

FERNANDA RODRIGUES RAMOS

**RECUPERAÇÃO DE INFORMAÇÃO BASEADA EM
CONTEÚDO: ANALISANDO IMAGENS
PRIORIZANDO A CARACTERÍSTICA COR**

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

CURITIBA

2004

FERNANDA RODRIGUES RAMOS

**RECUPERAÇÃO DE INFORMAÇÃO BASEADA EM
CONTEÚDO: ANALISANDO IMAGENS
PRIORIZANDO A CARACTERÍSTICA COR**

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

Área de Concentração: Metodologia e Técnicas de Computação.

Orientador: Prof. Dr. Díbio Leandro Borges.

CURITIBA

2004

Ramos, Fernanda Rodrigues

Recuperação de Informação Baseada em Conteúdo: analisando imagens priorizando a característica cor. Curitiba, 2004. 124p.

Dissertação – Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática Aplicada.

1. Recuperação de Informação 2. Extração de Características 3. Wavelets

I. Pontifícia Universidade Católica do Paraná. Centro de Ciências Exatas e de Tecnologia. Programa de Pós-Graduação em Informática Aplicada II-t

Agradecimentos

Quero agradecer a Deus pela vida, pela força e benção nos momentos difíceis.

Agradeço e dedico este trabalho aos meus queridos pais Alceu e Maria, e aos meus irmãos Ana Paula, André, Fabiano e Thiago, que mesmo de longe me encobriam de carinho e força para continuar essa caminhada. Devo a vocês tudo o que consegui.

Ao meu orientador, Prof. Dr. Díbio Leandro Borges que é um grande amigo, agradeço pela paciência, persistência e dedicação a mim oferecida. Foi mérito dele o conhecimento adquirido durante a realização deste trabalho.

Ao Prof. Dr. Alceu de Souza Britto pela ajuda com seu conhecimento e principalmente pelas horas gastas co-orientando e enriquecendo o trabalho, muito obrigada.

Não poderia deixar de agradecer as pessoas que acompanharam de perto esse sofrível, mas gratificante esforço: Aline, Alexandre, Cláudia, Cristiane, Gisele e William. Obrigada pelas palavras de carinho e grande ajuda a mim oferecida (que não foram poucas!!).

Claro que não esquecerei dos guris do LUCI: David, Carlos Magno, Éderson, Fausto, Islenho e Paulo. Obrigada pelas risadas e momentos de descontração.

Renata e Sirlene agradeço pela amizade adquirida durante o mestrado.

Agradeço as amigas de longe, mas que sempre estavam por perto: Fabiane e Juliane. Obrigada pelos e-mail's e pela amizade que fortaleceu com a distância.

Ao órgão financiador, CAPES, agradeço pela oportunidade da bolsa de estudos e a PUCPR pelo espaço físico por poder desenvolver o mestrado.

Ainda quero agradecer as amigas que foram surgindo durante a caminhada na PUCPR.

“A vida não é um corredor reto e tranqüilo que nós percorremos livres e sem empecilhos, mas um labirinto de passagens, pelas quais nós devemos procurar nosso caminho, perdidos e confusos, de vez em quando presos em um beco sem saída.

Porém, se tivermos fé, uma porta sempre será aberta para nós, não talvez aquela sobre a qual nós mesmos nunca pensamos, mas aquela que definitivamente se revelará boa para nós”.

A. J. Cronin

Sumário

Sumário	i
Lista de Figuras	iii
Lista de Tabelas	vii
Lista de Gráficos	viii
Lista de Símbolos	x
Lista de Abreviaturas	xi
Resumo	xii
Abstract	xiii
Capítulo 1	
Introdução	1
1.1 Motivação.....	2
1.2 Objetivo Geral e Específico.....	3
1.3 Organização do trabalho.....	4
Capítulo 2	
Trabalhos Relacionados	5
Capítulo 3	
Método Proposto	11
3.1 Imagem.....	13
3.2 Base de Imagens.....	14
3.3 Atributos Visuais.....	15
3.3.1 Atributo cor.....	15
3.3.2 Atributo textura.....	16
3.4 Modelos de Cor.....	17
3.4.1 Modelo RGB.....	18
3.4.2 Modelo HSV.....	19
3.4.3 Conversão de RGB para HSV.....	20
3.5 Distribuição Espacial.....	22
3.6 Decomposição <i>Wavelet</i>	23
3.7 Interseção.....	25
3.8 Distância Euclidiana.....	25
3.9 Somatório.....	26
3.10 Desempenho da Efetividade.....	26
3.11 Resultados do Desempenho.....	27
3.12 Descrição da abordagem proposta.....	30
3.12.1 Descrição dos passos.....	31
3.13 Matriz de Confusão.....	35

Capítulo 4	
Experimentos	38
4.1 Resultados da Matriz de Confusão.....	42
4.2 Análise da Combinação dos Resultados.....	65
4.3 Considerações Finais.....	71
Capítulo 5	
Conclusões	72
5.1 Trabalhos Futuros.....	73
Referências Bibliográficas	75
Apêndice A	78

Lista de Figuras

Figura 1.1	Esquematização de um sistema de recuperação de informação.....	1
Figura 3.1	Diagrama do método proposto.....	12
Figura 3.2	Imagem em nível de cinza e seu histograma.....	13
Figura 3.3	Exemplos de quatro diferentes elementos de textura (Bueno 2001).....	16
Figura 3.4	Espaço de cor RGB (Gonzalez & Woods 1992).....	18
Figura 3.5	Espaço de cor HSV (Gonzalez & Woods 1992).....	19
Figura 3.6	Canal S de uma imagem.....	21
Figura 3.7	Exemplos de figuras com mesma proporção de cor, mas diferente distribuição espacial (Schettini <i>et al</i> 2001).....	22
Figura 3.8	Diagrama ilustrando o processo de decomposição.....	23
Figura 3.9	Representação da decomposição <i>wavelet</i> em uma imagem.....	24
Figura 3.10	Imagem inicial de teste.....	31
Figura 3.11	Imagens dos canais convertidos H, S e V.....	32
Figura 3.12	Imagem dividida em nove partes.....	32
Figura 3.13	Histogramas do canal H (a), canal S (b) e canal V (c), baseados na imagem teste.....	34
Figura 4.1	Diagrama dos passos que são realizados para a execução dos experimentos.....	39
Figura 4.2	Resultados com a classificação.....	40
Figura 4.3	Imagens que representam as 27 classes usadas nos experimentos.....	43
Figura 4.4	Imagens da base de dados. (a) Imagens referentes a classe 5 (árvore) e (b) imagens da classe 19 (janela).....	44
Figura 4.5	Imagens da base de dados. (a) Imagens referentes a classe 10 (construção) e (b) imagens da classe 3 (animais).....	46
Figura 4.6	Imagens da base de dados. (a) Imagens referentes a classe 5 (árvore) e (b) imagens da classe 21 (mamíferos).....	46
Figura 4.7	Imagens da base de dados referentes a classe 22 (mosaico).....	47
Figura 4.8	Imagens da base de dados. (a) Imagens referentes a classe 16 (fruta) e (b) imagens da classe 17 (garotas).....	49
Figura 4.9	Imagens da base de dados. (a) Imagens referentes a classe 21 (mamíferos) e (b) imagens da classe 5 (árvores).....	49
Figura 4.10	Imagens da base de dados. (a) Imagens referentes a classe 16 (frutas) e (b) imagens da classe 5 (árvores).....	51
Figura 4.11	Imagens da base de dados. (a) Imagens referentes a classe 10 (construção) e (b) imagens da classe 26 (neve).....	51

Figura 4.12	Imagens da base de dados. (a) Imagens referentes a classe 22 (mosaico) e (b) imagens da classe 5 (árvore).....	54
Figura 4.13	Imagens da base de dados. (a) Imagens referentes a classe 24 (ponte) e (b) imagens da classe 26 (neve).....	54
Figura 4.14	Imagens da base de dados referentes a classe 1 (água).....	55
Figura 4.15	Imagens da base de dados. (a) Imagens referentes a classe 25 (céu) e (b) Imagens da classe 27 (textura).....	57
Figura 4.16	Imagens da base de dados referentes a classe 11 (amanhecer).....	57
Figura 4.17	Imagens da base de dados. (a) Imagens referentes a classe 22 (mosaico) e (b) Imagens da classe 5 (árvore).....	58
Figura 4.18	Imagens da base de dados. (a) Imagens referentes a classe 23 (montanha) e (b) Imagens da classe 26 (neve).....	60
Figura 4.19	Imagens da base de dados referentes a classe 10 (construção).....	60
Figura 4.20	Imagens da base de dados. (a) Imagens referentes a classe 25 (céu) e (b) Imagens da classe 11 (amanhecer).....	62
Figura 4.21	Imagens da base de dados. (a) Imagens referentes a classe 25 (céu) e (b) Imagens da classe 26 (neve).....	64
Figura 4.22	Imagens da base de dados. (a) Imagens referentes a classe 27 (textura) e (b) Imagens da classe 11 (amanhecer).....	64
Figura 4.23	Imagens da base de dados. (a) Imagens referentes a classe 9 (concreto) e (b) Imagens da classe 22 (mosaico).....	66
Figura 4.24	Imagens da base de dados referentes a classe 26 (neve).....	66
Figura A.1	Classe 1 referente as imagens Água – Imagens de 1 a 27.....	78
Figura A.2	Classe 1 referente as imagens Água – Imagens de 28 a 57.....	79
Figura A.3	Classe 1 referente as imagens Água – Imagens de 58 a 70.....	80
Figura A.4	Classe 2 referente as imagens Ar – Imagens de 1 a 30.....	81
Figura A.5	Classe 2 referente as imagens Ar – Imagens de 31 a 53.....	82
Figura A.6	Classe 3 referente as imagens Animais – Imagens de 1 a 30.....	83
Figura A.7	Classe 3 referente as imagens Animais – Imagens de 31 a 55.....	84
Figura A.8	Classe 4 referente as imagens Arame – Imagens de 1 a 6.....	85
Figura A.9	Classe 5 referente as imagens Árvore – Imagens de 1 a 30.....	86
Figura A.10	Classe 5 referente as imagens Árvore – Imagens de 31 a 60.....	87
Figura A.11	Classe 5 referente as imagens Árvore – Imagens de 61 a 90.....	88
Figura A.12	Classe 5 referente as imagens Árvore – Imagens de 91 a 103.....	89
Figura A.13	Classe 6 referente as imagens Caixa Postal – Imagens de 1 a 15.....	90
Figura A.14	Classe 7 referente as imagens Carro – Imagens de 1 a 30.....	91
Figura A.15	Classe 7 referente as imagens Carro – Imagens de 31 a 42.....	92
Figura A.16	Classe 8 referente as imagens Natal – Imagens de 1 a 30.....	93
Figura A.17	Classe 8 referente as imagens Natal – Imagens de 31 a 42.....	94
Figura A.18	Classe 9 referente as imagens Concreto – Imagens de 1 a 14.....	95
Figura A.19	Classe 10 referente as imagens Construção – Imagens de 1 a 30.....	96
Figura A.20	Classe 10 referente as imagens Construção – Imagens de 31 a 60.....	97
Figura A.21	Classe 10 referente as imagens Construção – Imagens de 61 a 79.....	98
Figura A.22	Classe 11 referente as imagens Amanhecer – Imagens de 1 a 30.....	99
Figura A.23	Classe 11 referente as imagens Amanhecer – Imagens de 31 a 60.....	100
Figura A.24	Classe 11 referente as imagens Amanhecer – Imagens de 61 a 63.....	101
Figura A.25	Classe 12 referente as imagens Pato – Imagens de 1 a 30.....	102
Figura A.26	Classe 12 referente as imagens Pato – Imagens de 31 a 37.....	103
Figura A.27	Classe 13 referente as imagens Bandeira – Imagens de 1 a 30.....	104

Figura A.28	Classe 13 referente as imagens Bandeira – Imagens de 31 a 43.....	105
Figura A.29	Classe 14 referente as imagens Flores – Imagens de 1 a 30.....	106
Figura A.30	Classe 14 referente as imagens Flores – Imagens de 31 a 43.....	107
Figura A.31	Classe 15 referente as imagens Futebol – Imagens de 1 a 30.....	108
Figura A.32	Classe 15 referente as imagens Futebol – Imagens de 31 a 44.....	109
Figura A.33	Classe 16 referente as imagens Frutas – Imagens de 1 a 19.....	110
Figura A.34	Classe 17 referente as imagens Garotas – Imagens de 1 a 27.....	111
Figura A.35	Classe 18 referente as imagens Árvore-Rosa – Imagens de 1 a 30....	112
Figura A.36	Classe 18 referente as imagens Árvore-Rosa – Imagens de 31 a 55..	113
Figura A.37	Classe 19 referente as imagens Janela – Imagens de 1 a 12.....	114
Figura A.38	Classe 20 referente as imagens Folha – Imagens de 1 a 30.....	115
Figura A.39	Classe 21 referente as imagens Mamíferos – Imagens de 1 a 16.....	116
Figura A.40	Classe 22 referente as imagens Mosaico – Imagens de 1 a 25.....	117
Figura A.41	Classe 23 referente as imagens Montanhas – Imagens de 1 a 30.....	118
Figura A.42	Classe 24 referente as imagens Ponte – Imagens de 1 a 8.....	119
Figura A.43	Classe 25 referente as imagens Céu – Imagens de 1 a 30.....	120
Figura A.44	Classe 25 referente as imagens Céu – Imagens de 31 a 43.....	121
Figura A.45	Classe 26 referente as imagens Neve – Imagens de 1 a 14.....	122
Figura A.46	Classe 27 referente as imagens Textura – Imagens de 1 a 30.....	123
Figura A.47	Classe 27 referente as imagens Textura – Imagens de 31 a 32.....	124

Lista de Tabelas

Tabela 3.1	Classes e suas respectivas quantidades de imagens.....	14
Tabela 3.2	Exemplo de arquivo gerado pelo classificador.....	29
Tabela 3.3	Exemplo de arquivo com o cálculo das medidas.....	30
Tabela 3.4	Matriz de confusão usada como ilustração.....	37

Lista de Gráficos

Gráfico 4.1	Gráficos dos resultados do <i>ranking</i> de classificação e porcentagem de acerto do canal H, coeficientes de aproximação. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.....	43
Gráfico 4.2	Gráficos dos resultados do <i>ranking</i> de classificação e porcentagem de acerto do canal S, coeficientes de aproximação. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.....	45
Gráfico 4.3	Gráficos dos resultados do <i>ranking</i> de classificação e porcentagem de acerto do canal V, coeficientes de aproximação. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.....	48
Gráfico 4.4	Gráficos dos resultados do <i>ranking</i> de classificação e porcentagem de acerto do canal H, coeficientes de detalhe horizontal. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.....	50
Gráfico 4.5	Gráficos dos resultados do <i>ranking</i> de classificação e porcentagem de acerto do canal S, coeficientes de detalhe horizontal. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.....	53
Gráfico 4.6	Gráficos dos resultados do <i>ranking</i> de classificação e porcentagem de acerto do canal V, coeficientes de detalhe horizontal. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.....	56
Gráfico 4.7	Gráficos dos resultados do <i>ranking</i> de classificação e porcentagem de acerto do canal H, coeficientes de detalhe vertical. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.....	59
Gráfico 4.8	Gráficos dos resultados do <i>ranking</i> de classificação e porcentagem de acerto do canal S, coeficientes de detalhe vertical. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.....	61
v		
Gráfico 4.9	Gráficos dos resultados do <i>ranking</i> de classificação e porcentagem de acerto do canal V, coeficientes de detalhe vertical. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.....	63
Gráfico 4.10	Gráfico com a combinação dos canais H, S e V com os coeficientes de aproximação e de detalhe vertical; (a) Gráfico do <i>ranking</i> de classificação; (b) Gráfico da porcentagem de acerto.....	67

Gráfico 4.11	Gráfico com os Coeficientes de Aproximação Canal V – Somatório.....	68
Gráfico 4.12	Gráfico com os Coeficientes Detalhe Horizontal Canal V – Somatório.....	69
Gráfico 4.13	Gráfico com os Coeficientes Detalhe Vertical Canal V – Somatório.....	70

Lista de Símbolos

$A_{2^{j+1}}^d f$	Sinal de entrada da decomposição <i>wavelet</i>
$A_{2^j}^d f$	Dados suavizados ou de baixa frequência, coeficientes de aproximação
$D^1_{2^j} f$	Dados de detalhe de alta frequência, vertical
$D^2_{2^j} f$	Dados de detalhe de alta frequência, horizontal
$D^3_{2^j} f$	Dados de detalhe de alta frequência, diagonal
S	Número de imagens recuperadas
$R_{i,q}$	Conjunto de imagens relevantes na base de dados
$R_{e,q}$	Conjunto de imagens recuperadas

Lista de Abreviaturas

CBIRS	<i>Content-Based Image Retrieval Systems</i>
G	Filtros passa-alta
H	Filtros passa-baixa
HSV	<i>Hue, Saturation, Value</i>
MCAG	<i>Modified Color Adjacency Graph</i>
QMF's	<i>Quadrature Mirror Filters</i>
RIC	Recuperar de Informação por Conteúdo
RGB	<i>Red, Green, Blue</i>
SVG	<i>Spatial Variance Graph</i>

Resumo

A pesquisa em recuperação de informação tem seu ponto chave na representação de características intrínsecas presentes em uma imagem. Isto tem sido um grande estímulo para pesquisas e para o desenvolvimento de ferramentas que armazenam, recuperam e representam as imagens de forma correta. Levando em consideração que a recuperação é feita por uma busca de imagens semelhantes a uma imagem de referência por meio de medidas de similaridade, o trabalho investiga uma nova forma de indexação de imagens baseada em cor e textura, na qual esses atributos aparecem combinados em um espaço de representação *wavelet*. É desenvolvido um método que soluciona a extração de cor e textura considerando a multirresolução nas imagens, e ainda uma função de similaridade para realizar o cálculo da semelhança entre a imagem teste e as imagens de uma base de dados.

Palavras-chave: 1. Recuperação de Informação. 2. Atributos cor e textura. 3. *Wavelet*. 4. Extração de Características.

Abstract

The research in information retrieval has its essential point in the representation of intrinsic features of an image. This has been a great incentive for researching and for developing of tools that store, retrieve and represent images of correct manner. Considering that the retrieval is done by a search of similar images to a reference image by means of measures of similarity, this work investigates a new manner to image indexing based in color and texture in which these attributes are combined in a wavelet representation framework. A method is developed to solve the extraction of color and texture considering multi-resolution in images and a similarity function to realize the calculation of similarity between a test image and the images of a database.

Keywords: 1. Information retrieval. 2. Color and texture attributes. 3. Wavelet. 4. Feature extraction.

Capítulo 1

Introdução

A pesquisa em recuperação de informação tem seu ponto chave na representação de características intrínsecas presentes em uma imagem. Isto tem sido um grande estímulo para pesquisas e para o desenvolvimento de ferramentas que armazenam, recuperam e representam as imagens de forma correta.

Recuperar informação por conteúdo (RIC) consiste na tarefa de buscar imagens não somente iguais, mas aquelas que sejam semelhantes entre si. Parte do princípio de uma imagem inicial e buscar em uma base de dados as imagens que contenham algum tipo de similaridade, como mostra a Figura 1.1.

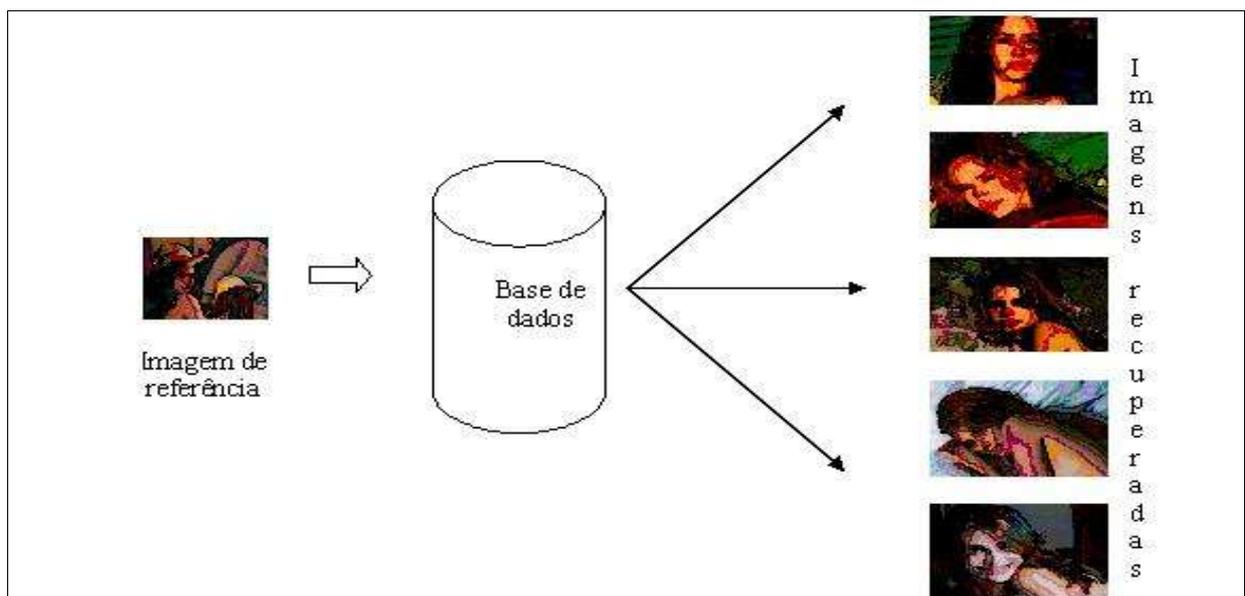


Figura 1.1 – Esquematização de um sistema de recuperação de informação.

Mas, como definir a similaridade na recuperação de informação? A recuperação de imagens é feita a partir de informações pictoriais. Isto é possível através da busca de imagens semelhantes a uma imagem de referência por meio de medidas de similaridade (GUDIVADA e RAGHAVAN 1995). Para isso, são desenvolvidos métodos que calculam a diferença entre as imagens que estão sendo comparadas. Atualmente, esses métodos utilizam cor, textura e forma como atributos de indexação, os quais são extraídos da imagem de maneira independente.

Na literatura estudada, as tarefas de recuperação de imagem podem ser aplicadas a técnicas de desenho na engenharia e arquitetura; em sistemas de informação geográfica; em base de imagens médicas; em sistemas para reconhecimento de face; em ferramentas de busca na internet; dentre outras.

Nesse contexto, muitos esforços têm sido direcionados em pesquisas e desenvolvimento dos chamados sistemas de recuperação de imagem baseados em conteúdo (*Content-Based Image Retrieval Systems – CBIRS*). Aplicados às imagens, esses sistemas devem organizar as informações relevantes, e permitir a recuperação de imagens de forma eficiente e conforme a necessidade do usuário.

O fato de priorizar a informação com um determinado atributo é um tópico complexo. São feitas muitas tentativas para capturar a similaridade perceptiva, pois diferentes pessoas possuem percepções variadas para um mesmo conteúdo.

Neste trabalho investiga-se uma nova forma de indexação de imagens baseada em cor e textura, na qual esses atributos aparecem combinados em um espaço de representação *wavelet*. A cor é uma característica que freqüentemente simplifica a identificação de objetos e a extração de uma cena, pois é um atributo que está presente em tudo o que é visível ao olho humano. A textura pode ser utilizada em combinação com a característica cor. Isso é possível pelo fato de estarem correlacionadas em níveis de multirresoluções diferentes.

1.1. Motivação

De acordo com (SCHETTINI *et al* 2001), em grande parte da literatura relacionada com a recuperação de informação por conteúdo, os métodos existentes não abordam o processo de extração baseado na combinação de cor e textura, e também não trabalham com a correlação existentes sobre os mesmos. Estes métodos abordam apenas as características cor, textura e forma em separado.

Os trabalhos relacionados nessa área buscam a identificação da similaridade através de histogramas e métricas de similaridade, não abordando a distribuição espacial, interseção de histogramas, combinação de resultados e a multirresolução das imagens. Levando em consideração esses itens, a proposta deste trabalho está relacionada com um conjunto de características que consigam manter a informação necessária à classificação das imagens, não precisando trabalhar com toda a informação presente nas mesmas, mas com parte dela.

Pretende-se, então, abordar a extração de características (cor e textura) como parâmetros correlacionados através de diferentes escalas. Logo, a contribuição deste trabalho está em oferecer um método que solucione a extração de cor e textura levando-se em consideração a multirresolução nas imagens, e ainda uma função de similaridade para realizar o cálculo da semelhança entre a imagem teste e as imagens da base de dados.

1.2. Objetivo Geral e Específico

O objetivo direto deste trabalho é apresentar um método de recuperação de imagens baseado nos atributos cor e textura, combinados em um espaço *wavelet*. Isso pode ser feito com a inclusão de técnicas de recuperação de informação que permitem tratar as imagens de forma natural.

Para que o sistema de recuperação possa manipular as imagens, o presente trabalho propõe correlacionar as características cor e textura em uma imagem. Representar a cor das imagens a serem indexadas em um modelo de cor HSV. Caracterizar os canais H, S, V no espaço *wavelet* utilizando a base Haar. Representar uma forma de associar a informação espacial ao atributo cor e textura. Analisar a recuperação de imagens em diferentes níveis de decomposição *wavelet* e analisar o comportamento das métricas distância euclidiana e somatório na fase de recuperação das imagens. A recuperação por similaridade é apresentada, como no exemplo: “Dada uma imagem X, quais são as imagens presentes no banco de dados

mais parecidas com ela”. Este é um exemplo de consulta que é respondida pelo método proposto através da manipulação das imagens.

1.3. Organização do trabalho

Este trabalho está organizado da seguinte forma. Após essa breve introdução, bem como a motivação para o desenvolvimento e os objetivos apresentados no Capítulo 1, no Capítulo 2 é apresentado um levantamento sobre os sistemas de RIC. São, também, enfatizados os pontos principais de tais sistemas. O Capítulo 3 descreve o método de recuperação de informação que está sendo desenvolvido, bem como sua arquitetura geral. As principais contribuições, que consistem da interseção de histogramas, e da distribuição espacial desenvolvida, são aqui apresentadas. O Capítulo 4 descreve os resultados dos experimentos realizados sobre a recuperação das imagens com a interseção dos histogramas e as métricas usadas. As conclusões gerais e linhas de futuras pesquisas são delineadas no Capítulo 5 deste trabalho.

Capítulo 2

Trabalhos Relacionados

Neste capítulo são apresentados trabalhos que se relacionam com o tema proposto, a recuperação de informação por conteúdo.

Em (OLIVEIRA *et al* 2002) foi proposto o protótipo de um sistema de recuperação de imagens com base na cor. Este armazena informações de regiões com a mesma cor na imagem (regiões cromáticas), como o tamanho, posição e limites com outras regiões. A recuperação da imagem é feita através de uma medida de similaridade baseada na comparação dos grafos *Modified Color Adjacency Graph* (MCAG) e *Spatial Variance Graph* (SVG). Esse protótipo permite realizar dois tipos de consultas: o usuário entra com uma imagem de exemplo para consulta ou o usuário define as cores de uma imagem de exemplo para consulta. Segundo os autores, os resultados utilizando esses grafos mostram-se robustos e o protótipo modelado mostrou-se viável, não apresentando resultados quantificados.

Melhorias podem ser feitas nesse protótipo como, por exemplo, a utilização do modelo de cor HSV. A utilização deste modelo visa diminuir o número de cores, isso é relatado em muitos trabalhos como um modelo que melhor enfatiza a recuperação de imagens (MATHIAS e CONCI 1998). E ainda analisar a possibilidade de reconhecimento de objetos usando grafos globais e locais, ou múltiplas representações de grafos, que podem empregar outros atributos, tais como, textura e forma.

Em (VRIES *et al* 2001) foi proposto um algoritmo de busca por similaridade em um sistema com uma base de dados de imagens coloridas do *Software Corel Draw* (COREL

2004). As características das imagens são extraídas usando o modelo de cor HSV. O método utilizado nessa abordagem armazena as informações em um vetor de características. Uma tabela é mantida em separado para cada dimensão de características desse vetor, contendo as informações relevantes de todas as imagens. A similaridade entre o vetor de consulta (*query*) que está sendo testada e cada um dos vetores de características das imagens que são comparadas é acumulada fazendo o mapeamento dessas tabelas. Esse método armazena a característica do vetor, e calcula a semelhança de todas as imagens dimensão por dimensão, este esquema é apresentado para duas métricas de similaridade, interseção de histograma e distância euclidiana. Os resultados experimentais são apresentados para ilustrar o desempenho do algoritmo na busca da similaridade de imagens, não apresentando dados quantitativos.

Este trabalho utiliza apenas o atributo cor para fazer a busca da similaridade entre as imagens. O método de armazenar as características extraídas em um vetor pode ser melhorado através da utilização da distribuição espacial e, ainda, usar um método de extração de características, como por exemplo, a decomposição *wavelet*. O uso do modelo de cor HSV enfatiza a relevância desse modelo em recuperação de imagens coloridas.

Das e Manmatha (2001) apresentam um trabalho para indexar imagens em uma base de dados de pássaros usando a cor para calcular as regiões de interesse. O problema principal nessa tarefa é a segmentação da região de interesse com o fundo, ou seja, eliminar o fundo e deixar a figura ou o objeto que interessa. Para encontrar o objeto colorido são usadas as bordas e o fundo, as regiões que não interessam são eliminadas. A indexação da base de pássaros é feita através dos histogramas das regiões de interesse determinados pela cor e pelas bordas do objeto. As imagens estão no espaço de cor RGB. Os resultados apresentados mostram que em 87% de 450 imagens de pássaros testadas, a segmentação é suficiente para determinar a recuperação dos pássaros corretamente. Esta técnica pode ser aplicada para outros tipos de base de dados de animais.

No trabalho de Das e Manmatha (2001) para solucionar o problema da segmentação a imagem pode ser dividida em partes, assim a região de interesse pode ser encontrada através da distribuição espacial. O modelo de cor HSV pode ser usado para destacar a cor nas regiões de interesse.

No trabalho de (STRICKER e DIMAI 1997 *apud* SCHETTINI *et al* 2001) foi proposto o método para avaliar e combinar a característica cor para buscar a similaridade em imagens. A imagem foi dividida em quatro partes, e para enfatizar a cena é atribuído um peso maior a parte central da região. Stricker (1992) usa histogramas para codificar as cores semelhantes, e computar a informação geométrica em indexação da imagem colorida. Esse método produz um espaço de cores distintas de 256 elementos, e não é robusto o suficiente para imagens relacionadas com textura.

Pode ser usada a interseção de histogramas para atribuir um significado a um conjunto de características e comparar as informações extraídas por uma função de similaridade.

Foi descrito por (FUNT e FINLAYSON 1995) um método para derivar descritores independentes da iluminação da cor, partindo da suposição da instância local da iluminação. O histograma derivado do logaritmo de uma imagem codifica o tamanho da borda entre cores, fornecendo uma descrição independente da iluminação da imagem. Em uma publicação subsequente (FINLAYSON *et al* 1998), propõe-se o uso de indexação angular colorida para descrever imagens semelhantes. No método de indexação angular, os objetos são representados por três ângulos de cores. Os autores adicionaram três ângulos de textura de cores, calculados após uma filtragem linear da imagem.

Em (GEVERS e SMEULDERS 1996 *apud* SCHETTINI *et al* 2001) foi apresentado um conjunto de modelos sobre cor e que registram a invariância sob condições de mudança de imagens, incluindo a presença de sombras e reflexos. Os testes foram feitos em uma base de dados de 500 imagens. Um método sobre o modelo de cor RGB proposto supõe a reflexão dicromática e a iluminação de cor branca é independente do ponto de vista, orientação da superfície, direção da iluminação e intensidade. Esse trabalho não apresenta resultados quantitativos.

Para registrar a invariância nas imagens pode ser usada a distribuição espacial, como está sendo usada nesta dissertação. Não influenciando as condições de mudança nas imagens, como sombras e reflexos.

A proposta de (GALDINO 2000) foi construir um modelo de atenção visual que utiliza a informação de cor para detectar e acompanhar regiões de cores em uma seqüência de

imagens. Na primeira fase do modelo, a informação de cor utilizada em uma imagem é extraída identificando quais são as principais regiões coloridas e monitorando as novas regiões que não existiam na imagem utilizada. Na segunda fase do modelo, resolve-se o problema da detecção espacial das regiões na imagem. Localizam-se as regiões através de uma extensão de correlograma de cores, onde foram incorporadas faixas de cores extraídas da divisão de cores RGB em 16 subcubos. Os resultados mostrados no trabalho de (GALDINO e BORGES 2000) demonstram que acompanhar objetos ou regiões pode ser feito usando cor como característica relevante. O modelo mostrou-se preciso na localização espacial das regiões, devido a faixa de cores em que foi dividido o espectro. Os resultados não foram apresentados de forma quantitativa.

Em (IQBAL e AGGARWAL 2002) foi avaliado o desempenho da combinação das formas, cor e textura para recuperação de imagem. A forma é extraída pela aplicação de agrupamentos de percepção, tais como, segmentos de linha, linhas longas, junções “L”, junções “U”, linhas paralelas e polígonos. A análise de cor é feita pelo mapeamento de todos os *pixels* de uma imagem em uma palheta de cores fixas que usa caracteres lingüísticos, que identifica uma cor com um nome, para descrever o conteúdo cor. A análise da textura é feita usando um banco de filtros de Gabor, que são freqüentemente usados como uma medida de similaridade baseado na textura em recuperação de imagem. Uma metodologia para avaliação do desempenho dessas análises é apresentada em uma base de dados de imagens coloridas. A classificação é apresentada para as classes separadamente, os resultados mostram que a classe mamíferos apresentou melhores resultados somente com o atributo forma, resultando em 70,34% de classificação.

No trabalho de (BRAMBILLA *et al* 1999) é usada a transformada de multirresolução *wavelet*, base Haar, num espaço de cor CIELUV para processar assinaturas de imagens com uso em aplicações de recuperação de informação baseado em conteúdo, ou seja, busca por similaridade. Os coeficientes usados são os de detalhes horizontal e vertical, fazendo a decomposição no 1º nível. As características das imagens são codificadas em assinaturas (características salientes), que são comparadas na fase de recuperação aplicando uma medida de similaridade. O desempenho da classificação da recuperação é avaliado pela *class precision*, que mede o grau de correspondência entre a resposta do sistema e avaliação do observador.

A análise do trabalho de (BRAMBILLA *et al* 1999) com os dados contendo a maior parte de características de bordas mostrou que a base Haar forneceu resultados bons com os coeficientes de detalhe horizontal e vertical. Ainda seria preciso melhorar os resultados em relação a assinatura de imagens, por não recuperar corretamente o conteúdo da imagem. Os resultados obtidos chegaram a 70% de classificação, com as 20 primeiras imagens.

Percebe-se que o método proposto por (BRAMBILLA *et al* 1999) pode ser melhorado usando a informação espacial, isto é, dividir a imagem em sub-imagens, indexar e trabalhar com cada parte da divisão em separado. A decomposição feita através da transformada *wavelet* pode ser testada com a multirresolução em mais de um nível, gerando assim resultados diferenciados com escalas diferentes.

Em Long e Leow (2000), foi abordado um trabalho usando texturas com escalas, orientações, intensidade e contraste. Para mostrar a aplicação desse espaço perceptual, características da textura são extraídas usando filtros de Gabor e uma rede neuronal *feedforward multi-layer* é treinada para mapear as características de Gabor nesse espaço perceptual. Os resultados dos testes mostraram que a rede neuronal pode realizar o mapeamento precisamente. Além disso, testes de recuperação mostram que o desempenho é melhorado pelo mapeamento computacional das características de textura. Os resultados foram apresentados através do *ranking* de classificação das imagens similares com textura.

Conforme foi possível observar, os trabalhos apresentados relacionam-se com os métodos CBIRS, explorando o atributo cor para caracterizar o conteúdo da imagem. O trabalho proposto aqui é relacionado com o modelo de (BRAMBILLA *et al* 1999) que se baseia na estratégia de recuperar um conjunto de imagens de uma base de dados que satisfaça o critério da similaridade.

Vale ressaltar que neste trabalho estão sendo usadas tarefas que diferenciam dos demais trabalhos, tais como a divisão da imagem em subimagens para se conseguir a informação espacial da mesma, as métricas distância euclidiana e somatório contribuem para a classificação das imagens e ainda está sendo usada a multirresolução nas imagens, isso é um fato que caracteriza a relação da textura com a característica cor, isto é, a textura está implícita em diferentes níveis de escalas usadas pela decomposição *wavelet*.

Nesse capítulo foi apresentada uma revisão bibliográfica sobre os trabalhos relacionados com recuperação de informação por conteúdo. No capítulo seguinte é proposto o método para recuperação de informação por conteúdo utilizando os canais da transformada *wavelet* para gerar as características significativas para a extração. Os passos principais que compõem o algoritmo de recuperação proposto também serão descritos no próximo capítulo, bem como a transformada e os modelos de cor.

Capítulo 3

Método Proposto

O método proposto consiste na recuperação de informação baseada em conteúdo através da extração das características cor e textura em espaço *wavelet*. O objetivo do espaço *wavelet* é obter as informações que sejam significativas para representar as características cor e textura.

A abordagem proposta segue o esquema apresentado na Figura 3.1.

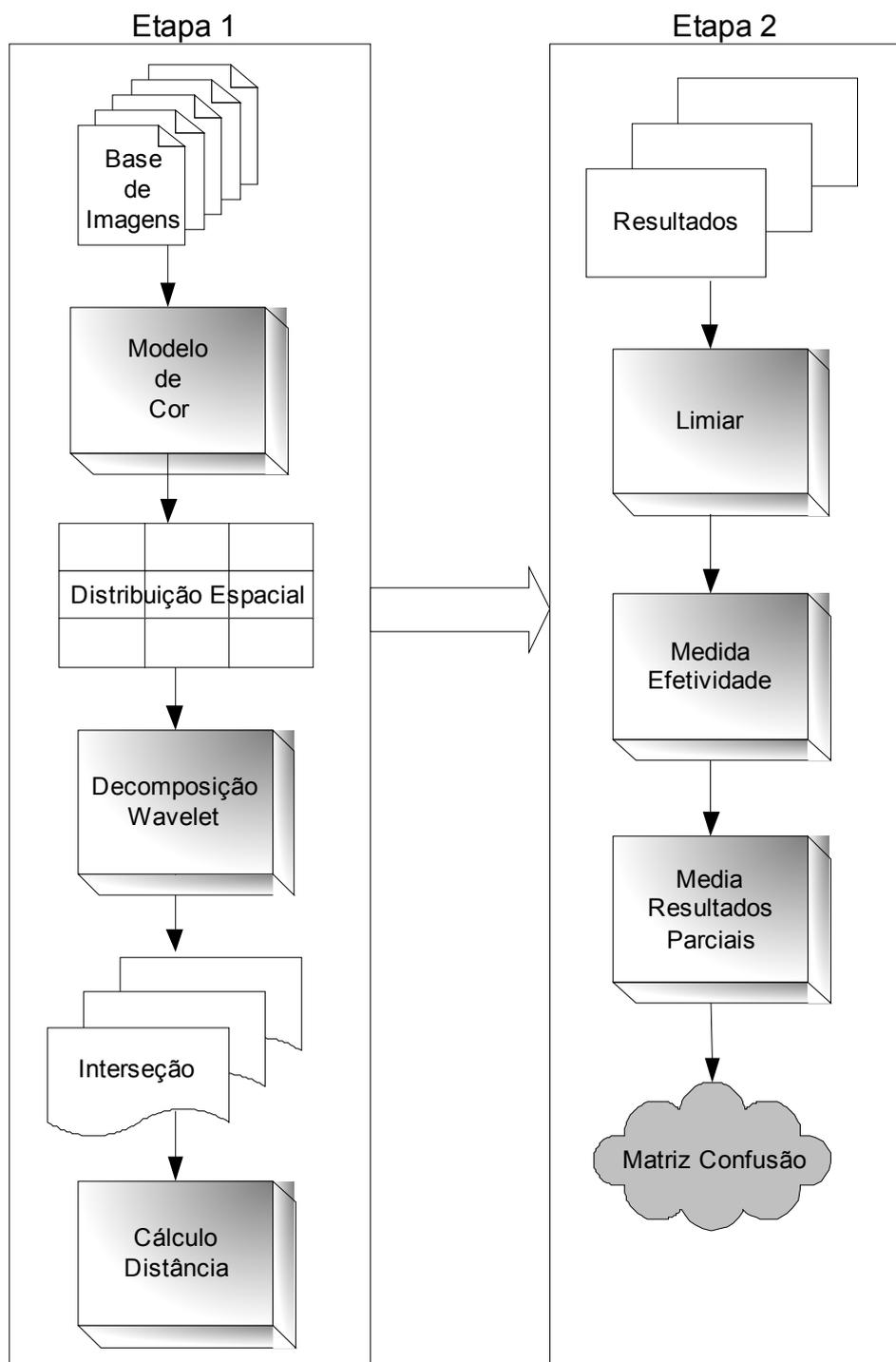


Figura 3.1 – Diagrama do método proposto.

3.1. Imagem

Uma imagem é descrita pela função $f(x,y)$ com intensidade de luz sobre a cena, sendo seu valor, em qualquer ponto de coordenadas espaciais (x,y) , proporcional ao brilho da imagem naquele ponto. Em uma imagem cujas informações são apresentadas em intervalos ou bandas distintas de frequência, é necessário uma função $f(x,y)$ para cada banda. É o caso de imagens coloridas padrão RGB, que são formadas pela informação de cores primárias, como o vermelho (*Red*), verde (*Green*) e azul (*Blue*). E as imagens no formato HSV que são formadas pela informação da matiz (*Hue*), saturação (*Saturation*) e intensidade de luz (*Value*). Para o processamento de uma imagem é fundamental representar sua informação num formato adequado ao tratamento computacional. Uma imagem pode ser representada por uma matriz, em que os índices de linha e coluna referenciam o brilho médio amostrado no ponto correspondente da cena (FACON 1996).

A visualização da distribuição das informações de uma imagem é feita por um histograma que dá a distribuição dos dados graficamente, veja exemplo na Figura 3.2.

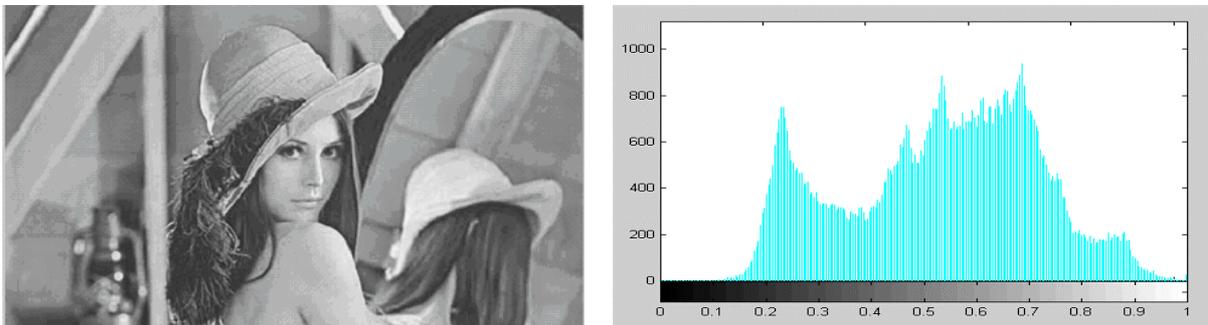


Figura 3.2 – Imagem em nível de cinza e seu histograma

O histograma de uma imagem corresponde a uma tabela que dá para cada nível de cinza, o número de *pixels* correspondentes na imagem. Quando o histograma fornece somente o número de *pixels* e não a localização desses, ele permite dar uma descrição global da imagem.

3.2. Base de Imagens

As imagens utilizadas neste trabalho foram obtidas de repositórios que possuem uma interessante base de dados para se trabalhar com recuperação de informação. No *site* CIRES - *Content Based Image Retrieval System* (CIRES 2004) foram obtidas boa parte das imagens, pois trabalha com o mesmo assunto apresentado nesta dissertação. Os *sites* da *University of Southern California - Signal and Image Processing Institute* (USC 2004) e da *Revista Trip* (TRIP 2004) serviram para aumentar a base de imagens.

A base de imagens contém cenas variadas como, paisagens, animais, carros, texturas, flores, garotas e natureza em geral, todas imagens coloridas. Esta foi separada em 27 classes e possui um total de 1.020 imagens. As figuras possuem o mesmo tamanho de 225x400 *pixels* com a finalidade de obter um padrão do conjunto de imagens com a mesma dimensão. A Tabela 3.1 apresenta as classes usadas nos experimentos e o número de imagens em cada classe.

Tabela 3.1 – Classes e suas respectivas quantidades de imagens.

Classe	Nome	Número de imagens	Classe	Nome	Número de imagens
1	Água	70	15	Flores	43
2	Ar	53	16	Futebol	44
3	Animais	55	17	Fruta	19
4	Arame	06	18	Garota	27
5	Árvore	103	19	Árvore-rosa	55
6	Caixa	15	20	Janela	12
7	Carro	42	21	Folhagem	30
8	Natal	42	22	Mamíferos	16
9	Concreto	14	23	Mosaico	25
10	Construção	79	24	Montanhas	31
11	Amanhecer	63	25	Ponte	07
12	Pato	37	26	Céu	43
13	Bandeira	43	27	Neve	14
14	Textura	32			

Cada uma dessas 1.020 imagens foi convertida do modelo de cor RGB para o modelo HSV, que será melhor explicado na Seção 3.4.

3.3. Atributos Visuais

A análise de imagem baseada no seu conteúdo tem seu processo modelado como uma hierarquia de abstrações, isto é, em um primeiro momento estão os *pixels* da imagem, com a informação sobre cores ou textura associada ao elemento. Em um segundo momento estão os atributos como bordas, cantos, linhas, curvas e regiões de cores. Em um terceiro momento procura-se combinar e interpretar os atributos do momento anterior, relacionando os objetos que possuam tais características. Em um último momento aproxima-se do mapeamento humano, a partir do qual busca-se compreender o relacionamento entre os objetos presentes na imagem (TRAINA 2001).

Existem muitos métodos para detectar e reconhecer automaticamente determinadas classes de objetos e atributos, desse modo, a seguir são apresentados alguns atributos visuais e suas características.

3.3.1. Atributo Cor

As cores em uma imagem possuem um papel importante na recuperação da mesma. Na realidade todas as cores exibidas são criadas por combinações de quantidades apropriadas de vermelho, verde e azul. Um *pixel* de 24 bits em padrão RGB representa 2^{24} ou aproximadamente 16,7 milhões de cores diferentes. Muitas vezes, para aumentar a eficiência no processamento, as cores da imagem são re-quantizadas de forma a diminuir o número de cores possível e facilitar o tratamento das mesmas através de seu histograma (TRAINA 2001).

A utilização de histogramas de cores em sistemas de recuperação de imagens em conteúdo deve-se, principalmente, a três fatores (PASS 1996):

- é computacionalmente simples e barata a resolução de imagens em histogramas de cores.
- pequenas alterações de movimentação de imagem pouco afetam os histogramas.
- objetos distintos frequentemente possuem histogramas diferentes.

Dessa forma é natural que os histogramas de cores venham sendo estudados e implementados em sistemas de recuperação de imagens baseados em conteúdo, tanto acadêmicos (HAFNER 1995) quanto comerciais, como o QBIC (IBM 2004).

3.3.2. Atributo Textura

A textura é considerada um padrão visual na qual há um grande número de elementos visíveis distribuídos de forma imparcial com variadas densidades. A caracterização de textura varia com a intensidade em uma janela como contraste, granularidade, direcionalidade e repetitividade. A análise da textura obtém os elementos presentes em uma imagem, determinando seu formato, e estimando as regras de posicionamento. Um exemplo de elementos de textura está na Figura 3.3.

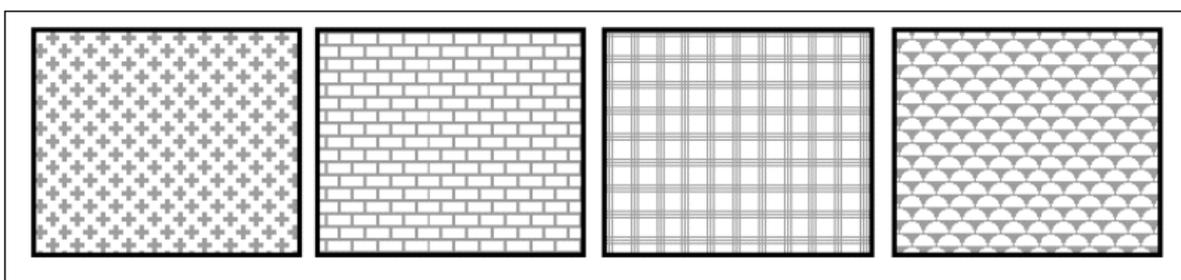


Figura 3.3 – Exemplos de quatro diferentes elementos de textura (BUENO 2001).

O tratamento de textura difere do realizado sobre cores devido ao fato de que as texturas são definidas sobre janelas ou regiões da imagem e não sobre *pixels* como as cores. A segmentação de uma imagem utilizando textura determina quais regiões da imagem possuem textura uniforme (BUENO 2001).

De acordo com Long e Leow (2000), modelos computacionais de característica de textura podem ser agrupados em três categorias: estrutural, estatístico e espectral. Modelos estruturais caracterizam texturas de acordo com o relacionamento local entre *pixels* de imagens. Modelos estatísticos categorizam texturas de acordo com medidas estatísticas de característica visual, tais como, grossura, granularidade, regularidade, entre outros. Modelos espectrais caracterizam textura como propriedades de transformadas de Fourier ou nos resultados de filtragem das texturas por filtros apropriados. Os modelos estatístico e espectral são freqüentemente usados em sistemas de recuperação de imagem baseados em textura.

Neste trabalho, a textura está sendo representada através da transformada *wavelet*, capturando as características em diferentes escalas.

3.4. Modelos de Cor

Na análise de imagens, o processamento da cor é de grande importância para identificar e extrair características, devido ao fato de estar presente em tudo o que é visível ao olho humano.

A cor está presente em tudo o que observamos. Este é um elemento essencial e desempenha funções múltiplas na visualização de imagens e cenas. A percepção da cor pelo homem é caracterizada pela interação da luz com seu sistema de visão. Desta forma, este interpreta as cores de maneira particular, dando-lhes um significado que depende de condições psicofísicas. Para padronizar a especificação das cores, foram criados sistemas para representação de cor.

Um modelo de cor é uma especificação de um sistema de coordenadas tridimensionais e um subespaço dentro deste sistema onde cada cor é representada por um único ponto. Os modelos de cores frequentemente usados para imagens coloridas são o R (*Red*), G (*Green*) e B (*Blue*) que são usados para monitores coloridos e câmeras de vídeo em cores; o YIQ que é o padrão para transmissão de TV colorida, onde o Y corresponde a luminância, o I e Q são dois componentes cromáticos chamados fase e quadratura e o modelo H (*Hue*), S (*Saturation*) e V (*Value*) usado para manipulação de imagens coloridas (GONZALEZ e WOODS, 1992).

A seguir, serão explicados os modelos de cor e o método de conversão usado neste trabalho.

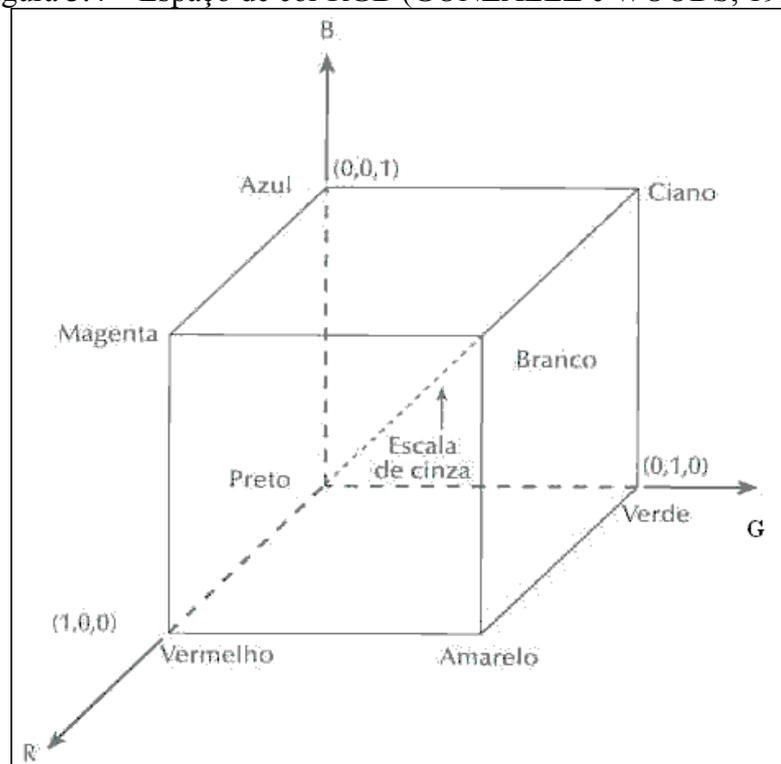
3.4.1. Modelo RGB

Imagens no modelo RGB consistem em três planos de imagens independentes, um para cada cor primária, vermelho, verde e azul. Essas três imagens combinam-se sobre a tela, quando visualizadas em um monitor RGB e produzem uma imagem de cores compostas.

O modelo RGB está associado às superfícies emissoras de luz. Por isso, é um modelo quase universal empregado pelos equipamentos que manipulam a emissão de luz, tais como os monitores e os televisores a cores.

A Figura 3.4 ilustra o modelo de cor RGB.

Figura 3.4 – Espaço de cor RGB (GONZALEZ e WOODS, 1992).



Os três parâmetros do modelo RGB definem um espaço tridimensional com direções ortogonais (R, G e B). Assim, está definido o espaço RGB. As cores deste espaço existem no subespaço em que $0 \leq (R, G, B) \leq 1$. Cada uma das cores primárias corresponde a um dos vértices do cubo localizados sobre os eixos do espaço, em que apenas uma das coordenadas não é nula (GONZALEZ e WOODS, 1992).

3.4.2. Modelo HSV

O modelo HSV (*Hue, Saturation, Value*) de cores deve sua utilidade a dois fatores principais. Um é o componente de intensidade V que é desacoplado da informação de cor na imagem. O outro fator são os componentes matiz e saturação que são relacionados à percepção humana de cores.

O termo *hue* distingue entre azul, verde, amarelo, vermelho; é a cor pura da imagem. A *saturation* da cor, por vezes denominada pureza ou saturação, indica o afastamento da cor. Uma cor vermelha ou azul puras são cores altamente saturadas, enquanto o rosa e as cores denominadas de pastéis são cores pouco saturadas. O *value* é a intensidade da luz refletida pela superfície nos objetos, o brilho é a quantidade de luz emitida pelas superfícies de objetos luminosos (GONZALEZ e WOODS, 1992). Um exemplo gráfico do modelo de cor HSV pode ser visualizado na Figura 3.5.

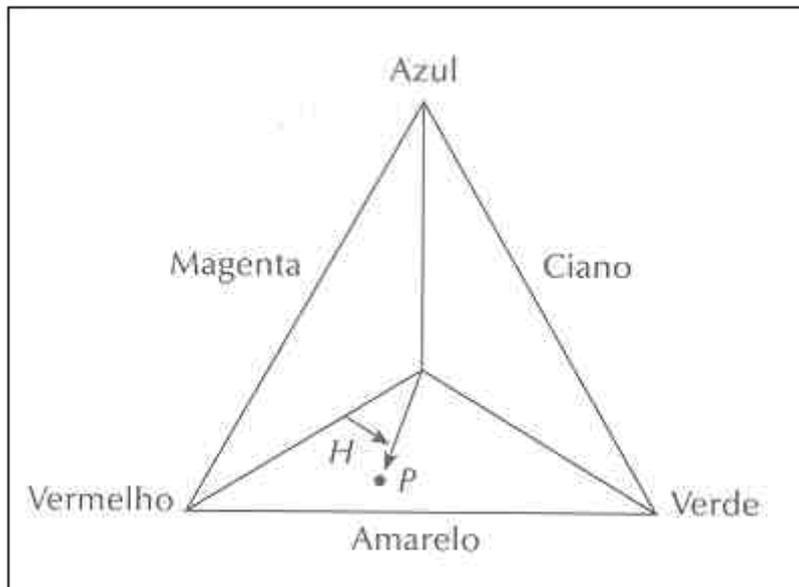


Figura 3.5 – Espaço de cor HSV (GONZALEZ e WOODS, 1992).

Quando se projetam as cores primárias e a cor branca sobre um mesmo plano, à distância de uma unidade da origem e escalando a pirâmide de forma que a distância entre o ponto representativo de qualquer cor primária e o ponto central correspondente à cor branca seja também unitária, obteremos uma estrutura piramidal de seis lados contendo todas as cores do espaço HSV.

A vantagem de se trabalhar com essa representação de cor encontra-se na possibilidade de separar a intensidade da informação tonalidade e saturação, bem como, na relação que existe entre essas componentes, muito próxima da forma na qual o homem percebe a cor.

3.4.3. Conversão de RGB para HSV

As cores nos modelos HSV são obtidas a partir das informações RGB com respeito aos valores normalizados do vermelho, verde e azul, dados por:

$$r = \frac{\text{Red}}{255} \quad (3.1)$$

$$g = \frac{\text{Green}}{255} \quad (3.2)$$

$$b = \frac{\text{Blue}}{255} \quad (3.3)$$

Após obter os canais r , g e b normalizados, as equações abaixo são utilizadas para o cálculo da intensidade, saturação e matiz.

$$V = \frac{(r + g + b)}{3} \quad (3.4)$$

$$S = 1 - \frac{3}{(r + g + b)} \times \min(r, g, b) \quad (3.5)$$

$$H = \arctan \left(\frac{\sqrt{(r - g) + (r - b)}}{\sqrt{(r - g)^2 + (r - g)(g - b)}} \right) \quad (3.6)$$

A vantagem dessa representação de cor está na possibilidade de separar os canais. Esse fato torna o modelo HSV uma ferramenta ideal para o desenvolvimento de algoritmos de processamento de imagens, Figura 3.6.



Figura 3.6 – Canal S de uma imagem.

A Figura 3.6 representa um exemplo do canal S de uma imagem da base de dados que foi convertida do modelo de cor RGB para o modelo de cor HSV.

A característica cor é utilizada como um atributo para sistemas de recuperação de informação por conteúdo. A descrição de cor em objetos do mundo real pode ser alterada por alguns efeitos como: textura da superfície, pela mudança da iluminação ou sombra de outros objetos na cena, pelo espectro da iluminação, pelo caminho da luz que interage com o objeto e posição do observador.

Em (SCHETTINI *et al* 2001) a utilização da característica cor é feita intensivamente para experimentos em correspondência de cor, encontrando imagens contendo uma cor específica em determinada proporção, buscas por regiões para encontrar imagens contendo regiões com cores como especificado na imagem inicial, recuperar imagens com um certo efeito de cor induzido, entre outros.

Em um contexto geral, as características selecionadas apresentam algumas propriedades básicas, tais como:

- similaridade perceptual: quanto maior a diferença de uma imagem com outra, menor será a sua semelhança;
- eficiência: as características podem ser rapidamente computadas;
- escalabilidade: o desempenho do sistema não é influenciado pelo tamanho da base de dados.

Essas tarefas dependem da definição eficiente e robusta da característica cor que pode representar o conteúdo da imagem.

3.5. Distribuição Espacial

Um ponto fraco em métodos de RIC é a falta de informação espacial, ou seja, as imagens possuem mesma proporção de cores, mas diferentes distribuições espaciais. Pode-se concluir que trabalhar somente com distribuição espacial não é o suficiente para representar o conteúdo *pictorial* de uma imagem.

Assim, uma maneira de fornecer informação espacial é dividir a imagem em sub-imagens, e indexar cada uma dessas partes (STRICKER e DIMAI 1996).

A Figura 3.7 mostra alguns exemplos de padrões tendo a mesma proporção de cor, mas diferente distribuição espacial.

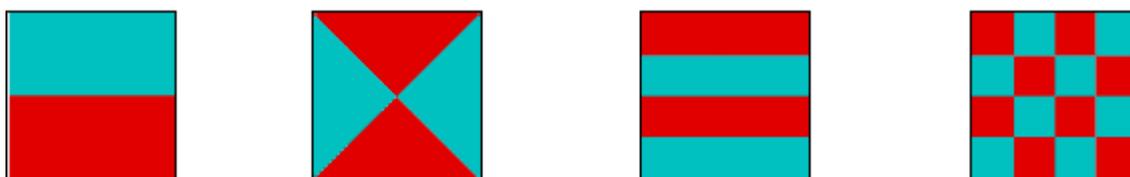


Figura 3.7 – Exemplos de figuras com mesma proporção de cor, mas diferente distribuição espacial (SCHETTINI *et al* 2001).

A vantagem de se trabalhar com distribuição espacial é que ela fornece informações significativas através dos *pixels* de cada porção da imagem que for particionada. Neste trabalho está se usando a distribuição espacial fazendo uma divisão das imagens em nove partes iguais.

Optou-se por essa divisão pelo fato de poder localizar elementos que mesmo dispostos em regiões diferentes da imagem possam ser comparados e assim um *ranking* pode ser feito. Como se propõe trabalhar com imagens quaisquer, a divisão em nove partes possibilita imagens com muitos tipos de elementos, e não só aquela com elemento central mais importante.

3.6. Decomposição *Wavelet*

A transformada *wavelet* utiliza funções *wavelet*, que tem uma boa localização tanto no tempo quanto no espaço, sendo possível desenvolver decomposições *wavelet* com uma grande variedade de funções básicas, e também enfatizar a redundância ou eliminá-las através dos níveis de decomposição.

A decomposição de Mallat (1989) define a aplicação de operações de convolução do sinal da imagem com filtros *QMF's* (*Quadrature Mirror Filters*) e subamostragens (*downsamplings*).

A Figura 3.8 ilustra o processo de decomposição.

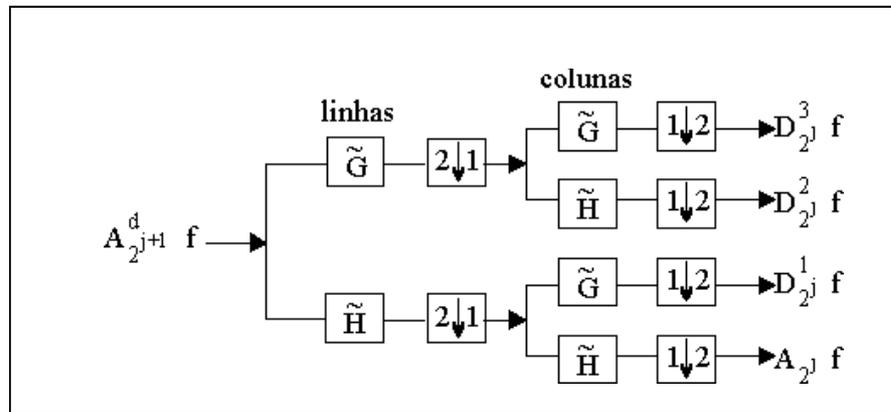


Figura 3.8 – Diagrama ilustrando o processo de decomposição.

A decomposição proposta por Mallat (1989) define a aplicação de operações de convolução do sinal com filtros passa-baixa (H) e passa-alta (G) nos dados originais, seguidos de operações de subamostragem de linhas ($1 \downarrow 2$) e colunas ($2 \downarrow 1$), dependendo do caso. A operação ($1 \downarrow 2$) mantém uma linha de duas e a ($2 \downarrow 1$) mantém uma coluna de duas.

Assim, o sinal de entrada $A_{2^{j+1}}^d f$ é decomposto em quatro conjuntos de coeficientes: $A_{2^j}^d f$, $D_{2^j}^1 f$, $D_{2^j}^2 f$ e $D_{2^j}^3 f$, onde as $D_{2^j}^i f$ representam as bandas de alta frequência contendo informações direcionais de detalhe na escala j , por isso referenciadas como imagens de detalhe e $A_{2^j}^d f$ é a banda de baixa frequência referenciada como uma imagem de baixa resolução na escala j .

Os filtros usados, H e G, para a decomposição *wavelet* no algoritmo de Mallat representam a função de base Haar, utilizada neste trabalho, escolhida experimentalmente e por apresentar uma menor complexidade algorítmica, isto comparada com a base Daubechies.

Para ilustração, a Figura 3.9 apresenta os quatro canais de saída do algoritmo de decomposição *wavelet*, esses quatro conjuntos representam um conjunto para dados suavizados ou de baixa frequência (aproximação), e mais três conjuntos direcionais de alta frequência que são: horizontal, vertical e diagonal.

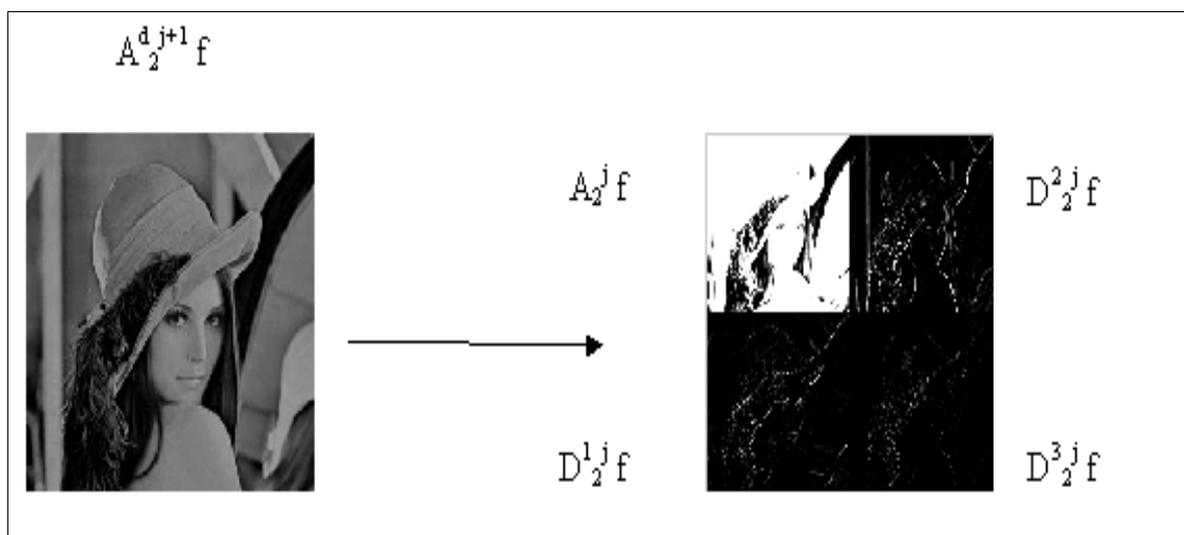


Figura 3.9 – Representação da decomposição *wavelet* em uma imagem.

Na Figura 3.9, os conjuntos de saída da decomposição *wavelet* são para dados suavizados ou de baixa frequência, $A_{2^j} f$, e os três conjuntos de detalhes, sendo de alta frequência, vertical $D_{2^j}^{1,j} f$, horizontal $D_{2^j}^{2,j} f$ e diagonal $D_{2^j}^{3,j} f$. Os conjuntos são subamostragens da imagem original, porém os coeficientes de detalhe contêm dados que somados a $A_{2^j} f$ reconstruem a imagem original. As imagens dos coeficientes vertical $D_{2^j}^{1,j} f$, horizontal $D_{2^j}^{2,j} f$ e diagonal $D_{2^j}^{3,j} f$, apresentadas na Figura 3.9 são representadas da seguinte forma: os coeficientes mais escuros representam as regiões de menor resposta aos filtros, enquanto que os coeficientes mais claros representam as regiões de alta resposta do filtro.

No algoritmo *wavelet*, usando a base Haar, a dimensão dos canais de saída é a metade da dimensão da imagem original. Esta redução de dimensionalidade é realizada através da

subamostragem da decomposição *wavelet*.

3.7. Interseção

Para atribuir um significado a um conjunto de características, comparam-se as informações extraídas por uma função de similaridade. Nesse sentido, medidas de distância entre histogramas são usadas para calcular a similaridade das características, onde o histograma é usado como um conjunto ordenado de características (SMEULDERS *et al* 2000).

Neste trabalho a medida de similaridade usada é a interseção de histogramas, como mostra a fórmula 3.7:

$$Dist(H, H') = \frac{\sum_{i=1}^n \min(h_i, h'_i)}{\sum_{i=1}^n h'_i} \quad (3.7)$$

h_i = histograma da imagem teste

h'_i = histograma da imagem testada

Um histograma é um vetor $H = [h_1, \dots, h_n]$ em que cada entrada contém o número de *pixels* havendo a cor da imagem distribuída em proporções, e é considerada uma função de densidade de probabilidade das cores.

3.8. Distância Euclidiana

A distância euclidiana é utilizada como métrica para cálculo da distância na etapa de classificação. Na classificação, cada *pixel* será incorporado a um agrupamento, através da análise da similaridade de distância euclidiana, que é dada por:

$$d_{Euclid} = \sqrt{\sum_{i,j} (A(i,j) - M(i,j))^2} \quad (3.8)$$

$A(i,j)$ = matriz que representa a imagem teste

$M(i,j)$ = matriz que representa a imagem testada

O *pixel* é incorporado ao agrupamento que apresenta a menor distância euclidiana. Este procedimento é feito até que toda a imagem seja classificada. Assim, quanto menor a distância euclidiana entre duas imagens, maior a proximidade entre elas.

3.9. Somatório

Para fazer o cálculo entre histogramas foi desenvolvido um algoritmo baseado no somatório entre *pixels*. Utilizando essa concepção pode-se concluir que quando duas imagens são similares o somatório entre elas é igual a 1 (um). O somatório é dado pela fórmula 3.9:

$$S(f, fl) = \frac{\sum (f_i, fl_i)}{n * m} \quad (3.9)$$

f_i : imagem referência

fl_i : imagem referida

O somatório é realizado para a distribuição espacial, como explicado na seção 3.5. O resultado é gerado para ser usado como referência para a classificação das imagens. Assim, quanto maior a resposta da métrica somatório, maior será a similaridade das imagens.

3.10. Desempenho da Efetividade

Na implementação final de um sistema de recuperação de informação por conteúdo, é necessário fazer uma avaliação do sistema. Sistemas de RIC necessitam da avaliação de quanto preciso é o conjunto de resposta. Assim, este tipo de avaliação é denominado Avaliação do Desempenho da Efetividade.

A medida chamada de Efetividade foi proposta por Mehtre *et al* (SMEULDERS *et al* 2000) para avaliar os métodos de recuperação de imagem.

Para avaliar o desempenho do algoritmo de recuperação de informação, a eficiência é medida em termos de revocação e precisão. Para entender melhor o significado dessas medidas:

- revocação – é a fração dos documentos relevantes que foi recuperado;
- precisão – é a fração dos documentos recuperados que é relevante.

Seja S o número de imagens recuperadas de uma lista de imagens semelhantes a partir de uma *query*, $R_{i,q}$ o conjunto de imagens relevantes na base de dados, ou seja, o número de imagens que cada classe possui e $R_{e,q}$ o conjunto de imagens recuperadas, ou seja, imagens recuperadas pelo sistema até o limiar (S), imagens de uma mesma classe.

$$\text{Revocação} = \frac{R_{i,q} \cap R_{e,q}}{R_{i,q}} \quad (3.10)$$

$$\text{Precisão} = \frac{R_{i,q} \cap R_{e,q}}{R_{e,q}} \quad (3.11)$$

Se $R_{i,q} \leq S$, a efetividade corresponde a revocação, ou, se $R_{i,q} > S$ a efetividade corresponde a precisão. O desempenho pode ser avaliado por diferentes valores do limiar S , neste trabalho está se usando 60% das primeiras imagens de um total de 1.020 imagens. Usou-se um limiar de 60% para fazer uma comparação com as seiscentas primeiras imagens. Poderia ter sido escolhido qualquer valor para o limiar, mas optou-se por usar um número aleatório de imagens, em nível de comparação.

3.11. Resultados do Desempenho

O processamento dos resultados é feito para cada imagem da base de dados, ou seja, o cálculo da distância euclidiana, o somatório e a ordenação dos resultados foram processados nas imagens. Na distância euclidiana, quanto menor a distância entre as imagens, maior será a proximidade entre elas e na métrica somatório quanto maior a distância entre as imagens,

menor será a proximidade entre elas. Fazendo a comparação da imagem com ela mesma, o resultado da distância euclidiana tem que ser zero, pois representa a própria imagem, e no somatório, a comparação da imagem com ela mesma, deve produzir um como resultado.

Na Tabela 3.2, é apresentado um exemplo de um arquivo gerado pelo classificador²⁹ processamento feito para cada imagem e sua respectiva classificação. No primeiro conjunto dos resultados ordenados, com a classe 1, obteve-se um resultado de classificação considerando a recuperação das cinco primeiras imagens, quatro foram classificadas corretamente, aplicando a métrica distância euclidiana com uma porcentagem de 80%, conforme Tabela 3.3. Os testes foram feitos apenas com duas classes contendo um total de dez imagens, para efeito de ilustração.

Tabela 3.2 – Exemplo de arquivo gerado pelo classificador.

```

Resultados Ordenados
=====
#ClasseImagem: 1
!Classificacao          Distancia  Classe
agua01.jpg----- 0.00000    1
agua02.jpg----- 4.67040    1
agua05.jpg----- 5.04757    1
agua04.jpg----- 5.56732    1
air01.jpg----- 5.60768    2
air05.jpg----- 6.08595    2
air04.jpg----- 6.56043    2
air02.jpg----- 6.78996    2
agua03.jpg----- 7.14666    1
air03.jpg----- 7.92337    2
-----
#ClasseImagem: 1
!Classificacao          Distancia  Classe
agua05.jpg----- 0.00000    1
agua04.jpg----- 4.81441    1
agua01.jpg----- 5.04757    1
air01.jpg----- 5.41212    2
agua02.jpg----- 5.45825    1
air04.jpg----- 5.65517    2
air02.jpg----- 6.28700    2
agua03.jpg----- 6.29530    1
air05.jpg----- 6.34942    2
air03.jpg----- 7.40276    2
-----
#ClasseImagem: 2
!Classificacao          Distancia  Classe
air01.jpg----- 0.00000    2
agua05.jpg----- 5.41212    1
agua01.jpg----- 5.60768    1
agua02.jpg----- 5.65238    1
air02.jpg----- 5.73722    2
agua04.jpg----- 5.90973    1
air04.jpg----- 6.25441    2
air03.jpg----- 6.27267    2
air05.jpg----- 6.35655    2
agua03.jpg----- 7.90793    1
-----
#ClasseImagem: 2
!Classificacao          Distancia  Classe
air02.jpg----- 0.00000    2
air03.jpg----- 4.75852    2
air04.jpg----- 4.97026    2
air01.jpg----- 5.73722    2
agua05.jpg----- 6.28700    1
agua04.jpg----- 6.50256    1
air05.jpg----- 6.60562    2
agua01.jpg----- 6.78996    1
agua03.jpg----- 7.29505    1
agua02.jpg----- 7.69567    1

```

Para cada bloco de imagem classificado é calculada a medida de efetividade, usando as equações 3.10 e 3.11. Na Tabela 3.3 é mostrado um exemplo de um arquivo processado pelo classificador com as medidas calculadas.

Tabela 3.3 – Exemplo de arquivo com o cálculo das medidas.

Resultado parcial -	80.00
Resultado parcial -	80.00
Resultado parcial -	100.00
Resultado parcial -	80.00
Resultado parcial -	80.00
Média da classe 1:	84.00
Resultado parcial -	40.00
Resultado parcial -	80.00
Resultado parcial -	100.00
Resultado parcial -	80.00
Resultado parcial -	40.00
Média da classe 2:	68.00

Conforme a classificação da Tabela 3.2, os resultados parciais são calculados para os blocos de cada imagem. Assim os resultados do desempenho são gerados para fazer a média final de cada classe, ou seja, obter a média final em que cada classe teve a sua classificação. A média de recuperação da classe 1 teve 84% de acerto, e a classe 2 teve 68% de recuperação.

3.12. Descrição da abordagem proposta

A abordagem proposta consiste na utilização da decomposição *wavelet* nas imagens da base de dados, extraindo os coeficientes mais significativos que são utilizados para calcular as similaridades e saliências das imagens.

O método pode ser dividido em 4 passos, que serão melhor explicados nas seções seguintes.

3.12.1. Descrição dos passos

Passo 1 – Inicialmente, a imagem está no formato RGB e em seguida é convertida para o sistema HSV usando as fórmulas 3.1, 3.2, 3.3. As imagens da base de dados possuem dimensão de 225x400 *pixel*, Figura 3.10.



Figura 3.10 – Imagem inicial de teste.

Essa conversão gera os canais do modelo de cor HSV individuais, tendo a vantagem de cada imagem ser trabalhada separadamente, Figura 3.11.

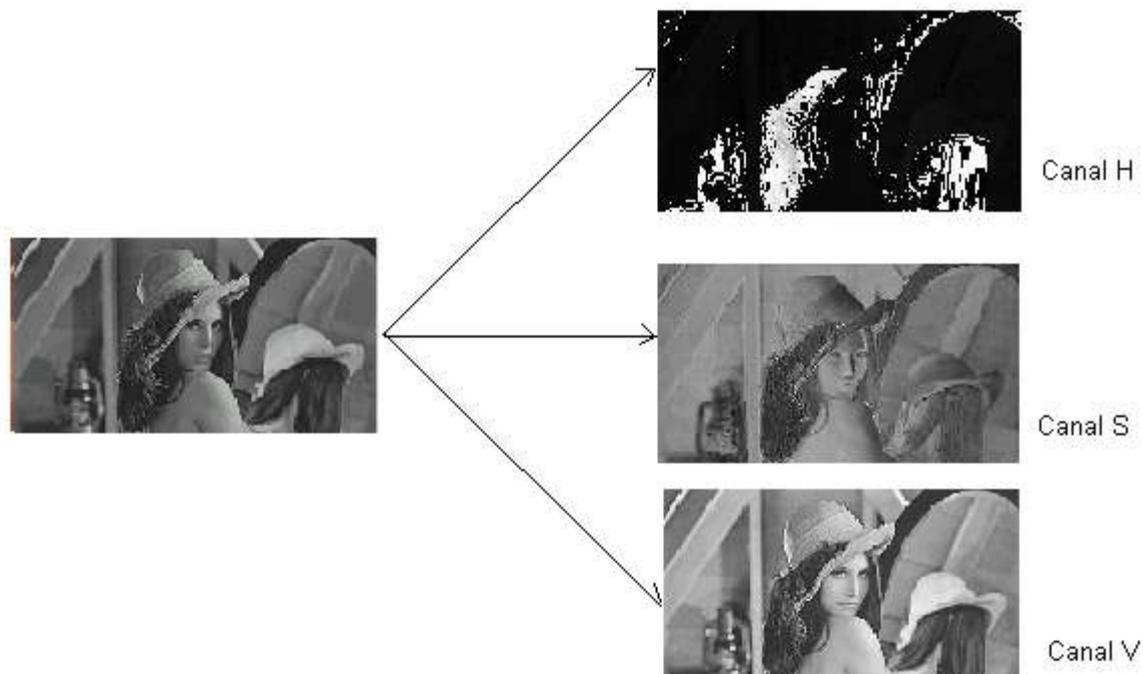


Figura 3.11 – Imagens dos canais convertidos H, S e V.

Passo 2 – A região de interesse da imagem é definida pela distribuição espacial, ou seja, entende-se como distribuição espacial, a região definida a partir de detalhes obtidos da própria imagem, onde será concentrado o processamento. A distribuição espacial é feita pela divisão da imagem em nove partes iguais, Figura 3.12. Para obter resultados de cada porção da imagem, a distribuição espacial é adquirida para observar o melhor nível de detalhes na visualização, assim, a medida que a distribuição aumenta o nível de detalhes da imagem também aumenta.

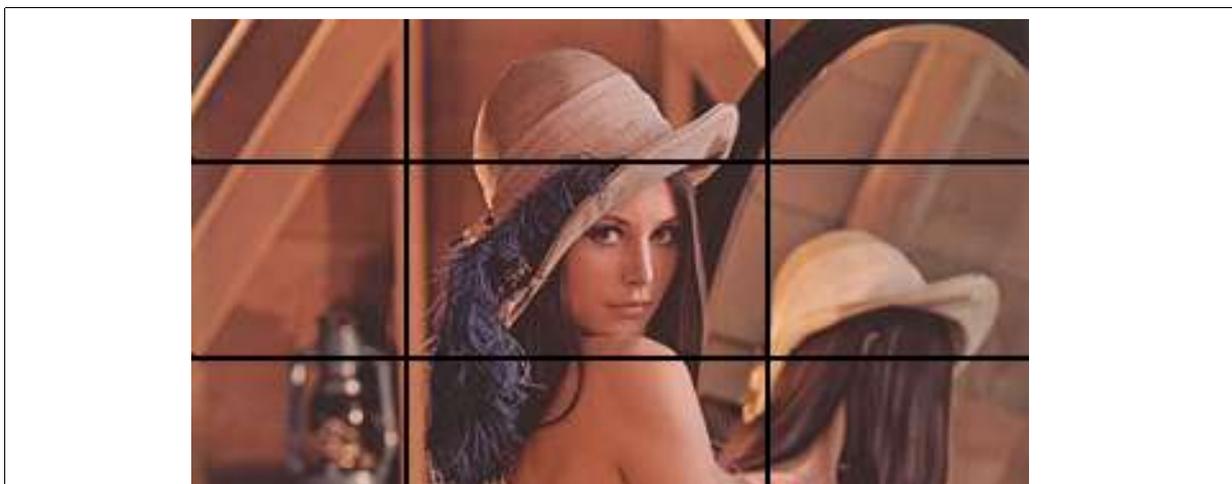


Figura 3.12 – Imagem dividida em nove partes.

Para cada distribuição é aplicado o algoritmo de decomposição *wavelet*, que será melhor explicado no passo 3.

Passo 3 – A base utilizada neste trabalho é a Haar, escolhida experimentalmente, essa base é usada para extrair os coeficientes de aproximação, detalhe horizontal e detalhe vertical. Ainda é gerado um histograma dos coeficientes de cada canal do modelo HSV. A Figura 3.13 apresenta os histogramas dos coeficientes de aproximação do canal H, S e V da imagem.

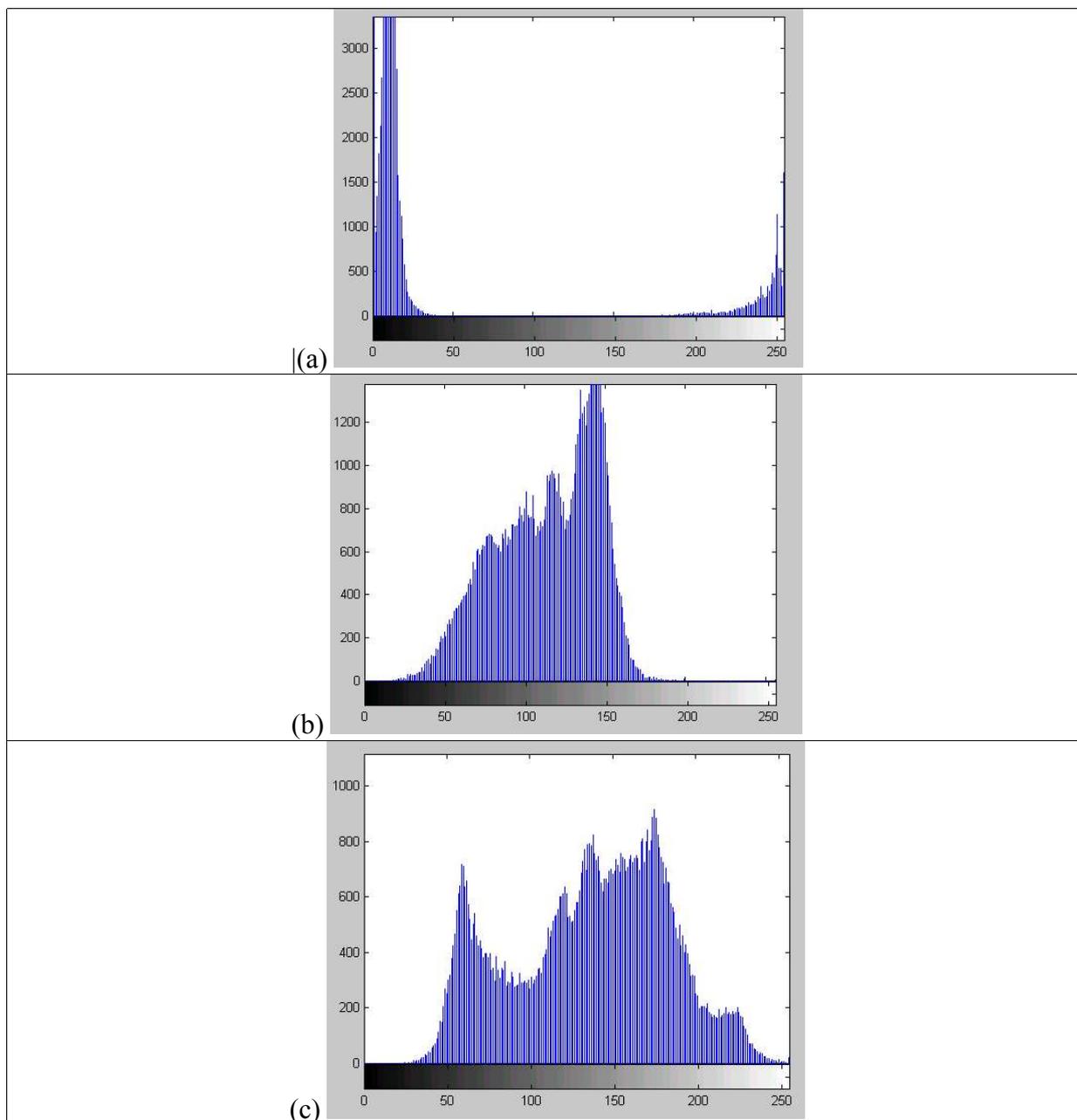


Figura 3.13 – Histogramas do canal H (a), canal S (b) e canal V (c), baseados na imagem teste.

O histograma é representado por um gráfico que para cada nível de cinza gera o número de *pixels* correspondentes na imagem. O histograma fornece o número de *pixels* e não a localização desses, descrevendo a imagem globalmente.

Passo 4 – Para identificar as características mais salientes são calculadas a interseção e a distância euclidiana dos coeficientes de aproximação, detalhe horizontal e detalhe vertical, pelas fórmulas 3.7 e 3.8, que são extraídos para calcular as similaridades e saliências das imagens. Os cálculos das fórmulas são processados internamente, gerando resultados numéricos.

Os 4 passos são repetidos para o processamento de todas as 1.020 imagens da base de dados. Gerando arquivos de resultados como os da Tabela 3.2 e Tabela 3.3.

3.13. Matriz de Confusão

É possível identificar dentre as imagens que não foram classificadas corretamente, a quais classes tais imagens foram atribuídas. Esses resultados podem ser extraídos da matriz de confusão através da relação de imagens classificadas, assim, é possível avaliar como as imagens foram erroneamente classificadas entre as classes.

A Tabela 3.4 representa a matriz de confusão relacionada com a base *wavelet* Haar, com 18 classes de imagens e os resultados obtidos com a métrica distância euclidiana, para efeito de ilustração. Essa tabela apresenta a classificação da classe 1 (água) com a própria classe 1 (água), em seguida a classificação da classe 1 (água) com a classe 2 (ar), e assim por diante. Por exemplo, as imagens da classe 1 (água) comparadas com as imagens da classe 1 (água), estima-se que 19 imagens foram recuperadas corretamente. A classe 1 (água) comparada com a classe 2 (ar) teve 9 imagens recuperadas como sendo da classe 2 (ar). Outro exemplo pode ser dado através da classe 8 (natal) que teve identificação com a classe 11 (amanhecer) em 7,2857 imagens. Como está se trabalhando com distribuição espacial nas imagens, os valores aparecem em números reais e não com valores de classificação inteiros. Esses resultados foram gerados com um limiar de 40% para efeitos de ilustração. No trabalho está sendo usado um limiar de 60% para analisar a classificação das imagens, esse número foi escolhido aleatoriamente como referência para recuperar um pouco mais da metade das imagens. Os resultados da matriz de confusão foram calculados segundo a medida de efetividade, apresentada na Seção 3.10, gerando as imagens das classes que foram identificadas erroneamente.

No Capítulo 3 foi apresentado o método proposto de solução para o problema, evidenciando as características essenciais da mesma, como a base de imagens usada e os atributos visuais usados neste trabalho. Além disso, os modelos de cor, a distribuição espacial, a decomposição *wavelet*, e as métricas de distâncias foram apresentados.

O Capítulo 4 mostra os experimentos realizados, juntamente com os resultados e a análise dos mesmos.

Tabela 3.4 – Matriz de confusão usada como ilustração.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	19,4286	6,8714	1,6286	0,9000	6,7857	0,0286	2,9000	1,3714	0,2571	4,4143	0,8714	0,4143	5,0286	0,4286	1,3143	0,3429	0,1714	3,0714
2	6,7736	9,2264	1,3208	1,2642	2,7547	0,1509	0,9623	1,7736	0,2642	4,5094	0,7925	0,1698	7,0000	0,5472	0,8868	0,7358	0,2642	2,8302
3	0,7091	0,8000	6,9818	0,1091	4,3273	0,6182	4,6182	2,1091	1,9273	5,4545	0,8727	1,9273	1,1636	5,3273	0,9455	3,5455	0,6545	1,4182
4	0,1667	1,5000	0,1667	2,3333	0,1667	0,0000	0,0000	0,0000	0,0000	0,0000	0,1667	0,0000	0,6667	0,0000	0,0000	0,0000	0,0000	0,0000
5	3,3725	2,2843	4,4118	0,1373	17,3725	0,4118	10,0686	3,8627	0,8137	9,4608	1,4510	2,8235	4,5882	6,9706	5,9804	3,2647	0,3725	7,5196
6	0,0000	0,2000	1,1333	0,0000	0,4667	4,0000	1,6667	1,5333	0,2000	0,2000	0,2000	0,1333	0,0667	1,0000	1,0000	0,2667	0,2000	0,0667
7	0,8095	0,2143	2,6429	0,0000	3,6667	0,0476	10,2381	2,8095	0,2857	3,7381	0,2857	0,8810	1,0000	3,4524	0,6429	2,0000	0,0952	1,2143
8	0,7619	1,2381	1,8571	0,0000	2,6905	0,1429	2,6667	7,0000	0,5476	3,9048	3,1667	0,2143	0,3333	1,8333	2,0000	1,3333	0,9048	2,9286
9	0,0714	0,0714	1,8571	0,0000	0,7143	0,0714	0,9286	0,6429	2,8571	0,8571	1,0000	0,5000	0,1429	0,6429	0,0714	0,0714	0,8571	0,0000
10	2,2658	2,6203	4,0127	0,2152	6,5823	0,3165	7,2532	4,3165	1,1899	16,7342	2,2025	0,7468	1,1139	3,1646	1,5063	3,1772	0,5316	4,2405
11	0,9365	1,8730	3,8571	0,0317	2,0159	0,6190	3,3810	7,2857	1,5238	6,5556	8,0159	0,4444	0,1905	2,1111	1,8254	2,3333	2,0635	3,7302
12	0,3514	0,2162	2,7838	0,0270	5,7297	0,1892	2,9189	0,8919	0,7568	1,5676	0,2162	5,5676	1,7027	5,3784	0,1351	2,7297	0,1351	0,5946
13	3,9070	7,3488	1,0000	0,8372	4,1163	0,0233	2,5581	0,6279	0,1163	1,0930	0,1628	0,8372	8,1860	1,8837	0,1860	0,4186	0,0465	1,3488
14	0,0930	0,0930	2,5116	0,0000	6,1860	0,0698	2,8837	1,4884	0,1860	0,8372	0,2558	1,8837	1,4419	11,6279	0,8140	4,7674	0,0465	2,5814
15	0,5000	0,7045	1,7500	0,0000	6,2045	0,2727	2,6591	2,0227	0,0455	1,1591	0,2500	0,4545	0,2955	0,8864	20,8636	0,0000	0,1136	3,2045
16	0,0000	0,0000	1,5263	0,0000	2,1053	0,0000	0,2632	0,4211	0,0000	0,6842	0,0000	0,3684	0,0000	5,1579	0,0000	5,7368	0,0000	0,2105
17	0,1852	0,5556	1,4815	0,0000	0,8889	0,1852	0,6667	3,3704	1,4815	1,8519	1,6667	0,0741	0,3333	0,8148	1,2222	0,8889	2,9259	1,0370
18	1,2909	1,5636	2,4545	0,2909	4,8727	0,0545	3,9455	3,3818	0,2182	3,5091	0,7455	0,2545	1,0182	3,5818	4,4000	1,0364	0,0000	15,8364

Capítulo 4

Experimentos

Os resultados apresentados nessa seção elucidam os experimentos realizados para validação do método de recuperação de informação proposto. O algoritmo implementado deve levar em consideração todas as possíveis informações existentes nas imagens trabalhadas. Os resultados da matriz de confusão são apresentados na Seção 4.1, a Seção 4.2 apresenta a análise da combinação dos resultados e as considerações finais são apresentadas na Seção 4.3.

Os experimentos realizados seguem a variação dos seguintes passos, conforme mostra a Figura 4.1. Em um primeiro momento a imagem lida é convertida do modelo atual (RGB) para um modelo de cor específico, no caso deste trabalho está se usando o modelo HSV pelo fato de os canais possuírem suas características independentes. A imagem é decomposta no espaço *wavelet*, utilizando o algoritmo de decomposição de Mallat (1989) com a base Haar. Esse algoritmo executa uma transformação não-redundante e ortogonal que resulta na saída de quatro canais de características. A transformada *wavelet*, base Haar, é calculada para cada divisão (nove partes) da imagem, onde são extraídos os coeficientes de aproximação, coeficientes de detalhe horizontal e os coeficientes de detalhe vertical em três níveis de decomposição, ou seja, usando a multirresolução das imagens. As imagens possuem tamanho de 225x400 *pixels*, quando a transformada *wavelet* é calculada as imagens diminuem de tamanho. A vantagem dessa redução de dimensionalidade é que os coeficientes relevantes ficam em evidência, sendo assim são usadas somente essas informações. As informações apresentadas nos histogramas são usadas para o cálculo da métrica de similaridade, onde são calculadas a interseção de histogramas, a distância euclidiana e a métrica somatório para se ter

as informações de semelhança das imagens. Esses resultados são colocados em uma ordem de prioridade que se possa verificar a similaridade das figuras.

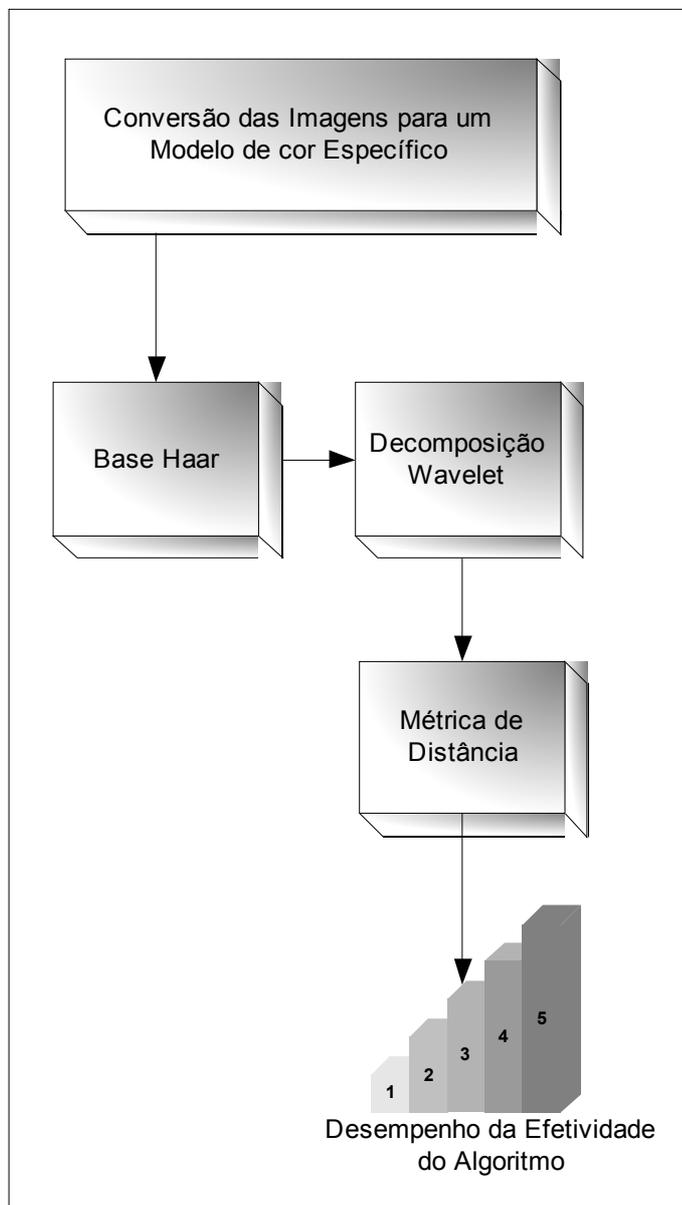


Figura 4.1. – Diagrama dos passos que são realizados para a execução dos experimentos.

Os resultados iniciais foram gerados contendo apenas três classes com um total de 15 imagens, para efeitos de ilustração. Na primeira rodada de testes o classificador faz um *ranking* das imagens mais semelhantes até as menos semelhantes, ou seja, cada imagem da base que está sendo lida é comparada com todas as outras imagens da base, fazendo uma classificação pela classe da imagem.

A Figura 4.2 mostra o exemplo da classificação contendo os resultados teste gerados pelo classificador com três classes (água, ar e animais) e cada classe com cinco imagens cada, usando apenas os coeficientes de aproximação no primeiro nível de decomposição.



Figura 4.2 – Resultados com a classificação.

Na segunda rodada de testes o classificador faz uma análise dos resultados gerados a partir da primeira rodada, ou seja, calcula a medida de efetividade para gerar o desempenho da recuperação da informação.

A cada processamento das imagens, o classificador produz médias de classificação independentes para as classes, considerando os canais H, S e V separadamente. Os resultados são calculados partindo de um limiar, ou seja, uma determinada quantidade de imagens recuperadas é usada para realizar o cálculo. Nesse caso, está se usando 60% do total de imagens da base de dados, esse número foi escolhido aleatoriamente para verificar os resultados que seriam melhores classificados com um pouco mais da metade da base, ou seja, as 600 primeiras imagens.

A menor distância entre a imagem da *query* e a imagem a ser comparada resulta na semelhança das imagens. A efetividade do método da recuperação de informação é avaliada com as medidas de revocação e precisão.

Os experimentos foram realizados sobre a base de dados de 27 classes com um total de 1.020 imagens. Todas as imagens foram processadas pelo método de recuperação de informação proposto. A Figura 4.3 apresenta algumas das imagens da base de dados, onde cada imagem representa a sua classe específica.

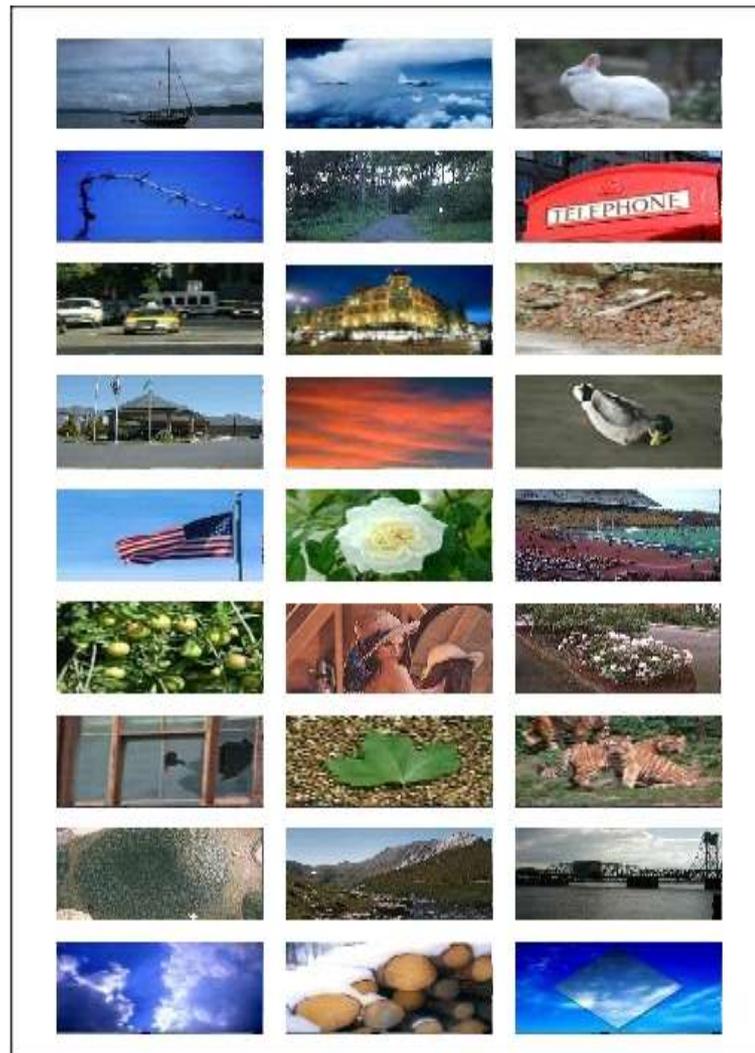


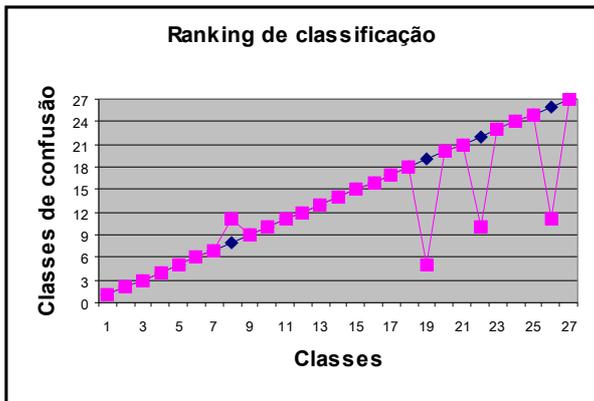
Figura 4.3 – Imagens que representam as 27 classes usadas nos experimentos.

A seguir, são apresentados resultados obtidos pelo método de recuperação proposto neste trabalho.

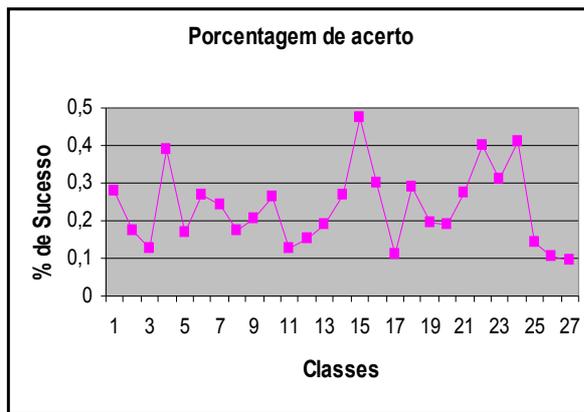
4.1. Resultados da Matriz de Confusão

Nesta subseção os resultados dos experimentos da matriz de confusão são apresentados e analisados.

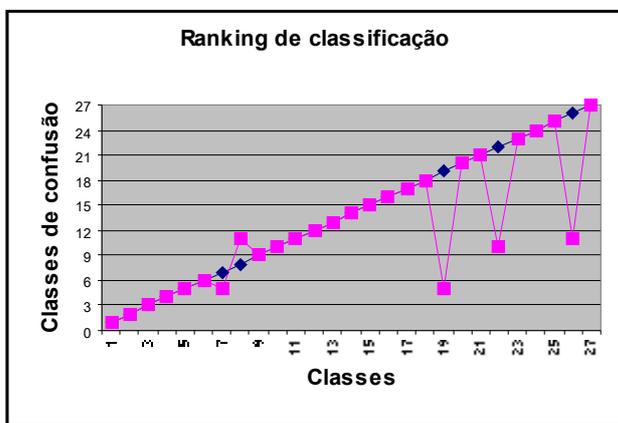
Os Gráficos 4.1 apresentam ilustrações para melhor elucidar a matriz de confusão feita para o canal H no 1º, 2º e 3º nível de decomposição, sendo o *ranking* de classificação que corresponde a quais classes foram identificadas erroneamente e o gráfico da porcentagem de acerto que corresponde a taxa de sucesso da classificação de cada classe.



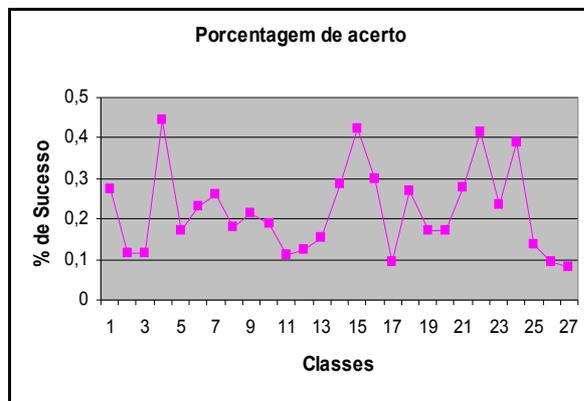
(a) Decomposição no 1º nível com os coeficientes de aproximação do canal H.



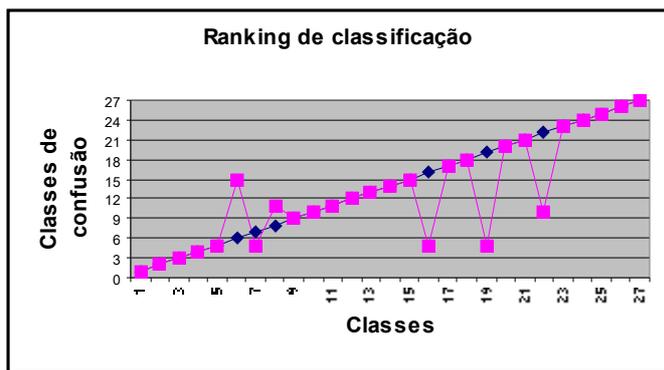
(b) Gráfico referente a porcentagem de acerto no 1º nível.



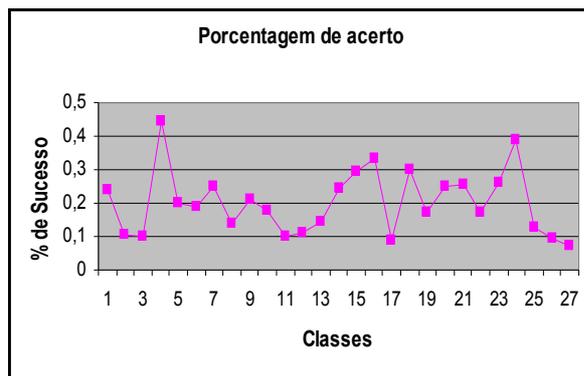
(c) Decomposição no 2º nível com os coeficientes de aproximação do canal H.



(d) Gráfico referente a porcentagem de acerto no 2º nível.



(e) Decomposição no 3º nível com os coeficientes de aproximação do canal H.



(f) Gráfico referente a porcentagem de acerto no 3º nível.

Gráfico 4.1 – Gráficos dos resultados do *ranking* de classificação e porcentagem de acerto do canal H, coeficientes de aproximação. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.

Pode-se observar através do Gráfico 4.1 (a), (c), (e), considerando o *ranking* de classificação, que as classes 19 (janela), 22 (mosaico) e 26 (neve) foram identificadas erroneamente nos três níveis de decomposição. Isso aconteceu devido as imagens apresentarem tonalidades de cores muito parecidas. Uma afirmação dessa classificação pode ser vista com as imagens das classes que tiveram a mesma identificação, como as classes 19 (janela) e 5 (árvore), Figura 4.4.

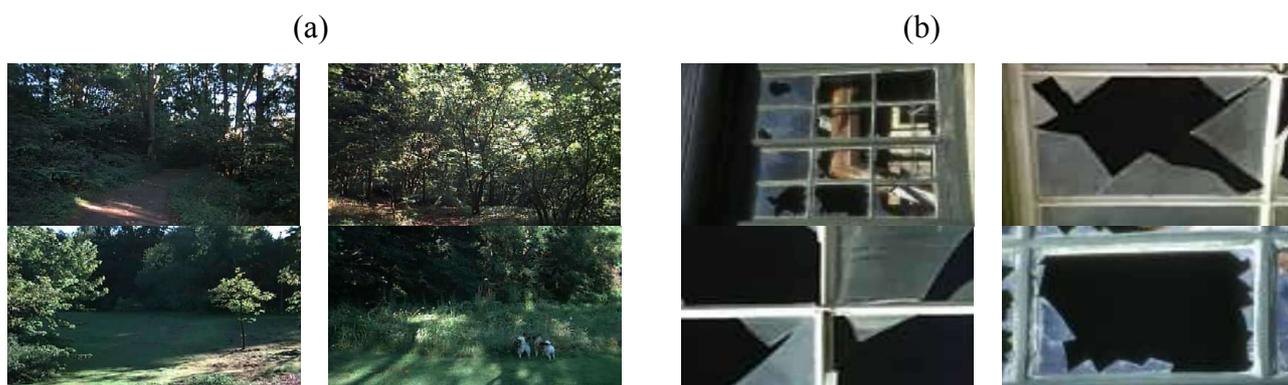
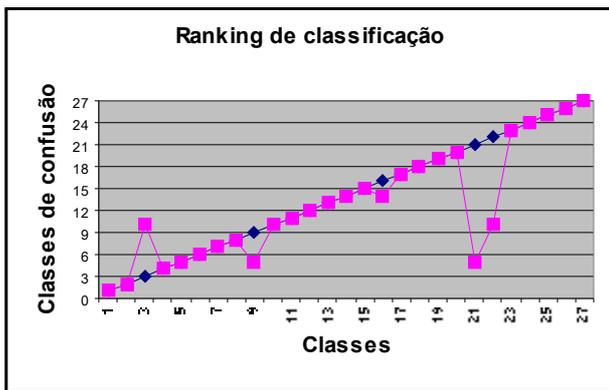


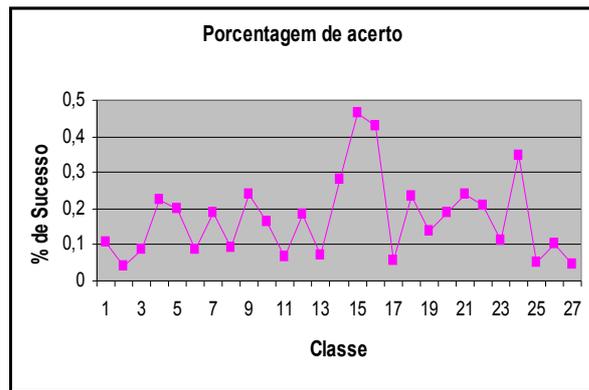
Figura 4.4 – Imagens da base de dados. (a) Imagens referentes a classe 5 (árvore) e (b) Imagens da classe 19 (janela).

Essas imagens apresentam cenas escuras, nesse caso o classificador identificou a mesma cor pura da imagem nas duas classes, visto que as características selecionadas através dos coeficientes *wavelet* de aproximação resultam em informações significativas para a interseção dos histogramas (Seção 3.7) realizada em cada parte da distribuição espacial (Seção 3.5), resultando em respostas para o cálculo da distância euclidiana (Seção 3.8). Essas respostas geram os resultados para o cálculo da medida de efetividade (Seção 3.10) visto no Gráfico 4.1 (b), (d), (f) da porcentagem de acerto. A classificação ficou acima dos 45% no 1º, 2º e 3º níveis de decomposição.

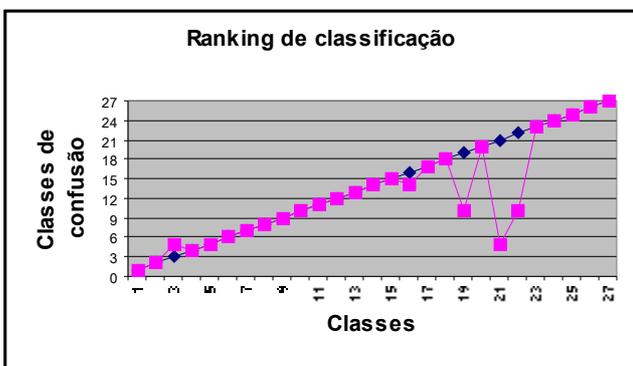
Os resultados dos experimentos com o canal S, nos três níveis de decomposição, são apresentados ilustrativamente pelo Gráfico 4.2.



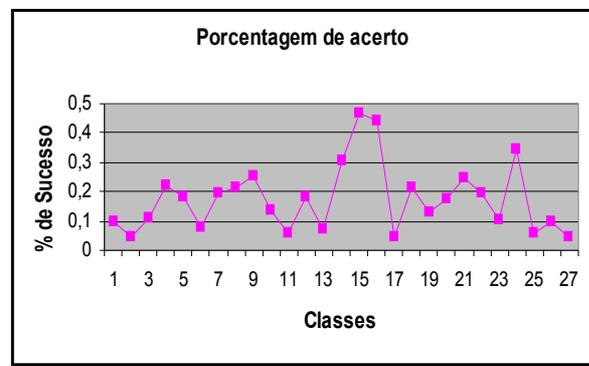
(a) Decomposição no 1º nível com os coeficientes de aproximação do canal S.



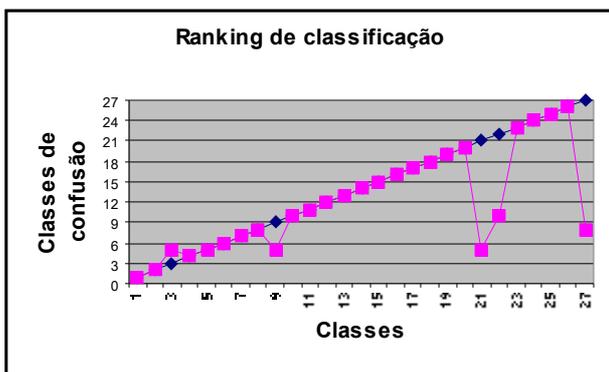
(b) Gráfico referente a porcentagem de acerto no 1º nível.



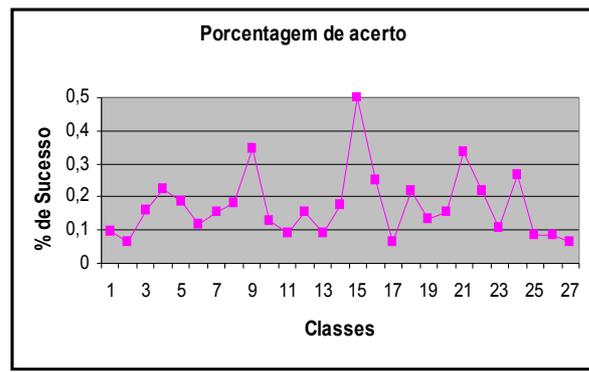
(c) Decomposição no 2º nível com os coeficientes de aproximação do canal S.



(d) Gráfico referente a porcentagem de acerto no 2º nível.



(e) Decomposição no 3º nível com os coeficientes de aproximação do canal S.



(f) Gráfico referente a porcentagem de acerto no 3º nível.

Gráfico 4.2 - Gráficos dos resultados do *ranking* de classificação e porcentagem de acerto do canal S, coeficientes de aproximação. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.

Existe uma pequena diferença entre as classes que foram identificadas erroneamente. A classe 3 (animais) foi identificada como sendo a classe 10 (construção) apenas no 1º nível, e os coeficientes de aproximação ficaram mais significativos nesse nível, isso aconteceu porque as imagens possuem uma combinação de cores muito parecidas, Figura 4.5. Já as classes 21 (mamíferos) e 22 (mosaico) foram identificadas como sendo as classes 5 (árvore) e 10 (construção), respectivamente, Figura 4.6 e Figura 4.7.



Figura 4.5 – Imagens da base de dados. (a) Imagens referentes a classe 10 (construção) e (b) Imagens da classe 3 (animais).

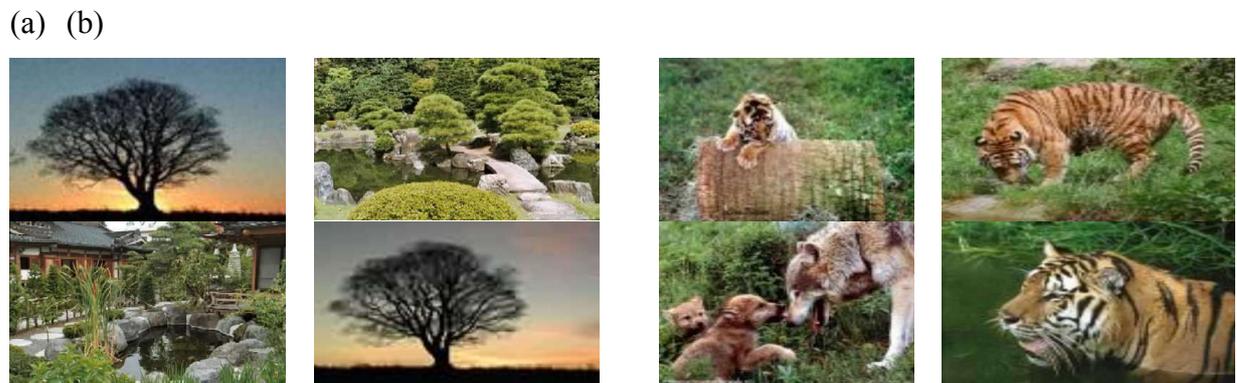


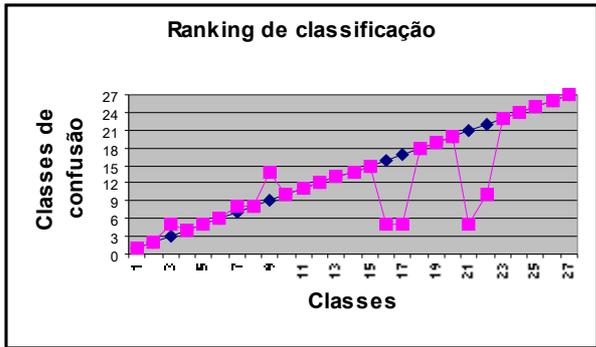
Figura 4.6 – Imagens da base de dados. (a) Imagens referentes a classe 5 (árvore) e (b) Imagens da classe 21 (mamíferos).



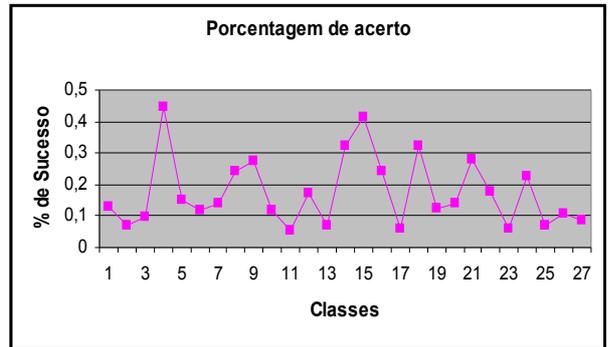
Figura 4.7 - Imagens da base de dados referentes a classe 22 (mosaico).

As características salientes das classes 21 (mamíferos) e 5 (árvore) apresentaram a mesma quantidade de tons em verde sendo que a saturação está mais significativa nessa área, fazendo com que os cálculos da interseção dos histogramas de cada distribuição espacial das imagens apontassem para a classificação dessas classes. A porcentagem de acerto da classificação ficou acima dos 47%, chegando a 50% no 3º nível, Gráfico 4.2 (f). Isso significa que trabalhando com um pequeno número de características obtêm-se resultados significativos, pois as características que não interessam são excluídas e utiliza-se somente as que são significativas.

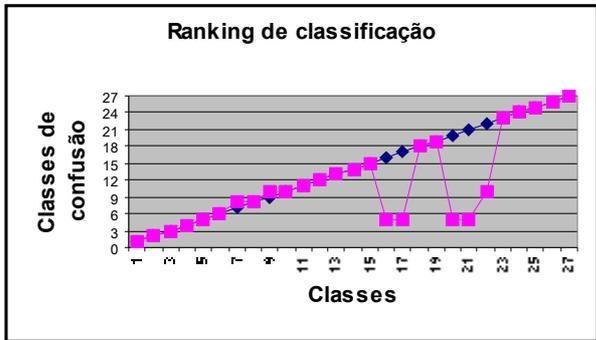
Os gráficos (Gráfico 4.3) correspondem ao canal V das imagens no 1º, 2º e 3º nível de decomposição com os coeficientes de aproximação.



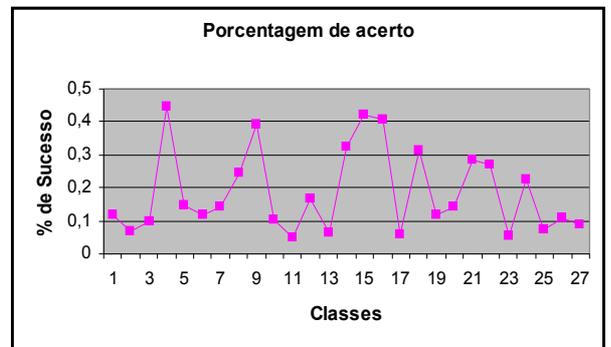
(a) Decomposição no 1º nível com os coeficientes de aproximação do canal V.



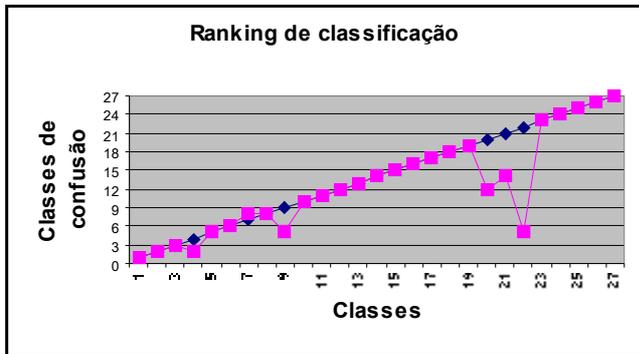
(b) Gráfico referente a porcentagem de acerto no 1º nível.



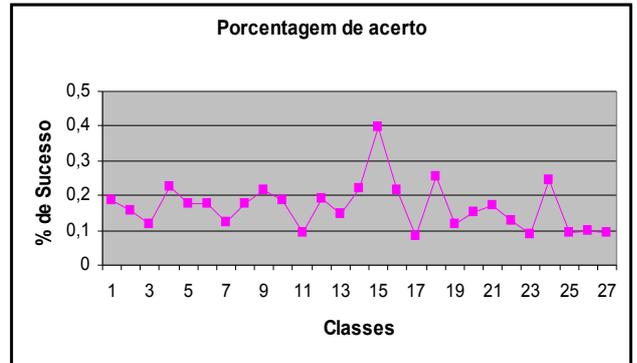
(c) Decomposição no 2º nível com os coeficientes de aproximação do canal V.



(d) Gráfico referente a porcentagem de acerto no 2º nível.



(e) Decomposição no 3º nível com os coeficientes de aproximação do canal V.



(f) Gráfico referente a porcentagem de acerto no 3º nível.

Gráfico 4.3 – Gráficos dos resultados do *ranking* de classificação e porcentagem de acerto do canal V, coeficientes de aproximação. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.

Neste estágio do algoritmo, a intensidade de luz presente no canal V influenciou as classes 16 (fruta), 17 (garota) e 21 (mamíferos) que foram identificadas com a classe 5 (árvore) que apresenta praticamente a mesma quantidade de luz nas imagens, isso influenciou a classificação, como pode ser visto na Figura 4.8 e Figura 4.9.



Figura 4.8 – Imagens da base de dados. (a) Imagens referentes a classe 16 (fruta) e (b) Imagens da classe 17 (garotas).

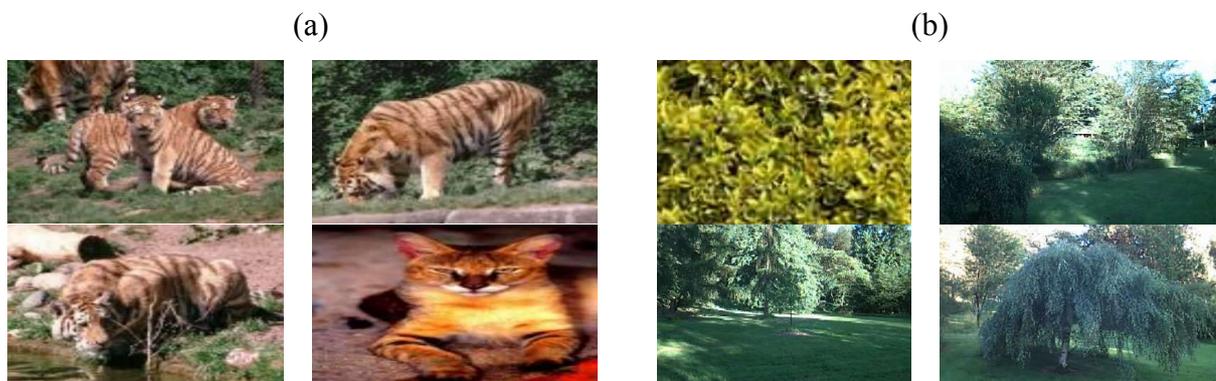
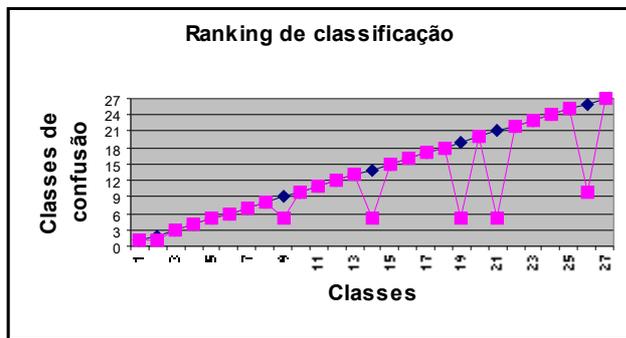


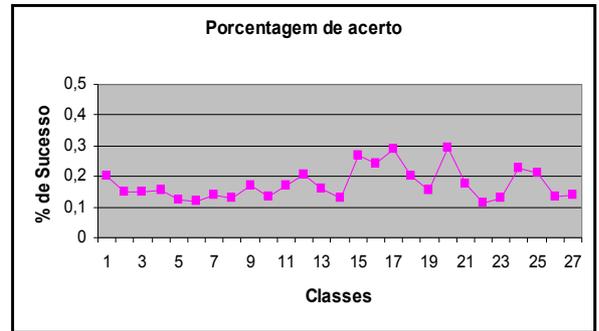
Figura 4.9 – Imagens da base de dados. (a) Imagens referentes a classe 21 (mamíferos) e (b) Imagens da classe 5 (árvores).

A porcentagem de acerto, Gráficos 4.3 (b) e (d), apresentou o sucesso da classificação ficando acima dos 43% no 1º e 2º nível de decomposição, as características extraídas com o canal V com os coeficientes de aproximação necessitam mais informações para ter uma taxa de classificação significativa. Com poucas informações, como mostra o Gráfico (f) no 3º nível de decomposição, os resultados chegaram até os 40%. Isso resulta que com os coeficientes de aproximação no canal V precisa-se de uma quantidade significativa de características para se recuperar as imagens corretamente.

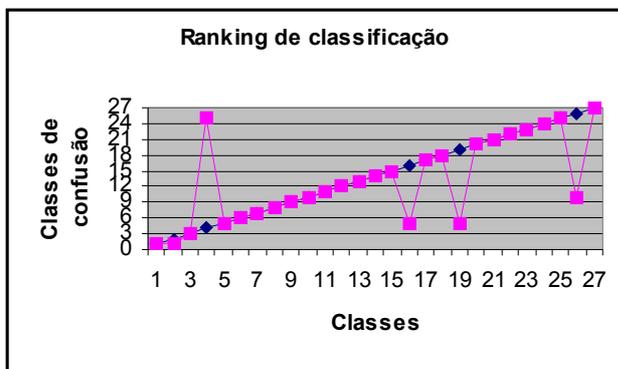
Os gráficos a seguir, apresentam os resultados com os experimentos referentes aos coeficientes de detalhe horizontal nos canais H, S e V. Nos três níveis de decomposição.



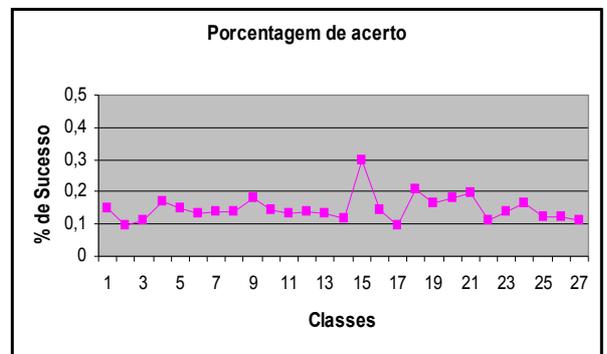
(a) Decomposição no 1º nível com os coeficientes de detalhe horizontal do canal H.



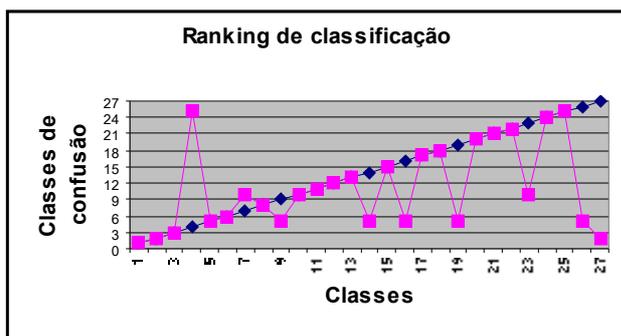
(b) Gráfico referente a porcentagem de acerto no 1º nível.



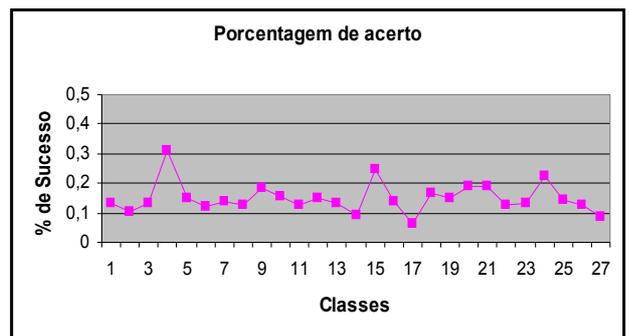
(c) Decomposição no 2º nível com os coeficientes de detalhe horizontal do canal H.



(d) Gráfico referente a porcentagem de acerto no 2º nível.



(e) Decomposição no 3º nível com os coeficientes de detalhe horizontal do canal H.



(f) Gráfico referente a porcentagem de acerto no 3º nível.

Gráfico 4.4 – Gráficos dos resultados do *ranking* de classificação e porcentagem de acerto do canal H, coeficientes de detalhe horizontal. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.

Os experimentos com os coeficientes de detalhe horizontal, do canal H correspondem ao Gráfico 4.4. Vale ressaltar que os coeficientes de alta frequência são subamostragens da imagem original, porém contêm detalhes que, somados aos coeficientes de aproximação reconstróem a imagem original. Sendo assim, as características salientes que foram extraídas nas imagens tiveram a classificação da classe 16 (fruta) como sendo a classe 5 (árvore), Figura 4.10.



Figura 4.10 – Imagens da base de dados. (a) Imagens referentes a classe 16 (frutas) e (b) Imagens da classe 5 (árvores).

Como mostra a Figura 4.10, as classes 16 (fruta) e 5 (árvore) são visivelmente semelhantes, devido possuírem tonalidades da cor verde nas imagens e texturas muito parecidas, com isso, o classificador identificou sendo imagens da mesma classe.

Pode-se perceber ainda que a classe 26 (neve) foi classificada como sendo a classe 10 (construção), Figura 4.11.

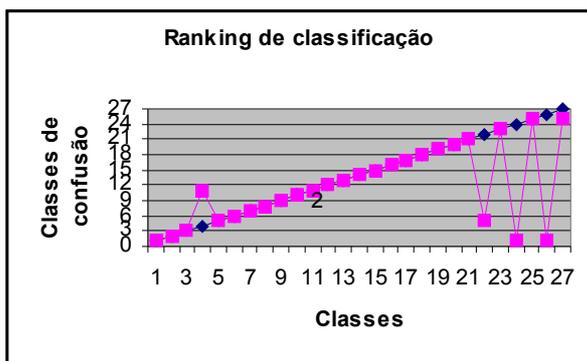


Figura 4.11 – Imagens da base de dados. (a) Imagens referentes a classe 10 (construção) e (b) Imagens da classe 26 (neve).

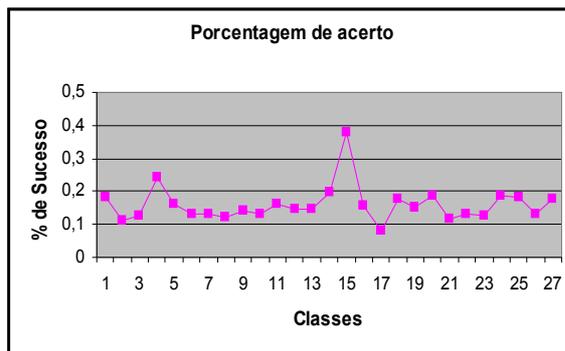
Essas imagens têm em comum a pureza da cor azul misturada com a cor branca, por isso as classes foram identificadas dessa forma. Já a porcentagem de acerto com esses

coeficientes não indicou muito sucesso, ficando com taxa de classificação em 30% nos três níveis de decomposição.

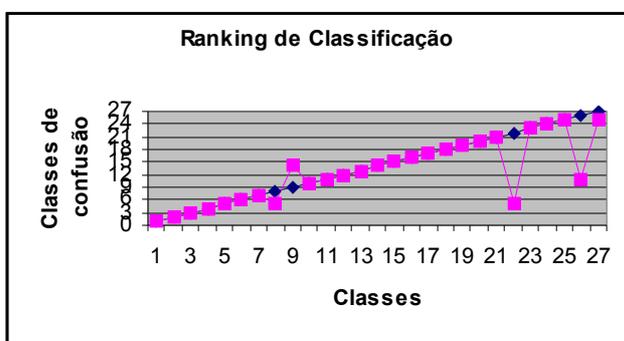
As ilustrações do Gráfico 4.5 correspondem aos experimentos realizados com os coeficientes de detalhe horizontal do canal S. Esse canal indica o afastamento da cor, ou seja, é a quantidade de cor branca misturada a outras cores.



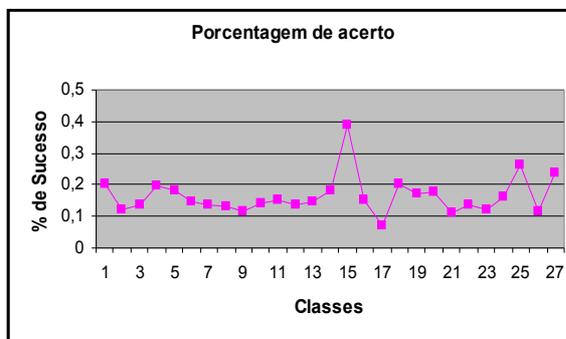
(a) Decomposição no 1º nível com os coeficientes de detalhe horizontal do canal S.



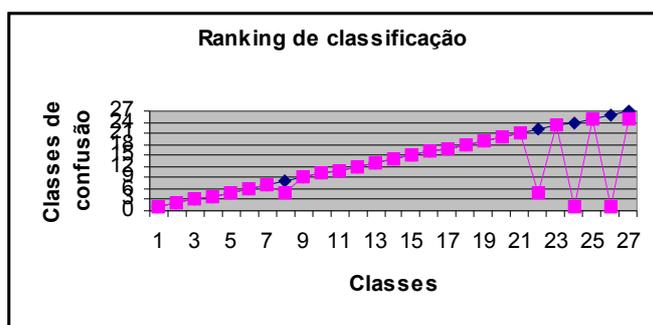
(b) Gráfico referente a porcentagem de acerto no 1º nível.



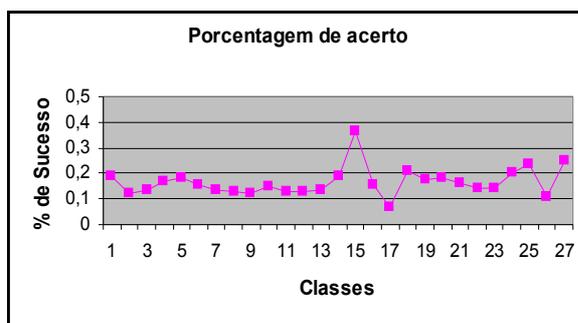
(c) Decomposição no 2º nível com os coeficientes de detalhe horizontal do canal S.



(d) Gráfico referente a porcentagem de acerto no 2º nível.



(e) Decomposição no 3º nível com os coeficientes de detalhe horizontal do canal S.



(f) Gráfico referente a porcentagem de acerto no 3º nível.

Gráfico 4.5 – Gráficos dos resultados do *ranking* de classificação e porcentagem de acerto do canal S, coeficientes de detalhe horizontal. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.

Pode-se perceber que a classe 22 (mosaico) foi identificada como sendo a classe 5 (árvore), Figura 4.12. E as classes 24 (ponte) e 26 (neve) foram identificadas como sendo a classe 1 (água), Figura 4.13 e Figura 4.14.



Figura 4.12 - Imagens da base de dados. (a) Imagens referentes a classe 22 (mosaico) e (b) Imagens da classe 5 (árvore).



Figura 4.13 – Imagens da base de dados. (a) Imagens referentes a classe 24 (ponte) e (b) Imagens da classe 26 (neve).

Imagens referentes a classe 1 (água)

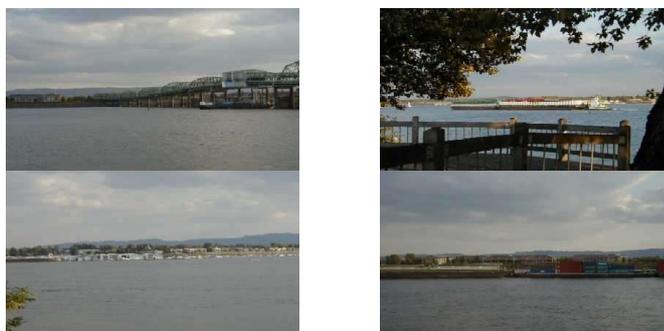
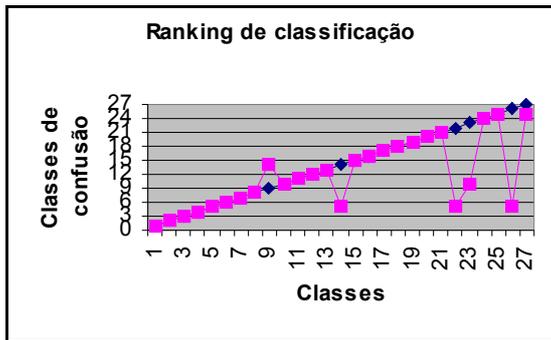


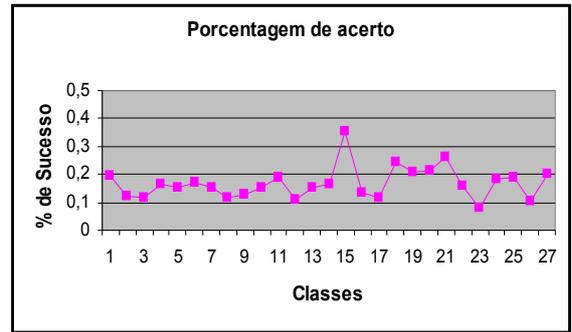
Figura 4.14 – Imagens da base de dados referentes a classe 1 (água).

A classificação das classes chegou aos 40% no 1º e 2º nível de decomposição, aumentando um pouco a porcentagem em relação ao canal H com os coeficientes de detalhe horizontal.

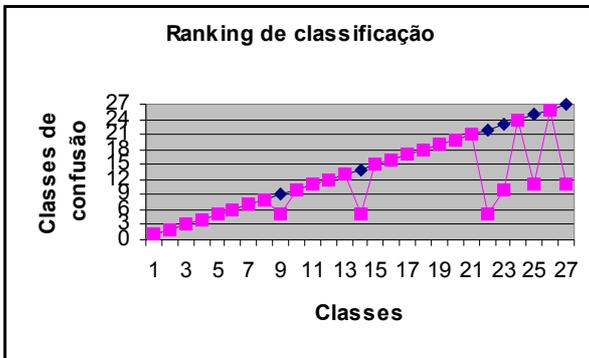
A ilustração seguinte apresenta os resultados dos experimentos do canal V com os coeficientes de detalhe horizontal, Gráfico 4.6.



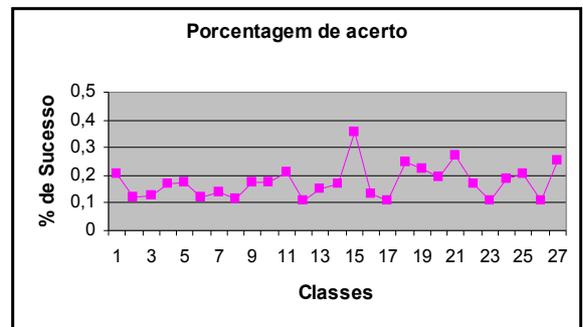
(a) Decomposição no 1º nível com os coeficientes de detalhe horizontal do canal V.



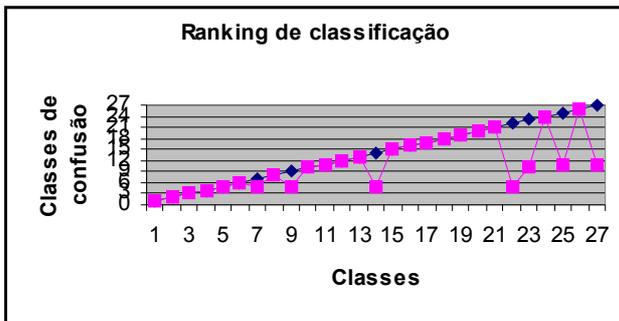
(b) Gráfico referente a porcentagem de acerto no 1º nível.



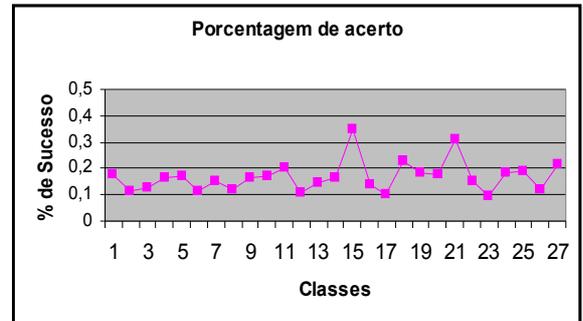
(c) Decomposição no 2º nível com os coeficientes de detalhe horizontal do canal V.



(d) Gráfico referente a porcentagem de acerto no 2º nível.



(e) Decomposição no 3º nível com os coeficientes de detalhe horizontal do canal V.



(f) Gráfico referente a porcentagem de acerto no 3º nível.

Gráfico 4.6 – Gráficos dos resultados do *ranking* de classificação e porcentagem de acerto do canal V, coeficientes de detalhe horizontal. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.

Observando esses gráficos é possível perceber que as imagens das classes 25 (céu) e 27 (textura) foram identificadas como sendo a classe 11 (amanhecer), Figura 4.15 e Figura 4.16. E a classe 22 (mosaico) foi identificada como sendo a classe 5 (árvore), Figura 4.17.

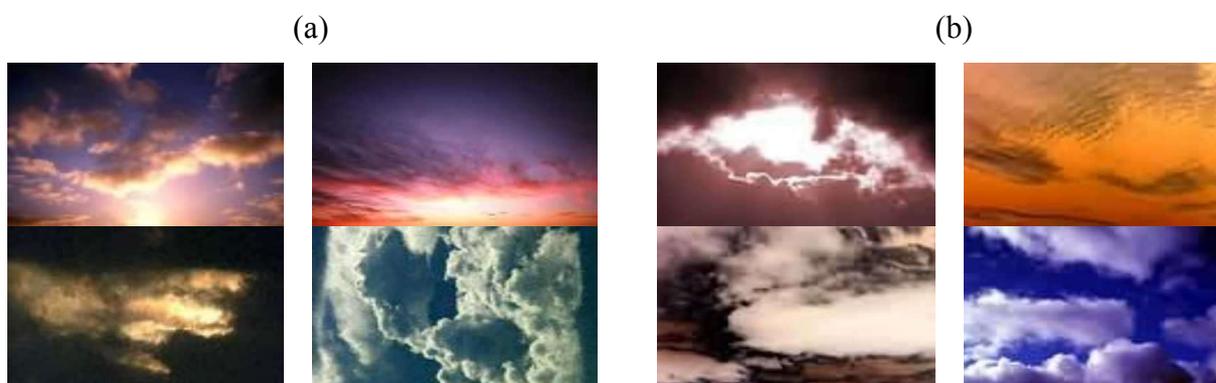


Figura 4.15 – Imagens da base de dados. (a) Imagens referentes a classe 25 (céu) e (b) Imagens da classe 27 (textura).



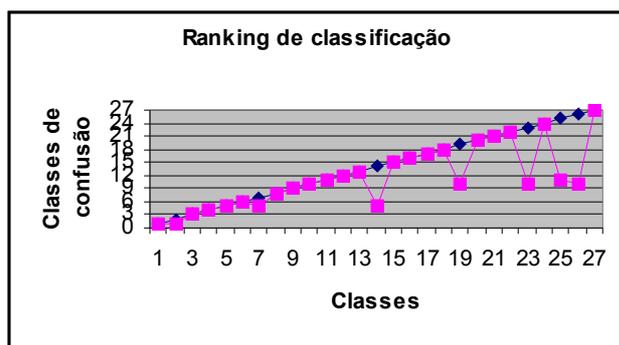
Figura 4.16 – Imagens da base de dados referentes a classe 11 (amanhecer).



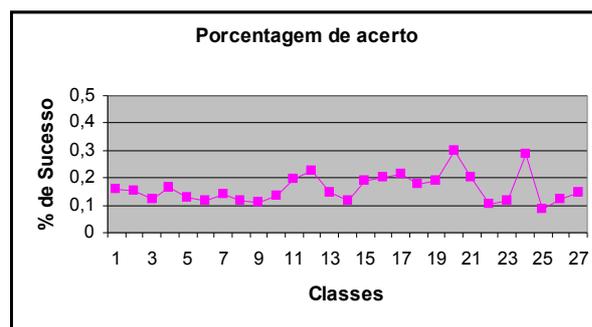
Figura 4.17 – Imagens da base de dados. (a) Imagens referentes a classe 22 (mosaico) e (b) Imagens da classe 5 (árvore).

Os coeficientes mais significativos, extraídos pela decomposição *wavelet*, levam em consideração a intensidade de luz que é representada pelo canal V, categorizando que essas imagens fossem da mesma classe. Com isso as imagens dessas classes foram identificadas conforme visto no Gráfico 4.6. Já a porcentagem de acerto da classificação se mostrou abaixo dos 40% no 1º, 2º e 3º nível de decomposição.

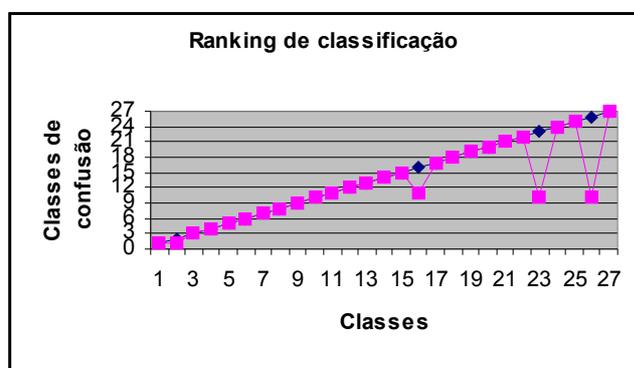
Os gráficos a seguir representam os resultados referentes a classificação de acerto do canal H com os coeficientes de detalhe vertical.



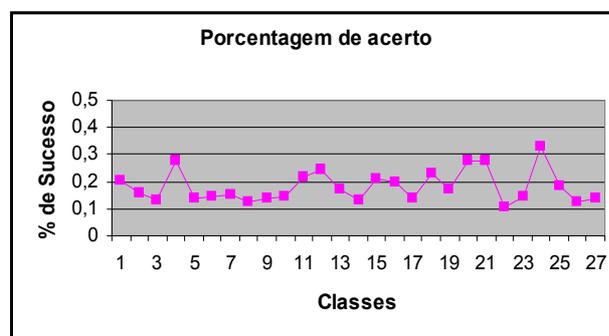
(a) Decomposição no 1º nível com os coeficientes de detalhe vertical do canal H.



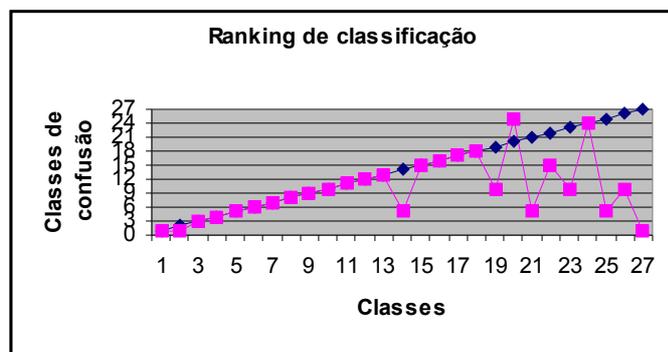
(b) Gráfico referente a porcentagem de acerto no 1º nível.



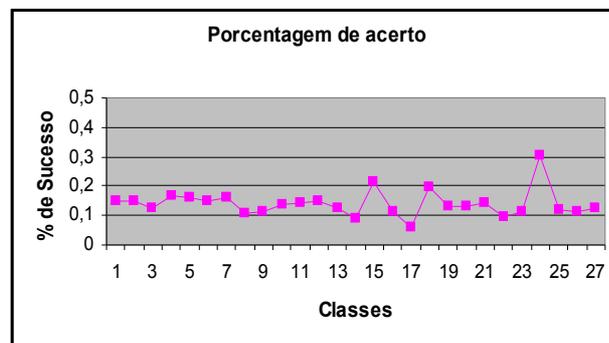
(c) Decomposição no 2º nível com os coeficientes de detalhe vertical do canal H.



(d) Gráfico referente a porcentagem de acerto no 2º nível.



(e) Decomposição no 3º nível com os coeficientes de detalhe vertical do canal H.



(f) Gráfico referente a porcentagem de acerto no 3º nível.

Gráfico 4.7 – Gráficos dos resultados do *ranking* de classificação e porcentagem de acerto do canal H, coeficientes de detalhe vertical. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.

O Gráfico 4.7 elucidava os resultados dos experimentos realizados do canal H com os coeficientes de detalhe vertical. Pode-se perceber nos resultados que as classes 23 (montanha) e 26 (neve) foram identificadas como sendo a classe 10 (construção), isso ocorreu nos três níveis de decomposição, Figura 4.18 e Figura 4.19.



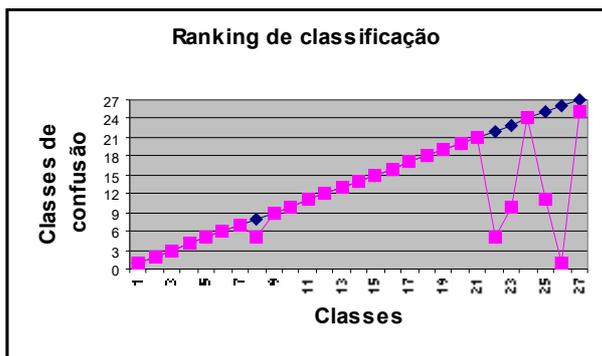
Figura 4.18 – Imagens da base de dados. (a) Imagens referentes a classe 23 (montanha) e (b) Imagens da classe 26 (neve).



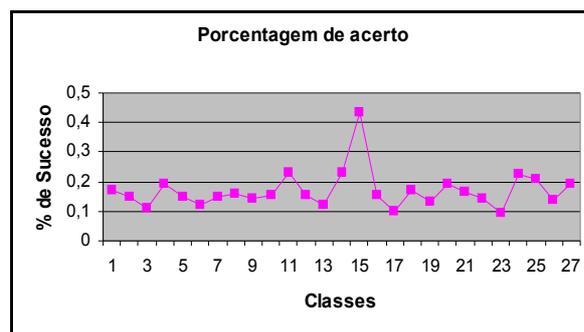
Figura 4.19 – Imagens da base de dados referentes a classe 10 (construção).

As imagens dessas classes apresentam semelhanças em suas cenas, visto que as imagens das casas presentes na classe 10 (construção) têm o mesmo formato que as montanhas na classe 23 (montanha) e algumas elevações na classe 26 (neve). Com isso, os coeficientes extraídos geraram respostas que foram caracterizadas havendo alguma semelhança com essas imagens. Esses resultados não foram considerados significativos, devido a baixa taxa de classificação das classes na porcentagem de acerto.

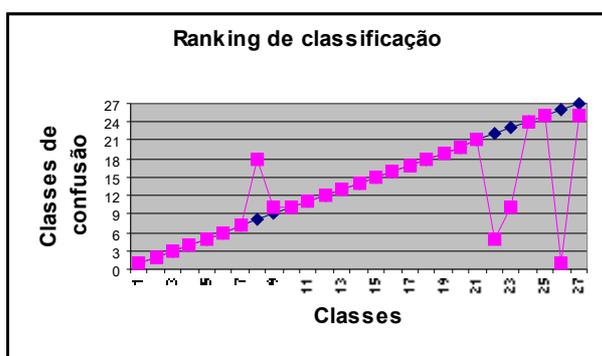
A seguir são apresentados os resultados do canal S com os coeficientes de detalhe vertical.



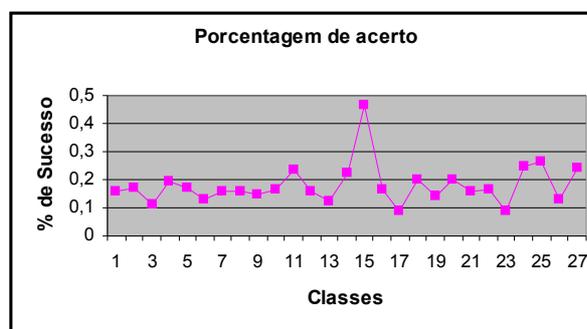
(a) Decomposição no 1º nível com os coeficientes de detalhe vertical do canal S.



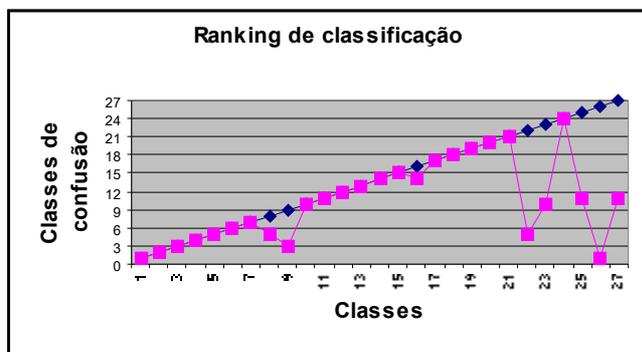
(b) Gráfico referente a porcentagem de acerto no 1º nível.



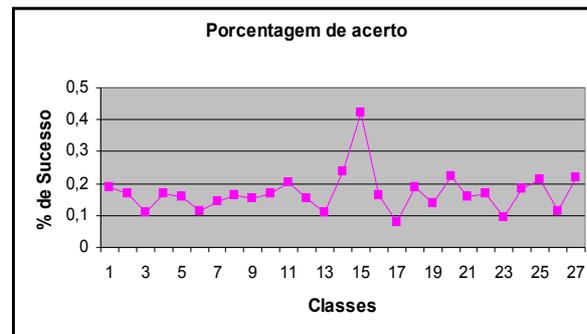
(c) Decomposição no 2º nível com os coeficientes de detalhe vertical do canal S.



(d) Gráfico referente a porcentagem de acerto no 2º nível.



(e) Decomposição no 3º nível com os coeficientes de detalhe vertical do canal S.



(f) Gráfico referente a porcentagem de acerto no 3º nível.

Gráfico 4.8 – Gráficos dos resultados do *ranking* de classificação e porcentagem de acerto do canal S, coeficientes de detalhe vertical. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.

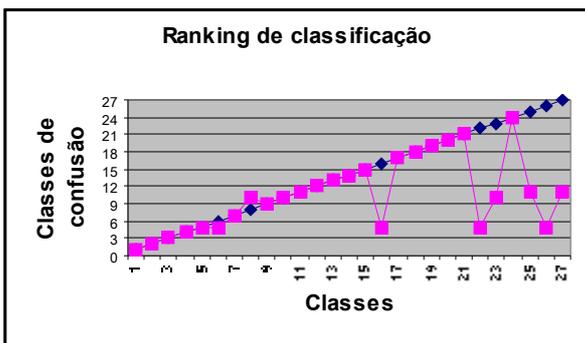
O *ranking* de classificação apresentado nos Gráficos 4.8 (a), (c), (e), apresentam os resultados dos experimentos do canal S com os coeficientes de detalhe vertical. As características significativas representadas pela saturação do canal S dessas imagens contribuíram para que a classificação da classe 25 (céu) fosse identificada como sendo parte da classe 11 (amanhecer), Figura 4.20. Isso ocorreu no 1º e 3º nível de decomposição.



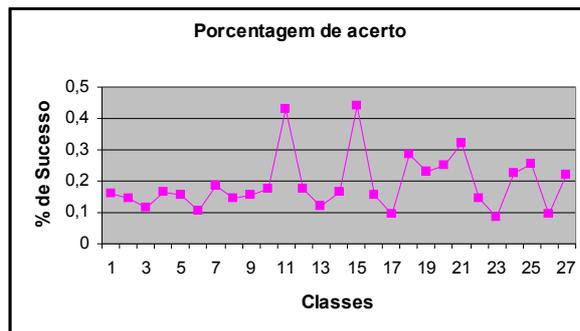
Figura 4.20 – Imagens da base de dados. (a) Imagens referentes a classe 25 (céu) e (b) Imagens da classe 11 (amanhecer).

O resultado da classificação aconteceu porque as imagens possuem muitos pontos em comum nas cenas, tais como nuvens e texturas variadas, isso influenciou na classificação. A porcentagem de acerto das classes apresentou melhores resultados comparados com os do canal H, ficando acima dos 40% de classificação apenas para a classe 15 nos três níveis de decomposição.

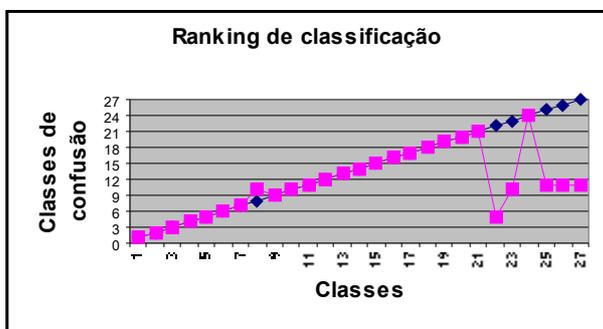
Os resultados do canal V com os coeficientes de detalhe vertical nos três níveis de decomposição são apresentados nos gráficos a seguir.



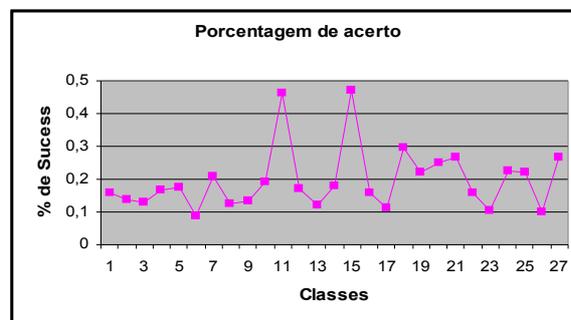
(a) Decomposição no 1º nível com os coeficientes de detalhe vertical do canal V.



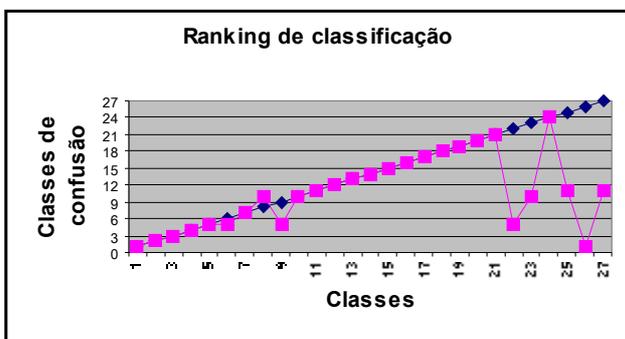
(b) Gráfico referente a porcentagem de acerto no 1º nível.



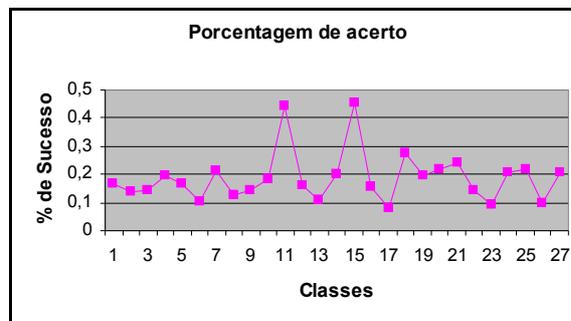
(c) Decomposição no 2º nível com os coeficientes de detalhe vertical do canal V.



(d) Gráfico referente a porcentagem de acerto no 2º nível.



(e) Decomposição no 3º nível com os coeficientes de detalhe vertical do canal V.



(f) Gráfico referente a porcentagem de acerto no 3º nível.

Gráfico 4.9 – Gráficos dos resultados do *ranking* de classificação e porcentagem de acerto do canal V, coeficientes de detalhe vertical. (a) e (b) Decomposição no 1º nível; (c) e (d) Decomposição no 2º nível; (e) e (f) Decomposição no 3º nível.

Os Gráficos 4.9 (a), (c) e (e) representam a classificação das imagens do canal V com os coeficientes de detalhe vertical. A classificação das imagens teve as classes 25 (céu), 26 (neve) e 27 (textura) classificadas como sendo a classe 11 (amanhecer), Figura 4.21 e Figura 4.22.

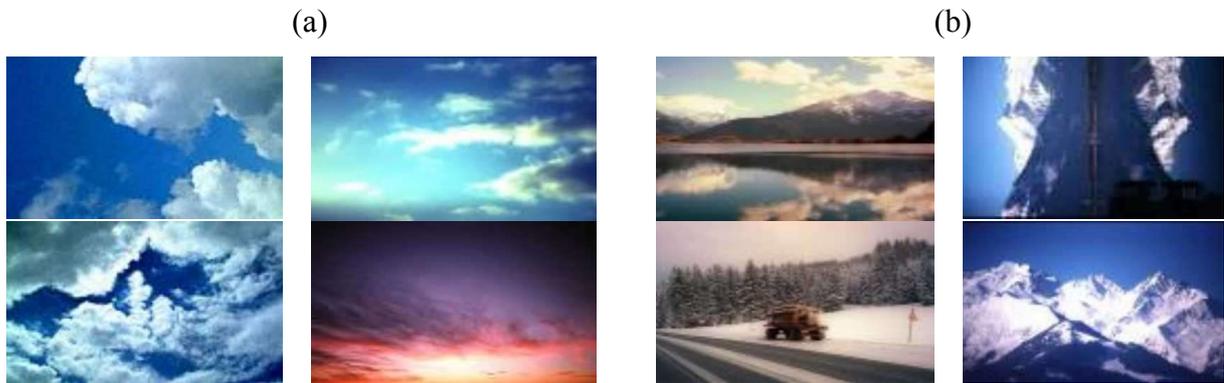


Figura 4.21 – Imagens da base de dados. (a) Imagens referentes a classe 25 (céu) e (b) Imagens da classe 26 (neve).

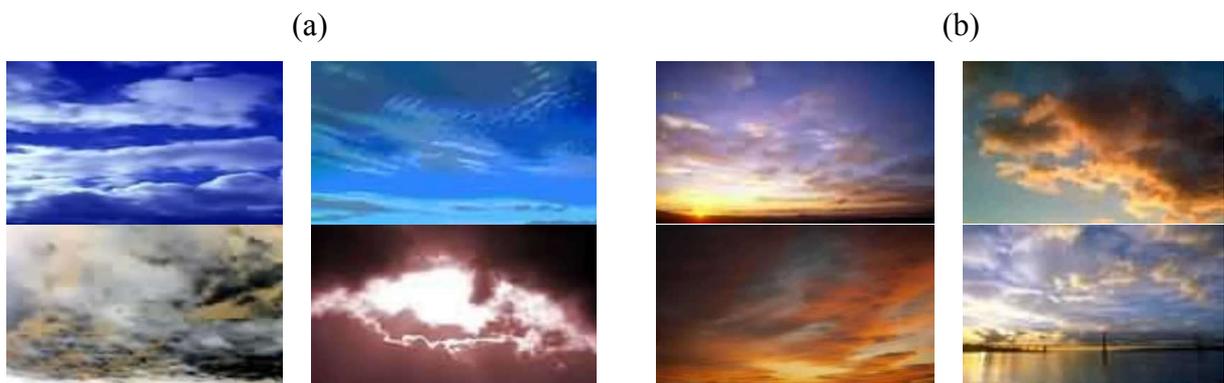


Figura 4.22 - Imagens da base de dados. (a) Imagens referentes a classe 27 (textura) e (b) Imagens da classe 11 (amanhecer).

Esse resultado ocorreu porque as imagens representadas pela intensidade de luz do canal V influenciaram para que os coeficientes mais significativos gerassem esse resultado. Já a porcentagem de acerto melhorou bastante comparada com os canais H e S anteriormente citados, chegando a classificação acima de 45% no 1º, 2º e 3º nível de decomposição.

4.2. Análise da Combinação dos Resultados

Para analisar com maior precisão a classificação das imagens, foi feita uma combinação dos conjuntos de coeficientes que obtiveram melhores resultados, sendo eles os coeficientes de aproximação dos canais H e V e os coeficientes de detalhe vertical no canal S, todos no 3º nível de decomposição.

A combinação dos conjuntos de coeficientes com os melhores resultados seguiu o mesmo processamento dos experimentos feito até o momento. Os resultados apresentados com essa combinação tiveram classificações satisfatórias comparados com os resultados anteriores.

De acordo com a análise da matriz de confusão, pode-se perceber que considerando o Gráfico 4.10, no *ranking* de classificação, com o somatório dos coeficientes de aproximação no canal H, coeficientes de aproximação no canal V e coeficientes de detalhe vertical no canal S, no 3º nível de decomposição, as classes 9 (concreto), 22 (mosaico) e 26 (neve) foram as que tiveram maior incerteza perante o classificador, Figura 4.23 e Figura 4.24.



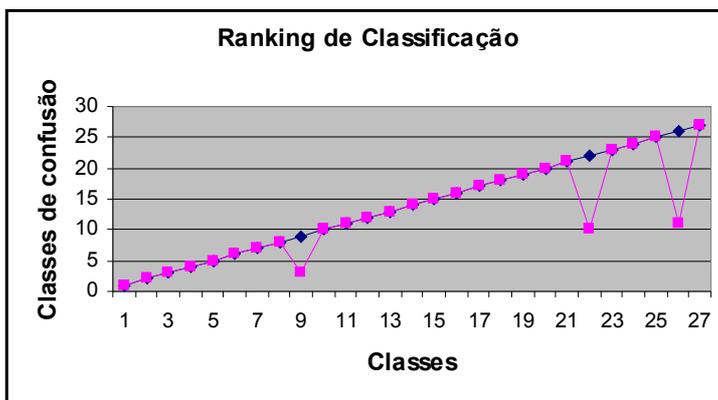
Figura 4.23 – Imagens da base de dados. (a) Imagens referentes a classe 9 (concreto) e (b) Imagens da classe 22 (mosaico).



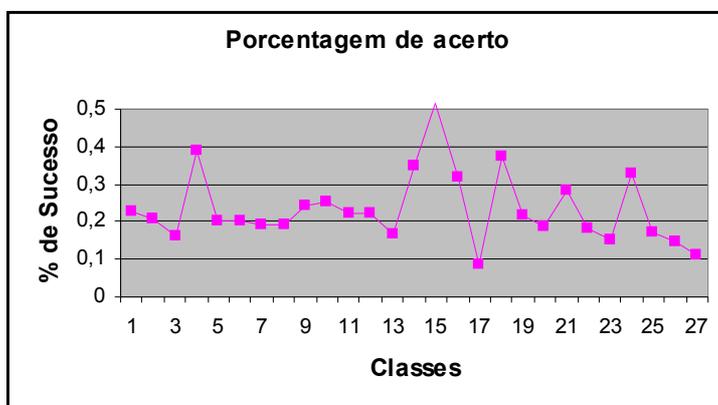
Figura 4.24 – Imagens da base de dados referentes a classe 26 (neve).

Na porcentagem de acerto, as classes obtiveram melhor porcentagem de sucesso, chegando a mais de 50%, quando combinando os coeficientes no 1º, 2º e 3º nível de decomposição.

Com a análise dos resultados do Gráfico 4.10 pode-se perceber que combinando o conjunto de coeficientes dos canais que tiveram melhores resultados, foram mais significativos que os resultados calculados em separado.



(a)



(b)

Gráfico 4.10 – Gráfico com a combinação dos canais H, S e V com os coeficientes de aproximação e de detalhe vertical; (a) Gráfico do ranking de classificação; (b) Gráfico da porcentagem de acerto.

Novos experimentos foram feitos e optou-se em fazer testes com uma parte dos coeficientes, porque até o momento os experimentos estavam sendo feitos usando todos os coeficientes das imagens, assim com os novos testes usou-se uma porção dos coeficientes mais altos, ou seja, as características com valores mais elevados. Para cada decomposição feita os coeficientes foram colocados em um vetor para se obter todas as características nos três níveis de decomposição, utilizando o limiar de 10% dos primeiros coeficientes. Esse limiar foi usado porque foram feitos testes com 20% e 30% das características e não apresentaram resultados significativos como os apresentados pelos 10% dos coeficientes.

Os resultados com esse novo teste foram considerados promissores, portanto, com a utilização de uma parte das características os resultados tiveram uma boa porcentagem na classificação das imagens. Esses resultados serão apresentados a seguir.

Nessa rodada de testes, os experimentos foram feitos com as métricas de distância euclidiana e somatório. Os resultados foram bem sucedidos com a métrica do somatório, tendo porcentagem de classificação acima da métrica distância euclidiana, com os coeficientes de aproximação, detalhe horizontal e detalhe vertical no canal V os resultados se mostraram acima de 60% na porcentagem de acerto para apenas algumas das classes, como mostra os gráficos abaixo.

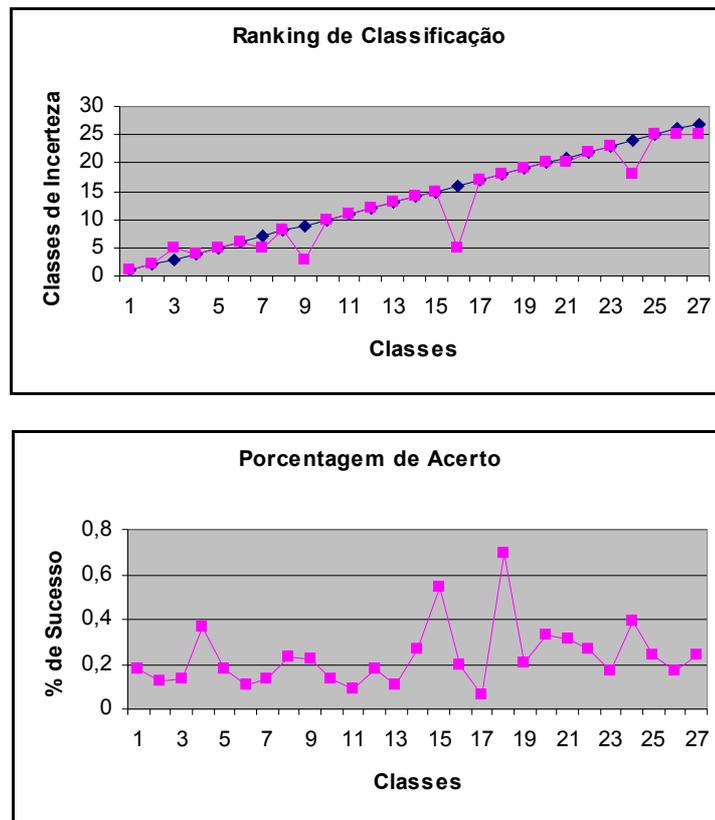


Gráfico 4.11 – Gráfico com os Coeficientes de Aproximação Canal V – Somatório.

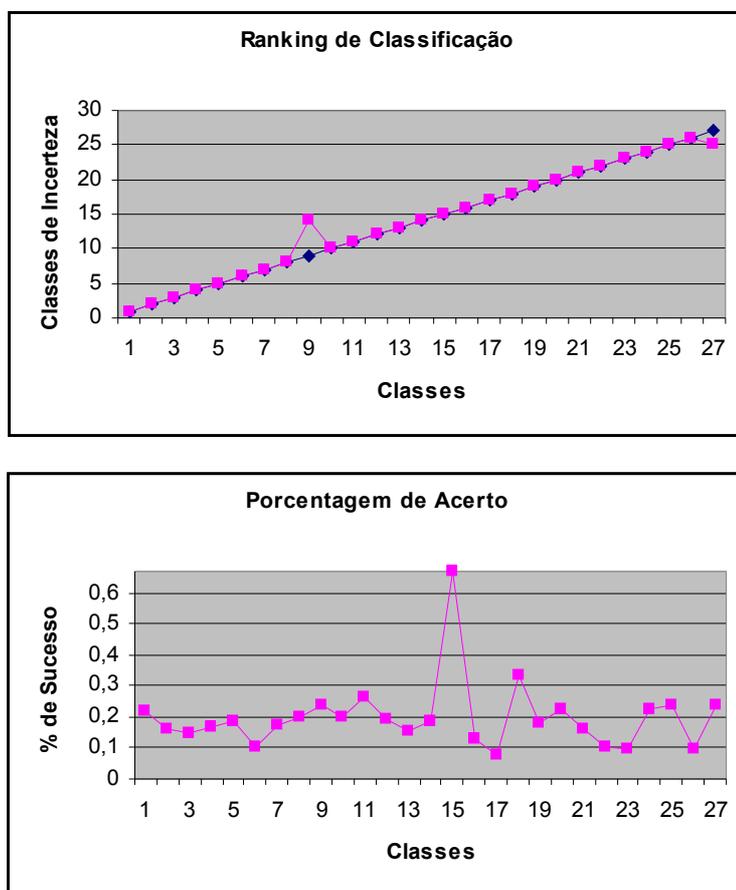


Gráfico 4.12 – Gráfico com os Coeficientes Detalhe Horizontal Canal V – Somatório.

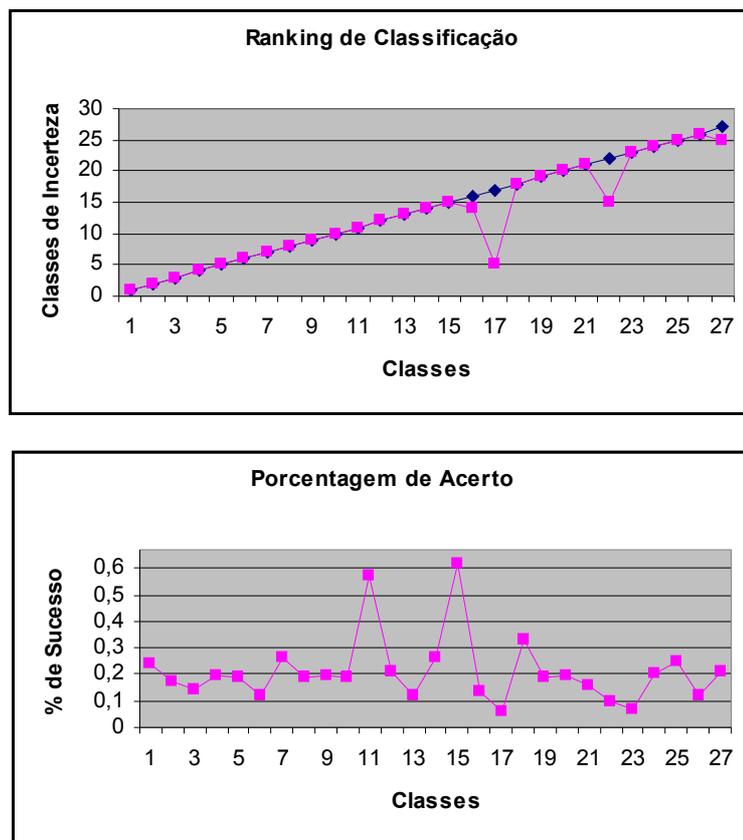


Gráfico 4.13 – Gráfico com os Coeficientes Detalhe Vertical Canal V – Somatório.

No *ranking* de classificação poucas classes foram recuperadas erroneamente, tanto com a métrica distância euclidiana quanto com o somatório, os resultados com os coeficientes de aproximação, detalhe horizontal e detalhe vertical, a classe 27 (textura) foram identificadas como sendo a classe 25 (céu). São classes de percepção muito parecidas, por isso houve erro de classificação.

Entretanto, com as taxas de acerto atingidas, os experimentos comprovam a viabilidade do método de recuperação de imagens, dada a vasta quantidade de tipos diferentes de imagens potencialmente disponíveis.

4.3. Considerações Finais

Analisando de uma forma geral os resultados da matriz de confusão, a classe 5 (árvore) é a que mais foi identificada com as demais classes nos três níveis de decomposição e com os três canais. Pois, apresenta imagens variadas, tais como cenas escuras, cenas claras, cenas com luminosidade acentuada, cores mais fortes em determinadas cenas, fazendo com que fosse classificada com a maioria das imagens da base de dados.

A recuperação das imagens utilizando o método proposto apresentou-se satisfatório, sendo que melhorias devem ser feitas permitindo assim uma busca mais refinada nas imagens. Poucas técnicas na literatura tratam a questão de recuperar informação priorizando combinar as características cor e textura e de utilizar corretamente as informações nas imagens, e este trabalho visa preencher essa questão. É importante ressaltar que o modelo de cor HSV está sendo usado, visto que é considerado um dos modelos eficientes para recuperar imagens, tendo a possibilidade de separar a intensidade da informação tonalidade e saturação, bem como, na relação que existe entre essas componentes, muito próxima da forma na qual o homem percebe a cor.

O Capítulo 4 apresentou os experimentos realizados, bem como as considerações sobre as imagens e classes definidas. O próximo capítulo traz as conclusões finais do trabalho, bem como algumas sugestões para trabalhos futuros.

Capítulo 5

Conclusões

Para definir a similaridade em recuperação de informação, são desenvolvidos métodos que calculam a diferença entre imagens que estão sendo comparadas. Como foi apresentada nesta dissertação, a tarefa de recuperar imagens pode ser aplicada em diferentes áreas, tais como: Engenharia e Arquitetura; Sistemas de informação Geográfica; em base de imagens médicas; dentre outras. Assim, muitos esforços foram direcionados em pesquisas e desenvolvimento dos CBIRS.

Este trabalho apresentou um método baseado em uma nova forma de indexação para recuperação de imagens, levando-se em consideração a extração das características cor e textura, combinados em um espaço *wavelet*. Isso foi realizado com a inclusão de técnicas de recuperação de informação que permitiram tratar as imagens de forma natural.

A escolha por essas características foi atribuída ao fato da cor representar a imagem como um todo, proporcionando informações relevantes para a tarefa de recuperação. Já a característica textura por estar intimamente relacionada com a característica cor, a qual foi tratada em diferentes níveis de multirresolução.

Levando em consideração os objetivos apresentados neste trabalho, concluiu-se que o modelo de cor HSV usado como forma de representar a cor das imagens tem a vantagem de ser um modelo que separa a intensidade da informação tonalidade e saturação, bem como, na relação que existe entre essas componentes, muito próxima da forma na qual o homem percebe a cor.

A base Haar usada buscou realçar as melhores características das imagens trabalhadas, representando o sinal de forma redundante, pois não se sabia quais características eram mais significativas para se fazer a classificação e isso foi evidenciado através da extração das características, apresentada na Seção 3.6.

A associação da informação espacial ao atributo cor e textura foi realizada através da multirresolução das imagens. Isso foi alcançado através da decomposição *wavelet* que permitiu a transformação das imagens em diferentes escalas.

A distância euclidiana influenciou nos resultados devido ao fato de trabalhar com a distância entre *pixels* de mesma posição espacial, resultando em um valor correspondente à diferença quadrática entre *pixels* de mesma localização espacial nas imagens. Já a métrica somatório mostrou-se um pouco melhor utilizando uma menor quantidade de informações dos coeficientes, comparada com a distância euclidiana, influenciando na classificação dos resultados.

Os experimentos com a combinação dos resultados, Seção 4.3 foram bem sucedidos com a métrica do somatório, tendo porcentagem de classificação acima da métrica distância euclidiana, com os coeficientes de aproximação, detalhe horizontal e detalhe vertical no canal V os resultados se mostraram acima de 60% na porcentagem de acerto.

A dimensionalidade do espaço de características foi reduzida e os experimentos realizados mostraram-se significativos. Trabalhando-se com o terceiro nível de decomposição os resultados apresentaram uma melhor classificação, ou seja, com poucas informações pôde-se obter melhores resultados.

A quantidade e a variedade de imagens que compõe a base de dados foi suficiente para fazer a classificação das imagens, bem como, validar as hipóteses levantadas durante os objetivos específicos.

5.1 Trabalhos Futuros

Apesar do método proposto auxiliar na recuperação de informação baseada em conteúdo através da extração dos atributos cor e textura em espaço *wavelet*, este trabalho não pretende esgotar as discussões sobre CBIRS. Isto porque o método desenvolvido neste trabalho considera um conjunto específico de características. Outros métodos podem ser propostos, quer incluindo novos atributos ou quer abordando as mesmas características, mas de maneiras diferenciadas, como por exemplo, a recuperação das imagens por similaridade através de pontos selecionados por regiões nas imagens.

A extensão da metodologia de recuperação de informação em imagens *on-line*, e o desenvolvimento de uma interface como método de ajuda à visualização das imagens recuperadas também podem ser consideradas como propostas de trabalhos futuros.

Referências Bibliográficas

BRAMBILLA, C.; VENTURA, D.; SCHETTINI, R. Multiresolution Wavelet Transform and Supervised Learning for Content-Based Image Retrieval. *IEEE Multimedia Systems 99*, IEEE Computer Society, Vol. I, 1999, páginas 183-188.

BUENO, J. M. *Suporte de Recuperação de Imagens Médicas Baseada em Conteúdo através de Histogramas Métricos*. Tese de Doutorado, Instituto de Ciências Matemáticas e de Computação Universidade de São Paulo, São Carlos, São Paulo, Brasil, 2001.

CIRES. Content Based Image Retrieval System. Disponível em <http://amazon.ece.utexas.edu/~qasim/research.htm>, acessado em 23 de abril de 2004.

COREL. Disponível em <http://www.corel.com>, acessado em 2 de Abril de 2004.

DAS, M. and MANMATHA, R.. Automatic Segmentation and Indexing in a Database of Bird Images. *IEEE Computer Society*, Vol. 2, 2001, páginas 351-358.

FACON, J. *Morfologia Matemática: Teoria e Exemplos*. Editora Universitária Champagnat da Pontifícia Universidade Católica do Paraná, Curitiba, 1996.

FINLAYSON, G. D., B. SCHIELE, J. L. CROWLEY, Comprehensive Colour Image Normalization, *Proc. 5th European Conference on Computer Vision*, 1998, páginas 475-490.

FUNT, B. V., G. D. FINLAYSON, Color Constant Color Indexing, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, 1995, páginas 522-529.

GALDINO, L. L. *Um Modelo de Atenção Visual para Detecção e Acompanhamento de Regiões*. Dissertação de Mestrado, Escola de Engenharia Elétrica Universidade Federal de Goiás, Goiânia, Goiás, Brasil, 2000.

GALDINO, L. L.; BORGES, D. L. *A Visual Attention Model for Tracking Regions Based on Color Correlograms*. Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens, Gramado, Brasil, 2000, páginas 36-43.

- GONZALEZ, R. e WOODS, R. E. *Processamento de Imagens Digitais*. Ed. Edgard Blücher Ltda, São Paulo, 2000.
- GUDIVADA, V. N. e RAGHAVAN, V.V. Content-Based Image Retrieval Systems. *IEEE Transactions on Patterns Analysis and Machine Intelligence*, vol.18, Setembro, 1995, páginas 18-22.
- HAFNER J., SAWHNEY H., EQUITZ W., FLICKNER M., NIBLACK W. Efficient Color Histogram Indexing for Quadratic Form Distance Function. *IEEE Transactions on Patterns Analysis and Machine Intelligence*, vol.17, Julho, 1995, páginas 729-736.
- IBM. IBM's Query by Image Content. Disponível em <http://www.qbic.almaden.ibm.com/>, acessado em 23 de abril de 2004.
- IQBAL, Q. e AGGARWAL, J. K. *Combining Structure, Color and Texture for Image Retrieval: A Performance Evaluation*. 16th International Conference on Pattern Recognition, Quebec, Canadá, vol. 2, 2002, páginas 438-443.
- LONG, H. Z. e LEOW, W. K. *Perceptual Texture Space for Content-Based Image Retrieval*. World Scientific, 2000.
- MALLAT, S. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, n. 7, julho de 1989, páginas 674-693.
- MATHIAS, E. e CONCI, A. *Comparing the Influence of Color Spaces and Metrics in Content-Based Image Retrieval*. Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens, Rio de Janeiro, Brasil, 1998, páginas 371-378.
- OLIVEIRA, C. J. S.; ARAUJO, A. A.; GOMES, D. R.; JUNIOR, C. A. S. Proposta de um Protótipo de um Sistema de Recuperação de Imagens com Base na Cor. *III Workshop em Tratamento de Imagens*, Belo Horizonte, Minas Gerais, junho 2002.
- PASS, G.; ZABIH, R.; MILLER, J. Comparing Images Using Color Coherence Vector. *ACM Multimedia*, Boston, MA, 1996.
- TRIP. Revista Trip. Disponível em <http://www2.uol.com.br/trip>, acessado em 23 de abril de 2004.
- SCHETTINI, R.; CIOCCA, G. e ZUFFI, S. *A Survey of Methods for Colour Image Indexing*

and Retrieval in Image Databases. Technical Report, Istituto Tecnologie Informatiche Multimediali, 2001.

USC. University of Southern California - Signal and Image Processing. Disponível em <http://sipi.usc.edu/services/database/>, acessado em 23 de abril de 2004.

SMEULDERS, A. W. M.; WORRING, M.; SANTINI, S.; GUPTA, A. e JAIN, R. Content-Based Image Retrieval at the End of the Early Years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, n. 12, 2000, páginas 1349-1380.

STRICKER, M. e DIMAI, A. Color Indexing with Weak Spatial Constraints. *Storage and Retrieval for image databases IV, Proc. SPIE 2670*, 1996, páginas 29-40.

TRAINA, A. J. M. *Suporte à Visualização de Consultas por Similaridade em Imagens Médicas através de Estrutura de Indexação Métrica*. Tese de livre docência apresentada ao ICMC-USP, setembro de 2001.

VRIES, A. P.; MAMOULIS, N. e NES, N. J. *Efficient Image Retrieval by Exploiting Vertical Fragmentation*. Centrum voor Wiskunde en Informatica, 2001, Report INS-R 0109.

Apêndice A

As figuras consideradas na realização dos experimentos do Apêndice A constituem as imagens da base de dados, e são apresentadas a seguir de acordo com as classes trabalhadas.

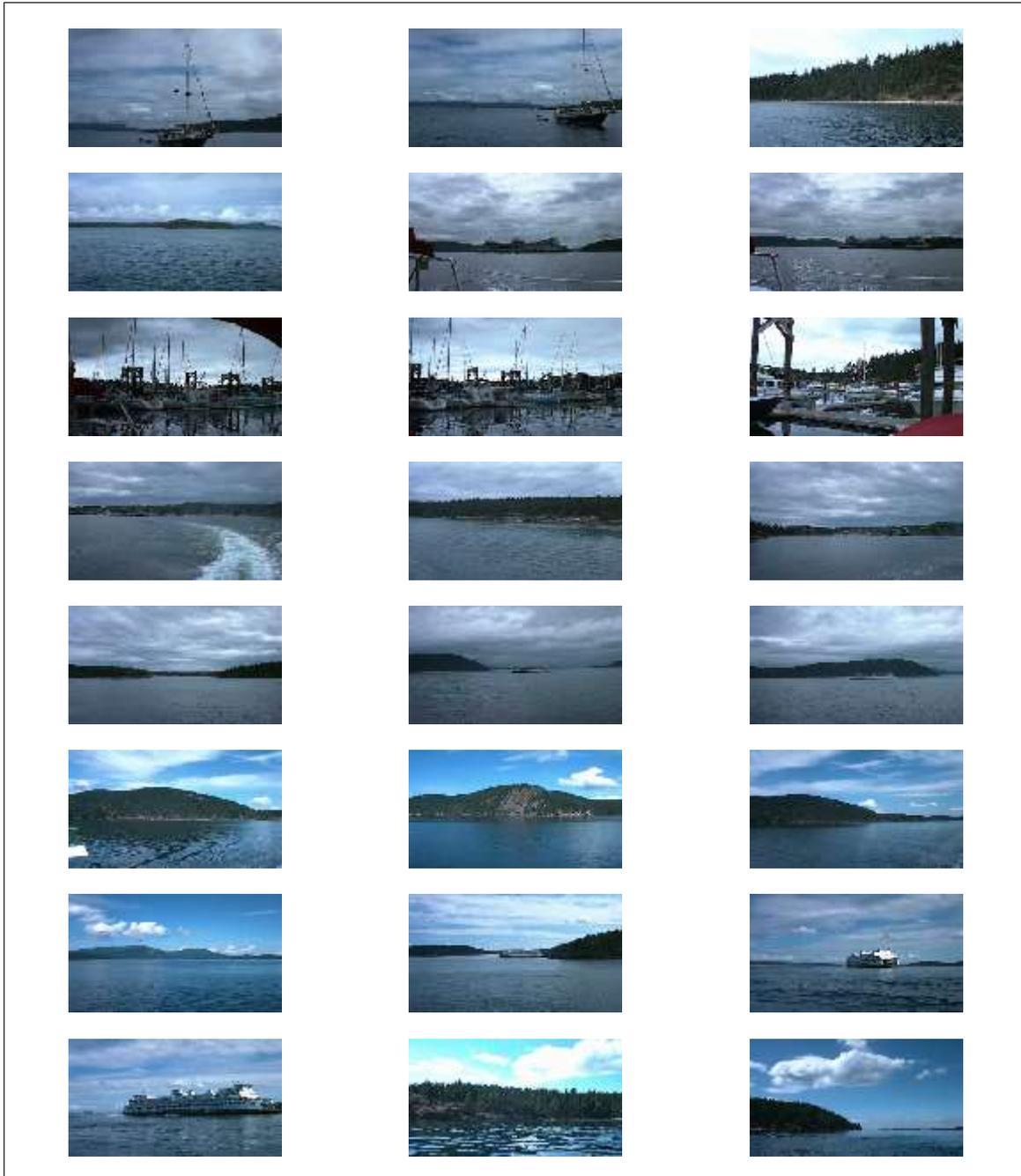




Figura A.1 –Classe 1 referente a Água – Imagens de 1 a 27.

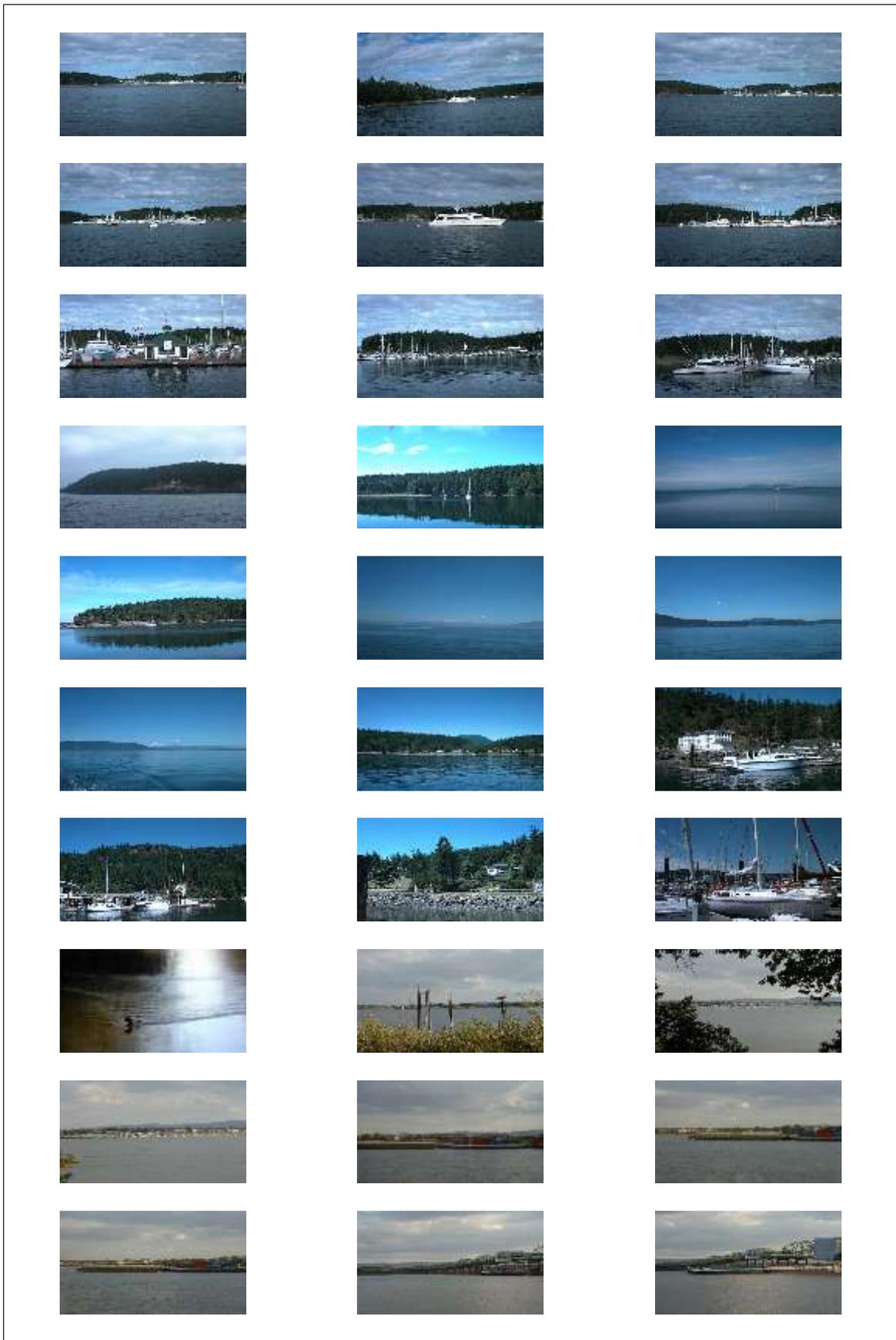


Figura A.2 - Classe 1 referente a Água – Imagens de 28 a 57.

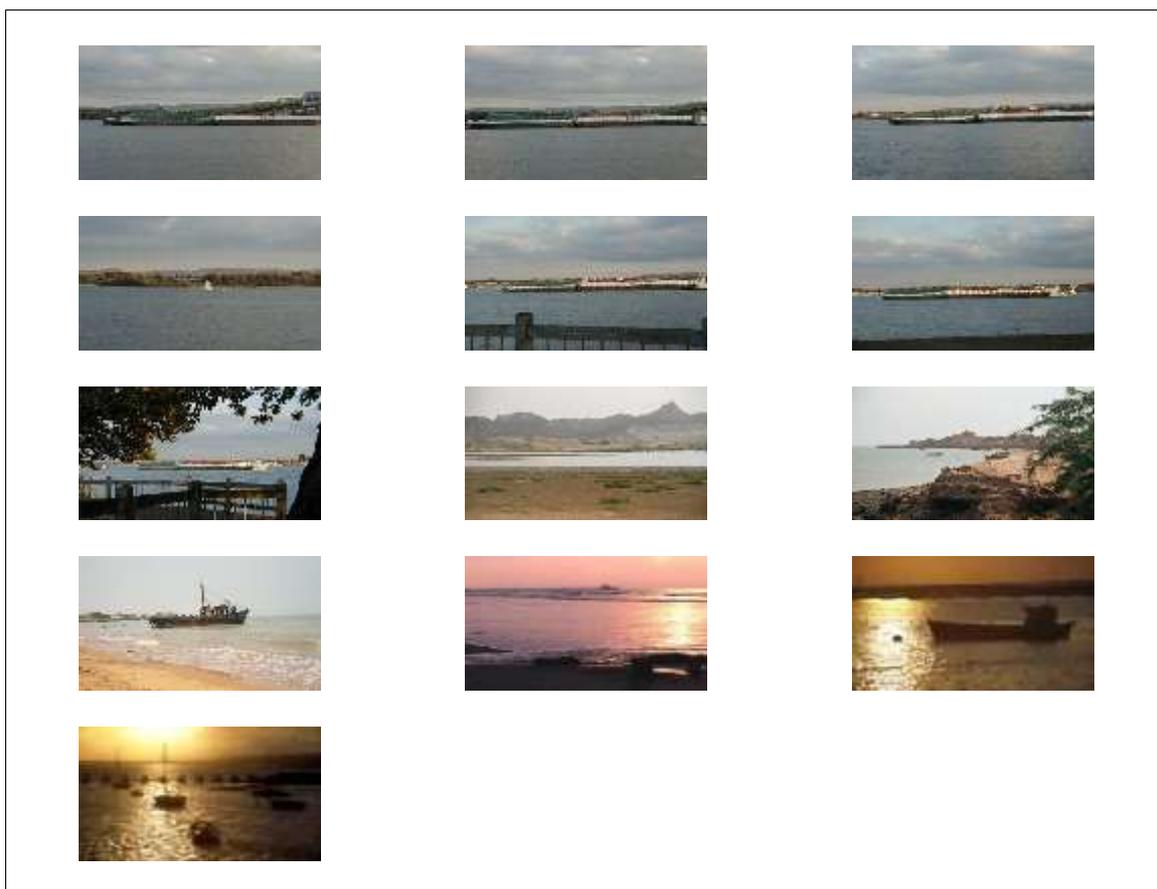


Figura A.3 - Classe 1 referente as imagens Água – Imagens de 58 a 70.

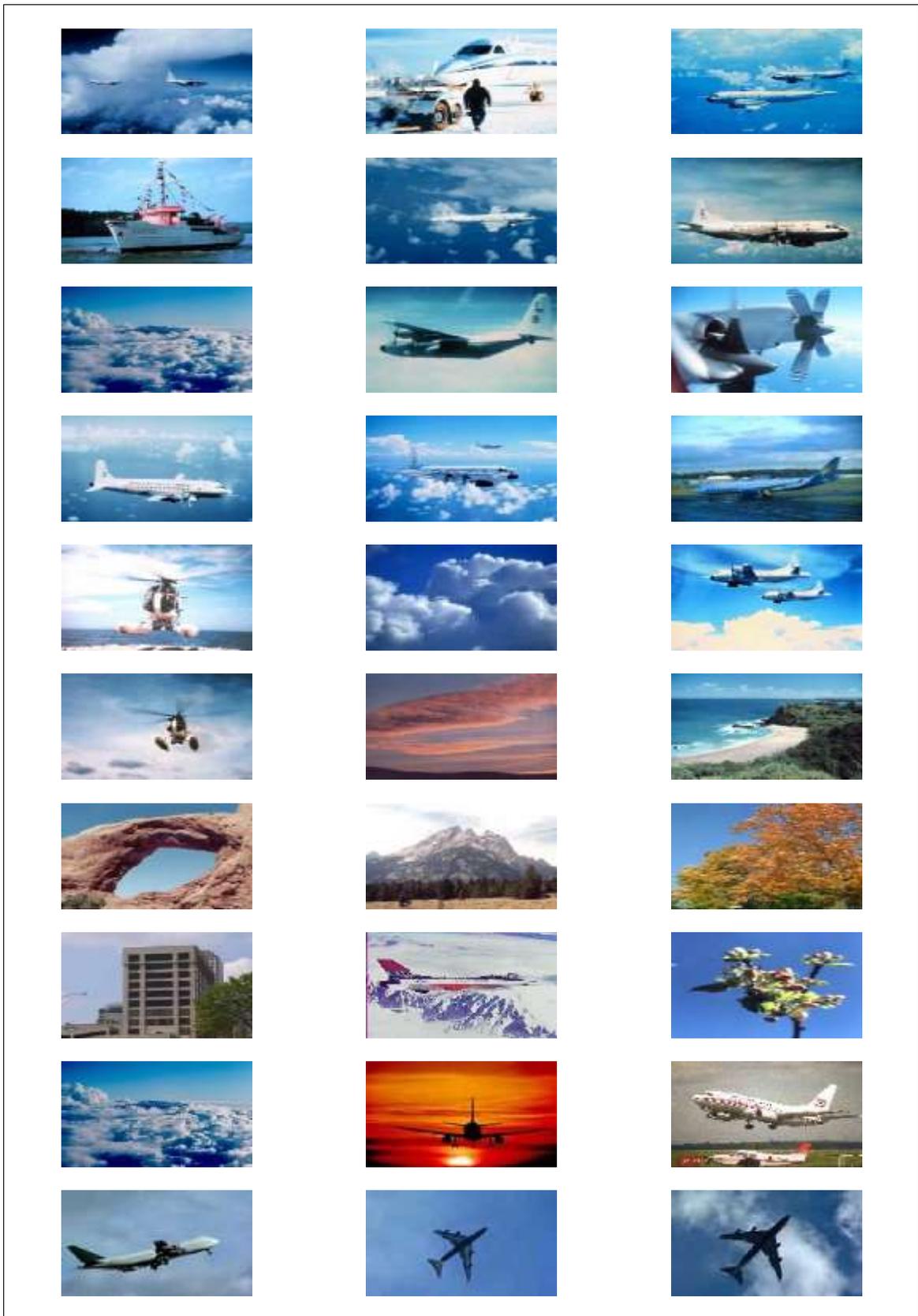


Figura A.4 - Classe 2 referente as imagens Ar – Imagens de 1 a 30.

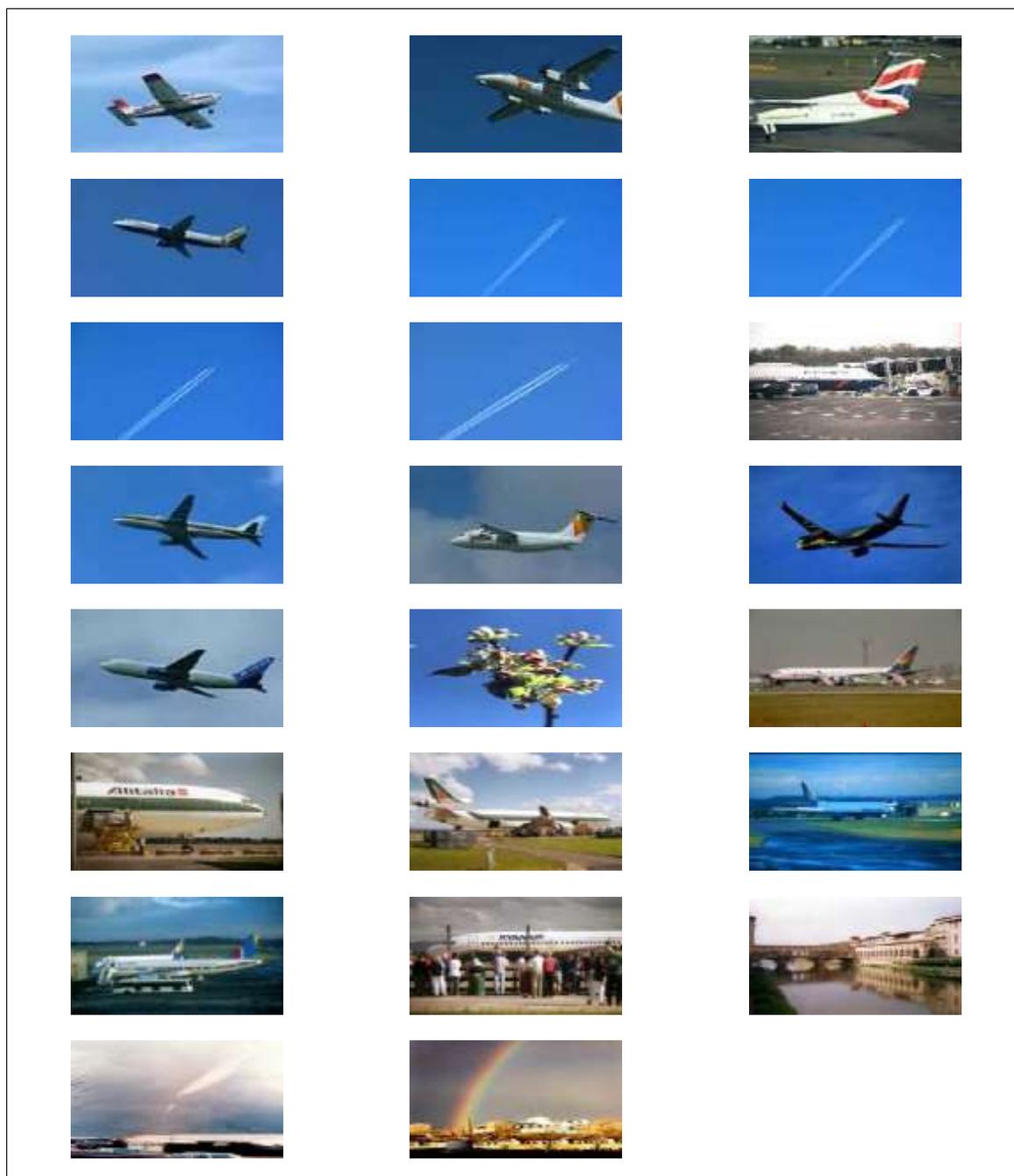


Figura A.5 - Classe 2 referente as imagens Ar – Imagens de 31 a 53.



Figura A.6 - Classe 3 referente as imagens Animais – Imagens de 1 a 30.



Figura A.7 - Classe 3 referente as imagens Animais – Imagens de 31 a 55.



Figura A.8 - Classe 4 referente as imagens Arame – Imagens de 1 a 6.



Figura A.9 - Classe 5 referente as imagens Árvore – Imagens de 1 a 30.

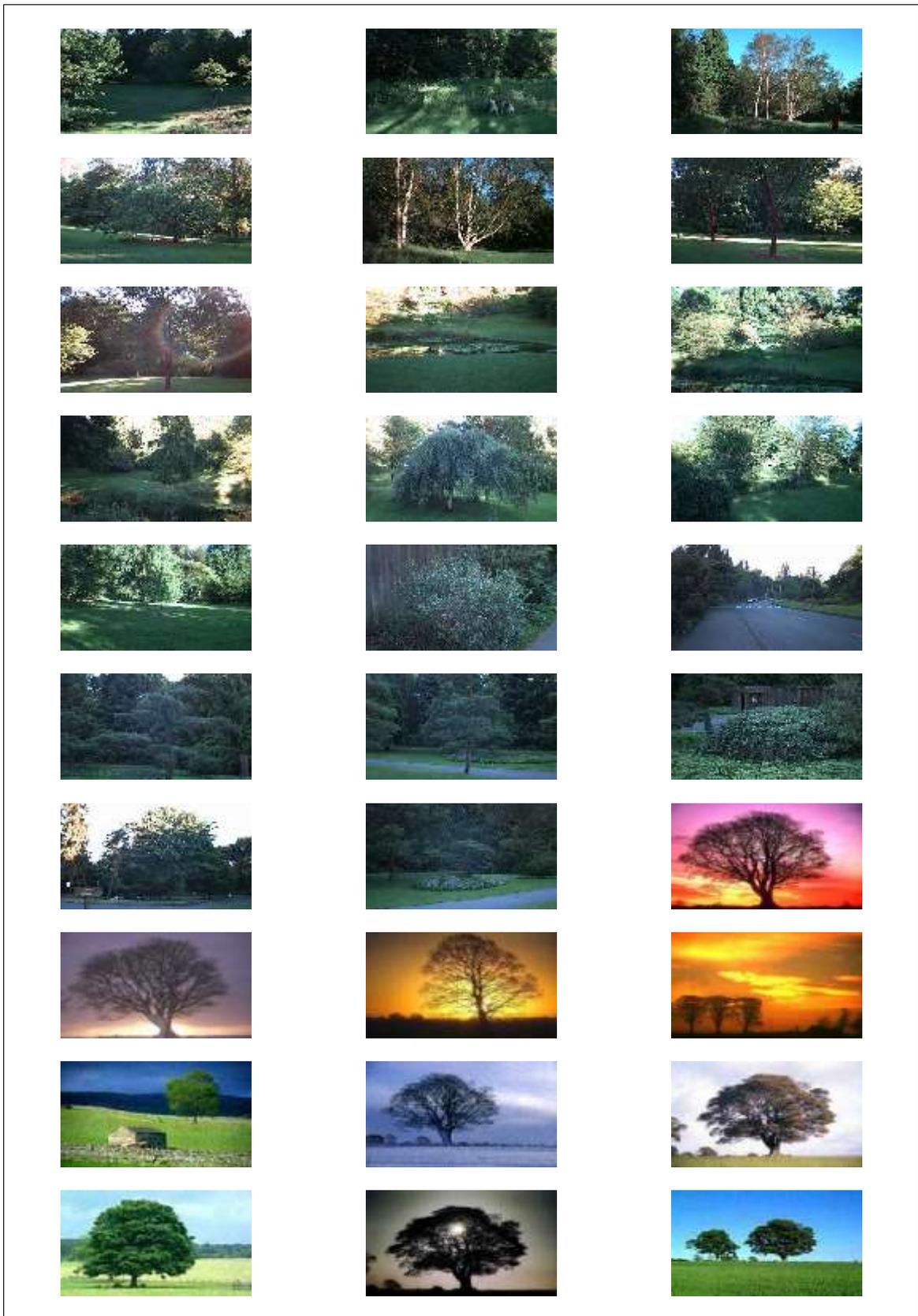


Figura A.10 - Classe 5 referente as imagens Árvore – Imagens de 31 a 60.

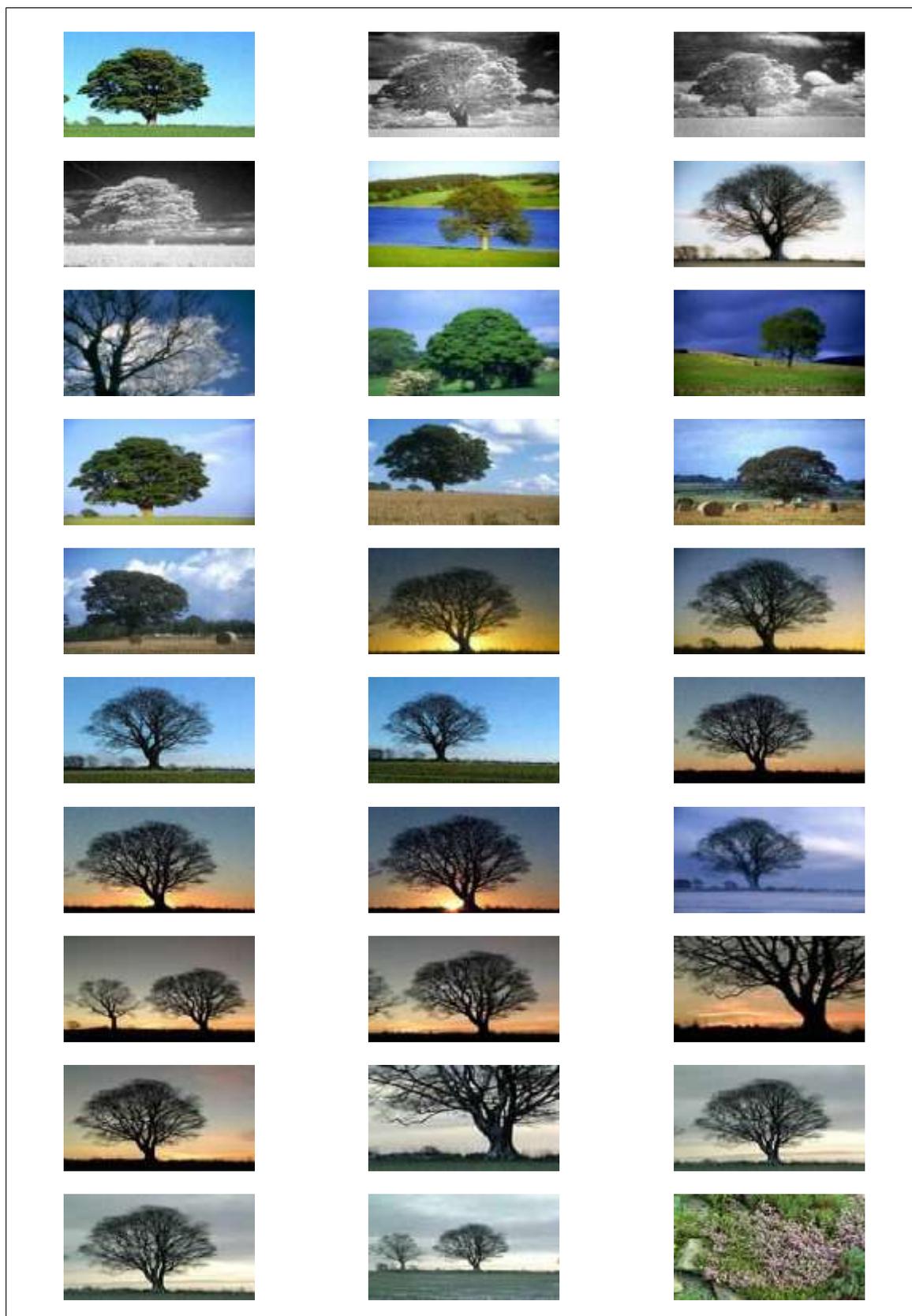


Figura A.11 - Classe 5 referente as imagens Árvore – Imagens de 61 a 90.



Figura A.12 - Classe 5 referente as imagens Árvore – Imagens de 91 a 103.

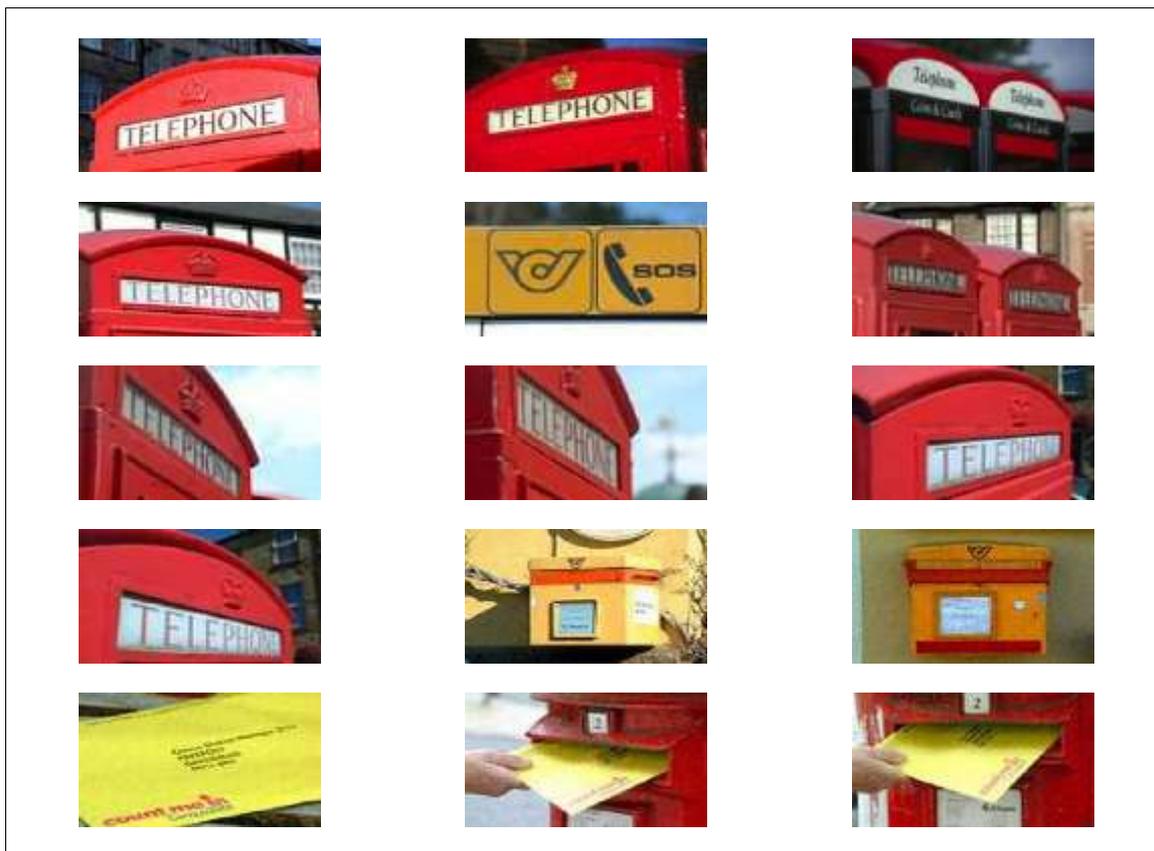


Figura A.13 - Classe 6 referente as imagens Caixa Postal – Imagens de 1 a 15.

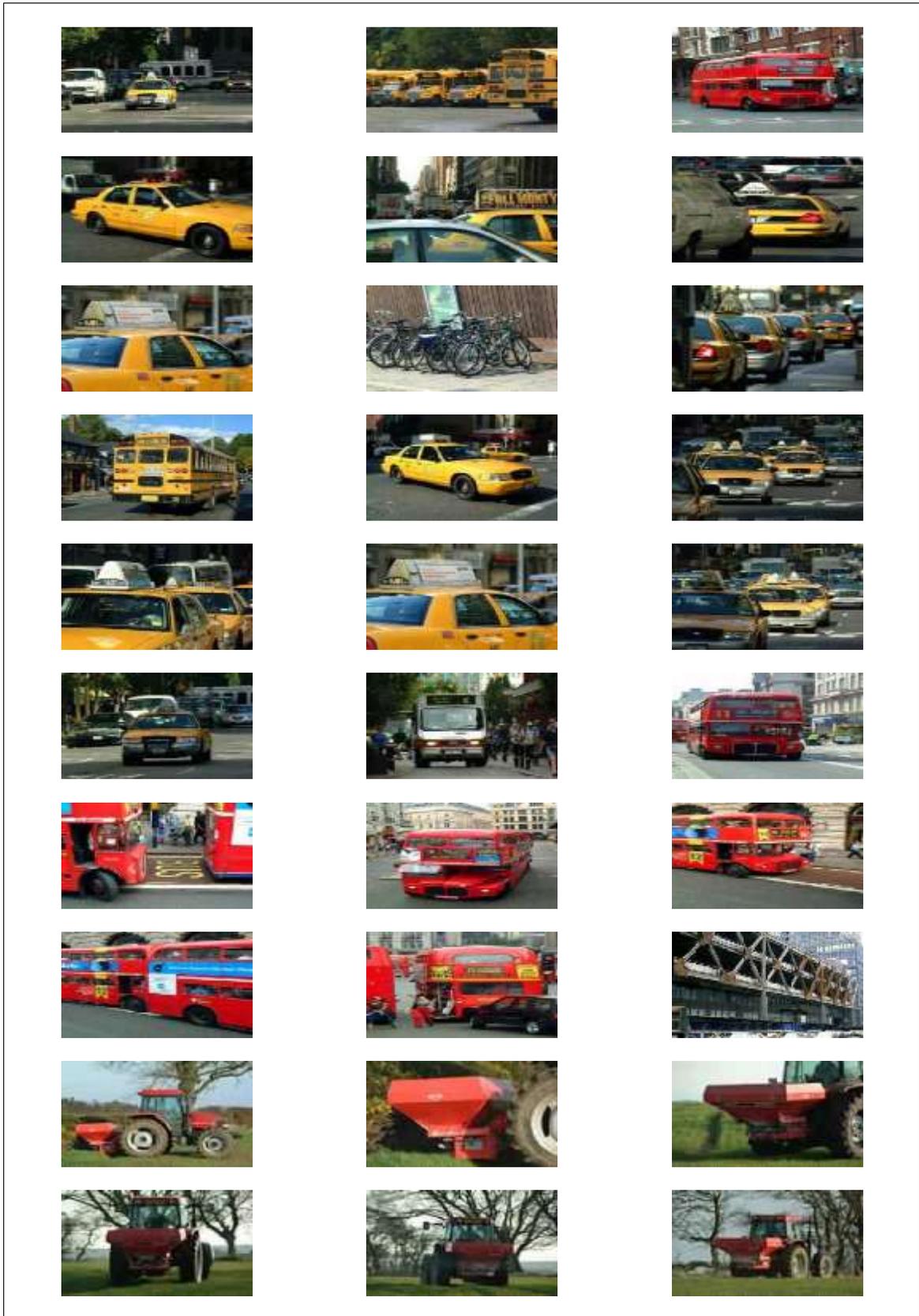


Figura A.14 - Classe 7 referente as imagens Carro – Imagens de 1 a 30.



Figura A.15 - Classe 7 referente as imagens Carro – Imagens de 31 a 42.

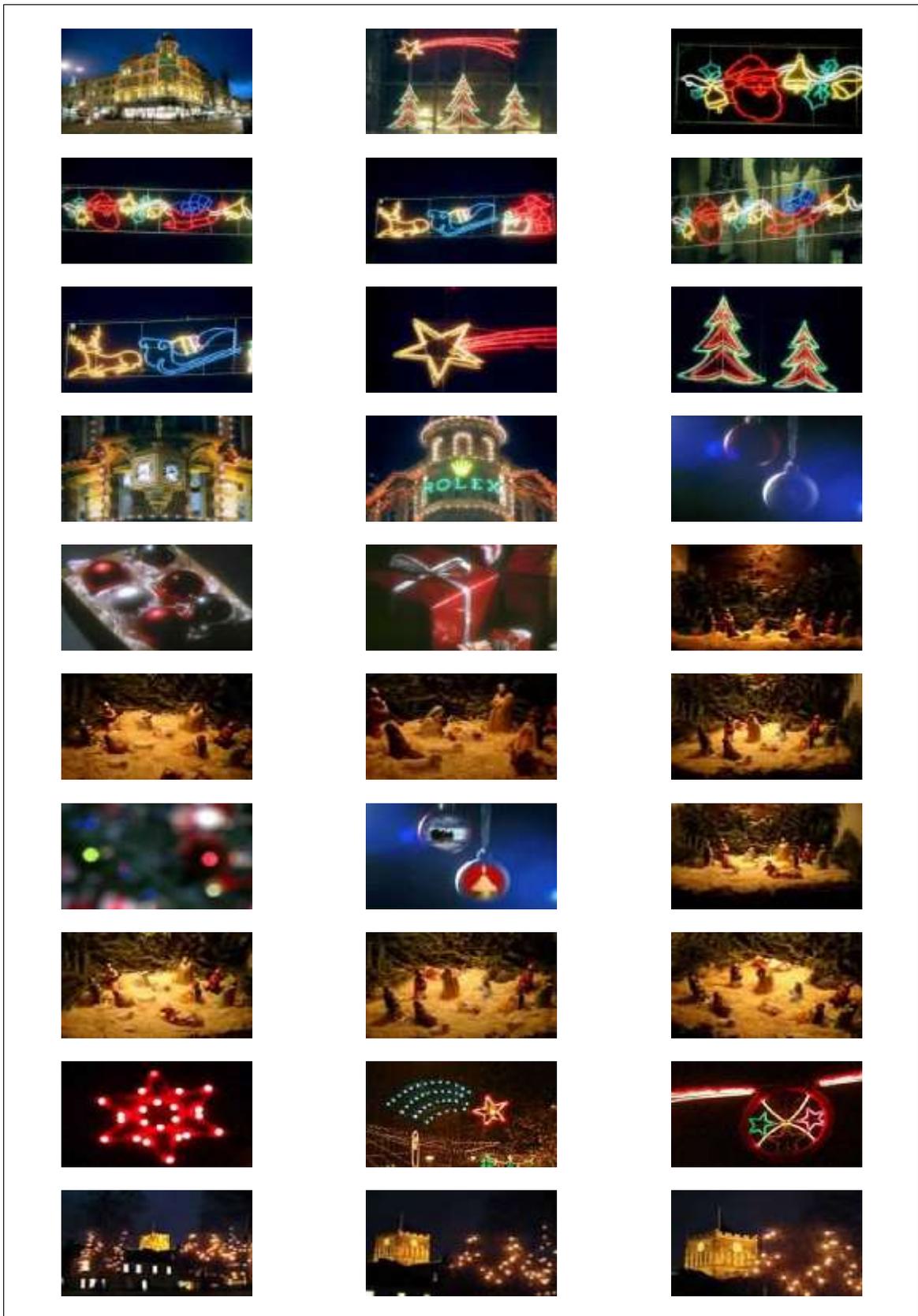


Figura A.16 - Classe 8 referente as imagens Natal – Imagens de 1 a 30.

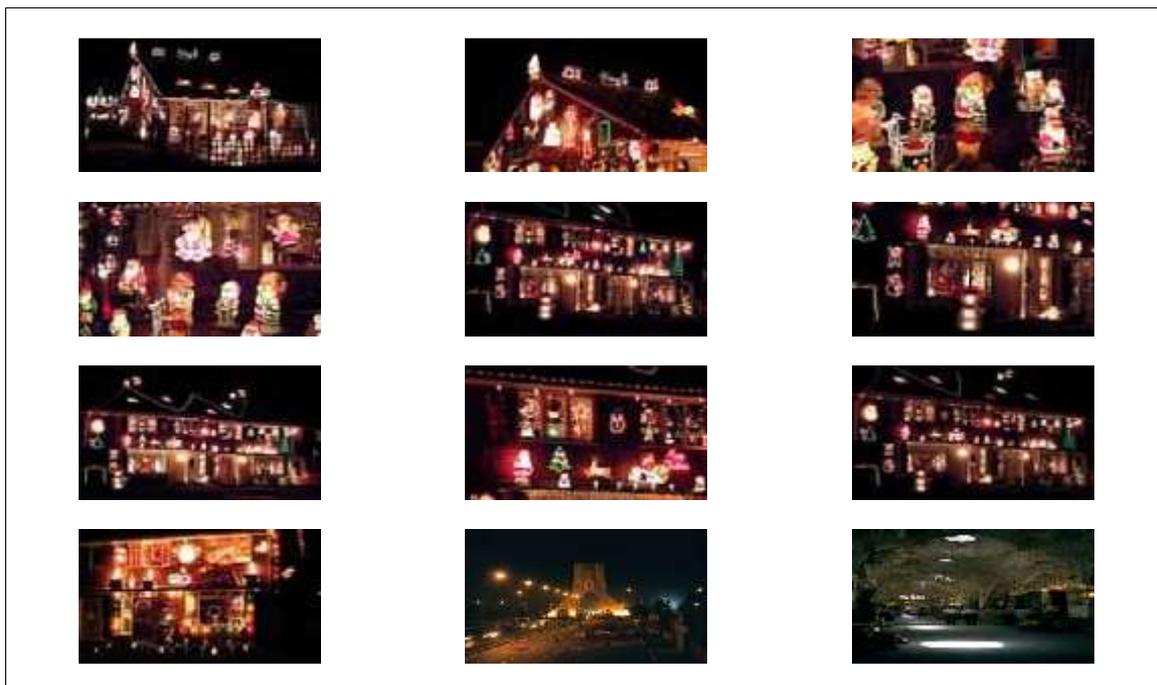


Figura A.17 - Classe 8 referente as imagens Natal – Imagens de 31 a 42.



Figura A.18 - Classe 9 referente as imagens Concreto – Imagens de 1 a 14.

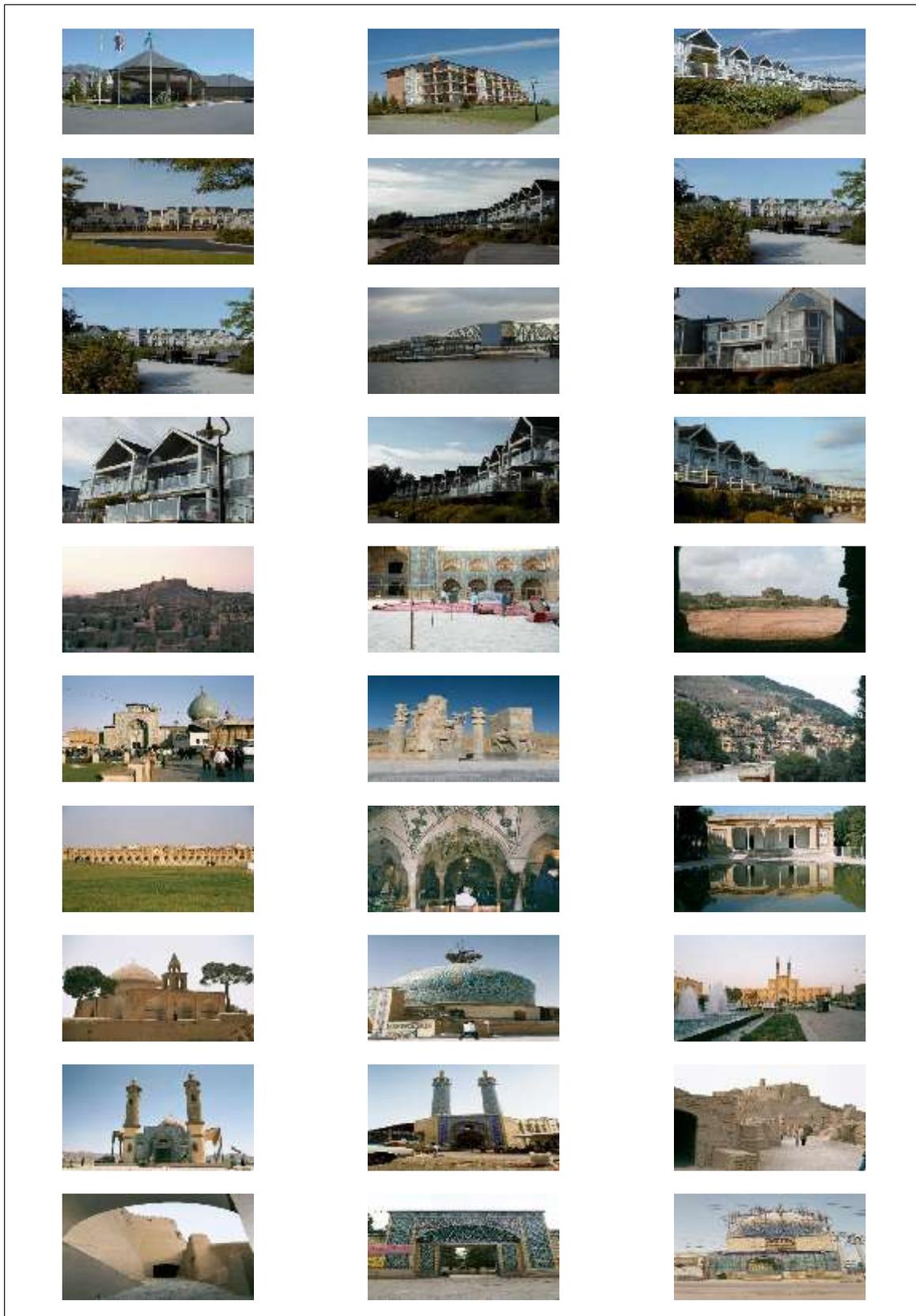


Figura A.19 - Classe 10 referente as imagens Construção – Imagens de 1 a 30.

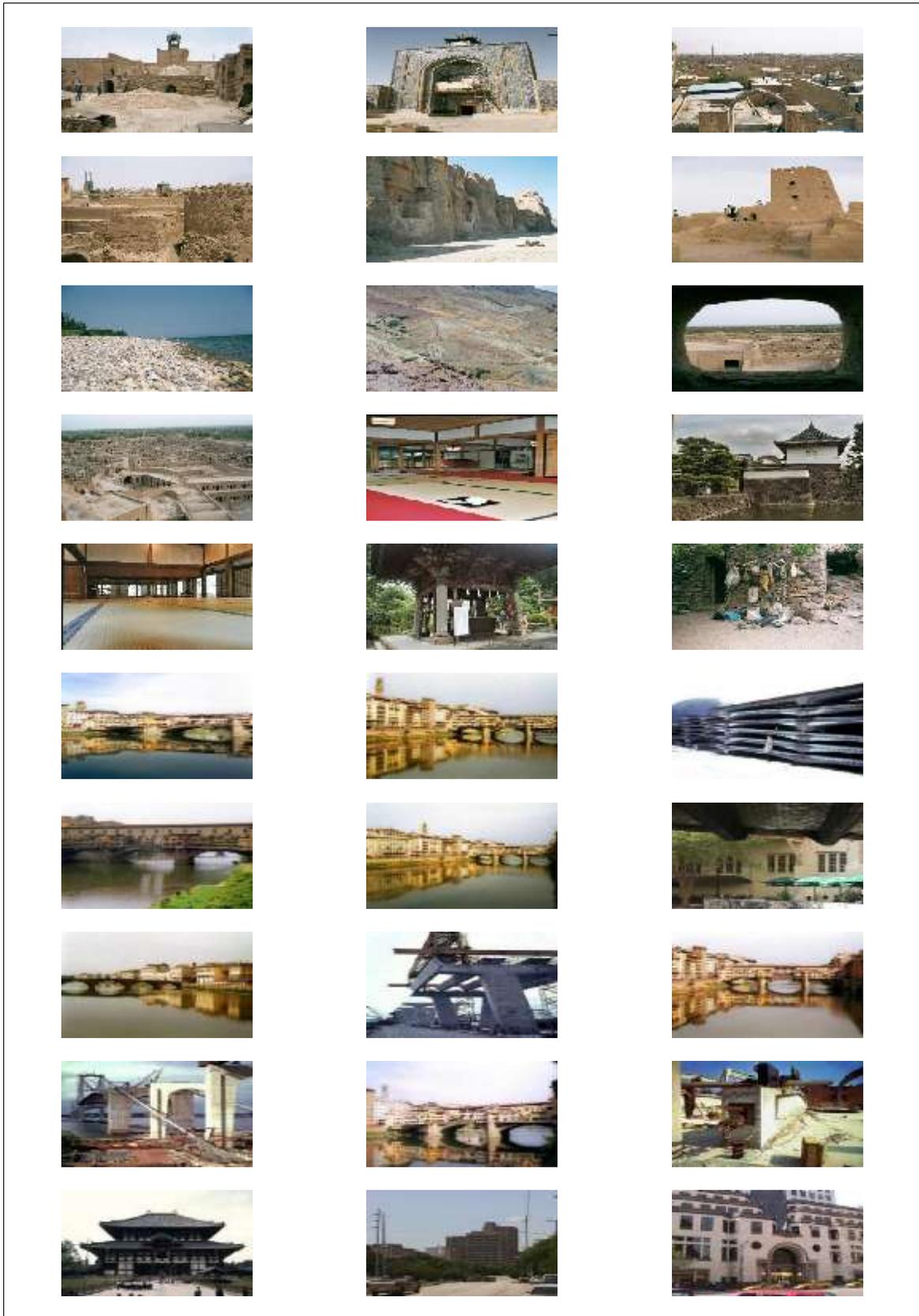


Figura A.20 - Classe 10 referente as imagens Construção – Imagens de 31 a 60.



Figura A.21 - Classe 10 referente as imagens Construção – Imagens de 61 a 79.

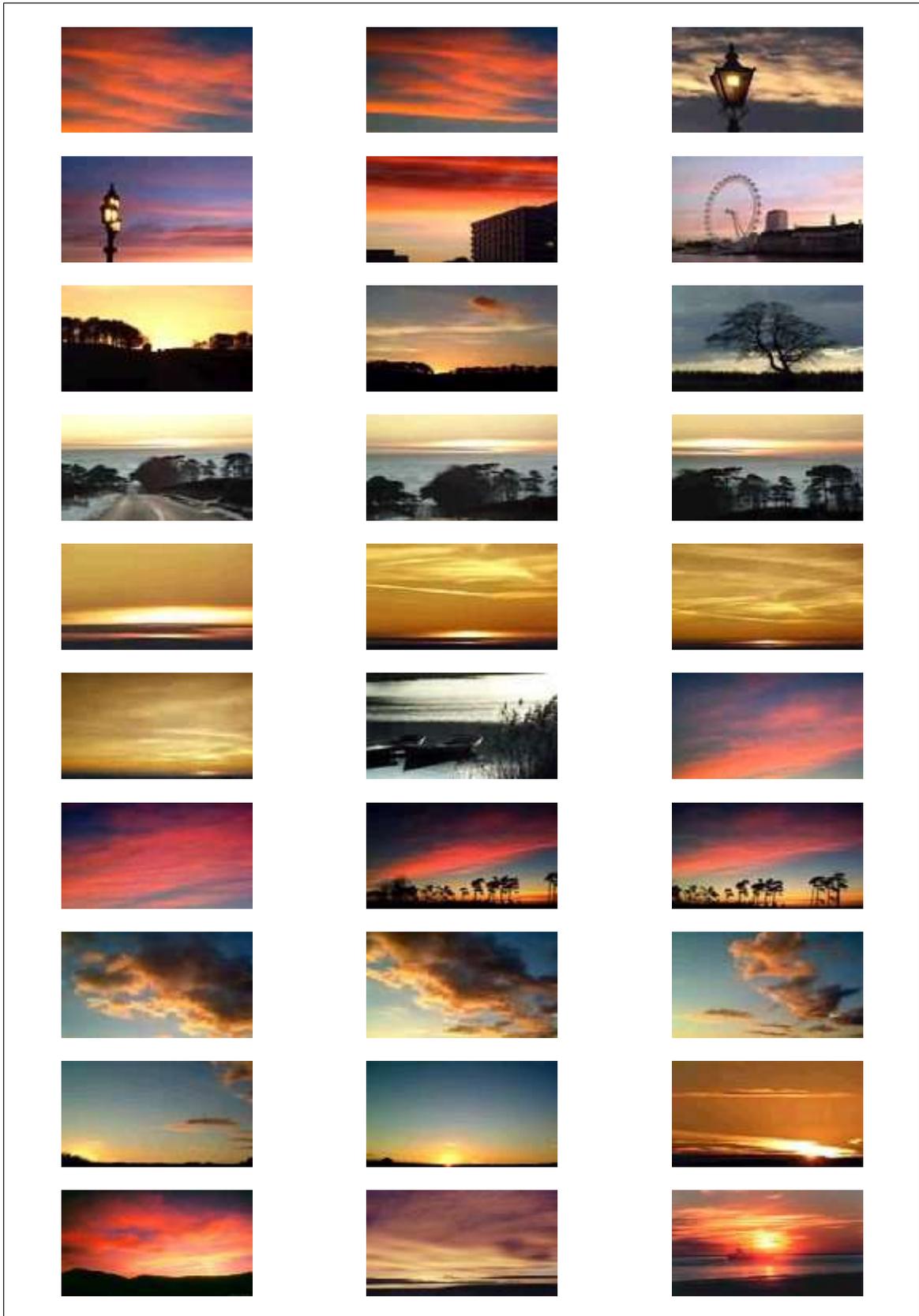


Figura A.22 - Classe 11 referente as imagens Amanhecer – Imagens de 1 a 30.

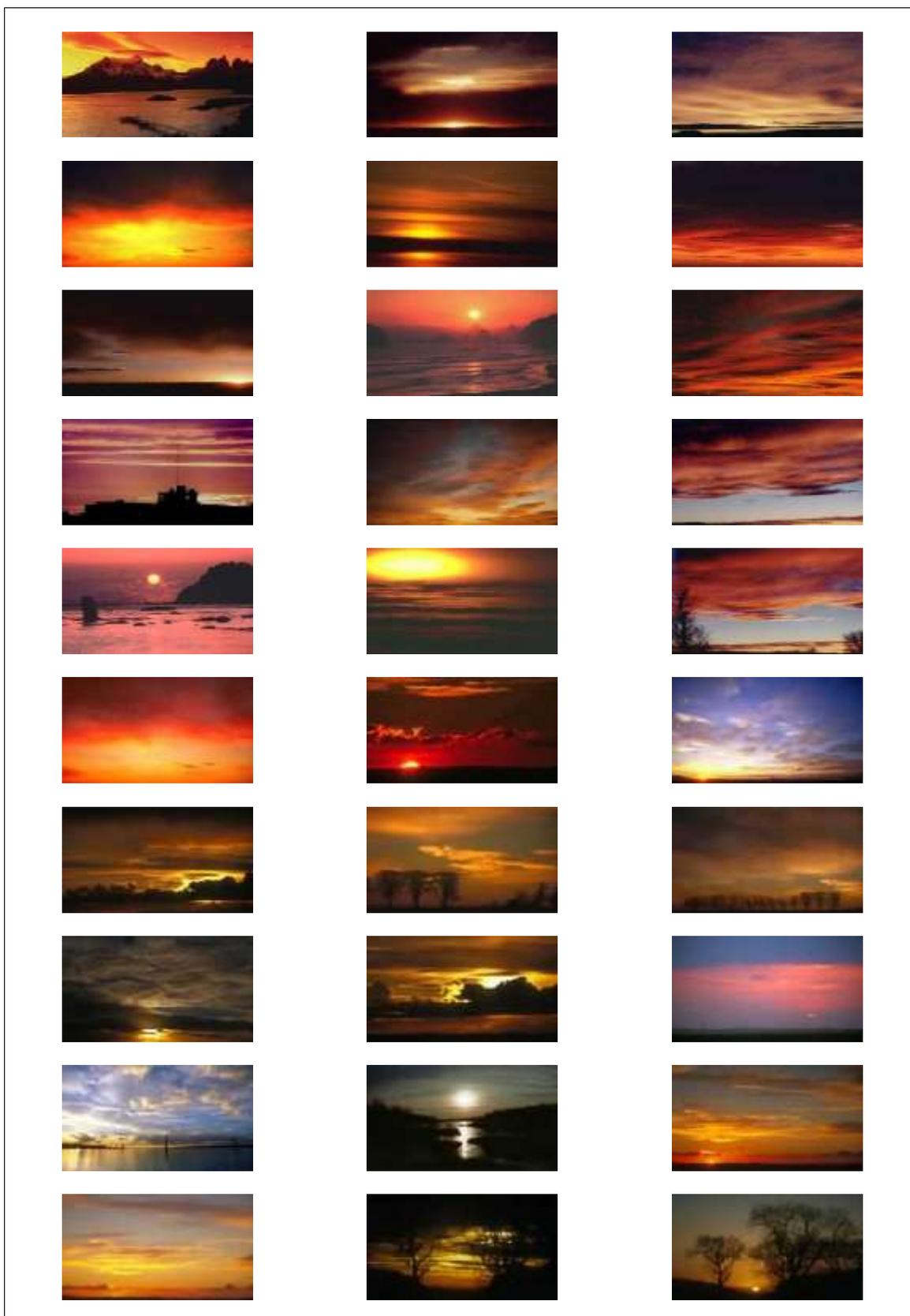


Figura A.23 - Classe 11 referente as imagens Amanhecer – Imagens de 31 a 60.



Figura A.24 - Classe 11 referente as imagens Amanhecer – Imagens de 61 a 63.

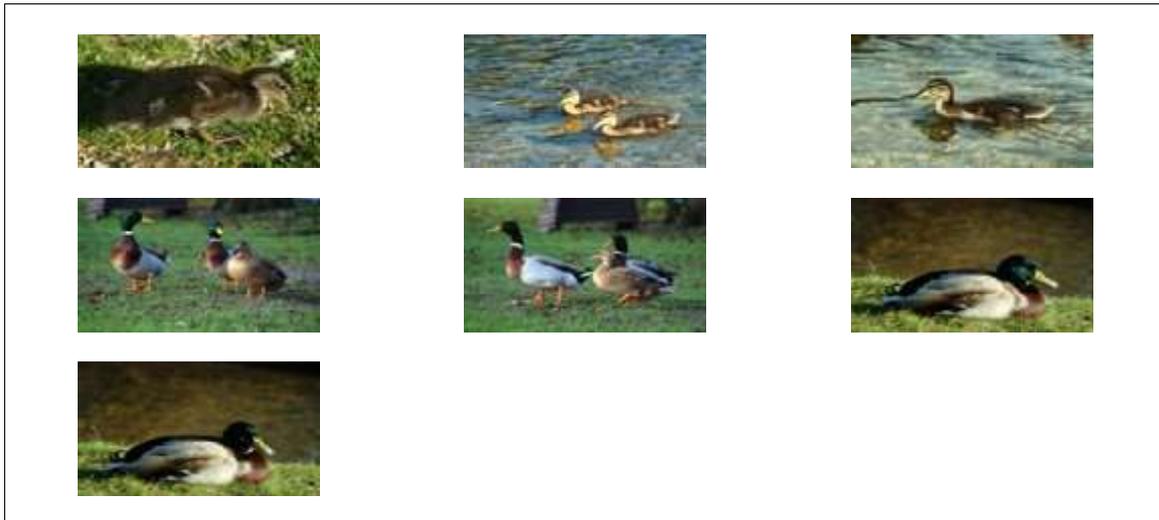


Figura A.26 - Classe 12 referente as imagens Pato – Imagens de 31 a 37.



Figura A.27 - Classe 13 referente as imagens Bandeira – Imagens de 1 a 30.



Figura A.29 - Classe 14 referente as imagens Flores – Imagens de 1 a 30.

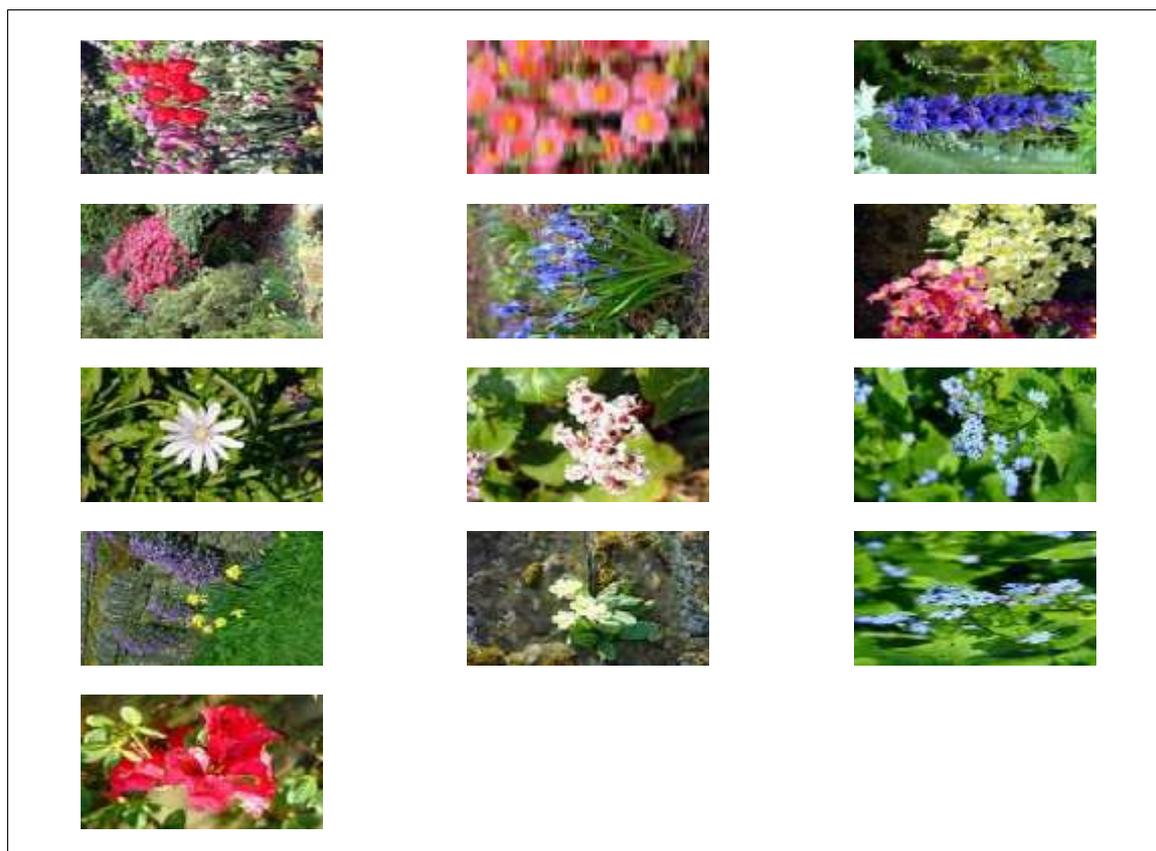


Figura A.30 - Classe 14 referente as imagens Flores – Imagens de 31 a 43.

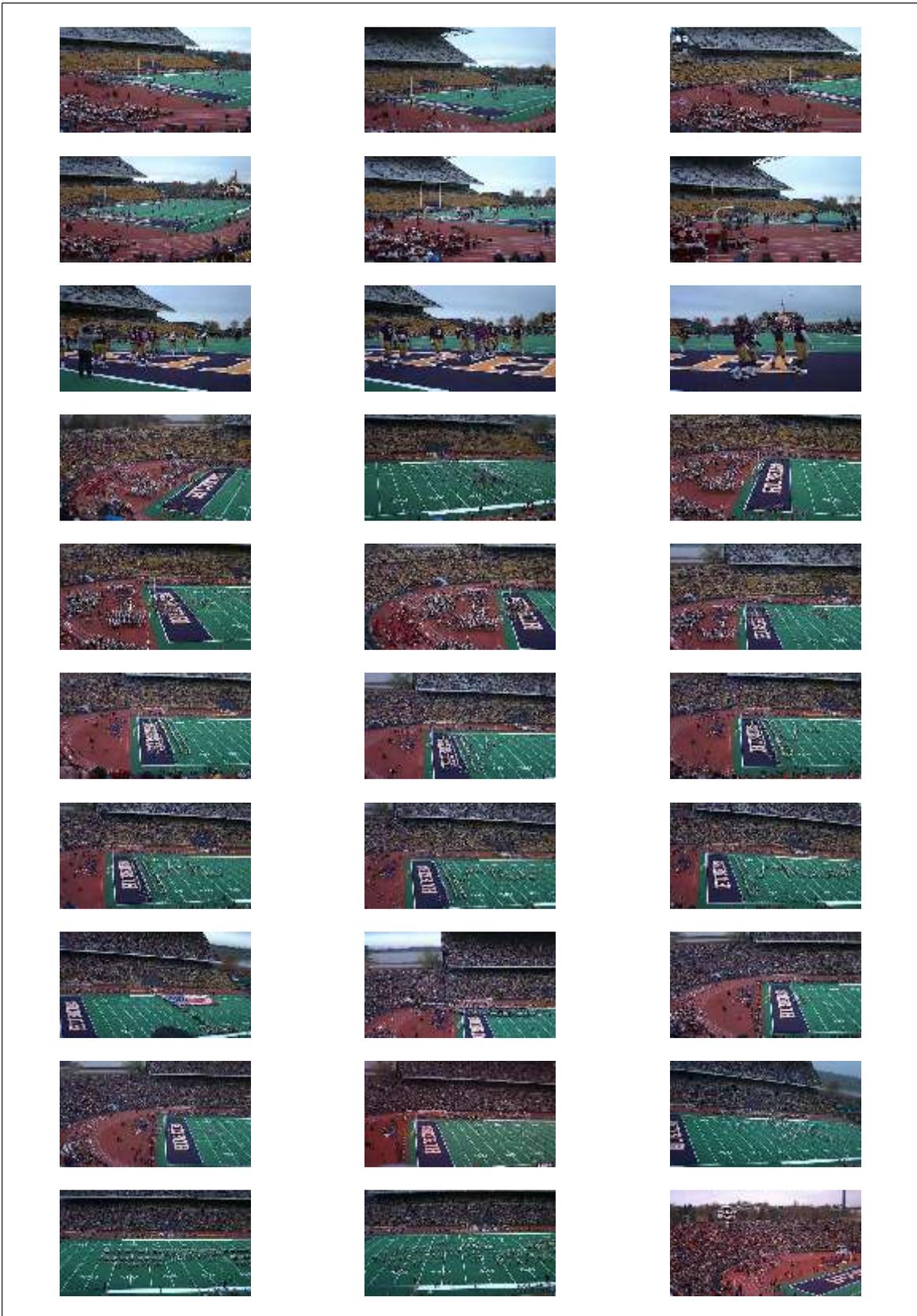


Figura A.31 - Classe 15 referente as imagens Futebol – Imagens de 1 a 30.



Figura A.32 - Classe 15 referente as imagens Futebol – Imagens de 31 a 44.



Figura A.33 - Classe 16 referente as imagens Frutas – Imagens de 1 a 19.

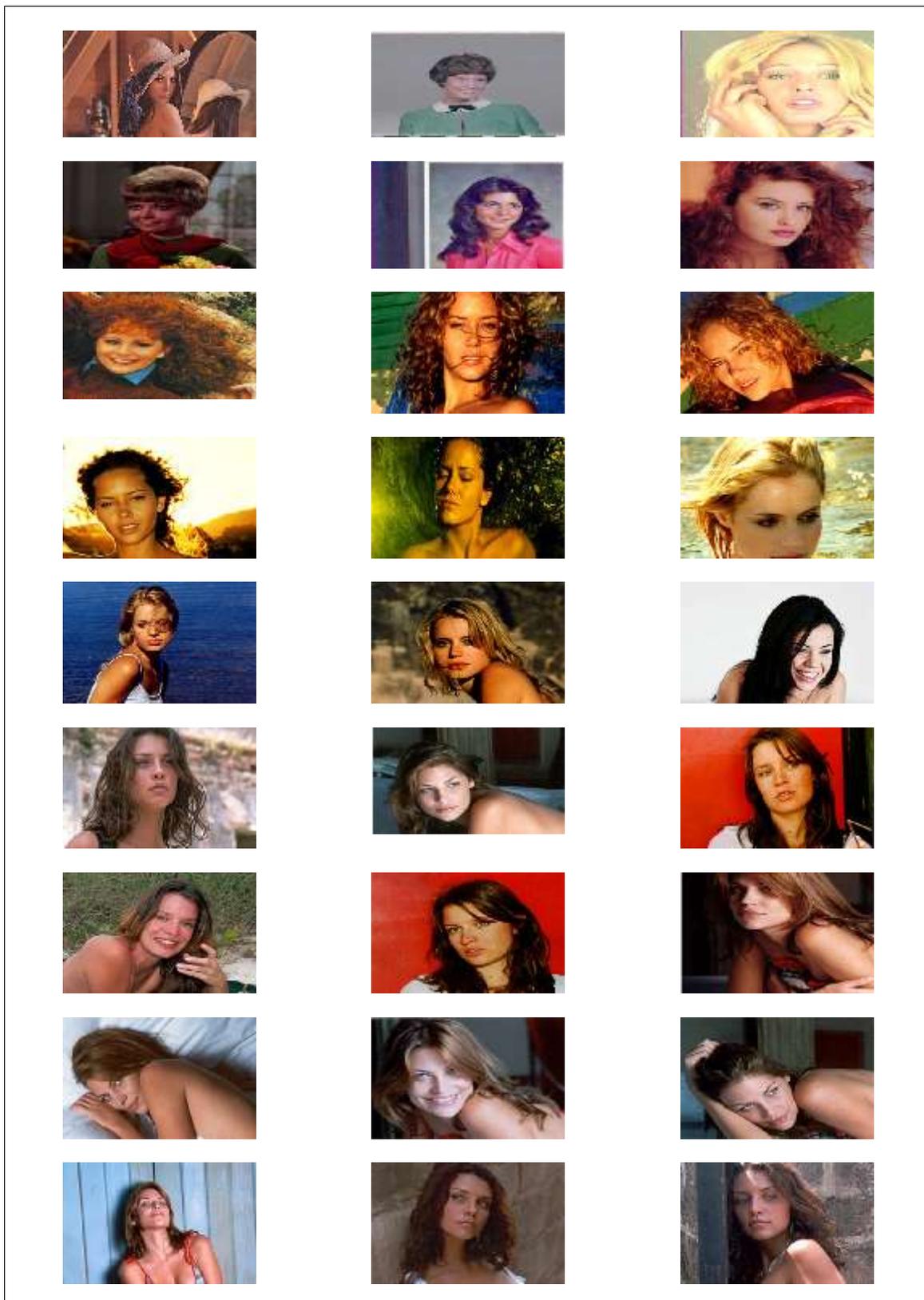


Figura A.34 - Classe 17 referente as imagens Garotas – Imagens de 1 a 27.



Figura A.35 - Classe 18 referente as imagens Árvore-Rosa – Imagens de 1 a 30.



Figura A.36 - Classe 18 referente as imagens Árvore-Rosa – Imagens de 31 a 55.

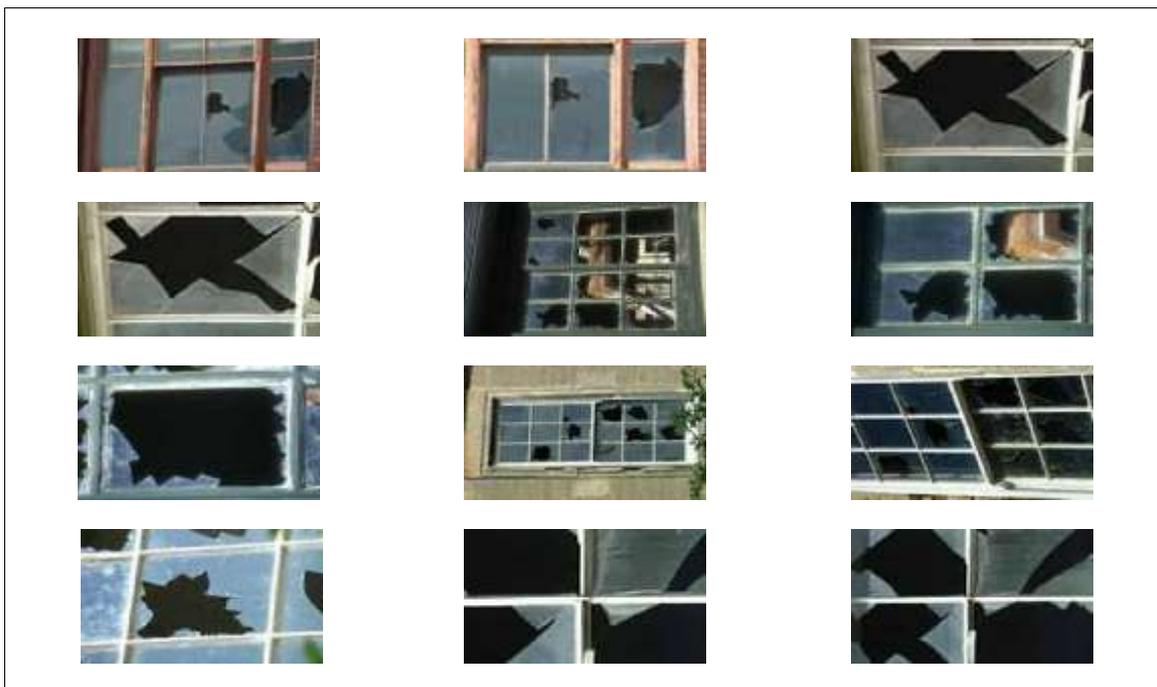


Figura A.37 - Classe 19 referente as imagens Janela – Imagens de 1 a 12.



Figura A.38 - Classe 20 referente as imagens Folha – Imagens de 1 a 30.

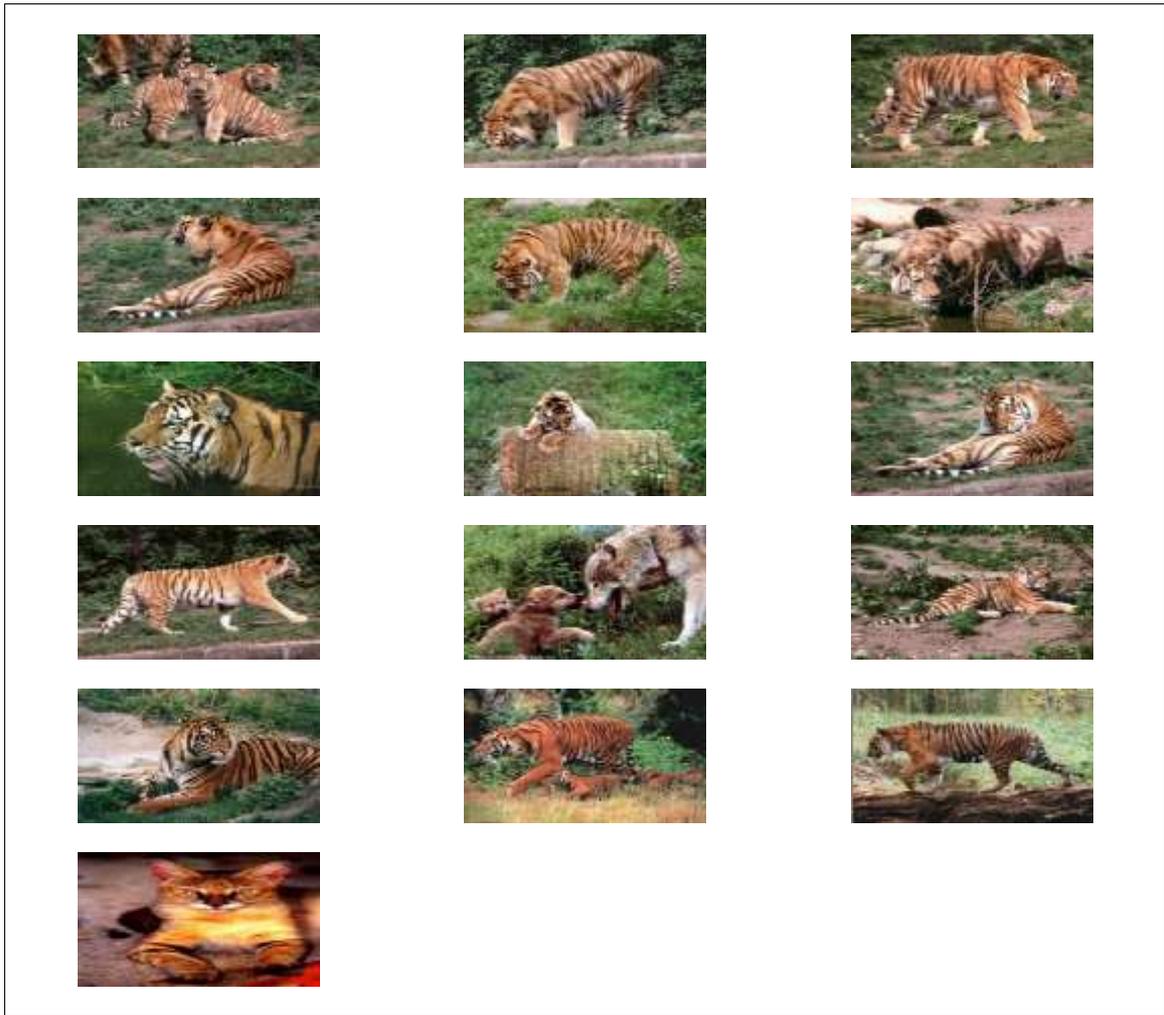


Figura A.39 - Classe 21 referente as imagens Mamíferos – Imagens de 1 a 16.

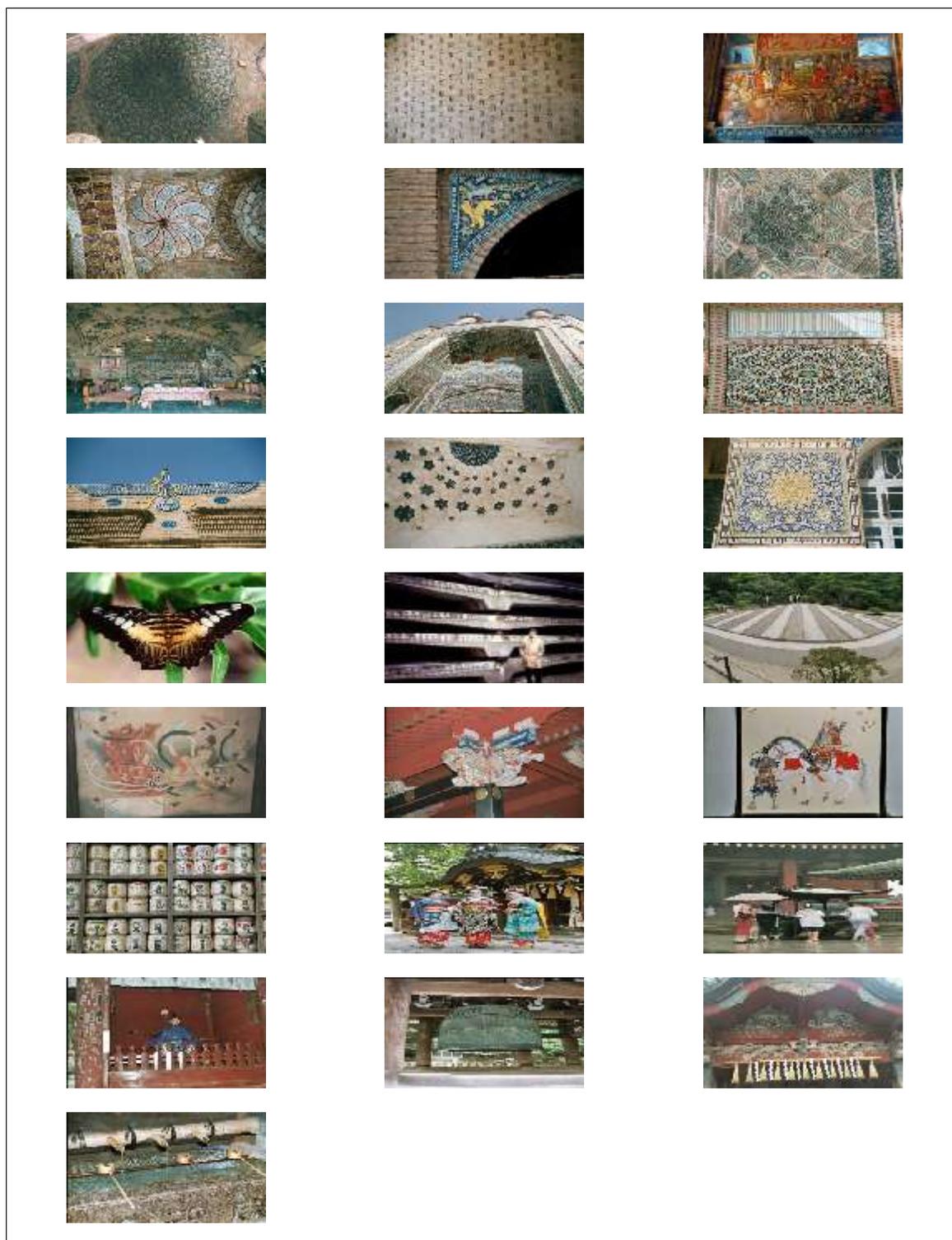


Figura A.40 - Classe 22 referente as imagens Mosaico – Imagens de 1 a 25.



Figura A.41 - Classe 23 referente as imagens Montanhas – Imagens de 1 a 30.

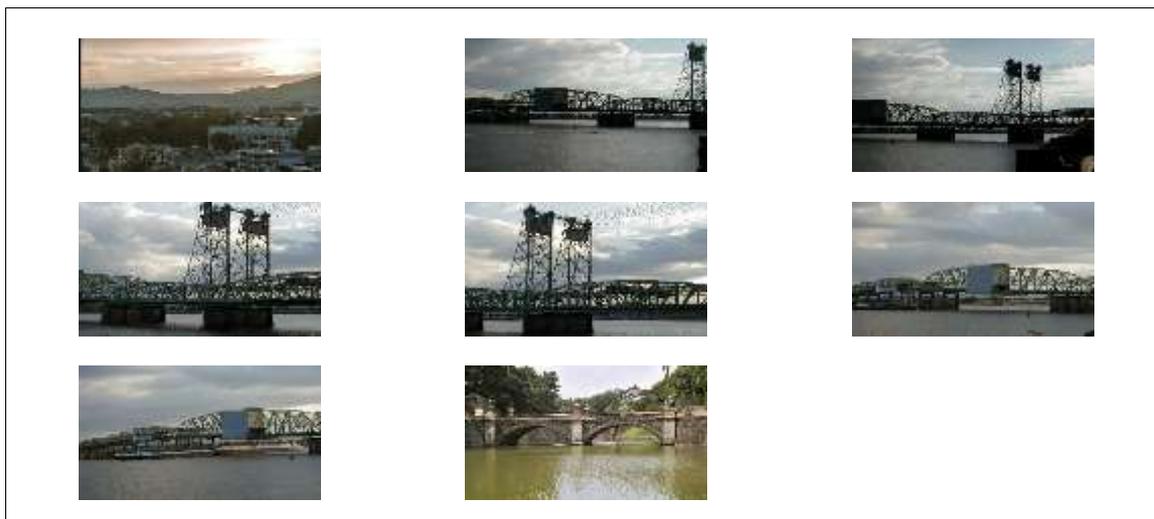


Figura A.42 - Classe 24 referente as imagens Ponte – Imagens de 1 a 8.

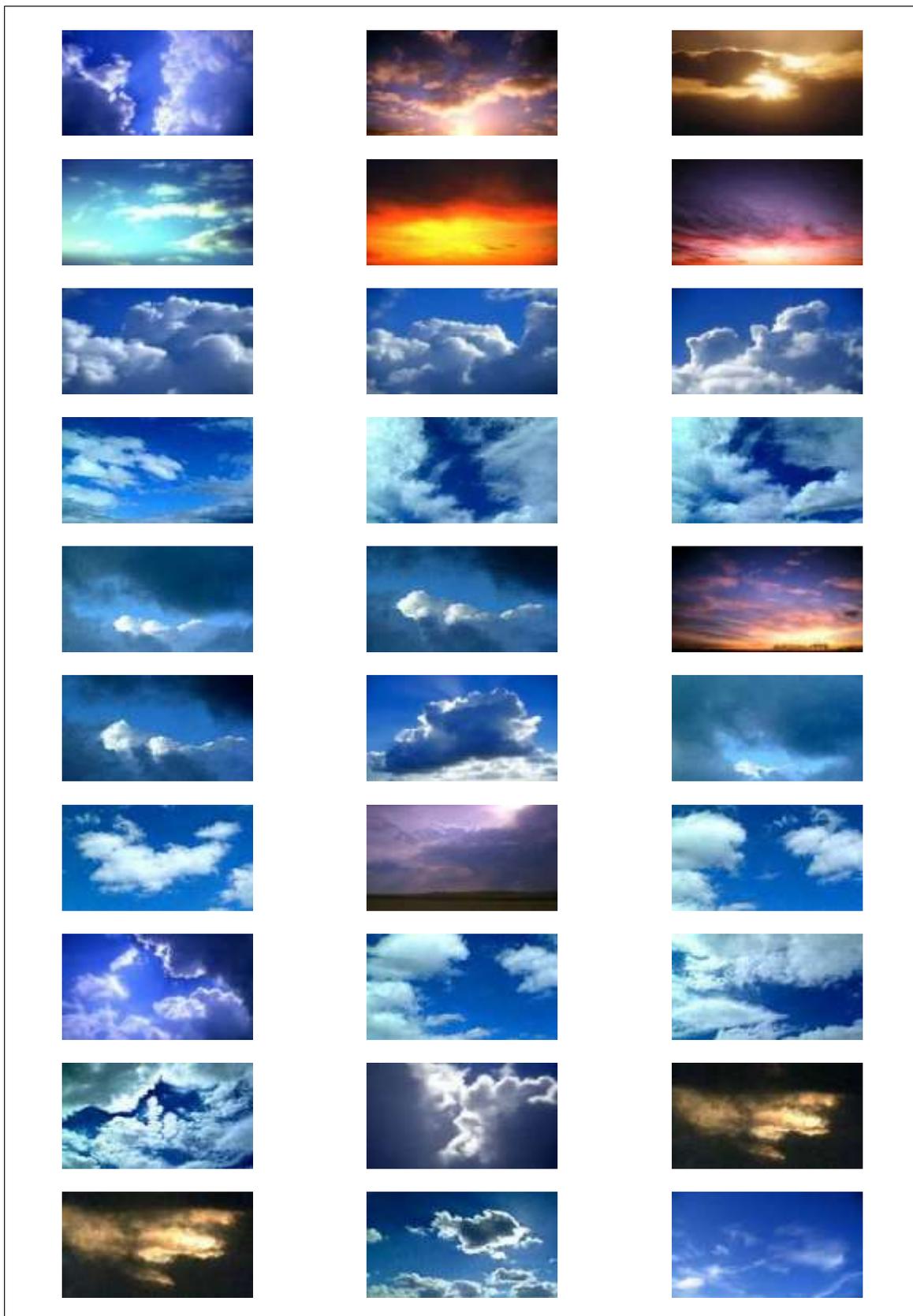


Figura A.43 - Classe 25 referente as imagens Céu – Imagens de 1 a 30.

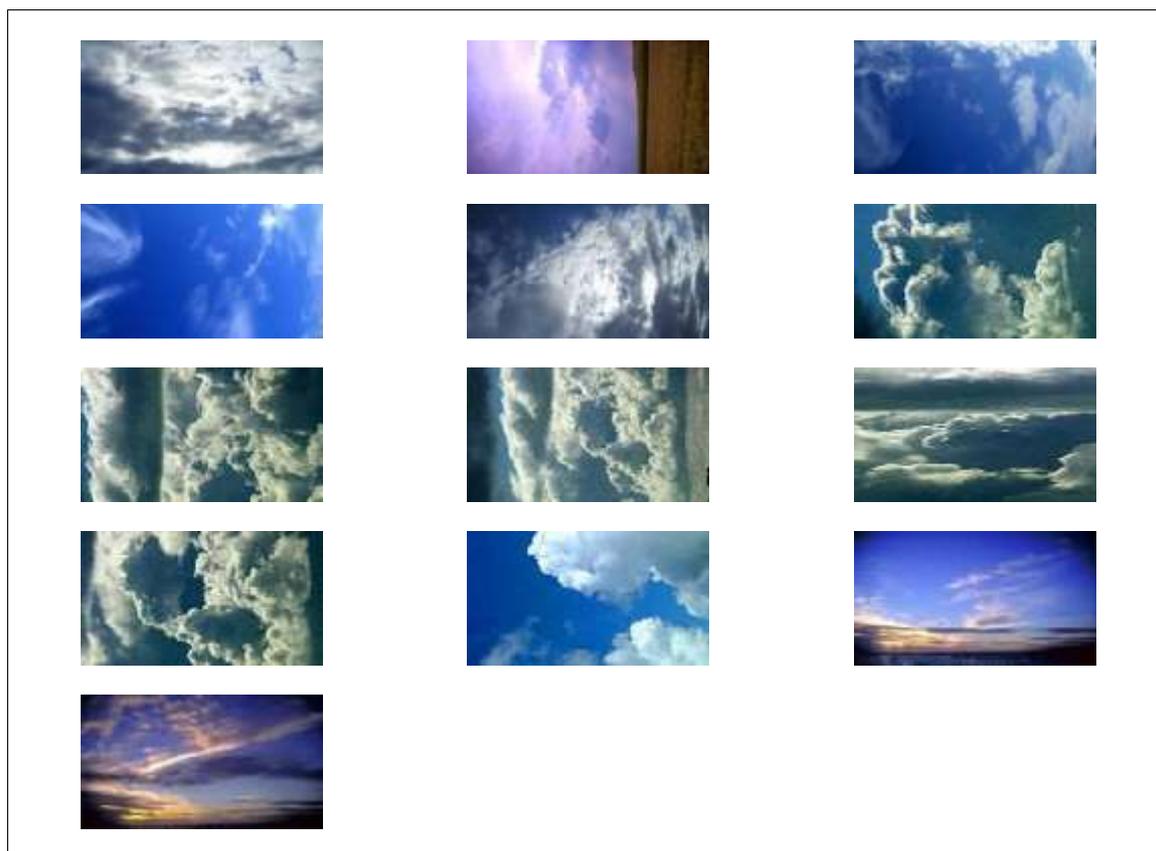


Figura A.44 - Classe 25 referente as imagens Céu – Imagens de 31 a 43.

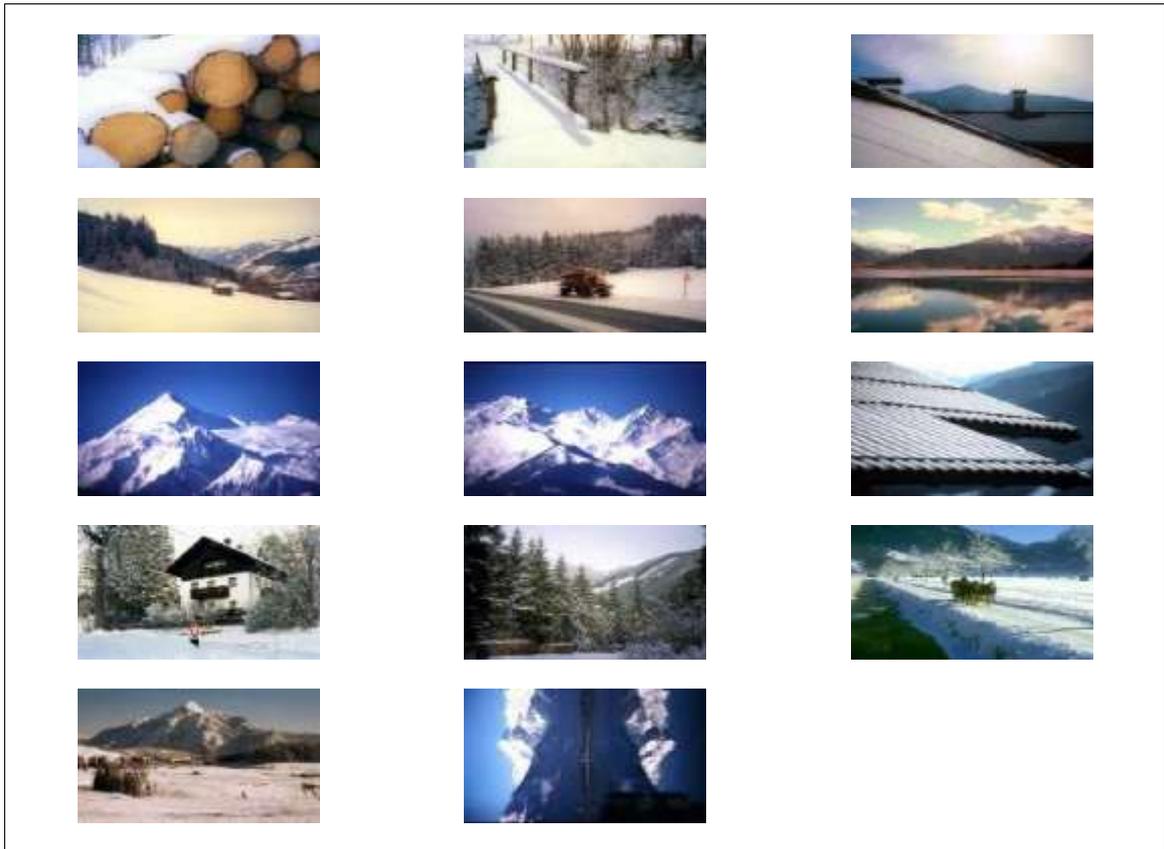


Figura A.45 - Classe 26 referente as imagens Neve – Imagens de 1 a 14.

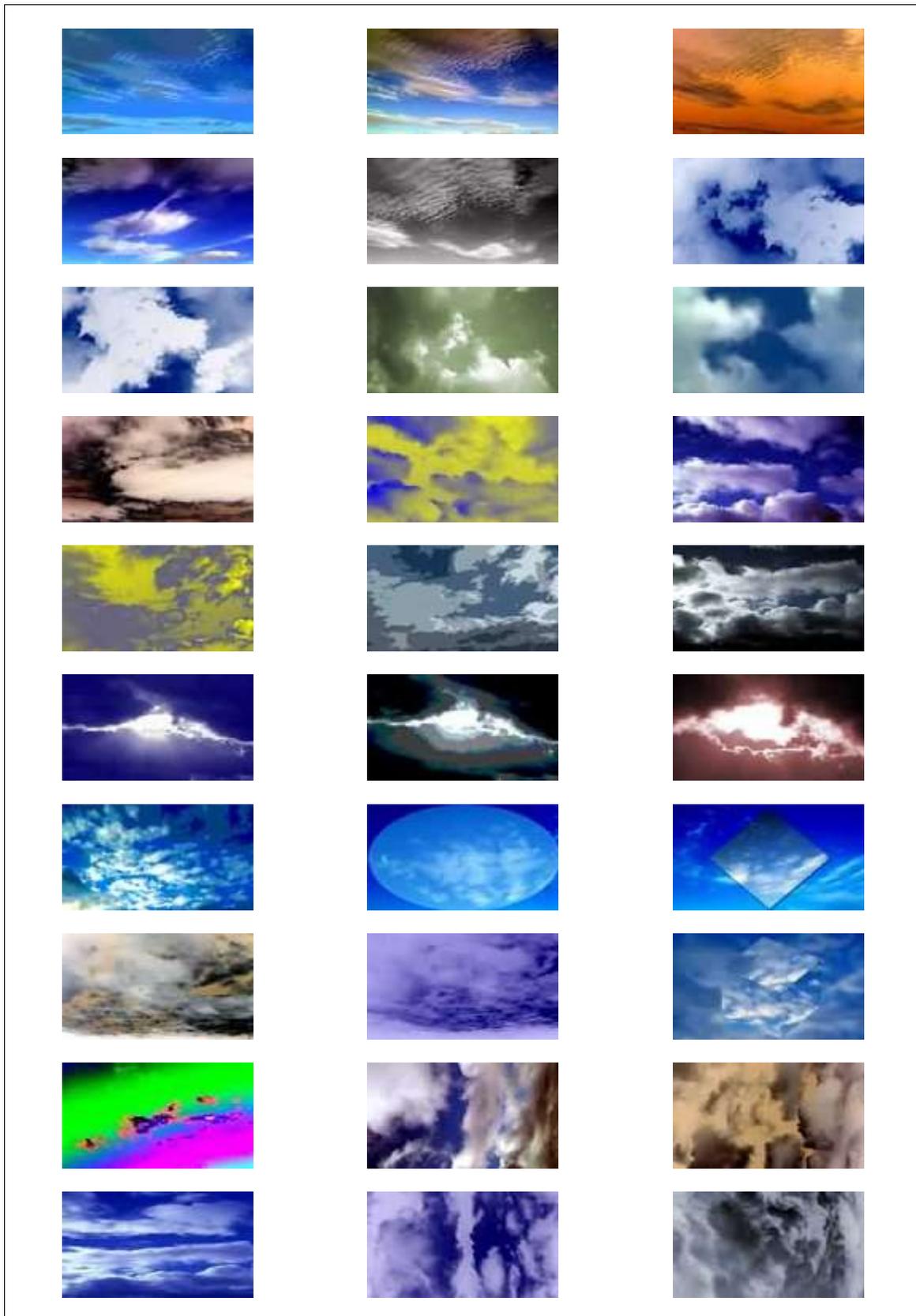


Figura A.46 - Classe 27 referente as imagens Textura – Imagens de 1 a 30.



Figura A.47 - Classe 27 referente as imagens Textura – Imagens de 31 a 32.