obtidos pelas RNAs apresentados na Seção 5.4, foi construída uma SVM Multiclasses. Os vetores de características de cada extrator,

utilizado no Modelo Hierárquico Perceptivo, foram unidos em um único vetor de características. Este vetor possui dimensão igual a 167, ou seja, composto de 128 características obtidas pela média dos *keypoints* SIFT, 24 características de Concavidade e Convexidade, oito características de Cadeias de Freeman e sete características de Momentos de Hu.

Foram construídos dois conjuntos: treinamento e teste. Para a classificação foram utilizadas quatro classes de extratores. Cada classe representa um extrator de características, sendo: Concavidade e Convexidade corresponde à Classe 1, Cadeias de Freeman à Classe 2, Momentos de Hu à Classe 3 e SIFT pertence à Classe 4.

A Tabela 5.14 apresenta os resultados obtidos para as MRVs contidas no conjunto de teste.

Tabela 5.14: SVM Multiclasses

MRV		Resultados <i>predict</i>	Resultados Tabela 1	Resultados <i>predict</i>	Resultados Tabela 5.11	R_n <i>Best</i>
1037814.bmp		1	1	CONC	CONC	0,92
1055261.bmp		4	4	SIFT	SIFT	0,93
1138103.bmp		4	4	SIFT	SIFT	1,00
1138293_.bmp		1	1	CONC	CONC	0,99
1190540.bmp		4	4	SIFT	SIFT	0,99
1259886.bmp		1	1	CONC	CONC	0,92
1267206.bmp		2	2	FCC	FCC	0,90
1279931.bmp		1	1	CONC	CONC	0,95
1289047.bmp		4	4	SIFT	SIFT	0,93
1376861.bmp		1	1	CONC	CONC	0,92
1439229_.bmp		3	3	HU	HU	0,87
1486213_.bmp		4	4	SIFT	SIFT	0,94
1575268_.bmp		4	4	SIFT	SIFT	0,98
2010916_.bmp		4	4	SIFT	SIFT	0,96
2016658_.bmp		2	2	FCC	FCC	0,85
2018809_.bmp		4	4	SIFT	SIFT	0,87
2042822_.bmp		1	4	CONC	SIFT	0,76
665322.bmp		2	2	FCC	FCC	0,89
914.bmp		2	2	FCC	FCC	0,97
967049.bmp		4	4	SIFT	SIFT	0,98
Média						0,93
Desvio Padrão						0,06

Observa-se nos resultados apresentados na Tabela 5.14, que a SVM conseguiu classificar melhor as MRVs em relação às RNAs. Das 20 MRVs contidas no conjunto de teste, 19 MRVs foram classificadas corretamente. A MRV 2042822_.bmp (\mathcal{V}) não obteve o resultado esperado, sua classe deveria ser SIFT, porém foi classificada como Concavidade e Convexidade. Tal fato ocorreu também para as RNAs, conforme explicado anteriormente, a SVM também foi treinada com as características sem aplicar os zoneamentos perceptivos ($Z=0$), logo conforme Tabela 5.10 o melhor extrator para essa MRV muda quando aplicado o zoneamento perceptivo.

A matriz de confusão obtida nos experimentos utilizando SVM, referentes às 20 (vinte) MRVs contidas no conjunto de teste, é apresentada na Tabela 5.15.

Tabela 5.15: Matriz de confusão Teste - predict

	CONC	FCC	Hu	SIFT
CONC	5	0	0	0
FCC	0	4	0	0
Hu	0	0	1	0
SIFT	1	0	0	8
				19 acertos
				1 erro
				Acurácia = 95%

Das 20 MRVs contidas no conjunto de teste, apenas uma MRV apresentou confusão (Tabela 5.15). No entanto, tal resultado não afeta a Média Geral *Best* para R_n , uma vez que a diferença entre os resultados obtidos pelos extratores é 0,01, sendo que Concavidade e Convexidade obteve $R_n = 0,76$ e SIFT obteve $R_n = 0,77$.

Tais resultados confirmam que as SVMs apresentam sucesso em sistemas de reconhecimento de padrões. De acordo com Rufino [RUF11] apud Bennet e Campbell [BEN00], as SVMs são eficientes em relação à velocidade e complexidade. Igualando-se a busca de mínimos de uma função convexa, isto é, sem mínimos locais. Assim, muitos problemas que ocorrem em RNAs e árvores de decisão são eliminados. Tal observação pode explicar os bons resultados obtidos pela SVM em relação às RNAs.

Nos experimentos realizados foi criada somente uma SVM para a tarefa de classificação. Para os experimentos com RNAs foram construídas 4 RNAs, uma para cada extrator para realizar tal tarefa. Assim, utilizando SVM a complexidade do sistema foi reduzida. Foram executados experimentos com uma única RNA com o objetivo de

classificar as quatro classes de extratores, porém não obteve melhores resultados que a RNA Múltiplos classificadores.

5.6 Comparativo dos Resultados Obtidos com outros Trabalhos Apresentados na Literatura

Para validar os resultados obtidos, a Tabela 5.16 apresenta uma comparação entre os melhores (*Best*) resultados do Modelo Hierárquico Perceptivo com o Sistema ARTISAN proposto por Eakins et. al. [EAK98]. Tal comparação é feita utilizando as 20 (vinte) MRVs contidas no "conjunto verdade", selecionadas pelos especialistas do escritório de patentes do Reino Unido. Os valores apresentados na Tabela 5.16 referem-se à medida *Normalized Recall* (R_n).

Tabela 5.16: Comparativo resultados por MRV

MRV		R_n <i>Best</i>	ARTISAN R_n
1037814.bmp		0,92	0,96
1055261.bmp		0,93	0,98
1138103.bmp		1,00	0,89
1138293_.bmp		0,99	1,00
1190540.bmp		0,99	1,00
1259886.bmp		0,92	0,84
1267206.bmp		0,90	0,92
1279931.bmp		0,95	0,90
1289047.bmp		0,93	1,00
1376861.bmp		0,92	0,90
1439229_.bmp		0,87	0,93
1486213_.bmp		0,94	0,84
1575268_.bmp		0,98	0,99
2010916_.bmp		0,96	0,90
2016658_.bmp		0,85	0,61
2018809_.bmp		0,87	0,90
2042822_.bmp		0,76	0,48
665322.bmp		0,89	0,98
914.bmp		0,97	0,94
967049.bmp		0,98	0,90
Média		0,93	0,89
Desvio Padrão		0,06	0,13

Observa-se na Tabela 5.16, que das 20 MRVs consultadas 10 (dez) obtiveram melhores resultados através do Modelo proposto. É importante ressaltar que duas MRVs obtiveram resultados muito superiores aos obtidos pelo Artisan. A MRV 2016658_.bmp () obteve $R_n = 0,85$ comparada ao Artisan quem obteve $R_n = 0,61$. É relevante mencionar que o Modelo Hierárquico não obteve valores abaixo de $R_n = 0,76$, a maioria dos resultados (16 MRVs) obteve valor para R_n maior que 0,90.

A Tabela 5.17 apresenta os resultados obtidos para *Normalized Precision* (P_n) conforme Tabela 5.11, tais valores foram obtidos considerando o melhor extrator e o melhor zoneamento perceptivo definido com base nos resultados obtidos para R_n conforme Tabelas 5.9, 5.14 e 5.16.

Tabela 5.17: Comparativo resultados por MRV

MRV		<i>Best P_n</i>	ARTISAN <i>P_n</i>
1037814.bmp		0,79	0,85
1055261.bmp		0,78	0,91
1138103.bmp		0,94	0,66
1138293_.bmp		0,90	0,88
1190540.bmp		0,91	0,97
1259886.bmp		0,70	0,59
1267206.bmp		0,55	0,67
1279931.bmp		0,72	0,64
1289047.bmp		1,00	0,94
1376861.bmp		0,60	0,49
1439229_.bmp		0,52	0,89
1486213_.bmp		0,67	0,47
1575268_.bmp		0,94	0,84
2010916_.bmp		0,78	0,61
2016658_.bmp		0,39	0,15
2018809_.bmp		0,41	0,44
2042822_.bmp		1,00	0,16
665322.bmp		0,68	0,84
914.bmp		0,90	0,69
967049.bmp		0,91	0,82
Média		0,75	0,68
Desvio Padrão		0,19	0,24

Na Tabela 5.17, das 20 MRVs consultadas 13 (treze) obtiveram melhores resultados através do Modelo proposto. É importante ressaltar que as MRVs 1289047.bmp () e 2042822_.bmp () obtiveram excelente resultado sendo $P_n = 1$. Ainda, as MRVs 2016658_.bmp () e 2042822_.bmp () atingiram resultados muito baixos através do Sistema Artisan, sendo respectivamente $P_n = 0,15$ e $P_n = 0,16$. Utilizando o Modelo proposto estas duas MRVs conseguiram melhores resultados, sendo respectivamente $P_n = 0,39$ e $P_n = 1$.

Os resultados individuais para R_n e P_n das 20 (vinte) MRVs de consulta, somente foram apresentados por Eakins et. al. [EAK98]. Cerri et. al. [CER06] apresentaram somente a Média Geral.

Cabe destacar que a Média Geral obtida pelo Modelo Hierárquico Perceptivo foi de $R_n = 0,93$, comparada a obtida pelo Artisan que foi de $R_n = 0,89$. Ainda, sem utilizar os zoneamentos perceptivos, selecionando somente o melhor extrator no primeiro nível, obtém-se $R_n = 0,91$. Um comparativo dos resultados obtidos na literatura para a base de MRVs *UK Patent Office* é apresentado na Tabela 5.18.

Tabela 5.18: Comparativo resultados Média Geral

Autores	R_n	P_n
Modelo Hierárquico Perceptivo	0,93	0,75
Eakins et al. [EAK98]	0,89	0,67
Cerri et al. [CER06]	0,81	0,56

O Modelo Hierárquico Perceptivo apresentado atingiu valores superiores aos contidos na literatura, tanto para R_n quanto para P_n . Deve-se considerar que poucos trabalhos publicados utilizam a mesma base de MRVs, dificultando a comparação dos resultados obtidos. Ainda, o Sistema Artisan é parâmetro de comparação para diversos autores quando se trata de sistemas de recuperação de MRVs. Durante os experimentos realizados neste trabalho foi solicitada autorização para uso da base de MRVs do MPEG-7, a resposta foi negativa informando que apenas membros poderiam utilizar a base de MRVs para experimentos. Existem bases de imagens disponibilizadas pelo MPEG-7, porém não de MRVs.

Os experimentos adicionais, utilizando cálculo de similaridade pela Distância de Histograma e Distância de Mahalanobis, são apresentados respectivamente no Apêndice

1 e Apêndice 2. Experimentos com a combinação de diferentes extratores de características, são apresentados no Apêndice 3.

5.7. Considerações Finais

Este capítulo apresentou uma análise detalhada dos resultados obtidos com os experimentos de validação do método proposto. Além disso, uma comparação dos resultados obtidos com outros trabalhos apresentados na literatura também foi realizada.

No próximo Capítulo são apresentadas as principais conclusões deste trabalho bem como sugestões de trabalhos futuros.

Capítulo 6

Conclusão

As MRVs são consideradas um dos ativos mais importantes de uma empresa. Elas representam produtos ou serviços e têm o objetivo de fixar a imagem da empresa na mente dos consumidores, associando a MRV à qualidade do produto e até mesmo produzindo sensações boas ou más dependendo da experiência dos consumidores.

Os escritórios de registro de patentes recebem anualmente milhares de solicitações de registro de novas MRVs. Conforme apresentado no Capítulo 2, no Brasil 85.810²⁰ novos registros de MRV foram concedidos no ano de 2014. Em face desta realidade, é necessária uma solução computacional para auxiliar na tarefa de recuperação de MRVs nas bases de dados dos escritórios de concessão de registros, visando identificar similaridades de modo a autorizar ou não o uso de uma nova MRV. Nesse contexto, esta pesquisa apresentou um Modelo Hierárquico Perceptivo para recuperação de MRVs incorporando aspectos da percepção humana das formas.

O modelo apresentado neste trabalho envolve a definição dinâmica de um conjunto de extratores de características, associado a mecanismos de zoneamentos perceptivos aplicados a um sistema de recuperação de MRVs.

Assim, foi estabelecido um protocolo experimental de modo que cenários de testes, para prova de conceito, pudessem ser realizados para validar o modelo proposto. Finalmente, realizou-se a análise dos resultados obtidos e com base nos resultados são apresentadas as principais conclusões.

O Modelo Hierárquico Perceptivo proposto cumpre com seus objetivos, uma vez que as taxas obtidas nos experimentos ($R_n = 0,93$ e $P_n = 0,75$) são superiores às

²⁰ <http://www.inpi.gov.br/portal/artigo/estatisticas>

encontradas na literatura pesquisada, conforme apresentado na Tabela 5.18 do Capítulo 5. O modelo desenvolvido é composto de dois níveis: no primeiro nível define-se dinamicamente o extrator de características e no segundo nível são aplicados os mecanismos de zoneamento perceptivo não-simétrico.

A seleção dinâmica do extrator de características (1º Nível) foi necessária, uma vez que os experimentos realizados, unindo características de diferentes extratores obtiveram resultados inferiores se comparados aos extratores individuais (Apêndice 3). Para a seleção do melhor extrator, dois classificadores foram testados: RNAs e SVM. Com base nas matrizes de confusão obtidas pelos classificadores, a SVM apresentou uma taxa de acerto superior se comparada com as obtidas pelas RNAs (Tabelas 5.14 e 5.15). Desta forma, definiu-se que a SVM compõe o Modelo Hierárquico Perceptivo para fazer a classificação do melhor extrator de características a ser utilizado.

No 2º. Nível, são utilizados os mecanismos de zoneamento perceptivo não-simétrico, os quais são aplicados ao melhor extrator de características definido no 1º. Nível. Os zoneamentos perceptivos permitiram uma análise de informações locais, evidenciando elementos-chave para a recuperação das MRVs. Conforme discutido na Tabela 5.10, os zoneamentos contribuíram para o aumento na taxa de recuperação. A média geral, considerando o melhor extrator de características, sem utilizar o zoneamento ($Z = 0$) é de $R_n = 0,91$ (Tabela 5.8). Aplicando os zoneamentos perceptivos, e definido o melhor zoneamento, a taxa atinge o valor de $R_n = 0,93$ (Tabela 5.9). Isso demonstra que o uso dos zoneamentos agrega informações complementares sobre o padrão analisado, e estas informações devem ser incorporadas ao modelo proposto.

Para calcular a similaridade entre as MRVs, foi utilizada a Distância Euclidiana entre os vetores de características. No caso do extrator SIFT, duas técnicas foram propostas: distância mínima entre pontos-chave e distância mínima entre os zoneamentos perceptivos. A distância entre pontos-chave obteve melhores resultados, sendo a média geral $R_n = 0,79$ e $P_n = 0,59$. A distância entre zoneamentos perceptivos obteve média geral $R_n = 0,77$ e $P_n = 0,66$ ($Z=5V$). Observa-se que 09 (nove) das 20 (vinte) MRVs obtiveram melhores resultados utilizando a similaridade calculada por meio dos zoneamentos perceptivos. Mesmo não obtendo resultados superiores, a distância entre os zoneamentos perceptivos pode ser uma técnica complementar ao extrator SIFT, uma vez que a similaridade calculada entre regiões que compõem a imagem obteve bons resultados para aproximadamente 50% das MRVs consultadas (Tabela 5.6). Esses resultados demonstram que a percepção de similaridade, por

intermédio da comparação dos zoneamentos perceptivos, pode auxiliar na recuperação das MRVs. Consta-se também que, utilizando o extrator SIFT e selecionando o melhor resultado para a distância mínima entre os pontos-chave e os zoneamentos perceptivos (Tabela 5.6), a taxa de recuperação melhora e a média geral obtém valores de $R_n = 0,84$ e $P_n = 0,66$.

Deve-se considerar ainda que a base de dados possui imagens de baixa qualidade, dificultando a extração adequada de características, este problema também foi detectado por Eakins et al. [EAK98] e Cerri et al. [CER06]. É importante destacar que a base não é dividida em classes de MRVs, e somente 20 MRVs são disponibilizadas para testes, trazendo as soluções em um nível de complexidade para ser resolvido com o uso de métodos de classificação.

Adicionalmente ao Modelo Hierárquico Perceptivo, o princípio da Gestalt Figura-Fundo é explorado através das Deficiências de Concavidade e Convexidade. Esta abordagem complementa os estudos apresentados por [ABE06] [HIR09] [BEF12] com base nas teorias da Gestalt aplicadas às MRVs. Avaliando os resultados obtidos nos experimentos de concavidade e convexidade (Figura-Fundo), apresentados no Capítulo 5, considera-se que houve êxito nesta abordagem uma vez que os valores obtidos são satisfatórios se comparáveis aos obtidos por Cerri et al. [CER06]. Nos experimentos, sem zoneamento perceptivo ($Z=0$) a média foi de $R_n = 0,78$ e $P_n = 0,55$. Entretanto, ao aplicar os zoneamentos perceptivos as taxas de recuperação melhoram, atingindo a média geral para o zoneamento de cinco regiões vertical ($Z=5V$), valor de $R_n = 0,81$ e $P_n = 0,59$. Deve-se considerar ainda que, ao selecionarmos o melhor zoneamento perceptivo, para cada MRV consultada, a média geral atinge o valor de $R_n = 0,83$ e $P_n = 0,61$.

Os valores obtidos para R_n e P_n , nos experimentos realizados para Figura-Fundo, são superiores ($R_n = 0,83$ e $P_n = 0,61$) aos apresentados por Cerri et al. [CER06] $R_n = 0,81$ e $P_n = 0,56$. Estes resultados demonstram que é possível incorporar o princípio Figura-Fundo nos sistemas CBIR, de maneira a complementar outras abordagens que utilizam os princípios da Gestalt na recuperação de MRVs.

Novamente, retorna-se aos Princípios de Gestalt exemplificando que um outdoor produzido pela empresa Ogilvy & Mather China, foi premiado no festival de

criatividade em Cannes 2012²¹. A MRV em questão, da empresa Coca-Cola, foi utilizada de maneira simples, conforme ilustrado na Figura 6.1.



Figura 6.1: Outdoor coca-cola

Fonte: <http://www.ogilvy.com>

Observa-se que os princípios da Gestalt são fortemente empregados na Figura 6.1, em especial o princípio do fechamento. Olhando globalmente vemos a famosa onda, que há muito tempo nos lembra da MRV da Coca-Cola. Porém, em uma segunda observação percebendo informações mais relevantes, sendo possível enxergar dois braços e mesmo sem detalhes suficientes enxergamos perfeitamente a garrafa de Coca-Cola. Para a empresa Ogilvy & Mather China, "esses detalhes vão além da garrafa de Coca-Cola, é um gesto de uma pessoa compartilhando sua Coca-Cola com outra, mais do que um produto, uma atitude. Uma riqueza de significados, sem a necessidade de qualquer palavra".

O grande desafio nos sistemas CBIR ainda é incorporar aspectos da percepção humana nos processos computacionais, sendo que neste trabalho foram apresentadas algumas técnicas (a exemplo, da extração de primitivas – Figura-Fundo) que podem auxiliar neste processo, no entanto ainda há muito a ser explorado.

²¹ <http://www.ogilvy.com/News/Press-Releases/June-2012-Cannes-Lions-Grand-Prix-honour-for-Ogilvy-Mather-China.aspx>
<http://propagandaelementar.com/tag/gestalt/>

Para trabalhos futuros, ainda considerando o problema da percepção humana, podem-se realizar experimentos adaptando o algoritmo para extração de concavidade e convexidade, de forma a considerar como fundo (*background*) os pixel pretos, invertendo a percepção de primeiro e segundo planos.

Outro trabalho futuro é aplicar os zoneamentos perceptivos antes de utilizar os extratores de características de contorno, desta forma é possível obter informações adicionais sobre a região interna da MRV.

Ainda, existe a necessidade de complementar o Modelo Hierárquico Perceptivo, estabelecendo um método para rejeitar falsos positivos.

Outra sugestão de pesquisa é a realização de experimentos utilizando o Modelo Hierárquico Perceptivo em outras bases de MRVs, por exemplo a base de MRVs do MPEG-7.

Considera-se, ainda, como trabalho futuro a possibilidade de se incorporar ao Modelo Hierárquico Perceptivo novos extratores, bem como retirar aqueles que não serão aplicados na recuperação em função de algumas particularidades, deixando o modelo mais flexível.

É necessário adicionar ao Modelo Hierárquico Perceptivo a seleção dinâmica dos mecanismos de zoneamento utilizados no segundo nível.

Finalmente, pode-se propor a aplicação do cálculo da distância mínima entre os zoneamentos perceptivos para outros extratores, como por exemplo o extrator de características SURF.

Referências

- [ABE06] ABE, K.; IGUCHI, H.; TIAN, H.; ROY, D. *Recognition of Plural Grouping Patterns in Trademarks for CBIR According to the Gestalt Psychology*. Institute of Electronics, Information on Communication Engineers, vol. E89-D, no. 6, p. 1798-1805, 2006.
- [AIR05] AIRES, S. B. K. *Reconhecimento de Caracteres Manuscritos Baseado em Regiões Perceptivas*. Dissertação de Mestrado. PUCPR, 2005.
- [AIR14] AIRES, S. B. K.; FREITAS, C. O. A. ; OLIVEIRA, L. S. *Feature Analysis for Content-based Trademark Retrieval*. 27th International Conference on Computer Applications in Industry and Engineering, p.245-249, New Orleans, USA, 2014.
- [ALW98] ALWIS, S.; AUSTIN, J. *A Novel Architecture for Trademark Image Retrieval Systems*. Challenge of Image Retrieval, 1998.
- [ALW99] ALWIS, S.; AUSTIN, J. *Trademark Image Retrieval Using Multiple Features*. Challenge of Image Retrieval, Newcastle, 1999.
- [AMA14] AMARAL, A. M. M. M. *Identificação de Autoria de Documentos Manuscritos Utilizando Características Grafométricas*. Tese de Doutorado. PUC-PR, 2014.
- [ANU13] ANUAR, F. M. A.; SETCHI, R.; YU-KUN, L. *Trademark image retrieval using an integrated shape descriptor*. Expert Systems with Applications, vol. 40, no. 1, p. 105-121, 2013.

- [AUM02] AUMONT, J. *A imagem*. Tradução Estela dos Santos Abreu. Campinas, SP: Papirus, 1993. 317p. Disponível em <http://www.nextimagem.com.br/wp-content/uploads/Aumont-Jacques-A-imagem.pdf>. Acessado em 11 de fevereiro de 2015.
- [AUR91] AURENHAMMER, F. *Voronoi diagrams: A survey of a fundamental geometric data structure*. ACM Comput. Surv., vol. 3, no. 3, pp. 345–405, 1991.
- [BAG07] BAGDANOV, A. D.; BALLAN, L.; BERTINI, M.; DEL BIMBO, A. D. *Trademark matching and retrieval in sports video databases*. In Proc. ACM Int. Workshop Multimedia Inf. Retr., p. 79–86, 2007.
- [BEF12] BEFANE, B.; ABE, K.; HAYASHI, T. *A Method for Extracting Grouping Areas of Good Continuity Parts in Trademark Images*. *Image and Vision Computing New Zealand (IVCNZ'12)*, Dunedin, New Zealand. p. 295-300, 2012.
- [BEL06] BELO, F. A. W. *Desenvolvimento de Algoritmos de Exploração e Mapeamento Visual para Robôs Móveis de Baixo Custo*. Dissertação de Mestrado. PUC-RJ, 2006.
- [BEN00] BENNETT, K. P.; CAMPBELL, C. *Support vector machines: Hype or Hallelujah?*. SIGKDD Explorations. vol 2, no. 2, p. 1-13, 2000.
- [BER00] BERG, M.; SCHWARZKOPF, O.; VAN KREVELD, M., OVERMARS, M.. *Computational Geometry: Algorithms and Applications*. Berlin, Germany: Springer-Verlag, 2000.
- [BER02] BERNS, R. M. *O desenvolvimento da criança*. Trad. Cecília C. Bartalotti e Marcos Bagno. São Paulo: Loyola, 2002.
- [BLU03] BLUMENSTEIN, M.; VERMA, B.; BASLI, H. *A novel feature extraction technique for the recognition of segmented handwritten characters*. 7th Int. Conf. Document Analysis and Recognition, ICDAR'03, p. 137–141, 2003.

- [BOB01] BOBER, M. *MPEG-7 Visual Shape Descriptors*. IEEE Transaction on Circuits and Systems for Video Technology, vol. 11, no. 6, p. 716-719, 2001.
- [BOR12] BORGES, H. B. *Classificador Hierárquico Multirótulo usando uma Rede Neural Competitiva*. Tese de Doutorado. PUCPR, 2012.
- [BRA96] BRASIL. Lei nº 9.276, de 14 de maio de 1996. Disponível em http://www.planalto.gov.br/ccivil_03/Leis/L9276.htm. Acessado em 02 de fevereiro de 2015.
- [BRE96] BREIMAN, L. *Bagging predictors*. Journal Machine Learning. vol. 24, p. 123–140, 1996.
- [BRI14] BRITTO JR. A.; SABOURIN, R.; OLIVEIRA, L. S. *Dynamic Selection of Classifiers - A Comprehensive Review*. Pattern Recognition, vol. 47, no.11, p. 3665-3680, 2014.
- [BRO02] BROWN, M.; LOWE, D. G. *Invariant features from interest point groups*. British Machine Vision Conference (BMVC), p. 656-665, 2002
- [BRO05] BROWN, G.; WYATT, J.; HARRIS, R.; YAO, X. *Diversity creation methods: a survey and categorization*. Inf. Fusion, vol. 6, no. 1, p. 5–20, 2005.
- [CAV13] CAVALIN, P. R.; SABOURIN, R. ; SUEN, C. Y. *Dynamic selection approaches for multiple classifier systems*. Neural Computing and Applications, vol. 22, no. 3-4, p.673-688, 2013.
- [CEA10] CEAMANOS, X.; WASKE, B.; BENEDIKTSSON, J. A.; CHANUSSOT, J.; FAUVEL, M.; SVEINSSON, J. R. *A classifier ensemble based on fusion of support vector machines for classifying hyperspectral data*. Int. J. Image Data Fusion, vol.1, no. 4, p.293–307, 2010.
- [CER06] CERRI, M.; FERRI, M.; GIORGI, D. *Retrieval of trademark images by means of size functions*, *Graphical Models*, vol. 68, no. 5-6, p. 451-471, 2006.

- [CHA11b] CHACON, G; GASTARDELLI, E; MORAES, F. D.; OLIVEIRA, G.; ALBUQUERQUE, M. P.; ALBUQUERQUE, M. P. *Aplicação da técnica de momentos invariantes no reconhecimento de padrões em imagens digitais*. Nota técnica. ISSN 0101-9201. Centro Brasileiro de Pesquisas Físicas, CBPF, 2011.
- [CHA99] CHAN, D. Y.; KING, I. *Genetic Algorithm for Weights Assignment in Dissimilarity Function for trademark Retrieval*. In Third International Conference on Visual Information and Information Systems (VISUAL'99), p. 557 – 565, 1999.
- [CHA11a] CHANG, W. H., CHENG, M. C., KUO, C.M.; YANG, N. C.; HUANG, D. S. *An Efficient Contour-Based Layered Shape Descriptor for Image Retrieval*. International Journal of Innovative Computing, Information and Control, vol. 7, no. 7(A), p. 3903-3922, july 2011.
- [CHA11c] CHANG, C.C.; LIN, C. J. *LIBSVM: A library for support vector machines*. ACM Trans. on Intelligent Systems and Technology, vol. 2, no. 27, p. 1–27:27, 2011.
- [CHA15] CHANG, C. C., LIN, C. J. *LIBSVM - A Library for Support Vector Machines*. Disponível em <http://www.csie.ntu.edu.tw/~cjlin/libsvm>. Acessado em 07 de fevereiro de 2015.
- [CHA06] CHAVES, A. C. F. *Extração de Regras Fuzzy para Máquinas de Vetor Suporte (SVM) para Classificação em Múltiplas Classes*. Tese de Doutorado. PUC-RJ, 2006.
- [COR02] CORREIA, S. E. N., CARVALHO, J.M., SABOURIN, R. *Human-Perception Handwritten Character Recognition using Wavelets*. In Proceedings of XVI Brazilian Symposium on Computer Graphics and Image Processing, IEEE Computer Society, p.404, 2002.

- [COS12] COSTA, Y. M. G.; OLIVEIRA, L. S.; KOERICH, A. L.; GOUYON, F.; MARTINS, J. G. *Music genre classification using LBP textural features*. *Signal Processing* 92, p. 2723-2737, 2012.
- [DAT08] DATTA, R.; JOSHI, D.; LI, J.; WANG, J. Z. *Image retrieval: Ideas, influences, and trends of the new age*. *ACM Computing Surveys*, vol. 40, no. 2, p.1–60, 2008.
- [DU_12] DU, P.; XIA, J.; ZHANG, W.; TAN, K.; LIU, Y.; LIU, S. *Multiple Classifier System for Remote Sensing Image Classification: A Review*, *Sensors*, vol.12, no. 4, p. 4764-4792, 2012.
- [DUH12] DUHIGG, C. *O poder do Hábito: Por que fazemos o que fazemos na vida e nos negócios*. Trad. Rafael Mantovani. Rio de Janeiro: Objetiva, 2012.
- [DUI00] DUIN, R. P. W., TAX, D. M. J. *Experiments with Classifier Combining Rules*. J. Kittler; F. ROLI (Eds). *Multiple Classifier System, Lecture Notes in Computer Science*, v. 1897. Berlin Heidelberg: Springer-Verlag, pp. 16-29, 2000.
- [EAK98] EAKINS, J. P.; BOARDMAN, J. M.; GRAHAM, M. E. *Similarity Retrieval of trademarks Image*. *IEEE Multimedia*, vol. 5, no. 2, p. 53-63, 1998.
- [FAC96] FACON, J. *Morfologia matemática: teoria e exemplos*. Editora Universitária Champagnat da Pontifícia Católica do Paraná: Curitiba, 1996.
- [FFW12] FASHION Forward. Gucci x Guess: similaridade, cópia, inspiração ou coincidência? Disponível em <http://ffw.com.br/noticias/moda/gucci-x-guess-similaridade-copia-inspiracao-ou-coincidencia/>. Acesso em 04 de julho de 2012.
- [FRE96] FREUND, Y.; SCHAPIRE, R. E. *Experiments with a New Boosting Algorithm*. In *Proceedings of the International Conference on Machine Learning*, Bari, Italy, 3–6 July, p. 148–156, 1996.

- [FRE74] FREEMAN, H. *Computer Processing of Line-Drawing Images*, ACM Computing Surveys, vol. 6, no.1, p.57-97, 1974.
- [FRE01] FREITAS, C. O. A. *Uso de modelos escondidos de Markov para reconhecimento de palavras manuscritas*. Tese de Doutorado, PUCPR/PPGIA, Curitiba-PR, 2001, 188p.
- [FRE02] FREITAS, C. O. A. *Percepção visual e reconhecimento de palavras manuscritas*. Monografia, Concurso de Promoção da Carreira Docente à Classe de professor Titular, PUCPR/PPGIA, Curitiba-PR, 2002, 95p.
- [FRE07] FREITAS, C. O. A., OLIVEIRA, L. E. S., BORTOLOZZI, F., AIRES, S. B. K.. *Handwritten Character Recognition using Non-Symmetrical Perceptual Zoning*. International Journal of Pattern Recognition and Artificial Intelligence, v. 21, no. 1, p. 135-155, 2007.
- [FRE07a] FREITAS, C. O. A. ; KAPP, Marcelo ; SABOURIN, Robert . *Methodology for the Design of NN-based Month-Word Recognizers Written on Brazilian Bank Checks*. Image and Vision Computing, v. 25, p. 40-49, 2007.
- [GIA01] GIACINTO, G.; ROLI, F. *Dynamic classifier selection based on multiple classifier Behaviour*, Pattern Recognition. vol. 34, p.1879–1881, 2001.
- [GOM04] GOMES FILHO, J. *Gestalt do Objeto: Sistema de Leitura Visual da Forma*. Editora Escrituras, 2a edição. São Paulo-SP, 2004, 95p.
- [GOM07] GOMES, O. F. M.. *Microscopia Co-Localizada: Novas Possibilidades na Caracterização de Minérios*. Tese de Doutorado. PUC-RJ, 2007.
- [GOM12] GOMBRICH, E. H.. *Gombrich Essencial: textos selecionados sobre arte e cultura*. Rev. Técnica Paula Ramos. Tradução Alexandre Salvaterra. Editora Bookman, 1a edição. Porto Alegre, 2012, 624p.

- [GON08] GONZALEZ, R.C.; WOODS, R.E. *Processamento de Imagens Digitais*. Ed. Pearson. 3 edição, 2008.
- [GON10] GONZALES, G. L. G. *Aplicação da Técnica SIFT para Determinação de Campos de Deformações de Materiais usando Visão Computacional*. Dissertação de Mestrado. PUC-RJ, 2010.
- [HAY01] HAYKIN, S. *Redes neurais: princípios e prática*. 2.ed. Tradução de, Paulo Martins Engel. Porto Alegre: Bookman, 2001.
- [HAI10] HAITAO, R.; YELI, L.; LIKUN, L. *Single Closed Contour Trademark Classification based on Support Vector Machine*. 3rd International Congress on Image and Signal Processing (CISP), p. 1942-1946, 2010.
- [HEG12] HEGADI, R. S. *Recognition of Printed Kannada Numerals based on Zoning Method*. International Journal of Computer Applications (0975–8878) on National Conference on Advanced Computing and Communications - NCACC, April 2012.
- [HIR11] HIRABARA, L.Y. *Reconhecimento de caracteres manuscritos baseado na seleção dinâmica de mecanismos de zoneamento*. Dissertação de Mestrado. PUCPR, 2011.
- [HIR09] HIROMASA, I.; KOJI, A.; TADANOBU, M.; HARUHIKO, K.; YOSHIMASA, D. *Recognition of Grouping Patterns in Trademarks Based on the Gestalt Psychology*. Electronics and Communications in Japan, vol. 92, no. 10, p. 844-853, 2009.
- [HU_62] HU, M. K. *Visual Pattern Recognition by Moment Invariants*, IRE Trans. Info. Theory, vol. IT-8, no. 2, p.179–187, 1962.
- [HUN04] HUNG, M. H.; HSIEH, C. H; KUO, C. M. *An Efficient Two-Stage trademark Retrieval System*. Int. Computer Symposium, Dec. 15-17, Taipei, Taiwan, p.214-219, 2004.

- [HUS06] HUSSAIN, M.; EAKINS, J. P. *Component-based visual clustering using the self-organizing map*, Neural Networks, vol. 20, no. 2, p. 260-273, 2006.
- [IMP10] IMPEDOVO, S.; MODUGNO, R. ; PIRLO, G. *Membership Functions for Zoning-based Recognition of Handwritten Digits*. Proc. International Conference on Pattern Recognition, Istanbul, Turkey, p. 1876 – 1879, 2010.
- [IMP12] IMPEDOVO, D.; PIRLO, G.; MODUGNO, R. *New Advancements in Zoning-Based Recognition of Handwritten Characters*. International Conference on Frontiers in Handwriting Recognition, p. 661-665, 2012.
- [INP12] INSTITUTO Nacional da Propriedade Industrial. Disponível em <http://www.inpi.gov.br>. Acesso em 15 de maio de 2012.
- [IPO12] INTELLECTUAL Property Office. Disponível em <http://www.ipo.gov.uk/>. Acesso em 03 de maio de 2012.
- [JAI98] JAIN, A. K.; VAILAYA, A. *Shape-Based Retrieval: A Case Study with Trademark Image Databases*. Pattern Recognition, vol. 31, no. 9, p. 1369-1390, 1998.
- [JAI00] JAIN, A. K.; DUIN, R. P. W.; MAO, J. *Statistical pattern recognition: a review*. IEEE Trans.Pattern Anal. Mach. Intell. vol. 22, no. 1, p 4–37, 2000.
- [JIA06] JIANG, H.; NGO, C. W.; TAN, H. K. *Gestalt-based Feature Similarity Measure in Trademark Database*. ACM, Pattern Recognition, v. 39, no. 5, p.988-1001, 2006.
- [JUS09] JUSOH, N. A; ZAIN, J. M. *Application of Freeman Chain Codes: An Alternative Recognition Technique for Malaysian Car*. International Journal of Computer Science and Network Security, IJCSNS. vol. 09, no. 11, p. 222-227, 2009.

- [KAT92] KATO, T. *Database architecture for content based image retrieval*. Proceedings of SPIE Image Storage and Retrieval Systems, vol. 1662, p. 112-123, 1992.
- [KIM00] KIM , W. Y.; KIM, Y. S. *A region-based shape descriptor using Zernike moments*. Signal Processing: Image Communication, vol. 16, no. 1-2, p.95-102, 2000.
- [KIT98] KITTLER, J.; HATEF, M.; DUIN, R.P.W.; MATAS, J. *On combining classifiers*. IEEE Trans. Pattern Anal.Mach.Intell. vol. 20, no. 3, p.226–239, 1998.
- [KO_08] KO, A.; SABOURIN, R.; BRITTO Jr, A. *From dynamic classifier selection to dynamic ensemble selection*. Pattern Recognition, vol. 41, no. 5, p. 1718–1731, 2008.
- [KOE05] KOERICH, A. L.; KALVA, P. R. *Unconstrained handwritten character recognition using metaclasses of characters*. IEEE Int. Conf. Image Processing (ICIP), p. 542–545, 2005.
- [KOR06] KORTINF, T.S. *Classificação de Imagem por Regiões*. Nota Técnica. INPE, 2006.
- [KUM11] KUMAR, R.; TRIPATHI; R. C.; TIWARI, M. D. *A Comprehensive Study on Content based Trademark Retrieval System*. International Journal of Computer Application, vol.13, no. 6, p. 18-22, 2011.
- [LEC00] LECCE, V.; DIMAURO, G.; GUERRERO, A.; IMPEDOVO, S.; PIRLO, G.; SALZO, A. *Zoning design for handwritten numerical recognition*. 7th Int. Workshop on Frontiers in Handwriting Recognition, p. 583–588, 2000.
- [LI_04] LI, M.; CHEN, X; LI, X.; MA, B.; VITANYI, P.M.B. *The Similarity Metric*. IEEE Transations on Information Theory, vol.50, no.12, p. 3250-3264, 2004.

- [LIN09] LINDEN, R. *Técnicas de Agrupamento*. Revista de Sistemas de Informação. FSMA, vol. n. 4, p.18-36, 2009.
- [LIU14] LIU, Y. *Sistema de recomendação dos amigos na rede social online baseado em Máquinas de Vetores Suporte*. Dissertação de Mestrado. UnB - Brasília, 2014.
- [LOW99] LOWE, D. G. *Object Recognition from Local Scale-Invariant Features*. Proc. of the International Conference on Computer Vision, p. 1-8, 1999.
- [LOW04] LOWE, D. *Distinctive image features from scale-invariant keypoints*. Int. J. Comput. Vis., vol. 60, no. 2, p. 91–110, 2004.
- [MOR04] MORITA, M. E.; SABOURIN, R. ; BORTOLOZZI, F; SUEN, C. Y. *Segmentation and recognition of handwritten dates: an HMM-MLP hybrid approach*, Int. J. Docum. Analy.Recogn. vol. 6, no. 8, p. 248–262, 2004.
- [MUR11] MURTHY, O. V. R.; HANMANDLU, M. *Zoning based Devanagari Character Recognition*. International Journal of Computer Applications, vol. 27, no.4, p. 975 – 8887, 2011.
- [NAB12] NABIHA, A.; NADIR, F. *New dynamic ensemble of classifiers selection approach based on confusion matrix for arabic handwritten recognition*. International Conference on Multimedia Computing and Systems (ICMCS), p.308–313, 2012.
- [NIE04] NIEVOLA, J. C. *Redes neurais artificiais*. In: XII Escola regional de informática – ERI 2004, Anais, p.01-50, ISBN 858844287-6, Universidade Estadual do Centro-Oeste – UNICENTRO, Guarapuava, 2004.
- [OKA92] OKABE, A.; BOOTS, B.; SUGIHARA, K. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Chichester, U.K.: Wiley, 1992.
- [OLI04] OLIVEIRA JR, J. J.; KAPP, M. N.; FREITAS, C.; DE CARVALHO, J. M. e SABOURIN, R., *Handwritten Month Word Recognition Using Multiple*

Classifier, XVII Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI), outubro 17-20, 2004.

- [OLI09] OLIVEIRA JR, J. J. ; CARVALHO, J. M. ; FREITAS, C. O. A. ; SABOURIN, R. *Handwritten Word Recognition Using Multi-View Analysis*. In: 14th Iberoamerican Congress on Pattern Recognition - CIARP, Guadalajara. Proc. of 14th Iberoamerican Congress on Pattern Recognition. Germany: Springer - Lecture Notes in Computer Science, p. 371-378, 2009.
- [OTS79] OTSU, N. *A threshold selection method from gray-level histograms*. IEEE Transactions Systems, Man. and Cybernetics, SMC 9, vol.1, p.63-66, 1979.
- [PAR97] PARKER, J. R. *Algorithms for Image Processing and Computer Vision*. Ed. John Wiley & Sons, p.310-315, 1997.
- [PEN09] PENATTI, O. A. B. *Estudo comparativo de descritores para recuperação de imagens por conteúdo na Web*. Dissertação de Mestrado. UNICAMP, 2009.
- [PIR12] PIRLO, G.; IMPEDOVO, D. *Adaptive Membership Functions for Handwritten Character Recognition by Voronoi-Based Image Zoning*. IEEE Transactions on Image Processing, vol. 21, no. 9, p. 227-231, 2012.
- [PON04] PONTI Jr, M. P. *Combinação de Múltiplos Classificadores para Identificação de Materiais em Imagens Ruidosas*. Dissertação de Mestrado. UFSCAR, São Carlos-SP, 2004.
- [QI_09] QI, Y.L. *A Relevance Feedback Method to Trademark Retrieval Based on SVM*. International Forum on Computer Science-Technology and Applications, vol 1. Washington(USA) , 2009.
- [RAD03] RADTKE, P. V. W; OLIVEIRA, L. S., SABOURIN, R. ; WONG, T. *Intelligent zoning design using multi-objective evolutionary algorithms*. 7th Int. Conf. Document Analysis and Recognition (ICDAR2003), p. 824–828, 2003.

- [RAH99] RAHMAN, A. F. R.; FAIRHURST, M. C. *Serial combination of multiple experts: A unified evaluation*. Pattern Anal. Appl., vol. 2, no. 4, p. 292–311, 1999.
- [RAJ09] RAJASHEKARARADHYA, S. V.; RANJAN, P. V. *A novel zone based feature extraction algorithm for handwritten numeral recognition of four Indian scripts*. Digital Technol. J., vol. 2, no. 1, p. 41–51, 2009.
- [RAN06] RANAWANA, R.; PALADE, V. *Multi-classifier systems: Review and a roadmap for developers*. Int. J. Hybrid Intell. Syst., vol. 3, no. 1, p. 1–41, 2006.
- [REZ05] REZENDE, S, O. *Sistemas Inteligentes: Fundamentos e Aplicações*. Barueri, SP: Manole, 2005.
- [ROS56] ROSENBLATT, F. *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*. Psychological Review, vol. 65, no. 6, p. 386-408, 1956.
- [ROY09] ROY, P. P.; PAL, U.; LLADÓS, J.; DELALANDRE, M. *Multi-Oriented and Multi-Sized Touching Character Segmentation using Dynamic Programming*. 10th International Conference on Document Analysis and Recognition, p. 11-15, 2009.
- [RUF11] RUFINO, H. L. P. *Algoritmo de Aprendizado Supervisionado - Baseado em Máquina de Vetores de Suporte - Uma contribuição Para o Reconhecimento de Dados Desbalanceados*. Tese de Doutorado. Universidade Federal de Uberlândia, 2011.
- [RUM86] RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. *Learning Representations of Back-Propagation Errors*. Nature (London), vol. 323, no. p. 533-536, 1986.
- [SAB93] SABOURIN, M.; MITICHE, A.; THOMAS, D.; NAGY, G. *Classifier combination for hand- printed digit recognition*. Proceedings of the Second

International Conference on Document Analysis and Recognition, p.163–166, 1993.

[SAF13] SAFI, L. A.; SRINIVASAGAN, K. G. *Offline Tamil Handwritten Character Recognition using Zone based Hybrid Feature Extraction Technique*. International Journal of Computer Applications, vol. 65, no.1, p. 975 – 8887, 2013.

[SAH13] SAHBI, H.; BALLAN, L.; SERRA, G.; DEL BIMBO, A. *Context-dependent logo matching and recognition*. In IEEE Transactions on Image Processing, vol. 22, no. 3, p. 1018-1030, 2013.

[SAL71] SALTON, G. *The Smart Retrieval System - Experiments in Automatic Document Processing*. Prentice-Hall, N.J., 1971.

[SAN06] DOS SANTOS. E. M.; SABOURIN, R.; MAUPIN, P. *Single and multi objective genetic algorithms for the selection of ensemble of classifiers*. Proceedings of International Joint Conference on Neural Networks, Vancouver-Canada, p. 3070–3077, 2006.

[SCA05] SCARPEL, R. A. *Utilização de Support Vector Machine em Previsão de Insolvência de Empresas*. XXXVII Simpósio Brasileiro de Pesquisa Operacional (SBPO). Gramado- RS, 2005

[SHA10] SHARMA, D.; GUPTA, D. *Isolated handwritten digit recognition using adaptive unsupervised incremental learning technique*. Int. J. Comput. Appl., vol. 7, no. 4, p. 27–33, 2010.

[SHA14] SHAABAN, Z. *Trademark Image Retrieval System using Neural Network*. International Journal of Computer Science and Network (IJCSN), vol. 3, no. p. 73-82, 2014.

- [SIL11] SILVA, A. T. *Recuperação de imagens por conteúdo baseada em realimentação de relevância e classificador por floresta de caminhos ótimos*. Tese de Doutorado. UNICAMP, 2011.
- [SNO10] SNOEK, C. G. M.; SMEULDERS, A. W. M. *Visual-concept search solved?* IEEE Computer, vol. 43, no.6, p. 76–78, 2010.
- [SOL05] SOLANA, C. D. O. *Reconstrução Digital de Documentos por Aproximação Poligonal*. Dissertação de Mestrado. PUC-PR, 2005.
- [VAP99] VAPNIK, V. N. *The nature of statistical learning theory*. Editora Springer-Verlag, 2a. edição. New York, 1999.
- [VER10] VERAS, Eduardo. *Afinal, o que são trademarks?* Disponível em http://www.fdc.org.br/pt/blog_marketing/Lists/Postagens/Post.aspx?ID=7, acessado em 04 de julho de 2012.
- [WAN12] WANG, Z.; HONG, K. *A Novel Approach for Trademark Image Retrieval by Combining Global Features and Local Features*. Journal of Computational Information Systems. vol.8, no. 4, p. 1633-1640, 2012.
- [WIP12] WIPO. World Intellectual Property Organization. *International Classification of the Figurative Elements of Marks – VIENNA Classification*. Disponível em <http://www.wipo.int/classifications/vienna/en>, acessado em 02 de julho de 2012.
- [XIA04] XIANG, P.; XIUZI, Y.; SANYUAN, Z. *A hybrid method for robust car plate character recognition*. IEEE Int. Conf. Syst. Man Cybern, p. 4377–4737, 2004.
- [XU_92] XU, L.; KRZYSAK, A.; SUEN, C.Y. *Methods of combining multiple classifiers and their applications to handwriting recognition*. IEEE Trans. Syst. Man Cybern. vol. 22, no. 3, p. 418–435, 1992.

[YOS99] YOSHITAKA, A.; ICHIKAWA, T. *A survey on content-based retrieval for multimedia databases*. IEEE Transactions on Knowledge and Data Engineering, vol. 11, no. 1, p.81–93, 1999.

[ZHA04] ZHANG, D.; LU, G. *Review of Shape Representation and Description*. Pattern Recognition, Pattern Recognition Society. vol. 37, p. 1-19, 2004.

APÊNDICE A

Foram realizados experimentos com os extratores de características aplicando o cálculo de similaridade através da Distância de Histograma. As Tabelas A.1 até A.3, apresentam os resultados obtidos. Tais resultados são apresentados individualmente conforme o extrator e o zoneamento perceptivo. Ao final as Tabelas A.4 até A.6, apresenta um comparativo das taxas alcançadas pela Distância Euclidiana (DE) e Distância de Histograma (DH).

Tabela A.1: Distância de Histogramas - Concavidade e Convexidade

MRV		Z=0		Z=4		Z=5h		Z=5v		Z=7	
		R _n	P _n								
1037814.bmp		0.888475	0.685977	0.901314	0.712824	0.903951	0.676858	0.895149	0.622653	0.902602	0.659945
1055261.bmp		0.829261	0.686390	0.871349	0.756150	0.852676	0.769176	0.825290	0.705556	0.829724	0.706759
1138103.bmp		0.712069	0.463848	0.692540	0.488121	0.676020	0.482985	0.684305	0.452899	0.741715	0.516343
1138293_.bmp		0.930500	0.797959	0.764313	0.743640	0.823755	0.755354	0.841643	0.723444	0.908155	0.778804
1190540.bmp		0.815472	0.622902	0.779262	0.547997	0.751129	0.411846	0.760951	0.456731	0.689656	0.365472
1259886.bmp		0.883003	0.536518	0.823721	0.458429	0.693327	0.386092	0.827035	0.473503	0.805036	0.610899
1267206.bmp		0.808365	0.525845	0.677793	0.481399	0.574924	0.389804	0.616828	0.404056	0.582518	0.402004
1279931.bmp		0.952610	0.702705	0.871865	0.593388	0.872639	0.526912	0.863034	0.566942	0.850702	0.542619
1289047.bmp		0.896326	0.566516	0.782556	0.453803	0.793048	0.491623	0.716152	0.430579	0.754418	0.537182
1376861.bmp		0.926645	0.535120	0.831164	0.360956	0.804187	0.402370	0.712112	0.319709	0.714085	0.370677
1439229_.bmp		0.925984	0.560975	0.576229	0.405746	0.744363	0.415288	0.763262	0.397980	0.806930	0.422501
1486213_.bmp		0.453577	0.203562	0.417803	0.193234	0.506247	0.197133	0.547422	0.220892	0.459546	0.173397
1575268_.bmp		0.657906	0.536011	0.684394	0.521904	0.678885	0.530202	0.699349	0.549302	0.694083	0.527085
2010916_.bmp		0.834566	0.515891	0.830309	0.575716	0.819073	0.526914	0.823684	0.564243	0.856388	0.585368
2016658_.bmp		0.807706	0.356301	0.678996	0.295479	0.677301	0.399568	0.662824	0.464200	0.583841	0.289174
2018809_.bmp		0.668845	0.304666	0.622960	0.267388	0.565869	0.248063	0.614407	0.306635	0.531253	0.282584
2042822_.bmp		0.740963	0.312449	0.475246	0.198634	0.662680	0.263593	0.677305	0.258093	0.709740	0.294368
665322.bmp		0.702436	0.486606	0.677152	0.448276	0.623489	0.423775	0.669806	0.454000	0.746228	0.612369
914.bmp		0.732687	0.348192	0.672664	0.271148	0.703366	0.312080	0.722255	0.325809	0.825847	0.440676
967049.bmp		0.917189	0.778217	0.871754	0.797609	0.870191	0.774328	0.900739	0.747262	0.912636	0.787568
Média		0.804229	0.526332	0.725169	0.478592	0.729856	0.469198	0.741178	0.472225	0.745255	0.495290
Desvio Padrão		0.124539	0.161759	0.133157	0.18287	0.109419	0.164292	0.099325	0.151006	0.128102	0.171579

Tabela A.2: Distância de Histogramas - Cadeias de Freeman

MRV		Z=0		Z=4		Z=5h		Z=5v		Z=7	
		R _n	P _n								
1037814.bmp		0.745588	0.659857	0.686389	0.528248	0.698420	0.538318	0.732855	0.561336	0.770534	0.600758
1055261.bmp		0.816441	0.657694	0.771516	0.636497	0.774969	0.637328	0.800154	0.693895	0.776240	0.566637
1138103.bmp		0.757233	0.440952	0.720041	0.376535	0.765074	0.341317	0.739258	0.384149	0.715283	0.366403
1138293_.bmp		0.884134	0.775630	0.753910	0.490643	0.811439	0.488335	0.782023	0.360040	0.607968	0.260740
1190540.bmp		0.898274	0.535914	0.965585	0.735555	0.968327	0.741502	0.955744	0.640375	0.638931	0.427493
1259886.bmp		0.624508	0.255422	0.589042	0.225255	0.576899	0.224681	0.560814	0.214821	0.524222	0.211205
1267206.bmp		0.829962	0.453158	0.888426	0.529012	0.738399	0.426165	0.869679	0.528305	0.874736	0.540055
1279931.bmp		0.645532	0.428032	0.733471	0.464780	0.743522	0.425220	0.772903	0.381940	0.704651	0.298629
1289047.bmp		0.817262	0.477387	0.916867	0.785134	0.898053	0.737806	0.864069	0.558314	0.758218	0.529811
1376861.bmp		0.774000	0.450082	0.607764	0.313947	0.591931	0.311602	0.567931	0.305887	0.560247	0.321500
1439229_.bmp		0.757247	0.329292	0.618944	0.322130	0.591139	0.316144	0.532987	0.306314	0.580151	0.319406
1486213_.bmp		0.721325	0.331965	0.749229	0.435412	0.663098	0.340882	0.759602	0.325854	0.721306	0.327467
1575268_.bmp		0.800234	0.540766	0.697126	0.406956	0.694341	0.420752	0.746240	0.493276	0.689311	0.459941
2010916_.bmp		0.735354	0.473419	0.851489	0.449078	0.888327	0.440762	0.879541	0.411909	0.883790	0.414844
2016658_.bmp		0.818369	0.372640	0.806876	0.414471	0.836294	0.388981	0.837907	0.439041	0.846736	0.446526
2018809_.bmp		0.596288	0.313536	0.461162	0.270903	0.447310	0.267796	0.448835	0.271027	0.443446	0.260699
2042822_.bmp		0.574160	0.210479	0.599605	0.227583	0.599201	0.228108	0.619137	0.301135	0.555968	0.245709
665322.bmp		0.867705	0.666343	0.605719	0.473796	0.611744	0.532650	0.602860	0.469281	0.579331	0.512041
914.bmp		0.925291	0.774154	0.789563	0.566967	0.793384	0.579540	0.769464	0.541975	0.741105	0.601319
967049.bmp		0.916840	0.807952	0.699850	0.332471	0.698507	0.300400	0.760330	0.386617	0.809081	0.651729
Média		0.775287	0.497734	0.725629	0.449269	0.719519	0.434414	0.730117	0.428775	0.689063	0.418146
Desvio Padrão		0.103893	0.176737	0.125213	0.153308	0.128533	0.154433	0.133149	0.129098	0.123578	0.135473

Tabela A.3: Distância de Histogramas - SIFT

MRV		Z=0		Z=4		Z=5h		Z=5v		Z=7	
		R _n	P _n								
1037814.bmp		0.868008	0.536385	0.786503	0.415386	0.809386	0.444543	0.821671	0.594080	0.822849	0.645571
1055261.bmp		0.863807	0.642272	0.800314	0.456984	0.824704	0.479169	0.852701	0.538778	0.774519	0.349636
1138103.bmp		0.735625	0.368049	0.566049	0.334803	0.526096	0.260642	0.476461	0.232383	0.445729	0.207176
1138293_.bmp		0.836988	0.455645	0.774608	0.331711	0.655113	0.260689	0.806676	0.350565	0.732354	0.352269
1190540.bmp		0.684193	0.311992	0.846524	0.457930	0.647845	0.284214	0.890297	0.480605	0.740676	0.331550
1259886.bmp		0.791581	0.395476	0.730511	0.329268	0.787646	0.369087	0.733332	0.311396	0.611506	0.247328
1267206.bmp		0.572010	0.259991	0.645526	0.272162	0.614569	0.275036	0.546333	0.255046	0.570746	0.232766
1279931.bmp		0.788087	0.402303	0.709441	0.307497	0.762399	0.347811	0.753018	0.318268	0.797276	0.412271
1289047.bmp		0.457191	0.275486	0.813593	0.407510	0.847964	0.354858	0.930436	0.518941	0.947332	0.531828
1376861.bmp		0.507916	0.215974	0.272593	0.163144	0.311763	0.173857	0.256428	0.159032	0.282652	0.168958
1439229_.bmp		0.786564	0.482628	0.764993	0.481831	0.814632	0.498671	0.748395	0.476072	0.698558	0.392183
1486213_.bmp		0.770745	0.391256	0.681222	0.284901	0.707030	0.350676	0.645350	0.246804	0.665164	0.255859
1575268_.bmp		0.911255	0.563356	0.509560	0.301760	0.716201	0.359164	0.504727	0.302156	0.572226	0.305619
2010916_.bmp		0.527337	0.234204	0.688513	0.286885	0.754348	0.324415	0.684034	0.295375	0.754052	0.308839
2016658_.bmp		0.566776	0.227773	0.527537	0.223906	0.502571	0.221602	0.493000	0.187422	0.501364	0.190814
2018809_.bmp		0.693088	0.277258	0.818872	0.416024	0.809512	0.356459	0.782500	0.370051	0.781065	0.319271
2042822_.bmp		0.449288	0.176073	0.567635	0.214115	0.548183	0.198690	0.515842	0.182557	0.577554	0.217772
665322.bmp		0.303727	0.206308	0.314367	0.278391	0.300335	0.271245	0.318489	0.218650	0.347628	0.277433
914.bmp		0.412404	0.152776	0.551987	0.199779	0.579143	0.208052	0.577611	0.217474	0.604420	0.220904
967049.bmp		0.800780	0.494012	0.777752	0.541350	0.753320	0.367984	0.816668	0.599933	0.772653	0.383427
Média		0.666369	0.353461	0.657405	0.335267	0.663638	0.320343	0.657698	0.342779	0.650016	0.317574
Desvio Padrão		0.177497	0.141098	0.163005	0.102872	0.161433	0.090117	0.189002	0.142451	0.166615	0.117474

Tabela A.4: Comparativo DE e DH - Deficiências de Concavidade e Convexidade - Rn Top-100

MRV	Z=0		Z=4		Z=5h		Z=5v		Z=7		Best
	DE	DH									
1037814.bmp	0.913979	0.888475	0.917314	0.903354	0.923221	0.903951	0.917584	0.895149	0.923100	0.902602	0.923221
1055261.bmp	0.897219	0.829261	0.911045	0.841965	0.909472	0.852676	0.894974	0.825290	0.865497	0.829724	0.911045
1138103.bmp	0.737104	0.712069	0.812021	0.671672	0.811405	0.676020	0.814224	0.684305	0.851087	0.741715	0.851087
1138293_.bmp	0.955971	0.930500	0.991243	0.748536	0.990405	0.823755	0.991243	0.841643	0.990652	0.908155	0.991243
1190540.bmp	0.851790	0.815472	0.880515	0.748486	0.921053	0.751129	0.885327	0.760951	0.905818	0.689656	0.921053
1259886.bmp	0.848089	0.883003	0.916820	0.784151	0.918092	0.693327	0.901978	0.827035	0.866385	0.805036	0.918092
1267206.bmp	0.677049	0.808365	0.729084	0.612283	0.718567	0.574924	0.738139	0.616828	0.730832	0.582518	0.808365
1279931.bmp	0.953005	0.952610	0.925494	0.893350	0.929644	0.872639	0.926351	0.863034	0.931582	0.850702	0.953005
1289047.bmp	0.747940	0.896326	0.825443	0.764412	0.840423	0.793048	0.827632	0.716152	0.813889	0.754418	0.896326
1376861.bmp	0.794684	0.926645	0.905872	0.783011	0.913152	0.804187	0.915806	0.712112	0.911718	0.714085	0.926645
1439229_.bmp	0.677129	0.925984	0.506026	0.795679	0.508710	0.744363	0.509959	0.763262	0.528769	0.806930	0.925984
1486213_.bmp	0.680433	0.453577	0.762888	0.572348	0.744234	0.506247	0.770837	0.547422	0.734799	0.459546	0.770837
1575268_.bmp	0.654612	0.657906	0.674113	0.674750	0.707316	0.678885	0.678999	0.699349	0.661593	0.694083	0.707316
2010916_.bmp	0.816665	0.834566	0.863391	0.818177	0.871914	0.819073	0.867837	0.823684	0.816451	0.856388	0.871914
2016658_.bmp	0.659382	0.807706	0.705319	0.679895	0.724799	0.677301	0.733135	0.662824	0.756254	0.583841	0.807706
2018809_.bmp	0.475659	0.668845	0.547983	0.590371	0.544648	0.565869	0.530832	0.614407	0.498198	0.531253	0.668845
2042822_.bmp	0.758863	0.740963	0.706144	0.642354	0.685935	0.662680	0.715997	0.677305	0.722987	0.709740	0.758863
665322.bmp	0.832462	0.702436	0.863169	0.692338	0.837925	0.623489	0.856050	0.669806	0.863386	0.746228	0.863386
914.bmp	0.753039	0.732687	0.842490	0.749086	0.816295	0.703366	0.834093	0.722255	0.863311	0.825847	0.863311
967049.bmp	0.885958	0.917189	0.865320	0.900618	0.855350	0.870191	0.865791	0.900739	0.864364	0.912636	0.917189
Média	0.778552	0.804229	0.807585	0.743342	0.808628	0.729856	0.808839	0.741178	0.805034	0.745255	0.862772
Desvio Padrão	0.120504	0.124539	0.12844	0.100305	0.129554	0.112262	0.127017	0.099325	0.129089	0.128102	0.084756

Tabela A.5: Comparativo DE e DH - Cadeias de Freeman - Rn Top-100

MRV	Geral		Z4		Z5h		Z5v		Z7		Best
	DE	DH									
1037814.bmp	0.786716	0.745588	0.735814	0.686389	0.726180	0.698420	0.704129	0.732855	0.712802	0.770534	0.786716
1055261.bmp	0.792322	0.816441	0.796275	0.771516	0.769863	0.774969	0.838733	0.800154	0.851924	0.776240	0.851924
1138103.bmp	0.789920	0.757233	0.805323	0.720041	0.791383	0.765074	0.745381	0.739258	0.744222	0.715283	0.805323
1138293_.bmp	0.795050	0.884134	0.795346	0.753910	0.738261	0.811439	0.795030	0.782023	0.566887	0.607968	0.884134
1190540.bmp	0.891746	0.898274	0.978533	0.965585	0.978661	0.968327	0.940430	0.955744	0.849088	0.638931	0.978661
1259886.bmp	0.385766	0.624508	0.474768	0.589042	0.493127	0.576899	0.514331	0.560814	0.453709	0.524222	0.624508
1267206.bmp	0.885593	0.829962	0.899220	0.888426	0.884902	0.738399	0.872736	0.869679	0.860588	0.874736	0.89922
1279931.bmp	0.576236	0.645532	0.697555	0.733471	0.662669	0.743522	0.704563	0.772903	0.674190	0.704651	0.772903
1289047.bmp	0.725937	0.817262	0.842939	0.916867	0.850733	0.898053	0.792996	0.864069	0.762702	0.758218	0.916867
1376861.bmp	0.902968	0.774000	0.545114	0.607764	0.579254	0.591931	0.547149	0.567931	0.587879	0.560247	0.902968
1439229_.bmp	0.466881	0.757247	0.543472	0.618944	0.523040	0.591139	0.540327	0.532987	0.585114	0.580151	0.757247
1486213_.bmp	0.766058	0.721325	0.752942	0.749229	0.761248	0.663098	0.785705	0.759602	0.785280	0.721306	0.785705
1575268_.bmp	0.855858	0.800234	0.753183	0.697126	0.711034	0.694341	0.777441	0.746240	0.773753	0.689311	0.855858
2010916_.bmp	0.788071	0.735354	0.867253	0.851489	0.905291	0.888327	0.911028	0.879541	0.924310	0.883790	0.92431
2016658_.bmp	0.852163	0.818369	0.824992	0.806876	0.806435	0.836294	0.802116	0.837907	0.770505	0.846736	0.852163
2018809_.bmp	0.616720	0.596288	0.518155	0.461162	0.522629	0.447310	0.543948	0.448835	0.590748	0.443446	0.61672
2042822_.bmp	0.452776	0.574160	0.592982	0.599605	0.579641	0.599201	0.678367	0.619137	0.629554	0.555968	0.678367
665322.bmp	0.892959	0.867705	0.697752	0.605719	0.699073	0.611744	0.724366	0.602860	0.756129	0.579331	0.892959
914.bmp	0.968185	0.925291	0.881323	0.789563	0.890291	0.793384	0.863989	0.769464	0.800398	0.741105	0.968185
967049.bmp	0.849690	0.916840	0.782888	0.699850	0.742386	0.698507	0.823368	0.760330	0.855509	0.809081	0.91684
Média	0.752081	0.775287	0.739291	0.725629	0.730805	0.719519	0.745307	0.730117	0.726765	0.689063	0.833579
Desvio Padrão	0.165368	0.103893	0.139602	0.125213	0.138098	0.128533	0.126652	0.133149	0.123532	0.123578	0.104448

Tabela A.6²²: Comparativo DE e DH - SIFT- Rn Top-100

MRV	Z=0		Z=4		Z=5h		Z=5v		Z=7		Best
	DE	DH									
1037814.bmp	0.912539	0.868008	0.823350	0.786503	0.790173	0.809386	0.833166	0.821671	0.916524	0.822849	0.916524
1055261.bmp	0.879483	0.863807	0.768679	0.800314	0.807764	0.824704	0.779002	0.852701	0.748625	0.774519	0.879483
1138103.bmp	0.947842	0.735625	0.680384	0.566049	0.712735	0.526096	0.708477	0.476461	0.636453	0.445729	0.947842
1138293_.bmp	0.807475	0.836988	0.793117	0.774608	0.797840	0.655113	0.804674	0.806676	0.854304	0.732354	0.854304
1190540.bmp	0.972074	0.684193	0.948161	0.846524	0.974440	0.647845	0.970871	0.890297	0.972054	0.740676	0.97444
1259886.bmp	0.748604	0.791581	0.796776	0.730511	0.831000	0.787646	0.846196	0.733332	0.817452	0.611506	0.846196
1267206.bmp	0.680267	0.572010	0.680312	0.645526	0.665663	0.614569	0.646046	0.546333	0.553371	0.570746	0.680312
1279931.bmp	0.864302	0.788087	0.828750	0.709441	0.757719	0.762399	0.803115	0.753018	0.709337	0.797276	0.864302
1289047.bmp	0.481493	0.457191	0.773131	0.813593	0.838297	0.847964	0.862370	0.930436	0.855706	0.947332	0.947332
1376861.bmp	0.532625	0.507916	0.660615	0.272593	0.698727	0.311763	0.656679	0.256428	0.728080	0.282652	0.72808
1439229_.bmp	0.600605	0.786564	0.551009	0.764993	0.546671	0.814632	0.580972	0.748395	0.525461	0.698558	0.814632
1486213_.bmp	0.606401	0.770745	0.438006	0.681222	0.347903	0.707030	0.420652	0.645350	0.490152	0.665164	0.770745
1575268_.bmp	0.980409	0.911255	0.871413	0.509560	0.845615	0.716201	0.884881	0.504727	0.724532	0.572226	0.980409
2010916_.bmp	0.917521	0.527337	0.945984	0.688513	0.944398	0.754348	0.945540	0.684034	0.918023	0.754052	0.945984
2016658_.bmp	0.576173	0.566776	0.557628	0.527537	0.539018	0.502571	0.473978	0.493000	0.488762	0.501364	0.576173
2018809_.bmp	0.587171	0.693088	0.433450	0.818872	0.411287	0.809512	0.429380	0.782500	0.421642	0.781065	0.818872
2042822_.bmp	0.814892	0.449288	0.482259	0.567635	0.535294	0.548183	0.520195	0.515842	0.590495	0.577554	0.814892
665322.bmp	0.514959	0.303727	0.766266	0.314367	0.819495	0.300335	0.802169	0.318489	0.717424	0.347628	0.819495
914.bmp	0.431992	0.412404	0.457754	0.551987	0.468374	0.579143	0.519558	0.577611	0.678416	0.604420	0.678416
967049.bmp	0.926195	0.800780	0.751635	0.777752	0.693066	0.753320	0.787889	0.816668	0.761909	0.772653	0.926195
Média	0.739151	0.666369	0.700434	0.657405	0.701274	0.663638	0.713791	0.657698	0.705436	0.650016	0.839231
Desvio Padrão	0.183865	0.177497	0.163674	0.163005	0.17447	0.161433	0.171951	0.189002	0.157658	0.166615	0.109668

²² A Distância Euclidiana na Tabela A.6, foi calculada diretamente sobre os vetores de características obtidos pelo extrator SIFT.

APÊNDICE B

Alguns dos experimentos com os extratores de características aplicando o cálculo de similaridade através da Distância de Mahalanobis são apresentados nas Tabelas B.1 e B.2.

Tabela B.1: SIFT - Rn Top-100

MRV		R_n	P_n	L_n
1037814.bmp		0.355145	0.171687	0.402963
1055261.bmp		0.402670	0.230657	0.148693
1138103.bmp		0.893957	0.696974	0.486537
1138293_.bmp		0.746820	0.687902	0.044078
1190540.bmp		0.890445	0.659990	1.000000
1259886.bmp		0.664618	0.245113	0.629366
1267206.bmp		0.572333	0.279596	0.678600
1279931.bmp		0.605861	0.234840	0.461903
1289047.bmp		0.340371	0.343651	0.011159
1376861.bmp		0.668675	0.275209	0.487574
1439229_.bmp		0.568363	0.429390	0.533228
1486213_.bmp		0.131198	0.107200	0.373162
1575268_.bmp		0.993809	0.955672	1.000000
2010916_.bmp		0.736028	0.426769	0.648190
2016658_.bmp		0.416286	0.176763	0.663805
2018809_.bmp		0.361180	0.174676	0.261538
2042822_.bmp		0.565092	0.225959	0.889042
665322.bmp		0.945666	0.559845	0.990533
914.bmp		0.256139	0.111564	0.375395
967049.bmp		0.796197	0.554490	0.581574
Média		0.595543	0.377397	0.533367
Desvio Padrão		0.243927	0.235458	0.294227

Tabela B.2: Conc.Conv. com Cadeias de Frreman - Top-100

MRV		Z=4		Z=5h		Z=5v		Z=7	
		R _n	P _n						
1037814.bmp		0.871670	0.624277	0.874948	0.630635	0.884220	0.637561	0.897083	0.703861
1055261.bmp		0.844752	0.727662	0.811526	0.711095	0.831543	0.725541	0.824975	0.717622
1138103.bmp		0.847898	0.439341	0.857810	0.453839	0.876138	0.462152	0.883116	0.465605
1138293_.bmp		0.831190	0.796339	0.851484	0.828642	0.845893	0.764733	0.818134	0.550547
1190540.bmp		0.624692	0.570335	0.612366	0.585368	0.604861	0.519773	0.601016	0.558046
1259886.bmp		0.744872	0.370402	0.720138	0.357262	0.728358	0.472141	0.674636	0.423472
1267206.bmp		0.950197	0.647041	0.919204	0.619128	0.918307	0.663989	0.911323	0.619777
1279931.bmp		0.681577	0.474388	0.593425	0.321594	0.672627	0.428264	0.585861	0.322513
1289047.bmp		0.762971	0.724686	0.775667	0.727170	0.767276	0.596823	0.739052	0.535439
1376861.bmp		0.712596	0.290104	0.672396	0.293749	0.620827	0.257443	0.685745	0.281144
1439229_.bmp		0.554493	0.323053	0.577664	0.330659	0.590307	0.329970	0.630272	0.345563
1486213_.bmp		0.583288	0.236026	0.586667	0.255202	0.668661	0.267896	0.549383	0.224656
1575268_.bmp		0.829241	0.594652	0.839067	0.576057	0.752599	0.638367	0.823869	0.651719
2010916_.bmp		0.950275	0.609210	0.959299	0.678104	0.956529	0.641320	0.950686	0.633405

APÊNDICE C

Alguns experimentos combinando características dos diferentes extratores foram realizados. As combinações foram realizadas unindo os vetores de características dos extratores e normalizando os dados com base no maior valor. A similaridade é calculada utilizando distância euclidiana. As Tabelas C.1 até C.4, apresentam resultados obtidos na recuperação das Top-100 MRVs.

Tabela C.1: Combinação 4 extratores $Z=0$ (único vetor normalizado) - Top-100

MRV		R_n	P_n
1037814.bmp		0.918435	0.672778
1055261.bmp		0.842211	0.742504
1138103.bmp		0.912426	0.793136
1138293_.bmp		0.796056	0.748464
1190540.bmp		0.963643	0.778298
1259886.bmp		0.774179	0.480230
1267206.bmp		0.663116	0.355286
1279931.bmp		0.854567	0.579918
1289047.bmp		0.411262	0.308571
1376861.bmp		0.587906	0.239334
1439229_.bmp		0.598271	0.446810
1486213_.bmp		0.709454	0.330450
1575268_.bmp		0.978770	0.880968
2010916_.bmp		0.917595	0.570852
2016658_.bmp		0.581734	0.280890
2018809_.bmp		0.552887	0.228608
2042822_.bmp		0.798444	0.319013
665322.bmp		0.542481	0.265948
914.bmp		0.382216	0.153948
967049.bmp		0.903827	0.781478
Média		0.734474	0.497874
Desvio Padrão		0.183945	0.233498

Tabela C.2: Combinação Concavidade e Convexidade e Cadeias de Freeman Top-100

MRV		Z=4		Z=5h		Z=5v		Z=7	
		R _n	P _n						
1037814.bmp		0.887973	0.632301	0.894298	0.636233	0.864783	0.707956	0.853987	0.706279
1055261.bmp		0.895412	0.758467	0.881093	0.737860	0.892606	0.788296	0.889609	0.773681
1138103.bmp		0.834763	0.492537	0.844807	0.489612	0.813550	0.469659	0.827350	0.486588
1138293_.bmp		0.879499	0.834451	0.924998	0.845820	0.882783	0.834715	0.906765	0.840419
1190540.bmp		0.954077	0.763264	0.979963	0.797570	0.948772	0.722330	0.935125	0.610182
1259886.bmp		0.641800	0.529884	0.656137	0.554712	0.688794	0.531041	0.653072	0.474183
1267206.bmp		0.790999	0.570005	0.776627	0.517421	0.803945	0.568235	0.811870	0.564399
1279931.bmp		0.855455	0.607661	0.856385	0.541726	0.872405	0.574566	0.883636	0.545653
1289047.bmp		0.857037	0.729240	0.870749	0.722766	0.838211	0.654610	0.817535	0.572539
1376861.bmp		0.724565	0.424615	0.770038	0.471698	0.758732	0.417833	0.826188	0.466553
1439229_.bmp		0.542606	0.360798	0.530927	0.342093	0.554613	0.346446	0.595543	0.388579
1486213_.bmp		0.785206	0.444396	0.777084	0.453203	0.799685	0.480827	0.770881	0.449560
1575268_.bmp		0.731991	0.507358	0.719600	0.507466	0.735185	0.513450	0.759936	0.553221
2010916_.bmp		0.896333	0.554710	0.921130	0.596968	0.927475	0.602421	0.913757	0.570244
2016658_.bmp		0.790542	0.453333	0.799806	0.449854	0.794303	0.465290	0.784296	0.467116
2018809_.bmp		0.512005	0.212658	0.537924	0.217836	0.545849	0.218372	0.543428	0.218809
2042822_.bmp		0.688690	0.271235	0.670691	0.256225	0.732760	0.289750	0.710417	0.299250
665322.bmp		0.828893	0.657901	0.806242	0.652288	0.834553	0.650425	0.855843	0.668644
914.bmp		0.925377	0.683694	0.925403	0.674907	0.909309	0.659072	0.908939	0.658254
967049.bmp		0.855039	0.794548	0.845145	0.765615	0.865442	0.829415	0.865199	0.831162
Média		0.793913	0.564153	0.799452	0.561594	0.803188	0.566235	0.805669	0.557266
Desvio Padrão		0.121369	0.172728	0.124198	0.172512	0.109968	0.171468	0.107840	0.162169

Tabela C.3: Combinação Concavidade/Convexidade e SIFT Top-100

MRV		Z=0		Z=4		Z=5h		Z=5v		Z=7	
		R _n	P _n								
1037814.bmp		0.914952	0.782868	0.917235	0.763298	0.920004	0.782143	0.920296	0.769704	0.906826	0.770031
1055261.bmp		0.894018	0.798900	0.908307	0.821783	0.907610	0.818163	0.897885	0.795614	0.877189	0.774266
1138103.bmp		0.728943	0.434806	0.806720	0.518165	0.813221	0.526172	0.808668	0.524948	0.807879	0.540002
1138293_.bmp		0.953959	0.858102	0.991105	0.899242	0.991608	0.900678	0.991076	0.899152	0.992861	0.904667
1190540.bmp		0.852026	0.691207	0.880288	0.642698	0.899329	0.659118	0.882240	0.641674	0.901667	0.649517
1259886.bmp		0.842568	0.538937	0.914237	0.694988	0.912618	0.693163	0.903557	0.687519	0.910319	0.705600
1267206.bmp		0.670943	0.495824	0.725524	0.528948	0.731594	0.489857	0.724485	0.499231	0.725193	0.517764
1279931.bmp		0.951353	0.716570	0.924954	0.677520	0.930315	0.705105	0.928517	0.682477	0.928548	0.682560
1289047.bmp		0.755370	0.552116	0.833778	0.505764	0.838692	0.506130	0.813182	0.496729	0.841782	0.535584
1376861.bmp		0.795248	0.462284	0.902690	0.585728	0.905792	0.582737	0.909127	0.589853	0.935297	0.603367
1439229_.bmp		0.672594	0.531621	0.505149	0.502882	0.484783	0.500077	0.509772	0.503548	0.504032	0.502718
1486213_.bmp		0.710884	0.355319	0.768531	0.417713	0.749241	0.368353	0.768728	0.416055	0.766872	0.423097
1575268_.bmp		0.657913	0.515212	0.674674	0.511931	0.676260	0.509902	0.681207	0.520298	0.675251	0.510972
2010916_.bmp		0.818753	0.497208	0.863465	0.556429	0.876722	0.572337	0.865569	0.558534	0.845407	0.559975
2016658_.bmp		0.662482	0.285848	0.708349	0.341123	0.740495	0.425943	0.734365	0.423295	0.767927	0.442389
2018809_.bmp		0.483486	0.218400	0.555926	0.307548	0.535046	0.248054	0.500681	0.232316	0.545732	0.242945
2042822_.bmp		0.759423	0.319025	0.706779	0.277395	0.695638	0.268478	0.701815	0.273928	0.702971	0.277963
665322.bmp		0.828597	0.591381	0.857430	0.696113	0.882556	0.742359	0.856523	0.694944	0.835677	0.684835
914.bmp		0.753640	0.527144	0.844718	0.574234	0.879383	0.585808	0.860670	0.596220	0.869398	0.600536
967049.bmp		0.889251	0.872501	0.867877	0.867787	0.864668	0.840654	0.869903	0.868145	0.866587	0.836235
Média		0.779820	0.552264	0.807887	0.584564	0.811779	0.586262	0.806413	0.583709	0.810371	0.588251
Desvio Padrão		0.118131	0.186981	0.12719	0.176048	0.134176	0.182424	0.131895	0.177539	0.127546	0.170697

Tabela C.4: Combinação Concavidade/Convexidade e M. Hu Z=0 Top-100

MRV		R_n	P_n
1037814.bmp		0.926256	0.659551
1055261.bmp		0.903885	0.820562
1138103.bmp		0.743630	0.409516
1138293_.bmp		0.945153	0.853873
1190540.bmp		0.829612	0.509904
1259886.bmp		0.857863	0.544208
1267206.bmp		0.675722	0.452194
1279931.bmp		0.930611	0.657134
1289047.bmp		0.713170	0.459683
1376861.bmp		0.809647	0.471761
1439229_.bmp		0.700662	0.537976
1486213_.bmp		0.679409	0.335362
1575268_.bmp		0.670622	0.521605
2010916_.bmp		0.804862	0.492473
2016658_.bmp		0.660334	0.290327
2018809_.bmp		0.469948	0.209673
2042822_.bmp		0.741517	0.310240
665322.bmp		0.833044	0.585728
914.bmp		0.746510	0.519722
967049.bmp		0.890928	0.872782
Média		0.776669	0.525714
Desvio Padrão		0.119144	0.180627