

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE

CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA

Coordenação de Pós-Graduação em Ciência da Computação

DISSERTAÇÃO DE MESTRADO

AGRUPAMENTO DE FACES EM COLEÇÕES DE FOTOGRAFIAS DIGITAIS

EDUARDO SANTIAGO MOURA

ORIENTADORES

HERMAN MARTINS GOMES

JOÃO MARQUES DE CARVALHO

CAMPINA GRANDE, PARAÍBA

FEVEREIRO – 2012

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE

CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA

Coordenação de Pós-Graduação em Ciência da Computação

AGRUPAMENTO DE FACES EM COLEÇÕES DE FOTOGRAFIAS DIGITAIS

EDUARDO SANTIAGO MOURA

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande – Campus I como parte dos requisitos necessários à obtenção do grau de Mestre em Ciência da Computação (MSc).

Área de concentração: Ciência da Computação

Linha de pesquisa: Metodologia de Técnicas de Computação

Herman Martins Gomes
João Marques de Carvalho
Orientadores

Campina Grande - Paraíba
Fevereiro – 2012

DIGITALIZAÇÃO:
SISTEMOTECA - UFCG

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

M929a Moura, Eduardo Santiago.

Agrupamento de faces em coleções de fotografias digitais / Eduardo Santiago Moura. – Campina Grande, 2012.

159f.: il. col.

Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática.

Orientadores: Prof. Ph.D. Herman Martins Gomes

Prof. Ph.D. João Marques de Carvalho.

Referências.

1. Visão Computacional. 2. Processamento de imagens. 3. Agrupamento de Faces. I. Título.

CDU 004.8(043)


"AGRUPAMENTO DE FACES EM COLEÇÕES DE FOTOGRAFIAS DIGITAIS"

EDUARDO SANTIAGO MOURA


DISSERTAÇÃO APROVADA EM 28/02/2012



HERMAN MARTINS GOMES, Ph.D
Orientador(a)



JOÃO MARQUES DE CARVALHO, Ph.D
Orientador(a)



JOSÉ EUSTAQUIO RANGEL DE QUEIROZ, D.Sc
Examinador(a)

GEORGE DARMITON DA CUNHA CAVALCANTI, Dr.
Examinador(a)

CAMPINA GRANDE - PB

Declaro para os devidos fins que participei como examinador externo da banca de mestrado de *Eduardo Santiago Moura*, intitulada *Agrupamento de Faces em Coleções de Fotografais Digitais*.

A banca ocorreu no dia 28 de fevereiro de 2011, na UFCG, e minha participação se deu por meio virtual usando a ferramenta Skype.

Após a análise da dissertação e da apresentação, considero Aprovado o trabalho de Eduardo Santiago Moura, estabelecendo um prazo de 30 dias para que as correções solicitadas sejam implementadas.

Recife-PE, 28 de fevereiro de 2012.



George Darmiton da Cunha Cavalcanti
Professor Adjunto
CIn-UFPE

AGRADECIMENTOS

Primeiramente, agradeço a Deus, por sempre iluminar minha caminhada, até mais essa conquista em minha vida.

À minha família e, principalmente, à minha esposa Luana, por todo apoio, força e confiança depositada para que esse objetivo de vida pudesse ser realizado. Em especial, ao meu querido avô José Milton Santiago (*in memoriam*).

Aos orientadores Herman Martins Gomes e João Marques de Carvalho, pelos ensinamentos, dedicação, paciência e orientação, durante toda minha caminhada, desde a graduação até, futuramente, o doutorado.

Ao Professor Eustáquio Rangel, pelos ensinamentos, pelo incentivo e pela frase motivacional semanal: "E aí? Está tudo pronto?".

Aos membros da banca examinadora, Professor Eustáquio Rangel e Professor George Darmiton, pelas críticas e sugestões que contribuíram para o enriquecimento deste trabalho.

Aos meus amigos do Laboratório de Visão Computacional – LVC (Claúdio, Eanes, Paulo, Xycho) e dissidentes (Bosco, Brito, Luciana, Odilon e Vinicius), pelos momentos de descontração durante a realização deste trabalho.

À equipe da COPIN (professores e funcionários), por garantir a infraestrutura necessária aos mestrandos em Ciência da Computação da UFCG.

O presente trabalho de dissertação foi alcançado em cooperação com a Hewlett-Packard Brasil Ltda. e com recursos provenientes da Lei de Informática (Lei nº 8.248, de 1991).

RESUMO

Como consequência da popularização da *Internet* (impulsionada pelas redes sociais, tais como *FaceBook*, *Orkut*, *MySpace*) e da crescente proliferação de câmeras digitais e dispositivos móveis, a organização automática de grandes coleções de fotografias digitais torna-se um recurso de extrema relevância. Sistemas tradicionais consideram apenas informações simples (e.g., data, nome de arquivo e diretório) para auxiliar na tarefa de organização. No entanto, para grandes coleções, tipicamente formadas por milhares de imagens, estas informações são insuficientes para produzir bons níveis de organização e de satisfação. A maioria das técnicas avançadas nesta área busca analisar o conteúdo das imagens e extrair informações de mais alto nível como, por exemplo, faces. Neste sentido, faces ocupam um papel preponderante, dada a importância que desempenham nas relações humanas. Portanto, no escopo de fotografias contendo faces, o agrupamento por faces é um tópico de grande relevância. Diante deste contexto, esta dissertação objetiva tratar do problema de agrupamento de faces, buscando: (i) a obtenção de melhor desempenho em relação ao estado da arte de técnicas para o agrupamento de faces no contexto de fotografia digital de consumo; e (ii) investigar formas de minimizar a degradação normalmente associada a variações nas imagens de faces (tais como iluminação, expressões faciais e pose). A abordagem proposta para atingir os objetivos supracitados é constituída por uma etapa de pré-processamento, seguida por etapas de extração de características SURF (*Speeded Up Robust Features*) e agrupamento. A partir de um estudo experimental e de testes estatísticos, nos quais se comparou a abordagem proposta com três sistemas comerciais, evidenciaram-se a existência de diferenças estatisticamente significativas entre os resultados gerados em favor da abordagem proposta.

Palavras-chave: agrupamento de faces, agrupamento aglomerativo hierárquico, métricas de avaliação de agrupamento, avaliação estatística.

ABSTRACT

As a result of the popularization of the *Internet* (motivated by social networks like *Facebook*, *Orkut*, *MySpace*) and the increasing proliferation of digital cameras and mobile devices, the automatic organization of large digital photo albums has become an extremely relevant resource. Traditional systems use only simple information (such as date, file and folder name) to help with the organization task. However, for large collections, typically formed by millions of images, this information is insufficient to achieve good levels of organization and satisfaction. Most advanced techniques in this area aim to analyze image content and to extract high level information, e.g., faces. In this sense, faces occupy a preponderant role, given their importance to human relations. Therefore, within the scope of photographs containing faces, face clustering is a very relevant topic. Within this context, this dissertation aims to address the problem of face clustering, while seeking: (i) to obtain better performance over the state of the art techniques in face clustering, and (ii) to investigate ways to minimize degradation usually associated with variations in face images (such as lighting, facial expressions and pose). The proposed approach to reach the above goals is composed of a preprocessing step followed by SURF (*Speeded Up Robust Features*) feature extraction and clustering steps. From an experimental study and statistical tests, in which the proposed approach and three commercial applications were compared, statistically significant differences between the generated results were inferred, with better results obtained by the proposed approach.

Keywords: face clustering, hierarchical agglomerative clustering, clustering validity measures, statistical evaluation.

CONTEÚDO

| | |
|---|-----------|
| 1. INTRODUÇÃO | 1 |
| 1.1. MOTIVAÇÕES | 2 |
| 1.2. DESCRIÇÃO DO PROBLEMA | 4 |
| 1.3. FORMALIZAÇÃO DO PROBLEMA | 5 |
| 1.4. OBJETIVOS E RELEVÂNCIA | 7 |
| 1.5. ESTRUTURA DA DISSERTAÇÃO | 8 |
| 2. TRABALHOS RELACIONADOS | 10 |
| 2.1. METODOLOGIA DE PESQUISA BIBLIOGRÁFICA..... | 10 |
| 2.2. RECONHECIMENTO SUPERVISIONADO E NÃO SUPERVISIONADO | 11 |
| 2.3. PROPÓSITO DO SISTEMA | 11 |
| 2.4. CARACTERÍSTICAS EXTRAÍDAS PARA IDENTIFICAÇÃO PESSOAL..... | 12 |
| 2.5. MÉTRICAS DE SIMILARIDADE..... | 22 |
| 2.6. TÉCNICAS DE AGRUPAMENTO/RECONHECIMENTO | 27 |
| 2.7. MÉTRICAS DE AVALIAÇÃO DE AGRUPAMENTO..... | 34 |
| 2.8. CONSIDERAÇÕES SOBRE OS TRABALHOS ANALISADOS | 37 |
| 2.9. CONSIDERAÇÕES FINAIS | 43 |
| 3. ABORDAGEM PROPOSTA | 45 |
| 3.1. VISÃO GERAL DA ARQUITETURA PROPOSTA | 45 |
| 3.2. PRÉ-PROCESSAMENTO | 47 |
| 3.2.1. Detecção e Correção da Orientação de Faces | 48 |
| 3.2.2. Equalização e Compensação de Iluminação..... | 49 |
| 3.2.2.1. Filtragem Homomórfica..... | 50 |
| 3.2.2.2. Equalização de Histograma | 51 |
| 3.3. PROCESSAMENTO PRINCIPAL..... | 53 |
| 3.3.1. Extração de Características Faciais | 53 |
| 3.3.1.1. Speeded Up Robust Features (SURF) | 55 |
| Imagem Integral..... | 55 |
| Matriz Hessiana baseada em Pontos de Interesse | 56 |
| Representação Espaço-Escala..... | 58 |
| Localização de Pontos de Interesse..... | 59 |
| Descritor de Pontos de Interesse | 61 |
| 3.3.2. Comparação e Determinação de Similaridade..... | 64 |
| 3.3.3. Técnica de Agrupamento..... | 67 |
| 3.4. DETALHES DE PROJETO E IMPLEMENTAÇÃO | 69 |
| 3.5. CONSIDERAÇÕES FINAIS | 70 |
| 4. AVALIAÇÃO EXPERIMENTAL | 72 |
| 4.1. BASE DE IMAGENS | 73 |

| | |
|--|------------|
| 4.2. EXPERIMENTO DE CALIBRAÇÃO | 75 |
| 4.2.1. <i>Teste de Normalidade e Teste Visual</i> | 77 |
| 4.2.2. <i>Teste F ANOVA e Teste de Tukey</i> | 82 |
| 4.3. EXPERIMENTO DE COMPARAÇÃO | 87 |
| 4.3.1. <i>Teste de Normalidade e Teste Visual</i> | 91 |
| 4.3.2. <i>Teste F ANOVA e Teste de Tukey</i> | 95 |
| 4.4. CONSIDERAÇÕES FINAIS | 99 |
| 5. CONSIDERAÇÕES FINAIS..... | 101 |
| 5.1. SÍNTESE DA PESQUISA | 101 |
| 5.2. CONTRIBUIÇÕES DO TRABALHO..... | 102 |
| 5.3. TRABALHOS FUTUROS | 104 |
| APÊNDICE A IMAGENS UTILIZADAS NOS EXPERIMENTOS | 115 |
| APÊNDICE B TESTES ESTATÍSTICOS..... | 120 |
| B.1. TESTE DE NORMALIDADE JARQUE-BERA | 120 |
| B.2. TESTE VISUAL BOXPLOT | 121 |
| B.3. TESTE F ANOVA FATOR ÚNICO | 122 |
| B.4. TESTE DE TUKEY | 124 |
| APÊNDICE C RESULTADOS NUMÉRICOS DO EXPERIMENTO DE CALIBRAÇÃO | 125 |
| APÊNDICE D RESULTADOS NUMÉRICOS DO EXPERIMENTO DE COMPARAÇÃO | 137 |
| APÊNDICE E RESULTADOS DETALHADOS OBTIDOS A PARTIR DO USO DA ABORDAGEM PROPOSTA..... | 143 |
| E.1. RESULTADOS DA PARTIÇÃO 1 | 144 |
| E.2. RESULTADOS DA PARTIÇÃO 2 | 146 |
| E.3. RESULTADOS DA PARTIÇÃO 3 | 148 |
| E.4. RESULTADOS DA PARTIÇÃO 4 | 150 |
| E.5. RESULTADOS DA PARTIÇÃO 5 | 152 |
| E.6. RESULTADOS DA PARTIÇÃO 6 | 154 |
| E.7. RESULTADOS DA PARTIÇÃO 7 | 156 |
| E.8. RESULTADOS DA PARTIÇÃO 8 | 158 |

LISTA DE FIGURAS

| | |
|---|----|
| Figura 1.1 – Agrupamento de faces: (A) Coleção de fotografias; e (B) Grupos de faces resultantes. | 4 |
| Figura 2.1 – Alguns pontos de interesse para o reconhecimento de faces: (A) Imagem de entrada (B) Pontos de interesse. | 12 |
| Figura 2.2 – Problemas em imagens de faces no tocante a variações de: (A) Pose; (B) Oclusão; (C) Iluminação; (D) Expressão facial; e (E) Condições de aquisição da imagem. | 13 |
| Figura 2.3 – Ilustração dos diferentes papéis das características globais e locais no reconhecimento de faces. | 14 |
| Figura 2.4 – Pontos fiduciais utilizados para composição da representação facial. | 15 |
| Figura 2.5 – Abordagem utilizada por Zhang et al. (2011): (A) Grade 3x3 e (B) Regiões de interesse. | 16 |
| Figura 2.6 – Abordagem utilizada por Palit et al. (2009): (A) Pontos fiduciais; (B) Regiões de interesse; e (C) Regiões extraídas. | 18 |
| Figura 2.7 – Transformada de <i>Hough</i> de uma face: (A) Imagem de entrada; (B) Bordas da face; e (C) Imagem de entrada com quatro FFL extraídas. | 21 |
| Figura 2.8 – Blocos de características extraídas. | 21 |
| Figura 2.9 – Exemplo de um dendrograma. | 29 |
| Figura 2.10 – Representação de $k = 10$ vizinhos para x_i e x_j , e 24 pares de vizinhos que possuem o mesmo nome e contribuem para o cálculo da probabilidade marginal. | 31 |
| Figura 2.11 – Três funções de pertinência criadas pela projeção dos grupos nos eixos dos dados de entrada. | 33 |
| Figura 2.12 – Menu <i>popup</i> do sistema de anotação de faces. | 34 |
| Figura 3.1 – Arquitetura da técnica proposta. | 46 |
| Figura 3.2 – Etapa 1 – Módulo de detecção e correção da orientação de faces. | 48 |
| Figura 3.3 – Etapa 2 – Módulo de equalização e compensação de iluminação. | 50 |
| Figura 3.4 – Resumo dos passos da filtragem homomórfica. | 51 |
| Figura 3.5 – Exemplo de compensação de iluminação: (A) Imagem normalizada; e (B) Imagem após filtragem homomórfica. | 51 |
| Figura 3.6 – Exemplo de equalização de histograma: (A) Imagem após filtragem homomórfica; e (B) Imagem após equalização de histograma. | 53 |
| Figura 3.7 – Etapa 3 – Módulo de extração de características faciais. | 54 |
| Figura 3.8 – Cálculo de área retangular utilizando imagens integrais. | 55 |

| | |
|--|-----|
| Figura 3.9 – Aproximação LoG. Linha superior: derivadas Gaussianas de segunda ordem, discretizadas e recortadas, nas direções x, y e xy, ou seja, L_{xx} , L_{yy} e L_{xy} , respectivamente. Linha inferior: aproximações de filtros representados por caixas nas direções x, y e xy, ou seja, D_{xx} , D_{yy} e D_{xy} , respectivamente. | 57 |
| Figura 3.10 – Ao invés de iterativamente reduzir o tamanho da imagem (esquerda), o uso de imagens integrais permite o aumento da escala do filtro em tempo constante (direita). | 59 |
| Figura 3.11 – Tamanhos dos filtros subseqüentes devem diferir por um mínimo de 6 pixels para preservar a estrutura. Filtros 9x9, 15x15 e 21x21, em seqüência. | 60 |
| Figura 3.12 – Supressão não-máxima. O pixel marcado com 'X' é selecionado como uma máxima se for maior do que os pixels adjacentes em seu intervalo e nos intervalos acima e abaixo. | 60 |
| Figura 3.13 – Determinação da orientação: uma janela de orientação deslizante de tamanho $\pi/3$ detecta a orientação dominante das respostas das <i>wavelets</i> de Haar em cada ponto da amostra dentro de uma vizinhança circular em torno do ponto de interesse. | 62 |
| Figura 3.14 – Janelas do descritor. O tamanho da janela do descritor é 20 vezes a dimensão do ponto detectado e está orientada ao longo da direção dominante ilustrada em verde. | 63 |
| Figura 3.15 – Para construir o descritor, uma grade orientada quadrática com 4x4 sub-regiões é colocada sobre o ponto de interesse (esquerda). Para cada subregião, as respostas <i>Wavelets</i> são computadas a partir de amostras de 5x5. Para cada subregião, são computadas as somas dx, dx , dy, e dy , calculadas relativamente à orientação da grade (direita). | 63 |
| Figura 3.16 – Etapa 4 – Módulo de comparação e determinação de similaridade. | 64 |
| Figura 3.17 – Etapa 5 – Módulo de agrupamento de faces. | 67 |
| Figura 3.18 – Diagrama de classes da aplicação desenvolvida. | 70 |
| Figura 4.1 – Amostra de imagens da base de faces IMM: variações de gênero, iluminação, expressão facial e pose. | 74 |
| Figura 4.2 – Gráficos <i>boxplot</i> das abordagens investigadas: (A) RI; (B) ARI; (C) P; (D) R; (E) F; (F) JI; e (G) FM. | 80 |
| Figura 4.3 – Tela principal do Adobe Photoshop Elements. | 90 |
| Figura 4.4 – Tela principal do PittPatt FaceSort. | 90 |
| Figura 4.5 – Tela principal do Google Picasa. | 91 |
| Figura 4.6 – Gráficos <i>boxplot</i> das ferramentas analisadas: (A) RI; (B) ARI; (C) P; (D) R; (E) F; (F) JI; e (G) FM. | 93 |
| Figura A.1 – Imagens contidas na Partição 1. | 115 |

| | |
|--|------------|
| Figura A.2 – Imagens contidas na Partição 2. | 116 |
| Figura A.3 – Imagens contidas na Partição 3. | 116 |
| Figura A.4 – Imagens contidas na Partição 4. | 117 |
| Figura A.5 – Imagens contidas na Partição 5. | 117 |
| Figura A.6 – Imagens contidas na Partição 6. | 118 |
| Figura A.7 – Imagens contidas na Partição 7. | 118 |
| Figura A.8 – Imagens contidas na Partição 8. | 119 |
| Figura B.1 – Representação de um gráfico <i>boxplot</i>. | 121 |
| Figura E.1 – Grupos de faces obtidos na Partição 1. | 144 |
| Figura E.2 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 1. | 144 |
| Figura E.3 – Dendrograma dos grupos de faces obtidos na Partição 1. | 145 |
| Figura E.4 – Grupos de faces obtidos na Partição 2. | 146 |
| Figura E.5 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 2. | 146 |
| Figura E.6 – Dendrograma dos grupos de faces obtidos na Partição 2. | 147 |
| Figura E.7 – Grupos de faces obtidos na Partição 3. | 148 |
| Figura E.8 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 3. | 148 |
| Figura E.9 – Dendrograma dos grupos de faces obtidos na Partição 3. | 149 |
| Figura E.10 – Grupos de faces obtidos na Partição 4. | 150 |
| Figura E.11 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 4. | 150 |
| Figura E.12 – Dendrograma dos grupos de faces obtidos na Partição 4. | 151 |
| Figura E.13 – Grupos de faces obtidos na Partição 5. | 152 |
| Figura E.14 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 5. | 152 |
| Figura E.15 – Dendrograma dos grupos de faces obtidos na Partição 5. | 153 |
| Figura E.16 – Grupos de faces obtidos na Partição 6. | 154 |
| Figura E.17 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 6. | 154 |
| Figura E.18 – Dendrograma dos grupos de faces obtidos na Partição 6. | 155 |
| Figura E.19 – Grupos de faces obtidos na Partição 7. | 156 |
| Figura E.20 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 7. | 156 |
| Figura E.21 – Dendrograma dos grupos de faces obtidos na Partição 7. | 157 |
| Figura E.22 – Grupos de faces obtidos na Partição 8. | 158 |
| Figura E.23 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 8. | 158 |
| Figura E.24 – Dendrograma dos grupos de faces obtidos na Partição 8. | 159 |

LISTA DE TABELAS

| | |
|---|------------|
| Tabela 2.1 – Resumo dos trabalhos analisados. | 39 |
| Tabela 4.1 – Teste de normalidade Jarque-Bera para a métrica RI das abordagens investigadas. | 78 |
| Tabela 4.2 – Teste de normalidade Jarque-Bera para a métrica ARI das abordagens investigadas. | 78 |
| Tabela 4.3 – Teste de normalidade Jarque-Bera para a métrica P das abordagens investigadas. | 78 |
| Tabela 4.4 – Teste de normalidade Jarque-Bera para a métrica R das abordagens investigadas. | 79 |
| Tabela 4.5 – Teste de normalidade Jarque-Bera para a métrica F das abordagens investigadas. | 79 |
| Tabela 4.6 – Teste de normalidade Jarque-Bera para a métrica JI das abordagens investigadas. | 79 |
| Tabela 4.7 – Teste de normalidade Jarque-Bera para a métrica FM das abordagens investigadas. | 80 |
| Tabela 4.8 – Teste F ANOVA para as sete métricas de avaliação das abordagens investigadas. | 82 |
| Tabela 4.9 – Teste de normalidade Jarque-Bera para a métrica RI das ferramentas comparadas. | 92 |
| Tabela 4.10 – Teste de normalidade Jarque-Bera para a métrica ARI das ferramentas comparadas. | 92 |
| Tabela 4.11 – Teste de normalidade Jarque-Bera para a métrica P das ferramentas comparadas. | 92 |
| Tabela 4.12 – Teste de normalidade Jarque-Bera para a métrica R das ferramentas comparadas. | 92 |
| Tabela 4.13 – Teste de normalidade Jarque-Bera para a métrica F das ferramentas comparadas. | 92 |
| Tabela 4.14 – Teste de normalidade Jarque-Bera para a métrica JI das ferramentas comparadas. | 93 |
| Tabela 4.15 – Teste de normalidade Jarque-Bera para a métrica FM das ferramentas comparadas. | 93 |
| Tabela 4.16 – Teste F ANOVA para as sete métricas de avaliação das ferramentas analisadas. | 95 |
| Tabela C.1 – Abordagens investigadas nos testes estatísticos do experimento de calibração. | 125 |

| | |
|--|------------|
| Tabela C.2 – Resultados numéricos da abordagem Histograma..... | 126 |
| Tabela C.3 – Resultados numéricos da abordagem LBP..... | 126 |
| Tabela C.4 – Resultados numéricos da abordagem Máximo..... | 127 |
| Tabela C.5 – Resultados numéricos da abordagem Norm(C). | 127 |
| Tabela C.6 – Resultados numéricos da abordagem Média..... | 128 |
| Tabela C.7 – Resultados numéricos da abordagem Correlação. | 128 |
| Tabela C.8 – Resultados numéricos da abordagem Abs(C)..... | 129 |
| Tabela C.9 – Resultados numéricos da abordagem Mínimo. | 129 |
| Tabela C.10 – Teste de Tukey para valores de RI das abordagens investigadas. | 130 |
| Tabela C.11 – Teste de Tukey para valores de ARI das abordagens investigadas. | 131 |
| Tabela C.12 – Teste de Tukey para valores de P das abordagens investigadas... | 132 |
| Tabela C.13 – Teste de Tukey para valores de R das abordagens investigadas. . | 133 |
| Tabela C.14 – Teste de Tukey para valores de F das abordagens investigadas... | 134 |
| Tabela C.15 – Teste de Tukey para valores de JI das abordagens investigadas.. | 135 |
| Tabela C.16 – Teste de Tukey para valores de FM das abordagens investigadas. | 136 |
| Tabela D.1 – Ferramentas comparadas nos testes estatísticos do experimento de calibração..... | 137 |
| Tabela D.2 – Resultados numéricos da ferramenta Photoshopel. | 138 |
| Tabela D.3 – Resultados numéricos da ferramenta Picasa..... | 138 |
| Tabela D.4 – Resultados numéricos da ferramenta FaceSort..... | 139 |
| Tabela D.5 – Resultados numéricos da abordagem proposta PhotoCluster. | 139 |
| Tabela D.6 – Teste de Tukey para valores de RI das ferramentas comparadas. . | 140 |
| Tabela D.7 – Teste de Tukey para valores de ARI das ferramentas comparadas. | 140 |
| Tabela D.8 – Teste de Tukey para valores de P das ferramentas comparadas. | 140 |
| Tabela D.9 – Teste de Tukey para valores de R das ferramentas comparadas.... | 141 |
| Tabela D.10 – Teste de Tukey para valores de F das ferramentas comparadas. . | 141 |
| Tabela D.11 – Teste de Tukey para valores de JI das ferramentas comparadas. | 141 |
| Tabela D.12 – Teste de Tukey para valores de FM das ferramentas comparadas. | 142 |
| Tabela E.1 – Resultados por classe dos grupos de faces obtidos na Partição 1. . | 145 |
| Tabela E.2 – Resultados por classe dos grupos de faces obtidos na Partição 2. . | 147 |
| Tabela E.3 – Resultados por classe dos grupos de faces obtidos na Partição 3. . | 149 |
| Tabela E.4 – Resultados por classe dos grupos de faces obtidos na Partição 4. . | 151 |
| Tabela E.5 – Resultados por classe dos grupos de faces obtidos na Partição 5. . | 153 |
| Tabela E.6 – Resultados por classe dos grupos de faces obtidos na Partição 6. . | 155 |
| Tabela E.7 – Resultados por classe dos grupos de faces obtidos na Partição 7. . | 157 |
| Tabela E.8 – Resultados por classe dos grupos de faces obtidos na Partição 8. . | 159 |

LISTA DE QUADROS

| | |
|--|-----------|
| Quadro 4.1 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica RI das abordagens investigadas. | 84 |
| Quadro 4.2 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica ARI das abordagens investigadas. | 84 |
| Quadro 4.3 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica P das abordagens investigadas. | 85 |
| Quadro 4.4 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica R das abordagens investigadas. | 85 |
| Quadro 4.5 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica F das abordagens investigadas. | 86 |
| Quadro 4.6 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica JI das abordagens investigadas. | 86 |
| Quadro 4.7 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica FM das abordagens investigadas. | 87 |
| Quadro 4.8 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica RI das ferramentas analisadas. | 96 |
| Quadro 4.9 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica ARI das ferramentas analisadas. | 97 |
| Quadro 4.10 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica P das ferramentas analisadas. | 97 |
| Quadro 4.11 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica R das ferramentas analisadas. | 97 |
| Quadro 4.12 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica F das ferramentas analisadas. | 98 |
| Quadro 4.13 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica JI das ferramentas analisadas. | 98 |
| Quadro 4.14 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica FM das abordagens investigadas. | 99 |

LISTA DE SIGLAS E ABREVIATURAS

| | |
|---------------|---|
| ANOVA | <i>AN</i>alysis <i>Of</i> <i>VA</i>riance |
| ARI | <i>Adjusted Rand Index</i> |
| DoG | <i>Difference of Gaussians</i> |
| EM | <i>Expectation-Maximization</i> |
| F | <i>F-Measure</i> |
| FM | <i>Folks and Mallows Index</i> |
| gl | graus de liberdade |
| GT | <i>Ground-Truth</i> |
| IC | Intervalo de Confiança |
| JI | <i>Jaccard Index</i> |
| JPEG | <i>Joint Photographic Experts Group</i> |
| HAC | <i>Hierarchical Agglomerative Clustering</i> |
| LDA | <i>Linear Discriminant Analysis</i> |
| LoG | <i>Laplacian of Gaussians</i> |
| MQ | Média de Quadrados |
| OpenCV | <i>Open Source Computer Vision (Intel Library)</i> |
| PCA | <i>Principal Component Analysis</i> |
| PDI | Processamento Digital de Imagens |
| PR | <i>Precision-Recall</i> |
| RI | <i>Rand Index</i> |
| VC | Visão Computacional |
| SIFT | <i>Scale-Invariant Feature Transform</i> |
| SQ | Soma de Quadrados |
| SURF | <i>Speeded Up Robust Features</i> |

Capítulo 1

Introdução

Neste capítulo, são introduzidos os elementos fundamentais para a realização desta pesquisa. Estes elementos são apresentados por meio da apresentação e formalização do problema em estudo – agrupamento de faces em coleções de fotografias digitais – bem como do contexto e das principais motivações que justificam a importância deste trabalho na área de *Visão Computacional (VC)*.

As faces humanas são algumas das entidades mais importantes e freqüentemente encontradas em fotografias e podem ser consideradas como características semânticas de alto nível (ANTONOPOULOS, NIKOLAIDIS e PITAS, 2007), sendo preponderantes para as relações humanas. Aliada à popularidade crescente de fotografias digitais, a organização de álbuns de fotografias pessoais tornou-se um tema de pesquisa ativa. Esforços existentes nesta área têm sido dedicados ao uso de reconhecimento facial para auxiliar no processo de anotação das pessoas. No entanto, algoritmos de agrupamento de faces atuais ainda não são robustos às variações de aparência presentes em fotografias reais (LIN et al., 2010).

Diante do exposto, nesta dissertação são discutidos os trabalhos no âmbito de agrupamento de faces no contexto de fotografia digital de consumo, dando-se ênfase à obtenção do melhor desempenho em relação ao estado da arte das técnicas destinadas ao agrupamento de faces, com o intuito de facilitar a organização e o compartilhamento de coleções de fotografias de usuários.

As seções remanescentes deste capítulo estão divididas como segue. Na Seção 1.1, argumenta-se sobre as motivações para a pesquisa. A descrição e a formalização do problema a ser resolvido são apresentadas

nas Seções 1.2 e 1.3, respectivamente. Os objetivos a serem alcançados e a relevância desta pesquisa são apresentadas na Seção 1.4. Finalmente, na Seção 1.5, é apresentada a estrutura desta dissertação.

1.1. Motivações

A face desempenha um papel importante na interação humana, visto que expressões faciais podem ser vistas como uma linguagem universal facilitadora do compartilhamento de informação e comunicação entre pessoas (WECHSLER, 2006). Além disto, com a popularização da *Internet*, impulsionada pelas redes sociais como *FaceBook*¹, *Orkut*², *MySpace*³, somada à crescente proliferação de câmeras digitais e dispositivos móveis multimídia, o cotidiano das pessoas, incluindo sobremaneira suas faces, passou a ser facilmente capturado, armazenado, pesquisado e compartilhado (CHOI et al., 2008).

Como conseqüência desta demanda, a quantidade de coleções de fotografias digitais vem crescendo consideravelmente (YOUSEF e O'NEILL, 2007). Como exemplo, a rede social considerada como a mais popular no mundo, *FaceBook*, no ano de 2007, contava com 70 milhões de usuários ativos e armazenava 1,7 bilhão de fotos, nas quais, 2,2 bilhões de anotações de pessoas estavam presentes nas fotografias (THE FACEBOOK BLOG, 2007; STONE, ZICKLER e DARRELL, 2008). Atualmente, a rede apresenta mais de 500 milhões de usuários ativos e armazena mais de 15 bilhões de fotos, com uma média de 220 milhões de novas fotos sendo postadas a cada semana, confirmando tal tendência de crescimento (OSTROW, 2009; FACEBOOK STATISTICS, 2010).

Os fatos supramencionados despertaram o interesse de inúmeras empresas para o desenvolvimento de aplicativos de *software* online, tais como, *Snapfish*⁴, *Face.com*⁵, *MyHeritage*⁶, *Flickr*⁷ (CAO et al., 2010). Devido ao aumento significativo das coleções de fotos pessoais, a organização

¹ FaceBook, disponível em: <http://www.facebook.com>

² Orkut, disponível em: <http://www.orkut.com>

³ MySpace, disponível em: <http://www.myspace.com>

⁴ Snapfish by HP, disponível em: <http://www.snapfish.com>

⁵ Face.com, disponível em: <http://face.com>

⁶ MyHeritage, disponível em: <http://www.myheritage.com>

⁷ Flickr by Yahoo!, disponível em: <http://www.flickr.com>

automática de grandes álbuns de fotografias digitais está se tornando uma necessidade extremamente importante (CHEN et al., 2003; ZHANG et al., 2003; CUI et al., 2007; KAPOOR et al., 2009; LIN et al., 2010).

Assim, para uma organização mais eficiente, uma simples solução está no uso de palavras-chave (anotações ou *tags*) semânticas para o gerenciamento das fotografias, tais como: "*quem*" está na foto, "*quando*" a foto foi tirada, "*onde*" foi tirada, "*qual o tipo*" da foto (retrato, grupo de pessoas ou paisagem) e quaisquer outras informações contextuais relevantes (LEI et al., 2003). Segundo Sharavanan e Azath (2009), devido a grandes mudanças ocasionadas na face por variações de iluminação, pose, expressão facial, oclusão e efeitos de envelhecimento, as técnicas e métodos existentes de reconhecimento de faces, de cena e de evento não podem garantir um resultado preciso se o processo de anotação for totalmente automático.

Choi et al. (2008) afirmaram que, para agrupar e indexar fotos pessoais com base no entendimento de "*quem*" está nas fotos, o uso de anotações das faces se faz essencial. A maioria dos aplicativos de *software* atual dá suporte à realização manual desta tarefa, permitindo que os usuários rotulem fotos uma a uma, ou selecionem manualmente as fotos que possivelmente possuem o mesmo rótulo para então, efetuar a aplicação do rótulo informado após a seleção (CUI et al., 2007).

Por outro lado, a anotação manual de faces por usuários é uma tarefa que nem sempre, as pessoas estão dispostas a realizar, por exigir uma grande quantidade de tempo e esforço. Por exemplo, Fu, Hsu e Lee (2009) constataram que usuários do *Flickr* levam em média 5 a 6 segundos para rotular uma fotografia.

Desta forma, a idéia de agrupamento de faces é motivada pela conveniência de fazer tal seleção automaticamente. Muitos aplicativos organizadores de coleções de fotografias digitais, tais como *Picasa*⁸, *FaceSort*⁹, *iPhoto*¹⁰, *Adobe Photoshop Elements*¹¹, *iView Media*¹², *ACDSee*¹³,

⁸ Google Picasa, disponível em: <http://picasa.google.com>

⁹ PittPatt FaceSort, disponível em: <http://demo.pittpatt.com/facesort>

¹⁰ Apple iPhoto, disponível em: <http://www.apple.com/br/ilife/iphoto>

Microsoft Digital Image Suite¹⁴, EasyAlbum¹⁵, digiKam¹⁶ e F-Spot¹⁷ recentemente adicionaram reconhecimento e agrupamento de faces para ajudar parcialmente a automatizar este processo.

1.2. Descrição do Problema

Nos métodos de agrupamento de faces, as características faciais são utilizadas como identificador de cada pessoa, de modo a agrupar várias imagens de faces de uma mesma pessoa em um determinado grupo (*cluster*). A fim de extrair características de identificação, para uma dada pessoa, em uma determinada fotografia, primeiramente a região da face é detectada e, em seguida, são aplicados extratores de características na região detectada, tentando produzir uma representação única da face.

O problema de agrupamento de faces pode ser formulado como segue: dada uma coleção arbitrária de imagens e um conjunto de faces humanas nesta coleção, deve-se agrupar as faces conforme a identidade das pessoas, sem um conhecimento prévio de qualquer uma das pessoas envolvidas. As faces encontradas podem ter sido obtidas sob diferentes condições de imageamento, e.g., pose, iluminação, expressão facial e oclusão parcial (Vide Figura 1.1).

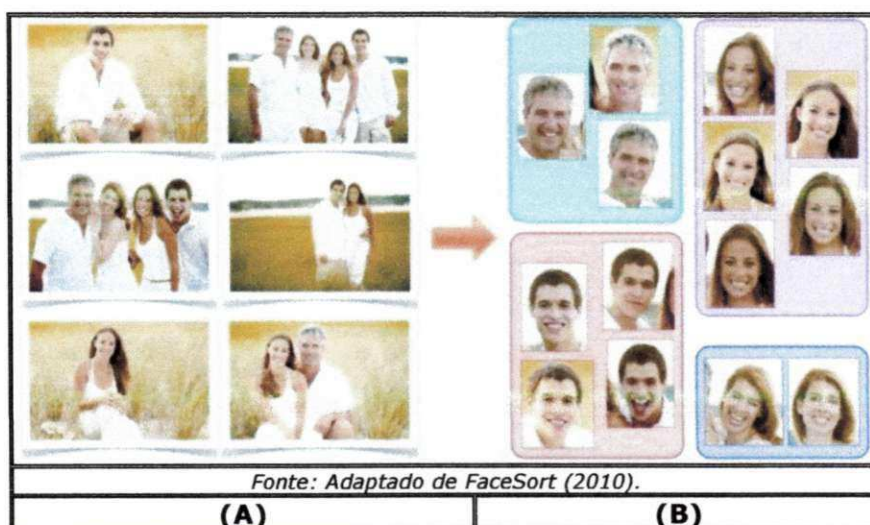


Figura 1.1 – Agrupamento de faces: (A) Coleção de fotografias; e (B) Grupos de faces resultantes.

¹¹ Adobe Photoshop Elements, disponível em: <http://www.adobe.com/br/products/photoshopel>

¹² iView Media, disponível em: <http://www.iview-multimedia.com>

¹³ ACDSee, disponível em: <http://www.acdsee.com>

¹⁴ Microsoft Digital Image Suite, disponível em: <http://www.microsoft.com/products/imaging>

¹⁵ EasyAlbum, disponível em: <http://research.microsoft.com/en-us/groups/vc/easyalbumdownload.aspx>

¹⁶ KDE digiKam, disponível em: <http://www.digikam.org>

¹⁷ GNOME F-Spot, disponível em: <http://f-spot.org>

Assim, o agrupamento de faces pode ser considerado como uma forma de classificação não supervisionada aplicada sobre um conjunto finito de objetos, cujo objetivo é agrupar os objetos em classes, de tal forma que objetos similares sejam colocados no mesmo grupo, enquanto objetos diferentes sejam colocados em grupos diferentes (ANTONOPOULOS, NIKOLAIDIS e PITAS, 2007).

1.3. Formalização do Problema

A estratégia de reconhecimento não supervisionado é adequada ao problema de agrupamento de faces já que, a princípio, não se tem informação prévia sobre as possíveis pessoas que aparecem em uma determinada coleção. Para formular matematicamente o problema, são apresentadas algumas definições e terminologias. Cao et al. (2010) argumentaram que no projeto de um sistema automático para o problema, as principais decisões a serem tomadas recaem na escolha de uma representação facial e de uma métrica de correspondência para a comparação de faces.

Desta forma, sejam f_{global}^i e f_{local}^i as características globais e locais extraídas de imagens da $i^{ésima}$ pessoa de uma coleção e seja $F^i = \{ f_{global}^i, f_{local}^i \}$ a representação facial composta pelo conjunto de características que identificam a $i^{ésima}$ pessoa incluída em um determinado grupo. Além disto, faz-se necessário definir a métrica de correspondência que é uma função de similaridade entre diferentes características de identidade das pessoas. A similaridade entre as representações faciais F^i e F^j da $i^{ésima}$ e $j^{ésima}$ pessoa, respectivamente, pode ser definida como:

$$S_{faces}(F^i, F^j) = w_{global} \cdot S_{global}(f_{global}^i, f_{global}^j) + w_{local} \cdot S_{local}(f_{local}^i, f_{local}^j) \quad (1.1)$$

em que $S_{global}(f_{global}^i, f_{global}^j)$ é a função que mede a similaridade entre as características faciais globais e $S_{local}(f_{local}^i, f_{local}^j)$ é a função que mede a

similaridade entre as características faciais locais. w_{global} e w_{local} são os valores dos pesos que controlam a importância das características globais e locais, respectivamente. Para determinar a eficiência do limiar de agrupamento, o valor da similaridade, $S_{faces}(F^i, F^j)$, pode ser normalizado no intervalo $[0,1]$ (SANTOS e EMBRECHTS, 2009).

Assim, o algoritmo de agrupamento de faces pode ser expresso conforme os seguintes passos:

- (1) Iniciar todos os grupos C_k isoladamente. Se existirem N representações faciais, então $C_k (k=1, \dots, N)$ grupos são criados no estado inicial, compostos unicamente por uma representação facial F^k ;
- (2) Calcular a similaridade, $S_{cluster}(C_m, C_n)$, entre dois grupos C_m e C_n , em que $1 \leq m, n \leq N$ e $m \neq n$, de acordo com a equação:

$$S_{cluster}(C_m, C_n) = \frac{1}{N_{c_m} N_{c_n}} \sum_{F^i \in C_m} \sum_{F^j \in C_n} S_{faces}(F^i, F^j) \quad (1.2)$$

onde N_{c_m} e N_{c_n} , denotam o número de representações faciais incluídas em cada um dos grupos, respectivamente;

- (3) Encontrar os grupos mais próximos, denotados por C'_m e C'_n , que:

$$(C'_m, C'_n) = \underset{C_m, C_n}{arg\ min} D_{cluster}(C_m, C_n), \text{ sendo } m \neq n; \quad (1.3)$$

- (4) Agrupar os dois grupos mais próximos em um único *cluster*, $C'_m = C'_m \cup C'_n$, e remover C'_n ;
- (5) Repetir os passos 2, 3 e 4, enquanto o valor de $D_{cluster}(C_m, C_n)$ não satisfizer um dado critério de parada, e.g., uma quantidade predefinida de grupos.

Considerando que, quanto maior a proximidade (semelhança) entre dois grupos, mais elevado é o valor da medida de similaridade entre eles, a medida de dissimilaridade pode ser interpretada como um conceito complementar à similaridade, ou seja, quanto maior a proximidade entre dois grupos, menor é o valor da medida de dissimilaridade entre eles.

1.4. Objetivos e Relevância

Torna-se evidente o contínuo interesse da indústria na área de agrupamento de faces, fato que pode ser observado pelo financiamento de pesquisas e investimentos em aplicativos de *software* comerciais por grandes empresas ao longo desta década, a exemplo da *Hewlett-Packard (HP)*, *Apple*, *Microsoft*, *Yahoo!* e *Google*. O recente projeto *FaceRec*, ocorrido em 2010 no âmbito da cooperação técnico-científica firmada entre a *Hewlett-Packard* e a Universidade Federal de Campina Grande, com incentivos da Lei de Informática, se insere no contexto do estudo de tais métodos.

Apesar de toda a pesquisa já desenvolvida, o agrupamento de faces humanas, de forma automática, preciso e robusto, ainda constitui um problema em aberto. Dificuldades surgem devido a vários fatores, dentre os quais podem ser destacados: (i) diferenças de resolução da imagem; (ii) variações na escala e orientação da face; (iii) variações nas condições de iluminação; (iv) variações de pose; (v) variações de expressões faciais; e (vi) geração dos grupos de forma automática.

Diante deste contexto, nesta dissertação é proposta uma técnica de agrupamento de faces em coleções de fotografias digitais que objetiva: (i) a obtenção do melhor desempenho em relação ao estado da arte, com o intuito de facilitar a organização e o compartilhamento de coleções de fotografias de usuários; e (ii) a investigação de formas de minimizar a queda de desempenho, normalmente associada a variações de iluminação, expressões faciais e pose.

A técnica tem como diferencial a agregação de módulos para atenuar os efeitos indesejados provocados pela falta de controle das condições do ambiente no momento da captura da imagem e a combinação, de forma ponderada, de abordagens já utilizadas atualmente no contexto de agrupamento de faces em fotografias digitais. Os objetivos específicos são:

- a) Elaborar um modelo de agrupamento de faces por meio de um processo de aprendizagem não-supervisionada para ser utilizado em tarefa de reconhecimento envolvendo um número indeterminado de classes;

- b) Aplicar técnicas de Processamento Digital de Imagens (PDI) ao problema de agrupamento de faces com variações de iluminação, expressão facial e pose;
- c) Investigar e extrair características faciais que auxiliem na resolução do problema supracitado;
- d) Investigar métricas de similaridade que levem a um bom desempenho do sistema;
- e) Avaliar os resultados produzidos pelo sistema a partir de métricas objetivas; e
- f) Validar o estudo, a partir de experimentos comparativos com aplicações concorrentes, mais especificamente, *Adobe Photoshop Elements* (PHOTOSHOPEL, 2010), *Google Picasa* (PICASA, 2010), e *PittPatt Facesort* (FACESORT, 2010).

1.5. Estrutura da Dissertação

O presente documento é composto por um total de cinco capítulos e cinco apêndices. No **Capítulo 2**, apresenta-se um panorama das técnicas de agrupamento de faces em coleções de fotografias digitais, a partir de uma revisão de trabalhos relevantes da área. Descrevem-se, também, técnicas usadas na tarefa específica de extração de características para identificação pessoal, métricas de similaridade, estratégias comumente empregadas para resolver o problema objeto de estudo, métricas de avaliação da qualidade de agrupamento, bem como se delimita o escopo do presente trabalho.

No **Capítulo 3**, descreve-se, de forma detalhada a abordagem proposta para a solução do problema de agrupamento de faces em coleções de fotografias digitais, incluindo-se a arquitetura geral, o fluxo de processamento e o funcionamento de cada módulo da técnica proposta. Adicionalmente, são apresentados alguns detalhes de implementação e de organização interna da aplicação desenvolvida.

No **Capítulo 4**, reúnem-se os experimentos realizados e uma discussão dos resultados obtidos. A apresentação dos experimentos contempla os diversos testes realizados até a obtenção da abordagem

descrita no Capítulo 3. Adicionalmente, o Capítulo 4 contém gráficos comparativos de resultados obtidos a partir de sete métricas objetivas de avaliação de agrupamento entre três aplicações concorrentes em face ao sistema proposto.

No **Capítulo 5**, são apresentadas as conclusões obtidas a partir deste estudo, as principais contribuições e trabalhos futuros que podem ser derivados do exposto ao longo desta dissertação.

Finalmente, no **Apêndice A**, são expostas amostras de imagens utilizadas nos experimentos realizados. No **Apêndice B**, descrevem-se os testes estatísticos adotados nesta dissertação. O **Apêndice C** contém os resultados numéricos obtidos das abordagens investigadas, da mesma forma que o **Apêndice D**, contém os resultados numéricos da aplicação final desenvolvida e das três ferramentas concorrentes utilizadas no experimento de comparação. Por fim, no **Apêndice E**, ilustram-se os resultados dos agrupamentos obtidos a partir do uso da abordagem proposta.

Capítulo 2

Trabalhos Relacionados

Neste capítulo, apresenta-se um levantamento e análise de trabalhos relacionados com o tema desta pesquisa, focando em trabalhos nos quais são investigados ou propostos métodos que auxiliaram no amadurecimento de uma nova solução para o problema em estudo – agrupamento de faces em coleções de fotografias digitais. Além disto, detalham-se também a metodologia de pesquisa bibliográfica realizada e dois principais ramos de aplicação da área de reconhecimento de faces.

2.1. Metodologia de Pesquisa Bibliográfica

A sistemática de pesquisa bibliográfica empregada na execução deste trabalho objetivou determinar o “estado da arte” sobre o problema em questão, de forma a identificar o que foi publicado sobre o assunto, que aspectos já foram abordados e quais as lacunas existentes na literatura, de maneira a possibilitar a delimitação do problema a ser estudado, além de prover uma estruturação conceitual que dará sustentação ao desenvolvimento da pesquisa.

A revisão de literatura foi resultado de um levantamento bibliográfico no qual foi utilizada uma das ferramentas de busca mais populares e abrangentes da *Internet*, *Google*¹⁸, de acordo com as seguintes palavras-chave: (i) *face clustering*; (ii) *face matching*; (iii) *face tagging*; (iv) *photo tagging*; (v) *face clustering survey*; (vi) *image clustering survey*; e (vii) *clustering algorithms*, descartando as ocorrências com aplicações em vídeo.

Como resultado da pesquisa, foram selecionados vinte artigos altamente relacionados com o trabalho em questão, publicados em

¹⁸ Google, disponível em: <http://www.google.com>

periódicos, conferências e *workshops* mais relevantes na área de Visão Computacional que serão detalhados nas seções posteriores.

2.2. Reconhecimento Supervisionado e Não Supervisionado

O reconhecimento de padrões objetiva a classificação de objetos (padrões) em um número de categorias ou classes (THEODORIDIS e KOUTROUMBAS, 1999). No caso do reconhecimento de faces, as imagens de faces são os objetos e as classes são seus nomes ou identificações. Assim, dado um padrão, seu reconhecimento/classificação pode ser categorizado como: *supervisionado*, em que o padrão de entrada é identificado como um membro de uma classe pré-definida pelos padrões de treinamento, que são rotulados com suas classes; e *não supervisionado* ou *clustering*, em que o padrão é associado (agrupado) a uma classe que é aprendida com base na similaridade entre os padrões de treinamento, ou seja, o próprio sistema toma a decisão de criar novas classes ou agrupar classes preexistentes.

Dentre as aplicações baseadas em reconhecimento supervisionado, destaca-se a identificação para controle de acesso, segurança e vigilância (ZHAO et al., 2003). No tocante a aplicações baseadas em reconhecimento não supervisionado, destacam-se: agrupamento de objetos, mineração de dados e diagnósticos médicos, dentre outros (JAIN e DUBES, 1988). As abordagens e técnicas, referentes aos vinte artigos selecionados, discutidas nas subseções seguintes, foram agrupadas de acordo com os seguintes critérios: (i) propósito do sistema; (ii) características extraídas para identificação pessoal; (iii) métricas de similaridade; (iv) técnicas de agrupamento/reconhecimento; e (v) métricas de avaliação de agrupamento.

2.3. Propósito do Sistema

O propósito-chave dos sistemas apresentados nos trabalhos selecionados foi classificado em duas categorias: *agrupamento de faces* e *reconhecimento de faces*, sendo dezessete trabalhos pertencentes à primeira categoria e três à segunda. Considerando que o problema-alvo desta dissertação é o agrupamento de faces, foi dada maior importância à busca de trabalhos

relacionados especificamente a este tema. Um sumário da revisão pode ser encontrado na Tabela 2.1, a qual é apresentada e discutida na Seção 2.9. As principais características faciais identificadas na literatura pesquisada serão detalhadas na seção seguinte.

2.4. Características Extraídas para Identificação Pessoal

A face humana possui inúmeras características (tais como os olhos, sobrancelhas, boca, nariz e cabelo) que são utilizadas, de maneira intuitiva, para o reconhecimento de pessoas. Para descrever ou caracterizar uma pessoa, pode-se analisar o posicionamento, a forma e o tamanho destas características, além da distância que separa cada uma delas.

Sistemas de reconhecimento de faces extraem informações para relacionar um determinado indivíduo à imagem que o identifica. Na Figura 2.1, é possível observar alguns pontos de interesse de uma face. Abate et al. (2007) afirmaram que as características mais exploradas são a boca, o nariz, os olhos e as sobrancelhas, as quais têm sido utilizadas com sucesso para o reconhecimento de faces frontais. Contudo, em algumas situações, as imagens não apresentam todas as informações necessárias, e.g., em imagens de perfil.

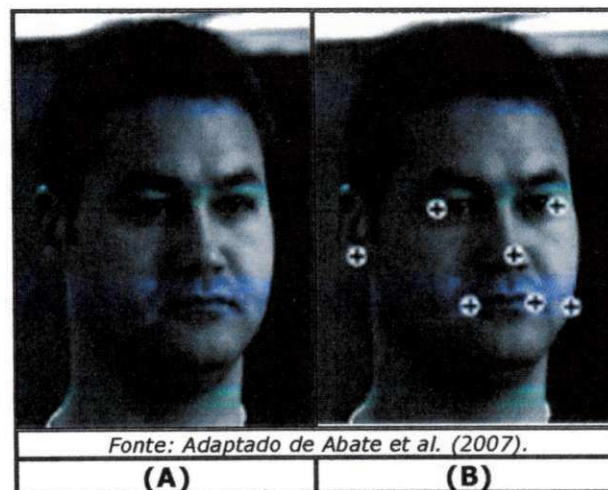


Figura 2.1 – Alguns pontos de interesse para o reconhecimento de faces: (A) Imagem de entrada (B) Pontos de interesse.

Os problemas listados a seguir representam desafios e servem de motivação para as pesquisas na área:

- a) **Pose** – As imagens apresentam faces em diversas poses, com diversos graus de variação, como frontal, 1/4 de perfil, semi-perfil e perfil, o que pode ocluir, total ou parcialmente, algumas características, tais como olhos e orelhas;
- b) **Oclusão** – Algumas imagens apresentam a oclusão parcial ou total de características da face por objetos da cena (óculos, chapéu, cabelo, etc.);
- c) **Iluminação** – Variações de iluminação podem ser consideradas uma forma de oclusão natural, em que regiões da face podem ficar muito escuras ou claras demais, dificultando a identificação;
- d) **Expressão facial** – A expressão facial (dúvida, raiva, surpresa, etc.) pode alterar as características e forma da face; e
- e) **Condições da imagem** – As condições de imageamento podem influenciar no reconhecimento. Isso se deve à possibilidade de haver ruídos e também devido à resolução de aquisição.



Figura 2.2 – Problemas em imagens de faces no tocante a variações de: (A) Pose; (B) Oclusão; (C) Iluminação; (D) Expressão facial; e (E) Condições de aquisição da imagem.

Esses problemas influenciam na tomada de decisão sobre o método de reconhecimento de faces do sistema, além da escolha das informações que serão analisadas, conforme ilustrado na Figura 2.2.

Representações faciais existentes podem ser divididas em duas categorias: *global* e *local*. Na representação facial global, cada componente do vetor de características contém as informações incorporadas em toda a imagem de face, o que corresponde a uma característica holística da face. Em contrapartida, para a representação facial local, cada componente do vetor de características representa apenas uma determinada região na face, ou seja, só codifica os traços detalhados naquela área específica. Embora muitos métodos de representação facial com base em características globais

ou locais tenham sido propostos com sucesso, continua sendo um problema em aberto determinar a representação mais adequada para o agrupamento de faces (SU et al., 2009).

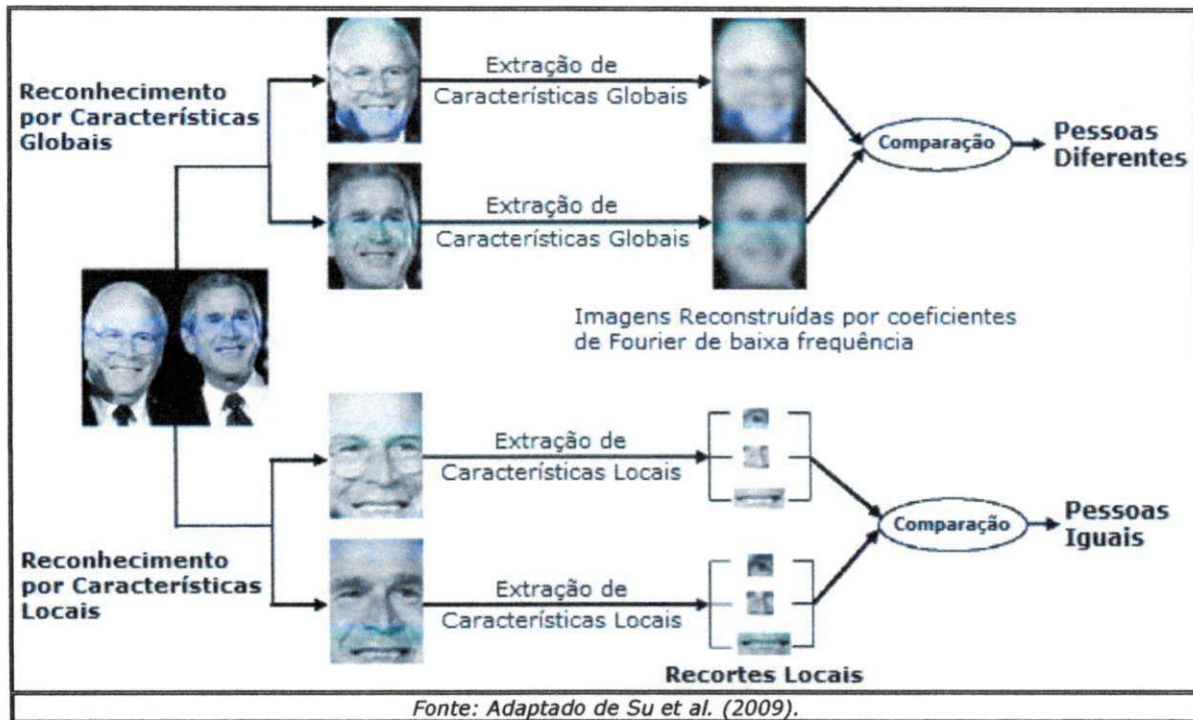


Figura 2.3 – Ilustração dos diferentes papéis das características globais e locais no reconhecimento de faces.

Na Figura 2.3, as duas faces de entrada são artificiais, ou seja, os principais componentes (olhos, nariz e boca) são na verdade de uma mesma pessoa. Todavia, as faces aparentam ser muito desiguais em termos da configuração estrutural global, por apresentarem diferentes contornos faciais e penteados. Assim, torna-se evidente que as características globais e locais desempenham papéis diferentes na percepção da face, sendo, portanto, necessário, combiná-las de maneira eficiente. Intuitivamente, a informação local é incorporada às variações detalhadas da aparência facial, enquanto a informação global representa a configuração estrutural holística dos órgãos da face, e.g., o contorno facial.

Enquanto, inicialmente, as representações faciais globais eram populares para o reconhecimento de faces, recentemente, cada vez mais tentativas estão sendo feitas para desenvolver sistemas de reconhecimento de faces baseados em características locais, que se acredita serem mais robustos às variações de expressão facial, iluminação e oclusão (SU et al., 2009). Assim, é natural esperar um melhor desempenho destes sistemas

através da combinação de informações globais e locais, ao considerarmos que ambos os tipos de características são cruciais para a percepção da face humana.

Zhang et al. (2011) adotaram, em seu trabalho, um método de agrupamento baseado em formas da face. A forma é normalmente utilizada para distinguir pessoas pela distância e dimensão de suas características faciais. No referido trabalho, as características faciais, incluindo os contornos da face, olhos, nariz e boca são extraídos de imagens frontais para classificação. Há um total de 68 pontos fiduciais extraídos por meio de modelos ativos de forma, denominados *Active Shape Models* – ASM (COOTES et al., 1995) para a composição da representação facial, conforme pode ser observado na Figura 2.4.



Figura 2.4 – Pontos fiduciais utilizados para composição da representação facial.

Zhu, Wen e Sun (2011) e Antonopoulos, Nikolaidis e Pitas (2007) fizeram uso do descritor SIFT (*Scale-invariant feature transform*), originalmente proposto por Lowe (1999), para a tarefa de agrupamento de faces. Trata-se de um método para a extração de características invariantes, muito utilizado para verificar a correspondência entre diferentes perspectivas de um objeto ou de uma cena. SIFT avalia características denominadas pontos de interesse (*keypoints*), as quais são invariantes a níveis significativos de distorção da forma, escala, perspectiva e mudanças de iluminação. A cada ponto de interesse é atribuído um vetor de 128 elementos, que expressa a orientação, escala e localização de uma região de pixels em torno do ponto de interesse.

Jayech e Mahjoub (2011) propuseram o uso de regiões de interesse

(*Region of interest* – ROI) para a extração de características locais, visando extrair informações espaciais das faces. A abordagem é baseada na modelagem da imagem em uma grade 3x3, refletindo uma descrição local da imagem. Assim, a imagem a ser processada é dividida em nove blocos, conforme pode ser observado na Figura 2.5. Em cada bloco, são extraídos descritores para compor o vetor de características, calculados a partir dos valores dos pixels, a saber: (i) média; (ii) desvio padrão; (iii) energia; (iv) entropia; (v) contraste; e (vi) homogeneidade.

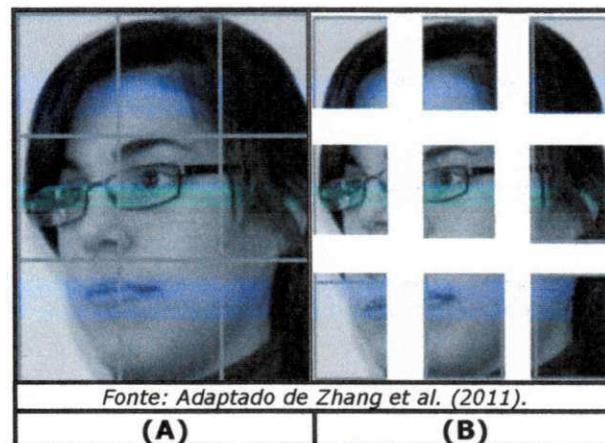


Figura 2.5 – Abordagem utilizada por Zhang et al. (2011): (A) Grade 3x3 e (B) Regiões de interesse.

Zhang, Zhang e Tretter (2010) adotaram em seu trabalho um conjunto de filtros, originalmente proposto por Winn, Criminisi e Minka (2005), composto por três gaussianas, quatro derivadas de primeira ordem de Gaussianas e quatro filtros Laplaciano de Gaussianas (*Laplacian of the Gaussian* – LoG) para a extração de informações de cor e textura das imagens de faces, no contexto de agrupamento de faces. A fim de reduzir a influência de ruídos nas características e capturar os padrões visuais mais relevantes, as características extraídas são agrupadas com uma modelagem de mistura de Gaussianas (Gaussian Mixture Modeling – GMM), treinada por meio do algoritmo *Expectation-Maximization* (EM).

Prince e Elder (2010) propuseram um novo método para o reconhecimento de faces robusto a variações de pose, baseado em uma versão probabilística de *Fisherfaces* (BELHUMEUR, HESPANHA e KRIEGMAN, 1997), denominado PLDA (*Probabilistic Linear Discriminant Analysis*). Um modelo composto de sinal e ruído foi desenvolvido para explicar a variação nos dados da face, em que a principal variação representava uma identidade

facial. Ao invés de basear o reconhecimento em comparações de distância, a comparação foi calculada como sendo a probabilidade de componentes de identidade subjacentes serem os mesmos, independentemente do valor da pose. Baseados em experimentos realizados em um trabalho anterior (PRINCE e ELDER, 2007), os autores afirmaram que este método produziu bons resultados, no que diz respeito à taxa de erro de reconhecimento, tanto em reconhecimento de faces frontais quanto em reconhecimento de faces com variações de pose, com taxa de erro de 0,3% se comparado com algoritmos contemporâneos, e.g., Análise de Componentes Principais (*Principal Component Analysis – PCA*) (JOLLIFFE, 2002) com taxa de erro 33,9% e Análise Discriminante Linear (*Linear Discriminant Analysis – LDA*) (MCLACHLAN, 2004) com taxa de erro de 11,9% na base de imagens XM2VTS (MESSER et al., 1999).

Kapoor et al. (2009) empregaram em seu trabalho algoritmos de extração de características faciais, propostos por Wright e Hua (2009), no âmbito de agrupamento de faces. A face foi dividida em regiões de 64x64 pixels, em tons de cinza com média zero e variância unitária e um descritor local foi extraído de cada região. A localização de cada uma dessas regiões foi incorporada ao descritor correspondente para formar um descritor de aparência espacial conjunta. Cada um destes descritores foi então quantizado por um conjunto de árvores de projeção randomizadas (*RPTrees*). A representação facial final consistia de um histograma esparsa formado pela coleção de índices quantizados dos descritores de aparência espacial. Wright e Hua (2009) argumentaram que seu método era invariante à pose e que apresentava melhores taxas de reconhecimento (96,5%) do que o PCA (88,1%), LDA (93,9%) e LPP (*Locality Preserving Projection*) (HE e NIYOGI, 2003), na base de imagens ORL (SAMARIA e HARTER, 1994).

Palit et al. (2009) utilizaram uma abordagem, originalmente proposta por Perlmutter et al. (2006) no contexto de agrupamento de faces. O sistema fundamenta-se na detecção e localização de dezenove pontos fiduciais (também conhecidos como *landmarks*), correspondentes às características faciais. Para localizar tais características, foi utilizado um modelo que combinava uma representação generativa das posições das

mesmas com uma representação discriminativa de sua aparência. Os pontos fiduciais serviram como centro de extração de regiões faciais de tamanho 11x11 pixels, as quais compuseram a representação facial do modelo, conforme pode ser observado na Figura 2.6.

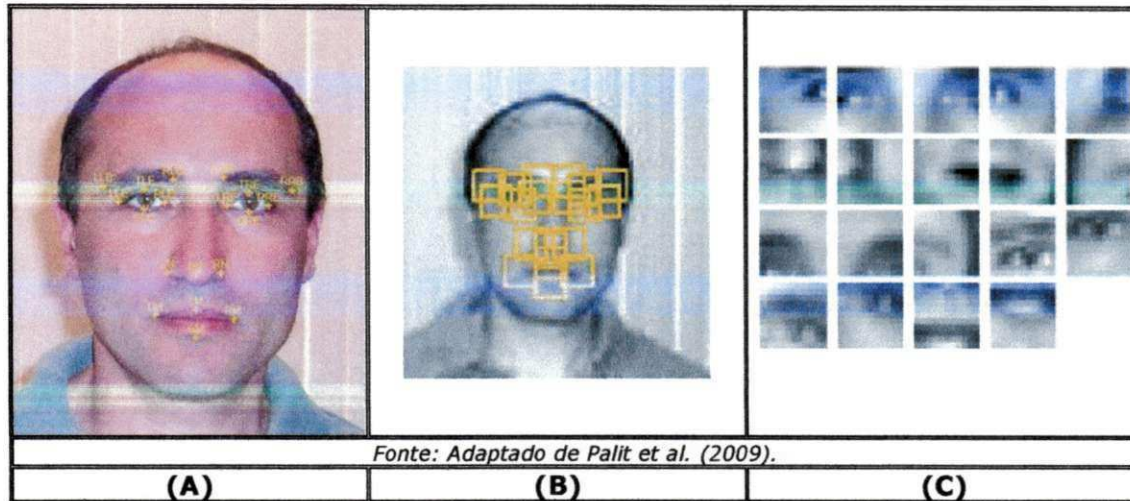


Figura 2.6 – Abordagem utilizada por Palit et al. (2009): (A) Pontos fiduciais; (B) Regiões de interesse; e (C) Regiões extraídas.

Pinto, Dicarlo e Cox (2009) empregaram em seu trabalho a característica do tipo V1 (PINTO, COX e DICARLO, 2008) para a tarefa de reconhecimento. Tal característica consistia de uma representação inspirada nas propriedades conhecidas da área cortical V1. Os modelos V1, compostos por *Wavelets* de Gabor, abrangiram uma série de orientações e frequências espaciais, representando apenas descritores de primeira ordem do córtex visual primário e não contendo qualquer representação particularmente sofisticada de forma, nem possuindo qualquer mecanismo projetado para tolerar variações na imagem, e.g., variação de iluminação e pose. Entretanto, os autores apresentaram evidências de que essa abordagem simples é capaz de obter um bom desempenho de reconhecimento: 79,35% de taxa de reconhecimento, nível de estado-da-arte, sobre a base de faces *Labeled Faces in the Wild* – LFW (HUANG et al., 2007).

Hua e Akbarzadeh (2009) destacaram que uma simples diferença entre Gaussianas (*Difference of Gaussians* – DoG) supera outros métodos mais utilizados em situações com variações de iluminação para aplicações de reconhecimento de faces. No que diz respeito à taxa de erro de reconhecimento, o método apresentou uma taxa de erro de 1,6% na base

de imagens ORL, a qual foi inferior ao resultado de vários outros métodos na mesma base de imagens, como LDA, com taxa 7,2%, *Laplacian-Face* (HE et al., 2005), com taxa 6,8%, *Spatially Smooth Fisherface* – SLDA (CAI, HE e HAN, 2007), com taxa 2,3% e *Regularized Fisherface* – RLDA (CAI et al., 2007), com taxa 3,6%. O método apresentou uma taxa de reconhecimento de 78,64% na base de imagens LFW (HUANG et al., 2007).

Guillaumin, Verbeek e Schmid (2009) experimentaram os vários conjuntos de características de faces utilizados no trabalho de Wolf, Hassner e Taigman (2008) na tarefa de agrupamento de faces, com a adição do descritor SIFT, computados em pontos fixos na face (cantos da boca, olhos e nariz) pelo uso de um detector de características faciais. Em seu trabalho, os autores comprovaram que as características baseadas em SIFT produziram melhores resultados (acréscimo de cerca de 1% na taxa de reconhecimento) do que os descritores relatados em Wolf, Hassner e Taigman (2008).

Analogamente, Wolf, Hassner e Taigman (2009) fizeram uso dos descritores reportados em seu trabalho anterior (WOLF, HASSNER e TAIGMAN, 2008), com a adição do descritor SIFT, para o problema de reconhecimento de faces. Foram comparados os resultados com os descritores: (i) LBP (OJALA, PIETIKAINEN e HARWOOD, 1996); (ii) *Three-patch* LBP (TPLBP) (WOLF, HASSNER e TAIGMAN, 2008); (iii) *Four-patch* LBP (FPLBP) (WOLF, HASSNER e TAIGMAN, 2008); (iv) Gabor C1 (RIESENHUBER e POGGIO, 1999); e (v) SIFT (LOWE, 1999). Todos os descritores foram concatenados em um único vetor de características. Em comparação com as variantes do LBP, TPLBP e FPLBP em termos de taxa de reconhecimento, o descritor SIFT é menos sensível ao desalinhamento das faces, no entanto, é facilmente perturbado por bordas causadas por variações de iluminação. Individualmente o SIFT apresentou melhor taxa de reconhecimento 75,76% que os demais descritores, LBP (73,43%), TPLBP (71,63%) e FPLBP (71,75%). Entretanto, uma combinação global gerou um resultado superior ao do SIFT isoladamente, 83,98%, na base de dados LFW (HUANG et al., 2007).

Kyperountas, Tefas e Pitas (2008) demonstraram que o LDA pode ser considerado uma poderosa ferramenta para a redução de dados e a extração de características em abordagens baseadas em aparência para a tarefa de agrupamento de faces. Os autores argumentaram que, quando se trata de resolver problemas de classificação de padrões, algoritmos baseados em LDA superam os baseados em PCA, pelo fato de que o primeiro otimiza representações de baixa dimensionalidade de objetos com foco na extração de características mais discriminantes, enquanto o segundo visa apenas à reconstrução do objeto.

No trabalho desenvolvido por Tian et al. (2007), as imagens de faces identificadas por um detector de faces baseado em características de *Haar* (VIOLA e JONES, 2001) foram alinhadas e normalizadas por meio de um detector de olhos (MA et al., 2004). Então, características LBP (OJALA, PIETIKAINEN e HARWOOD, 1996) foram extraídas para a tarefa de agrupamento de faces.

Chacon, Rivas e Ramírez et al. (2006) utilizaram como característica, linhas de características faciais, denominadas *Face Feature Lines* (FFL), conforme pode ser observado na Figura 2.7. FFL são linhas proeminentes e podem ser extraídas a partir da transformada de *Hough* sobre imagens de faces de baixa resolução. Há destaque para o fato das FFL serem consideradas características importantes em estudos recentes de reconhecimento de faces. Experimentos realizados demonstraram que o coeficiente de correlação da combinação KLT+FFL (0,01099) era superior aos resultados individuais do KLT (0,01102) e PCA (0,09548), respectivamente.

Perronnin e Dugelay (2005) aplicaram as mesmas características adotadas em um trabalho anterior (PERRONNIN, DUGELAY e ROSE, 2003). Nele foram empregadas *Wavelets* de Gabor para cada imagem de entrada. Nos experimentos realizados, foram utilizadas imagens com grande variabilidade em termos de expressões faciais e as imagens foram pré-processadas para a extração de regiões faciais de tamanho 128x128 pixels, que foram normalizadas pelas coordenadas dos olhos e da ponta do nariz.

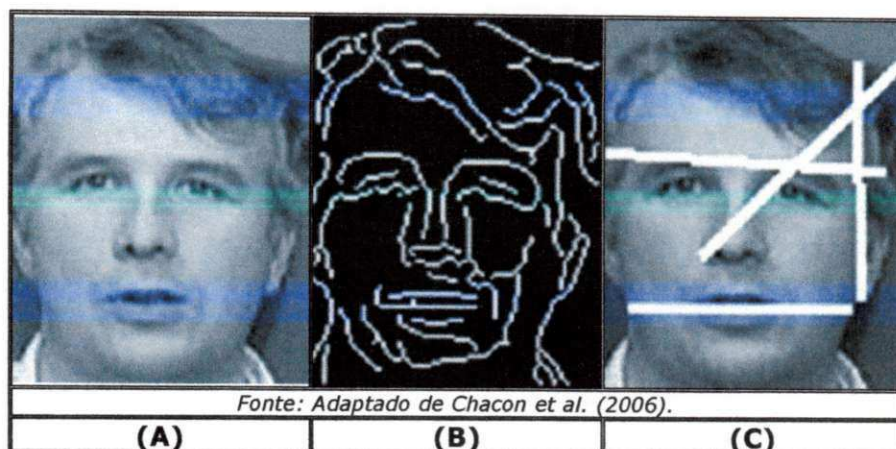


Figura 2.7 – Transformada de Hough de uma face: (A) Imagem de entrada; (B) Bordas da face; e (C) Imagem de entrada com quatro FFL extraídas.

Li e Tang (2004) empregaram em seu trabalho um método bayesiano baseado em subespaço probabilístico (MOGHADDAM, JEBARA e PENTLAND, 2000). Para melhor avaliar o desempenho do sistema, as imagens foram pré-processadas de acordo com as seguintes etapas: (i) alinhamento e orientação das faces; (ii) redimensionamento das faces, de modo que as distâncias entre os dois olhos fossem as mesmas para todas as imagens; (iii) recorte das faces para a remoção do fundo e da região do cabelo; e (iv) equalização do histograma para a normalização fotométrica.



Figura 2.8 – Blocos de características extraídas.

Analogamente aos trabalhos de Jayech e Mahjoub (2011) e Tian et al. (2007), Chen et al. (2003) fizeram o uso de regiões para a extração de características locais. A área de face foi representada em blocos de altura e largura 4x2 e as características de cor e textura foram extraídas desses blocos, conforme pode ser observado na Figura 2.8. Como característica de cor foi utilizado o auto-correlograma (HUANG et al., 1997), considerado

muito eficaz na combinação de características de cor e textura. Para extrair a informação de textura baseada em *Wavelets*, a imagem original foi decomposta por uma transformada *Wavelet* de três níveis.

Zhang et al. (2003) também empregaram auto-correlogramas, além de momentos cor-textura (YU et al., 2002). Cada região da face foi estendida e dividida em dois blocos que incluía a face e o corpo, respectivamente, e, em seguida, foram extraídos desses dois blocos dois vetores de características de cor e textura. Experimentos demonstraram que a combinação de informações do corpo e da face proporcionou a obtenção de melhores resultados que ambos individualmente.

Diante do exposto, as características exploradas nesta seção, consideradas robustas a um ou mais dos problemas descritos, foram selecionadas como candidatas para a composição da representação facial do sistema desenvolvido nesta dissertação. As principais métricas de similaridade identificadas na literatura pesquisada serão detalhadas na próxima seção.

2.5. Métricas de Similaridade

A métrica de similaridade é uma medida que avalia numericamente o grau de semelhança entre dois objetos (WANGENHEIM e WANGENHEIM, 2003), mais especificamente, faces, no âmbito desta dissertação. Desta forma, a abordagem mais direta para verificar a semelhança entre pares de descritores de imagens é a de considerar a distância entre os vetores de características que codificam a aparência das duas imagens. Considerando duas imagens de face I_1 e I_2 , que são codificadas utilizando um descritor g como sendo $g(I_1)$ e $g(I_2)$, o par de imagens é considerado similar se $d(g(I_1), g(I_2)) < T$, em que d é uma medida de distância e T um limiar. Deste modo, as métricas detalhadas a seguir podem ser usadas como a medida de distância d .

Zhang et al. (2011) adotaram em seu trabalho a distância de *Hausdorff*. A distância de Hausdorff é mais adequada para calcular a distância entre dois conjuntos de pontos, sendo aplicada para medir a

similaridade entre dois vetores de características faciais. Conceitualmente, a distância *Hausdorff* mede o quão distante dois subconjuntos do espaço de medida estão um do outro.

Zhu, Wen e Sun (2011) e Kapoor et al. (2009) empregaram como métrica de similaridade à distância L1 que, por sua vez, é invariante à translação e à rotação, segundo os autores. No trabalho de Wright e Hua (2009), outra medida foi utilizada, a métrica frequência de documento inversa (*Inverse document frequency* – IDF), uma variante ponderada da L1 que visa medir a similaridade entre dois histogramas esparsos. Apesar de existirem outras métricas de distância para comparar histogramas, os resultados de reconhecimento mostraram melhor desempenho com o uso da medida IDF (89,4%) em comparação com a L1 (89,3%) e a L2 (86,3%).

Jayech e Mahjoub (2011) fizeram uso da distância Tangente, uma ferramenta matemática que permite comparar duas imagens, tendo em conta pequenas transformações (e.g., rotações, translações). Introduzida no início dos anos 90, por Simard et al. (1998), a distância Tangente foi combinada, posteriormente, com diferentes classificadores, para reconhecimento de caracteres, detecção e reconhecimento de faces e reconhecimento de fala. A distância entre um par de imagens (I_1 , I_2) é calculada como sendo a distância entre os espaços de parâmetros via I_1 e I_2 , respectivamente. Estes espaços localmente modelam todas as formas geradas pelas transformações possíveis entre as duas imagens.

Zhang, Zhang e Tretter (2010) utilizaram em seu trabalho a *Earth Mover's Distance* – EMD (RUBNER, TOMASI e GUIBAS, 2000), para medir a distância entre assinaturas de duas regiões contextuais. A distância EMD é uma medida para avaliar a dissimilaridade entre dois espaços de características com base no custo mínimo para transformar uma característica na outra.

Palit et al. (2009) elencaram quatro métricas de distância para verificar a similaridade entre os vetores de face: (i) a distância Euclidiana; (ii) a distância média; (iii) a distância da média ponderada em torno da mediana e (iv) a distância média ponderada. Os autores citaram que, em

experimentos realizados, foi verificado que a distância da média ponderada em torno da mediana produziu os melhores resultados do que as demais, no entanto, não foram dados detalhes de tais experimentos.

Pinto, Dicarlo e Cox (2009) adotaram uma medida de *pair-matching* entre cada um dos seis algoritmos variantes da representação de pixels com cada um dos oitos algoritmos variantes da representação do tipo V1. Assim, para cada par de treinamento, o vetor de características resultante foi rotulado como "mesmo" ou "diferente", e a tarefa de rotulagem (teste) de novos exemplos foi tratada como um problema de classificação com duas categorias.

Hua e Akbarzadeh (2009) propuseram uma métrica denominada *Robust Elastic and Partial Matching* para o reconhecimento de faces. A métrica segue o princípio fundamental da distância de *Hausdorff*, sendo definida no espaço de características, ou seja, no espaço dos descritores locais de imagens. Além disto, restrições garantem que cada descritor local seja combinado apenas com seus vizinhos no espaço da imagem. Esta restrição espacial é essencial para a comparação de duas imagens, conforme detalhado nos experimentos.

Guillaumin, Verbeek e Schmid (2009) elaboraram uma métrica de distância, *Logistic Discriminant based Metric Learning* (LDML), baseada no aprendizado em métricas de *Mahalanobis* sobre uma dada representação espacial. O método usa discriminante logístico para aprender uma métrica a partir de um conjunto de pares de imagens rotuladas. Seu objetivo é encontrar uma métrica tal que a distância entre pares positivos seja menor que as distâncias correspondentes aos pares negativos e obter uma estimativa probabilística para identificar se as duas imagens retratam o mesmo objeto.

Wolf, Hassner e Taigman (2009) utilizaram o sistema *One-Shot Similarity* (OSS). O valor da OSS é calculado pelo treinamento de um modelo discriminativo exclusivo para os dois sinais que estão sendo comparados, usando um conjunto de amostras de fundo. Assim, o sistema *Two-Shot Similarity* (TSS) é então definido como uma medida de quão bem

esse modelo discrimina os dois conjuntos. OSS e TSS são consideradas meta-medidas de similaridade que podem ser ajustadas para operar com a maioria dos algoritmos de aprendizagem discriminativa.

Wolf, Hassner e Taigman (2008) fizeram uso da distância Euclidiana e distância Euclidiana quadrada. A motivação para a segunda distância é que os vetores do descritor consistiam principalmente de histogramas e que aplicar a raiz quadrada previamente aos cálculos de distância corresponde à distância *Hellinger* entre probabilidades. *Support Vector Machines – SVM* linear binário (VAPNIK, 1999) é considerado pelos autores como um método que oferece desempenho semelhante e pode ser generalizado para mais de uma medida de distância por par. Uma maneira simples de combinar múltiplos descritores de imagens e múltiplas distâncias é criar um vetor de distâncias e executar um SVM linear sobre este vetor.

Kyperountas, Teas e Pitas (2008) argumentaram que o agrupamento com a distância Euclidiana quadrada é mais rápido do que agrupamento com distância Euclidiana regular e que o resultado de alguns métodos de agrupamento, e.g., *Jarvis-Patrick*, *K-Means*, não é alterado se a distância Euclidiana quadrada é substituída pela regular. No entanto, a saída do agrupamento hierárquico é susceptível a alteração.

Antonopoulos, Nikolaidis e Pitas (2007) empregaram, em seu trabalho, o seguinte procedimento para avaliar a similaridade entre pares de imagens de faces. Primeiramente, pontos de interesse SIFT, juntamente com seus correspondentes vetores de características, foram extraídos de pares de imagens. Em seguida, pontos de interesse foram comparados, a partir da distância Euclidiana entre seus vetores de características. O resultado foi a quantidade de correspondências (C) entre *keypoints* para um par de imagens de faces. Como a comparação de *keypoints* do par P_{12} era diferente do resultado para o par P_{21} , o valor máximo de correspondências, $\max(C(P_{12}), C(P_{21}))$, foi considerado.

Tian et al. (2007) adotaram a distância Qui-quadrado, na qual cada quadrado da distância Euclidiana é ponderado pelo inverso da frequência correspondente a cada termo. A divisão de cada termo quadrado pela

freqüência normaliza a maior variação nas freqüências altas e a menor variação nas baixas freqüências. Se não houver normalização, as diferenças entre grandes proporções tendem a ser maior e, assim, dominar o cálculo da distância, enquanto as diferenças entre pequenas proporções tendem a ser dominadas.

Chacon, Rivas e Ramírez (2006) empregaram em seu trabalho a distância D , proposta no trabalho de Abonyi e Szeifert (2003). Esta medida de distância consiste em dois termos. O primeiro termo é baseado no algoritmo Gath-Gheva – GG (ABONYI e SZEIFERT, 2003) para agrupamento não-supervisionado. O segundo termo permite o uso de rótulos de classe e é definido como uma probabilidade *a posteriori*.

Perronnin e Dugelay (2005) realizaram um conjunto de experimentos para determinar qual a distância mais adequada para medir a similaridade entre vetores de características. É sabido que a distância entre os vetores de características deve ser baseada em métricas relativamente simples. Foram testadas as medidas L1, L2 (Euclidiana) e cosseno. Os experimentos comprovaram que a medida do cosseno (90%) apresentou melhores resultados de reconhecimento do que as demais, L1 (63%) e L2(69%), na base de imagens FERET (PHILLIPS et al., 2000).

Li e Tang (2004) decomuseram a distância entre uma determinada imagem e uma classe de faces em duas componentes ortogonais à distância no espaço de características – *Distance-in-feature-space* (DIFS) (correspondente à projeção sobre o menor espaço de dimensão) e a distância do espaço de características estimada a partir da reconstrução de erro – *Distance-from-feature-space* (DFFS). Chen et al. (2003) adotaram a taxa H-Hit para avaliar o desempenho da anotação de faces. A taxa de H-Hit dá uma estimativa de previsão do desempenho para um álbum de fotografias, de modo que o H-Hit é significativo somente para números pequenos, como 1, 2, 5, 10.

De maneira similar, Zhang et al. (2003) utilizaram a taxa H-Hit para avaliar o desempenho de um algoritmo proposto para anotação de faces. Suponha-se que as faces de álbum foram anotadas seqüencialmente, em

uma ordem aleatória. Dada uma face de entrada não anotada, o sistema gera uma lista H de nomes candidatos com base no histórico de faces anotadas. Se o verdadeiro nome dessa face estiver na lista considera-se que a face foi encontrada (*hit*) pela lista de nomes. Assim, foi possível calcular a taxa média de sucesso de todas as faces do álbum, dado o comprimento da lista H de nomes candidatos. Na próxima seção, serão comentadas as técnicas de agrupamento e reconhecimento mais relevantes da literatura revisada.

2.6. Técnicas de Agrupamento/Reconhecimento

Diariamente, o ser humano se depara com uma grande quantidade de informações e as armazenam ou representam como dados, para posterior análise e gerenciamento. Um dos meios principais para lidar com esses dados é classificá-los ou agrupá-los em um conjunto de categorias ou grupos (XU e WUNSCH, 2005). A fim de aprender sobre um novo objeto ou compreender um novo fenômeno, as pessoas sempre tentam identificar as características que podem descrevê-lo ou ainda, compará-lo com outros objetos ou fenômenos conhecidos, com base na similaridade ou dissimilaridade, generalizada como proximidade, de acordo com certas normas ou regras.

Neste contexto, técnicas de agrupamento particionam os dados em certo número de grupos (subconjuntos ou categorias), considerando a homogeneidade interna e externa de cada partição, ou seja, padrões em um mesmo grupo devem ser semelhantes entre si, enquanto padrões em grupos diferentes devem ser diferentes entre si. A seguir, as estratégias de agrupamento/reconhecimento de faces identificadas na revisão da literatura serão apresentadas e discutidas.

Zhang et al. (2011) elaboraram um método de agrupamento modificado baseado no algoritmo *Iterative Self-Organizing Data Analysis Technique* – ISODATA (JAIN e DUBES, 1988) para categorizar faces em classes diferentes, utilizando características de formas faciais. Inicialmente, as imagens de faces foram agrupadas em sete classes, de acordo com o contorno da face. Em seguida, para cada classe de contorno de face,

características faciais de forma incluindo os olhos, nariz, boca e suas localizações relativas foram usadas para agrupar faces de diferentes pessoas. Apesar dos autores considerarem sua abordagem promissora, tal afirmação pode ser criticada, devido à fraca avaliação experimental realizada, dado que nenhuma métrica objetiva de avaliação foi utilizada.

Zhu, Wen e Sun (2011) elaboraram um novo algoritmo de agrupamento, denominado *Rank-Order*, cujo núcleo é baseado em uma comparação entre pares de faces, utilizando informações de seus respectivos vizinhos, no conjunto. Tal estratégia é motivada por uma observação de que faces de uma mesma pessoa normalmente compartilham as mesmas faces vizinhas. Especificamente, para cada face, uma lista de ordem de classificação é gerada, classificando todas as outras faces do conjunto de dados pela distância absoluta da face em relação às demais, determinando, assim, seus vizinhos. Experimentos realizados indicaram taxas de precisão (*precision*) de 97% e revocação (*recall*) de 86% para a base de imagens *Gallagher* (GALLAGHER e CHEN, 2008).

Jayech e Mahjoub (2011) propuseram uma nova abordagem combinando a distância Tangente, o algoritmo *K-Means* e redes bayesianas. Primeiro, a técnica de distância Tangente foi usada para calcular espaços tangentes que representassem a mesma imagem, com o objetivo de reduzir o erro na fase de classificação. Segundo, a imagem de face foi dividida em um conjunto de blocos. Para cada bloco, foi calculado um vetor de descritores. Então, o algoritmo *K-Means* foi usado para agrupar características de baixo nível, incluindo informações de cor e textura para construir uma representação facial. Finalmente, classificadores de redes bayesianas foram responsáveis pelo agrupamento das imagens de faces.

Zhang, Zhang e Tretter (2010), Antonopoulos, Nikolaidis e Pitas (2007) e Li e Tang (2004) empregaram em seus trabalhos a técnica de agrupamento aglomerativo hierárquico (*Hierarchical Agglomerative Clustering* – HAC), a partir da qual os grupos são construídos pela combinação de grupos existentes com base em sua proximidade. O processo básico do HAC pode ser resumido pelas seguintes etapas: (i) inicializar um

conjunto de grupos; (ii) encontrar o par de grupos mais próximo que têm a maior métrica de similaridade e combiná-los em um novo cluster; e (iii) estimar a medida de semelhança entre o novo grupo e todos os demais grupos, resultando em um dendrograma, conforme ilustrado na Figura 2.9. Experimentos realizados no trabalho de Zhang, Zhang e Tretter (2010) evidenciaram taxas médias de precisão (*precision*) de 99% e revocação (*recall*) de 80%, superiores àquelas obtidas pelas ferramentas Picasa (PICASA, 2010), 100%-15% (P-R) e EasyAlbum (CUI et al., 2007), 49%-9% (Precisão-Revocação) em na base de imagens CMU (GALLAGHER e CHEN, 2009).

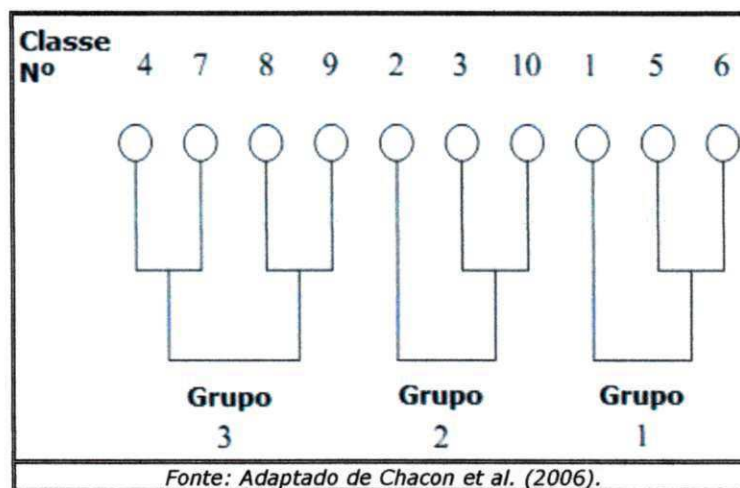


Figura 2.9 – Exemplo de um dendrograma.

Prince e Elder (2010) desenvolveram uma abordagem bayesiana em que na fase de treinamento um modelo generativo probabilístico foi aprendido a partir das faces. A identidade individual foi representada como uma variável latente, cujo resultado é idêntico quando faces candidatas são semelhantes. Esse modelo foi usado para calcular a similaridade do conjunto inteiro para cada agrupamento hipotético usando um processo equivalente ao modelo de seleção bayesiano, sendo possível comparar modelos com diferentes números de pessoas. Nos experimentos realizados, foram utilizadas cinco métricas de avaliação de agrupamentos em imagens frontais e com variações de pose, além dos resultados computados das métricas que variam no intervalo $[0, 1]$, ou seja, quanto mais próximo de 1, melhor avaliado era o método. Os autores demonstraram que o método é promissor, gerando valores acima de 0,9 para faces frontais e 0,85 para faces com variações de pose.

Pinto, Dicarlo e Cox (2009) mostraram que um modelo simples, *Multiple Kernel Learning* (MKL), produz desempenho comparável ao estado da arte na conceituada base de imagens para o reconhecimento de faces, LFW (HUANG et al., 2007). O modelo utilizado foi associado a uma máquina de vetor de suporte (SVM). Experimentos evidenciaram que essa abordagem simples é capaz de produzir um bom desempenho com uma taxa de reconhecimento de 79,35%, sobre a base de imagens LFW (HUANG et al., 2007).

Kapoor et al. (2009) utilizaram um modelo discriminativo probabilístico que modela distribuições *a posteriori* ao propagar informações por meio de um esquema de passagem de anotações. A incerteza da estimativa fornecida pelo modelo, naturalmente, permite o uso de paradigmas de aprendizagem ativos em que o usuário é consultado a cada iteração para anotar faces adicionais. Experimentos mostram que, ao realizar uma aprendizagem ativa incorporando restrições *a priori*, obteve-se uma melhora significativa em tarefas de reconhecimento.

Guillaumin, Verbeek e Schmid (2009) fizeram uso de uma abordagem de vizinhos mais próximos que calculava a probabilidade de duas imagens pertencerem à mesma classe. Os autores denominaram esta abordagem de *Marginalized k-Nearest Neighbour* (MkNN). O método MkNN calcula a probabilidade marginal de que duas faces são da mesma pessoa, ou seja, a marginalização sobre quem é exatamente, conforme pode ser observado na Figura 2.10.

Palit et al. (2009) apresentaram uma nova técnica de agrupamento de faces com base em um algoritmo de agrupamento espectral, *Ncut*, que é um critério de otimização que minimiza o corte em um grafo. O método proposto foi comparado com uma versão modificada do *K-Means* (PERLMUTTER et al., 2006). A versão modificada do *K-Means* formava dois grupos iniciais. Em seguida, um algoritmo de divisão era utilizado para aumentar o número de grupos para a quantidade desejada. Grupos poderiam ser posteriormente incorporados, se a distância entre seus centróides estivesse dentro de um determinado limiar. Experimentos

demonstraram que a técnica proposta agrupou corretamente 98 de 164 faces, resultado superior àquele apresentado pelo algoritmo modificado baseado no *K-Means*, que agrupou corretamente apenas 45 as 164 faces.

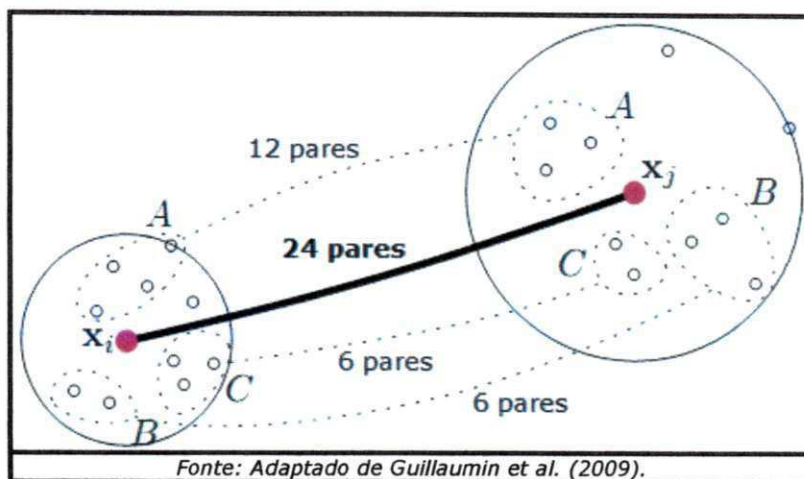


Figura 2.10 – Representação de $k = 10$ vizinhos para x_i e x_j , e 24 pares de vizinhos que possuem o mesmo nome e contribuem para o cálculo da probabilidade marginal.

Wolf, Hassner e Taigman (2009) propuseram duas técnicas similaridade, *One-Shot Similarity* (OSS) e *Two-Shot Similarity* (TSS), que foram combinados com classificadores SVM. Há indícios no trabalho de um aumento considerável no desempenho de reconhecimento com a combinação de SVM e OSS ou TSS, isso é especialmente verificado na região de baixo falso-positivo, que é a região crucial para a maioria das aplicações de reconhecimento de faces. Experimentos realizados com a combinação destas técnicas atingiram uma taxa de reconhecimento de 83,98% na base de imagens LFW (HUANG et al., 2007).

Wolf, Hassner e Taigman (2008) empregaram o algoritmo de aprendizado *One-Vs-All* SVM, no qual, para cada objeto, um SVM foi treinado utilizando imagens de treinamento do objeto como exemplos positivos e todas as imagens dos outros objetos como contra-exemplos. Dado um par de imagens, um modelo da primeira pessoa foi treinado para tentar classificar outra imagem. Em seguida, os papéis foram invertidos e o processo repetido. A média das duas classificações foi definida como a pontuação de similaridade. Em seguida, foi aplicado o limiar SVM na média para obter uma previsão. Dado um novo par de testes, foi construído um classificador baseado em OSS para cada uma das imagens usando o mesmo procedimento e média dos dois valores de classificação.

Kyperountas, Tefas e Pitas (2008) apresentaram uma metodologia que aborda o problema de reconhecimento de faces por meio de vários passos de agrupamento. Em cada etapa de agrupamento, as faces de teste e treinamento foram projetadas para um espaço discriminante e os dados projetados de treinamento foram divididos em grupos usando o algoritmo *K-Means*. Em seguida, um subconjunto dos grupos treinados foi selecionado, com base no grau de semelhança das faces nesses agrupamentos com a face de teste. Na etapa de agrupamento que se seguia, um novo espaço discriminante era definido pelo processamento desse subconjunto e ambos os conjuntos de teste e treinamento são projetados neste espaço. Este processo era repetido até que um grupo final fosse selecionado como o mais similar a face de teste.

Tian et al. (2007) propuseram uma abordagem interativa de anotação de faces que combina aprendizagem não supervisionada e interativa. No estágio não supervisionado, um algoritmo de agrupamento parcial foi responsável por encontrar os grupos mais evidentes, em vez de agrupar todas as faces em grupos individuais, o que leva a uma boa marcação inicial para uma posterior interação com o usuário. Na etapa interativa, um procedimento de anotação baseado na minimização tanto da incerteza global do sistema quanto na estimativa do número de operações do usuário foi proposto para reduzir a interação do usuário o máximo possível.

Chacon, Rivas e Ramírez (2006) fizeram uso de uma nova abordagem que visava projetar um sistema *fuzzy* para o problema de agrupamento de faces. O sistema foi baseado no algoritmo de agrupamento *Gath-Gheva* (GG) e no método *Abonyi-Szeifert* de classificação (ABONYI e SZEIFERT, 2003). Agrupamento no espaço de entrada-saída é uma técnica amplamente utilizada para criar funções de pertinência de um sistema *fuzzy* (RAMÍREZ, 2005; ABONYI, BABUSKA e SZEIFERT, 2002). Uma vez que os grupos foram gerados, as funções de pertinência do sistema foram criadas pela projeção dos grupos nos eixos dos dados de entrada, como pode ser observado na Figura 2.11. Essas funções de pertinência foram usadas para definir os antecedentes das regras *fuzzy*. O algoritmo *Gath-Gheva* (GG) é um algoritmo de agrupamento *fuzzy* que utiliza o rótulo dos dados para criar

grupos *fuzzy* para construção dos antecedentes de um sistema de inferência *fuzzy*.

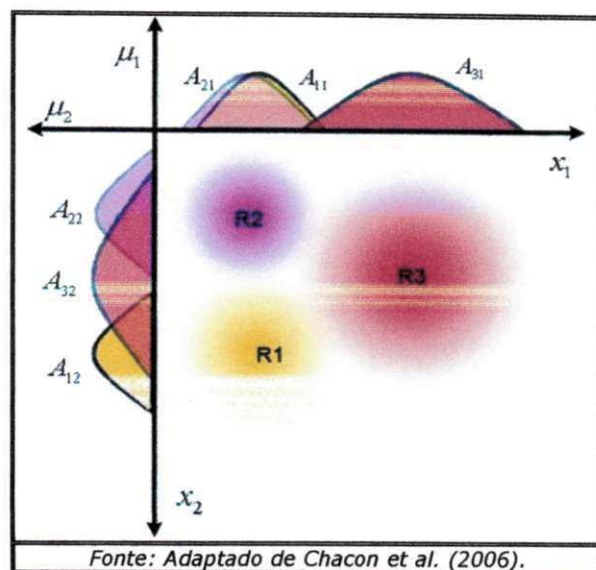


Figura 2.11 – Três funções de pertinência criadas pela projeção dos grupos nos eixos dos dados de entrada.

Perronnin e Dugelay (2005) adotaram um método baseado no algoritmo *Expectation-Maximization* (EM) como técnica de agrupamento, por utilizar uma métrica de similaridade probabilística. Durante o passo-E, a distância entre cada observação e cada centróide de um grupo era calculada e cada observação é atribuída ao grupo mais próximo (ou probabilisticamente a todos os clusters). Durante o passo-M, o centróide do cluster era atualizado com as observações atribuídas. A etapa de atualização também dependia da distância escolhida, desde que o centróide fosse definido como o ponto que minimiza a distância média entre as observações atribuídas e do próprio centróide.

Chen et al. (2003) e Zhang et al. (2003) utilizaram uma técnica de aprendizagem de anotação de faces, com o intuito de gerar uma lista de nomes candidatos para uma nova face não anotada de acordo com sua similaridade em relação às demais previamente anotadas, conforme pode ser observado na Figura 2.12. O método tem como base o algoritmo *K*-vizinhos mais próximos (*K*-NN). Dada uma face não anotada, entre as *K* faces anotadas mais próximas, a lista de nomes candidatos era gerada pela classificação dos nomes, de acordo com a soma de similaridades da face não anotada.

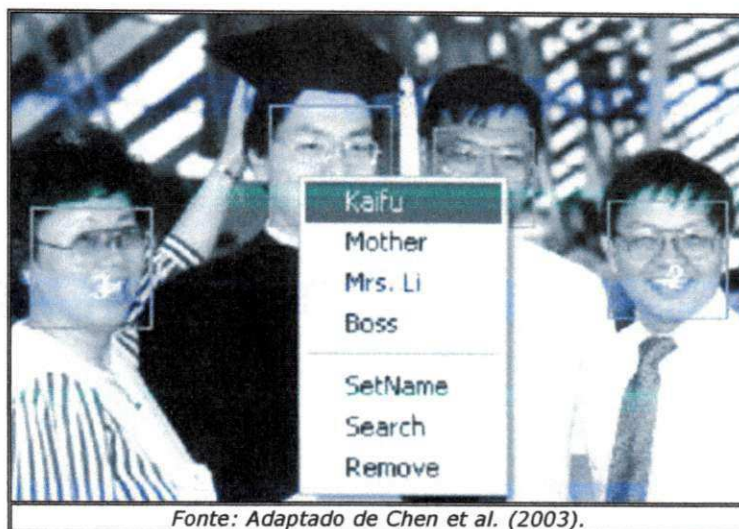


Figura 2.12 – Menu *popup* do sistema de anotação de faces.

Na próxima seção, serão detalhadas as principais métricas de avaliação de agrupamento identificadas durante a revisão bibliográfica que foram empregadas no experimentos comparativos realizados.

2.7. Métricas de Avaliação de Agrupamento

Uma das principais dificuldades em problemas de classificação consiste na correta avaliação do desempenho do classificador. Isto geralmente é feito a partir da aplicação de uma medida de desempenho comum, e.g., erro quadrado médio (MSE), área em porcentagem sob a curva ROC (*Receiver Operating Characteristic*), dentre outras. Todas estas métricas permitem comparar o resultado rotulado pelo algoritmo de classificação supervisionada com rótulos previamente conhecidos (*ground truth*).

Diferentemente, para algoritmos de agrupamento não supervisionado torna-se útil a introdução de uma medida que permita avaliar o desempenho do algoritmo em relação à qualidade da divisão dos dados de entrada em classes ou grupos diferentes, focando na relação entre os elementos de cada classe e não em rótulos previamente fornecidos. Esta é a razão principal da utilização de métricas de validação de agrupamentos em problemas de classificação não supervisionada. As métricas descritas a seguir foram escolhidas conforme resultado da análise de trabalhos revisados anteriormente.

Rand Index (RI) é uma métrica da similaridade entre dois

agrupamentos de dados. Dados um conjunto de n elementos $S = \{o_1, \dots, o_t\}$ e duas partições de S a comparar, $X = \{x_1, \dots, x_r\}$ e $Y = \{y_1, \dots, y_n\}$ (em que x_i e y_i denotam agrupamentos de elementos de S), consideram-se as seguintes propriedades:

- (i) a , número de pares de elementos em S que estão num mesmo conjunto em X e num mesmo conjunto em Y ;
- (ii) b , número de pares de elementos em S que estão num mesmo conjunto em X e em conjuntos distintos em Y ;
- (iii) c , número de pares de elementos em S que estão em conjuntos distintos em X e num mesmo conjunto em Y ; e
- (iv) d , número de pares de elementos em S que estão em conjuntos distintos em X e em conjuntos distintos em Y .

A métrica *Rand Index* (RAND, 1971) é definida como sendo:

$$RI = \frac{a + d}{a + b + c + d} = \frac{a + d}{\binom{n}{2}} \quad (2.1)$$

Intuitivamente, pode-se considerar a soma $a + d$ como sendo o número de concordâncias entre X e Y e $b + c$ como sendo o número de discordâncias entre X e Y . O índice Rand varia no intervalo $[0,1]$, com 0 indicando que os dois conjuntos de dados não concordam com qualquer par de pontos e 1 indicando que os agrupamentos de dados são exatamente iguais.

De fato, é desejável que o índice de similaridade entre as duas partições aleatórias assuma valor próximo de zero. O problema com o índice de Rand é que seus valores esperados para pares de partições aleatórias nem sempre são discriminantes, ou seja, geralmente são gerados valores próximos. Hubert e Arabie (1985), ao tomar a distribuição hipergeométrica generalizada como o modelo de aleatoriedade, encontraram o valor esperado para $a + d$ e sugeriram uma correção na métrica *Rand*, no sentido de aumentar o poder de diferenciação entre duas partições aleatórias.

Adjusted Rand Index (ARI) é um melhoramento da métrica *Rand* e

se tornou um dos índices de avaliação de agrupamentos mais utilizados e recomendados para a medição de concordância entre duas partições com diferentes números de grupos. ARI (SANTOS e EMBRECHTS, 2009) é formulado como sendo:

$$ARI = \frac{Index - Expected_Index}{Max_Index - Expected_Index} = \frac{\binom{n}{2} (a+d) - [(a+b)(a+c) + (c+d)(b+d)]}{\binom{n}{2} - [(a+b)(a+c) + (c+d)(b+d)]} \quad (2.2)$$

De maneira similar à métrica RI, a métrica ARI assume valores de 0 (agrupamentos totalmente diferentes) a 1 (agrupamentos idênticos).

Precision-Recall (PR), ou precisão e revocação, são duas métricas utilizadas para avaliar a corretude de um algoritmo de reconhecimento de padrões. Elas podem ser vistas como versões estendidas da acurácia, uma métrica simples que calcula a fração de instâncias na qual o resultado correto é retornado (DAVIS e GOADRICH, 2006). Precisão mede a habilidade de o sistema retornar apenas resultados relevantes (OLSEN e DELEN, 2008), podendo ser definida como:

$$P = \frac{Verdadeiro\ Positivo}{Verdadeiro\ Positivo + Falso\ Positivo} \quad (2.3)$$

Cobertura mede a habilidade de o sistema retornar todos os resultados relevantes (OLSEN e DELEN, 2008) e pode ser formulada como:

$$R = \frac{Verdadeiro\ Positivo}{Verdadeiro\ Positivo + Falso\ Negativo} \quad (2.4)$$

F-Mesure (F) é a média harmônica de precisão e cobertura, também conhecida como *F1 Score*. A métrica F (KANDEFER e SHAPIRO, 2009), pode ser definida como:

$$F = 2 \cdot \frac{P \cdot R}{P + R} \quad (2.5)$$

Similarmente às métricas *Rand*, assume valores no intervalo [0,1].

Jaccard Index (JI), também conhecido como o coeficiente de

similaridade de *Jaccard*, é uma métrica utilizada para comparação de similaridade e diversidade entre conjuntos de amostras, e é definido como o tamanho da interseção dividido pelo tamanho da união do conjunto de amostras. O índice de *Jaccard* (IVCHENKO e HONOV, 1998), pode ser expresso por:

$$J = \frac{a}{a+b+c} \quad (2.6)$$

Da mesma forma que as métricas *Rand* e *F*, o índice de *Jaccard* varia no intervalo [0,1].

Folkes and Mallows Index (FM) é um método de avaliação externa utilizado para determinar a similaridade entre os dois agrupamentos. Esta métrica de similaridade pode mensurada entre dois agrupamentos hierárquicos ou entre um agrupamento e uma classificação de referência. Quanto maior o valor do índice Fowlkes-Mallows, mais próximos são os agrupamentos (ou *clusters*). O índice de *Folkes and Mallows* (HALKIDI, BATISTAKIS e VAZIRGIANNIS, 2011), pode ser definido como sendo:

$$FM = \frac{a}{\sqrt{m_1 m_2}} = \sqrt{\frac{a}{a+b} \cdot \frac{a}{a+c}} \quad (2.7)$$

sendo $m_1 = (a+b)$ e $m_2 = (a+c)$. Da mesma forma que as métricas *Rand*, *F* e *Jaccard*, o índice de *Folkes e Mallows* varia no intervalo [0,1].

Na próxima seção, são sumariados e analisados os principais pontos de cada uma das seções anteriores, com o objetivo de destacar o estado da arte para o problema objeto de estudo, bem como permitir a identificação de estratégias relevantes para compor a abordagem proposta.

2.8. Considerações sobre os Trabalhos Analisados

Os trabalhos analisados anteriormente foram agrupados em ordem cronológica decrescente, conforme ilustrado na Tabela 2.1. Nesta tabela, a coluna *Propósito do Sistema* indica qual a aplicação do sistema proposto, a coluna *Característica* indica que características foram utilizadas para

construção do modelo, a coluna *Métrica de Similaridade* expressa que medida de distância foi utilizada para comparação de faces candidatas, a coluna *Técnica de Agrupamento/Reconhecimento* especifica os métodos desenvolvidos ou empregados para a composição do sistema, a coluna *Taxa de Avaliação* especifica os valores dos resultados experimentais obtidos em cada trabalho, a coluna *Robusto a Variações* indica que tipo de variação o método é robusto, e finalmente, a coluna *Base de Imagens* explicita as bases de faces adotadas para verificar a taxa de reconhecimento ou agrupamento do sistema proposto.

Do resumo apresentado na Tabela 2.1 e da exposição anterior, evidencia-se o fato de que a área de reconhecimento de faces aplicada ao agrupamento de faces tem despertado muito interesse, tanto de pesquisadores como de grandes empresas, e.g., *Google, Apple, HP e Yahoo!*, dentre outras. Este fato este é reforçado pela crescente demanda por métodos para a organização automática de fotografias, devido ao aumento significativo da quantidade de imagens acumuladas por usuários de câmeras digitais e dispositivos móveis multimídia, as quais, em sua grande maioria, contêm pessoas como principal tema.

No tocante à característica, constata-se uma grande diversidade de técnicas e métodos, sendo as características SIFT e LBP as mais presentes. Por outro lado, Chacon, Rivas e Ramirez (2006) em seu trabalho afirmaram que as duas características mais importantes para o reconhecimento utilizadas atualmente são PCA e LDA. Para sustentar tal afirmação, os autores apresentaram uma análise baseada em quatro métricas de avaliação: (i) decorrelação; (ii) confiança; (iii) capacidade discriminativa; e (iv) tempo de computação.

Tabela 2.1 – Resumo dos trabalhos analisados.

| Artigo | Propósito do Sistema | Característica | Métrica de Similaridade | Técnica de Agrupamento/Reconhecimento | Taxa de Avaliação | Robusto a Variações | Base de Imagens |
|------------------------------|----------------------|--|--|--|---------------------------------|---------------------|-----------------|
| Zhang et al., 2011 | Agrupamento | ASM, Pontos fiduciais | Hausdorff | ISODATA | - | - | MUCT |
| Zhu, Wen e Sun, 2011 | Agrupamento | SIFT | L1 | Rank-Order | Precision: 0,97 Recall: 0,86 | Orientação, Escala | Gallagher |
| Jayech e Mahjoub, 2011 | Agrupamento | Pixel-wise | Tangente | K-Means | - | - | ORL |
| Zhang, Zhang e Tretter, 2010 | Agrupamento | Gaussianas, LoG | EMD | HAC | Precision: 0,99 Recall: 0,80 | Pose | CMU |
| Prince e Elder, 2010 | Agrupamento | PLDA | - | Agrupamento Bayesiano, Modelo Generativo | Precision: 0,90 Recall: 0,85 | Pose | XM2VTS |
| Kapoor et al., 2009 | Agrupamento | PCA, Implicit Elastic Matching with Random Projections | L1, IDF | Aprendizagem ativa com restrições | - | Pose | ORL |
| Palit et al., 2009 | Agrupamento | PCA, Pontos fiduciais | Euclidiana, Média, Média ponderada | Agrupamento Spectral, N-Cut | Precision: 0,73 | - | Própria |
| Pinto et al., 2009 | Reconhecimento | V1 | Pair Matching | MKL | Reconhecimento: 0,79 | Nenhuma | LFW |
| Hua e Akbarzadeh, 2009 | Reconhecimento | DoG | Robust Elastic and Partial Matching Metric | - | - | Iluminação | ORL, LFW |
| Guillaumin et al., 2009 | Agrupamento | LBP, TPLBP, FPLBP, SIFT | LDML | MkNN | - | - | LFW |

Tabela 2.1 – Resumo dos trabalhos analisados (continuação).

| Artigo | Propósito do Sistema | Característica | Métrica de Similaridade | Técnica de Agrupamento/Reconhecimento | Taxa de Avaliação | Robusto a Variações | Base de Imagens |
|---|-----------------------------|--|---------------------------------|--|--|----------------------------|------------------------|
| Wolf, Hassner e Taigman, 2009 | Reconhecimento | LBP, TPLBP, FPLBP, Gabor C1, SIFT | OSS, TSS | SVM | Reconhecimento: 0,83 | - | LFW |
| Wolf, Hassner e Taigman, 2008 | Agrupamento | LBP, TPLBP, FPLBP, Gabor C1 | Euclidiana, Euclidiana Quadrada | One-Vs-All SVM | - | - | LFW |
| Kyperountas et al., 2008 | Agrupamento | LDA | Euclidiana Quadrada | Agrupamento Discriminante Adaptativo | - | - | UMIST, XM2VTS |
| Antonopoulos, Nikolaidis e Pitas (2007) | Agrupamento | SIFT | Euclidiana | HAC | F-Measure: 0,87 Overall Entropy: 0,79 | Orientação, Escala | Própria |
| Tian et al., 2007 | Agrupamento | LBP | Chi-square | Agrupamento Parcial, Anotação Interativa | Precision: 0,86 | - | Riya |
| Chacon, Rivas e Ramirez, 2006 | Agrupamento | FFL | Euclidiana, D | Agrupamento Fuzzy Gath-Gheva | Precision: 0,90 | Iluminação | ORL, YALE |
| Perronnin e Dugelay, 2005 | Agrupamento | Wavelets de Gabor | L1, Euclidiana, Cosseno | EM | - | - | FERET |
| Li e Tang, 2004 | Agrupamento | Método Bayesiano | DIFS, DFFS | HAC | Reconhecimento: 0,97 | - | FERET, XM2VTS |
| Chen et al., 2003 | Agrupamento | Auto-correlograma, Wavelets | H-Hit rate | K-NN | - | - | - |
| Zhang et al., 2003 | Agrupamento | Auto-correlograma, Momentos de textura-cor | H-Hit rate | K-NN | - | - | FERET |

De maneira similar, Perronnin e Dugelay (2005) também reforçaram tal análise ao argumentar que o PCA é uma técnica de redução de dimensão que é ótima em relação à compressão de dados, em geral é considerada subótima para o reconhecimento e que para tal tarefa, o LDA deve ser preferido ao PCA.

No que diz respeito à tarefa de reconhecimento propriamente dita, as características do tipo V1 e DoG são computacionalmente mais simples do que aquelas adotadas no trabalho de Wolf, Hassner e Taigman et al. (2009), no entanto, quando comparadas às taxas de reconhecimento na base LFW (HUANG et al., 2007), a combinação das características proposta por Wolf, Hassner e Taigman et al. (2009) atingiu uma taxa de 83,98% que supera ambas V1-like com taxa de 79,35% (PINTO, COE e DICARLO, 2009) e DoG com taxa de 78,64% (HUA e AKBARZADEH, 2009).

Por outro lado, em relação à tarefa de agrupamento, tal comparação não pode ser efetuada, devido à grande variabilidade das bases de faces utilizadas e pelo fato de que a resposta resultante dos sistemas são grupos. Vale salientar também que apenas 25% dos autores dos trabalhos relacionados (ZHU, WEN e SUN, 2011; PRINCE e ELDER, 2010; PALIT et al., 2009; CHACON, RIVAS e RAMIREZ, 2006; PERRONNIN e DUGELAY, 2005) usaram métricas objetivas para avaliação do agrupamento, em sua maioria apenas apresentaram resultados de avaliações subjetivas.

Em relação à métrica de similaridade, constata-se que a distância Euclidiana foi a mais freqüentemente empregada nos trabalhos analisados, seguida das métricas L1 e H-Hit. Entretanto, Palit et al. (2009) verificaram em seus experimentos que a distância da média ponderada em torno da mediana produziu melhores resultados do que a distância Euclidiana. Da mesma forma, Perronnin e Dugelay (2005) evidenciaram que a medida Cosseno apresentou melhores resultados do que medidas L1 e Euclidiana. Assim, não existe uma métrica de similaridade que seja de consenso geral na literatura, ou seja, faz-se necessária a realização de um conjunto de experimentos para determinar qual a distância mais adequada para medir a similaridade entre as características escolhidas.

Quanto à técnica de agrupamento/reconhecimento, mais uma vez percebe-se uma grande variação nos métodos propostos, sendo o algoritmo HAC o mais presente. Adicionalmente, foi possível verificar que a base para as técnicas de agrupamento foi o algoritmo *K-Means*, ou seja, a maioria das técnicas desenvolvidas são variações que aperfeiçoaram o algoritmo de acordo com a necessidade do sistema, quer seja em relação às características adotadas quanto à métrica de similaridade empregada.

A estratégia de limitação de escopo das técnicas que poderiam ser aplicadas ao sistema proposto de agrupamento de faces, foi a eliminação de técnicas que necessitam de um treinamento prévio supervisionado e/ou da necessidade de intervenção do usuário quer seja no treinamento quanto no teste do sistema.

Outro ponto importante identificado é o fato 35% dos autores dos trabalhos analisados (ZHU, WEN e SUN, 2011; PRINCE e ELDER, 2010; KAPOOR et al., 2009; PINTO, COX e DICARLO, 2009; HUA e AKBARZADEH, 2009; CHACON, RIVAS e RAMIREZ, 2006) indicarem claramente se seus métodos são robustos a algum tipo variação. No entanto, uma possível indicativa para tal suposição, seria em relação às bases de imagens utilizadas nos experimentos, por exemplo, ao considerar que uma dada abordagem ser robusta a variações de pose, iluminação e expressão facial por utilizar a base de imagens LFW (HUANG et al., 2007) como teste e apresentar resultados consideráveis.

Diante deste contexto, pode-se concluir que ainda não foi desenvolvido um sistema para o agrupamento de faces que seja robusto à uma variabilidade de condições tais como iluminação, pose, expressão facial e oclusão de partes da face. Assim, uma opção para alcançar este objetivo seria criar um sistema híbrido que explorasse o melhor de cada técnica.

Portanto, a partir das evidências de todas as discussões e comentários anteriores, uma possível composição inicial para o futuro sistema seria a escolha das características SIFT, LBP e LDA, as distâncias Euclidiana e Cosseno como métrica de similaridade e os algoritmos HAC e *K-Means* como técnicas para o método de agrupamento.

Durante o desenvolvimento da abordagem proposta, foi verificado que a característica SURF, apesar de apresentar comportamento similar à característica SIFT, apresentou melhores resultados do que a característica SIFT, além dos experimentos comparativos elaborados no trabalho de Juan e Gwon (2009) evidenciarem que o SIFT deve ser preterido ao SURF em relação a desempenho na geração dos descritores (cerca de 3 vezes mais lento) e na taxa de reconhecimento de 78,1% contra 85,7% do SURF na base de imagens Caltech (CALTECH FACE DATABASE, 2010).

Adicionalmente, em relação à estratégia de comparação de faces, foi verificado que a utilização de distâncias como a Euclidiana e Cosseno, não serem compatíveis com os descritores SURF, para tanto, utilizou-se um procedimento baseado no algoritmo *Fast Approximate Nearest Neighbors* – FANN (MUJA e LOEW, 2009), para determinar o grau de similaridade entre duas faces em função da correspondência entre descritores e inspirado na abordagem proposta por Antonopoulos, Nikolaidis e Pitas (2007).

2.9. Considerações Finais

Diante do exposto, constata-se que apesar dos grandes avanços alcançados na área de reconhecimento de faces, ainda há muito trabalho a ser feito. De acordo com a revisão bibliográfica, os sistemas destinados ao agrupamento de faces obtiveram elevadas taxas de reconhecimento quando imagens de faces foram adquiridas em condições controladas. Todavia, tal situação está muito distante da realidade, na qual existe uma necessidade crescente de sistemas cada vez mais robustos, i.e., mais tolerantes a variações de pose, iluminação, oclusão e expressão facial.

Assim, neste capítulo foram apresentados métodos, métricas e técnicas que serviram de base para a idealização do trabalho ora desenvolvido, um sistema de agrupamento de faces robusto a variações de iluminação, expressão facial e pose, no contexto de fotografia digital de consumo, com o intuito de facilitar a organização e o compartilhamento de coleções de fotografias de usuários.

No próximo capítulo, apresenta-se, de forma detalhada, a abordagem

proposta para a solução do problema objeto de estudo, incluindo-se a arquitetura geral, o fluxo de processamento e o funcionamento de cada módulo. Adicionalmente, são apresentados alguns detalhes de implementação e de organização interna da aplicação desenvolvida para validar a abordagem proposta.

Capítulo 3

Abordagem Proposta

Neste capítulo, é apresentada a abordagem proposta para a solução do problema de agrupamento de faces em coleções de fotografias digitais. A abordagem proposta foi inspirada em características de alguns trabalhos relevantes apresentados no Capítulo 2. Como exemplos, podem-se citar a abordagem adotada por Zhu, Wen e Sun (2011), no tocante à característica empregada para identificação pessoal, assim como a abordagem proposta por Antonopoulos, Nikolaidis e Pitas (2007), no tocante à estratégia de comparação das faces detectadas, e aos trabalhos de Zhang, Zhang e Tretter (2010) e Li e Tang (2004), em relação à técnica de agrupamento.

Desta forma, o propósito deste capítulo é apresentar, de forma detalhada, a abordagem proposta para a solução do problema objeto de estudo, incluindo-se uma visão geral de arquitetura do sistema, seguida do fluxo de processamento de cada fase do processo e o funcionamento de cada módulo da técnica elaborada. Adicionalmente, são apresentados alguns detalhes de implementação e de organização interna da aplicação desenvolvida para validar a abordagem proposta.

3.1. Visão Geral da Arquitetura Proposta

A arquitetura da técnica desenvolvida foi inspirada em abordagens encontradas na revisão bibliográfica, sendo ilustrada na Figura 3.1 e descrita sucintamente como segue:

- (i) A imagem de entrada é submetida, primeiramente, ao módulo de detecção de faces utilizado para buscar todas as faces contidas na imagem e, posteriormente, ao módulo de correção da orientação das faces detectadas pela determinação e localização da posição

- dos olhos, gerando faces normalizadas;
- (ii) As imagens de faces normalizadas são submetidas a duas etapas de *pré-processamento*, a saber: (a) compensação de iluminação; e (b) equalização de histograma, originando faces tratadas;
 - (iii) Após a etapa de pré-processamento, descritores SURF são extraídos para cada uma das faces normalizadas e tratadas, originando suas respectivas representações faciais;
 - (iv) Em seguida, todas as representações faciais são confrontadas (*matching*), de forma a comporem uma matriz de similaridade; e
 - (v) Por fim, a partir da matriz de similaridade das representações faciais, a técnica de agrupamento é aplicada, para a geração final dos grupos de faces semelhantes.

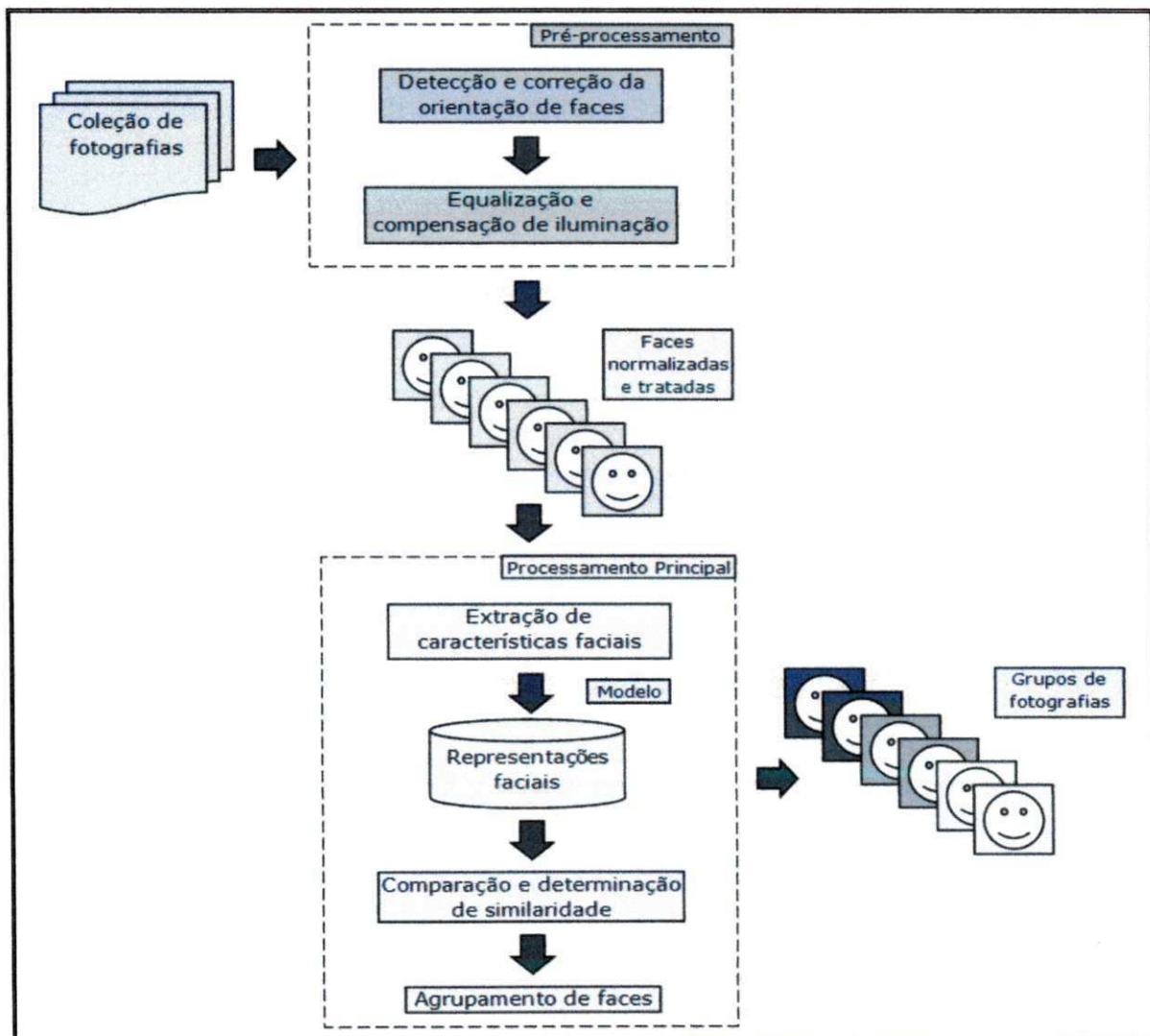


Figura 3.1 – Arquitetura da técnica proposta.

Tomando como base artigos relevantes da área de agrupamento de faces revisados no capítulo anterior, algumas técnicas da área de Processamento Digital de Imagens e Visão Computacional foram estudadas e implementadas, com o intuito de serem usadas nos módulos de pré-processamento e no cerne do processamento da abordagem proposta. Cada técnica implementada é apresentada nas seções que seguem.

O propósito do presente trabalho é a definição de uma abordagem para o agrupamento de faces em coleções de fotografias digitais, usando, dentre as técnicas analisadas, os algoritmos que atingiram melhores resultados. Neste sentido, para as etapas de pré-processamento ou de processamento principal, apenas os melhores foram escolhidos para serem avaliados.

Portanto, algumas técnicas de PDI estudadas podem ser substituídas por técnicas similares em cada um dos módulos propostos, e.g., a compensação da iluminação por uma técnica diferente da filtragem homomórfica; o uso de características SIFT ou LDA, ao invés do SURF; algoritmos de agrupamento *K-Means*, EM, dentre outros, ao invés do HAC. O cerne da abordagem proposta está, portanto, no arcabouço elaborado para o agrupamento de faces, bem como na proposição do fluxo de etapas a serem seguidas, até a geração final dos grupos de faces similares.

Os resultados apresentados neste capítulo são resultantes da utilização de técnicas em cujos experimentos foram verificados os melhores resultados, considerando-se cada uma das etapas da abordagem proposta. Tais experimentos são apresentados e discutidos no próximo capítulo. A seguir, cada etapa do fluxo de execução mostrado na Figura 3.1 é detalhada e exemplos de imagens processadas, após a aplicação de cada módulo, são apresentadas.

3.2. Pré-Processamento

A etapa inicial do sistema é responsável por um *pré-processamento* que engloba duas sub-etapas, a saber: (i) *detecção e correção da orientação de faces*; (ii) *equalização e compensação da iluminação*, conforme ilustrado anteriormente (Figura 3.1). Cada uma das sub-etapas foram novamente

divididas em duas partes que serão detalhadas na seqüência.

3.2.1. Detecção e Correção da Orientação de Faces

Considerando que o foco principal desta dissertação são faces humanas, a etapa inicial do sistema consiste na detecção de faces, seguida, de uma correção da orientação, pela determinação e localização da posição dos olhos, originando faces normalizadas com o objetivo de facilitar a posterior comparação.

O módulo de detecção e correção da orientação de faces tem como objetivo, localizar, recortar e normalizar o tamanho e a orientação de faces humanas, como preparação para a etapa de equalização e compensação de iluminação. A correção da orientação é feita pela rotação da imagem em torno do ponto central. O ângulo de rotação é aquele formado pela reta que une o centro dos olhos e o eixo horizontal. Uma ilustração de aplicação deste módulo é apresentada na Figura 3.2 e as descrições de cada um de seus componentes são apresentadas ao longo desta seção.

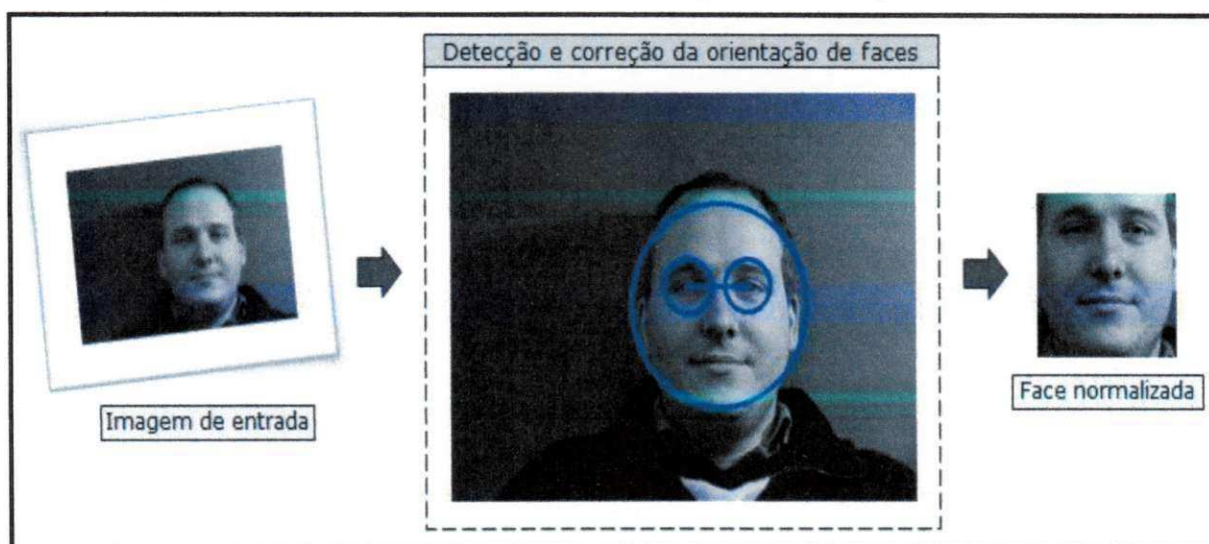


Figura 3.2 – Etapa 1 – Módulo de detecção e correção da orientação de faces.

A etapa de detecção de faces e olhos é realizada com o uso da biblioteca OpenCV (OPENCV, 2010). Esta biblioteca é composta de funções e classes, escritas nas linguagens C e C++, que implementam diversos algoritmos de PDI e VC. A escolha dos detectores da biblioteca OpenCV foi baseada no estudo comparativo entre sete diferentes bibliotecas elaborado por Degtyarev e Seredin (2010), no qual a OpenCV foi classificada como a

segunda melhor biblioteca, em termos de precisão e velocidade de processamento, sendo a única de código aberto.

De acordo com Bradski, Kaehler e Pisarevsky (2005), os detectores de faces e olhos da OpenCV foram desenvolvidos por meio da técnica de detecção de objetos proposta por Viola e Jones (2001), que utiliza características de *Haar* e um conjunto de classificadores interligados em cascata.

As características de *Haar* foram originalmente propostas por Papageorgiou, Oren e Poggio (1998), sendo assim denominadas porque são calculadas de forma similar aos coeficientes das Transformadas *Wavelets* de *Haar*. Estas características avaliam diferenças de intensidade média entre duas regiões retangulares adjacentes da imagem.

O detector de faces foi treinado com imagens de tamanho fixo (24x24 pixels) e a detecção foi feita deslocando-se uma janela de busca deste mesmo tamanho sobre a imagem de entrada e verificando-se a região correspondente da imagem a cada deslocamento contém uma face. O processo utilizado para detectar os olhos assumiu que a região de busca era composta de uma face previamente detectada. O detector de olhos foi treinado com imagens de tamanho 18x12 pixels e seguiu a mesma estratégia de janela deslizante do detector de faces (CASTRILLÓN et al., 2008).

3.2.2. Equalização e Compensação de Iluminação

Após a aplicação dos detectores de faces e olhos, mencionada na seção anterior, a imagem de entrada é submetida a um processo de correção. Este processo é baseado em dois algoritmos de PDI, utilizados na seguinte ordem: (i) *filtragem homomórfica*, visando aprimorar a qualidade da imagem, por meio da compressão da faixa dinâmica de brilho, simultaneamente ao aumento do contraste; e (ii) *equalização de histograma*, buscando realçar a imagem em termos de contraste.

As imagens de face de saída, denominadas faces normalizadas e tratadas, servirão de entrada para o fluxo principal da abordagem proposta.

Uma ilustração de aplicação deste módulo é apresentada na Figura 3.3 e as descrições de cada componente do mesmo são apresentadas nas subseções que seguem.



Figura 3.3 – Etapa 2 – Módulo de equalização e compensação de iluminação.

3.2.2.1. Filtragem Homomórfica

Uma imagem é formada a partir da luz refletida pelos componentes da cena retratada, de forma que a imagem percebida pelo observador pode ser caracterizada por dois parâmetros: (i) a quantidade de iluminação que incide na cena; e (ii) a quantidade de iluminação refletida pelos componentes da cena. Desta forma, uma imagem $f(x,y)$ pode ser expressa como o produto dos componentes de iluminação, $i(x,y)$, e refletância, $r(x,y)$ (GONZALEZ e WOODS, 2010), ou seja:

$$f(x,y) = i(x,y)r(x,y) \quad (3.1)$$

Este modelo de formação de imagens é conhecido como modelo de iluminação-refletância e tem sido usado para o melhoramento da qualidade de imagens adquiridas sob condições de iluminação não adequadas para a análise automática (DELAC, GRGIC e KOS, 2006).

O objetivo do módulo de compensação de iluminação utilizado neste trabalho é reduzir o componente de iluminação de uma imagem de face, de forma que a imagem final se apresente como uma aproximação para a refletância da face, a qual é independente das condições de iluminação. A filtragem homomórfica foi o procedimento utilizado para separar estes componentes, conforme apresentado no diagrama da Figura 3.4.

O processo de filtragem homomórfica permite manipular sinais

combinados através de operações não-lineares. O algoritmo consiste basicamente em transformar o problema da combinação não-linear em um problema de combinação linear, utilizando o princípio da convolução por separabilidade e operações matemáticas apropriadas, em função de um filtro homomórfico $H(u,v)$ (GONZALEZ e WOODS, 2010).

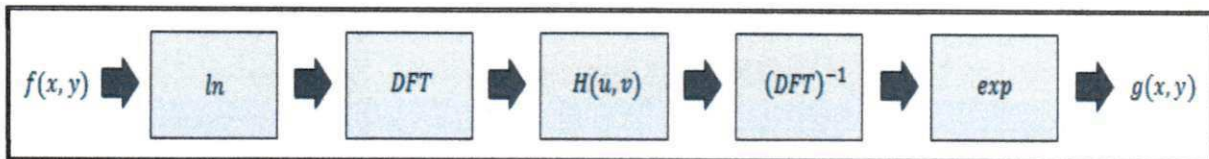


Figura 3.4 – Resumo dos passos da filtragem homomórfica.

O componente de iluminação de uma imagem geralmente é caracterizado por variações espaciais suaves, enquanto o componente de refletância tende a variar abruptamente, particularmente nas junções de diferentes objetos. Essas características levam a associar as componentes espectrais de baixas frequências de uma imagem à iluminação e aquelas de altas frequências à reflectância. Assim, a função da filtragem homomórfica é atenuar o componente de iluminação de uma imagem e ressaltar sua a refletância. O resultado da aplicação da filtragem homomórfica, presente na segunda etapa da abordagem proposta, é ilustrado na Figura 3.5B.

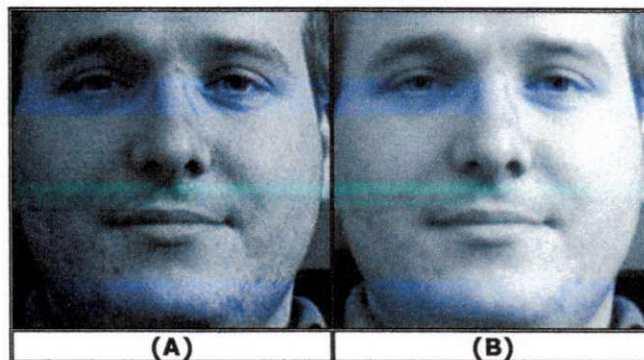


Figura 3.5 – Exemplo de compensação de iluminação: (A) Imagem normalizada; e (B) Imagem após filtragem homomórfica.

3.2.2.2. Equalização de Histograma

A equalização de histograma é uma técnica a partir da qual se procura redistribuir os valores de tons de cinza dos pixels em uma imagem, de modo a obter uma imagem com histograma que aproxime uma distribuição uniforme, na qual o número (percentual) de pixels de qualquer nível de cinza é, idealmente, o mesmo (MARQUES FILHO e VIEIRA NETO, 1999).

Para uma imagem I , contendo n , pixels com um histograma h , o histograma normalizado (PARKER, 1997), pode ser definido como:

$$h_{norm}(i) = \frac{1}{n} h(i); i = 0, \dots, k-1 \quad (3.2)$$

O histograma normalizado (h_{norm}) corresponde à densidade de probabilidade de valores de intensidade da imagem. O histograma acumulado é definido em termos do histograma normalizado como:

$$h_{acum}(r) = \sum_{i=0}^r h_{norm}(i); r = 0, \dots, k-1 \quad (3.3)$$

em que a função h_{acum} é uma função de distribuição de probabilidade acumulada. Portanto, trata-se uma função não decrescente com $h_{acum}(k-1) = 1$. A partir da Equação 3.3, o histograma normalizado é expresso como:

$$h_{norm}(q) = h_{acum}(q) - h_{acum}(q-1); q = 0, \dots, k-1 \quad (3.4)$$

Tratando h_{norm} e h_{acum} como funções de uma variável contínua x , estes vetores podem ser vistos como uma função de densidade de probabilidade (FDP) e uma função de distribuição acumulada (FDA), considerando o relacionamento $h_{norm} = dh_{acum}(x)/dx$. Portanto, a imagem produzida pela operação de equalização pode ser descrita por:

$$E(x, y) = eq(I(x, y)) = h_{acum}(I(x, y)) \quad (3.5)$$

em que a imagem E tem um histograma com distribuição que aproxima da distribuição uniforme. O resultado da aplicação da equalização de histograma, presente na segunda etapa da abordagem proposta, é ilustrado na Figura 3.5. Pode-se observar que na imagem resultante do processamento os detalhes das características faciais são realçados.

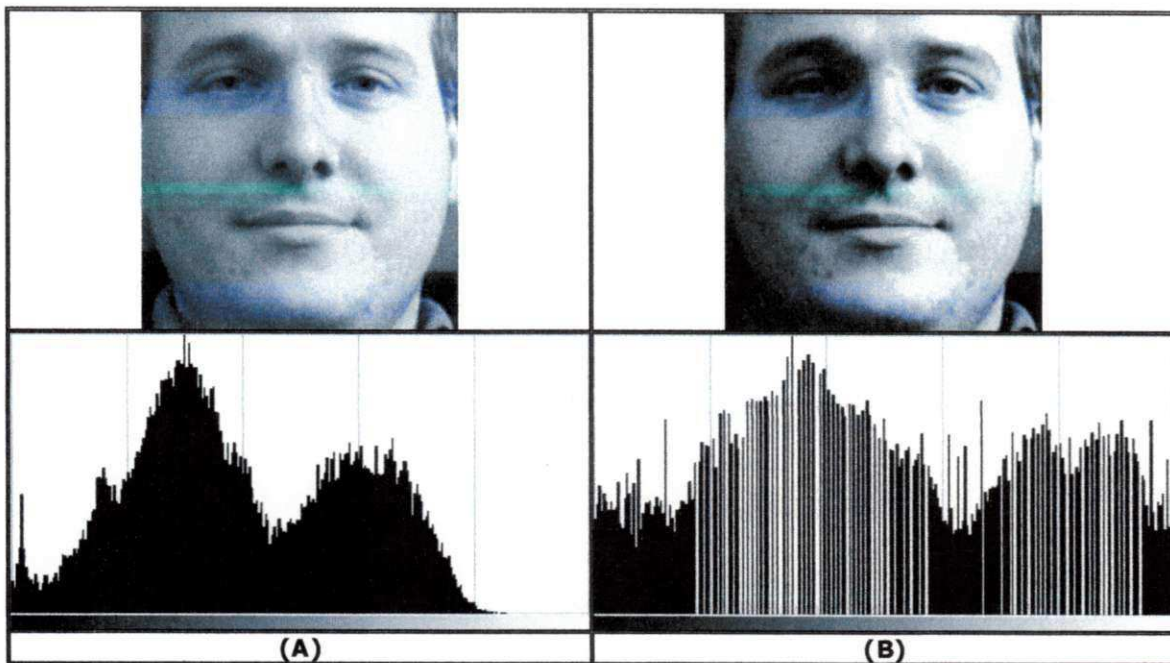


Figura 3.6 – Exemplo de equalização de histograma: (A) Imagem após filtragem homomórfica; e (B) Imagem após equalização de histograma.

3.3. Processamento Principal

O cerne de processamento da abordagem proposta compreende três módulos, a saber: (i) *extração de características faciais*, visando à formação de representações faciais por meio do uso de descritores SURF; (ii) *comparação e determinação de similaridade*, para a determinação da semelhança entre as representações faciais e geração de uma matriz de similaridade; e (iii) *técnica de agrupamento*, objetivando a composição dos grupos de faces semelhantes a partir de um agrupamento aglomerativo hierárquico.

3.3.1. Extração de Características Faciais

Ao tentar comparar características entre imagens diferentes, freqüentemente se é confrontado com o problema das mudanças de escala, ou seja, as diferentes imagens a serem analisadas podem ser tomadas a uma distância diferente dos objetos de interesse e, conseqüentemente, esses objetos serão retratados em diferentes tamanhos. Ao se tentar extrair uma mesma característica de duas imagens usando uma vizinhança de tamanho fixo, por causa da mudança de escala, os seus padrões de intensidade não corresponderão (LAGANIÈRE, 2011).

Para resolver este problema, o conceito de características invariantes à escala foi introduzido na área de VC. Nos últimos anos, duas abordagens promissoras para detectar regiões salientes em imagens foram desenvolvidas: *Scale Invariant Feature Transform* – SIFT (LOWE, 1999) e *Speeded Up Robust Features* – SURF (BAY, TUYTELAARS e VAN GOOL, 2006). Ambas as abordagens não apenas detectam pontos de interesse (*keypoints*), como também propõem um método para criação de um descritor invariante. Este descritor pode ser usado para identificar pontos de interesse únicos e compará-los, mesmo sob uma variedade de condições que dificultam o reconhecimento, tais como mudanças de escala, rotação, iluminação e pose (BAUER, SÜNDERHAUF e PROTZEL, 2007).

Experimentos comparativos elaborados no trabalho de Juan e Gwon (2009) evidenciaram que o SIFT deve ser preterido ao SURF, para a tarefa de reconhecimento, por ter apresentado as piores de taxas de desempenho na geração dos descritores (cerca de 3 vezes mais lento) e uma taxa de reconhecimento de 78,1% contra 85,7% do SURF, na base de imagens Caltech (CALTECH FACE DATABASE, 2010). Baseado nesse resultado, o descritor SURF foi escolhido como a técnica para extração de características faciais utilizada neste trabalho. Uma ilustração de aplicação deste módulo é apresentada na Figura 3.7 e um detalhamento do descritor SURF é apresentado na próxima subseção.

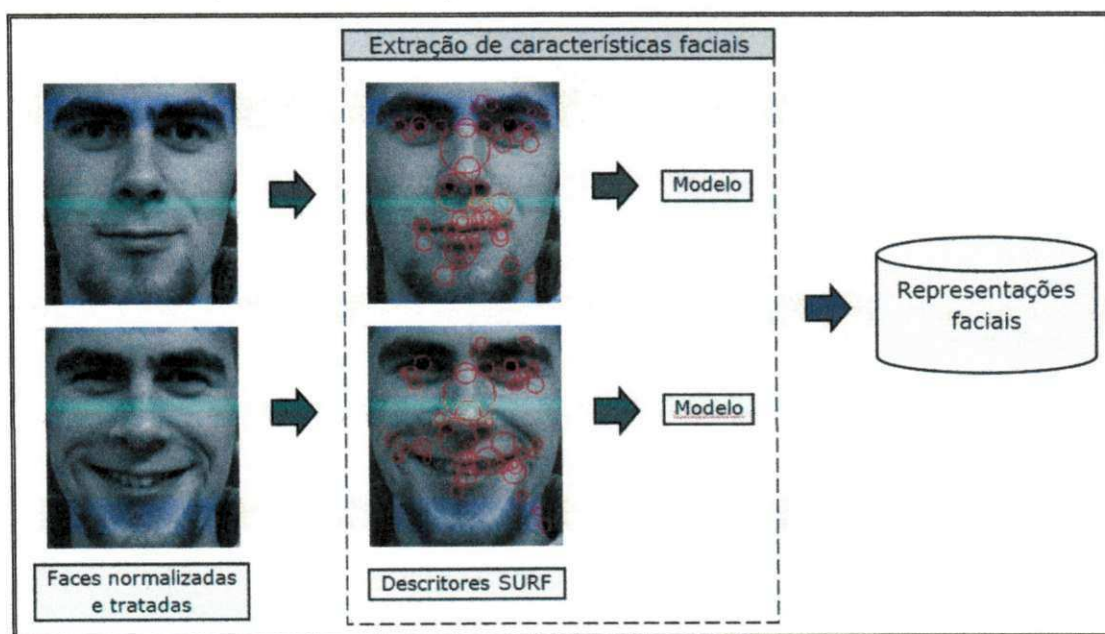


Figura 3.7 – Etapa 3 – Módulo de extração de características faciais.

3.1.1.1. Speeded Up Robust Features (SURF)

Nesta seção, o detector-descritor SURF é analisado do ponto de vista teórico para fornecer uma visão detalhada de seu funcionamento. Grande parte do aumento de desempenho em comparação ao SIFT pode ser atribuído ao uso de uma representação de imagem intermediária, conhecida como *Imagem Integral* (VIOLA e JONES, 2001).

Imagem Integral

A imagem integral é rapidamente calculada a partir de uma imagem de entrada e é utilizada para acelerar o cálculo da área de regiões retangulares nela contidas. Dada uma imagem de entrada $I(x,y)$, cada pixel $I_x(i,j)$ da imagem integral I_x é calculado como o somatório dos valores de intensidade dos pixels de I , da origem até ao pixel $I(i,j)$, ou seja:

$$I_{\Sigma}(x,y) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i,j) \quad (3.6)$$

Usando a imagem integral, a tarefa de calcular a área de uma região retangular é reduzida a operações de soma e subtração. Ao considerar o retângulo delimitado por vértices A, B, C e D (vide Figura 3.8), a soma das intensidades dos pixels daquela região na imagem original é calculada por:

$$\Sigma = A - B - C + D \quad (3.7)$$

Dado que tempo de computação é invariante à mudança no tamanho da região retangular, esta abordagem é particularmente útil quando grandes áreas são consideradas. SURF faz bom uso dessa propriedade para realizar rápidas convoluções de diferentes filtros de tamanhos variados em tempo próximo de constante.

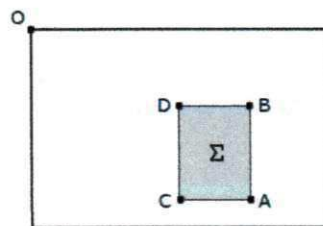


Figura 3.8 – Cálculo de área retangular utilizando imagens integrais.

Matriz Hessiana baseada em Pontos de Interesse

O detector SURF é baseado no determinante da matriz Hessiana. Assumindo uma função contínua de duas variáveis $f(x, y)$, a matriz Hessiana, H , é definida como a matriz das derivadas parciais da função f , de acordo com Laganière (2011):

$$H(f(x, y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \quad (3.8)$$

O determinante desta matriz, conhecido como discriminante, pode ser calculado como:

$$\det(H) = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 \quad (3.9)$$

O valor do discriminante é usado para classificar os máximos e mínimos da função, por meio do teste da derivada de segunda ordem. Uma vez que o determinante é o produto dos autovalores da matriz Hessiana, podem-se classificar os pontos com base no sinal do resultado. Se o determinante for negativo, então os autovalores possuem sinais diferentes e, portanto, o ponto não é um extremo local. Caso seja positivo, então ambos os autovalores são positivos ou ambos são negativos. Assim, em ambos os casos, o ponto é classificado como um extremo.

Traduzir esta teoria para trabalhar com imagens, em vez de uma função contínua, é uma tarefa relativamente simples. Primeiro, substituí-se os valores da função $f(x, y)$ pelas intensidades de pixel da imagem $I(x, y)$. Em seguida, faz-se necessário um método para calcular as derivadas parciais de segunda ordem da imagem, o que é feito pela convolução com um filtro com *kernel* apropriado.

No caso de SURF, a Gaussiana de segunda ordem é o filtro escolhido, uma vez que permite a análise em escala, bem como em espaço. O uso da Gaussiana permite variar a quantidade de suavização durante o estágio de convolução, para que o determinante possa ser calculado em diferentes

escalas. Além disso, como a Gaussiana é uma função isotrópica, ou seja, circularmente simétrica, a convolução com o *kernel* permite a invariância à rotação. Assim, dado um ponto p da imagem I com coordenadas (x,y) , a matriz Hessiana $H(p,\sigma)$ em p na escala σ , pode ser definida como:

$$H(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{xy}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix} \quad (3.10)$$

na qual $L_{xx}(p,\sigma)$ é a aproximação para a derivada Gaussiana de segunda ordem $\frac{\partial^2 g(\sigma)}{\partial x^2}$ da imagem $I(x,y)$ no ponto p e, similarmente, para L_{yy} e L_{xy} , que aproximam as demais derivadas parciais de segunda ordem. Essas derivadas são conhecidas como Laplacianos de Gaussianas (*Laplacian of Gaussians* – LoG).

Bay, Tuytelaars e Van Gool (2006) descobriram em seu trabalho um ganho de desempenho na aproximação do Laplaciano de Gaussianas (LoG) por uma Diferença de Gaussianas (DoG), ao utilizar filtros representados por caixas em seus respectivos *kernels*, conforme pode ser observado na Figura 3.9. Adicionalmente, um aumento considerável no desempenho é verificado, quando esses filtros são utilizados em conjunto com a imagem integral descrita anteriormente.

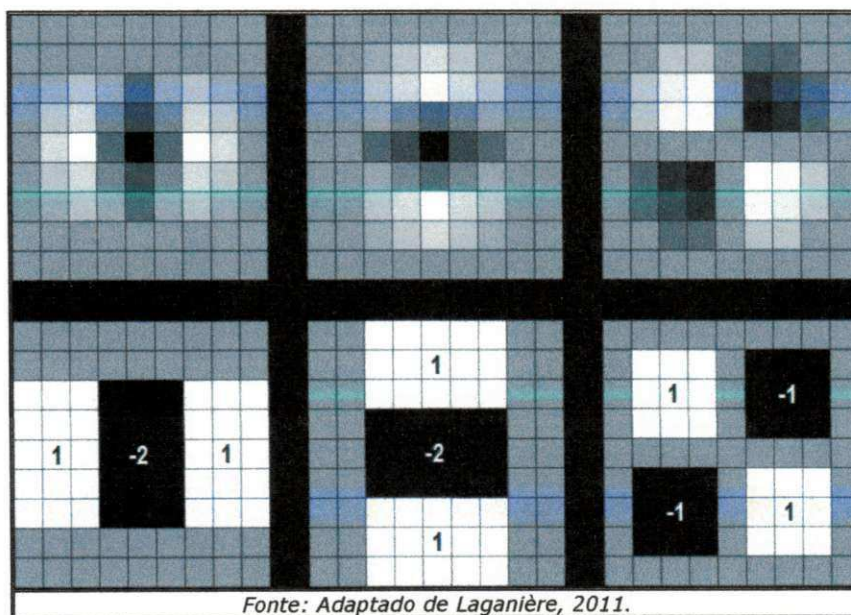


Figura 3.9 – Aproximação LoG. Linha superior: derivadas Gaussianas de segunda ordem, discretizadas e recortadas, nas direções x , y e xy , ou seja, L_{xx} , L_{yy} e L_{xy} , respectivamente. Linha inferior: aproximações de filtros representados por caixas nas direções x , y e xy , ou seja, D_{xx} , D_{yy} e D_{xy} , respectivamente.

Na Figura 3.9, os pesos aplicados às regiões pretas assumem valor -1, enquanto às regiões brancas são atribuídos peso 1 e às regiões cinza peso 0. Desta forma, Bay, Tuytelaars e Van Gool (2006) propuseram a seguinte fórmula, como uma aproximação para o cálculo do determinante da matriz Hessiana usando as Gaussianas aproximadas:

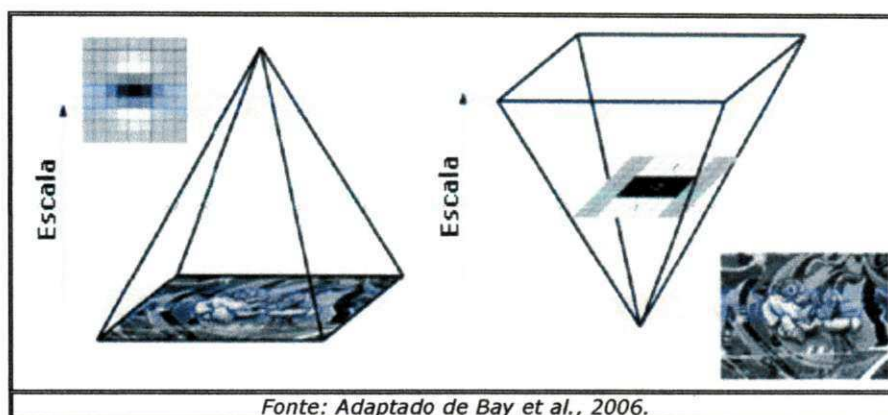
$$\det(\mathbf{H}_{\text{aproximado}}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (3.11)$$

No trabalho de Bay, Tuytelaars e Van Gool (2006), os dois filtros, LoG e DoG, são comparados em detalhe e os resultados indicam que a perda de precisão é praticamente desprezível pelo uso da representação em caixas, sendo largamente compensada pelo aumento considerável na velocidade de processamento. Os máximos locais desta função, em relação a espaço e escala, corresponde aos pontos de interesse da imagem. O método exato para a extração de pontos de interesse é explicado em seqüência.

Representação Espaço-Escala

A fim de detectar pontos de interesse, fazendo uso do determinante da matriz Hessiana, é necessário, primeiramente, introduzir a noção de espaço-escala. Espaço-escala pode ser definido como uma função contínua que pode ser utilizada para encontrar extremos em todas as escalas possíveis (BAY, TUYTELAARS e VAN GOOL, 2006). Na área de VC, espaço-escala pode ser tipicamente implementado como uma imagem-pirâmide, ou seja, a imagem de entrada é iterativamente sub-amostrada em tamanhos menores por meio de repetidas convoluções Gaussianas (BOSWORTH e ACTON, 2003).

Considerando que o tempo de processamento dos *kernels* usados no SURF é constante, o espaço-escala pode ser criado por meio da aplicação de *kernels* de tamanho crescente na imagem original. Isto permite que múltiplas camadas da pirâmide espaço-escala possam ser processadas simultaneamente tornando desnecessária a necessidade de sub-amostragem da imagem, proporcionando aumento de desempenho. Na Figura 3.10, ilustra-se a diferença entre a abordagem tradicional da estrutura espaço-escala e a abordagem utilizada pelo SURF.



Fonte: Adaptado de Bay et al., 2006.

Figura 3.10 – Ao invés de iterativamente reduzir o tamanho da imagem (esquerda), o uso de imagens integrais permite o aumento da escala do filtro em tempo constante (direita).

O espaço-escala é dividido em oitavos (*octaves*). Um oitavo representa uma série de respostas do filtro obtido pela convolução da imagem de entrada com um filtro de tamanho crescente. No SURF, o nível mais baixo da representação espaço-escala é obtido a partir de um filtro 9x9, correspondendo a uma Gaussiana com valor de $\sigma = 1,2$. As camadas subseqüentes são obtidas pelo aumento da escala dos filtros. À medida que o tamanho do filtro aumenta, o mesmo acontece com o valor da escala Gaussiana associadas, uma vez que relações entre os tamanhos permanecem constantes, mantendo as mesmas proporções. Esta escala pode ser calculada de acordo com a equação:

$$\sigma_{\text{aproximado}} = \text{TamanhoFiltro} \frac{1.2}{9} \quad (3.12)$$

Assim, para dois níveis sucessivos, deve-se aumentar este tamanho por um mínimo de 1 pixel em cada lado da região, a fim de manter o tamanho desigual e, assim, garantir a presença do pixel central. Isso corresponde a filtros de tamanho 15x15 e 21x21 (a partir do filtro 9x9) conforme pode ser observado na Figura 3.11.

Localização de Pontos de Interesse

A tarefa de localizar pontos de interesse invariantes a escala e rotação na imagem de entrada pode ser dividida em três etapas. Primeiramente, as respostas dos filtros, obtidas pela convolução da imagem de entrada com filtros de tamanho crescente são limiarizadas, de tal forma que todos os valores abaixo de um limiar pré-determinado são removidas. O aumento do

limiar implica a diminuição do número de pontos de interesse detectados, deixando apenas os mais fortes, enquanto a diminuição do limiar permite muito mais detecções de pontos de interesse. Portanto, o limiar pode ser ajustado para adequar a detecção, de acordo com a necessidade da aplicação.

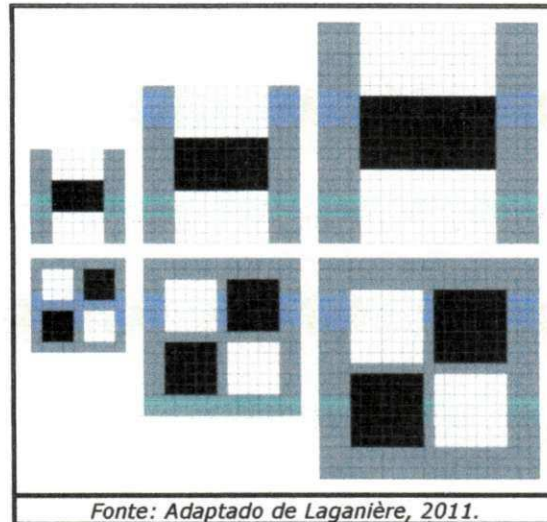


Figura 3.11 – Tamanhos dos filtros subsequentes devem diferir por um mínimo de 6 pixels para preservar a estrutura. Filtros 9x9, 15x15 e 21x21, em seqüência.

Depois da limiarização, uma supressão não-máxima é realizada para encontrar um conjunto de pontos candidatos. Para isto, cada pixel no espaço-escala é comparado com um conjunto de 26 vizinhos, composto de 8 pontos na escala nativa e 9 pontos em cada uma das escalas acima e abaixo de cada uma das escalas, de forma a encontrar o pixel de maior resposta na região, conforme ilustrado na Figura 3.12. Deste modo, obtém-se um conjunto de pontos de interesse com resposta mínima, determinada pelo valor do limiar.

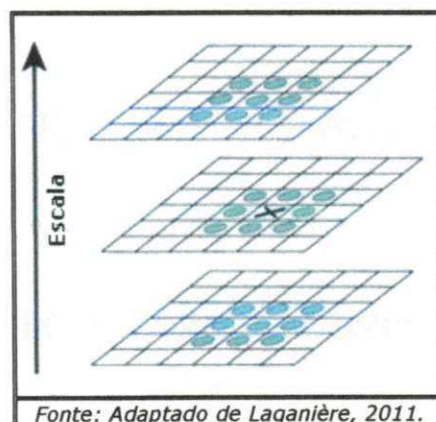


Figura 3.12 – Supressão não-máxima. O pixel marcado com 'X' é selecionado como uma máxima se for maior do que os pixels adjacentes em seu intervalo e nos intervalos acima e abaixo.

A etapa final da localização dos pontos de interesse envolve a interpolação dos dados próximos para encontrar a localização em ambos espaço e escala utilizando uma precisão sub-pixel. Isso é feito pela aplicação de uma função quadrática 3D proposta por Brown e Lowe (2002). Assim, o determinante da função Hessiana, $H(p, \sigma)$, pode ser expresso como uma expansão de Taylor, até os termos quadráticos, centrada na localização detectada:

$$H(x) = H + \frac{\partial H}{\partial x} x + \frac{1}{2} x' \frac{\partial^2 H}{\partial x^2} x \quad (3.13)$$

A localização interpolada do extremo, $\hat{x} = (x, y, \sigma)$, é encontrada tomando a derivada desta função e igualando-a a zero, tal que:

$$\hat{x} = - \frac{\partial H^{-1}}{\partial x^2} \frac{\partial H}{\partial x} \quad (3.14)$$

As derivadas são aproximadas por diferenças finitas de pixels vizinhos. Se \hat{x} é maior do que 0,5 nas direções x , y ou σ , a localização deve ser ajustada e a interpolação deve ser realizada novamente. Este procedimento é repetido, até que \hat{x} se torne inferior a 0,5 em todas as direções ou até que o número predeterminado de etapas de interpolação seja excedido. Aqueles pontos que não convergem são descartados do conjunto de pontos de interesse, deixando apenas aqueles mais estáveis e repetíveis. O conjunto das etapas anteriores compõe o denominado detector Hessiano.

Descritor de Pontos de Interesse

O descritor SURF descreve como as intensidades dos pixels estão distribuídas dentro de uma vizinhança de escala de cada ponto de interesse detectado pelo detector Hessiano. Esta abordagem é semelhante àquela do SIFT. No entanto, as imagens integrais em conjunto com filtros *Wavelets de Haar* (VIOLA e JONES, 2001), são empregadas para aumentar a robustez e diminuir o tempo de processamento. O procedimento para a extração do descritor é apresentado a seguir.

Para atingir a invariância à rotação, a cada ponto de interesse

detectado é atribuída uma orientação reprodutível, conforme explicado a seguir. A extração dos componentes do descritor é realizada em relação a esta direção, por isso é importante que essa direção encontrada seja repetível sob condições variáveis. Para determinar a orientação, *Wavelets de Haar* de tamanho 4σ são calculadas em um conjunto de pixels com raio de 6σ em torno do ponto detectado, sendo σ a escala na qual o ponto foi detectado.

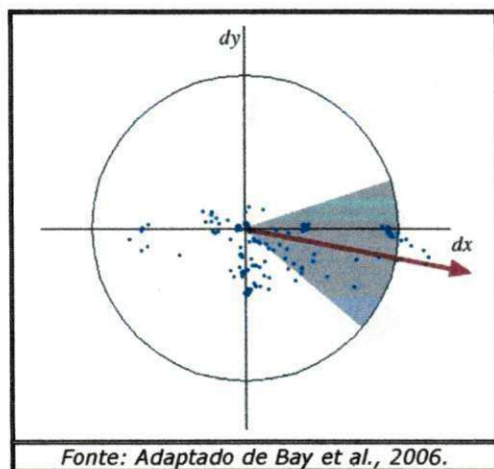


Figura 3.13 – Determinação da orientação: uma janela de orientação deslizante de tamanho $\pi/3$ detecta a orientação dominante das respostas das *wavelets de Haar* em cada ponto da amostra dentro de uma vizinhança circular em torno do ponto de interesse.

As respostas são ponderadas com uma Gaussiana centrada no ponto de interesse. Uma vez ponderadas, as respostas são representadas como pontos de um espaço vetorial bidimensional definido por dx e dy (Figura 3.13). A orientação dominante é selecionada pela rotação de um segmento de círculo cobrindo um ângulo de $\frac{\pi}{3}$ radianos em torno da origem. A cada rotação, os vetores correspondentes aos pontos dentro do segmento são somados, determinando um vetor resultante. O maior vetor resultante determina a orientação final do ponto de interesse, conforme pode ser observado na Figura 3.13.

O segundo passo para a extração do descritor SURF requer a construção de uma janela quadrada ao redor do ponto de interesse. Esta janela contém os pixels a partir dos quais são calculados os componentes do vetor descritor, sendo dividida em subregiões de tamanho 4×4 pixels, conforme ilustrado na Figura 3.14.



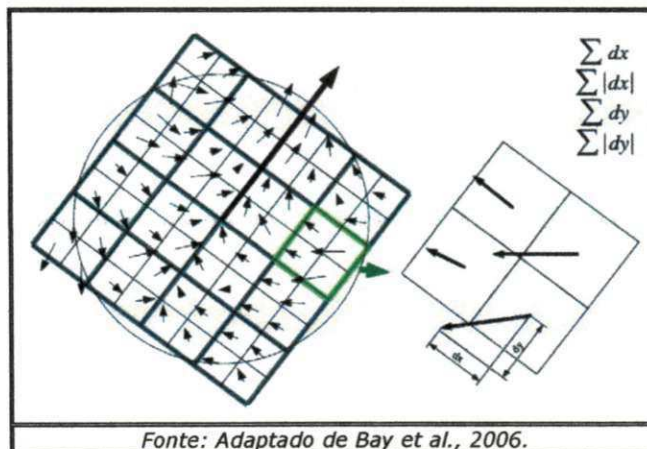
Fonte: Adaptado de Laganière, 2011.

Figura 3.14 – Janelas do descritor. O tamanho da janela do descritor é 20 vezes a dimensão do ponto detectado e está orientada ao longo da direção dominante ilustrada em verde.

Em cada uma das subregiões, *wavelets* de *Haar* de tamanho 2σ são calculadas em 25 pontos da janela distribuídos uniformemente. Ao considerar dx e dy as respostas das *wavelets* em termos de x e y , respectivamente, para os 25 pontos de amostragem (para cada subregião), o vetor de cada subregião, pode ser definido como:

$$V_{\text{região}} = [\sum dx, \sum dy, \sum |dx|, \sum |dy|] \quad (3.15)$$

Portanto, cada subregião contribui com quatro valores para o vetor do descritor, levando a um vetor global de 64 dimensões (vide Figura 3.15). O descritor SURF resultante é invariante à rotação, à escala, ao brilho e, após normalização unitária, ao contraste.



Fonte: Adaptado de Bay et al., 2006.

Figura 3.15 – Para construir o descritor, uma grade orientada quadrática com 4x4 sub-regiões é colocada sobre o ponto de interesse (esquerda). Para cada subregião, as respostas *Wavelets* são computadas a partir de amostras de 5x5. Para cada subregião, são computadas as somas dx , $|dx|$, dy , e $|dy|$, calculadas relativamente à orientação da grade (direita).

Na próxima seção, será detalhado o módulo de *comparação e determinação de similaridade*, para a determinação da semelhança entre as representações faciais e geração da matriz de similaridade. Em seguida, o módulo de *técnica de agrupamento*, objetivando a composição dos grupos de faces semelhantes utilizando um agrupamento aglomerativo hierárquico, será descrito.

3.3.2. Comparação e Determinação de Similaridade

De posse das representações faciais extraídas por meio dos descritores SURF de cada imagem de face normalizada e tratada, o módulo de *comparação e determinação de similaridade* é responsável pela determinação da semelhança do conjunto de todos possíveis pares de faces a serem agrupadas. Este procedimento é baseado no algoritmo proposto por Muja e Loew (2009), denominado *Fast Approximate Nearest Neighbors – FANN*, para determinar o grau de similaridade entre duas faces em função da correspondência entre descritores e inspirado na abordagem proposta por Antonopoulos, Nikolaidis e Pitas (2007), a qual é utilizada para mapear esta correspondência em uma matriz de similaridade. Na Figura 3.16, ilustra-se o detalhamento funcional deste módulo.

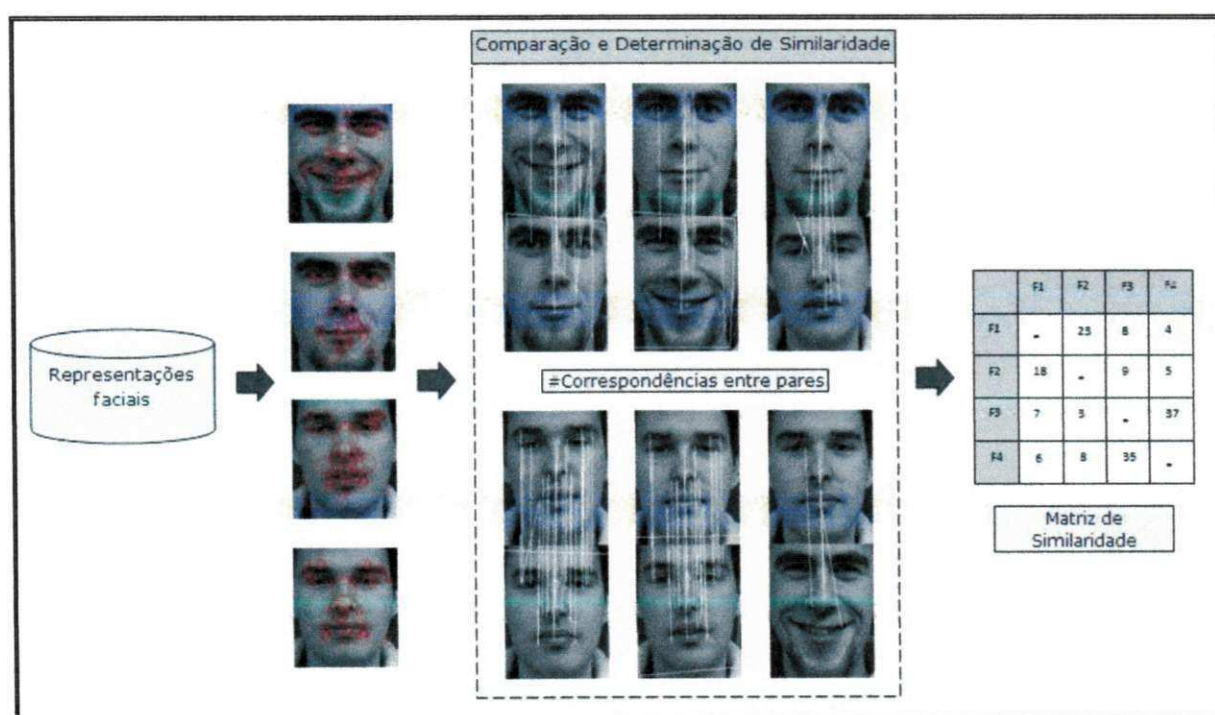


Figura 3.16 – Etapa 4 – Módulo de comparação e determinação de similaridade.

O problema do vizinho mais próximo pode ser formulado como segue. Dado um conjunto P de n pontos em um espaço métrico X , pré-processar P , tal que, para um ponto de consulta $q \in X$, pode-se encontrar (rapidamente) o ponto $n_q \in P$ mais próximo de q . A busca de vizinhos mais próximos é uma tarefa fundamental utilizado em vários domínios, incluindo aprendizagem de máquina, o agrupamento de objetos, a recuperação de documentos, bancos de dados, estatísticas, dentre outros (HAR-PELED e KUMAR, 2010).

Em algumas aplicações, pode ser aceitável a recuperação de um "bom palpite" do vizinho mais próximo. Nessas situações, pode-se usar um algoritmo que não garanta encontrar o vizinho mais próximo real em todos os casos, em troca de maior eficiência em termos de velocidade ou memória. Geralmente, tal algoritmo vai encontrar o vizinho mais próximo, na maioria dos casos, mas isso depende muito do conjunto de dados que está sendo consultado. Neste contexto, está inserido o FANN, algoritmo que aplica busca de prioridade em árvores *K-Means* hierárquica (*Hierarchical K-Means Tree*), proposto por Muja e Loew (2009).

A árvore *K-Means* hierárquica é construída dividindo-se os pontos de dados de cada nível em K regiões distintas usando um agrupamento *K-Means* e, em seguida, aplicando-se o mesmo método recursivamente para os pontos em cada região. A recursão é parada quando o número de pontos em uma região for menor do que K . O algoritmo FANN explora a árvore *K-Means* hierárquica da forma *best-bin-first* (com base em árvores *kd*), ou seja, retorna o vizinho mais próximo para uma grande fração de consultas e um vizinho muito próximo para os demais casos (HAR-PELED e KUMAR, 2010).

O algoritmo realiza inicialmente um único percurso pela árvore, adicionando uma fila de prioridades a todos os ramos inexplorados em cada nó ao longo do caminho. Em seguida, extrai da fila de prioridades o ramo que possui o centro mais próximo ao ponto de consulta e reinicia a travessia da árvore a partir daquele ramo. Em cada passagem, o algoritmo continua adicionando à fila de prioridades os ramos inexplorados ao longo do caminho. O grau de aproximação é especificado da mesma forma como é

definido para árvores *kd* randomizadas, parando a pesquisa logo após um número predeterminado de nós folha (pontos de dados) serem examinados (MUJA e LOEW, 2009).

O algoritmo FANN tem como saída o número de correspondências entre os pontos de interesse (*keypoints*), calculados pelo descritor SURF do par de descritores comparados. A partir destas correspondências, histogramas de cinza são extraídos dos pontos de interesse que compõem cada correspondência e ponderados por sua interseção, conforme a métrica proposta por Swain e Ballard (1991):

$$d(H_p, H_q) = \frac{\sum_i \min(H_p^i, H_q^i)}{\sum_i H_p^i} \quad (3.16)$$

na qual H_p e H_q são os histogramas de cinza dos pontos de interesse (*keypoints*) de cada uma das correspondências. Essa função assume um valor real que varia de 0, para situações em que os histogramas são totalmente diferentes; até 1, quando idênticos.

Devido ao fato de que a correspondência entre os descritores A e B não produz o mesmo resultado da correspondência entre B e A e considerando que a matriz de similaridade deve ser simétrica, faz-se necessário o cálculo da correspondência duas vezes para cada par de descritores, ou seja, uma vez para o par (A, B) e outra vez para o par (B, A). Assim, o número máximo de correspondências encontradas entre (A, B) e (B, A) é considerado como a correspondência final para este par de descritores. Desta forma, a matriz de similaridade pode ser definida de acordo com a seguinte função de similaridade proposta por Antonopoulos, Nikolaidis e Pitas (2007):

$$S(A, B) = S(B, A) = 100 \left(1 - \frac{M_{AB}}{\min(K_A, K_B)} \right), \quad (3.17)$$

em que M_{AB} é o número máximo de correspondências entre (A, B) e (B, A) e K_A e K_B são o número de pontos de interesse dos descritores A e B, respectivamente. Essa função assume valores no intervalo $[0,100]$, em que

valores próximos de 0 indicam grande similaridade entre os descritores. A matriz de similaridade computada servirá de entrada para a etapa final de processamento do sistema proposto, o módulo de agrupamento de faces.

3.3.3. Técnica de Agrupamento

A partir da matriz de similaridade, computada pelo módulo anterior, o módulo final do sistema proposto, denominado *Agrupamento de Faces*, objetiva a composição dos grupos de faces semelhantes utilizando um agrupamento aglomerativo hierárquico de modo que e os pares de faces sejam arranjados de acordo com os respectivos graus de similaridade, agrupados segundo uma disposição hierárquica. Na Figura 3.17, ilustra-se o detalhamento funcional do módulo em questão.

Os métodos hierárquicos são técnicas simples onde os dados são particionados sucessivamente, produzindo uma representação hierárquica dos agrupamentos (EVERITT, LANDAU e MORVEN, 2001). Essa representação facilita visualizar a formação dos agrupamentos em cada estágio no qual ela ocorreu, assim como o grau de semelhança entre eles. Os métodos hierárquicos não requerem que seja definido um número *a priori* de agrupamentos. A partir da análise do dendograma (diagrama que mostra a hierarquia e a relação dos agrupamentos em uma estrutura), pode-se inferir o número de agrupamentos adequados.

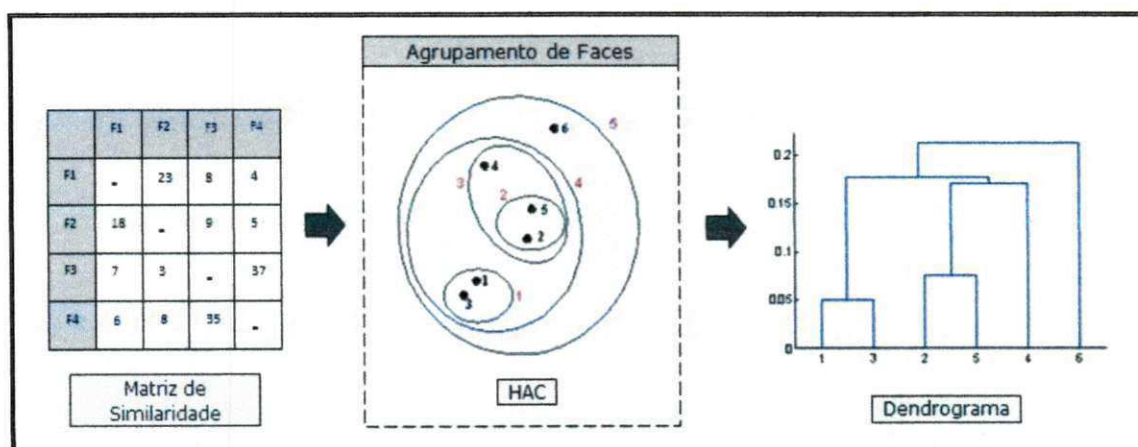


Figura 3.17 – Etapa 5 – Módulo de agrupamento de faces.

Nesse método, inicia-se com cada padrão formando seu próprio agrupamento e, gradualmente, os grupos são unidos, até que um único agrupamento contendo todos os dados seja gerado. Logo no início do

processo, os agrupamentos são pequenos e os elementos de cada grupo possuem um alto grau de dissimilaridade. Ao final do processo, têm-se poucos agrupamentos, cada um podendo conter muitos elementos e mais similares entre si (SILVA, 2005).

Depois de criada a matriz de similaridade, o próximo passo é encontrar o menor valor da matriz. Esse valor identifica os dois agrupamentos mais similares entre si. Feito isso, esses dois agrupamentos identificados são agrupados, formando assim um novo agrupamento. Logo em seguida, a matriz de similaridade é atualizada, contendo agora um agrupamento a menos. Esse procedimento é repetido, até restar apenas um único agrupamento. De acordo com Matteucci (2011), o procedimento geral pode ser formalizado como segue:

- (1) Inicialmente cada agrupamento contém um único padrão;
- (2) Calcula-se/atualiza-se a matriz de similaridade;
- (3) Forma-se um novo agrupamento pela união dos agrupamentos com maior grau de similaridade; e
- (4) Os passos 2 e 3 são executados (N-1) vezes, até que todos os objetos estejam em um único agrupamento.

Diversos algoritmos hierárquicos foram propostos, dentro os quais destacam-se: (i) Agrupamento por Ligação Simples; (ii) Agrupamento por Ligação Completa; (iii) Agrupamento por Centróide; e (iv) Agrupamento de Ward (MATTEUCCI, 2011). No entanto, o algoritmo adotado neste trabalho foi o Agrupamento por Média Aritmética Ponderada (*Unweighted Pair-Group Method using Arithmetic Average* – UPGMA) que produziu melhores resultados em comparação aos anteriores nos experimentos realizados (Capítulo 4).

O agrupamento UPGMA é baseado na decisão de fundir dois agrupamentos em um cálculo que envolve as similaridades (ou proximidades) entre todos os objetos de ambos os agrupamentos analisados. Neste caso, a maior similaridade é encontrada pelo cálculo da média aritmética de todas as distâncias entre os objetos de um dos

agrupamentos em relação, aos objetos do outro agrupamento (JAIN e DUBES, 1988), conforme a equação:

$$d(A, B) = \frac{1}{|A||B|} \sum_{x \in A} \sum_{y \in B} \text{dist}(x, y) \quad (3.18)$$

Os agrupamentos que apresentarem o resultado de maior similaridade (ou proximidade) são agrupados. Na próxima seção, serão descritos as ferramentas e linguagens que deram suporte ao desenvolvimento da abordagem proposta neste capítulo.

3.4. Detalhes de Projeto e Implementação

Para a validação da abordagem proposta, foi necessário desenvolver um sistema de agrupamento de faces, seguindo as etapas e o fluxo de execução apresentado e discutido anteriormente. Nesta seção, são apresentados detalhes de projeto e implementação da aplicação desenvolvida.

A implementação foi realizada utilizando a linguagem C++, a partir do ambiente de desenvolvimento integrado (*Integrated Development Environment - IDE*) *Microsoft Visual Studio 2005* e do sistema de *software* estatístico *XLStat* (XLSTAT, 2011). A linguagem C++ foi escolhida, dentre vários aspectos, por: (i) apresentar um desempenho superior para a tarefa, em relação a outras linguagens; (ii) ser muito utilizada na área de Visão Computacional; e (iii) dispor de bibliotecas tais como a Intel OpenCV, dotada de facilidades destinadas ao suporte no desenvolvimento do trabalho.

Há, no entanto, algumas desvantagens no uso da linguagem e do ambiente escolhidos, tais como a dependência em relação ao sistema operacional *Microsoft Windows*, problemas de gerenciamento de memória e apontadores, além de peculiaridades do ambiente de desenvolvimento. O ambiente de desenvolvimento *Microsoft Visual Studio 2005* foi escolhido pela disponibilidade de utilização, a partir do programa *MSDN Academic Alliance* e pela familiaridade do desenvolvedor com suas funcionalidades.

A aplicação de *software* desenvolvida foi estruturada em três grupos

de classes, a saber: (i) classes relacionadas ao controle do processo, responsáveis por operações de leitura e escrita em disco e controle do fluxo de execução do processo; (ii) classes relacionadas ao pré-processamento e à fase de representação da imagem; e (iii) classes relacionadas à fase de agrupamento de faces. As classes que fazem parte da aplicação desenvolvida são exibidas na Figura 3.18 e a separação entre os grupos funcionais destas classes é ilustrada pelas cores verde (controle), amarela (representação da imagem) e azul (agrupamento de faces).

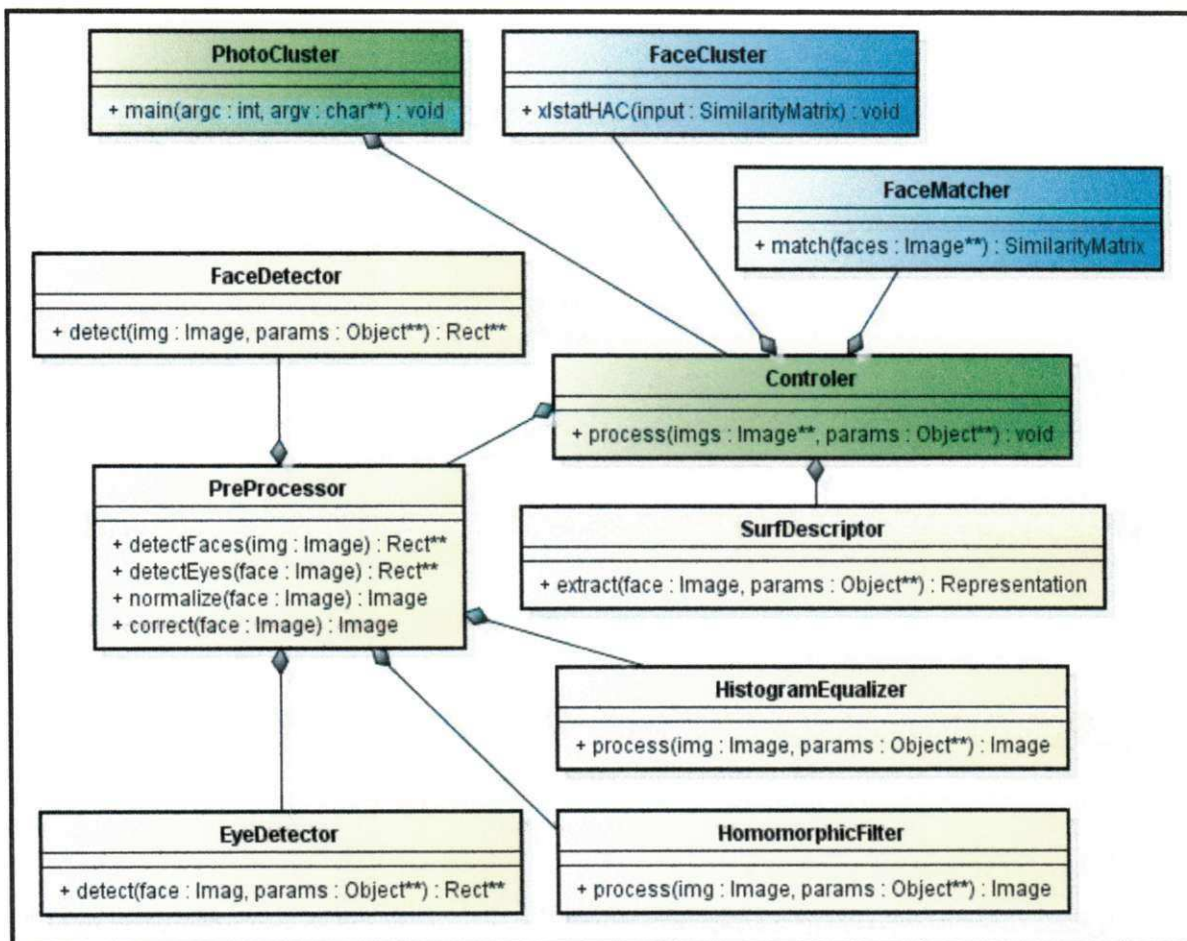


Figura 3.18 – Diagrama de classes da aplicação desenvolvida.

3.5. Considerações Finais

Neste capítulo, foi proposta uma abordagem para o agrupamento de faces em coleções de fotografias digitais, a qual é dividida nas seguintes etapas: (i) detecção e correção da orientação de faces; (ii) equalização e compensação da iluminação; (iii) extração de características faciais; (iv) comparação e determinação de similaridade; e (v) agrupamento.

Na etapa inicial, a imagem de entrada é submetida ao módulo de detecção, o qual produz como saída recortes de faces normalizadas que passarão por um processo de melhoria da iluminação, mediante uma filtragem homomórfica e uma equalização de histograma. Após o processo de melhoramento, a partir de cada face em questão, descritores SURF são extraídos transformando as imagens de face em representações faciais na forma de pontos de interesse. Em seguida, ocorre uma comparação em termos de similaridade entre todos os pares de representações faciais gerando uma matriz de similaridade. Por fim, a partir da matriz de similaridade das representações faciais, a técnica de agrupamento é aplicada para a geração final dos grupos de faces semelhantes.

O próximo capítulo contém uma descrição dos experimentos realizados e discussão dos resultados obtidos. A apresentação dos experimentos contempla os testes realizados nos vários caminhos disponíveis até que fosse atingida a abordagem descrita nesse capítulo. O próximo capítulo também contém uma análise comparativa entre a abordagem proposta e três ferramentas concorrentes por meio de métricas objetivas com outras ferramentas concorrentes de agrupamento de faces.

Capítulo 4

Avaliação Experimental

Neste capítulo, são descritos e comentados os experimentos realizados para validação da abordagem proposta para a solução do problema de agrupamento de faces em coleções de fotografias digitais, apresentada no Capítulo 3. A abordagem proposta é avaliada, inicialmente, por meio de testes visuais, a partir da confrontação gráfica de seus resultados em um processo de calibração. Posteriormente, os resultados são analisados estatisticamente, utilizando os testes ANOVA e *Tukey*, sendo comparados com outros resultados provenientes de abordagens concorrentes. Vários experimentos foram conduzidos em duas etapas distintas do estudo, sendo cada uma dessas etapas descrita em seções separadas, a saber:

- (i) Na Seção 4.2, são apresentados resultados de comparação objetiva, usando as sete métricas numéricas de avaliação de agrupamento, apresentadas no Capítulo 2, dos enfoques estudados para a montagem da abordagem proposta, visando determinar a melhor combinação de técnicas para o agrupamento de faces. Este procedimento é denominado *Experimento de Calibração*.
- (ii) Na Seção 4.4, relata-se a realização de um experimento comparativo entre a abordagem de melhor desempenho determinada no *Experimento de Calibração* e três ferramentas concorrentes. Este experimento é denominado *Experimento de Comparação*. Em ambos os experimentos, a avaliação dos resultados obtidos foi realizada a partir dos testes visuais e estatísticos, mencionados anteriormente.

As demais seções deste capítulo estão estruturadas como segue. Na

Seção 4.1, descreve-se a base de imagens utilizada nos experimentos realizados, bem como, a organização das imagens em cada um dos experimentos. As ferramentas concorrentes avaliadas no experimento comparativo são detalhadas na Seção 4.3. Por fim, na Seção 4.5, os resultados obtidos são analisados e discutidos na forma de considerações finais.

4.1. Base de Imagens

A escolha da base de imagens para a solução do problema em questão foi uma tarefa relativamente dispendiosa, devido a diversos fatores, dentre os quais destacam-se: (i) a grande variedade de bases empregadas nos trabalhos analisados no Capítulo 2; (ii) a não disponibilidade pública de algumas bases utilizadas em trabalhos presentes na revisão bibliográfica; (iii) a complexidade do procedimento para obter acesso a algumas bases (e.g., assinatura de termos de compromisso e confidencialidade, necessidade de envio de mídia digital por correio). A base LFW (HUANG et al., 2007) que é uma base pública e utilizada em 25% do trabalhos analisados, não apresenta grande quantidade de amostras de uma mesma pessoa, apenas cerca de 2 amostras para cada indivíduo, em média. Devido a este fato, a base LFW apresenta baixa variação de expressões faciais, iluminação e pose para um mesmo indivíduo, o que faz com que a mesma não seja a mais adequada para o problema objeto de estudo.

Diante deste contexto, a base de imagens adotada nos experimentos realizados foi a IMM (STEGMANN, ERSBOLL e LARSEN, 2003), primeiramente, por ser uma base pública e de fácil acesso, e segundo, por apresentar 6 amostras para cada um dos indivíduos presentes nas imagens. A base de imagens IMM engloba 240 imagens de faces de 40 indivíduos, sendo 7 mulheres e 33 homens. As imagens foram capturadas em Janeiro de 2001 com resolução de 640 x 480 pixels usando o formato JPEG. As seguintes variações podem ser encontradas nesta base de imagens: pose (frontal ou rotacionada fora do plano da imagem em aproximadamente 30 graus), expressão facial (alegre, neutra ou arbitrária) e iluminação (difusa ou direcional). As Figuras de 4.1A a 4.1H ilustram o conteúdo desta base.

Para os experimentos descritos nas Seções 4.2 e 4.4, as 240 imagens de faces foram sequencialmente organizadas em 8 subconjuntos de 30 imagens, de acordo com a numeração das mesmas. Desta forma, os experimentos puderam ser realizados com diferentes repetições e com uma amostra de imagens estatisticamente significativa (JAIN, 1991).

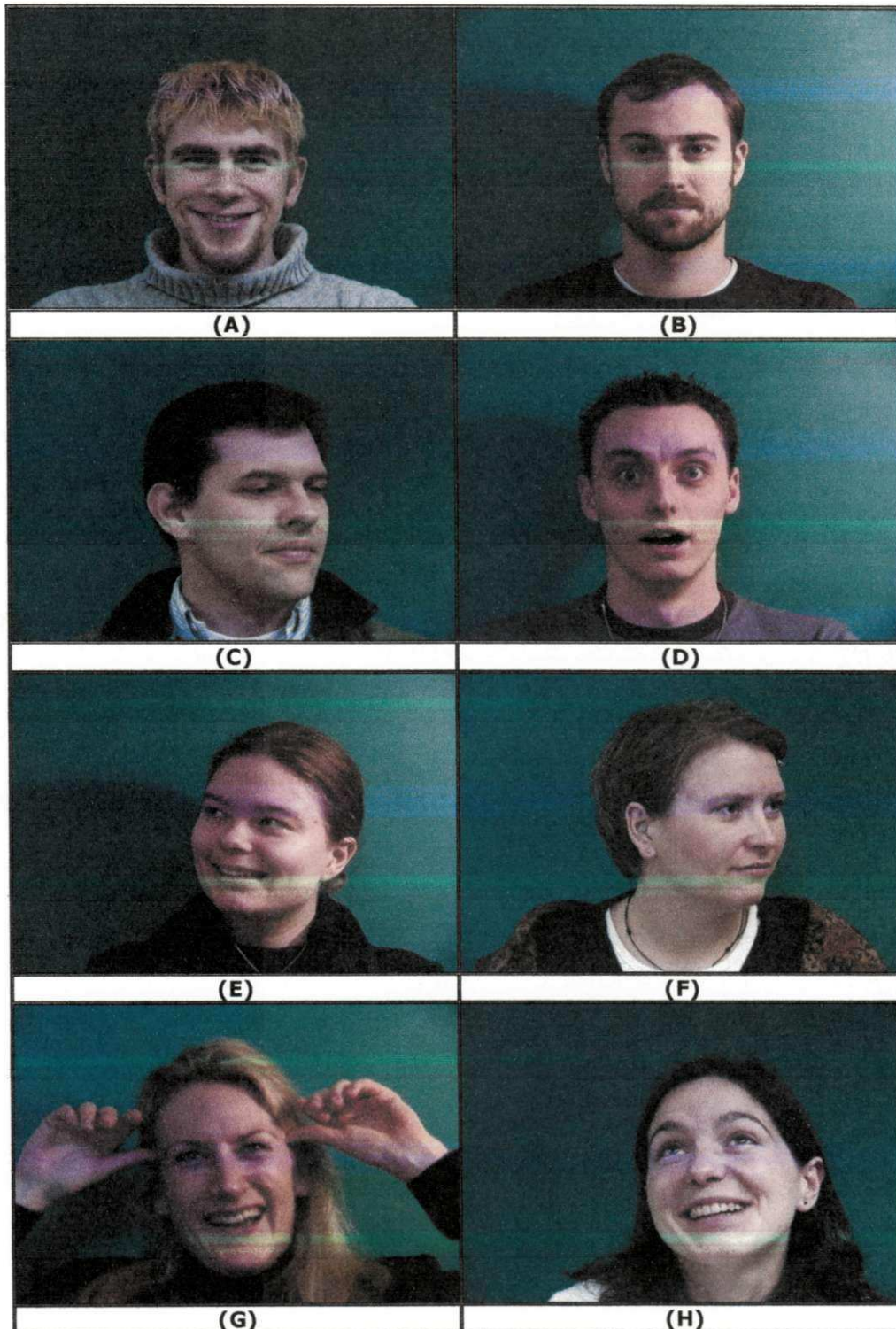


Figura 4.1 – Amostra de imagens da base de faces IMM: variações de gênero, iluminação, expressão facial e pose.

4.2. Experimento de Calibração

Nesta seção, são apresentados resultados comparativos (usando cada uma das sete métricas descritas no Capítulo 2) envolvendo oito estratégias estudadas para a construção da abordagem proposta, visando determinar a melhor combinação de algoritmos/técnicas para a geração dos grupos de faces semelhantes. As estratégias investigadas são sucintamente descritas como:

- (i) **Histograma**: corresponde à abordagem final detalhada no Capítulo 3, composta do módulo de pré-processamento, da característica SURF, da função de similaridade proposta por Antonopoulos, Nikolaidis e Pitas (2007) com ponderação por histograma de cinza e do algoritmo HAC;
- (ii) **LBP**: mesma configuração da abordagem Histograma, exceto pelo tipo de histograma utilizado na ponderação do módulo de *comparação e determinação de similaridade*, no qual foram utilizados histogramas LBP;
- (iii) **Máximo**: mesma configuração da abordagem Histograma, exceto pela estratégia de ponderação por meio de histogramas desenvolvida, utilizando apenas a função de similaridade proposta por Antonopoulos, Nikolaidis e Pitas (2007);
- (iv) **Norm(C)**: alteração na estratégia de ponderação com o uso da correlação normalizada entre os descritores, pela conversão de valores no intervalo $[-1,1]$ para o intervalo $[0,1]$;
- (v) **Média**: alteração na função de similaridade proposta por Antonopoulos, Nikolaidis e Pitas (2007), com o uso da média das correspondências entre os descritores;
- (vi) **Correlação(C)**: alteração na estratégia de ponderação com o uso da correlação entre os descritores, que assume valores no intervalo $[-1,1]$;
- (vii) **Abs(C)**: alteração na estratégia de ponderação com o uso do

valor absoluto da correlação entre os descritores; e

- (viii) **Mínimo**: alteração na função de similaridade proposta por Antonopoulos, Nikolaidis e Pitas (2007), com o uso do mínimo das correspondências entre os descritores.

A comparação objetiva realizada serviu, primeiramente, para fundamentar a tomada de diferentes decisões, tal como, por exemplo, qual o tipo de algoritmo de pré-processamento mais adequado ou qual a estratégia de comparação e determinação de similaridade mais adequada para a técnica de agrupamento empregada. Em segundo, lugar, a avaliação objetiva foi usada para comparar a abordagem proposta com outras ferramentas concorrentes de agrupamento de faces.

Conforme afirmado por Jain (1991), a comparação simples das médias dos resultados pode levar a conclusões não confiáveis. Devido a este fato, após a medição dos valores numéricos de cada métrica, foi realizada uma análise estatística dos resultados.

O propósito da análise estatística foi investigar evidências de variações estatisticamente significativas nos grupos considerados, tendo em vista tratar-se de um experimento ao longo do qual várias categorias de algoritmos foram consideradas.

Inicialmente, verificou-se a normalidade dos dados obtidos por meio de teste de *Jarque-Bera* (JAIN, 1991), seguido de um teste visual de Intervalo de Confiança (IC) conhecido por *boxplot* (JAIN, 1991). Posteriormente, foram realizados testes F ANOVA de fator único (LEVINE, BERENSON e STEPHAN, 2000) com grau de confiança de 95% para cada uma das sete métricas, pois há apenas um único fator a ser testado: o valor resultante de cada métrica em relação ao *ground-truth*.

Uma vez que este teste evidencia apenas diferenças entre médias, não possibilitando comparações entre pares de grupos, utilizou-se, por fim, o teste de *Tukey* (LEVINE, BERENSON e STEPHAN, 2000) para tal finalidade. Todos os testes de hipóteses adotados são detalhados no Apêndice B desta dissertação.

Assim, a hipótese nula de nenhuma diferença nas médias, definida como:

$$H_0 : \mu_{A_1} = \mu_{A_2} = \mu_{A_3} = \dots = \mu_{A_k} \quad (4.1)$$

foi testada, em relação à hipótese alternativa de haver diferenças entre as referidas médias (H_1 : nem todas as médias são iguais). A média μ_{A_i} significa a média dos valores obtidos da i -ésima abordagem investigada (A_i), para cada uma das sete métricas. Os resultados numéricos obtidos de cada métrica para cada uma das abordagens investigadas encontram-se nas Tabelas C.2 a C.9 (ver Apêndice C).

4.2.1. Teste de Normalidade e Teste Visual

Testes de normalidade são usados para determinar se um conjunto de dados de uma dada variável aleatória pode ser modelado por uma distribuição normal (JAIN, 1991). Este tipo de teste se faz necessário devido ao fato do Teste F ANOVA ser um teste paramétrico, no qual a variável de interesse deve seguir uma distribuição normal (LEVINE, BERENSON e STEPHAN, 2000).

Desta forma, o teste de normalidade pode ser definido pelas hipóteses:

$$\begin{cases} H_0: A \text{ amostra segue uma distribuição Normal;} \\ H_1: A \text{ amostra não segue uma distribuição Normal.} \end{cases} \quad (4.2)$$

Nas Tabelas 4.1 a 4.7, são exibidos os resultados do teste de normalidade *Jarque-Bera* (JAIN, 1991), obtidos com o auxílio do sistema de *software* estatístico *XLStat* (XLSTAT, 2011), para os dados associados à calibração da abordagem proposta para cada uma das sete métricas consideradas.

Conforme explicitado nessas tabelas, os valores **p-valor** são maiores do que o nível de significância α , indicando que não se rejeita a hipótese nula H_0 , o que significa que todas as amostras obtidas seguem uma distribuição Normal.

Tabela 4.1 – Teste de normalidade Jarque-Bera para a métrica RI das abordagens investigadas.

| Teste de normalidade Jarque-Bera para a métrica RI | | | | $\alpha=0,05$ |
|---|------------------------|----------------------|-----------|----------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| Histograma | 1,1823 | 5,9914 | 2 | 0,5536 |
| LBP | 0,6656 | 5,9914 | 2 | 0,7168 |
| Máximo | 2,3757 | 5,9914 | 2 | 0,3048 |
| Norm(C) | 1,8342 | 5,9914 | 2 | 0,3996 |
| Média | 0,7644 | 5,9914 | 2 | 0,6823 |
| Correlação | 1,4591 | 5,9914 | 2 | 0,4821 |
| Abs(C) | 2,1574 | 5,9914 | 2 | 0,3400 |
| Mínimo | 1,7180 | 5,9914 | 2 | 0,4235 |

Tabela 4.2 – Teste de normalidade Jarque-Bera para a métrica ARI das abordagens investigadas.

| Teste de normalidade Jarque-Bera para a métrica ARI | | | | $\alpha=0,05$ |
|--|------------------------|----------------------|-----------|----------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| Histograma | 1,1914 | 5,9914 | 2 | 0,5511 |
| LBP | 0,6221 | 5,9914 | 2 | 0,7326 |
| Máximo | 2,0592 | 5,9914 | 2 | 0,3571 |
| Norm(C) | 1,5525 | 5,9914 | 2 | 0,4601 |
| Média | 0,5565 | 5,9914 | 2 | 0,7571 |
| Correlação | 1,4591 | 5,9914 | 2 | 0,4821 |
| Abs(C) | 1,1041 | 5,9914 | 2 | 0,5757 |
| Mínimo | 0,7412 | 5,9914 | 2 | 0,6902 |

Tabela 4.3 – Teste de normalidade Jarque-Bera para a métrica P das abordagens investigadas.

| Teste de normalidade Jarque-Bera para a métrica P | | | | $\alpha=0,05$ |
|--|------------------------|----------------------|-----------|----------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| Histograma | 1,6772 | 5,9914 | 2 | 0,4323 |
| LBP | 1,4178 | 5,9914 | 2 | 0,4921 |
| Máximo | 1,4247 | 5,9914 | 2 | 0,4904 |
| Norm(C) | 1,8002 | 5,9914 | 2 | 0,4065 |
| Média | 0,9536 | 5,9914 | 2 | 0,6207 |
| Correlação | 2,0176 | 5,9914 | 2 | 0,3646 |
| Abs(C) | 2,7278 | 5,9914 | 2 | 0,2556 |
| Mínimo | 1,5348 | 5,9914 | 2 | 0,4642 |

Tabela 4.4 – Teste de normalidade Jarque-Bera para a métrica R das abordagens investigadas.

| Teste de normalidade Jarque-Bera para a métrica R | | | | $\alpha=0,05$ |
|--|------------------------|----------------------|-----------|---------------------------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| Histograma | 1,1823 | 5,9914 | 2 | 0,5536 |
| LBP | 0,5936 | 5,9914 | 2 | 0,7431 |
| Máximo | 2,1305 | 5,9914 | 2 | 0,3446 |
| Norm(C) | 1,5104 | 5,9914 | 2 | 0,4699 |
| Média | 0,7912 | 5,9914 | 2 | 0,6732 |
| Correlação | 0,5172 | 5,9914 | 2 | 0,7721 |
| Abs(C) | 0,7974 | 5,9914 | 2 | 0,6711 |
| Mínimo | 1,1138 | 5,9914 | 2 | 0,5729 |

Tabela 4.5 – Teste de normalidade Jarque-Bera para a métrica F das abordagens investigadas.

| Teste de normalidade Jarque-Bera para a métrica F | | | | $\alpha=0,05$ |
|--|------------------------|----------------------|-----------|---------------------------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| Histograma | 1,1943 | 5,9914 | 2 | 0,5503 |
| LBP | 0,6082 | 5,9914 | 2 | 0,7377 |
| Máximo | 1,9647 | 5,9914 | 2 | 0,3744 |
| Norm(C) | 1,4555 | 5,9914 | 2 | 0,4829 |
| Média | 0,5146 | 5,9914 | 2 | 0,7731 |
| Correlação | 0,9445 | 5,9914 | 2 | 0,6235 |
| Abs(C) | 1,1805 | 5,9914 | 2 | 0,5541 |
| Mínimo | 0,6048 | 5,9914 | 2 | 0,7390 |

Tabela 4.6 – Teste de normalidade Jarque-Bera para a métrica JI das abordagens investigadas.

| Teste de normalidade Jarque-Bera para a métrica JI | | | | $\alpha=0,05$ |
|---|------------------------|----------------------|-----------|---------------------------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| Histograma | 1,1823 | 5,9914 | 2 | 0,5536 |
| LBP | 0,6659 | 5,9914 | 2 | 0,7167 |
| Máximo | 2,3709 | 5,9914 | 2 | 0,3056 |
| Norm(C) | 2,0454 | 5,9914 | 2 | 0,3596 |
| Média | 0,7140 | 5,9914 | 2 | 0,6997 |
| Correlação | 0,9605 | 5,9914 | 2 | 0,6186 |
| Abs(C) | 1,1806 | 5,9914 | 2 | 0,5541 |
| Mínimo | 0,5920 | 5,9914 | 2 | 0,7437 |

Tabela 4.7 – Teste de normalidade Jarque-Bera para a métrica FM das abordagens investigadas.

| Teste de normalidade Jarque-Bera para a métrica FM | | | | $\alpha=0,05$ |
|--|-----------------|---------------|----|---------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| Histograma | 1,1879 | 5,9914 | 2 | 0,5521 |
| LBP | 0,7281 | 5,9914 | 2 | 0,6948 |
| Máximo | 2,1373 | 5,9914 | 2 | 0,3434 |
| Norm(C) | 1,5301 | 5,9914 | 2 | 0,4652 |
| Média | 0,6539 | 5,9914 | 2 | 0,7211 |
| Correlação | 1,0287 | 5,9914 | 2 | 0,5978 |
| Abs(C) | 1,1217 | 5,9914 | 2 | 0,5707 |
| Mínimo | 0,6142 | 5,9914 | 2 | 0,7355 |

Testes visuais são utilizados para verificar graficamente o comportamento dos dados obtidos, i.e, avaliar a simetria dos dados, sua dispersão e a existência ou não de *outliers* nos mesmos, sendo especialmente adequado para a comparação de dois ou mais conjuntos de dados correspondentes às categorias de uma variável qualitativa (LEVINE, BERENSON e STEPHAN, 2000). A representação gráfica de dados mais utilizada em testes visuais é conhecida por *boxplot* (JAIN, 1991). Na Figura 4.2, são exibidos os gráficos *boxplot*, para os dados associados à calibração da abordagem proposta para cada uma das sete métricas consideradas. Uma descrição sucinta da construção e da utilização dos gráficos *boxplot* encontra-se no Apêndice B.

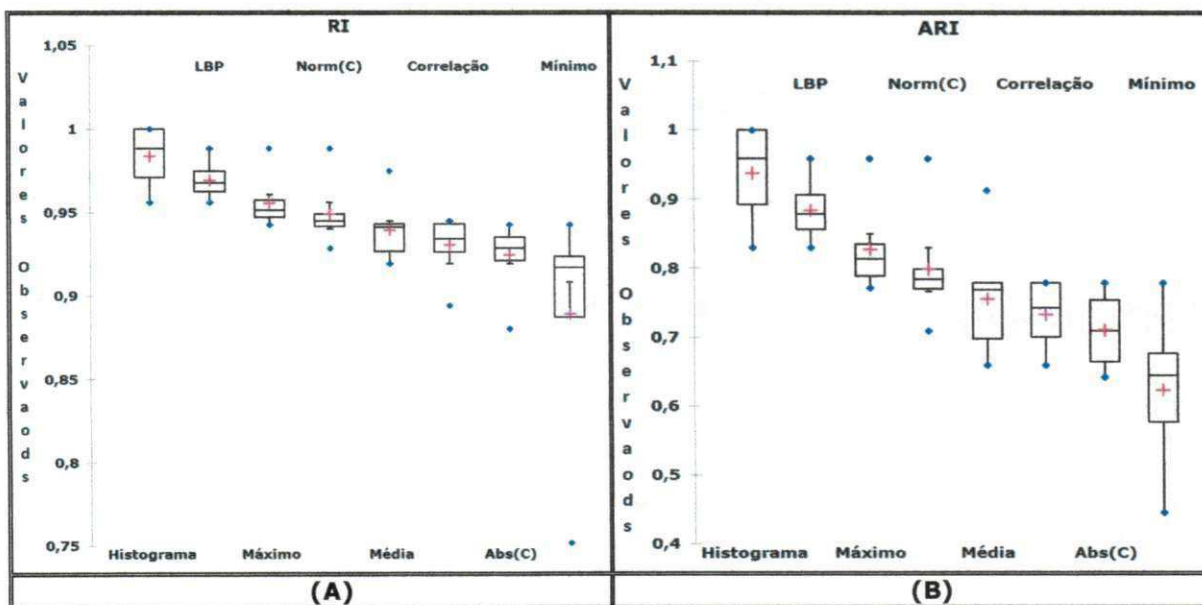


Figura 4.2 – Gráficos *boxplot* das abordagens investigadas: (A) RI; (B) ARI; (C) P; (D) R; (E) F; (F) JI; e (G) FM.

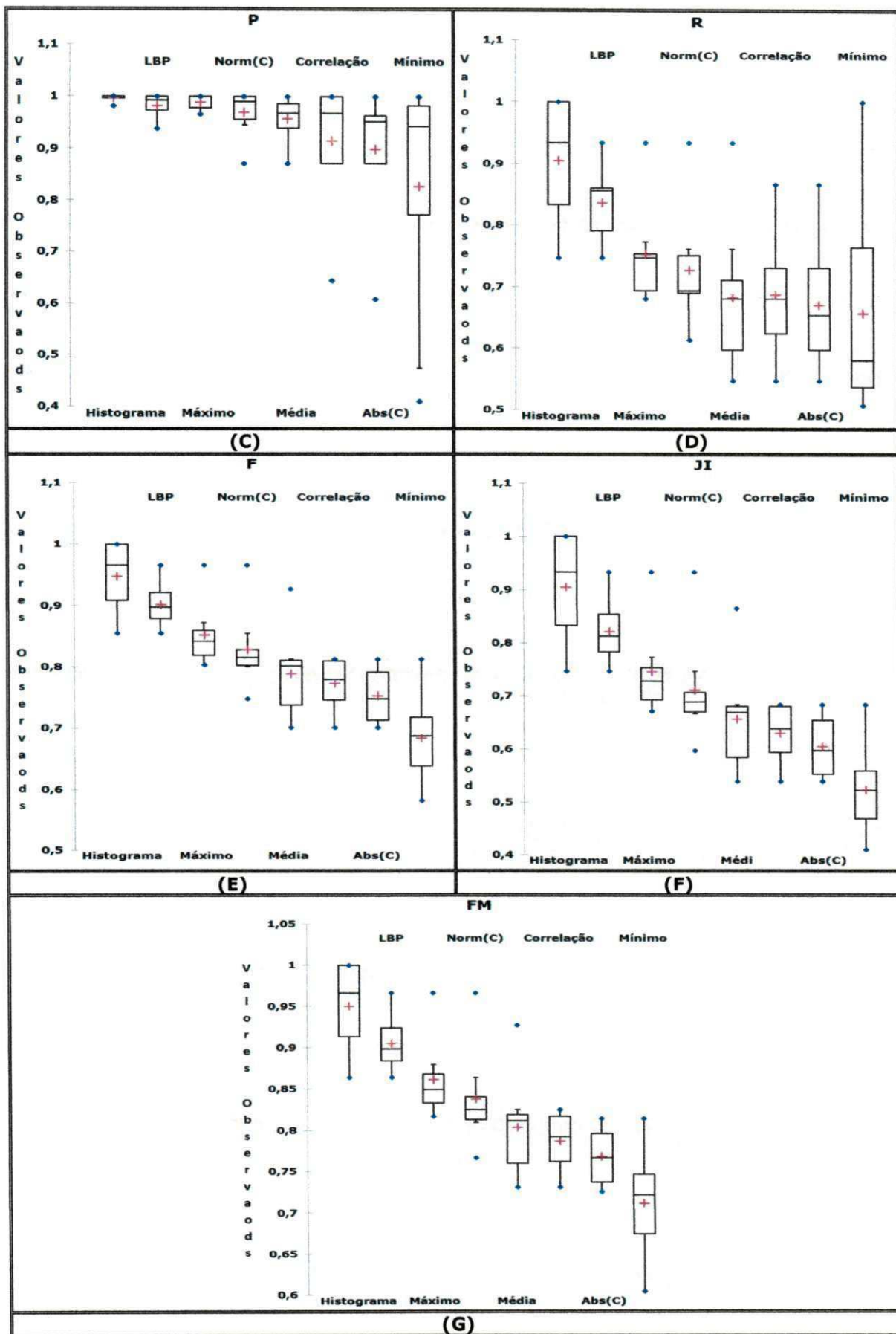


Figura 4.2 – Gráficos *boxplot* das abordagens investigadas: (A) RI; (B) ARI; (C) P; (D) R; (E) F; (F) JI; e (G) FM (continuação).

Analisando os gráficos *boxplot* da Figura 4.2, foi verificado que para todas as métricas, os respectivos IC de cada abordagem investigada possuem sobreposição e que a média de um está contida no IC do outro e vice-versa. Portanto, fez-se necessária a realização de um teste capaz de mensurar numericamente se existem variações estatisticamente significativas nas abordagens investigadas. Para atingir tal objetivo, foram realizados testes F ANOVA de fator único seguidos de testes de *Tukey* para definição da melhor abordagem a ser utilizada.

4.2.2. Teste F ANOVA e Teste de Tukey

Diante do exposto anteriormente, foram conduzidos testes F ANOVA de fator único com grau de confiança de 95%, seguidos de testes de *Tukey* para cada uma das sete métricas de avaliação de agrupamento, com a intenção de verificar a existência de variação significativa ao longo dos grupos de abordagens investigadas. Na Tabela 4.8, são exibidos os resultados do teste F ANOVA fator único, obtidos com o auxílio da aplicação estatística *XLStat* (XLSTAT, 2011), para os dados associados à calibração de cada uma das abordagens investigadas.

Tabela 4.8 – Teste F ANOVA para as sete métricas de avaliação das abordagens investigadas.

| F ANOVA de fator único das abordagens investigadas | | | | | | $\alpha=0,05$ |
|--|-------------------|----|--------|--------|----------------|----------------------|
| Métrica | Fonte de Variação | gl | SQ | MQ | F | F _{crítico} |
| RI | Intergrupos | 7 | 0,0476 | 0,0068 | 8,7220 | 2,2490 |
| | Intragrupos | 56 | 0,0437 | 0,0007 | | |
| | Total | 63 | 0,0913 | | | |
| ARI | Intergrupos | 7 | 0,5657 | 0,0808 | 16,6896 | 2,2490 |
| | Intragrupos | 56 | 0,2711 | 0,0048 | | |
| | Total | 63 | 0,8368 | | | |
| P | Intergrupos | 7 | 0,1872 | 0,0267 | 2,2626 | 2,2490 |
| | Intragrupos | 56 | 0,6619 | 0,0118 | | |
| | Total | 63 | 0,8491 | | | |
| R | Intergrupos | 7 | 0,4389 | 0,0627 | 5,1246 | 2,2490 |
| | Intragrupos | 56 | 0,6852 | 0,0122 | | |
| | Total | 63 | 1,1241 | | | |

Tabela 4.8 – Teste F ANOVA para as sete métricas de avaliação das abordagens investigadas (continuação).

| F ANOVA de fator único das abordagens investigadas | | | | | | $\alpha=0,05$ |
|--|-------------------|----|--------|--------|----------------|----------------------|
| Métrica | Fonte de Variação | gl | SQ | MQ | F | F _{crítico} |
| F | Intergrupos | 7 | 0,4033 | 0,0576 | 17,7562 | 2,2490 |
| | Intragrupos | 56 | 0,1817 | 0,0032 | | |
| | Total | 63 | 0,5850 | | | |
| JI | Intergrupos | 7 | 0,8494 | 0,1213 | 17,5212 | 2,2490 |
| | Intragrupos | 56 | 0,3878 | 0,0069 | | |
| | Total | 63 | 1,2372 | | | |
| FM | Intergrupos | 7 | 0,3317 | 0,0473 | 18,4734 | 2,2490 |
| | Intragrupos | 56 | 0,1436 | 0,0025 | | |
| | Total | 63 | 0,4753 | | | |

Conforme elucidado pelos resultados dos testes F ANOVA fator único na Tabela 4.8, os valores de F são maiores do que os valores de F_{crítico}. Portanto, há diferenças estatisticamente significativas entre as médias dos vários caminhos propostos para todas as sete métricas de avaliação de agrupamento calculadas.

Entretanto, o teste F ANOVA fator único evidencia apenas a existência de diferenças entre médias, não possibilitando comparações entre pares de grupos. Para melhor inferir o conjunto de algoritmos mais adequados a serem usados na abordagem proposta, foram conduzidos testes de *Tukey* com um IC de 95% para múltiplas comparações de médias. Os resultados dos testes de *Tukey* obtidos são apresentados nas Tabelas C.10 a C.16 situadas no Apêndice C. Nos Quadros 4.1 a 4.7, são apresentados os resultados dos grupos de similaridades obtidos por meio dos testes de *Tukey* realizados.

Pelos resultados dos grupos obtidos a no teste de *Tukey*, conforme exibidos no Quadro 4.1, pode ser verificado que **não há**, estatisticamente, diferença significativa entre as abordagens Histograma, LBP, Máximo e Norm(C) para a métrica RI, apesar da abordagem Histograma ter apresentado a maior média.

Quadro 4.1 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica RI das abordagens investigadas.

| Classificação do Teste de Tukey para a métrica RI | | | | | |
|---|-------|--------|---|---|---|
| Abordagem | Média | Grupos | | | |
| Histograma | 0,984 | A | | | |
| LBP | 0,969 | A | B | | |
| Máximo | 0,956 | A | B | C | |
| Norm(C) | 0,949 | A | B | C | |
| Média | 0,939 | | B | C | |
| Correlação | 0,930 | | B | C | D |
| Abs(C) | 0,924 | | | C | D |
| Mínimo | 0,889 | | | | D |

Analisando os resultados dos grupos obtidos a partir do teste de *Tukey* para a métrica ARI, exibidos no Quadro 4.2, pode ser verificado que **não há**, estatisticamente, diferença significativa entre as abordagens Histograma e LBP apesar de a abordagem Histograma ter apresentado a melhor média em comparação com as demais abordagens.

Quadro 4.2 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica ARI das abordagens investigadas.

| Classificação do Teste de Tukey para a métrica ARI | | | | | |
|--|-------|--------|---|---|-----|
| Abordagem | Média | Grupos | | | |
| Histograma | 0,938 | A | | | |
| LBP | 0,883 | A | B | | |
| Máximo | 0,827 | | B | C | |
| Norm(C) | 0,799 | | B | C | D |
| Média | 0,755 | | | C | D |
| Correlação | 0,733 | | | C | D |
| Abs(C) | 0,710 | | | | D E |
| Mínimo | 0,622 | | | | E |

No Quadro 4.3, pode ser verificado que **não há**, estatisticamente, diferença significativa entre nenhum das abordagens investigadas para a métrica P, demonstrando que todas as abordagens foram consideradas similares para esta métrica. Pode-se observar também que as médias das abordagens estão mais próximas entre si do que as médias evidenciadas

anteriormente.

Quadro 4.3 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica P das abordagens investigadas.

| Classificação do Teste de Tukey para a métrica P | | |
|--|-------|--------|
| Abordagem | Média | Grupos |
| Histograma | 0,996 | A |
| LBP | 0,989 | A |
| Máximo | 0,981 | A |
| Norm(C) | 0,969 | A |
| Média | 0,957 | A |
| Correlação | 0,915 | A |
| Abs(C) | 0,898 | A |
| Mínimo | 0,827 | A |

Pelos resultados dos grupos obtidos com o teste de *Tukey* para a métrica R, conforme exibidos no Quadro 4.4, pode ser verificado que **há**, estatisticamente, diferença significativa entre a abordagem Histograma e as demais, além do fato de sua média ser consideravelmente superior a todas as outras médias das abordagens investigadas, evidenciando que esta abordagem apresentou melhores resultados do que as demais, para a métrica em questão.

Quadro 4.4 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica R das abordagens investigadas.

| Classificação do Teste de Tukey para a métrica R | | | | |
|--|-------|--------|---|---|
| Abordagem | Média | Grupos | | |
| Histograma | 0,905 | A | | |
| LBP | 0,836 | | B | |
| Máximo | 0,752 | | B | C |
| Norm(C) | 0,727 | | B | C |
| Média | 0,687 | | B | C |
| Correlação | 0,682 | | B | C |
| Abs(C) | 0,670 | | B | C |
| Mínimo | 0,656 | | | C |

Similarmente ao resultado ilustrado no Quadro 4.2, os resultados dos

grupos obtidos pelo teste de *Tukey* para a métrica F, exibidos no Quadro 4.5, evidenciam que **não há**, estatisticamente, diferença significativa entre as abordagens Histograma e LBP, apesar de a abordagem Histograma ter apresentado a melhor média em comparação com as demais abordagens.

Quadro 4.5 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica F das abordagens investigadas.

| Classificação do Teste de Tukey para a métrica F | | | | | | |
|--|-------|--------|---|---|---|---|
| Abordagem | Média | Grupos | | | | |
| Histograma | 0,947 | A | | | | |
| LBP | 0,901 | A | B | | | |
| Máximo | 0,852 | | B | C | | |
| Norm(C) | 0,828 | | B | C | D | |
| Média | 0,788 | | | C | D | |
| Correlação | 0,772 | | | C | D | E |
| Abs(C) | 0,752 | | | | D | E |
| Mínimo | 0,683 | | | | | E |

Semelhantemente ao resultado apresentado no Quadro 4.4, os resultados dos grupos obtidos pelo teste de *Tukey* para a métrica JI, exibidos no Quadro 4.6, evidenciam que **há**, estatisticamente, diferença significativa entre a abordagem Histograma e as demais, evidenciando que está abordagem apresentou melhores resultados do que as demais para as métricas R e JI.

Quadro 4.6 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica JI das abordagens investigadas.

| Classificação do Teste de Tukey para a métrica JI | | | | | | |
|---|-------|--------|---|---|---|---|
| Abordagem | Média | Grupos | | | | |
| Histograma | 0,905 | A | | | | |
| LBP | 0,822 | | B | | | |
| Máximo | 0,746 | | B | C | | |
| Norm(C) | 0,711 | | B | C | D | |
| Média | 0,656 | | | C | D | |
| Correlação | 0,630 | | | C | D | E |
| Abs(C) | 0,605 | | | | D | E |
| Mínimo | 0,523 | | | | | E |

Analisando os resultados exibidos no Quadro 4.7, pode ser verificado que **não há**, estatisticamente, diferença significativa entre as abordagens Histograma e LBP, pelo teste de *Tukey* para a métrica FM, apesar de a abordagem Histograma ter apresentado a maior média.

Quadro 4.7 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica FM das abordagens investigadas.

| Classificação do Teste de Tukey para a métrica FM | | | | | | |
|---|-------|--------|---|---|---|---|
| Abordagem | Média | Grupos | | | | |
| Histograma | 0,950 | A | | | | |
| LBP | 0,905 | A | B | | | |
| Máximo | 0,861 | | B | C | | |
| Norm(C) | 0,838 | | B | C | D | |
| Média | 0,804 | | | C | D | |
| Correlação | 0,786 | | | C | D | E |
| Abs(C) | 0,768 | | | | D | E |
| Mínimo | 0,889 | | | | | E |

De acordo com os resultados dos grupos de similaridade obtidos, foi verificado, estatisticamente, que em 100% dos testes de *Tukey*, a abordagem Histograma apresentou a melhor média; em 28,57% (2 de 7) dos testes esta abordagem apresentou o melhor resultado isoladamente; e em 71,42% (5 de 7) dos testes a mesma apresentou o melhor resultado isoladamente e/ou em conjunto com a abordagem LBP, segunda melhor abordagem classificada. Assim, torna-se evidente a escolha da abordagem Histograma como o arcabouço a ser utilizado no experimento comparativo com ferramentas concorrentes a ser descrito em seqüência.

4.3. Experimento de Comparação

Nesta seção, são apresentados resultados comparativos entre o sistema proposto e implementado neste trabalho, denominado *PhotoCluster* e três ferramentas concorrentes, avaliados por cada uma das sete métricas utilizadas anteriormente. As ferramentas utilizadas no experimento em questão são:

- (i) **Adobe Photoshop Elements:** ferramenta proprietária

desenvolvida pela empresa *Adobe*, é um programa de edição de imagens *raster* dirigido a entusiastas da fotografia, portanto, não tendo muitas características úteis em um ambiente de produção de impressão. A ferramenta permite aos usuários, criar, editar, organizar e compartilhar suas imagens e fotografias (PHOTOSHOPEL, 2010);

- (ii) **Google Picasa**: ferramenta grátis desenvolvida pela empresa *Google* é um programa organizador de coleções de fotografias digitais cuja principal função é facilitar a procura por fotografias. Ao ser iniciada a ferramenta é realizada uma localização automática de todas as fotos e classificação das mesmas em álbuns visuais organizados pelas faces detectadas (PICASA, 2010);
- (iii) **PittPatt FaceSort**: ferramenta proprietária desenvolvida pelo projeto *Pittsburgh Pattern Recognition (PittPatt)*, auxilia o usuário na organização e pesquisa de suas fotos e vídeos, realiza reconhecimento de faces em fotos e vídeos, além de detectar faces de semi-perfil. Foi desenvolvida com uma tecnologia capaz de agrupar automaticamente faces em grupos de um mesmo assunto (FACESORT, 2010); e
- (iv) **PhotoCluster**: sistema desenvolvido neste trabalho, com emprego da abordagem Histograma (detalhada no Capítulo 3), que apresentou os melhores resultados no experimento de calibração.

De uma maneira geral, apesar de uma grande quantidade de trabalhos serem relacionados com o problema em questão, optou-se pela utilização de ferramentas concorrentes devido os seguintes fatores:

- (1) **Necessidade de implementação**: dado que existem inúmeras técnicas e algoritmos que poderiam ser empregados para a comparação com a abordagem ora proposta, a implementação de uma certa quantidade dos trabalhos iria demandar um tempo considerável, pois nenhum dos autores forneceu a implementação

de sua abordagem, o que poderia afetar o cronograma planejado;

- (2) **Necessidade de validação:** a partir da uma hipotética implementação dos trabalhos relacionados, testes de validação seriam necessários para verificar a corretude do código, a fim de evitar que resultados incorretos pudessem ser utilizados erroneamente;
- (3) **Crítérios de avaliação diferentes:** partindo do pressuposto de que os fatores 1 e 2 não pudessem ser realizados, uma alternativa seria a utilização de resultados reportados nos trabalhos; no entanto, devido ao fato de que a grande maioria dos trabalhos analisados adotou critérios e métricas de avaliação diferentes, além de inúmeras bases de imagens, isto impossibilitou uma comparação com a abordagem proposta;
- (4) **Utilização prévia:** as três ferramentas consideradas no experimento de comparação também foram utilizadas com sucesso em experimentos preliminares, realizados na proposta desta dissertação, o que favoreceu seu uso, por conta da familiarização com suas funcionalidades;
- (5) **Ferramentas utilizadas na prática:** outro fator decisivo está relacionado ao fato das ferramentas terem sido desenvolvidas exatamente com o propósito de fornecerem suporte à tarefa de agrupamento de faces, além de pertencerem a empresas ou universidades com grande reconhecimento internacional no tema da pesquisa (e.g., *Adobe*, *Google* e a *Universidade de Pittsburgh*). Portanto, o somatório de todos os fatores descritos anteriormente contribuíram para a escolha de ferramentas concorrentes para o experimento comparativo;

As Figuras 4.3 a 4.5 ilustram a tela principal de cada uma das três ferramentas concorrentes utilizadas no experimento, *Adobe Photoshop Elements* (Photoshopel), *PittPatt FaceSort* (FaceSort) e *Google Picasa* (Picasa), respectivamente.

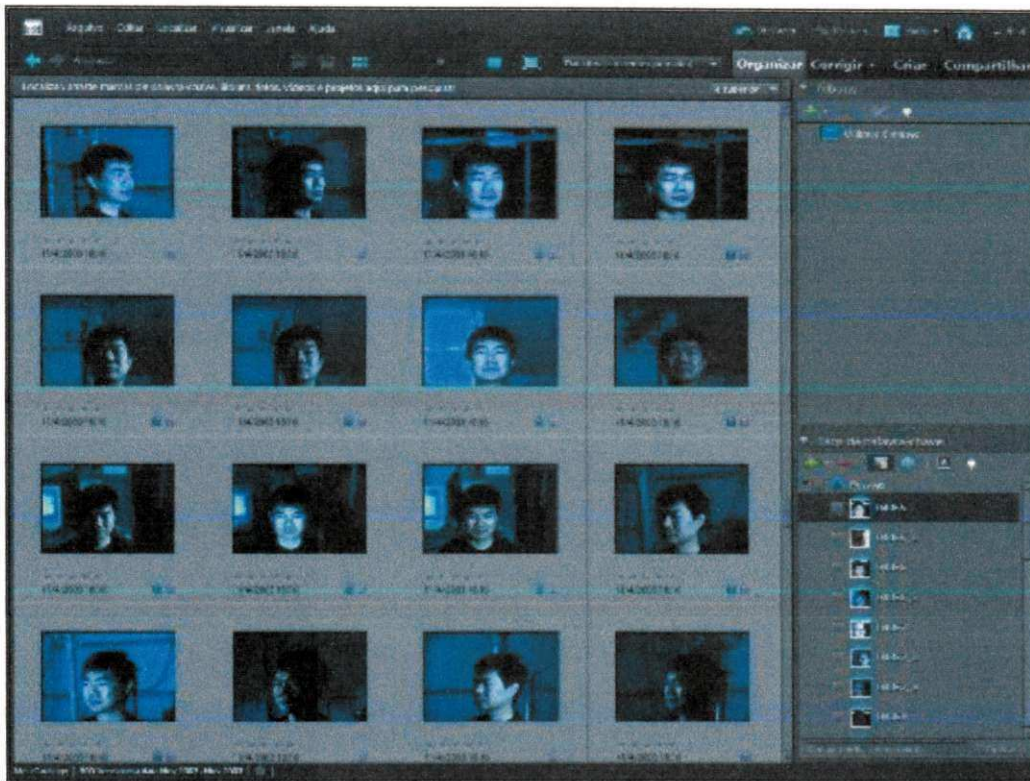


Figura 4.3 – Tela principal do Adobe Photoshop Elements.

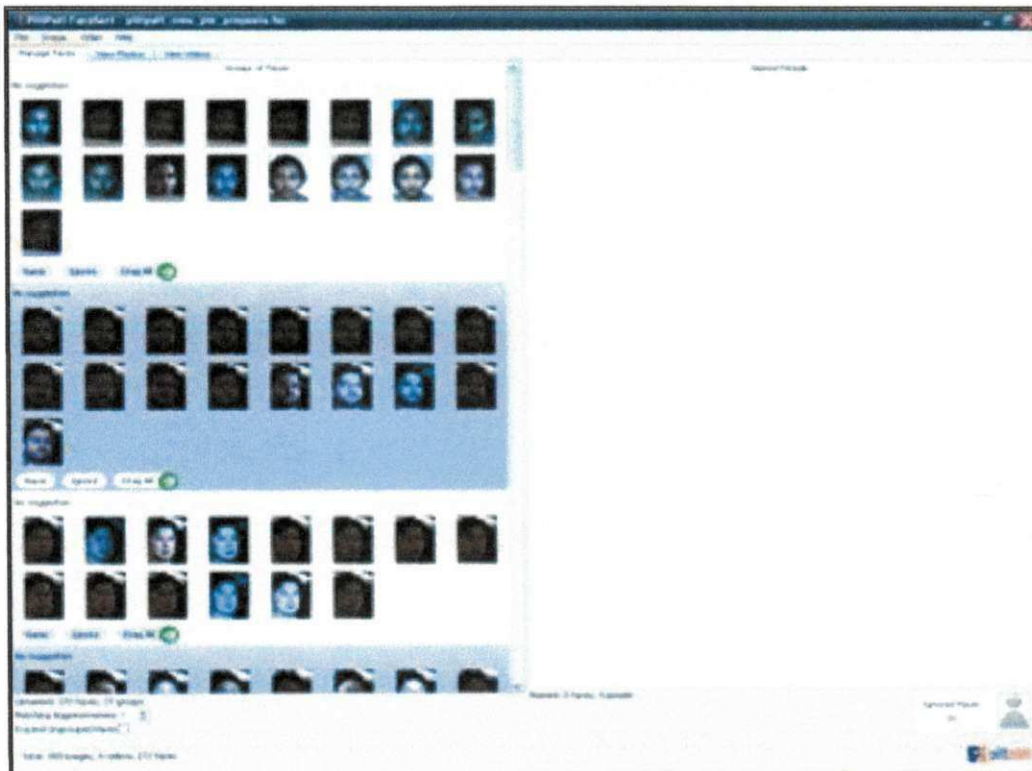


Figura 4.4 – Tela principal do PittPatt FaceSort.

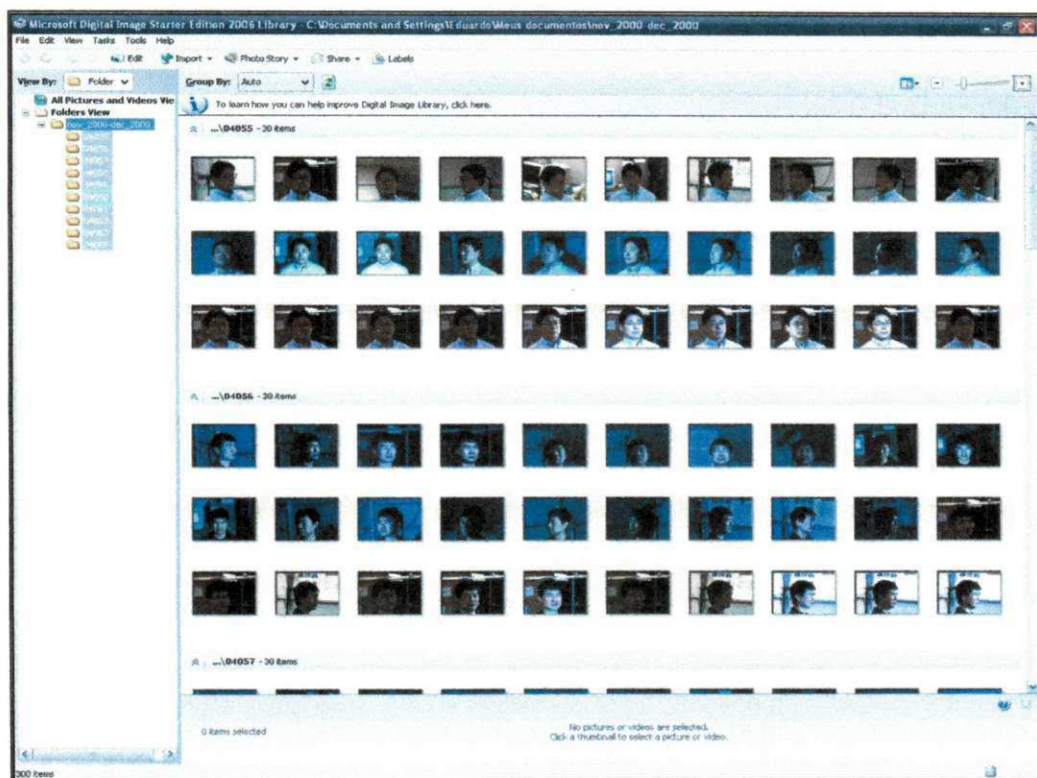


Figura 4.5 –Tela principal do Google Picasa.

Similarmente ao experimento anterior, verificou-se a normalidade dos dados obtidos por meio de teste de normalidade *Jarque-Bera*, seguido de um teste visual de IC, utilizando gráficos *boxplot*. Posteriormente, foram realizados testes F ANOVA de fator único e testes *Tukey* com grau de confiança de 95% para cada uma das sete métricas detalhadas no Capítulo 2. Os resultados numéricos obtidos de cada métrica para cada uma das ferramentas comparadas encontram-se nas Tabelas D.2 a D.5 (Apêndice D).

4.3.1. Teste de Normalidade e Teste Visual

Nas Tabelas de 4.9 a 4.15, são exibidos os resultados do teste de normalidade *Jarque-Bera* (JAIN, 1991), obtidos com o auxílio da aplicação estatística *XLStat* (XLSTAT, 2011), para os dados obtidos por meio da abordagem proposta e por cada uma três ferramentas concorrentes avaliadas. Conforme explicitado nessas tabelas, os valores **p-valor** são maiores do que o nível de significância α , indicando que não se rejeita a hipótese nula H_0 , o que significa que todas as amostras obtidas seguem uma distribuição Normal.

Tabela 4.9 - Teste de normalidade Jarque-Bera para a métrica RI das ferramentas comparadas.

| Teste de normalidade Jarque-Bera para a métrica RI | | | | $\alpha=0,05$ |
|---|------------------------|----------------------|-----------|----------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| FaceSort | 4,1617 | 5,9914 | 2 | 0,1248 |
| Picasa | 0,6574 | 5,9914 | 2 | 0,7198 |
| Photoshopel | 1,0091 | 5,9914 | 2 | 0,6037 |
| PhotoCluster | 1,1823 | 5,9914 | 2 | 0,5536 |

Tabela 4.10 - Teste de normalidade Jarque-Bera para a métrica ARI das ferramentas comparadas.

| Teste de normalidade Jarque-Bera para a métrica ARI | | | | $\alpha=0,05$ |
|--|------------------------|----------------------|-----------|----------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| FaceSort | 1,0029 | 5,9914 | 2 | 0,6056 |
| Picasa | 0,7589 | 5,9914 | 2 | 0,6842 |
| Photoshopel | 1,1523 | 5,9914 | 2 | 0,5620 |
| PhotoCluster | 1,1914 | 5,9914 | 2 | 0,5511 |

Tabela 4.11 - Teste de normalidade Jarque-Bera para a métrica P das ferramentas comparadas.

| Teste de normalidade Jarque-Bera para a métrica P | | | | $\alpha=0,05$ |
|--|------------------------|----------------------|-----------|----------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| FaceSort | 5,3640 | 5,9914 | 2 | 0,0684 |
| Picasa | 1,5824 | 5,9914 | 2 | 0,4532 |
| Photoshopel | 1,3677 | 5,9914 | 2 | 0,5046 |
| PhotoCluster | 1,6772 | 5,9914 | 2 | 0,4323 |

Tabela 4.12 - Teste de normalidade Jarque-Bera para a métrica R das ferramentas comparadas.

| Teste de normalidade Jarque-Bera para a métrica R | | | | $\alpha=0,05$ |
|--|------------------------|----------------------|-----------|----------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| FaceSort | 0,8156 | 5,9914 | 2 | 0,6650 |
| Picasa | 0,6419 | 5,9914 | 2 | 0,7254 |
| Photoshopel | 1,1790 | 5,9914 | 2 | 0,5546 |
| PhotoCluster | 1,1823 | 5,9914 | 2 | 0,5536 |

Tabela 4.13 - Teste de normalidade Jarque-Bera para a métrica F das ferramentas comparadas.

| Teste de normalidade Jarque-Bera para a métrica F | | | | $\alpha=0,05$ |
|--|------------------------|----------------------|-----------|----------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| FaceSort | 0,5304 | 5,9914 | 2 | 0,7670 |
| Picasa | 0,8032 | 5,9914 | 2 | 0,6692 |
| Photoshopel | 1,1769 | 5,9914 | 2 | 0,5551 |
| PhotoCluster | 1,1943 | 5,9914 | 2 | 0,5503 |

Tabela 4.14 – Teste de normalidade Jarque-Bera para a métrica JI das ferramentas comparadas.

| Teste de normalidade Jarque-Bera para a métrica JI | | | | $\alpha=0,05$ |
|--|-----------------|---------------|----|---------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| FaceSort | 0,3118 | 5,9914 | 2 | 0,8556 |
| Picasa | 0,6263 | 5,9914 | 2 | 0,7311 |
| Photoshopel | 1,1888 | 5,9914 | 2 | 0,5518 |
| PhotoCluster | 1,1823 | 5,9914 | 2 | 0,5536 |

Tabela 4.15 – Teste de normalidade Jarque-Bera para a métrica FM das ferramentas comparadas.

| Teste de normalidade Jarque-Bera para a métrica FM | | | | $\alpha=0,05$ |
|--|-----------------|---------------|----|---------------|
| Abordagem | Valor Observado | Valor Crítico | gl | p-valor |
| FaceSort | 0,4025 | 5,9914 | 2 | 0,8176 |
| Picasa | 0,8333 | 5,9914 | 2 | 0,6592 |
| Photoshopel | 1,1265 | 5,9914 | 2 | 0,5693 |
| PhotoCluster | 1,1879 | 5,9914 | 2 | 0,5521 |

Analisando os gráficos *boxplot* da Figura 4.6, para os dados obtidos neste experimento, foi verificado que para todas as métricas, os IC de cada uma das ferramentas investigadas possuem sobreposição e que a média de um está contida no IC do outro e vice-versa. Portanto, fez-se necessária a realização de testes F ANOVA de fator único seguidos de testes de *Tukey* para definição da melhor ferramenta avaliada.

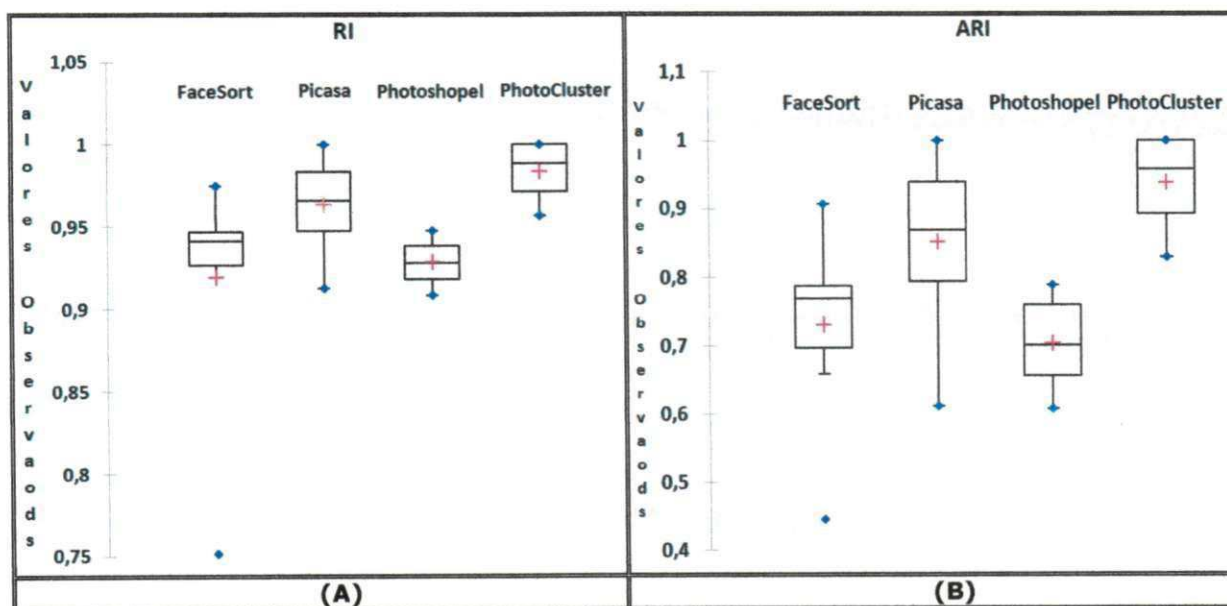


Figura 4.6 – Gráficos *boxplot* das ferramentas analisadas: (A) RI; (B) ARI; (C) P; (D) R; (E) F; (F) JI; e (G) FM.

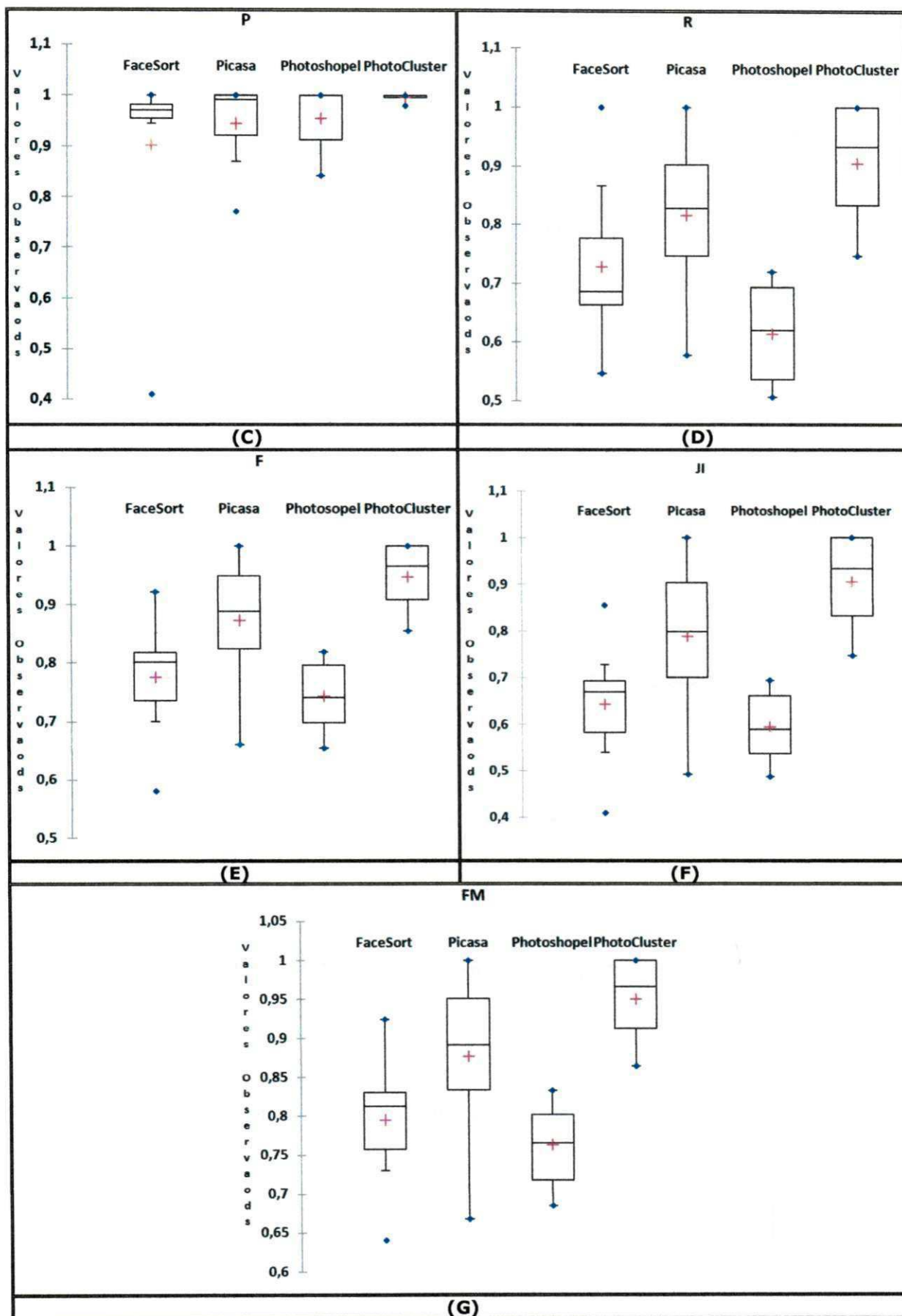


Figura 4.6 – Gráficos *boxplot* das ferramentas analisadas: (A) RI; (B) ARI; (C) P; (D) R; (E) F; (F) JI; e (G) FM (continuação).

4.3.2. Teste F ANOVA e Teste de Tukey

Diante do exposto na seção anterior, foram conduzidos testes F ANOVA de fator único com grau de confiança de 95%, seguidos de testes de *Tukey* para cada uma das sete métricas de avaliação de agrupamento, com a intenção de verificar se havia variação significativa entre as estratégias investigadas. Na Tabela 4.16, são exibidos os resultados do teste F ANOVA fator único, obtidos com o auxílio da aplicação estatística *XLStat* (XLSTAT, 2011), para os dados associados à comparação de cada uma das ferramentas analisadas.

Tabela 4.16 – Teste F ANOVA para as sete métricas de avaliação das ferramentas analisadas.

| F ANOVA de fator único das ferramentas analisadas | | | | | | $\alpha=0,05$ |
|---|-------------------|----|--------|--------|----------------|----------------------|
| Métrica | Fonte de Variação | gl | SQ | MQ | F | F _{crítico} |
| RI | Intergrupos | 3 | 0,0216 | 0,0072 | 4,6864 | 2,9490 |
| | Intragrupos | 28 | 0,0431 | 0,0015 | | |
| | Total | 31 | 0,0647 | | | |
| ARI | Intergrupos | 3 | 0,2842 | 0,0947 | 8,7094 | 2,9490 |
| | Intragrupos | 28 | 0,3046 | 0,0108 | | |
| | Total | 31 | 0,5888 | | | |
| P | Intergrupos | 3 | 0,0350 | 0,0116 | 0,9088 | 2,9490 |
| | Intragrupos | 28 | 0,3598 | 0,0128 | | |
| | Total | 31 | 0,3948 | | | |
| R | Intergrupos | 3 | 0,3721 | 0,1240 | 8,7235 | 2,9490 |
| | Intragrupos | 28 | 0,3981 | 0,0142 | | |
| | Total | 31 | 0,7702 | | | |
| F | Intergrupos | 3 | 0,2089 | 0,0696 | 9,4313 | 2,9490 |
| | Intragrupos | 28 | 0,2067 | 0,0073 | | |
| | Total | 31 | 0,4156 | | | |
| JI | Intergrupos | 3 | 0,4793 | 0,1597 | 10,5258 | 2,9490 |
| | Intragrupos | 28 | 0,4250 | 0,0151 | | |
| | Total | 31 | 0,9043 | | | |
| FM | Intergrupos | 3 | 0,1697 | 0,0565 | 9,2696 | 2,9490 |
| | Intragrupos | 28 | 0,1709 | 0,0061 | | |
| | Total | 31 | 0,3406 | | | |

Conforme os resultados sumariados na Tabela 4.16, os valores de F são maiores do que os valores de $F_{\text{crítico}}$, exceto para a métrica P. Portanto, há diferenças estatisticamente significativas entre as médias das ferramentas avaliadas para seis das sete (85,71%) métricas de avaliação de agrupamento adotadas.

Considerando que o teste F ANOVA fator único evidencia apenas a existência de diferenças entre médias, não possibilitando comparações entre pares de grupos, foram conduzidos testes de *Tukey* com um IC de 95% para múltiplas comparações de médias. Os resultados dos testes de *Tukey* obtidos são apresentados nas Tabelas D.6 a D.12 (ver Apêndice D). A seguir, nos Quadros 4.8 a 4.14, são apresentados os resultados dos grupos de similaridades, obtidos dos testes de *Tukey* realizados.

Pelos resultados dos grupos obtidos do teste de *Tukey* para a métrica RI, conforme exibidos no Quadro 4.8, pode ser verificado que **não há**, estatisticamente, diferença significativa entre as ferramentas *PhotoCluster* e *Picasa* para a métrica RI, apesar da ferramenta *PhotoCluster* ter apresentado a melhor média.

Quadro 4.8 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica RI das ferramentas analisadas.

| Classificação do Teste de Tukey para a métrica RI | | | |
|---|-------|--------|---|
| Abordagem | Média | Grupos | |
| PhotoCluster | 0,984 | A | |
| Picasa | 0,963 | A | B |
| Photoshopel | 0,928 | | B |
| FaceSort | 0,919 | | B |

Analisando os resultados dos grupos obtidos do teste de *Tukey* para a métrica ARI, exibidos no Quadro 4.9, pode ser verificado que **há**, estatisticamente, diferença significativa entre a ferramenta *PhotoCluster* e as demais, além do fato de sua média ser consideravelmente superior a todas as outras médias das ferramentas analisadas, evidenciando que esta ferramenta apresenta melhor desempenho do que as demais para a métrica em questão.

Quadro 4.9 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica ARI das ferramentas analisadas.

| Classificação do Teste de Tukey para a métrica ARI | | | | |
|--|-------|--------|---|---|
| Abordagem | Média | Grupos | | |
| PhotoCluster | 0,938 | A | | |
| Picasa | 0,852 | | B | |
| Photoshopel | 0,731 | | B | C |
| FaceSort | 0,704 | | | C |

No Quadro 4.10, os resultados do teste de *Tukey* para a métrica P mostram que **não há**, estatisticamente, diferença significativa entre nenhum das ferramentas analisadas. Pode-se observar também que as médias das abordagens estão mais próximas entre si do que nos casos anteriores.

Quadro 4.10 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica P das ferramentas analisadas.

| Classificação do Teste de Tukey para a métrica P | | |
|--|-------|--------|
| Abordagem | Média | Grupos |
| PhotoCluster | 0,996 | A |
| Picasa | 0,955 | A |
| Photoshopel | 0,945 | A |
| FaceSort | 0,903 | A |

Os resultados dos grupos obtidos do teste de *Tukey* para a métrica R, exibidos no Quadro 4.11, mostram que **há**, estatisticamente, diferença significativa entre a ferramenta PhotoCluster e as demais. Também observa-se que a média da *PhotoCluster* é consideravelmente superior a todas as outras médias das ferramentas analisadas, evidenciando que esta ferramenta apresenta melhor desempenho do que as demais para a métrica em questão.

Quadro 4.11 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica R das ferramentas analisadas.

| Classificação do Teste de Tukey para a métrica R | | | | |
|--|-------|--------|---|---|
| Abordagem | Média | Grupos | | |
| PhotoCluster | 0,905 | A | | |
| Picasa | 0,816 | | B | |
| Photoshopel | 0,728 | | B | C |
| FaceSort | 0,613 | | | C |

Os resultados dos grupos obtidos do teste de *Tukey* para a métrica *F*, exibidos no Quadro 4.12, evidenciam que **não há**, estatisticamente, diferença significativa entre as ferramentas *PhotoCluster* e *Picasa*, apesar da ferramenta *PhotoCluster* ter apresentado a melhor média, em comparação com as demais ferramentas analisadas.

Quadro 4.12 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica F das ferramentas analisadas.

| Classificação do Teste de Tukey para a métrica F | | | | |
|--|-------|--------|---|---|
| Abordagem | Média | Grupos | | |
| PhotoCluster | 0,947 | A | | |
| Picasa | 0,873 | A | B | |
| Photoshopel | 0,776 | | B | C |
| FaceSort | 0,743 | | | C |

Semelhantemente aos apresentados nos Quadros 4.9 e 4.11, os resultados dos grupos obtidos do teste de *Tukey* para a métrica *JI*, exibidos no Quadro 4.13, evidenciam que **há**, estatisticamente, diferença significativa entre a ferramenta *PhotoCluster* e as demais, evidenciando que esta ferramenta apresenta melhor desempenho do que as demais também para a métrica *JI*.

Quadro 4.13 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica JI das ferramentas analisadas.

| Classificação do Teste de Tukey para a métrica JI | | | | |
|---|-------|--------|---|---|
| Abordagem | Média | Grupos | | |
| PhotoCluster | 0,905 | A | | |
| Picasa | 0,789 | | B | |
| Photoshopel | 0,643 | | B | C |
| FaceSort | 0,594 | | | C |

Analisando os resultados no Quadro 4.14 produzidos pelo teste de *Tukey* para a métrica *FM*, pode ser verificado que **não há**, estatisticamente, diferença significativa entre as ferramentas *PhotoCluster* e *Picasa*, apesar da *PhotoCluster* ter apresentado a melhor média para esta métrica.

Quadro 4.14 – Grupos de similaridade obtidos por meio do teste de Tukey para a métrica FM das abordagens investigadas.

| Classificação do Teste de Tukey para a métrica FM | | | | |
|---|-------|--------|---|---|
| Abordagem | Média | Grupos | | |
| PhotoCluster | 0,950 | A | | |
| Picasa | 0,877 | A | B | |
| Photoshopel | 0,795 | | B | C |
| FaceSort | 0,763 | | | C |

Portanto, de acordo com os resultados dos grupos de similaridade produzidos pelos testes de *Tukey*, foi verificado, estatisticamente, que em 100% dos casos a ferramenta *PhotoCluster* apresentou a melhor média, em relação às demais ferramentas analisadas. Em 42,85% (3 de 7) dos testes, esta ferramenta apresentou o melhor resultado isoladamente e em 85,71% (6 de 7) dos testes apresentou o melhor resultado, isoladamente ou em conjunto com a ferramenta *Picasa*, a segunda melhor classificada. Assim, as ferramentas analisadas foram classificadas na seguinte ordem, em função de seu desempenho no agrupamento de faces:

- (1) ***PhotoCluster***;
- (2) *Picasa*;
- (3) *Photoshopel*; e
- (4) *FaceSort*.

4.4. Considerações Finais

Neste capítulo, foi descrita a parte experimental do trabalho desenvolvido no âmbito desta dissertação. Tal parte experimental teve como propósito a validação da abordagem proposta e descrita no Capítulo 3. A abordagem proposta foi avaliada, inicialmente, por meio de testes visuais, a partir da análise gráfica de seus resultados em um processo de calibração. Posteriormente, os resultados foram analisados estatisticamente pelos testes ANOVA fator único e *Tukey*, em comparação com outros resultados provenientes das ferramentas concorrentes consideradas.

A partir da análise dos resultados obtidos no experimento objetivo,

pôde-se concluir que existe suporte estatístico para afirmar que a abordagem proposta, *PhotoCluster*, apresenta um desempenho superior a todas as demais ferramentas concorrentes (*Picasa*, *Photoshopel* e *FaceSort*), em 42,85% (3 de 7) dos testes. Em 85,71% (6 de 7) dos testes, a ferramenta *PhotoCluster* também apresentou o melhor resultado, em conjunto com a segunda melhor ferramenta classificada, *Picasa*.

Foi descrita neste capítulo a avaliação experimental realizada, composta de testes visuais, testes de hipóteses e métricas de avaliação de agrupamento, que poderá ser utilizada em experimentos futuros com outras ferramentas e abordagens para o problema de agrupamento de faces em coleções de fotografias digitais.

No próximo capítulo, é apresentada uma síntese da pesquisa reportada nesta dissertação. Além disto, são apresentadas as contribuições e formuladas propostas para trabalhos futuros, as quais levam em conta as dificuldades e os problemas encontrados durante o desenrolar da pesquisa.

Capítulo 5

Considerações Finais

Neste capítulo, são apresentadas: (i) uma síntese dos principais tópicos abordados nesta dissertação; (ii) as contribuições resultantes da pesquisa desenvolvida; e (iii) proposições para trabalhos futuros, levando em conta as dificuldades e os problemas encontrados durante o desenrolar da pesquisa.

5.1. Síntese da Pesquisa

A pesquisa desenvolvida nesta dissertação teve como meta central a concepção de uma abordagem para a solução do problema de agrupamento de faces em coleções de fotografias digitais. Como motivação, tem-se o crescente interesse nesta área pela indústria de tecnologias da informação (*Adobe, Google e HP*, por exemplo), assim como pela lacuna de trabalhos sobre este tema, fundamentados em abordagens bem documentadas, com avaliações experimentais e estatísticas e que apresentem um arcabouço bem definido, visando *benchmarking*. O preenchimento desta lacuna é uma das contribuições desta dissertação, conforme destacado na próxima seção.

Durante o desenvolvimento da dissertação, foram conduzidos estudos relacionados a cada um dos principais tópicos identificados para o problema investigado: (i) propósito do sistema; (ii) características extraídas para a identificação pessoal; (iii) métricas de similaridade; (iv) técnicas de agrupamento/reconhecimento; e (v) métricas de avaliação de agrupamento. Tais estudos foram apresentados no Capítulo 2 desta dissertação e serviram como fundamentação para a concepção da arquitetura da abordagem proposta. Além disto, os referidos estudos também fundamentaram a implementação da aplicação de *software* utilizada no processo de validação

desta abordagem.

Com base nos estudos iniciais, foi desenvolvida uma abordagem original, fundamentada em técnicas e métodos do estado da arte em agrupamento de faces, relatado no Capítulo 2. Os detalhes desta extensão, da arquitetura, dos módulos, das técnicas e dos algoritmos que compõem a abordagem proposta foram apresentados no Capítulo 3.

O sistema computacional desenvolvido com base na abordagem proposta foi submetido a experimentos para a avaliação de sua viabilidade. Os resultados obtidos foram comparados com os resultados de outras aplicações de *software* destinadas ao agrupamento de faces em coleções de fotografias digitais. Estes resultados, juntamente com a descrição de todo o procedimento experimental utilizado, foram apresentados no Capítulo 4.

Por fim, foi possível concluir, a partir dos experimentos objetivos conduzidos e do processamento estatístico dos dados coletados, que os resultados produzidos pela abordagem proposta apresentaram melhores taxas na maioria das métricas de avaliação de agrupamento adotadas, quando comparadas aos resultados equivalentes produzidos pelas demais aplicações de *software* consideradas. Considerações a respeito desta avaliação e conclusões são apresentadas na próxima seção.

5.2. Contribuições do Trabalho

A partir dos resultados obtidos nos experimentos relatados no Capítulo 4, é possível concluir que objetivo de desenvolver uma abordagem inovadora e robusta para o problema de agrupamento de faces, visando a obtenção de um melhor desempenho em relação ao estado da arte, com o intuito de facilitar a organização e o compartilhamento de coleções de fotografias, foi alcançado. Para validar a abordagem proposta, um estudo experimental e estatístico foi conduzido em uma base de imagens pública.

A abordagem proposta tem como principal diferencial a agregação de módulos para atenuar os efeitos da queda de desempenho de agrupamento, normalmente associada a variações de iluminação, expressões faciais e pose. A implementação do sistema proposto atingiu o resultado esperado.

As principais contribuições da pesquisa podem ser enumeradas como segue:

- (1) Compilação de um corpo bibliográfico relacionado às principais técnicas e métodos utilizados para o problema de agrupamento de faces;
- (2) Formalização do problema de agrupamento de faces, não identificada no corpo bibliográfico compilado;
- (3) Proposição de uma abordagem inovadora e robusta, fundamentada no corpo bibliográfico compilado;
- (4) Constatação experimental de benefícios obtidos com a utilização de técnicas de pré-processamento, visando à compensação de iluminação;
- (5) Utilização do método SURF para agrupamento de faces, dado que nenhum trabalho relacionado utilizou esta característica para o problema em questão;
- (6) Elaboração de um arcabouço modular e extensível para a tarefa de agrupamento de faces;
- (7) Realização de um experimento objetivo para a validação da abordagem proposta, utilizando métricas de avaliação de agrupamento fundamentados no corpo bibliográfico compilado e um conjunto de testes estatísticos formais bem definido;
- (8) Desenvolvimento de um procedimento de comparação objetiva, baseada em resultados experimentais, que possibilitou a confrontação da abordagem proposta com ferramentas concorrentes de grandes empresas da área, revelando a existência de diferenças estatisticamente relevantes entre as mesmas;

Desta forma, pode-se verificar que os dois objetivos principais levantados foram atingidos com sucesso: (i) a obtenção de um melhor desempenho em relação ao estado da arte, com o intuito de facilitar a organização e o compartilhamento de coleções de fotografias de usuários; e (ii) a investigação de formas de minimizar a queda de desempenho,

normalmente associada à variações de iluminação, expressões faciais e pose.

Primeiramente, em relação ao primeiro objetivo, a partir dos resultados obtidos no *Experimento de Comparação*, pôde-se concluir que a utilização de testes estatísticos simples, por exemplo, avaliação de IC através de gráficos *boxplot*, não é suficiente para avaliar diferenças estatísticas entre sistemas. Para tanto, fez-se necessário a administração de testes estatísticos mais elaborados, a saber: (i) Teste ANOVA fator único; e (ii) Teste de *Tukey*. A partir dos resultados destes testes, pôde-se verificar estatisticamente que a abordagem proposta, PhotoCluster, apresenta um desempenho superior a todas as demais ferramentas concorrentes consideradas no experimento (*Picasa*, *Photoshop* e *FaceSort*), em 42,85% (3 de 7) dos testes. Em 85,71% (6 de 7) dos testes, a ferramenta PhotoCluster também apresentou o melhor resultado, em conjunto com a segunda melhor ferramenta classificada, *Picasa*.

Adicionalmente, em relação ao segundo principal objetivo, destaca-se uma melhoria no desempenho do sistema proposto devido ao uso da filtragem homomórfica e da equalização de histograma que auxiliaram na minimização da perda de desempenho associada à variações de iluminação. Para minimizar as dificuldades de reconhecimento ocasionadas por variações de expressões faciais e pose, foi adotado o método SURF na abordagem proposta, cujos resultados experimentais comprovaram sua robustez em relação a tais problemas.

A seguir, são apresentadas sugestões para a melhoria e/ou para a complementação dos resultados obtidos nesta pesquisa.

5.3. Trabalhos Futuros

Esta seção apresenta algumas sugestões para trabalhos futuros visando à obtenção do melhor desempenho e novas funcionalidades do sistema proposto. Algumas técnicas e métodos adotados na abordagem proposta podem ser substituídas por algoritmos similares em cada um dos módulos propostos, e.g., *Gradientfaces* (ZHANG et al., 2009a) para compensação da

iluminação; uso de características SIFT ou LDA, ao invés do SURF; algoritmos de agrupamento *K-Means*, EM, dentre outros, ao invés do HAC. Portanto, propõe-se como trabalho futuro investigar a substituição de alguns componentes principais da abordagem para a verificação de uma possível melhora no desempenho do sistema.

Outra proposta para trabalho futuro é a ampliação da base de imagens, considerando uma base massiva com grande variações de indivíduos, expressões faciais e pose, com o objetivo de avaliar o comportamento do sistema e das demais ferramentas concorrentes neste cenário um pouco mais próximo da realidade dos usuários.

Adicionalmente, outra importante proposta para trabalho futuro diz respeito a realização de um experimento para avaliação da eficiência computacional em termos de velocidade de processamento (*performance*) da ferramenta desenvolvida em comparação com as três ferramentas concorrentes consideradas nos experimentos realizados: *Adobe Photoshop Elements* (PHOTOSHOPEL, 2010), *Google Picasa* (PICASA, 2010), e *PittPatt Facesort* (FACESORT, 2010).

No tocante às limitações da abordagem desenvolvida, dois pontos a serem investigados foram identificados, a saber:

- (1) Limitação do SURF em relação a pose: foi verificado na literatura que o SURF é capaz de extrair características em faces com no máximo 30° de rotação fora do plano, o que inviabiliza sua utilização em cenários onde faces com grandes variações de pose (e.g., *full profile*) fossem consideradas;
- (2) Fragmentação de alguns grupos obtidos, percebida nos experimentos. Esta fragmentação, apesar de constituir uma limitação, é menos grave e mais fácil de corrigir do que seria o agrupamento incorreto de faces de diferentes pessoas. Na situação atual, tal fragmentação pode ser corrigida por um processo de anotação manual realizada pelo usuário final.

Adicionalmente, vislumbram-se extensões, a fim de permitir a aplicação desta abordagem a vídeos digitais, considerados uma fonte de

informação mais rica do que imagens estáticas, além de apresentar maior quantidade de dados (quadros) e fornecer informações temporais úteis para rastreamento e reconhecimento de objetos.

Dentre as potenciais aplicações de agrupamento de faces em vídeos digitais destacam-se: (i) a sumarização de vídeos de segurança (SONY et al., 2011); (ii) a identificação de pessoas específicas em vídeos noticiários (GAO, EKENEL e STIEFELHAGEN, 2011); (iii) a catalogação de cenas por atores presentes no vídeo (YAMAMOTO, YAMAGUCHI e AOKI, 2010); e (iv) a determinação do elenco de atores (*cast list*) presente no vídeo (ZHANG et al., 2009b).

Referências Bibliográficas

- ABATE, A., NAPPI, M., RICCIO, D. e SABATINO, D. 2D and 3D face recognition: A survey. *Pattern Recognition Letters*, 28:1885-1903, 2007.
- ABONYI, J., BABUSKA, R., e SZEIFERT, F. Modified Gath-Geva fuzzy clustering for identification of Takagi-Sugeno fuzzy models. *IEEE Transactions on Systems, Man and Cybernetics*, 32:612-621, 2002.
- ABONYI, J. e SZEIFERT, F. Supervised fuzzy clustering for the identification of fuzzy classifiers. *Pattern Recognition Letters*, 24:2195-2207, 2003.
- ANTONOPOULOS, P., NIKOLAIDIS, N. e PITAS, I. Hierarchical Face Clustering using SIFT Image Features. *IEEE Symposium on Computational Intelligence in Image and Signal Processing*, 325-329, 2007.
- BAY, H., TUYTELAARS, T. e VAN GOOL, L. Surf: Speeded up robust features. *European Conference on Computer Vision*, 404-417, 2006.
- BAUER, J., SÜNDERHAUF, N. e PROTZEL, P. Comparing Several Implementations of Two Recently Published Feature Detectors. *In Proc. of the International Conference on Intelligent and Autonomous Systems*, IAV, Toulouse, France, 22:481-494, 2007.
- BELHUMEUR, P. N., HESPANHA, J. e KRIEGMAN, D. J. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:711-720, 1997.
- BOSWORTH, J.H. e ACTON, S.T. Morphological scale-space in image processing. *Digital Signal Processing*, 13:338-367, 2003.
- BROWN, M. e LOWE, D.G. Invariant features from interest point groups. *British Machine Vision Conference*, Cardiff, Wales, 656-665, 2002.
- BRADSKI, G. R., KAEHLER, A. e PISAREVSKY, V. Learning-based computer vision with Intel's open source computer vision library. *Intel Technology Journal-Compute-Intensive, Highly Parallel Applications and Uses*, 9:119-130, 2005.
- CAI, D., HE, X. e HAN, J. Spectral regression for efficient regularized subspace learning. *In Proceedings of the 11th International Conference on Computer Vision*, 1-8, 2007.
- CAI, D., HE, X., HU, Y., HUANG, T. e HAN, J. Learning a spatially smooth subspace for face recognition. *In Proceedings of the 20th International Conference on Computer Vision and Pattern Recognition*, 7:1-7, 2007.
- Caltech Face Database. Disponível em: <http://www.vision.caltech.edu/archive.html>, 1999. Acesso em 27/12/2010.
- CAO, Z., YIN, Q., TANG, X. e SUN, J. Face recognition with learning-based descriptor. *In Proceedings of the 23th International Conference on*

- Computer Vision and Pattern Recognition*, 10:2707-2714, 2010.
- CASTRILLÓN, M., DÉNIZ, O., ANTÓN-CANALÍS, L. e LORENZO, J. Face and Facial Feature Detection Evaluation. *International Conference on Computer Vision Theory and Applications – VISAPP*, 167-172, 2008.
- CHACON, M., RIVAS, P. e RAMIREZ, A. A Fuzzy Clustering Approach for Face Recognition Based on Face Feature Lines and Eigenvectors. *Engineering Letters*, 35-44, 2006.
- CHEN, L., HU, B., ZHANG, L., LI, M. e ZHANG, H.J. Face annotation for family photo album management. *International Journal of Image and Graphics*, 3:01-14, 2003.
- CHOI, J., YANG, S., RO, Y. e PLATANIOTIS, K. Face annotation for personal photos using context-assisted face recognition. *In Proceedings of the 1st ACM International Conference on Multimedia information retrieval*, 1:44-51, 2008.
- COOTES, T. F., TAYLOR, C. J., COOPER, D. H. e GRAHAM, J. Active Shape Models - Their Training and Application, *Computer Vision and Image Understanding*, 61:38-59, 1995.
- CUI, J., WEN, F., XIAO, R., TIAN, O. e TANG, X. Easyalbum: an interactive photo annotation system based on face clustering and re-ranking. *In Proceedings of the SIGCHI conference on Human factors in computing systems*, 1222-1228, 2007.
- DAVIS, J. e GOADRICH, M. The relationship between Precision-Recall and ROC curves. *In Proceedings of the 23th International Conference on Machine Learning*, 233-240, 2006.
- DEGTYAREV, N. e SEREDIN, O. Comparative Testing of Face Detection Algorithms. ICISP, *International Conference on Image and Signal Processing*, 200-209, 2010.
- DELAC, K., GRGIC, M. e KOS, T. Sub-image homomorphic filtering technique for improving facial identification under difficult illumination conditions. *In Proceedings of International Conference on Systems, Signals and Image Processing*, 95-98, 2006.
- EVERITT, B. S., LANDAU, S. e MORVEN, L. Cluster Analysis. *A Hodder Arnold Publication*, 4:50-58, 2001.
- FACEBOOK STATISTICS. Facebook Press. Disponível em: <<http://www.facebook.com/press/info.php?statistics>>, 2010. Acesso em 15/10/2010.
- FACESORT. Pittsburgh Pattern Recognition – FaceSort. Disponível em: <<http://demo.pittpatt.com/facesort/>>, 2010. Acesso em 01/10/2010.
- FU, Y., HSU, W. e LEE, L. A latent semantic retrieval and clustering system for personal photos with sparse speech annotation. *In Proceedings of the 3rd Workshop on Searching spontaneous conversational speech*, 39-40, 2009.
- GALLAGHER, A. e CHEN, T. Clothing cosegmentation for recognizing people. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,

1-8, 2008.

- GALLAGHER, A. e CHEN, T. Using Context to Recognize People in Consumer Images. *IPSI Transactions on Computer Vision and Applications*, 1:115-126, 2009.
- GAO, H., EKENEL, H. K. e STIEFELHAGEN, R. Identifying Important People in Broadcast News Videos. *In Proceedings of the Conference on Machine Vision Applications (IAPR)*, Nara, Japan, 127-136, 2011.
- GONZALEZ, R. C. e WOODS, R. E. Digital Image Processing. *Addison-Wesley Pub (Sd)*, vol. 3, 2010.
- GUILLAUMIN, M., VERBEEK, J. e SCHMID, C. Is that you? Metric learning approaches for face identification. *In Proceedings of the 22th International Conference on Computer Vision and Pattern Recognition*, 498-505, 2009.
- HALKIDI, M., BATISTAKIS, Y. e VAZIRGIANNIS, M. On Clustering Validation Techniques. *Journal of Intelligent Information Systems*, 17:107-145, 2011.
- HAR-PELED, S. e KUMAR, N. Approximate Nearest Neighbor Search for Low Dimensional Queries. *Computational Geometry, Data Structures and Algorithms*, 2010.
- HE, X., YAN, S., HU, Y., NIYOGI, P. e ZHANG, H. Face recognition using laplacian faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27:328-340, 2005.
- HE, X. e NIYOGI, P. Locality Preserving Projections. *Advances in Neural Information Processing Systems, NIPS*, 2003.
- HUA, G.; AKBARZADEH, A. A robust elastic and partial matching metric for face recognition. *In Proceedings of the 13th International Conference on Computer Vision*, 2082-2089, 2009.
- HUANG, J., KUMAR, S., MITRA, M., ZHU, W. e ZABIH, R. Image indexing using color correlograms. *In Proceedings of the 10th International Conference on Computer Vision and Pattern Recognition*, 1997.
- HUANG, G., RAMESH, M., BERG, T. e LEARNED-MILLER, E. Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. University of Massachusetts, Amherst, Technical Report, 07-49, October, 2007.
- HUBERT, L. e ARABIE, P. Comparing partitions. *Journal of Classification*, 2:10-22, 1985.
- IVCHENKO, G. I. e HONOV, S. A. On the Jaccard Similarity Test. *Journal of Mathematical Sciences*, vol. 88, 1998.
- JAIN, A. K. e DUBES, R.C. Algorithms For Clustering Data. *Prentice Hall*, New Jersey, 1988.
- JAIN, R. The Art of Computer Systems Performance Analysis: techniques for experimental design, measurement, simulation, and modeling. *Wiley & Sons*, 1991.
- JARQUE, C. e BERA, A. Efficient test for normality, heterocedasticity and

- serial independence of regression residuals. *Econometrics Letters*, 6:255–259, 1980.
- JAYECH, K. e MAHJOUB, M. Clustering and Bayesian network for image of faces classification. *International Journal of Advanced Computer Science and Applications, Special Issue on Image Processing and Analysis (IJACSA)*, 2011.
- JO LLIFFE, I.T. Principal Component Analysis, second edition, New York: Springer-Verlag New York, Inc., 2002.
- JUAN, L. e GWON, O. A Comparison of SIFT, PCA-SIFT and SURF. *International Journal of Image Processing (IJIP)*, 3:143-152, 2009.
- KANDEFER, M. e SHAPIRO, S. An F-Measure for Context-based Information Retrieval. *In Proceedings of the 9th International Symposium on Logical Formalizations of Commonsense Reasoning*, 2009.
- KAPOOR, A., HUA, G., AKBARZADEH, A. e BAKER, S. Which Faces to Tag: Adding Prior Constraints into Active Learning. *In Proceedings of the 13th International Conference on Computer Vision*, 2009.
- KYPEROUNTAS, M., TEFAS, A. e PITAS, I. Face recognition via adaptive discriminant clustering. *In Proceedings of the 15th International Conference on Image Processing*, 2744-2747, 2008.
- LAGANIÈRE, R. OpenCV 2 Computer Vision Application Programming Cookbook, Packt Publishing, UK, 2011.
- LEI, Z., CHEN, L., LI, M. e ZHANG, H. Automated annotation of human faces in family albums. *In Proceedings of the 9th ACM International Conference on Multimedia*, 355-358, 2003.
- LEVINE, D. M., BERENSON, M. L., e STEPHAN, D. Estatística: Teoria e aplicações. LTC Editora, Rio de Janeiro, 2000.
- LI, Z. e TANG, X. Bayesian face recognition using support vector machine and face clustering. *In Proceedings of the 17th International Conference on Computer Vision and Pattern Recognition*, 2:374-380, 2004.
- LIN, D., KAPOOR, A., HUA, G. e BAKER, S. Joint People, Event, and Location Recognition in Personal Photo Collections Using Cross-Domain Context. *In Proceedings of the 11th European Conference on Computer Vision*, 2010.
- LOWE, D. G. Object recognition from local scale-invariant features. *In Proceedings of the International Conference on Computer Vision*, 1150-1157, 1999.
- MA, Y., DING, X., WANG, Z. e WANG, N. Robust precise eye location under probabilistic framework. *In Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, 2004.
- MARQUES FILHO, O. e VIEIRA NETO, H. Processamento Digital de Imagens, Rio de Janeiro. Brasport, 1999.
- MATTEUCCI, M. A Tutorial on Clustering Algorithms. Disponível em <http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/hierarchical.html>. Acesso em 09/06/2011.

- MCLACHLAN, G. J. Discriminant Analysis and Statistical Pattern Recognition In: *Wiley Interscience*, 2004.
- MESSER, K., MATAS, J., KITTLER, J., LUETTIN, J. e MAITRE, G. XM2VTSDB: The Extended M2VTS Database. In *Proceedings of the 2nd International Conference on Audio and Videobased Biometric Person Authentication*, 1999.
- MOGHADDAM, B., JEBARA, T. e PENTLAND, A. Bayesian face recognition. *Pattern Recognition Letters*, 33:1771-1782, 2007.
- MUJA, M. e LOWE, D. G. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. *International Conference on Computer Vision Theory and Applications, VISAPP*, 2009.
- OJALA, T., PIETIKAINEN, M. e HARWOOD, D. A comparative-study of texture measures with classification based on feature distributions. *Pattern Recognition*, vol. 29, 1996.
- OLSEN, D.L. e DELEN, D. Advanced Data Mining Techniques. Springer, ed. 1, USA, 2008.
- OPENCV. Intel® Open Computer Vision Library. Disponível em: <<http://sourceforge.net/projects/opencvlibrary/>>, 2006. Acesso em 27/10/2010.
- OSTROW, A. How Facebook Serves Up Its 15 Billion Photos. Disponível em: <<http://mashable.com/2009/04/30/facebook-photo-sharing/>>, 2009. Acesso em 15/10/2010.
- PALIT, B., NIGAM, R., PERLMUTTER, K. e PERLMUTTER, S. Spectral face clustering. In *Proceedings of the 12th International Conference on Computer Vision Workshops*, 186-191, 2009.
- PAPAGEORGIU, C. P., OREN, M. e POGGIO, T. A general framework for object detection, *Proc. of ICCV*, 555-562, 1998.
- PARKER, J. R. Algorithms for Image Processing and Computer Vision. *John Wiley & Sons*, 1997.
- PERLMUTTER, K., PERLMUTTER, S., ALSPECTOR, J., EVERINGHAM, M., HOLUB, A., ZISSERMAN, A. e PERONA, E. Using relevance feedback in face recognition. United States Patent Application ID, 20090034805, 2006.
- PERRONNIN, F., DUGELAY, J-L. e ROSE, K. Deformable face mapping for person identification. In *Proceedings of the International Conference on Image Processing*, 1:661-664, 2003.
- PERRONNIN, F. e DUGELAY, J-L. Clustering face images with application to image retrieval in large databases. In *Proceedings of the Biometric Technology for Human Identification*, 5779: 256-264, 2005.
- PHILLIPS, P., MOON, H, RAUSS, P. e RIZVI, S. The FERET evaluation methodology for face recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, n.10, 2000.
- PHOTOSHOPEL. Adobe Photoshop Elements. Disponível em: <<http://www.adobe.com/br/products/photoshopel/>>, 2010. Acesso em

01/10/2010.

PICASA. Google - Picasa. Disponível em:
<<http://picasa.google.com/features.html>>, 2010. Acesso em
01/10/2010.

PINTO, N., COX, D. e DICARLO, J. Why is Real-World Visual Object Recognition Hard. In *PLoS Computational Biology*, ed. 27, 2008.

PINTO, N., DICARLO, J. e COX, D. How far can you get with a modern face recognition test set using only simple features?. In *Proceedings of the 22th International Conference on Computer Vision and Pattern Recognition*, 2591-2598, 2009.

PRINCE, S. e HELDER, J. Probabilistic linear discriminant analysis for inferences about identity. In *Proceedings of the 11th International Conference on Computer Vision*, 2007.

PRINCE, S. e HELDER, J. Bayesian Identity Clustering. In *Proceedings of the Canadian Conference on Computer and Robot Vision*, 32-39, 2010.

RAMÍREZ, G. Estudio de métodos de reconocimiento de patrones basado en redes neurales y lógica difusa. In Chihuahua Institute of Technology, *Dissertação de Mestrado*, 2005.

RAND, W.M. Objective criteria for the evaluation of clustering methods. *Journal of American Statistics Association*, 66: 846-850, 1971.

RIESENHUBER, M. e POGGIO, T. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2: 1019-1025, 1999.

RUBNER, Y., TOMASI, C. e GUIBAS, L. The Earth Mover's distance as a metric for image retrieval. *International Journal of Computer Vision (IJCV)*, vol. 40. Hingham, MA, USA, 2000.

SAMARIA, F. e HARTER, A. Parameterization of a stochastic model for human face identification. In *Proceeding of the IEEE Workshop on Applications of Computer Vision*, 1994.

SANTOS, J.M. e EMBRECHTS, M. On the Use of the Adjusted Rand Index as a Metric for Evaluating Supervised Classification. In *Proceedings of the 9th International Conference on Artificial Neural Networks*, 2009.

SHARAVANAN, S. e AZATH, M. LDA based Face Recognition by using Hidden Markov Model in current trends. *International Journal of Engineering and Technology*, 77-85, 2009.

SILVA, ALICE. Análise Classificatória. *Universidade Nova de Lisboa - Faculdade de Ciências e Tecnologia*, 2005. Disponível em
<<http://ferrari.dmat.fct.unl.pt/services/AnaliseDados/Cluster.pdf>>. Acesso em 28/12/2011.

SIMARD, P., LE CUN, Y., DENKER, J. e VICTORRI, B. Transformation Invariance in Pattern Recognition — Tangent Distance and Tangent Propagation. Neural networks: tricks of the trade, *Lecture Notes in Computer Science*, Springer, Heidelberg, 1524:239-274, 1998.

SONY, A., AJITH, K., THOMAS, K., THOMAS, T. e OEEPA P. L. Video Summarization By Clustering Using Euclidean Distance. In *Proceedings*

- of the 2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies (ICSCCN), India, 642-646, 2011.
- SU, Y., SHAN, S., CHEN, X. e GAO, W. Hierarchical Ensemble of Global and Local Classifiers for Face Recognition. *IEEE Transactions on Image Processing*, vol. 18, no. 8, 2009.
- STEGMANN, M. B., ERSBOLL, B. K. e LARSEN, R. FAME - a flexible appearance modeling environment. *IEEE Transactions on Medical Imaging*, 22:1319-1331, 2003.
- STONE, Z., ZICKLER, T. e DARRELL, T. Autotagging facebook: Social network context improves photo annotation. *In Proceedings of the 21th International Conference on Computer Vision and Pattern Recognition*, 2008.
- SWAIN, M. J. e BALLARD, B. H. Color indexing. *International Journal of Computer Vision*, 7:11-32, 1991.
- THE FACEBOOK BLOG. Facebook Press. Disponível em: <<http://blog.facebook.com/blog.php?post=2406207130>>, 2007. Acesso em 15/10/2010.
- THEODORIDIS, S. e KOUTROUMBAS, K. Pattern Recognition. Academic Press, 1999.
- TIAN, Y., LIU, W., XIAO, R. e WEN, F. A Face Annotation Framework with Partial Clustering and Interactive Labeling. *In Proceedings of the 20th International Conference on Computer Vision and Pattern Recognition*, 1-8, 2007.
- VAPNIK, V. N. The Nature of Statistical Learning Theory. Springer, 2nd edition, 1999.
- VIOLA, P. e JONES, M. Rapid object detection using a boosted cascade of simple features. *In Proceedings of the 14th International Conference on Computer Vision and Pattern Recognition*, 511-518, 2001.
- WANGNHEIM, C.G.; WANGNHEIM, A. Raciocínio Baseado em Casos. Editora Manole, 2003.
- WECHSLER, H. Reliable Face Recognition Methods: System Design, Implementation and Evaluation (International Series on Biometrics). Springer-Verlag New York, Secaucus, NJ, USA, 2006.
- WINN, J., CRIMINISI, A. e MINKA, T. Object categorization by learned universal visual dictionary. *In Proceedings of the 10th International Conference on Computer Vision (ICCV)*, 2:1800-1807, 2005.
- WOLF, L., HASSNER, T. e TAIGMAN, Y. Descriptor based methods in the wild. *In Faces in Real-Life Images Workshop - ECCV*, 2008.
- WOLF, L., HASSNER, T. e TAIGMAN, Y. Similarity Scores based on Background Samples. *In Asian Conference on Computer Vision*, 2009.
- WRIGHT, J. e HUA, G. Implicit elastic matching with randomized projections for pose-variant face recognition. *In Proceedings of the 22th International Conference on Computer Vision and Pattern Recognition*,

1502-1509, 2009.

XLStat. Statistics package for Excel. Disponível em: <<http://www.xlstat.com/en/>>, 2011. Acesso em 13/11/2011.

XU, R. e WUNSCH, D. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16:645-678, 2005.

YAMAMOTO, K., YAMAGUCHI, O. e AOKI, H. Fast face clustering based on shot similarity for browsing video. In *Proceedings of the Progress in Informatics*, Special issue: 3D image and video technology, 7:53-62, 2010.

YOUSEF, K. e O'NEILL, E. Sunrise: Towards Location Based Clustering For Assisted Photo Management. In *Workshop on Tagging, Mining and Retrieval of Human-Related Activity Information*, 2007.

YU, H., LI, M., ZHANG, H. e FENG, J. Color texture moment for content-based image retrieval. In *Proceedings of the 9th International Conference on Image Processing*, 3:929-932, 2002.

ZHANG, L., CHEN, L., LI, M. e ZHANG, H. Automated annotation of human faces in family album. In *Proceedings of the 9th ACM International Conference on Multimedia*, 355-358, 2003.

ZHANG, S-C., FANG, B., LIANG, Y-Z., WEN, J. e WU, L. A face clustering method based on facial shape information. *International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)*, 44-49, 2011.

ZHANG, T.; TANG, Y. Y.; FANG, B.; SHANG, Z.; LIU, X. Face Recognition Under Varying Illumination Using Gradientfaces. In *Proceedings of the IEEE Transactions on Image Processing*, vol. 18, no. 11, 2599-2606, 2009a.

ZHANG W., ZHANG, T. e TRETTER, D. Beyond face: Improving person clustering in consumer photos by exploring contextual information. *IEEE International Conference on Multimedia and Expo (ICME)*, 1540-1545, 2010.

ZHANG, Y-F.; XU, C.; LU, H.; HUANG, Y-M. Character Identification in Feature-Length Films Using Global Face-Name Matching. In *Proceedings of the IEEE Transactions on Multimedia*, vol. 11, no. 7, 2009b.

ZHAO, W., CHELLAPPA, R., PHILLIPS, P. e ROSENFELD, A. Face recognition: A literature survey. *ACM Computing Surveys*, 35:339-458, 2003.

ZHU, C., WEN, F. e SUN, J. A rank-order distance based clustering algorithm for face tagging. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 481-488, 2011.

Apêndice A

Imagens Utilizadas nos Experimentos

Neste apêndice, são apresentadas as imagens utilizadas nos experimentos de calibração e de comparação. As imagens foram obtidas da base de imagens IMM (STEGMANN, ERSBOLL e LARSEN, 2003), a qual foi subdividida em 8 partições de imagens, compostas de 30 imagens cada, em um total de 240 imagens obtidas de 40 indivíduos, com 6 imagens por indivíduo. Apenas as faces recortadas são mostradas, devido ao seu tamanho original (640x480 pixels) ser elevado para apresentação.



Figura A.1 – Imagens contidas na Partição 1.

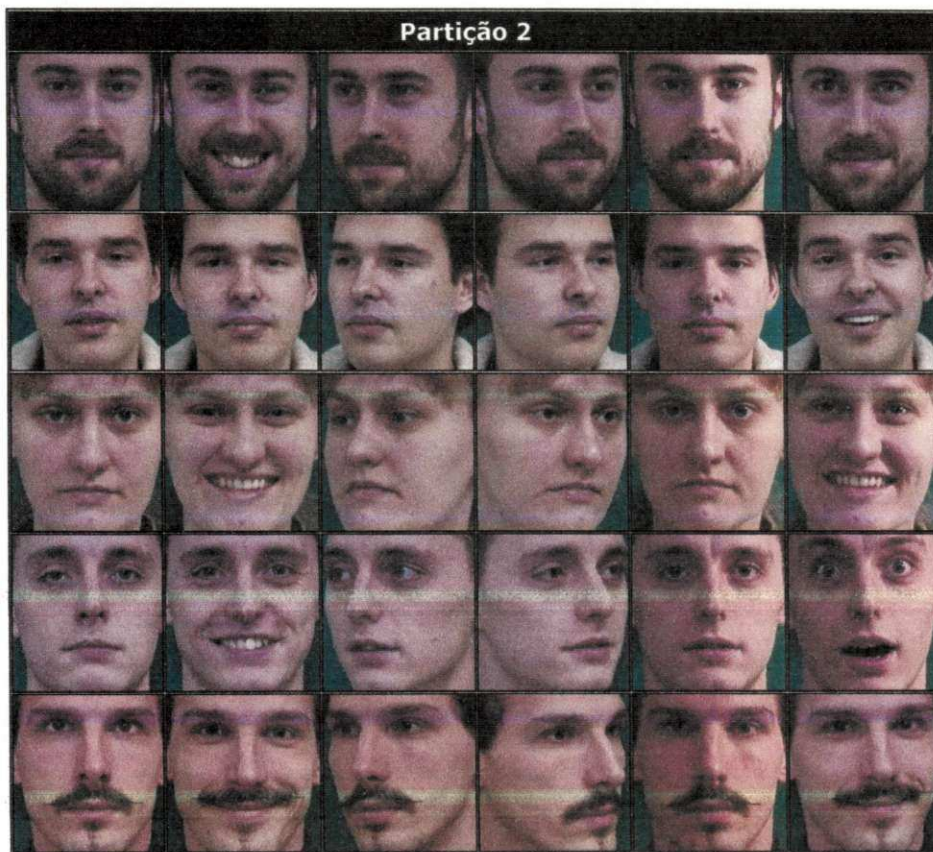


Figura A.2 – Imagens contidas na Partição 2.

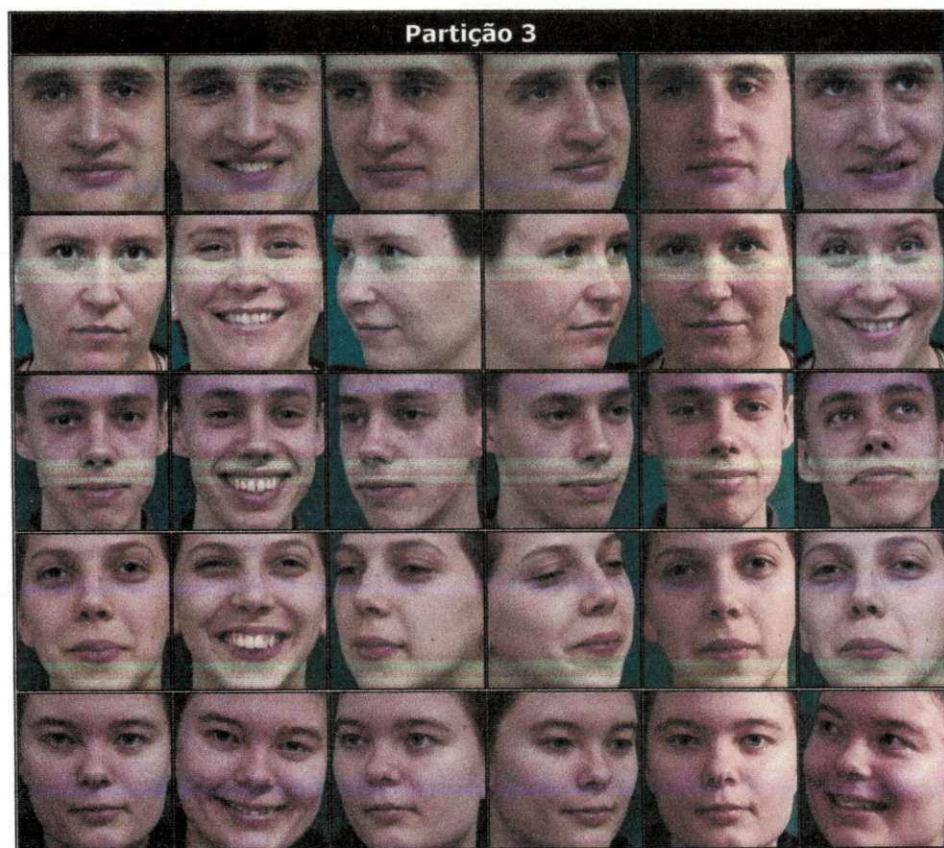


Figura A.3 – Imagens contidas na Partição 3.

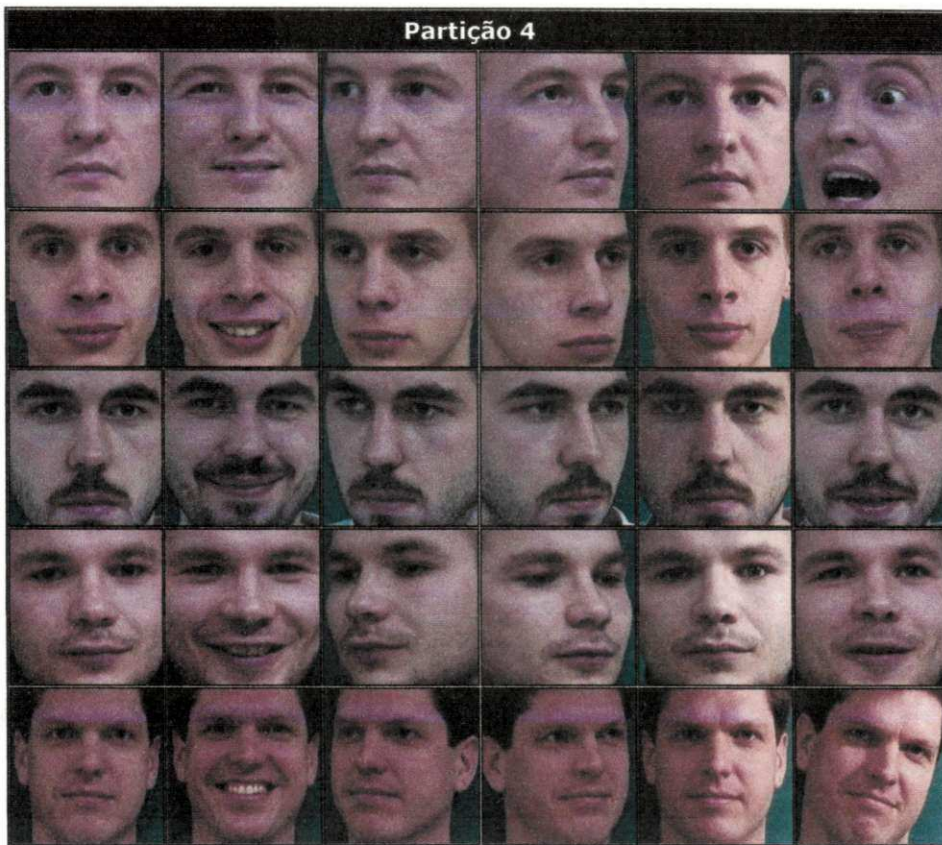


Figura A.4 – Imagens contidas na Partição 4.

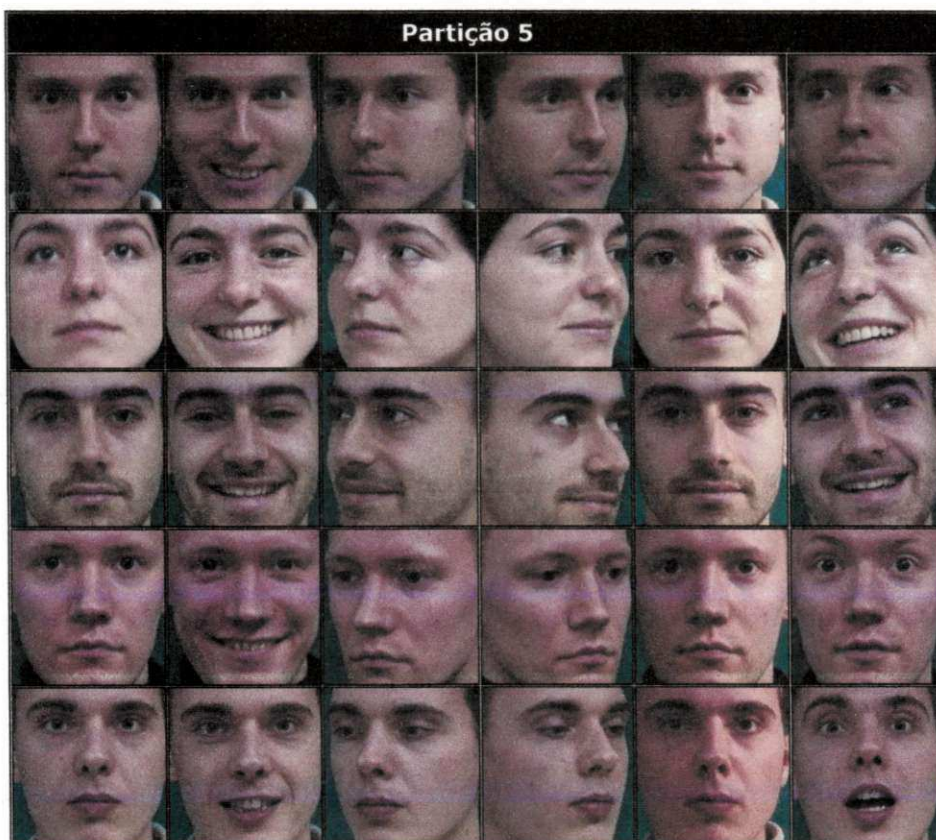


Figura A.5 – Imagens contidas na Partição 5.

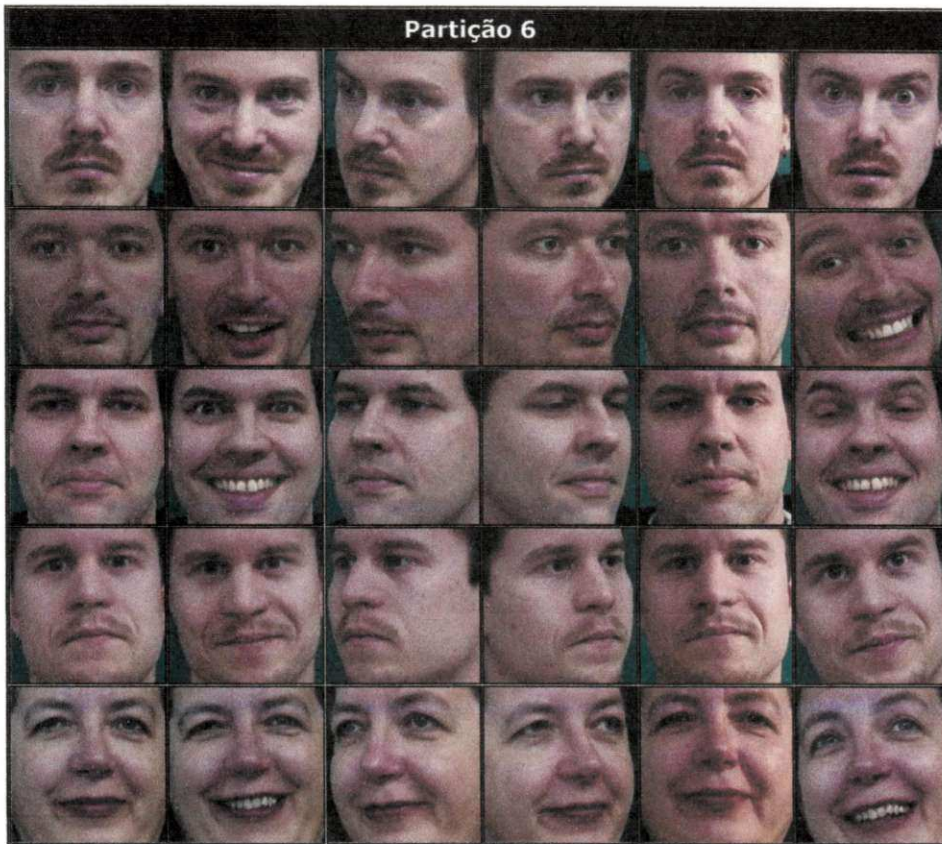


Figura A.6 – Imagens contidas na Partição 6.

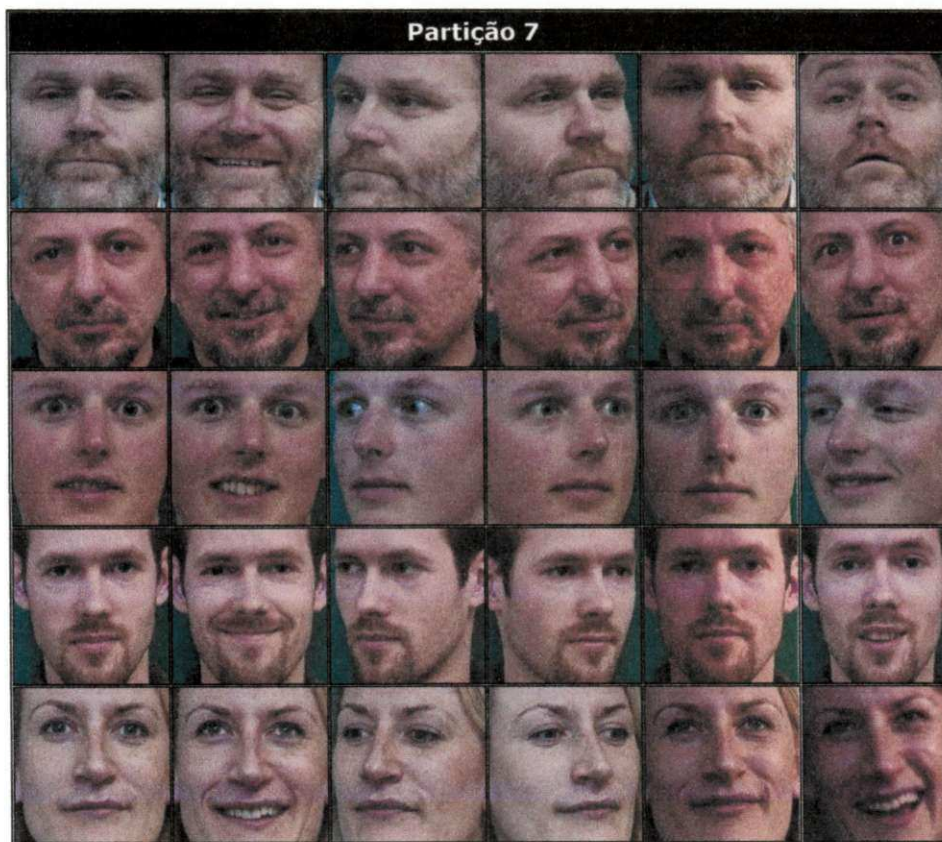


Figura A.7 – Imagens contidas na Partição 7.

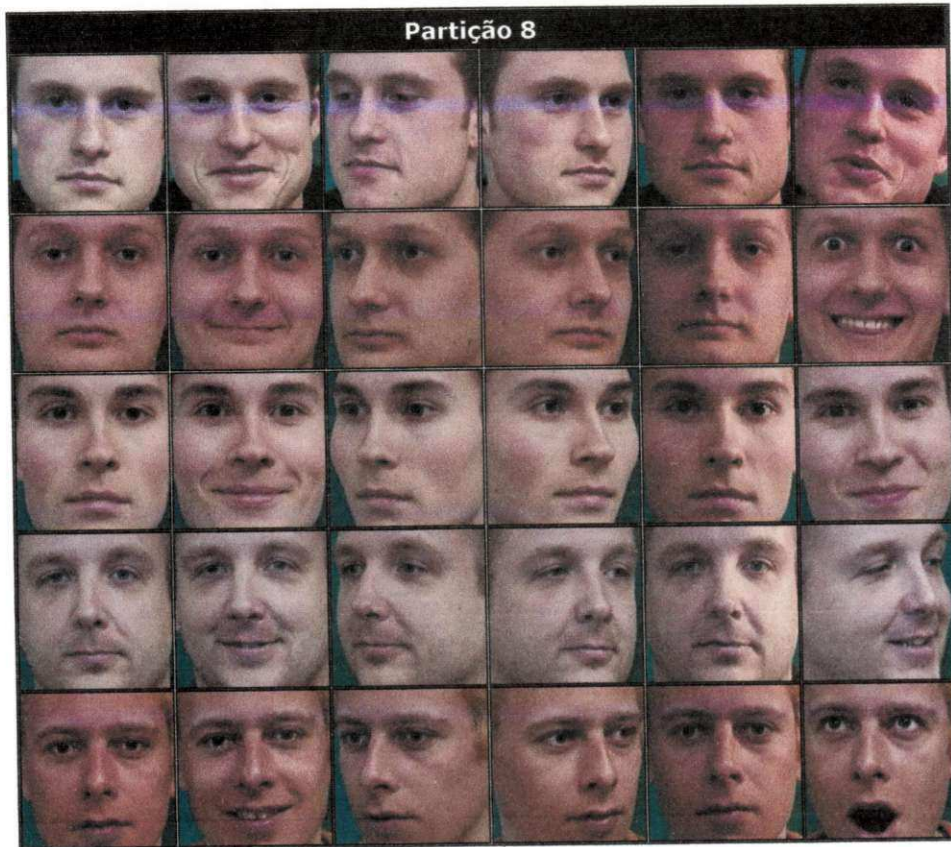


Figura A.8 – Imagens contidas na Partição 8.

Apêndice B

Testes Estatísticos

Este apêndice contém uma descrição sucinta dos testes estatísticos utilizados nesta dissertação, a saber: (i) teste de normalidade *Jarque-Bera* (JAIN, 1991); (ii) teste visual de IC conhecido por *boxplot* (JAIN, 1991); (iii) teste de hipótese F ANOVA de fator único (LEVINE, BERENSON e STEPHAN, 2000); e (iv) teste de hipótese de *Tukey* (LEVINE, BERENSON e STEPHAN, 2000).

B.1. Teste de Normalidade Jarque-Bera

Proposto por *Jarque e Bera* (1980), baseia-se na diferença entre os coeficientes de assimetria e curtose dos dados y_1, y_2, \dots, y_n e àqueles da distribuição assumida normal. Trata-se de um teste assintótico, no qual as hipóteses a serem testadas são:

$$\begin{cases} H_0: y_1, y_2, \dots, y_n \sim N(\mu, \sigma^2); \\ H_1: \text{não } H_0; \end{cases} \quad (\text{B.1})$$

O procedimento do teste consiste em calcular os valores da assimetria e da curtose da amostra e utilizar a seguinte estatística de teste, conforme Jain (1991):

$$JB = n \left(\frac{\alpha_3^2}{6} + \frac{(\alpha_4 - 3)^2}{24} \right) \quad (\text{B.2})$$

em que:

$$\alpha_3 = \frac{\sum_{i=1}^n (y_i - \bar{y})^3}{ns^3}, \alpha_4 = \frac{\sum_{i=1}^n (y_i - \bar{y})^4}{ns^4}, s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}.$$

\bar{y} a média amostral e s^2 , α_3 (assimetria) e α_4 (curtose), o segundo, terceiro

e quarto momentos centrais, respectivamente. A estatística JB segue uma distribuição qui-quadrado (χ^2) com dois graus de liberdade sob a hipótese nula. Assim, rejeita-se H_0 se $JB > \chi_{\alpha,2}^2$, onde $\chi_{\alpha,2}^2$ é o quantil de nível de confiança $(1-\alpha)$ da distribuição χ^2 com dois graus de liberdade. O teste de Jarque-Bera é conhecido por ter boas propriedades para a verificação de normalidade, é claro e simples de calcular e muito utilizado no contexto de regressão em econometria (JAIN, 1991).

B.2. Teste Visual *Boxplot*

O *boxplot* é um gráfico que possibilita representar a distribuição de um conjunto de dados com base em alguns de seus parâmetros descritivos, a saber: o quartil inferior ($Q1$), a mediana ($Q2$), o quartil superior ($Q3$) e o intervalo interquartil, $IQR = Q3 - Q1$ (JAIN, 1991). A Figura B.1 contém a ilustração de um *boxplot* e seus componentes principais.

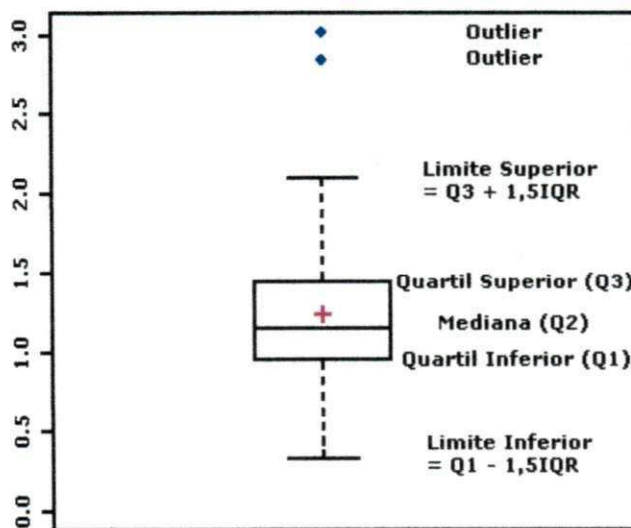


Figura B.1 – Representação de um gráfico *boxplot*.

A linha central da caixa marca a mediana do conjunto de dados. A parte inferior da caixa é delimitada pelo quartil inferior ($Q1$) e a parte superior pelo quartil superior ($Q3$). As hastes inferiores e superiores se estendem, respectivamente, do quartil inferior até o menor valor não inferior a $Q1 - 1,5IQR$ e do quartil superior até o maior valor não superior a $q3 + 1,5IQR$. Os valores inferiores a $Q1 - 1,5IQR$ e superiores a $Q3 + 1,5IQR$ são representados individualmente no gráfico, sendo caracterizados como *outliers*.

Testes visuais são utilizados para verificar graficamente o comportamento dos dados obtidos, i.e, avaliar a simetria dos dados, sua dispersão e a existência ou não de *outliers*, sendo especialmente adequado para a comparação de dois ou mais conjuntos de dados correspondentes às categorias de uma variável qualitativa (JAIN, 1991).

B.3. Teste F ANOVA fator único

O teste F ANOVA fator único verifica a hipótese nula ($H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$) de não existir nenhuma diferença nas médias aritméticas ($\mu_1, \mu_2, \mu_3, \dots, \mu_n$) obtidas em n grupos que representam populações cujas medidas foram retiradas aleatória e independentemente, seguem uma distribuição normal e têm variâncias equivalentes. Em caso de rejeição da hipótese nula, assume-se a hipótese alternativa como verdadeira, i.e., H_1 : nem todas as médias aritméticas são iguais (LEVINE, BERENSON e STEPHAN, 2000).

A probabilidade de que o teste de hipóteses rejeite a hipótese nula quando esta for verdadeira é definida pelo nível de confiança $(1-\alpha)$ do teste. Quanto maior o nível de confiança, i.e., quanto menor o valor de α , mais rigorosa será a condução do teste. Em contrapartida, quanto menor o nível de confiança, maior será a eficácia do teste, uma vez que se rejeitará mais freqüentemente a hipótese nula.

Para o teste da hipótese nula, calcula-se a estatística F, a partir da razão entre as variâncias ou termos médias de quadrados entre grupos (MQE) e dentro do grupo (MQD), definida como:

$$F = \frac{MQE}{MQD} \quad (\text{B.3})$$

As variâncias MQE e MQD são valores gerados a partir da divisão da soma dos quadrados das diferenças por seus respectivos graus de liberdade. A variação entre grupos, representada pela soma dos quadrados entre grupos (SQE), é definida como a soma dos quadrados das diferenças entre a média aritmética da amostra de cada grupo e a média:

$$SQE = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2 \quad (\text{B.4})$$

em que \bar{X} é a média geral dos valores obtidos em todas as classes, \bar{X}_i corresponde à média aritmética do grupo i , n_i é o número de observações no grupo i e k representa a quantidade de grupos. Visto que k é a quantidade de grupos comparados, a SQE é associada a $k-1$ graus de liberdade (LEVINE, BERENSON e STEPHAN, 2000), logo:

$$MQE = \frac{SQE}{k-1} \quad (\text{B.5})$$

Similarmente, a soma dos quadrados dentro do grupo (SQD) mede a diferença entre cada observação e a média aritmética de seu próprio grupo, acumulando os quadrados dessas diferenças sobre todos os grupos, sendo calculada a partir da equação:

$$SQD = \sum_{j=1}^k \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2, \quad (\text{B.6})$$

em que X_{ij} corresponde à i -ésima observação no grupo j e \bar{X}_j é a média aritmética do grupo j . No cálculo da variação dentro do grupo, cada um dos k níveis contribui em n_j-1 graus de liberdade, resultando em $n-k$ graus de liberdade associados à SQD:

$$\sum_{j=1}^k (n_j - 1) = (n_1 - 1) + (n_2 - 1) + \dots + (n_k - 1) = n - k, \quad (\text{B.7})$$

em que n é o somatório de observações de todos os grupos, i.e., $n = n_1 + n_2 + \dots + n_k$, assim:

$$MQD = \frac{SQD}{n - k} \quad (\text{B.8})$$

Caso o valor da estatística F (Equação B.3) exceda um valor crítico ($F_{\text{crítico}}$) para o nível de confiança escolhido, a hipótese nula H_0 de não existir nenhuma diferença nas médias aritméticas é rejeitada.

Os resultados de um teste ANOVA são geralmente sintetizados em uma tabela reduzida, cujas entradas incluem as fontes de variação

(intergrupos, intragrupos), os *graus de liberdade* (gl), as *somas de quadrados* (SQ), as *médias de quadrados* (MQ) e o valor da *estatística F* calculada (XLSTAT, 2011).

B.4. Teste de Tukey

O teste proposto por *Tukey* é também conhecido como teste de *Tukey* da diferença honestamente significativa (*Honestly Significant Difference* – HSD) e teste de *Tukey* da diferença totalmente significativa (*Wholly Significant Difference* – WSD). O teste de *Tukey* permite verificar as diferenças entre as médias obtidas em vários grupos. Todavia, neste procedimento são realizadas comparações múltiplas, permitindo evidenciar quais os pares de grupos que são diferentes entre si (LEVINE, BERENSON e STEPHAN, 2000).

No teste de *Tukey*, a hipótese nula é aceita para um par de médias (*i* e *j*) quando o módulo da diferença entre estas médias é menor que um valor crítico ($t_{i,j}$), definido por:

$$t_{i,j} = \frac{q \sqrt{MQD \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}{\sqrt{2}} \quad (\text{B.9})$$

em que *MQD* é a variância definida pela Equação B.8, n_i e n_j são a quantidade de elementos do grupo *i* e *j*, respectivamente, e *q* (estatística *Q*) o valor crítico associado ao nível de confiança escolhido para o teste de hipótese.

Os resultados de um teste de *Tukey* são geralmente sintetizados em uma tabela na qual são ordenadas as médias de forma crescente ou decrescente, para facilitar as comparações. Coloca-se uma letra do alfabeto na primeira média (normalmente a letra "A") e, em seguida, compara-se a diferença com as médias seguintes. Se a diferença for superior ao nível de confiança ($1-\alpha$) do teste, a diferença entre duas médias será considerada significativa atribuindo uma nova letra, e.g., "B", a média inferior, caso contrário, atribui-se a mesma letra anterior (XLStat, 2011).

Apêndice C

Resultados Numéricos do Experimento de Calibração

Este apêndice contém os resultados numéricos obtidos pelas abordagens investigadas nas 8 partições de imagens, compostas de 30 imagens cada, de um total de 240 imagens da base de imagens IMM (STEGMANN, ERSBOLL e LARSEN, 2003), além dos resultados dos testes de *Tukey* para as sete métricas de avaliação de agrupamento.

Com base nos dados das Tabelas C.2 a C.9, são fundamentadas a escolha de um grupo de algoritmos a serem usados na abordagem proposta, por meio do teste de *Tukey* representados pelas Tabelas C.10 a C.16. As abordagens investigadas estão listadas na Tabela C.1.

Tabela C.1 – Abordagens investigadas nos testes estatísticos do experimento de calibração.

| Abordagem |
|------------|
| Histograma |
| LBP |
| Máximo |
| Norm(C) |
| Média |
| Correlação |
| Abs(C) |
| Mínimo |

Na sequência, serão listados os resultados obtidos das abordagens investigadas para cada uma das sete métricas de avaliação de agrupamento levantadas na revisão bibliográfica, a saber: (i) *Rand Index* (RI); (ii) *Adjusted Rand Index* (ARI); (iii) *Precision* (P); (iv) *Recall* (R); (v) *F-Measure* (F); (vi) *Jaccard Index* (JI); e (vii) *Folkes and Mallows Index* (FM).

Os resultados numéricos obtidos da abordagem Histograma para as sete métricas de avaliação de agrupamento são sumarizados na Tabela C.2.

Tabela C.2 – Resultados numéricos da abordagem Histograma.

| Resultados Numéricos da Abordagem Histograma | | | | | | | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| Partição 2 | 0,9885 | 0,9586 | 0,9848 | 0,9333 | 0,9655 | 0,9333 | 0,9661 |
| Partição 3 | 0,9563 | 0,8299 | 1,0000 | 0,7467 | 0,8550 | 0,7467 | 0,8641 |
| Partição 4 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| Partição 5 | 0,9609 | 0,8496 | 1,0000 | 0,7733 | 0,8722 | 0,7733 | 0,8794 |
| Partição 6 | 0,9747 | 0,9059 | 1,0000 | 0,8533 | 0,9209 | 0,8533 | 0,9238 |
| Partição 7 | 0,9885 | 0,9586 | 0,9808 | 0,9333 | 0,9655 | 0,9333 | 0,9661 |
| Partição 8 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| Média (\bar{x}) | 0,9836 | 0,9378 | 0,9957 | 0,9050 | 0,9474 | 0,9050 | 0,9499 |
| Variância (s^2) | 0,0003 | 0,0047 | 0,0001 | 0,0105 | 0,0034 | 0,0105 | 0,0030 |
| Desvio Padrão (σ) | 0,0177 | 0,0686 | 0,0080 | 0,1026 | 0,0584 | 0,1026 | 0,0549 |

Os resultados numéricos obtidos da abordagem LBP para as sete métricas de avaliação do agrupamento são sumarizados na Tabela C.3.

Tabela C.3 – Resultados numéricos da abordagem LBP.

| Resultados Numéricos da Abordagem LBP | | | | | | | |
|--|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 0,9885 | 0,9586 | 1,0000 | 0,9333 | 0,9655 | 0,9333 | 0,9661 |
| Partição 2 | 0,9632 | 0,8591 | 0,9833 | 0,7973 | 0,8806 | 0,7867 | 0,8854 |
| Partição 3 | 0,9678 | 0,8780 | 0,9385 | 0,8592 | 0,8971 | 0,8133 | 0,8979 |
| Partição 4 | 0,9563 | 0,8299 | 1,0000 | 0,7467 | 0,8550 | 0,7467 | 0,8641 |
| Partição 5 | 0,9609 | 0,8496 | 1,0000 | 0,7733 | 0,8722 | 0,7733 | 0,8794 |
| Partição 6 | 0,9747 | 0,9059 | 1,0000 | 0,8533 | 0,9209 | 0,8533 | 0,9238 |
| Partição 7 | 0,9747 | 0,9070 | 0,9848 | 0,8667 | 0,9220 | 0,8553 | 0,9239 |
| Partição 8 | 0,9678 | 0,8780 | 0,9385 | 0,8592 | 0,8971 | 0,8133 | 0,8979 |
| Média (\bar{x}) | 0,9693 | 0,8833 | 0,9806 | 0,8361 | 0,9013 | 0,8219 | 0,9048 |
| Variância (s^2) | 0,0001 | 0,0016 | 0,0007 | 0,0036 | 0,0012 | 0,0034 | 0,0010 |
| Desvio Padrão (σ) | 0,0101 | 0,0403 | 0,0269 | 0,0600 | 0,0346 | 0,0585 | 0,0322 |

Os resultados numéricos obtidos da abordagem Máximo para as sete métricas de avaliação do agrupamento são sumarizados na Tabela C.4.

Tabela C.4 – Resultados numéricos da abordagem Máximo.

| Resultados Numéricos da Abordagem Máximo | | | | | | | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 0,9885 | 0,9586 | 1,0000 | 0,9333 | 0,9655 | 0,9333 | 0,9661 |
| Partição 2 | 0,9517 | 0,8142 | 0,9655 | 0,7467 | 0,8421 | 0,7273 | 0,8491 |
| Partição 3 | 0,9471 | 0,7891 | 1,0000 | 0,6933 | 0,8189 | 0,6933 | 0,8327 |
| Partição 4 | 0,9471 | 0,7891 | 1,0000 | 0,6933 | 0,8189 | 0,6933 | 0,8327 |
| Partição 5 | 0,9517 | 0,8142 | 0,9655 | 0,7467 | 0,8421 | 0,7273 | 0,8491 |
| Partição 6 | 0,9425 | 0,7708 | 0,9808 | 0,6800 | 0,8032 | 0,6711 | 0,8167 |
| Partição 7 | 0,9563 | 0,8299 | 1,0000 | 0,7467 | 0,8550 | 0,7467 | 0,8641 |
| Partição 8 | 0,9609 | 0,8496 | 1,0000 | 0,7733 | 0,8722 | 0,7733 | 0,8794 |
| Média (\bar{x}) | 0,9557 | 0,8269 | 0,9890 | 0,7517 | 0,8522 | 0,7457 | 0,8612 |
| Variância (s^2) | 0,0002 | 0,0035 | 0,0003 | 0,0065 | 0,0026 | 0,0068 | 0,0022 |
| Desvio Padrão (σ) | 0,0144 | 0,0588 | 0,0159 | 0,0806 | 0,0508 | 0,0826 | 0,0467 |

Os resultados numéricos obtidos da abordagem Norm(C) para as sete métricas de avaliação do agrupamento são sumarizados na Tabela C.5.

Tabela C.5 – Resultados numéricos da abordagem Norm(C).

| Resultados Numéricos da Abordagem Norm(C) | | | | | | | |
|--|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 0,9885 | 0,9586 | 1,0000 | 0,9333 | 0,9655 | 0,9333 | 0,9661 |
| Partição 2 | 0,9425 | 0,7708 | 0,9808 | 0,6800 | 0,8032 | 0,6711 | 0,8167 |
| Partição 3 | 0,9471 | 0,7891 | 1,0000 | 0,6933 | 0,8189 | 0,6933 | 0,8327 |
| Partição 4 | 0,9563 | 0,8299 | 1,0000 | 0,7467 | 0,8550 | 0,7467 | 0,8641 |
| Partição 5 | 0,9425 | 0,7783 | 0,8710 | 0,7606 | 0,8120 | 0,6835 | 0,8139 |
| Partição 6 | 0,9287 | 0,7088 | 0,9583 | 0,6133 | 0,7480 | 0,5974 | 0,7667 |
| Partição 7 | 0,9402 | 0,7658 | 0,9455 | 0,6933 | 0,8000 | 0,6667 | 0,8096 |
| Partição 8 | 0,9471 | 0,7891 | 1,0000 | 0,6933 | 0,8189 | 0,6933 | 0,8327 |
| Média (\bar{x}) | 0,9491 | 0,7988 | 0,9694 | 0,7267 | 0,8277 | 0,7107 | 0,8378 |
| Variância (s^2) | 0,0003 | 0,0053 | 0,0020 | 0,0089 | 0,0040 | 0,0098 | 0,0034 |
| Desvio Padrão (σ) | 0,0177 | 0,0728 | 0,0451 | 0,0946 | 0,0631 | 0,0990 | 0,0586 |

Os resultados numéricos obtidos da abordagem Média para as sete métricas de avaliação do agrupamento são sumarizados na Tabela C.6.

Tabela C.6 – Resultados numéricos da abordagem Média.

| Resultados Numéricos da Abordagem Média | | | | | | | |
|--|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 0,9747 | 0,9119 | 0,9211 | 0,9333 | 0,9272 | 0,8642 | 0,9272 |
| Partição 2 | 0,9195 | 0,6586 | 0,9762 | 0,5467 | 0,7009 | 0,5395 | 0,7305 |
| Partição 3 | 0,9287 | 0,7088 | 0,9583 | 0,6133 | 0,7480 | 0,5974 | 0,7667 |
| Partição 4 | 0,9425 | 0,7708 | 0,9808 | 0,6800 | 0,8032 | 0,6711 | 0,8167 |
| Partição 5 | 0,9402 | 0,7658 | 0,9455 | 0,6933 | 0,8000 | 0,6667 | 0,8096 |
| Partição 6 | 0,9425 | 0,7783 | 0,8710 | 0,7606 | 0,8120 | 0,6835 | 0,8139 |
| Partição 7 | 0,9218 | 0,6662 | 1,0000 | 0,5467 | 0,7069 | 0,5467 | 0,7394 |
| Partição 8 | 0,9448 | 0,7786 | 1,0000 | 0,6800 | 0,8095 | 0,6800 | 0,8246 |
| Média (\bar{x}) | 0,9394 | 0,7549 | 0,9566 | 0,6817 | 0,7884 | 0,6561 | 0,8036 |
| Variância (s^2) | 0,0003 | 0,0065 | 0,0019 | 0,0158 | 0,0052 | 0,0106 | 0,0038 |
| Desvio Padrão (σ) | 0,0174 | 0,0805 | 0,0438 | 0,1258 | 0,0723 | 0,1028 | 0,0619 |

Os resultados numéricos obtidos da abordagem Correlação para as sete métricas de avaliação do agrupamento são sumarizados na Tabela C.7.

Tabela C.7 – Resultados numéricos da abordagem Correlação.

| Resultados Numéricos da Abordagem Correlação | | | | | | | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 0,8943 | 0,6742 | 0,6436 | 0,8667 | 0,7386 | 0,5856 | 0,7468 |
| Partição 2 | 0,9195 | 0,6586 | 0,9762 | 0,5467 | 0,7009 | 0,5395 | 0,7305 |
| Partição 3 | 0,9287 | 0,7088 | 0,9583 | 0,6133 | 0,7480 | 0,5974 | 0,7667 |
| Partição 4 | 0,9425 | 0,7783 | 0,8710 | 0,7606 | 0,8120 | 0,6835 | 0,8139 |
| Partição 5 | 0,9448 | 0,7786 | 1,0000 | 0,6800 | 0,8095 | 0,6800 | 0,8246 |
| Partição 6 | 0,9333 | 0,7492 | 0,8710 | 0,7200 | 0,7883 | 0,6506 | 0,7919 |
| Partição 7 | 0,9356 | 0,7353 | 1,0000 | 0,6267 | 0,7705 | 0,6267 | 0,7916 |
| Partição 8 | 0,9448 | 0,7786 | 1,0000 | 0,6800 | 0,8095 | 0,6800 | 0,8246 |
| Média (\bar{x}) | 0,9305 | 0,7327 | 0,9150 | 0,6867 | 0,7722 | 0,6304 | 0,7863 |
| Variância (s^2) | 0,0003 | 0,0023 | 0,0150 | 0,0097 | 0,0016 | 0,0028 | 0,0013 |
| Desvio Padrão (σ) | 0,0170 | 0,0479 | 0,1224 | 0,0983 | 0,0404 | 0,0528 | 0,0355 |

Os resultados numéricos obtidos da abordagem Abs(C) para as sete métricas de avaliação do agrupamento são sumarizados na Tabela C.8.

Tabela C.8 – Resultados numéricos da abordagem Abs(C).

| Resultados Numéricos da Abordagem Abs(C) | | | | | | | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 0,8805 | 0,6416 | 0,6075 | 0,8667 | 0,7143 | 0,5556 | 0,7256 |
| Partição 2 | 0,9195 | 0,6586 | 0,9762 | 0,5467 | 0,7009 | 0,5395 | 0,7305 |
| Partição 3 | 0,9287 | 0,7088 | 0,9583 | 0,6133 | 0,7480 | 0,5974 | 0,7667 |
| Partição 4 | 0,9425 | 0,7783 | 0,8710 | 0,7606 | 0,8120 | 0,6835 | 0,8139 |
| Partição 5 | 0,9333 | 0,7492 | 0,8710 | 0,7200 | 0,7883 | 0,6506 | 0,7919 |
| Partição 6 | 0,9287 | 0,7088 | 0,9583 | 0,6133 | 0,7480 | 0,5974 | 0,7667 |
| Partição 7 | 0,9402 | 0,7658 | 0,9455 | 0,6933 | 0,8000 | 0,6667 | 0,8096 |
| Partição 8 | 0,9218 | 0,6662 | 1,0000 | 0,5467 | 0,7069 | 0,5467 | 0,7394 |
| Média (\bar{X}) | 0,9244 | 0,7097 | 0,8985 | 0,6701 | 0,7523 | 0,6047 | 0,7680 |
| Variância (s^2) | 0,0004 | 0,0027 | 0,0160 | 0,0124 | 0,0019 | 0,0032 | 0,0012 |
| Desvio Padrão (σ) | 0,0195 | 0,0515 | 0,1264 | 0,1113 | 0,0436 | 0,0565 | 0,0347 |

Os resultados numéricos obtidos da abordagem Mínimo para as sete métricas de avaliação do agrupamento são sumarizados na Tabela C.9.

Tabela C.9 – Resultados numéricos da abordagem Mínimo.

| Resultados Numéricos da Abordagem Mínimo | | | | | | | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 0,8253 | 0,4841 | 0,4737 | 0,7714 | 0,5870 | 0,4154 | 0,6045 |
| Partição 2 | 0,7517 | 0,4459 | 0,4098 | 1,0000 | 0,5814 | 0,4098 | 0,6402 |
| Partição 3 | 0,9195 | 0,6586 | 0,9762 | 0,5467 | 0,7009 | 0,5395 | 0,7305 |
| Partição 4 | 0,9287 | 0,7088 | 0,9583 | 0,6133 | 0,7480 | 0,5974 | 0,7667 |
| Partição 5 | 0,9425 | 0,7783 | 0,8710 | 0,7606 | 0,8120 | 0,6835 | 0,8139 |
| Partição 6 | 0,9149 | 0,6296 | 1,0000 | 0,5067 | 0,6726 | 0,5067 | 0,7118 |
| Partição 7 | 0,9218 | 0,6662 | 1,0000 | 0,5467 | 0,7069 | 0,5467 | 0,7394 |
| Partição 8 | 0,9080 | 0,6073 | 0,9268 | 0,5067 | 0,6552 | 0,4872 | 0,6853 |
| Média (\bar{X}) | 0,8891 | 0,6223 | 0,8270 | 0,6565 | 0,6830 | 0,5233 | 0,7115 |
| Variância (s^2) | 0,0043 | 0,0122 | 0,0586 | 0,0304 | 0,0060 | 0,0083 | 0,0046 |
| Desvio Padrão (σ) | 0,0659 | 0,1106 | 0,2421 | 0,1744 | 0,0776 | 0,0911 | 0,0675 |

Os resultados numéricos obtidos pelo teste de *Tukey* para a métrica *Rand Index* (RI) das abordagens investigadas são sumarizados na Tabela C.10.

Tabela C.10 – Teste de Tukey para valores de RI das abordagens investigadas.

| Teste de Tukey para métrica RI | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Constraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| Histograma vs Mínimo | 0,095 | 6,765 | 3,148 | < 0,0001 | Sim |
| Histograma vs Abs(C) | 0,059 | 4,236 | 3,148 | 0,002 | Sim |
| Histograma vs Correlação | 0,053 | 3,804 | 3,148 | 0,008 | Sim |
| Histograma vs Média | 0,044 | 3,167 | 3,148 | 0,048 | Sim |
| Histograma vs Norm(C) | 0,034 | 2,468 | 3,148 | 0,231 | Não |
| Histograma vs Máximo | 0,028 | 1,995 | 3,148 | 0,495 | Não |
| Histograma vs LBP | 0,014 | 1,028 | 3,148 | 0,968 | Não |
| LBP vs Mínimo | 0,080 | 5,737 | 3,148 | < 0,0001 | Sim |
| LBP vs Abs(C) | 0,045 | 3,208 | 3,148 | 0,043 | Sim |
| LBP vs Correlação | 0,039 | 2,776 | 3,148 | 0,122 | Não |
| LBP vs Média | 0,030 | 2,139 | 3,148 | 0,404 | Não |
| LBP vs Norm(C) | 0,020 | 1,439 | 3,148 | 0,835 | Não |
| LBP vs Máximo | 0,014 | 0,966 | 3,148 | 0,977 | Não |
| Máximo vs Mínimo | 0,067 | 4,771 | 3,148 | 0,000 | Sim |
| Máximo vs Abs(C) | 0,031 | 2,241 | 3,148 | 0,344 | Não |
| Máximo vs Correlação | 0,025 | 1,810 | 3,148 | 0,616 | Não |
| Máximo vs Média | 0,016 | 1,172 | 3,148 | 0,936 | Não |
| Máximo vs Norm(C) | 0,007 | 0,473 | 3,148 | 1,000 | Não |
| Norm(C) vs Mínimo | 0,060 | 4,298 | 3,148 | 0,002 | Sim |
| Norm(C) vs Abs(C) | 0,025 | 1,768 | 3,148 | 0,643 | Não |
| Norm(C) vs Correlação | 0,019 | 1,337 | 3,148 | 0,881 | Não |
| Norm(C) vs Média | 0,010 | 0,699 | 3,148 | 0,997 | Não |
| Média vs Mínimo | 0,050 | 3,599 | 3,148 | 0,015 | Sim |
| Média vs Abs(C) | 0,015 | 1,069 | 3,148 | 0,961 | Não |
| Média vs Correlação | 0,009 | 0,637 | 3,148 | 0,998 | Não |
| Correlação vs Mínimo | 0,041 | 2,961 | 3,148 | 0,080 | Não |
| Correlação vs Abs(C) | 0,006 | 0,432 | 3,148 | 1,000 | Não |
| Abs(C) vs Mínimo | 0,035 | 2,529 | 3,148 | 0,205 | Não |
| Valor crítico de Tukey: | | | 4,452 | | |

Os resultados numéricos obtidos pelo teste de *Tukey* para a métrica *Adjusted Rand Index* (ARI) das abordagens investigadas são sumarizados na Tabela C.11.

Tabela C.11 – Teste de Tukey para valores de ARI das abordagens investigadas.

| Teste de Tukey para métrica ARI | | | | | $\alpha=0,05$ |
|---------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Constraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| Histograma vs Mínimo | 0,315 | 9,067 | 3,148 | < 0,0001 | Sim |
| Histograma vs Abs(C) | 0,228 | 6,558 | 3,148 | < 0,0001 | Sim |
| Histograma vs Correlação | 0,205 | 5,896 | 3,148 | < 0,0001 | Sim |
| Histograma vs Média | 0,183 | 5,258 | 3,148 | < 0,0001 | Sim |
| Histograma vs Norm(C) | 0,139 | 3,996 | 3,148 | 0,004 | Sim |
| Histograma vs Máximo | 0,111 | 3,187 | 3,148 | 0,045 | Sim |
| Histograma vs LBP | 0,055 | 1,568 | 3,148 | 0,767 | Não |
| LBP vs Mínimo | 0,261 | 7,499 | 3,148 | < 0,0001 | Sim |
| LBP vs Abs(C) | 0,174 | 4,990 | 3,148 | 0,000 | Sim |
| LBP vs Correlação | 0,151 | 4,328 | 3,148 | 0,002 | Sim |
| LBP vs Média | 0,128 | 3,690 | 3,148 | 0,011 | Sim |
| LBP vs Norm(C) | 0,084 | 2,428 | 3,148 | 0,249 | Não |
| LBP vs Máximo | 0,056 | 1,619 | 3,148 | 0,737 | Não |
| Máximo vs Mínimo | 0,205 | 5,880 | 3,148 | < 0,0001 | Sim |
| Máximo vs Abs(C) | 0,117 | 3,370 | 3,148 | 0,028 | Sim |
| Máximo vs Correlação | 0,094 | 2,708 | 3,148 | 0,142 | Não |
| Máximo vs Média | 0,072 | 2,071 | 3,148 | 0,446 | Não |
| Máximo vs Norm(C) | 0,028 | 0,808 | 3,148 | 0,992 | Não |
| Norm(C) vs Mínimo | 0,176 | 5,072 | 3,148 | 0,000 | Sim |
| Norm(C) vs Abs(C) | 0,089 | 2,562 | 3,148 | 0,192 | Não |
| Norm(C) vs Correlação | 0,066 | 1,900 | 3,148 | 0,557 | Não |
| Norm(C) vs Média | 0,044 | 1,263 | 3,148 | 0,909 | Não |
| Média vs Mínimo | 0,133 | 3,809 | 3,148 | 0,008 | Sim |
| Média vs Abs(C) | 0,045 | 1,299 | 3,148 | 0,895 | Não |
| Média vs Correlação | 0,022 | 0,637 | 3,148 | 0,998 | Não |
| Correlação vs Mínimo | 0,110 | 3,172 | 3,148 | 0,047 | Sim |
| Correlação vs Abs(C) | 0,023 | 0,662 | 3,148 | 0,998 | Não |
| Abs(C) vs Mínimo | 0,087 | 2,510 | 3,148 | 0,213 | Não |
| Valor crítico de Tukey: | | | 4,452 | | |

Os resultados numéricos obtidos pelo teste de *Tukey* para a métrica *Precision* (P) das abordagens investigadas são sumarizados na Tabela C.12.

Tabela C.12 – Teste de Tukey para valores de P das abordagens investigadas.

| Teste de Tukey para métrica P | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Constraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| Histograma vs Mínimo | 0,169 | 3,104 | 3,148 | 0,056 | Não |
| Histograma vs Abs(C) | 0,097 | 1,789 | 3,148 | 0,630 | Não |
| Histograma vs Correlação | 0,081 | 1,485 | 3,148 | 0,812 | Não |
| Histograma vs Média | 0,039 | 0,719 | 3,148 | 0,996 | Não |
| Histograma vs Norm(C) | 0,026 | 0,483 | 3,148 | 1,000 | Não |
| Histograma vs LBP | 0,015 | 0,277 | 3,148 | 1,000 | Não |
| Histograma vs Máximo | 0,007 | 0,124 | 3,148 | 1,000 | Não |
| Máximo vs Mínimo | 0,162 | 2,980 | 3,148 | 0,076 | Não |
| Máximo vs Abs(C) | 0,091 | 1,665 | 3,148 | 0,709 | Não |
| Máximo vs Correlação | 0,074 | 1,361 | 3,148 | 0,871 | Não |
| Máximo vs Média | 0,032 | 0,596 | 3,148 | 0,999 | Não |
| Máximo vs Norm(C) | 0,020 | 0,359 | 3,148 | 1,000 | Não |
| Máximo vs LBP | 0,008 | 0,153 | 3,148 | 1,000 | Não |
| LBP vs Mínimo | 0,154 | 2,827 | 3,148 | 0,109 | Não |
| LBP vs Abs(C) | 0,082 | 1,512 | 3,148 | 0,798 | Não |
| LBP vs Correlação | 0,066 | 1,207 | 3,148 | 0,926 | Não |
| LBP vs Média | 0,024 | 0,442 | 3,148 | 1,000 | Não |
| LBP vs Norm(C) | 0,011 | 0,206 | 3,148 | 1,000 | Não |
| Norm(C) vs Mínimo | 0,142 | 2,621 | 3,148 | 0,170 | Não |
| Norm(C) vs Abs(C) | 0,071 | 1,306 | 3,148 | 0,893 | Não |
| Norm(C) vs Correlação | 0,054 | 1,001 | 3,148 | 0,972 | Não |
| Norm(C) vs Média | 0,013 | 0,236 | 3,148 | 1,000 | Não |
| Média vs Mínimo | 0,130 | 2,384 | 3,148 | 0,269 | Não |
| Média vs Abs(C) | 0,058 | 1,069 | 3,148 | 0,961 | Não |
| Média vs Correlação | 0,042 | 0,765 | 3,148 | 0,994 | Não |
| Correlação vs Mínimo | 0,088 | 1,619 | 3,148 | 0,737 | Não |
| Correlação vs Abs(C) | 0,017 | 0,304 | 3,148 | 1,000 | Não |
| Abs(C) vs Mínimo | 0,071 | 1,315 | 3,148 | 0,889 | Não |
| Valor crítico de Tukey: | | | 4,452 | | |

Os resultados numéricos obtidos pelo teste de *Tukey* para a métrica *Recall* (R) das abordagens investigadas são sumarizados na Tabela C.13.

Tabela C.13 – Teste de Tukey para valores de R das abordagens investigadas.

| Teste de Tukey para métrica R | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| Histograma vs Mínimo | 0,249 | 4,493 | 3,148 | 0,001 | Sim |
| Histograma vs Abs(C) | 0,235 | 4,248 | 3,148 | 0,002 | Sim |
| Histograma vs Média | 0,223 | 4,037 | 3,148 | 0,004 | Sim |
| Histograma vs Correlação | 0,218 | 3,946 | 3,148 | 0,005 | Sim |
| Histograma vs Norm(C) | 0,178 | 3,223 | 3,148 | 0,041 | Sim |
| Histograma vs Máximo | 0,153 | 2,772 | 3,148 | 0,002 | Sim |
| Histograma vs LBP | 0,069 | 1,245 | 3,148 | < 0,0001 | Sim |
| LBP vs Mínimo | 0,180 | 3,248 | 3,148 | 0,039 | Sim |
| LBP vs Abs(C) | 0,166 | 3,002 | 3,148 | 0,072 | Não |
| LBP vs Média | 0,154 | 2,791 | 3,148 | 0,118 | Não |
| LBP vs Correlação | 0,149 | 2,701 | 3,148 | 0,144 | Não |
| LBP vs Norm(C) | 0,109 | 1,978 | 3,148 | 0,506 | Não |
| LBP vs Máximo | 0,084 | 1,527 | 3,148 | 0,790 | Não |
| Máximo vs Mínimo | 0,095 | 1,721 | 3,148 | 0,674 | Não |
| Máximo vs Abs(C) | 0,082 | 1,475 | 3,148 | 0,817 | Não |
| Máximo vs Média | 0,070 | 1,264 | 3,148 | 0,908 | Não |
| Máximo vs Correlação | 0,065 | 1,174 | 3,148 | 0,936 | Não |
| Máximo vs Norm(C) | 0,025 | 0,451 | 3,148 | 1,000 | Não |
| Norm(C) vs Mínimo | 0,070 | 1,270 | 3,148 | 0,906 | Não |
| Norm(C) vs Abs(C) | 0,057 | 1,025 | 3,148 | 0,969 | Não |
| Norm(C) vs Média | 0,045 | 0,814 | 3,148 | 0,992 | Não |
| Norm(C) vs Correlação | 0,040 | 0,723 | 3,148 | 0,996 | Não |
| Correlação vs Mínimo | 0,030 | 0,547 | 3,148 | 0,999 | Não |
| Correlação vs Abs(C) | 0,017 | 0,301 | 3,148 | 1,000 | Não |
| Correlação vs Média | 0,005 | 0,090 | 3,148 | 1,000 | Não |
| Média vs Mínimo | 0,025 | 0,456 | 3,148 | 1,000 | Não |
| Média vs Abs(C) | 0,012 | 0,211 | 3,148 | 1,000 | Não |
| Abs(C) vs Mínimo | 0,014 | 0,245 | 3,148 | 1,000 | Não |
| Valor crítico de Tukey: | | | 4,452 | | |

Os resultados numéricos obtidos pelo teste de *Tukey* para a métrica *F-Measure* (F) das abordagens investigadas são sumarizados na Tabela C.14.

Tabela C.14 – Teste de Tukey para valores de F das abordagens investigadas.

| Teste de Tukey para métrica F | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| Histograma vs Mínimo | 0,264 | 9,283 | 3,148 | < 0,0001 | Sim |
| Histograma vs Abs(C) | 0,195 | 6,850 | 3,148 | < 0,0001 | Sim |
| Histograma vs Correlação | 0,175 | 6,152 | 3,148 | < 0,0001 | Sim |
| Histograma vs Média | 0,159 | 5,580 | 3,148 | < 0,0001 | Sim |
| Histograma vs Norm(C) | 0,120 | 4,203 | 3,148 | 0,002 | Sim |
| Histograma vs Máximo | 0,095 | 3,341 | 3,148 | 0,030 | Sim |
| Histograma vs LBP | 0,046 | 1,619 | 3,148 | 0,737 | Não |
| LBP vs Mínimo | 0,218 | 7,665 | 3,148 | < 0,0001 | Sim |
| LBP vs Abs(C) | 0,149 | 5,231 | 3,148 | < 0,0001 | Sim |
| LBP vs Correlação | 0,129 | 4,533 | 3,148 | 0,001 | Sim |
| LBP vs Média | 0,113 | 3,962 | 3,148 | 0,005 | Sim |
| LBP vs Norm(C) | 0,074 | 2,584 | 3,148 | 0,184 | Não |
| LBP vs Máximo | 0,049 | 1,722 | 3,148 | 0,673 | Não |
| Máximo vs Mínimo | 0,169 | 5,942 | 3,148 | < 0,0001 | Sim |
| Máximo vs Abs(C) | 0,100 | 3,509 | 3,148 | 0,019 | Sim |
| Máximo vs Correlação | 0,080 | 2,811 | 3,148 | 0,113 | Não |
| Máximo vs Média | 0,064 | 2,239 | 3,148 | 0,345 | Não |
| Máximo vs Norm(C) | 0,025 | 0,862 | 3,148 | 0,988 | Não |
| Norm(C) vs Mínimo | 0,145 | 5,080 | 3,148 | 0,000 | Sim |
| Norm(C) vs Abs(C) | 0,075 | 2,647 | 3,148 | 0,161 | Não |
| Norm(C) vs Correlação | 0,056 | 1,949 | 3,148 | 0,524 | Não |
| Norm(C) vs Média | 0,039 | 1,377 | 3,148 | 0,864 | Não |
| Média vs Mínimo | 0,105 | 3,703 | 3,148 | 0,011 | Sim |
| Média vs Abs(C) | 0,036 | 1,269 | 3,148 | 0,906 | Não |
| Média vs Correlação | 0,016 | 0,572 | 3,148 | 0,999 | Não |
| Correlação vs Mínimo | 0,089 | 3,131 | 3,148 | 0,052 | Não |
| Correlação vs Abs(C) | 0,020 | 0,698 | 3,148 | 0,997 | Não |
| Abs(C) vs Mínimo | 0,069 | 2,434 | 3,148 | 0,246 | Não |
| Valor crítico de Tukey: | | | 4,452 | | |

Os resultados numéricos obtidos pelo teste de *Tukey* para a métrica *Jaccard Index* (JI) das abordagens investigadas são sumarizados na Tabela C.15.

Tabela C.15 – Teste de Tukey para valores de JI das abordagens investigadas.

| Teste de Tukey para métrica JI | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| Histograma vs Mínimo | 0,382 | 9,174 | 3,148 | < 0,0001 | Sim |
| Histograma vs Abs(C) | 0,300 | 7,218 | 3,148 | < 0,0001 | Sim |
| Histograma vs Correlação | 0,275 | 6,599 | 3,148 | < 0,0001 | Sim |
| Histograma vs Média | 0,249 | 5,981 | 3,148 | < 0,0001 | Sim |
| Histograma vs Norm(C) | 0,194 | 4,670 | 3,148 | 0,000 | Sim |
| Histograma vs Máximo | 0,159 | 3,828 | 3,148 | 0,007 | Sim |
| Histograma vs LBP | 0,083 | 1,997 | 3,148 | < 0,0001 | Sim |
| LBP vs Mínimo | 0,299 | 7,177 | 3,148 | < 0,0001 | Sim |
| LBP vs Abs(C) | 0,217 | 5,221 | 3,148 | < 0,0001 | Sim |
| LBP vs Correlação | 0,191 | 4,602 | 3,148 | 0,001 | Sim |
| LBP vs Média | 0,166 | 3,984 | 3,148 | 0,005 | Sim |
| LBP vs Norm(C) | 0,111 | 2,673 | 3,148 | 0,153 | Não |
| LBP vs Máximo | 0,076 | 1,831 | 3,148 | 0,602 | Não |
| Máximo vs Mínimo | 0,222 | 5,346 | 3,148 | < 0,0001 | Sim |
| Máximo vs Abs(C) | 0,141 | 3,389 | 3,148 | 0,026 | Sim |
| Máximo vs Correlação | 0,115 | 2,771 | 3,148 | 0,124 | Não |
| Máximo vs Média | 0,090 | 2,153 | 3,148 | 0,395 | Não |
| Máximo vs Norm(C) | 0,035 | 0,842 | 3,148 | 0,990 | Não |
| Norm(C) vs Mínimo | 0,187 | 4,504 | 3,148 | 0,001 | Sim |
| Norm(C) vs Abs(C) | 0,106 | 2,547 | 3,148 | 0,198 | Não |
| Norm(C) vs Correlação | 0,080 | 1,929 | 3,148 | 0,538 | Não |
| Norm(C) vs Média | 0,055 | 1,311 | 3,148 | 0,891 | Não |
| Média vs Mínimo | 0,133 | 3,193 | 3,148 | 0,045 | Sim |
| Média vs Abs(C) | 0,051 | 1,237 | 3,148 | 0,917 | Não |
| Média vs Correlação | 0,026 | 0,618 | 3,148 | 0,998 | Não |
| Correlação vs Mínimo | 0,107 | 2,575 | 3,148 | 0,187 | Não |
| Correlação vs Abs(C) | 0,026 | 0,619 | 3,148 | 0,998 | Não |
| Abs(C) vs Mínimo | 0,081 | 1,956 | 3,148 | 0,520 | Não |
| Valor crítico de Tukey: | | | 4,452 | | |

Os resultados numéricos obtidos pelo teste de *Tukey* para a métrica *Folkes and Mallows* (FM) das abordagens investigadas são sumarizados na Tabela C.16.

Tabela C.16 – Teste de Tukey para valores de FM das abordagens investigadas.

| Teste de Tukey para métrica FM | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| Histograma vs Mínimo | 0,238 | 9,414 | 3,148 | < 0,0001 | Sim |
| Histograma vs Abs(C) | 0,182 | 7,183 | 3,148 | < 0,0001 | Sim |
| Histograma vs Correlação | 0,164 | 6,460 | 3,148 | < 0,0001 | Sim |
| Histograma vs Média | 0,146 | 5,780 | 3,148 | < 0,0001 | Sim |
| Histograma vs Norm(C) | 0,112 | 4,428 | 3,148 | 0,001 | Sim |
| Histograma vs Máximo | 0,089 | 3,503 | 3,148 | 0,019 | Sim |
| Histograma vs LBP | 0,045 | 1,782 | 3,148 | 0,635 | Não |
| LBP vs Mínimo | 0,193 | 7,633 | 3,148 | < 0,0001 | Sim |
| LBP vs Abs(C) | 0,137 | 5,402 | 3,148 | < 0,0001 | Sim |
| LBP vs Correlação | 0,118 | 4,679 | 3,148 | 0,000 | Sim |
| LBP vs Média | 0,101 | 3,998 | 3,148 | 0,004 | Sim |
| LBP vs Norm(C) | 0,067 | 2,647 | 3,148 | 0,161 | Não |
| LBP vs Máximo | 0,044 | 1,722 | 3,148 | 0,673 | Não |
| Máximo vs Mínimo | 0,150 | 5,911 | 3,148 | < 0,0001 | Sim |
| Máximo vs Abs(C) | 0,093 | 3,680 | 3,148 | 0,012 | Sim |
| Máximo vs Correlação | 0,075 | 2,957 | 3,148 | 0,080 | Não |
| Máximo vs Média | 0,058 | 2,276 | 3,148 | 0,325 | Não |
| Máximo vs Norm(C) | 0,023 | 0,925 | 3,148 | 0,982 | Não |
| Norm(C) vs Mínimo | 0,126 | 4,986 | 3,148 | 0,000 | Sim |
| Norm(C) vs Abs(C) | 0,070 | 2,755 | 3,148 | 0,128 | Não |
| Norm(C) vs Correlação | 0,051 | 2,032 | 3,148 | 0,470 | Não |
| Norm(C) vs Média | 0,034 | 1,352 | 3,148 | 0,875 | Não |
| Média vs Mínimo | 0,092 | 3,635 | 3,148 | 0,013 | Sim |
| Média vs Abs(C) | 0,036 | 1,403 | 3,148 | 0,852 | Não |
| Média vs Correlação | 0,017 | 0,681 | 3,148 | 0,997 | Não |
| Correlação vs Mínimo | 0,075 | 2,954 | 3,148 | 0,081 | Não |
| Correlação vs Abs(C) | 0,018 | 0,723 | 3,148 | 0,996 | Não |
| Abs(C) vs Mínimo | 0,057 | 2,231 | 3,148 | 0,350 | Não |
| Valor crítico de Tukey: | | | 4,452 | | |

Apêndice D

Resultados Numéricos do Experimento de Comparação

Este apêndice contém os resultados numéricos obtidos pelas ferramentas concorrentes soma a abordagem proposta, denominada *PhotoCluster*, nas 8 partições de imagens, compostas de 30 imagens cada, de um total de 240 imagens da base de imagens IMM (STEGMANN, ERSBOLL e LARSEN, 2003), além dos resultados dos testes de *Tukey* para as sete métricas de avaliação de agrupamento.

Com base nos dados das Tabelas D.2 a D.5, são fundamentadas a escolha de um grupo de algoritmos classificados como a melhor escolha, por meio do teste de *Tukey* representados pelas Tabelas D.6 a D.12. As ferramentas comparadas estão listadas na Tabela D.1.

Tabela D.1 – Ferramentas comparadas nos testes estatísticos do experimento de calibração.

| Abordagem |
|--|
| Adobe Photoshop Elements (Photoshopel) |
| Google Picasa |
| Pittpatt FaceSort |
| PhotoCluster |

Na sequência, serão listados os resultados obtidos das abordagens comparadas para cada uma das sete métricas de avaliação de agrupamento levantadas na revisão bibliográfica, a saber: (i) *Rand Index* (RI); (ii) *Adjusted Rand Index* (ARI); (iii) *Precision* (P); (iv) *Recall* (R); (v) *F-Measure* (F); (vi) *Jaccard Index* (JI); e (vii) *Folkes and Mallows Index* (FM).

Os resultados numéricos obtidos da ferramenta *Photoshopel* para as sete métricas de avaliação do agrupamento são sumarizados na Tabela D.2.

Tabela D.2 – Resultados numéricos da ferramenta Photoshopel.

| Resultados Numéricos da Ferramenta Photoshopel | | | | | | | |
|--|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 0,9471 | 0,7891 | 1,0000 | 0,6933 | 0,8189 | 0,6933 | 0,8327 |
| Partição 2 | 0,9333 | 0,7492 | 0,8710 | 0,7200 | 0,7883 | 0,6506 | 0,7919 |
| Partição 3 | 0,9195 | 0,6653 | 0,8431 | 0,6143 | 0,7107 | 0,5513 | 0,7197 |
| Partição 4 | 0,9356 | 0,7353 | 1,0000 | 0,6267 | 0,7705 | 0,6267 | 0,7916 |
| Partição 5 | 0,9218 | 0,6662 | 1,0000 | 0,5467 | 0,7069 | 0,5467 | 0,7394 |
| Partição 6 | 0,9149 | 0,6296 | 1,0000 | 0,5067 | 0,6726 | 0,5067 | 0,7118 |
| Partição 7 | 0,9471 | 0,7891 | 1,0000 | 0,6933 | 0,8189 | 0,6933 | 0,8327 |
| Partição 8 | 0,9080 | 0,6073 | 0,9268 | 0,5067 | 0,6552 | 0,4872 | 0,6853 |
| Média (\bar{x}) | 0,9284 | 0,7039 | 0,9551 | 0,6135 | 0,7427 | 0,5945 | 0,7631 |
| Variância (s^2) | 0,0002 | 0,0050 | 0,0044 | 0,0074 | 0,0042 | 0,0067 | 0,0032 |
| Desvio Padrão (σ) | 0,0146 | 0,0710 | 0,0660 | 0,0858 | 0,0648 | 0,0820 | 0,0566 |

Os resultados numéricos obtidos da ferramenta *Picasa* para as sete métricas de avaliação do agrupamento são sumarizados na Tabela D.3.

Tabela D.3 – Resultados numéricos da ferramenta Picasa.

| Resultados Numéricos da Ferramenta Picasa | | | | | | | |
|--|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| Partição 2 | 0,9816 | 0,9327 | 1,0000 | 0,8933 | 0,9437 | 0,8933 | 0,9452 |
| Partição 3 | 0,9126 | 0,6118 | 0,7708 | 0,5781 | 0,6607 | 0,4933 | 0,6676 |
| Partição 4 | 0,9678 | 0,8780 | 0,9385 | 0,8592 | 0,8971 | 0,8133 | 0,8979 |
| Partição 5 | 0,9425 | 0,7783 | 0,8710 | 0,7606 | 0,8120 | 0,6835 | 0,8139 |
| Partição 6 | 0,9632 | 0,8591 | 0,9833 | 0,7973 | 0,8806 | 0,7867 | 0,8854 |
| Partição 7 | 0,9885 | 0,9586 | 1,0000 | 0,9333 | 0,9655 | 0,9333 | 0,9661 |
| Partição 8 | 0,9494 | 0,7995 | 1,0000 | 0,7067 | 0,8281 | 0,7067 | 0,8406 |
| Média (\bar{x}) | 0,9632 | 0,8523 | 0,9454 | 0,8161 | 0,8735 | 0,7888 | 0,8771 |
| Variância (s^2) | 0,0008 | 0,0152 | 0,0071 | 0,0182 | 0,0116 | 0,0260 | 0,0111 |
| Desvio Padrão (σ) | 0,0281 | 0,1233 | 0,0840 | 0,1349 | 0,1076 | 0,1613 | 0,1051 |

Os resultados numéricos obtidos da ferramenta *FaceSort* para as sete métricas de avaliação do agrupamento são sumarizados na Tabela D.4.

Tabela D.4 – Resultados numéricos da ferramenta FaceSort.

| Resultados Numéricos da Ferramenta FaceSort | | | | | | | |
|--|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 0,7517 | 0,4459 | 0,4098 | 1,0000 | 0,5814 | 0,4098 | 0,6402 |
| Partição 2 | 0,9747 | 0,9070 | 0,9848 | 0,8667 | 0,9220 | 0,8553 | 0,9239 |
| Partição 3 | 0,9425 | 0,7708 | 0,9808 | 0,6800 | 0,8032 | 0,6711 | 0,8167 |
| Partição 4 | 0,9287 | 0,7088 | 0,9583 | 0,6133 | 0,7480 | 0,5974 | 0,7667 |
| Partição 5 | 0,9195 | 0,6586 | 0,9762 | 0,5467 | 0,7009 | 0,5395 | 0,7305 |
| Partição 6 | 0,9402 | 0,7658 | 0,9455 | 0,6933 | 0,8000 | 0,6667 | 0,8096 |
| Partição 7 | 0,9517 | 0,8142 | 0,9655 | 0,7467 | 0,8421 | 0,7273 | 0,8491 |
| Partição 8 | 0,9448 | 0,7786 | 1,0000 | 0,6800 | 0,8095 | 0,6800 | 0,8246 |
| Média (\bar{x}) | 0,9193 | 0,7312 | 0,9026 | 0,7283 | 0,7759 | 0,6434 | 0,7952 |
| Variância (s^2) | 0,0048 | 0,0186 | 0,0399 | 0,0208 | 0,0103 | 0,0174 | 0,0071 |
| Desvio Padrão (σ) | 0,0696 | 0,1363 | 0,1998 | 0,1441 | 0,1017 | 0,1320 | 0,0845 |

Os resultados numéricos obtidos da abordagem proposta *PhotoCluster* para as sete métricas de avaliação do agrupamento são sumarizados na Tabela D.5.

Tabela D.5 – Resultados numéricos da abordagem proposta PhotoCluster.

| Resultados Numéricos da Abordagem Proposta PhotoCluster | | | | | | | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Imagens | RI | ARI | P | R | F | JI | FM |
| Partição 1 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| Partição 2 | 0,9885 | 0,9586 | 0,9848 | 0,9333 | 0,9655 | 0,9333 | 0,9661 |
| Partição 3 | 0,9563 | 0,8299 | 1,0000 | 0,7467 | 0,8550 | 0,7467 | 0,8641 |
| Partição 4 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| Partição 5 | 0,9609 | 0,8496 | 1,0000 | 0,7733 | 0,8722 | 0,7733 | 0,8794 |
| Partição 6 | 0,9747 | 0,9059 | 1,0000 | 0,8533 | 0,9209 | 0,8533 | 0,9238 |
| Partição 7 | 0,9885 | 0,9586 | 0,9808 | 0,9333 | 0,9655 | 0,9333 | 0,9661 |
| Partição 8 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| Média (\bar{x}) | 0,9836 | 0,9378 | 0,9957 | 0,9050 | 0,9474 | 0,9050 | 0,9499 |
| Variância (s^2) | 0,0003 | 0,0047 | 0,0001 | 0,0105 | 0,0034 | 0,0105 | 0,0030 |
| Desvio Padrão (σ) | 0,0177 | 0,0686 | 0,0080 | 0,1026 | 0,0584 | 0,1026 | 0,0549 |

Os resultados obtidos pelo teste de *Tukey* para a métrica RI das ferramentas comparadas são sumarizados na Tabela D.6.

Tabela D.6 – Teste de Tukey para valores de RI das ferramentas comparadas.

| Teste de Tukey para métrica RI | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| PhotoCluster vs FaceSort | 0,064 | 3,280 | 2,730 | 0,014 | Sim |
| PhotoCluster vs Photoshopel | 0,055 | 2,811 | 2,730 | 0,042 | Sim |
| PhotoCluster vs Picasa | 0,020 | 1,040 | 2,730 | 0,728 | Não |
| Picasa vs FaceSort | 0,044 | 2,240 | 2,730 | 0,137 | Não |
| Picasa vs Photoshopel | 0,035 | 1,772 | 2,730 | 0,308 | Não |
| Photoshopel vs FaceSort | 0,009 | 0,469 | 2,730 | 0,965 | Não |
| Valor crítico de Tukey: | | | 3,861 | | |

Os resultados obtidos pelo teste de *Tukey* para a métrica ARI das ferramentas comparadas são sumarizados na Tabela D.7.

Tabela D.7 – Teste de Tukey para valores de ARI das ferramentas comparadas.

| Teste de Tukey para métrica ARI | | | | | $\alpha=0,05$ |
|---------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| PhotoCluster vs FaceSort | 0,234 | 4,485 | 2,730 | 0,001 | Sim |
| PhotoCluster vs Photoshopel | 0,207 | 3,962 | 2,730 | 0,002 | Sim |
| PhotoCluster vs Picasa | 0,086 | 1,641 | 2,730 | < 0,0001 | Sim |
| Picasa vs FaceSort | 0,148 | 2,844 | 2,730 | 0,039 | Sim |
| Picasa vs Photoshopel | 0,121 | 2,321 | 2,730 | 0,117 | Não |
| Photoshopel vs FaceSort | 0,027 | 0,523 | 2,730 | 0,953 | Não |
| Valor crítico de Tukey: | | | 3,861 | | |

Os resultados obtidos pelo teste de *Tukey* para a métrica P das ferramentas comparadas são sumarizados na Tabela D.8.

Tabela D.8 – Teste de Tukey para valores de P das ferramentas comparadas.

| Teste de Tukey para métrica P | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| PhotoCluster vs FaceSort | 0,093 | 1,642 | 2,730 | 0,372 | Não |
| PhotoCluster vs Photoshopel | 0,050 | 0,887 | 2,730 | 0,812 | Não |
| PhotoCluster vs Picasa | 0,041 | 0,716 | 2,730 | 0,890 | Não |
| Picasa vs FaceSort | 0,052 | 0,926 | 2,730 | 0,791 | Não |
| Picasa vs Photoshopel | 0,010 | 0,171 | 2,730 | 0,998 | Não |
| Photoshopel vs FaceSort | 0,043 | 0,756 | 2,730 | 0,873 | Não |
| Valor crítico de Tukey: | | | 3,861 | | |

Os resultados obtidos pelo teste de *Tukey* para a métrica R das ferramentas comparadas são sumarizados na Tabela D.9.

Tabela D.9 – Teste de Tukey para valores de R das ferramentas comparadas.

| Teste de Tukey para métrica R | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| PhotoCluster vs FaceSort | 0,292 | 4,890 | 2,730 | 0,000 | Sim |
| PhotoCluster vs Photoshopel | 0,177 | 2,963 | 2,730 | 0,030 | Sim |
| PhotoCluster vs Picasa | 0,089 | 1,492 | 2,730 | < 0,0001 | Sim |
| Picasa vs FaceSort | 0,203 | 3,398 | 2,730 | 0,010 | Sim |
| Picasa vs Photoshopel | 0,088 | 1,471 | 2,730 | 0,468 | Não |
| Photoshopel vs FaceSort | 0,115 | 1,927 | 2,730 | 0,240 | Não |
| Valor crítico de Tukey: | | | 3,861 | | |

Os resultados obtidos pelo teste de *Tukey* para a métrica F das ferramentas comparadas são sumarizados na Tabela D.10.

Tabela D.10 – Teste de Tukey para valores de F das ferramentas comparadas.

| Teste de Tukey para métrica F | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| PhotoCluster vs FaceSort | 0,205 | 4,763 | 2,730 | 0,000 | Sim |
| PhotoCluster vs Photoshopel | 0,172 | 3,992 | 2,730 | 0,002 | Sim |
| PhotoCluster vs Picasa | 0,074 | 1,720 | 2,730 | 0,332 | Não |
| Picasa vs FaceSort | 0,131 | 3,042 | 2,730 | 0,025 | Sim |
| Picasa vs Photoshopel | 0,098 | 2,271 | 2,730 | 0,129 | Não |
| Photoshopel vs FaceSort | 0,033 | 0,771 | 2,730 | 0,867 | Não |
| Valor crítico de Tukey: | | | 3,861 | | |

Os resultados obtidos pelo teste de *Tukey* para a métrica JI das ferramentas comparadas são sumarizados na Tabela D.11.

Tabela D.11 – Teste de Tukey para valores de JI das ferramentas comparadas.

| Teste de Tukey para métrica JI | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| PhotoCluster vs FaceSort | 0,311 | 5,041 | 2,730 | 0,000 | Sim |
| PhotoCluster vs Photoshopel | 0,262 | 4,247 | 2,730 | 0,001 | Sim |
| PhotoCluster vs Picasa | 0,116 | 1,887 | 2,730 | < 0,0001 | Sim |
| Picasa vs FaceSort | 0,194 | 3,154 | 2,730 | 0,019 | Sim |
| Picasa vs Photoshopel | 0,145 | 2,360 | 2,730 | 0,109 | Não |
| Photoshopel vs FaceSort | 0,049 | 0,794 | 2,730 | 0,857 | Não |
| Valor crítico de Tukey: | | | 3,861 | | |

Os resultados obtidos pelo teste de *Tukey* para a métrica FM das ferramentas comparadas são sumarizados na Tabela D.12.

Tabela D.12 – Teste de Tukey para valores de FM das ferramentas comparadas.

| Teste de Tukey para métrica FM | | | | | $\alpha=0,05$ |
|--------------------------------|-----------|-----------------------|---------------|----------|---------------|
| Contraste | Diferença | Diferença Padronizada | Valor crítico | Pr > Dif | Significante |
| PhotoCluster vs FaceSort | 0,187 | 4,782 | 2,730 | 0,000 | Sim |
| PhotoCluster vs Photoshopel | 0,155 | 3,962 | 2,730 | 0,002 | Sim |
| PhotoCluster vs Picasa | 0,073 | 1,864 | 2,730 | 0,266 | Não |
| Picasa vs FaceSort | 0,114 | 2,917 | 2,730 | 0,033 | Sim |
| Picasa vs Photoshopel | 0,082 | 2,097 | 2,730 | 0,179 | Não |
| Photoshopel vs FaceSort | 0,032 | 0,820 | 2,730 | 0,845 | Não |
| Valor crítico de Tukey: | | | 3,861 | | |

Apêndice E

Resultados Detalhados Obtidos a partir do Uso da Abordagem Proposta

Este apêndice contém um detalhamento dos resultados obtidos a partir do uso da abordagem proposta para a solução do problema de agrupamento de faces em coleções de fotografias digitais na base de imagens IMM (STEGMANN, ERSBOLL e LARSEN, 2003), na qual foi subdividida em 8 partições de imagens, compostas de 30 imagens cada, de um total de 240 imagens obtidas de 40 indivíduos, sendo com 6 imagens de amostra por indivíduo.

Os dendrogramas correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta são ilustrados a seguir. A linha pontilhada mais a direita do gráfico representa o limiar de corte do dendrograma, limiar este que define a composição final dos grupos a partir do nível de dissimilaridade entre os grupos. Quando o aumento no nível de dissimilaridade é forte, ou seja, o limiar de corte é definido quando o intervalo entre duas combinações sucessivas de grupos é a maior.

A adição de mais um grupo diminui significativamente a qualidade do grupo. Portanto, o limiar de corte é definido no nível anterior a este ponto de diminuição da qualidade do agrupamento. Desta forma, chega-se a um nível em que agrupam-se grupos que já estão homogêneos, assim, o truncamento automático, usa este critério para decidir quando parar de agregar as observações (ou os grupos de observações). Nesta dissertação este limiar foi configurado para ser determinado automaticamente pela aplicação *XLStat* (XLSTAT, 2011).

E.1. Resultados da Partição 1

Resultado visual dos grupos obtidos a partir do uso da abordagem proposta na Partição 1.

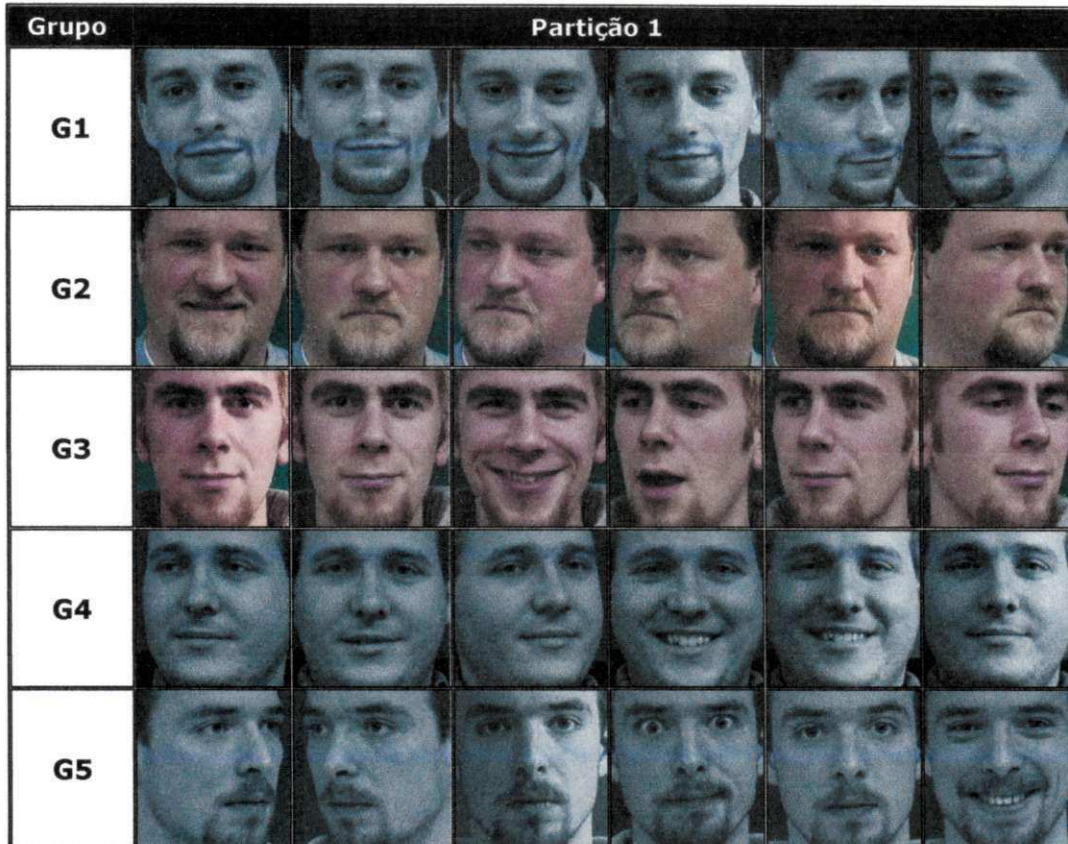


Figura E.1 – Grupos de faces obtidos na Partição 1.

Diagrama dos níveis de similaridade correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 1.

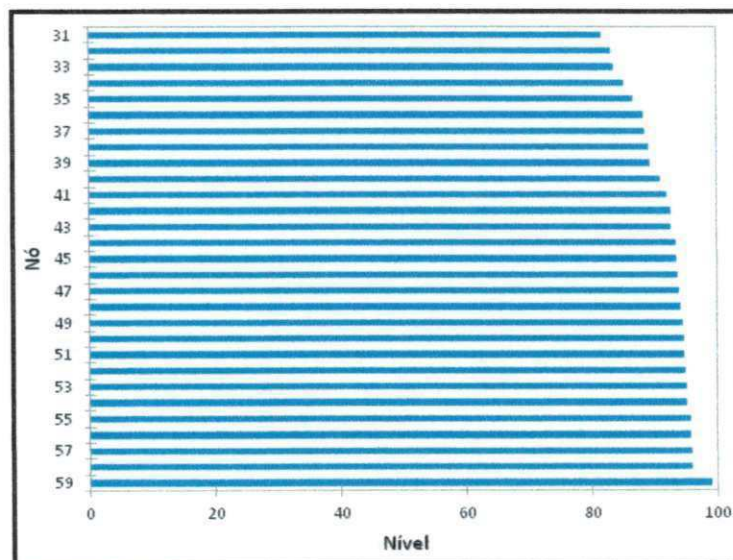


Figura E.2 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 1.

Dendrograma correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 1.

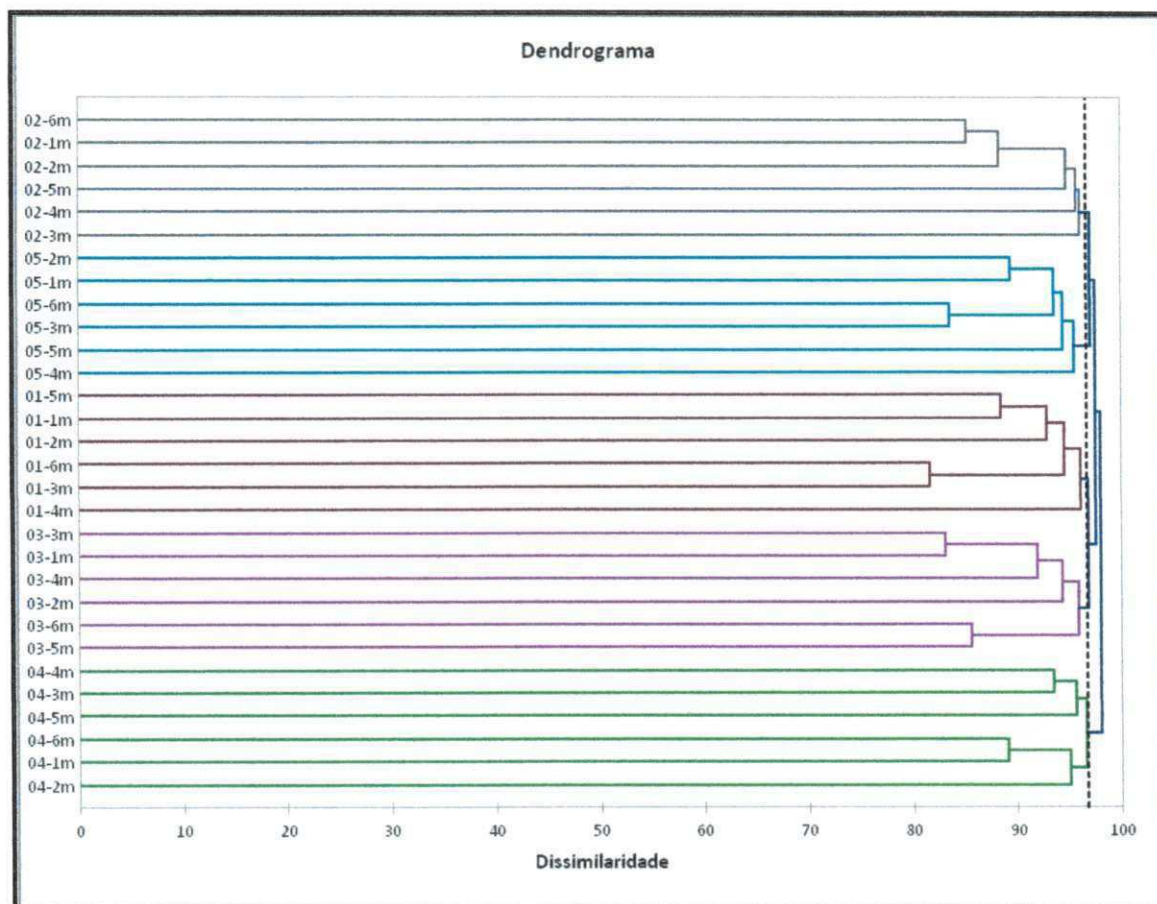


Figura E.3 – Dendrograma dos grupos de faces obtidos na Partição 1.

Resultados por classe correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 1. Os objetos ilustrados na tabela representam os nomes das imagens da base e os dois primeiros dígitos (em negrito) representam o indivíduo correspondente.

Tabela E.1 – Resultados por classe dos grupos de faces obtidos na Partição 1.

| Classe | Objeto |
|---------------|--|
| 1 (G1) | 02 -6m.jpg 02 -1m.jpg 02 -2m.jpg 02 -5m.jpg 02 -4m.jpg 02 -3m.jpg |
| 2 (G2) | 05 -2m.jpg 05 -1m.jpg 05 -6m.jpg 05 -3m.jpg 05 -5m.jpg 05 -4m.jpg |
| 3 (G3) | 01 -5m.jpg 01 -1m.jpg 01 -2m.jpg 01 -6m.jpg 01 -3m.jpg 01 -4m.jpg |
| 4 (G4) | 03 -3m.jpg 03 -1m.jpg 03 -4m.jpg 03 -2m.jpg 03 -6m.jpg 03 -5m.jpg |
| 5 (G5) | 04 -4m.jpg 04 -3m.jpg 04 -5m.jpg 04 -6m.jpg 04 -1m.jpg 04 -2m.jpg |

E.2. Resultados da Partição 2

Resultado visual dos grupos obtidos a partir do uso da abordagem proposta na Partição 2.

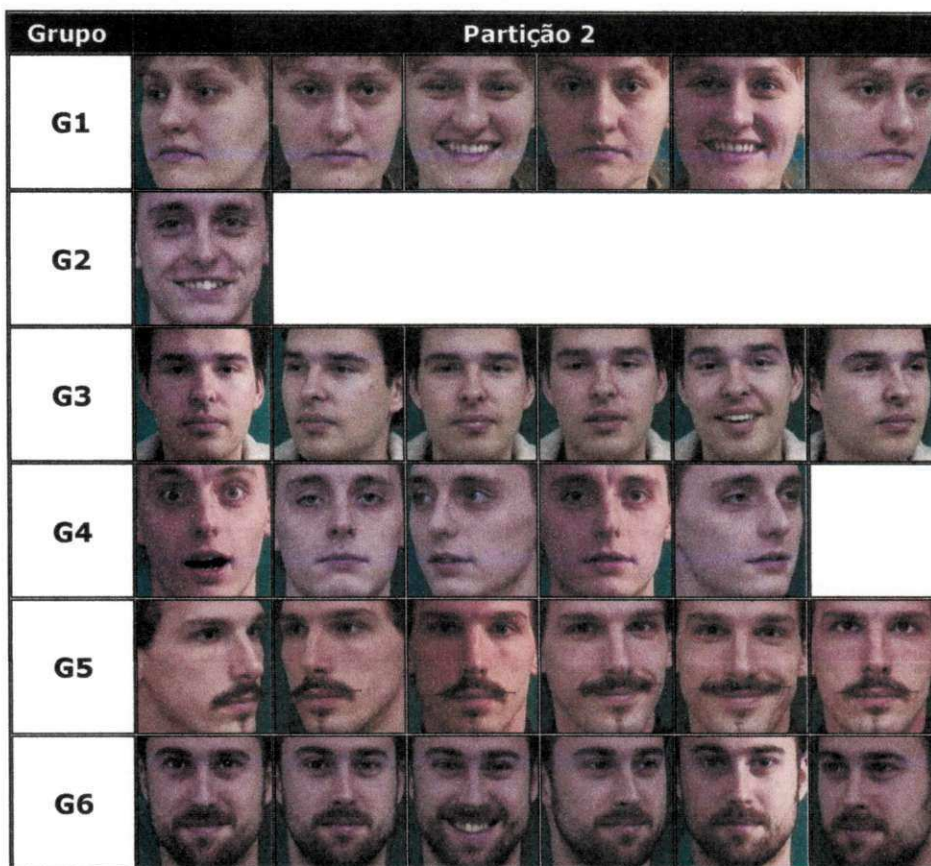


Figura E.4 – Grupos de faces obtidos na Partição 2.

Diagrama dos níveis de similaridade correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 2.

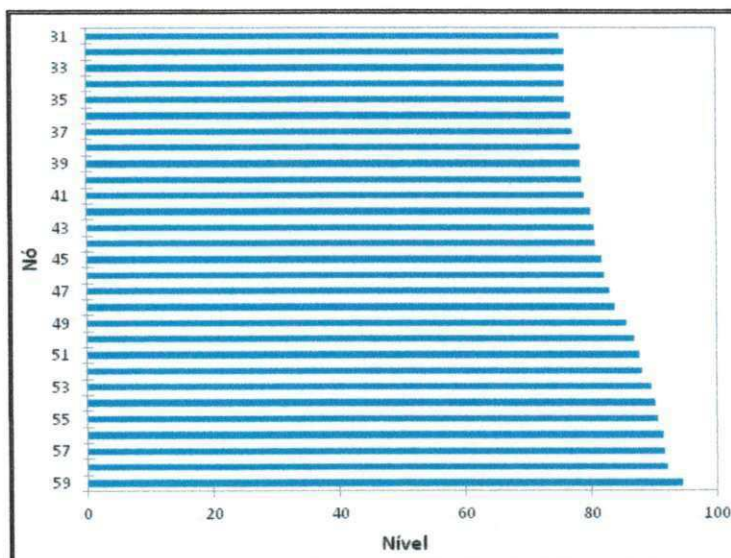


Figura E.5 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 2.

Dendrograma correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 2.

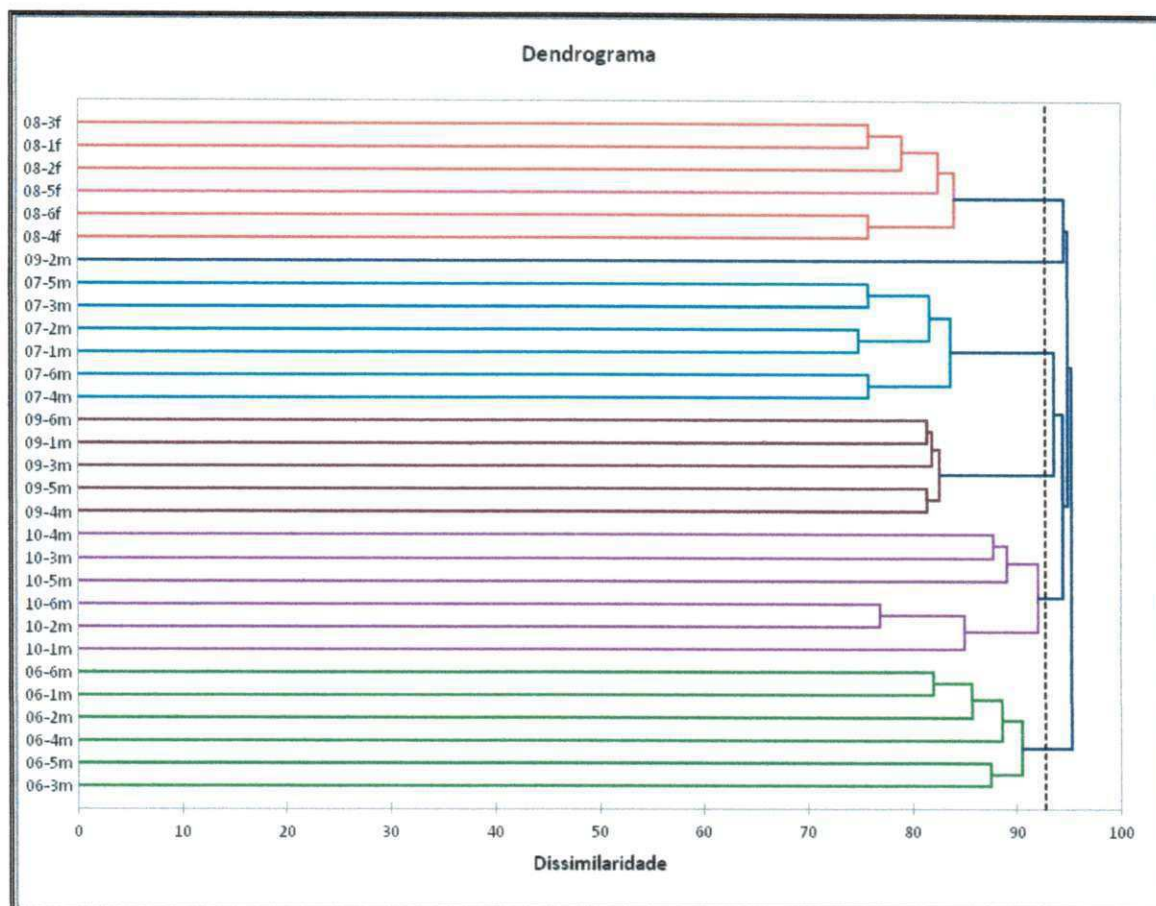


Figura E.6 – Dendrograma dos grupos de faces obtidos na Partição 2.

Resultados por classe correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 2. Os objetos ilustrados na tabela representam os nomes das imagens da base e os dois primeiros dígitos (em negrito) representam o indivíduo correspondente.

Tabela E.2 – Resultados por classe dos grupos de faces obtidos na Partição 2.

| Classe | Objeto |
|---------------|--|
| 1 (G1) | 08-3f.jpg 08-1f.jpg 08-2f.jpg 08-5f.jpg 08-6f.jpg 08-4f.jpg |
| 2 (G2) | 09-2m.jpg |
| 3 (G3) | 07-5m.jpg 07-3m.jpg 07-2m.jpg 07-1m.jpg 07-6m.jpg 07-4m.jpg |
| 4 (G4) | 09-6m.jpg 09-1m.jpg 09-3m.jpg 09-5m.jpg 09-4m.jpg |
| 5 (G5) | 10-4m.jpg 10-3m.jpg 10-5m.jpg 10-6m.jpg 10-2m.jpg 04-1m.jpg |
| 6 (G6) | 06-6m.jpg 06-1m.jpg 06-2m.jpg 06-4m.jpg 06-5m.jpg 06-3m.jpg |

E.3. Resultados da Partição 3

Resultado visual dos grupos obtidos a partir do uso da abordagem proposta na Partição 3.

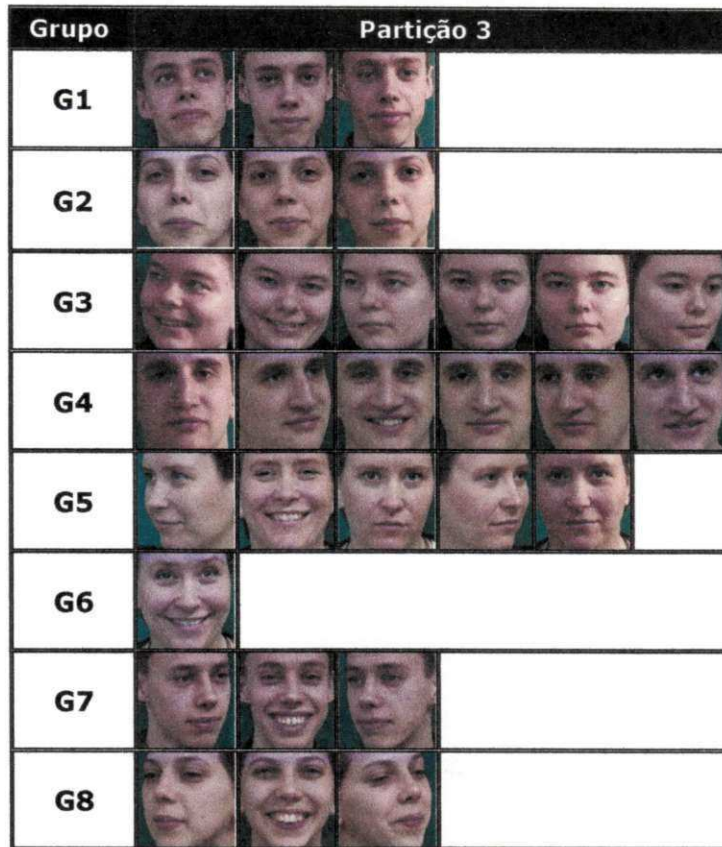


Figura E.7 – Grupos de faces obtidos na Partição 3.

Diagrama dos níveis de similaridade correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 3.

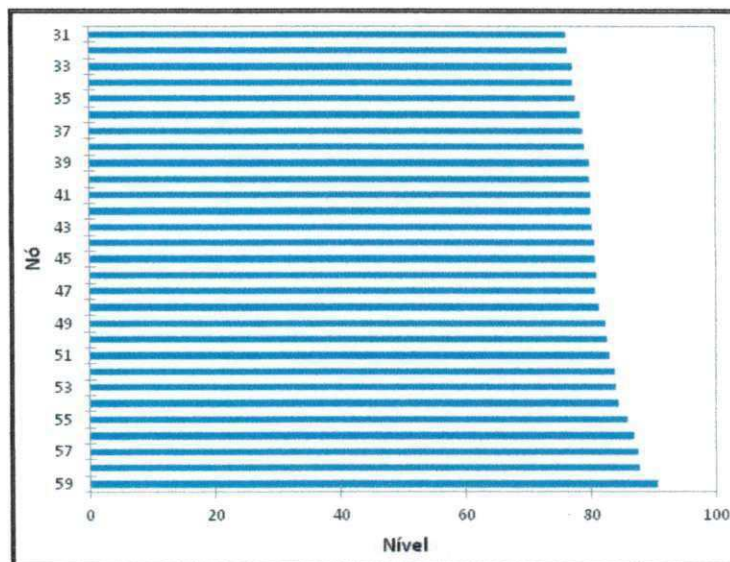


Figura E.8 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 3.

Dendrograma correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 3.

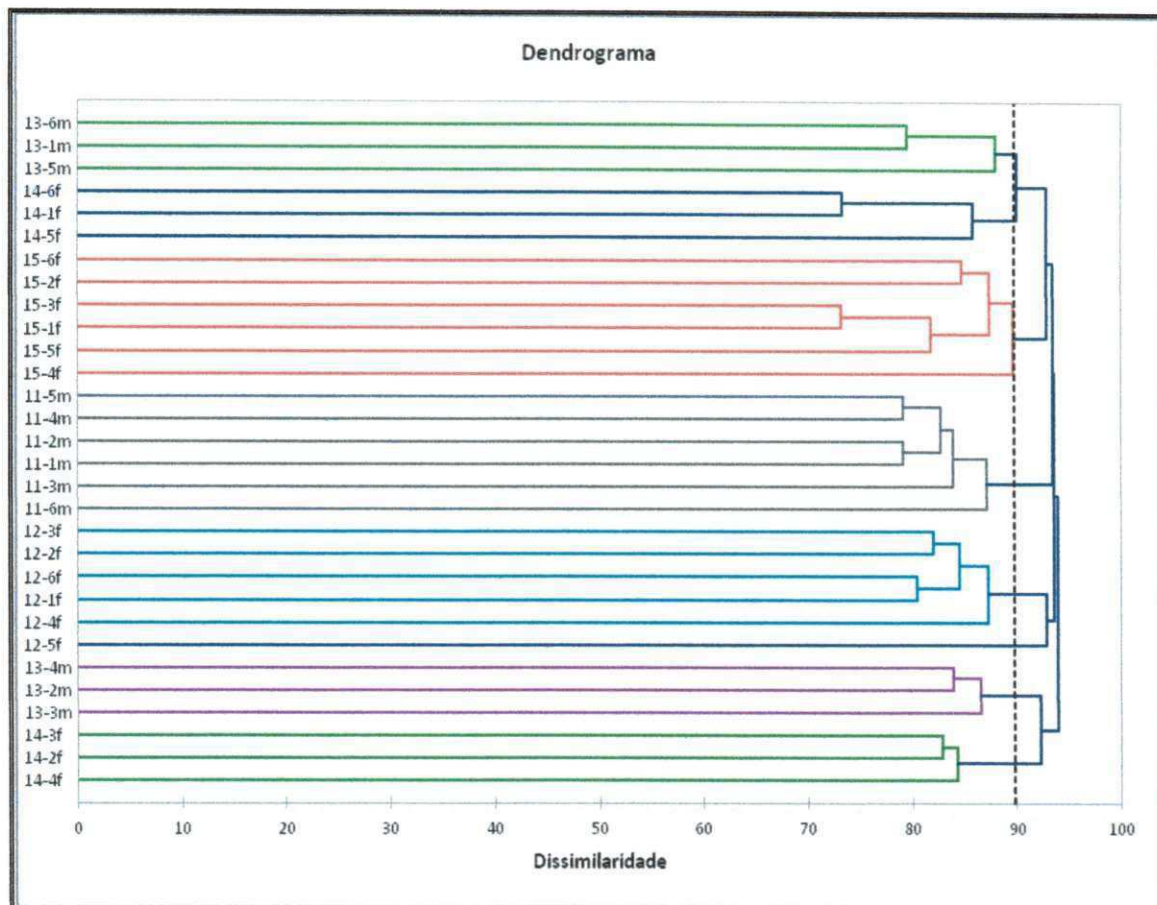


Figura E.9 – Dendrograma dos grupos de faces obtidos na Partição 3.

Resultados por classe correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 3. Os objetos ilustrados na tabela representam os nomes das imagens da base e os dois primeiros dígitos (em negrito) representam o indivíduo correspondente.

Tabela E.3 – Resultados por classe dos grupos de faces obtidos na Partição 3.

| Classe | Objeto |
|---------------|--|
| 1 (G1) | 13 -6m.jpg 13 -1m.jpg 13 -5m.jpg |
| 2 (G2) | 14 -6f.jpg 14 -1f.jpg 14 -5f.jpg |
| 3 (G3) | 15 -6f.jpg 15 -2f.jpg 15 -3f.jpg 15 -1f.jpg 15 -5f.jpg 15 -4f.jpg |
| 4 (G4) | 11 -5m.jpg 11 -4m.jpg 11 -2m.jpg 11 -1m.jpg 11 -3m.jpg 11 -6m.jpg |
| 5 (G5) | 12 -3f.jpg 12 -2f.jpg 12 -6f.jpg 12 -1f.jpg 12 -4f.jpg |
| 6 (G6) | 12 -5f.jpg |
| 7 (G7) | 13 -4m.jpg 13 -2m.jpg 13 -3m.jpg |
| 8 (G8) | 14 -3f.jpg 14 -2f.jpg 14 -4f.jpg |

E.4. Resultados da Partição 4

Resultado visual dos grupos obtidos a partir do uso da abordagem proposta na Partição 4.

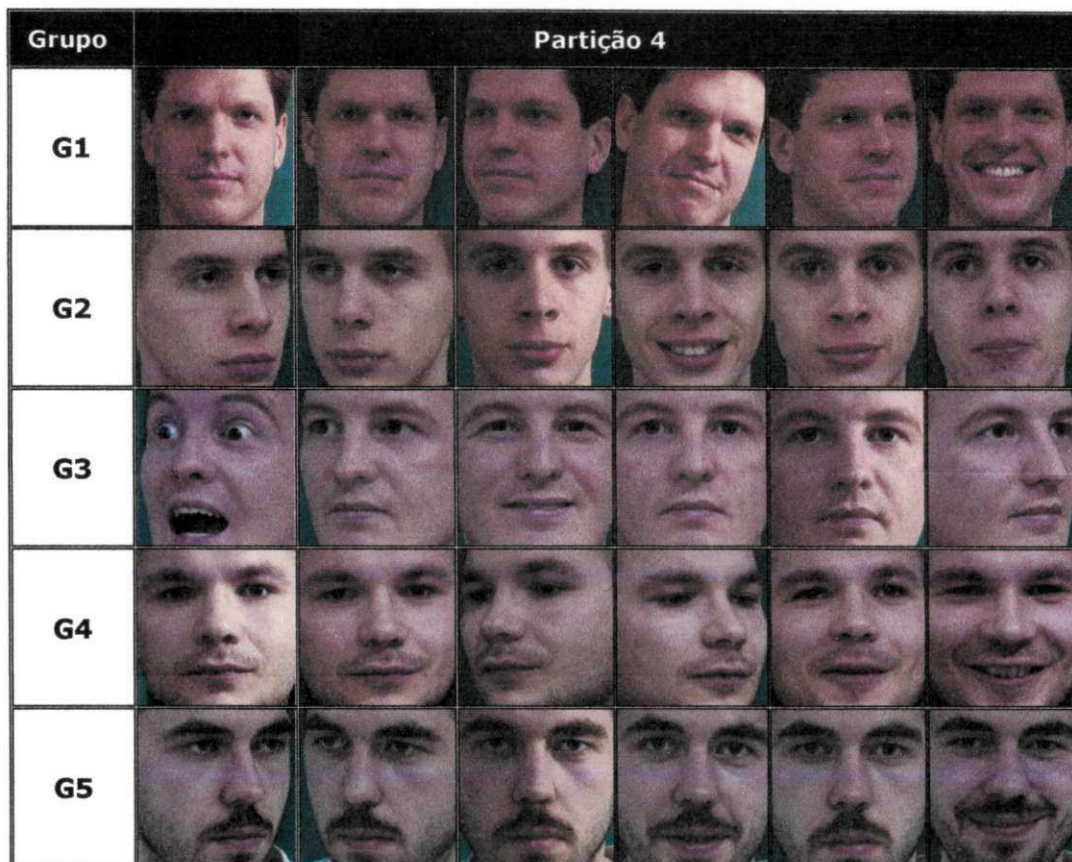


Figura E.10 – Grupos de faces obtidos na Partição 4.

Diagrama dos níveis de similaridade correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 4.

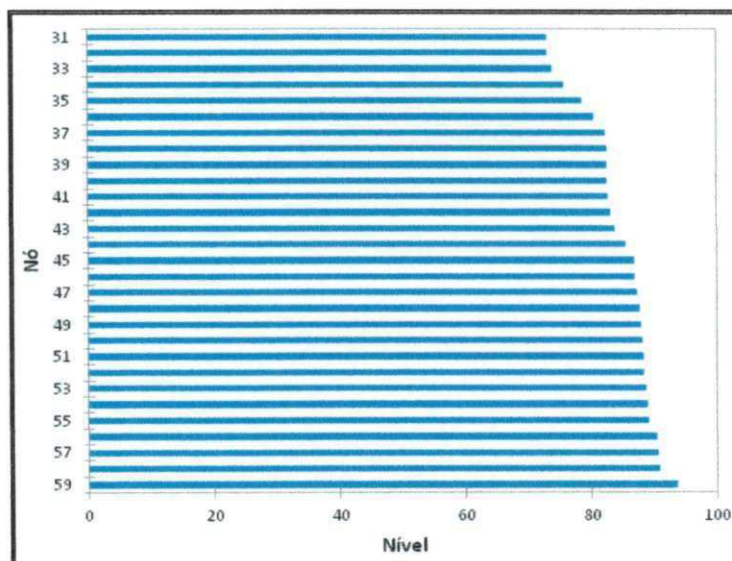


Figura E.11 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 4.

Dendrograma correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 4.

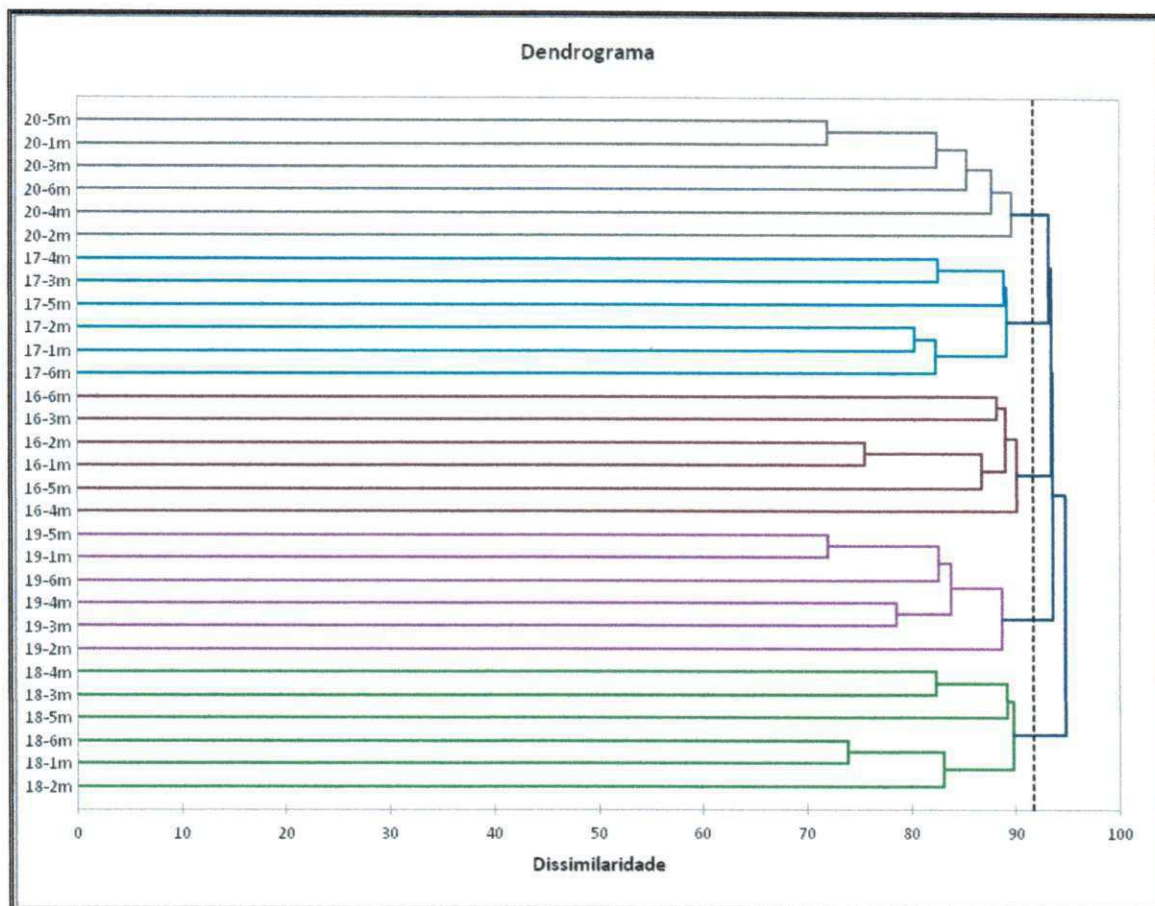


Figura E.12 – Dendrograma dos grupos de faces obtidos na Partição 4.

Resultados por classe correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 4. Os objetos ilustrados na tabela representam os nomes das imagens da base e os dois primeiros dígitos (em negrito) representam o indivíduo correspondente.

Tabela E.4 – Resultados por classe dos grupos de faces obtidos na Partição 4.

| Classe | Objeto |
|---------------|--|
| 1 (G1) | 20 -5m.jpg 20 -1m.jpg 20 -3m.jpg 20 -6m.jpg 20 -4m.jpg 20 -2m.jpg |
| 2 (G2) | 17 -4m.jpg 17 -3m.jpg 17 -5m.jpg 17 -2m.jpg 17 -1m.jpg 17 -6m.jpg |
| 3 (G3) | 16 -6m.jpg 16 -3m.jpg 16 -2m.jpg 16 -1m.jpg 16 -5m.jpg 16 -4m.jpg |
| 4 (G4) | 19 -5m.jpg 19 -1m.jpg 19 -6m.jpg 19 -4m.jpg 19 -3m.jpg 19 -2m.jpg |
| 5 (G5) | 18 -4m.jpg 18 -3m.jpg 18 -5m.jpg 18 -6m.jpg 18 -1m.jpg 18 -2m.jpg |

E.5. Resultados da Partição 5

Resultado visual dos grupos obtidos a partir do uso da abordagem proposta na Partição 5.

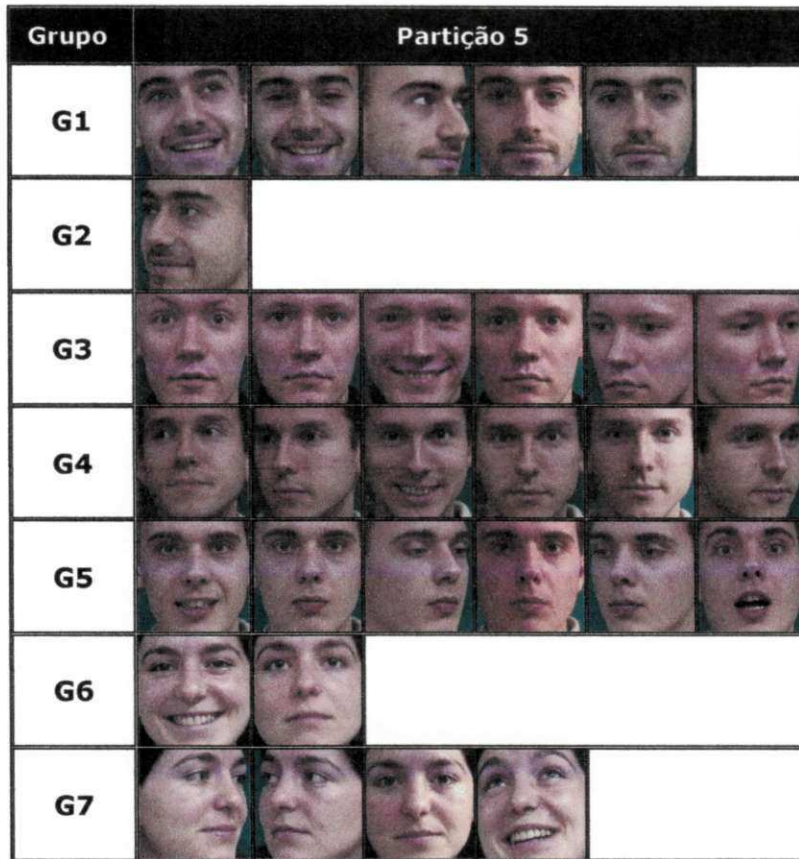


Figura E.13 – Grupos de faces obtidos na Partição 5.

Diagrama dos níveis de similaridade correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 5.

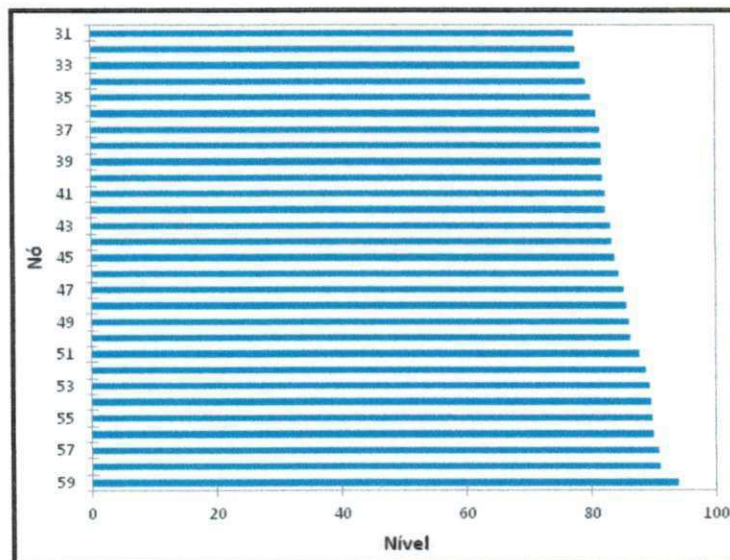


Figura E.14 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 5.

Dendrograma correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 5.

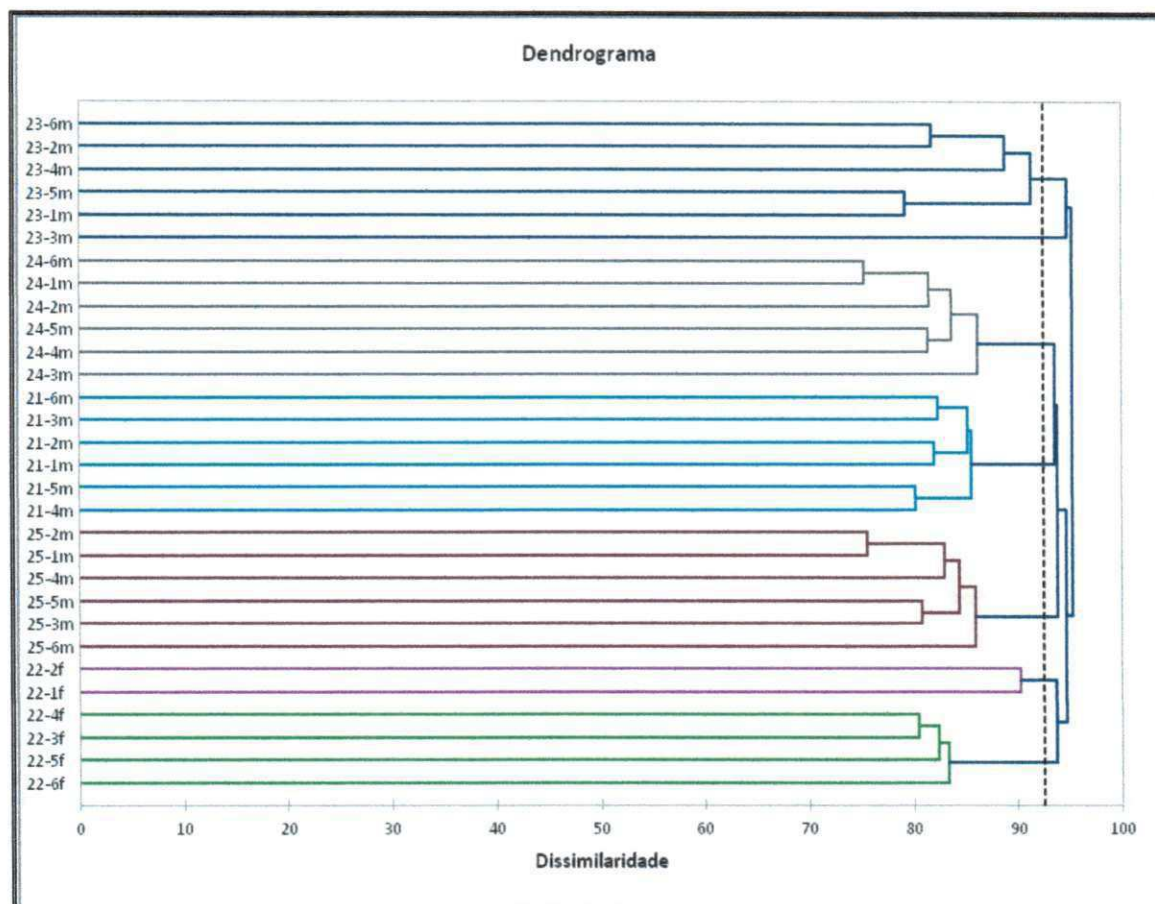


Figura E.15 – Dendrograma dos grupos de faces obtidos na Partição 5.

Resultados por classe correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 5. Os objetos ilustrados na tabela representam os nomes das imagens da base e os dois primeiros dígitos (em negrito) representam o indivíduo correspondente.

Tabela E.5 – Resultados por classe dos grupos de faces obtidos na Partição 5.

| Classe | Objeto |
|----------|--|
| 1 | 23-6m.jpg 23-2m.jpg 23-4m.jpg 23-5m.jpg 23-1m.jpg |
| 2 | 23-3m.jpg |
| 3 | 24-6m.jpg 24-1m.jpg 24-2m.jpg 24-5m.jpg 24-4m.jpg 24-3m.jpg |
| 4 | 21-6m.jpg 21-3m.jpg 21-2m.jpg 21-1m.jpg 21-5m.jpg 21-4m.jpg |
| 5 | 25-2m.jpg 25-1m.jpg 25-4m.jpg 25-5m.jpg 25-3m.jpg 25-6m.jpg |
| 6 | 22-2f.jpg 22-1f.jpg |
| 7 | 22-4f.jpg 22-3f.jpg 22-5f.jpg 22-6f.jpg |

E.6. Resultados da Partição 6

Resultado visual dos grupos obtidos a partir do uso da abordagem proposta na Partição 6.

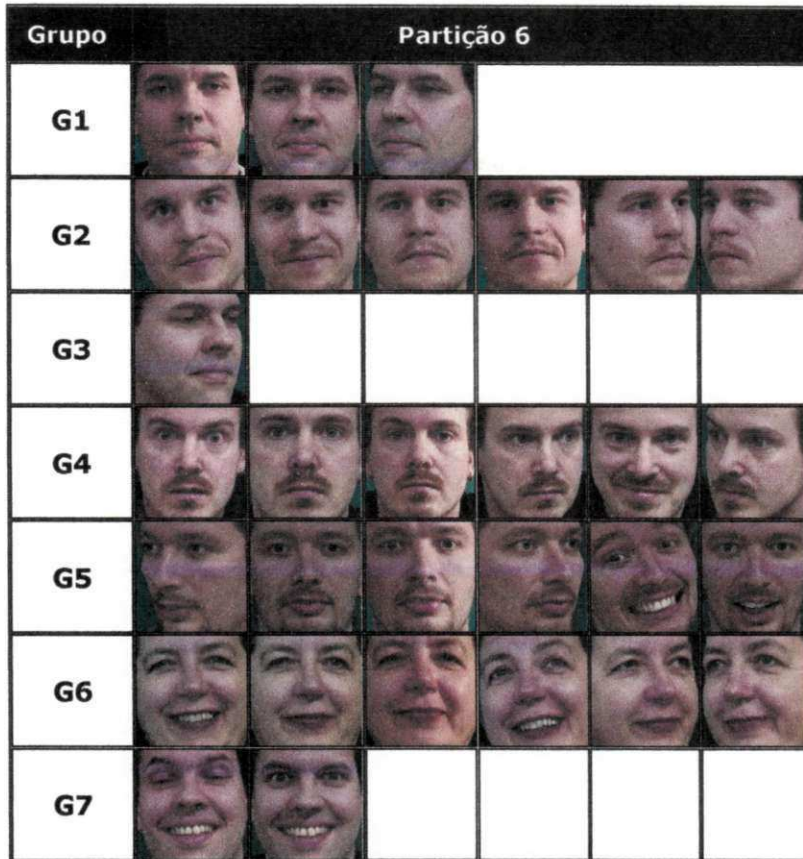


Figura E.16 – Grupos de faces obtidos na Partição 6.

Diagrama dos níveis de similaridade correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 6.

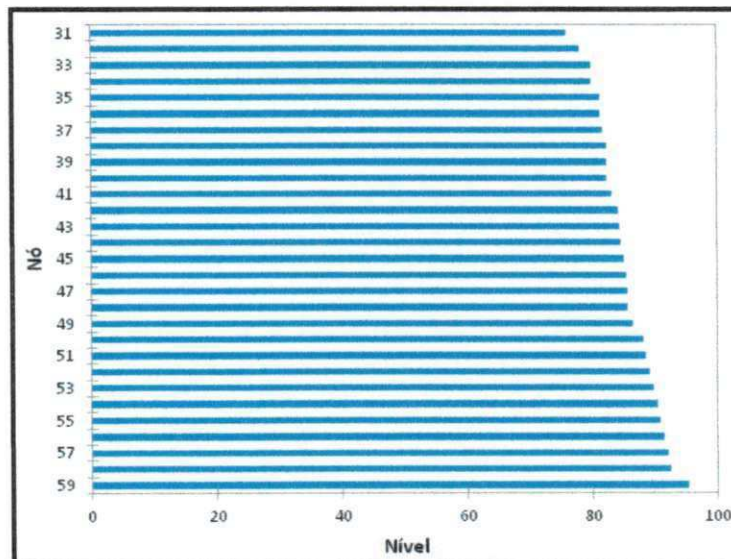


Figura E.17 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 6.

Dendrograma correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 6.

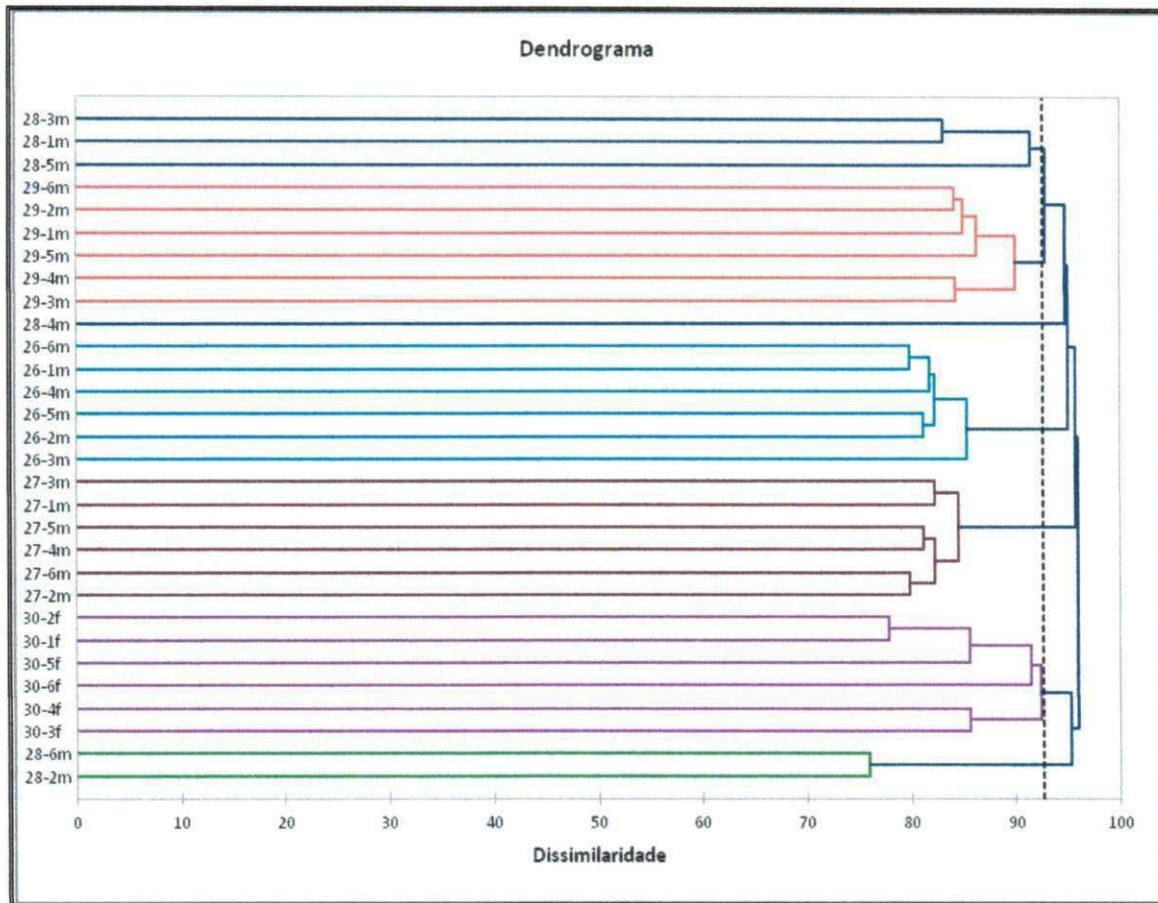


Figura E.18 – Dendrograma dos grupos de faces obtidos na Partição 6.

Resultados por classe correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 6. Os objetos ilustrados na tabela representam os nomes das imagens da base e os dois primeiros dígitos (em negrito) representam o indivíduo correspondente.

Tabela E.6 – Resultados por classe dos grupos de faces obtidos na Partição 6.

| Classe | Objeto |
|---------------|--|
| 1 (G1) | 28-3m.jpg 28-1m.jpg 28-5m.jpg |
| 2 (G2) | 29-6m.jpg 29-2m.jpg 29-1m.jpg 29-5m.jpg 29-4m.jpg 29-3m.jpg |
| 3 (G3) | 28-4m.jpg |
| 4 (G4) | 26-6m.jpg 26-1m.jpg 26-4m.jpg 26-5m.jpg 26-2m.jpg 26-3m.jpg |
| 5 (G5) | 27-3m.jpg 27-1m.jpg 27-5m.jpg 27-4m.jpg 27-6m.jpg 27-2m.jpg |
| 6 (G6) | 30-2f.jpg 30-1f.jpg 30-5f.jpg 30-6f.jpg 30-4f.jpg 30-3f.jpg |
| 7 (G7) | 28-6m.jpg 28-2m.jpg |

E.7. Resultados da Partição 7

Resultado visual dos grupos obtidos a partir do uso da abordagem proposta na Partição 7.

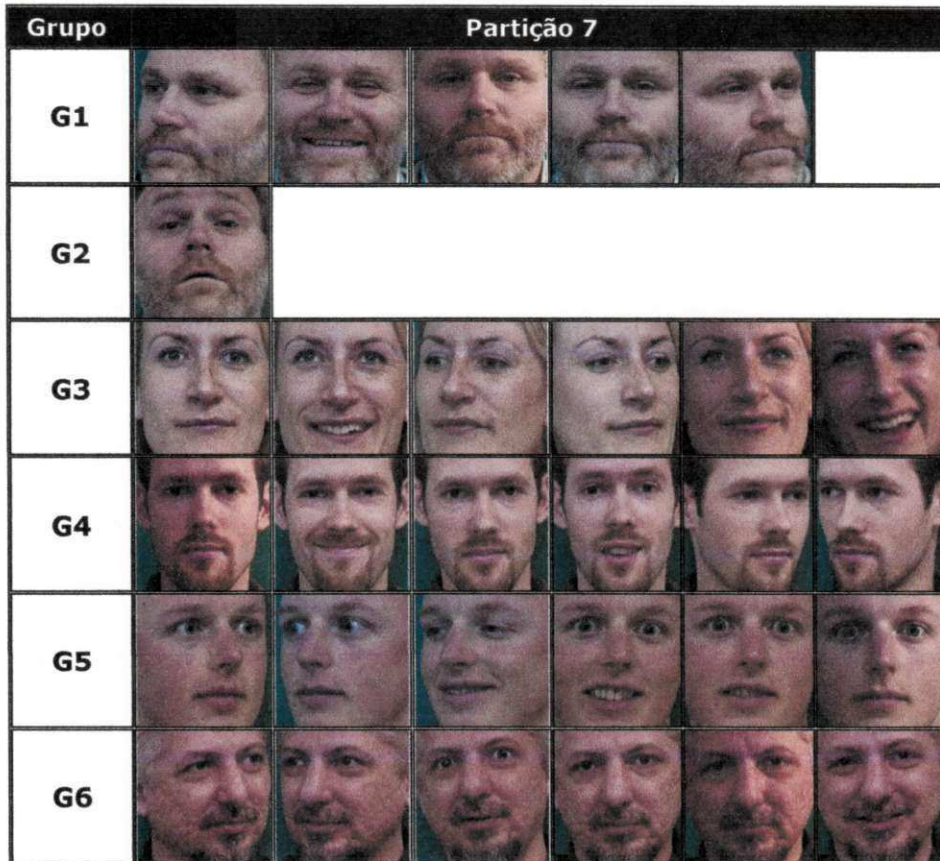


Figura E.19 – Grupos de faces obtidos na Partição 7.

Diagrama dos níveis de similaridade correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 7.

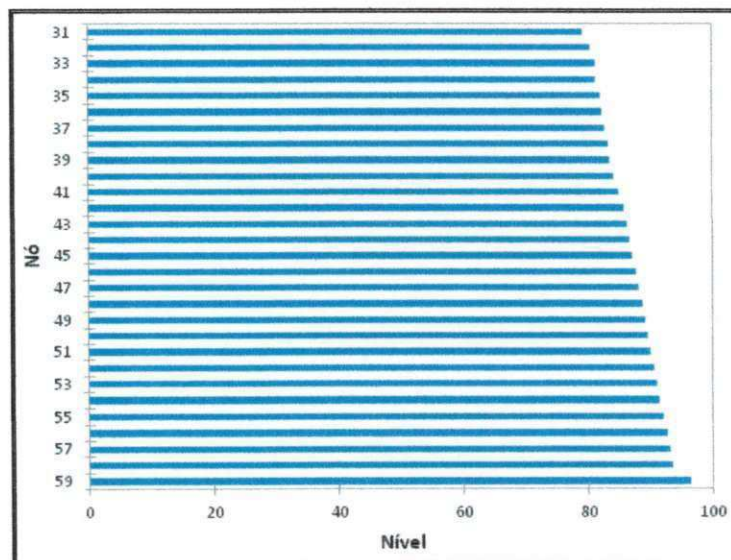


Figura E.20 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 7.

Dendrograma correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 7.

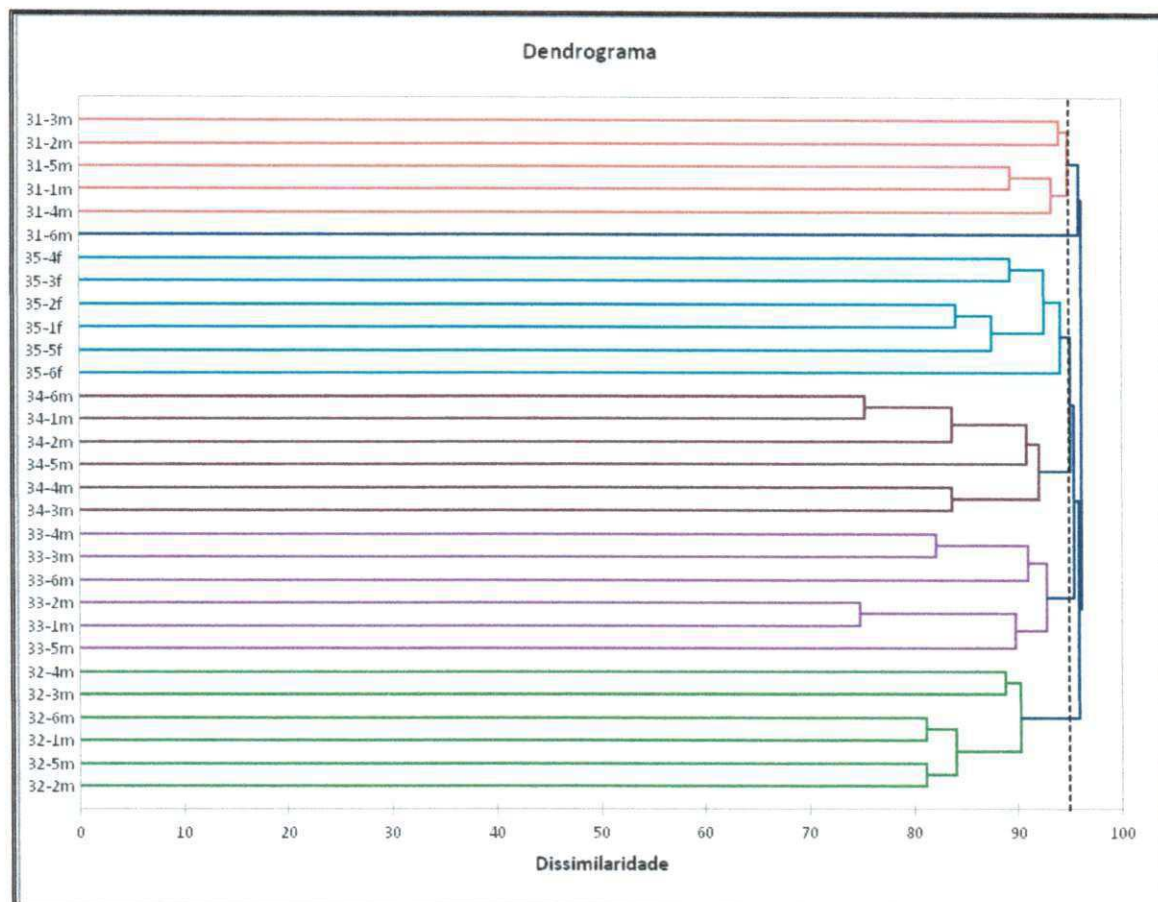


Figura E.21 – Dendrograma dos grupos de faces obtidos na Partição 7.

Resultados por classe correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 7. Os objetos ilustrados na tabela representam os nomes das imagens da base e os dois primeiros dígitos (em negrito) representam o indivíduo correspondente.

Tabela E.7 – Resultados por classe dos grupos de faces obtidos na Partição 7.

| Classe | Objeto |
|---------------|--|
| 1 (G1) | 31 -3m.jpg 31 -2m.jpg 31 -5m.jpg 31 -1m.jpg 31 -4m.jpg |
| 2 (G2) | 31 -6m.jpg |
| 3 (G3) | 35 -4f.jpg 35 -3f.jpg 35 -2f.jpg 35 -1f.jpg 35 -5f.jpg 35 -6f.jpg |
| 4 (G4) | 34 -6m.jpg 34 -1m.jpg 34 -2m.jpg 34 -5m.jpg 34 -4m.jpg 34 -3m.jpg |
| 5 (G5) | 33 -4m.jpg 33 -3m.jpg 33 -6m.jpg 33 -2m.jpg 33 -1m.jpg 33 -5m.jpg |
| 6 (G6) | 32 -4m.jpg 32 -3m.jpg 32 -6m.jpg 32 -1m.jpg 32 -5m.jpg 32 -2m.jpg |

E.8. Resultados da Partição 8

Resultado visual dos grupos obtidos a partir do uso da abordagem proposta na Partição 8.

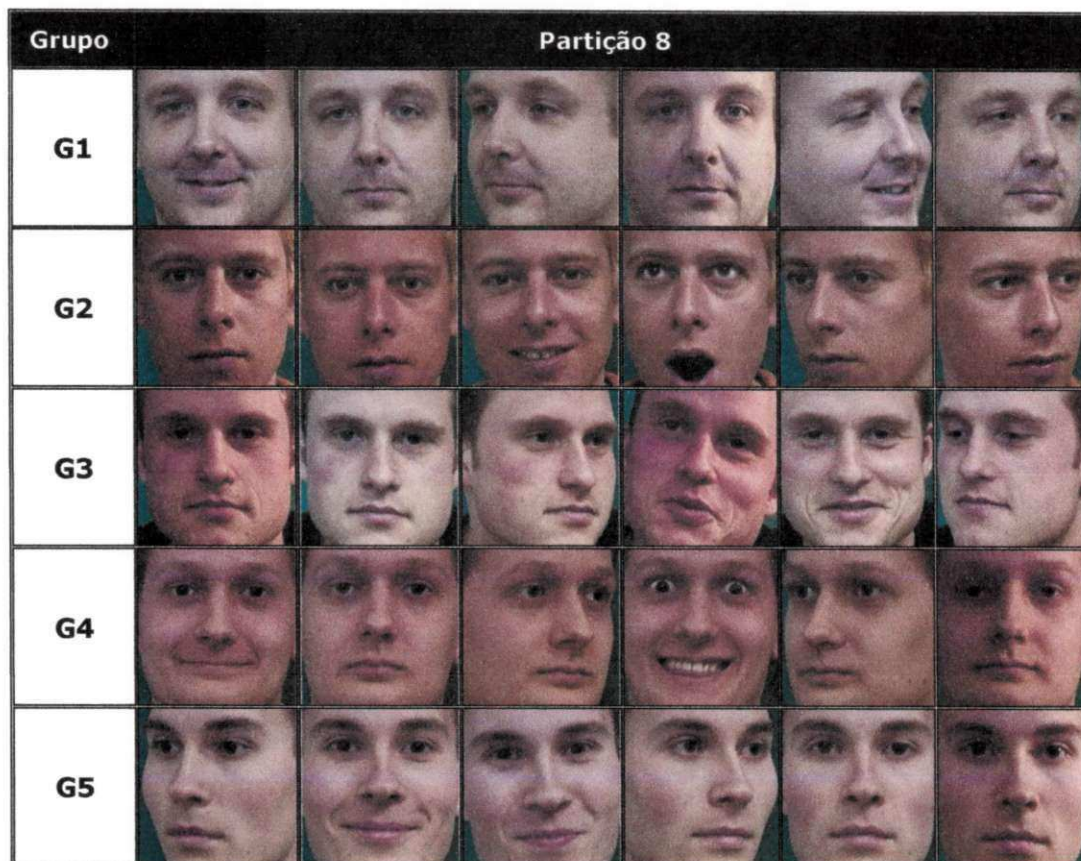


Figura E.22 – Grupos de faces obtidos na Partição 8.

Diagrama dos níveis de similaridade correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 8.

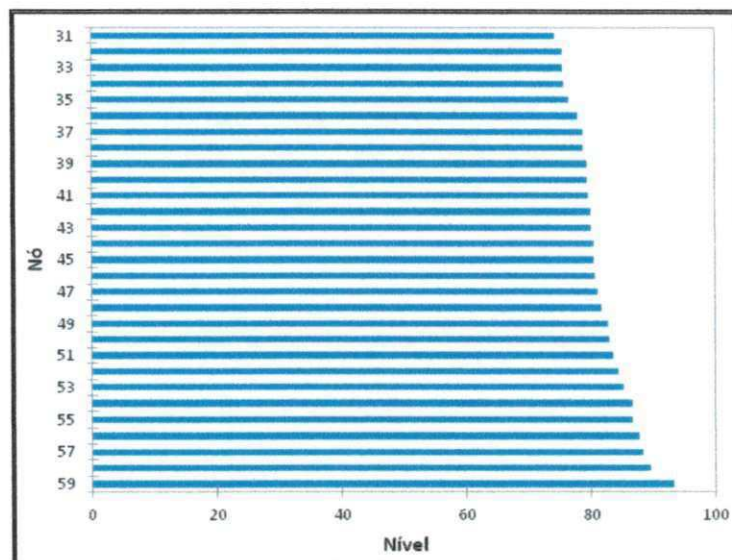


Figura E.23 – Diagrama dos níveis de similaridade dos grupos de faces obtidos na Partição 8.

Dendrograma correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 8.

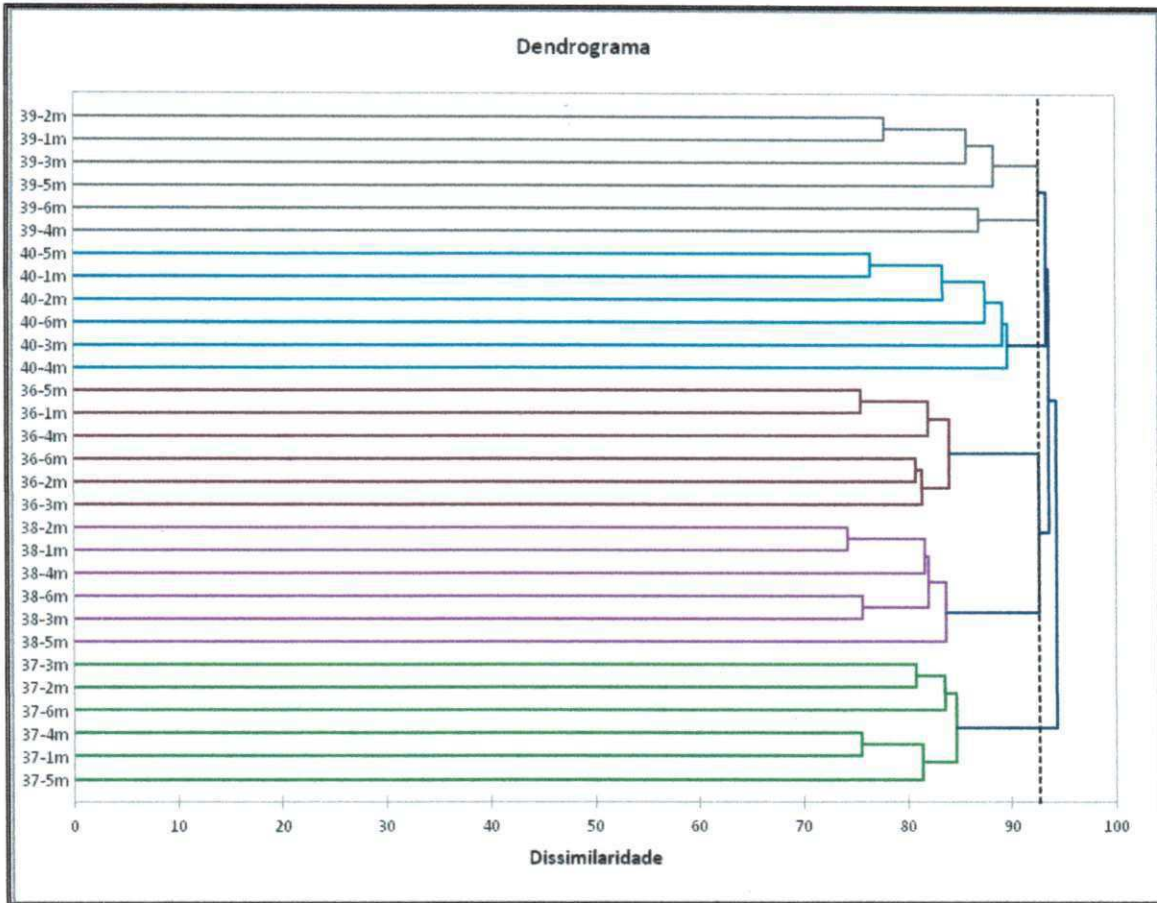


Figura E.24 – Dendrograma dos grupos de faces obtidos na Partição 8.

Resultados por classe correspondente aos grupos de faces obtidos a partir do uso da abordagem proposta na Partição 8. Os objetos ilustrados na tabela representam os nomes das imagens da base e os dois primeiros dígitos (em negrito) representam o indivíduo correspondente.

Tabela E.8 – Resultados por classe dos grupos de faces obtidos na Partição 8.

| Classe | Objeto |
|---------------|--|
| 1 (G1) | 39 -2m.jpg 39 -1m.jpg 39 -3m.jpg 39 -5m.jpg 39 -6m.jpg 39 -4m.jpg |
| 2 (G2) | 40 -5m.jpg 40 -1m.jpg 40 -2m.jpg 40 -6m.jpg 40 -3m.jpg 40 -4m.jpg |
| 3 (G3) | 36 -5m.jpg 36 -1m.jpg 36 -4m.jpg 36 -6m.jpg 36 -2m.jpg 36 -3m.jpg |
| 4 (G4) | 38 -2m.jpg 38 -1m.jpg 38 -4m.jpg 38 -6m.jpg 38 -3m.jpg 38 -5m.jpg |
| 5 (G5) | 37 -3m.jpg 37 -2m.jpg 37 -6m.jpg 37 -4m.jpg 37 -1m.jpg 37 -5m.jpg |