



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO  
UNIVERSIDADE DO ESTADO DO RIO GRANDE DO NORTE  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA  
COMPUTAÇÃO



Antonio Fernandes da Silva Filho

**AVALIAÇÃO DO USO DE TÉCNICAS DE  
AGRUPAMENTO NA BUSCA E RECUPERAÇÃO  
DE IMAGENS**

MOSSORÓ  
2016

ANTONIO FERNANDES DA SILVA FILHO

**AVALIAÇÃO DO USO DE TÉCNICAS DE  
AGRUPAMENTO NA BUSCA E RECUPERAÇÃO  
DE IMAGENS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação - associação ampla entre a Universidade do Estado do Rio Grande do Norte e a Universidade Federal Rural do Semi-Árido, para a obtenção do título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Daniel Sabino Amorim de Araújo

**Mossoró  
2016**

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

S586a Silva Filho, Antonio Fernandes da.  
Avaliação do uso de técnicas de agrupamento na  
busca e recuperação de imagens / Antonio Fernandes  
da Silva Filho. - 2016.  
89 f. : il.

Orientador: Daniel Sabino Amorim de Araújo.  
Dissertação (Mestrado) - Universidade Federal  
Rural do Semi-árido, Programa de Pós-graduação em  
Ciência da Computação, 2016.

1. Reconhecimento facial. 2. Agrupamento de  
imagens. 3. Extração de características. I. Araújo,  
Daniel Sabino Amorim de, orient. II. Título.

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

Antonio Fernandes da Silva Filho

**AVALIAÇÃO DO USO DE TÉCNICAS DE  
AGRUPAMENTO NA BUSCA E RECUPERAÇÃO  
DE IMAGENS**

Dissertação apresentada ao Programa de Pós-  
Graduação em Ciência da Computação para  
a obtenção do título de Mestre em Ciência da  
Computação.

APROVADA EM: 26 / 08 / 2016.

BANCA EXAMINADORA



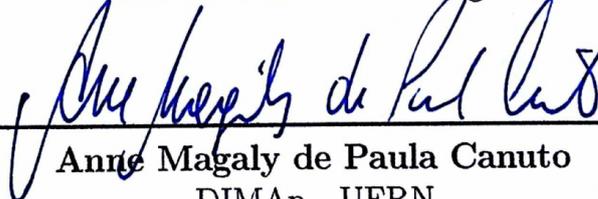
---

**Daniel Sabino Amorim de Araújo**  
IMD - UFRN  
Presidente



---

**Araken de Medeiros Santos**  
DCETH - UFERSA  
Examinador Interno ao Programa



---

**Anne Magaly de Paula Canuto**  
DIMAp - UFRN  
Examinadora Externa ao Programa

*Este trabalho é dedicado especialmente aos meus pais, pelo apoio constante e incondicional; aos demais familiares, por acreditarem no meu potencial.*

# Agradecimentos

Os meus principais agradecimentos vão para meus pais, pelo apoio incondicional, e por sempre terem acreditado em meu potencial. Sem eles, não seria a pessoa que sou hoje. Minha gratidão é eterna. Meu anseio de sucesso está destinado a lhes proporcionar sempre o melhor. Obrigado ao meu Pai, Antonio Fernandes, e minha Mãe, Paula Francinete. Amo vocês!

Agradeço à CAPES, pelo financiamento e apoio a este trabalho.

As instituições UERN e UFERSA, pelo desenvolvimento do programa de pós-graduação ao qual estou envolvido, bem como pelo apoio intelectual e material, fornecido pelas mesmas, fundamentais para o surgimento deste projeto de dissertação.

A meu orientador, Prof. Dr. Daniel Sabino, pelo desenvolvimento da ideia principal desse projeto, por todas as contribuições, conselhos, ajuda, sabedoria, paciência e sensibilidade, oferecidos no desenvolvimento desse trabalho e no decorrer do curso.

Aos professores da minha vida, que me ajudaram a alcançar o nível intelectual em que me encontro, além de ensinar a sempre buscar a evolução, seja como estudante ou ser humano. Pra todos, meu sinceros agradecimentos!

Agradeço também a minha irmã, minhas tias e tios, a minha eterna avó (*in memoriam*), meus primos, e todos aqueles familiares que me são próximos e que sempre torceram por mim. Sou grato por todo apoio e pela confiança depositada em minha pessoa. Muito obrigado! Amo vocês!

Aos meus amigos, de dentro e fora do âmbito acadêmico, que também torcem por mim, ajudando-me a superar desafios e me acompanhando nos momentos de descontração.

Aos demais, que não tenha lembrado, que torceram por mim ou tem algum sentimento bom em relação a minha pessoa, meus agradecimentos.

Muito obrigado!

*“A vida é aquilo que acontece enquanto você está ocupado fazendo outros planos.”*  
*(John Lennon)*

# Resumo

Atualmente quase todos os serviços e tarefas cotidianas envolvem algum aparato computacional, acarretando a criação e o conseqüente acúmulo de dados. Essa progressiva quantidade de dados representa uma importante oportunidade de exploração para os ramos científico e comercial, que passaram a valorizar e utilizar essas informações de forma mais intensa e objetiva. Aliado a isso, o processo natural de exposição da vida pública e privada através das redes sociais e dos dispositivos eletrônicos tende a gerar uma quantidade expressiva de imagens que podem e devem ser aproveitadas com os mais diversos fins, como na área de segurança pública. Nesse contexto, o reconhecimento facial tem avançado e atraído estudos e aplicações específicas, que visam a identificação de indivíduos por meio de características paramétricas. No entanto, alguns entraves ainda são encontrados, dificultando a realização eficiente dessa operação, como o custo computacional relativo ao tempo de busca e recuperação de imagens em bases de dados de grandes proporções. Baseado nisso, este trabalho propõe a utilização de algoritmos de agrupamento na organização dos dados de imagens, proporcionando assim um direcionamento e “encurtamento” nas buscas de imagens faciais. Mais especificamente, é feita uma análise relacionada à otimização imposta pelo uso de técnicas de agrupamento aplicadas na organização automatizada das imagens, como etapa preparativa para realização de buscas. O método proposto foi aplicado em bases de dados de imagens faciais reais, e utilizou dois algoritmos de agrupamento (*k-means* e EM) com variações para as medidas de similaridade (distância euclidiana e correlação de *Pearson*). Os resultados obtidos revelam que o emprego do agrupamento na organização dos dados mostrou-se eficiente, levando a uma redução significativa no tempo de busca, e sem prejuízos na precisão do processo.

**Palavras-chaves:** Reconhecimento facial, Agrupamento de imagens, Extração de características.

# Abstract

Nowadays, almost all services and daily tasks involve some computational apparatus, leading to creation and further accumulation of data. This progressive amount of data is an important opportunity of exploration for scientific and commercial branches, which started to value and to use this information more intense and objectively. Allied to this, the natural process of public and private life exposure through social networks and electronic devices tend to generate a significant amount of images that can and should be utilized with various purposes, such as in public security. In this context, facial recognition has advanced and attracted specific studies and applications, aiming at the identification of individuals through parametric features. However, some barriers are still found, making difficult the efficient performance of the operation, such as the computational cost on the search time and recovery of large proportions in image databases. Based on this, this paper proposes the use of clustering algorithms in the organization of image data, thus providing a direction and “shortening” in facial images searches. More specifically, an analysis related to the optimization is conducted imposed by the use of clustering techniques applied in the automated organization of images, the preparative step for performing searches. The proposed method was applied to real facial images databases and used two clustering algorithms k-means and EM with variations for the similarity measures (euclidean distance and Pearson correlation). The results show that the use of clustering in data organization has proved to be efficient, leading to a significant reduction in search time and without losses in process accuracy.

**Keywords:** Facial recognition, Image clustering, Search algorithms, Feature extraction, Clustering.

# Lista de ilustrações

Figura 1 – Representação de conjunto de dados. (Fonte: Xu e Wunsch II (2007)) .	21
Figura 2 – Representação de processo de agrupamento. (Adaptação: Xu e Wunsch II (2007)) . . . . .	22
Figura 3 – Exemplo de execução do <i>K-Means</i> . (Fonte: Linden (2009)) . . . . .	23
Figura 4 – Exemplo de execução do EM com dois componentes Gaussianos. (Fonte: Aggarwal e Reddy (2013)) . . . . .	24
Figura 5 – Passos de execução da técnica HOG. (Adaptação: Dalal e Triggs (2005))	28
Figura 6 – Demonstração do processo HOG (Adaptação: Nixon e Aguado (2012))	28
Figura 7 – Busca direta no conjunto de dados. . . . .	38
Figura 8 – Primeira etapa da busca com auxílio de agrupamento no conjunto de dados: divisão do conjunto de dados em grupos. . . . .	39
Figura 9 – Segunda etapa da busca com auxílio de agrupamento no conjunto de dados: determinação da proximidade dos grupos com a imagem teste. .	40
Figura 10 – Terceira etapa da busca com auxílio de agrupamento no conjunto de dados: busca local dentro dos grupos. . . . .	41
Figura 11 – Imagem em seu estado original. (Fonte: Thomaz e Giraldi (2010)) . . .	42
Figura 12 – Seleção da face. (Adaptação: Thomaz e Giraldi (2010)) . . . . .	42
Figura 13 – Mudança de proporção e transformação em escala de cinza (Adaptação: Thomaz e Giraldi (2010)). . . . .	42
Figura 14 – Melhoramento de contraste (Adaptação: Thomaz e Giraldi (2010)). . .	43
Figura 15 – Visualização da características na imagem (Adaptação: Thomaz e Giraldi (2010)). . . . .	43
Figura 16 – Características HOG extraídas (Adaptação: Thomaz e Giraldi (2010)).	44
Figura 17 – Exemplos de imagens da base de dados FEI. . . . .	46
Figura 18 – Exemplos de imagens da base de dados JAFFE. . . . .	47
Figura 19 – Gráfico relacionado a totalidade dos resultados da buscas. . . . .	62
Figura 20 – Gráfico relacionado a totalidade dos resultados da buscas. . . . .	76

# Lista de tabelas

Tabela 1 – Descrição das bases de imagens abordadas. . . . .	45
Tabela 2 – Legendas e as imagens correspondentes nas três bases abordadas na pesquisa . . . . .	49
Tabela 3 – Acurácia aferida nas experimentações . . . . .	52
Tabela 4 – Busca direta baseada em distância Euclidiana e correlação de <i>Pearson</i> nas bases completas. . . . .	53
Tabela 5 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e distância euclidiana na base <i>faces94</i> . . . . .	55
Tabela 6 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e distância euclidiana na base FEI. . . . .	55
Tabela 7 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e distância euclidiana na base JAFFE. . . . .	56
Tabela 8 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e correlação de <i>Pearson</i> na base <i>faces94</i> . . . . .	57
Tabela 9 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e correlação de <i>Pearson</i> na base FEI. . . . .	57
Tabela 10 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e correlação de <i>Pearson</i> na base JAFFE. . . . .	58
Tabela 11 – Busca com auxílio de agrupamento usando o EM e distância euclidiana na base <i>faces94</i> . . . . .	59
Tabela 12 – Busca com auxílio de agrupamento usando o EM e distância euclidiana na base FEI. . . . .	59
Tabela 13 – Busca com auxílio de agrupamento usando o EM e distância euclidiana na base JAFFE. . . . .	60
Tabela 14 – Busca com auxílio de agrupamento usando o EM e correlação de <i>Pearson</i> na base <i>faces94</i> . . . . .	60
Tabela 15 – Busca com auxílio de agrupamento usando o EM e correlação de <i>Pearson</i> na base FEI. . . . .	61
Tabela 16 – Busca com auxílio de agrupamento usando o EM e correlação de <i>Pearson</i> na base JAFFE. . . . .	61
Tabela 17 – Resultados sintetizados para os métodos de busca. . . . .	62
Tabela 18 – Ganho percentual de tempo de execução em detrimento da busca “gulosa”. . . . .	64
Tabela 19 – Acurácia aferida nas experimentações. . . . .	66

Tabela 20 – Valores limiares utilizados para determinar que duas imagens pertencem a uma mesma pessoa. . . . .	67
Tabela 21 – Busca direta de imagem não existente na base de dados baseada em distância Euclidiana e correlação <i>Pearson</i> . . . . .	68
Tabela 22 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e distância euclidiana na base <i>faces94</i> . . . . .	69
Tabela 23 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e distância euclidiana na base FEI. . . . .	69
Tabela 24 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e distância euclidiana na base JAFFE. . . . .	70
Tabela 25 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e correlação de <i>Pearson</i> na base <i>faces94</i> . . . . .	70
Tabela 26 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e correlação de <i>Pearson</i> na base FEI. . . . .	71
Tabela 27 – Busca com auxílio de agrupamento usando o <i>K-Means</i> e correlação de <i>Pearson</i> na base JAFFE. . . . .	71
Tabela 28 – Busca com auxílio de agrupamento usando o algoritmo EM e distância euclidiana na base <i>faces94</i> . . . . .	72
Tabela 29 – Busca com auxílio de agrupamento usando o algoritmo EM e distância euclidiana na base FEI. . . . .	73
Tabela 30 – Busca com auxílio de agrupamento usando o algoritmo EM e distância euclidiana na base JAFFE. . . . .	73
Tabela 31 – Busca com auxílio de agrupamento usando o EM e correlação de <i>Pearson</i> na base <i>faces94</i> . . . . .	74
Tabela 32 – Busca com auxílio de agrupamento usando o EM e correlação de <i>Pearson</i> na base FEI. . . . .	74
Tabela 33 – Busca com auxílio de agrupamento usando o EM e correlação de <i>Pearson</i> na base JAFFE. . . . .	75
Tabela 34 – Resultados sintetizados para os métodos de busca. . . . .	75
Tabela 35 – Ganho percentual de tempo de execução em detrimento da busca “gulosa”. . . . .	77
Tabela 36 – Tempo de execução dos agrupamentos com o algoritmo <i>K-Means</i> usando a distância euclidiana e correlação de <i>Pearson</i> nas três bases de dados completas. . . . .	78
Tabela 37 – Tempo de execução dos agrupamentos com o algoritmo EM nas três bases completas. . . . .	79

# Lista de Abreviaturas e Siglas

CMYK	<i>Cyan, Magenta, Yellow and Black</i>
CP	Correlação de <i>Pearson</i>
DE	Distância Euclidiana
DNA	<i>Deoxyribonucleic Acid</i>
DWT	<i>Discret Wavelet Transform</i>
EM	<i>Expectation-Maximization</i>
FEI	Fundação Educacional Inaciana
HOG	<i>Histogram Oriented Gradients</i>
HSV	<i>Hue, Saturation and Value</i>
JAFFE	<i>The Japanese Female Facial Expression</i>
kNN	<i>K-Nearest-Neighbor</i>
LSH	<i>Locality Sensitive Hashing</i>
PSO	<i>Particle Swarm Optimization</i>
RGB	<i>Red, Green and Blue</i>
RIBC	Recuperação de Imagem Baseada em Conteúdo
SIFT	<i>Scale Invariant Feature Transform</i>
SVM	<i>Support Vector Machine</i>
WLB	<i>Weak Labeled Database</i>

# Sumário

	<b>Lista de ilustrações</b> . . . . .	<b>10</b>
	<b>Lista de tabelas</b> . . . . .	<b>11</b>
<b>1</b>	<b>INTRODUÇÃO</b> . . . . .	<b>16</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b> . . . . .	<b>20</b>
<b>2.1</b>	<b>Análise de Agrupamento</b> . . . . .	<b>20</b>
<b>2.2</b>	<b>Algoritmos de Agrupamento</b> . . . . .	<b>22</b>
2.2.1	Algoritmos Particionais . . . . .	23
2.2.2	Algoritmos Probabilísticos . . . . .	23
2.2.3	Medidas de Proximidade . . . . .	24
2.2.3.1	Correlação de <i>Pearson</i> . . . . .	25
2.2.3.2	Distância Euclidiana . . . . .	25
<b>2.3</b>	<b>Processamento Digital de Imagens</b> . . . . .	<b>26</b>
2.3.1	Representação de Imagens Digitais . . . . .	26
<b>2.4</b>	<b><i>Histogram Oriented Gradients</i></b> . . . . .	<b>27</b>
2.4.1	Agrupamento de Imagens . . . . .	28
<b>2.5</b>	<b>Considerações Finais</b> . . . . .	<b>29</b>
<b>3</b>	<b>REVISÃO BIBLIOGRÁFICA</b> . . . . .	<b>31</b>
<b>3.1</b>	<b>Considerações Finais</b> . . . . .	<b>34</b>
<b>4</b>	<b>RECUPERAÇÃO DE IMAGENS COM AUXÍLIO DE TÉCNICAS DE AGRUPAMENTO</b> . . . . .	<b>35</b>
<b>4.1</b>	<b>Descrição da Problemática</b> . . . . .	<b>35</b>
<b>4.2</b>	<b>Solução Implementada</b> . . . . .	<b>37</b>
4.2.1	Sistema Proposto . . . . .	39
4.2.1.1	Passo 1 - Agrupamento da base . . . . .	39
4.2.1.2	Passo 2 - Verificação da semelhança entre imagem teste e grupos . . . . .	39
4.2.2	Passo 3 - Checagem interna nos grupos . . . . .	40
4.2.3	Descritores HOG . . . . .	41
<b>4.3</b>	<b>Considerações Finais</b> . . . . .	<b>43</b>
<b>5</b>	<b>MATERIAIS E EXPERIMENTOS</b> . . . . .	<b>45</b>

<b>5.1</b>	<b>Bases de Dados</b> . . . . .	<b>45</b>
5.1.1	Faces94 . . . . .	46
5.1.2	FEI . . . . .	46
5.1.3	JAFFE . . . . .	47
<b>5.2</b>	<b>Algoritmos e parâmetros</b> . . . . .	<b>47</b>
<b>5.3</b>	<b>Descrição dos experimentos</b> . . . . .	<b>48</b>
<b>5.4</b>	<b>Considerações finais</b> . . . . .	<b>50</b>
<b>6</b>	<b>RESULTADOS E DISCUSSÕES</b> . . . . .	<b>51</b>
<b>6.1</b>	<b>Busca de Imagem Existente na Base de Dados</b> . . . . .	<b>51</b>
6.1.1	Análise de Acurácia . . . . .	51
6.1.2	Análise do Tempo de Busca . . . . .	52
6.1.2.1	Desempenho do Algoritmo <i>k-means</i> . . . . .	54
6.1.2.2	Desempenho do Algoritmo EM . . . . .	58
6.1.2.3	<i>k-means</i> vs EM . . . . .	60
<b>6.2</b>	<b>Busca de Imagens Inexistentes na Base de Dados</b> . . . . .	<b>65</b>
6.2.1	Análise de Acurácia . . . . .	65
6.2.2	Análise do Tempo de Busca . . . . .	67
6.2.2.1	Desempenho do Algoritmo <i>k-means</i> . . . . .	69
6.2.2.2	Desempenho do Algoritmo EM . . . . .	72
6.2.3	<i>k-means</i> vs EM . . . . .	73
<b>6.3</b>	<b>Análise do Tempo de Execução do Processo de Agrupamento</b> . . . . .	<b>78</b>
<b>6.4</b>	<b>Considerações finais</b> . . . . .	<b>80</b>
<b>7</b>	<b>CONSIDERAÇÕES FINAIS</b> . . . . .	<b>82</b>
<b>7.1</b>	<b>Trabalhos Futuros</b> . . . . .	<b>83</b>
	<b>Referências</b> . . . . .	<b>85</b>

# 1 Introdução

O nosso cotidiano se encontra cada dia mais informatizado. Praticamente todos os muitos serviços e tarefas diárias envolvem algum esforço computacional. Com isso, a quantidade de dados que surgem todos os dias é imensa. A principal fonte de geração e armazenamento desses dados é a Internet, em especial com o provimento de redes sociais, bem como a evolução dos *gadgets*, que fomentam a comunicação interpessoal e em comunidade, e da computação em nuvem, que fornece um ambiente de armazenamento *web* e favorece o aumento de dados e informações pessoais/comerciais na rede. Podemos destacar, em especial, a quantidade crescente de imagens disponíveis na grande rede, e sua massiva utilização no desenvolvimento de novas pesquisas e técnicas altamente utilizáveis em diversos fins, como presentes em Saffer (2011), Alattab e Kareem (2012), Obulesu et al. (2015) e Desai e Sonawane (2017).

As redes sociais geram diariamente um montante considerável de dados pessoais, com capacidade de disseminação em níveis assustadores. Existem precedentes de uso positivo das redes sociais como utilidade pública. Dentre esses casos, vemos a identificação de criminosos e situações de perigo (Castiglione et al. (2011) e Shen et al. (2015)), diagnóstico de enfermidades (Hemmati e Chung (2014) e Huang et al. (2014)), melhorias em setores como de educação (Klimova et al. (2014) e Poulouva e Simonova (2014)).

Uma importante vertente do Aprendizado de Máquina trata de dados sem informações de classes ou rótulos (responsáveis por fornecer uma descrição sucinta dos padrões de semelhanças e diferenças nos dados): a Análise de Agrupamentos. Sua função principal é proporcionar uma organização dos dados a partir de semelhanças diversas. Essa característica possibilita promover uma separação de dados, fazendo com que os componentes do processo possam ser avaliados mais claramente, seja pelos aglomerados (individualmente) ou sob uma visão mais geral, além de promover ferramentas de predição (Everitt et al., 2011).

A Visão Computacional, por sua vez, é importante área da computação que busca explorar de forma eficiente a identificação de objetos e cenários baseados em imagens sensoriais, por meio do desenvolvimento de modelos e descrições (Shapiro e Stockman, 2000). Dentro dessa grande área de estudo, a detecção e reconhecimento de objetos e faces desponta como uma das fontes de pesquisas mais desafiadoras e abordadas. Encontrar faces em um cenário, bem como classificá-las, pode ser uma tarefa bastante árdua, devido às diversas características compartilhadas entre os indivíduos. Sendo assim, o Reconhecimento Facial mostra-se de extrema importância no desenvolvimento de diversas ferramentas

computacionais.

Adicionalmente aos crescentes avanços da ciência que se baseiam na geração e reconhecimento de padrões de imagens, existe ainda o processo natural de exposição da vida pública e privada. Por exemplo, é de conhecimento geral que o uso de câmeras para fins de segurança e monitoramento público cresce de forma exponencial nas cidades de médio e grande porte, como exemplificados em Chowdhry et al. (2013), K e Reshmi (2015), Satari et al. (2014). As redes sociais geram um montante de imagens pessoais diárias com capacidade de disseminação que atingiu níveis assustadores nos últimos tempos.

Essa progressiva quantidade de dados representa uma importante oportunidade de exploração. Munidos desse potencial, os ramos científico, acadêmico e comercial passaram a valorizar e utilizar essas informações de forma mais intensa e objetiva. Essa exploração é realizada, muitas vezes, através de técnicas relacionadas ao Aprendizado de Máquina, subárea da Inteligência Artificial, que pode ser definido como a capacidade de conceder ao computador o aprendizado a partir de um conjunto exemplos, e por meio de princípios de inferência (denominado indução), gerar conclusões genéricas ao problema (Faceli et al., 2014). Por exemplo, muitas empresas buscam formas de prever o comportamento de seus clientes em relação a algum serviço proposto, através de informações extraídas de dados dos históricos de compras dos próprios consumidores/prospectos, ou mesmo em redes sociais, sites de avaliação, entre outros meios.

É importante ressaltar que o uso de técnicas de agrupamento no contexto de imagens pode ter muitos fins relevantes. A separação em aglomerados semelhantes entre si e o refinamento de subgrupos é uma necessidade comum a outras áreas de aplicação emergentes, como a Bioinformática (muitas técnicas de sequenciamento geram imagens na ordem dos gigabytes), redes sociais (não a toa o crescimento de redes voltadas exclusivamente para imagens, como o *Instagram* e *Pinterest*), visão computacional (identificação de objetos semelhantes), dentre muitas outras.

Dentro desse contexto, os métodos de agrupamento podem ser usados para requintar o processo de extração de informações úteis dos meios que geram e usam imagens constantemente. Por exemplo, existe a possibilidade da utilização das técnicas de agrupamento e suas respectivas implementações na otimização de processos de recuperação de imagens faciais (que intenta o retorno da imagem requerida na busca), baseada nas características extraídas do rosto dos indivíduos. Essa proposição vai de encontro ao processamento de busca “gulosa”, forma mais básica e usual de pesquisa em bases de dados, que checa item por item até que o resultado requerido seja alcançado. Obviamente, esse método é mais rudimentar, ficando aquém quando a avaliação tem como base o desempenho temporal, mesmo que se faça uso de processamento massivo e paralelo. Se

uma técnica de agrupamento consegue separar o conjunto total de faces em grupos, essa busca pode ter um direcionamento mais refinado, diminuindo consideravelmente o tempo de resposta do sistema, e facilitando diretamente a vida do usuário final.

Apoiado no crescimento exponencial das bases de imagens e na sua preponderante utilização na realização de pesquisas científicas e implementações comerciais, seja no esforço pela velocidade do processamento de buscas, ou até mesmo agrupamento automático de imagens semelhantes, este trabalho propõe o desenvolvimento de um estudo acerca do uso técnicas de agrupamento com o objetivo de organizar os dados, como uma etapa de pré-processamento, aspirando assim direcionar e otimizar o processo de busca de imagens em grandes conjuntos de dados.

Essa análise engloba os aspectos referentes à análise de agrupamentos e busca de imagens, ou seja:

- formas de extração/representação dos dados;
- impacto que o pré-processamento pode trazer na qualidade final dos grupos gerados;
- escolha de algoritmos de agrupamento e medidas de similaridade/dissimilaridade;
- avaliação da eficiência e eficácia do processo de busca.

É importante notar que a análise da estratégia sugerida, bem como o seu contexto, trazem consigo uma boa contribuição para o campo do aprendizado de máquina envolvendo processamento de imagens, e principalmente na área de *Bigdata*. A possibilidade de explorar um processamento de busca otimizado, que traga consigo mais agilidade e dinamização ao regime tradicional, representa a principal contribuição deste projeto.

Este trabalho está organizado como descrito a seguir.

- O Capítulo 2 apresenta os fundamentos teóricos e definições acerca da análise de agrupamento e visão computacional, em sua abrangência;
- O Capítulo 3 apresenta algumas pesquisas relacionados ao tema abordado;
- No Capítulo 4, é demonstrada a problemática ligada ao trabalho e a seguinte proposta de solução;
- O Capítulo 5 aborda os materiais usados na resolução do problema, bem como o experimentos desenvolvidos;
- No Capítulo 6, são apresentados os resultados advindos das experimentações;

- E no Capítulo 7, as considerações finais.

## 2 Referencial Teórico

Neste capítulo são descritos os conceitos relativos à Análise de Agrupamento, focando em algoritmos com a finalidade ligada ao agrupamento de imagens. Além disso, são demonstrados aspectos da Visão Computacional e seus métodos ligados ao processamento de imagens e reconhecimento facial..

### 2.1 Análise de Agrupamento

Pode-se definir Análise de Agrupamento, de acordo com Mooi e Sarstedt (2011), como: “um método conveniente para identificar grupos homogêneos de objetos, qualificados como “grupos”. Objetos (ou casos, observações) em um grupo específico compartilham muitas características, mas são muito diferentes dos objetos que não pertencem a esse grupo”.

Quanto a definição de grupo, não existe um consenso entre a comunidade científica. Entretanto, Everitt et al. (2011) apresenta três definições que englobam o sentido geral do termo:

- “Um grupo é um conjunto de entidades que são parecidas, e entidades de grupos diferentes não são parecidas”;
- “Um agregado de pontos no espaço de ensaio de tal modo que a distância entre quaisquer dois pontos no grupo é menor do que a distância entre qualquer ponto no agrupamento e qualquer ponto em que não se encontra no agrupamento”;
- “Grupos podem ser descritos como regiões contínuas deste espaço (espaço de características  $d$ -dimensionais) contendo uma densidade relativamente elevada de pontos, separada das outras regiões por regiões que contêm uma densidade de pontos relativamente baixa”.

Para realização deste trabalho, foi considerada a segunda definição em detrimento das demais, pela consideração de que a mesma engloba o sentido mais geral e, ao mesmo tempo, mais “simplista” do conceito. A Figura 1 mostra exemplos de conjunto de dados contendo aglomerados de pontos, formando possíveis grupos.

A Análise de Agrupamento pode ser utilizada com diversas finalidades, dentre elas a sumarização de grandes bases de dados, facilitando o processo de estudo, além da

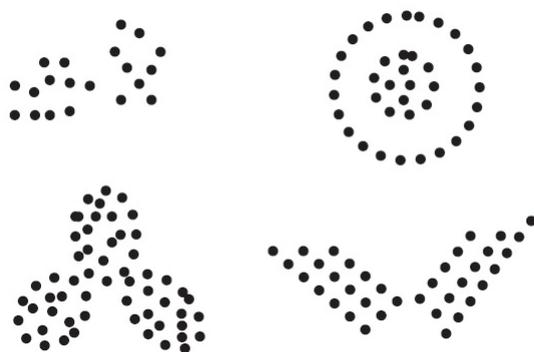


Figura 1 – Representação de conjunto de dados. (Fonte: Xu e Wunsch II (2007))

possibilidade de identificação de informações implícitas, o que propicia a formação de conhecimento que vai além da base de dados que está disponível.

De acordo com Xu e Wunsch II (2007), o processo de análise de agrupamento deve seguir quatro passos básicos:

1. Seleção ou extração de características: tecnicamente, seleção e extração representam técnicas diferentes. Na seleção ocorre a escolha de características distintas dentre um grupo de candidatos; já na extração são geradas novas informações a partir de transformações dos dados presentes. Diante dessas constatações, as duas técnicas são bastante importantes na formação de um agrupamento eficaz. Devido às questões técnicas, como dificuldade de interpretação, a extração é menos utilizada. A seleção de características também pode acarretar em problemas, como a perda de informação original. Sendo assim, as duas técnicas possuem vantagens e desvantagens. O que vai dizer quando cada uma deve ser utilizada é a natureza do problema.
2. Planejamento ou seleção do algoritmo de agrupamento: consiste na definição das medidas de proximidades, que serão utilizadas na formulação de uma função critério. Essas medidas irão determinar como os objetos serão alocados, de acordo com suas singularidades. Os algoritmos de agrupamento estão ligados a essas medidas de proximidade, seja explícita ou implicitamente. A escolha de um algoritmo de agrupamento depende decisivamente das características do problema.
3. Validação do agrupamento: corresponde às normas e critérios de avaliação necessários para disponibilizar um grau de confiança nos resultados do agrupamento. Uma boa validação deve apontar questões como se os agrupamentos obtidos foram significativos, ou qual algoritmo escolher. Os critérios de testes utilizados na validação, podem ser divididos em três categorias: índices externos, que se baseiam em estruturas pré-definidas; índices internos, que levam em consideração a estrutura do agrupamento

diretamente dos dados originais; índices relativos, que visam a comparação entre diferentes estruturas de agrupamento, de modo que seja escolhida a que melhor represente os objetos.

4. Interpretação dos resultados: representa o processo de interpretação dos dados gerados, visando a resolução dos problemas propostos inicialmente. Para resultados de qualidade é necessário que os três primeiros passos tenham sido executados de forma correta. A interpretação dos resultados geralmente demanda análises e experimentos mais profundos, o que pode envolver a integração de outras informações e provas de domínio experimental.

A Figura 2 mostra uma representação do processo de agrupamento, consistindo em quatro passos básicos.

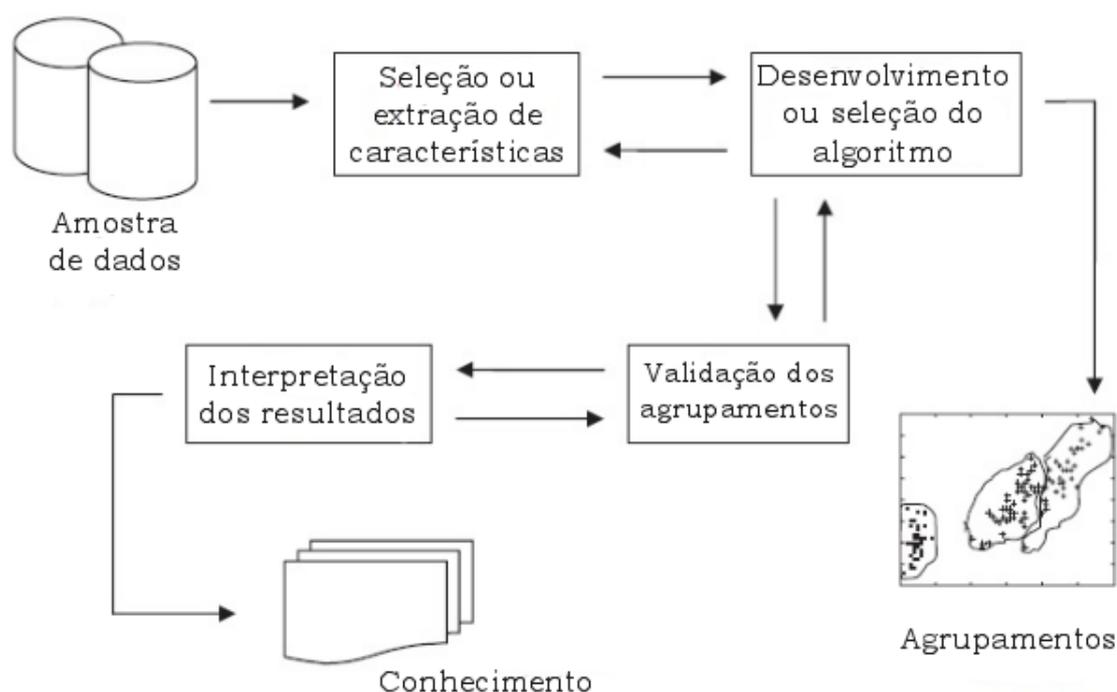


Figura 2 – Representação de processo de agrupamento. (Adaptação: Xu e Wunsch II (2007))

## 2.2 Algoritmos de Agrupamento

Nesta seção são apresentados alguns dos principais tipos de algoritmos de agrupamento, bem como seu funcionamento lógico e características. Eles são comumente definidos pelo modo como os grupos são formados. Há uma grande diversidade de tipos

de algoritmos de agrupamentos, como descritos em (Faceli et al., 2014). A seguir, são apresentados os algoritmos particionais e probabilísticos, integrantes do escopo da pesquisa.

### 2.2.1 Algoritmos Particionais

Essa técnica utiliza métodos iterativos na formação do agrupamento. Inicialmente, são criadas partições. Posteriormente, os objetos são deslocados para outros agrupamentos. Apesar de representar um método computacionalmente eficiente, pode convergir para ótimos locais - o erro quadrático (critério de agrupamento) garante a compactação dos agrupamentos. Partindo de um número fixo de agrupamentos, que é pré-definido, busca-se diminuir a variação existente dentro do agrupamento e, conseqüentemente, aumentar a variação entre agrupamentos distintos. O principal algoritmo dessa categoria é o *K-Means*, devido a sua simplicidade, em termos computacionais, e abrangente aplicação no meio científico. Nesse algoritmo são definidos os centros, os quais além de descrever a quantidade de agrupamentos do problema, também servirão de base para a alocação dos objetos, por meio de alguma medida de similaridade, formando agrupamentos compactos e esféricos.

A Figura 3 mostra exemplos de formações de grupos. (a) Os elementos são atribuídos para seus grupos aleatoriamente após o cálculo dos centróides (círculos maiores). (b) Os elementos são redistribuídos aos grupos de centróides mais próximos. (c) Centróides recalculados e grupos em sua forma final. Se não for o caso, repetem-se os passos (b) e (c) até que não haja mudanças nos centróides (Linden, 2009).

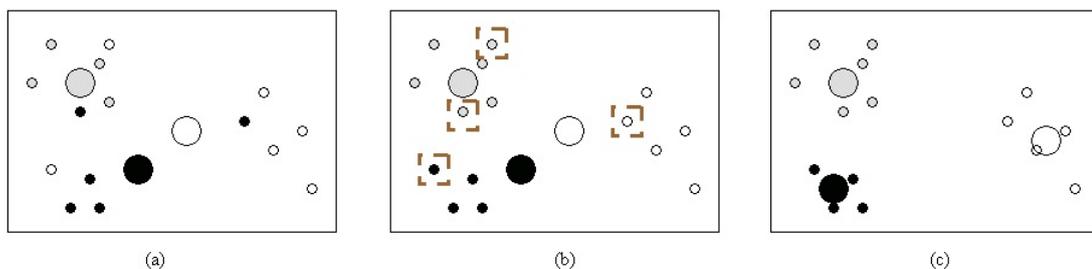


Figura 3 – Exemplo de execução do *K-Means*. (Fonte: Linden (2009))

### 2.2.2 Algoritmos Probabilísticos

As técnicas probabilísticas estão mais comumente norteadas pela geração de dados através da mistura de distribuições de probabilidades subjacentes. De forma mais simples, os agrupamentos podem ser representados por Gaussianas ou distribuições de *Poisson*, descritas como distribuições probabilísticas paramétricas. Sendo assim, o processo de desenvolvimento e distribuição de agrupamentos torna-se uma questão de estimativa de

parâmetros, visto que os dados são modelados através da mistura de  $k$  distribuições de componentes (Aggarwal e Reddy, 2013).

Um dos mais importantes algoritmos probabilísticos é o EM, descrito inicialmente por Dempster et al. (1977) em seu trabalho. A técnica *Expectation-Maximization* é voltada para a descoberta de estimativas de coerência máxima entre parâmetros presentes em modelos estocásticos, os quais necessitam de variáveis latentes. Esse algoritmo segue dois passos básicos: etapa de “expectativa”, seguida pela etapa de “maximização”. Sua importância pode ser comprovada pela ampla possibilidade de aplicação da teoria, bem como por sua simplicidade (Dempster et al., 1977).

Para realização do agrupamento é utilizado um modelo de Misturas Finitas de Gaussianas. Esse modelo generativo atua no cálculo iterativo do conjunto de parâmetros, visando alcançar um valor de convergência desejado. A mistura é descrita como um conjunto de distribuições de probabilidade  $K$ , onde cada uma das distribuições concerne a um dos aglomerados. Dessa forma, uma instância é atribuída com a probabilidade de união para cada agrupamento (Sammut e Webb, 2011).

O algoritmo EM é iniciado com a suposição de parâmetros, médias, covariâncias e mistura de probabilidades. Após isso, os dois passos principais - “Expectativa” e “Maximização” - são repetidos alternadamente. Na etapa de “Expectativa”, os parâmetros atuais são usados na determinação das probabilidades posteriores; já na etapa de “Maximização”, a verossimilhança é maximizada baseada nas probabilidades atualizadas e, em paralelo, as médias, covariâncias e coeficientes são recalculados (Aggarwal e Reddy, 2013). A Figura 4 ilustra o funcionamento básico do algoritmo EM com dois componentes Gaussianos.

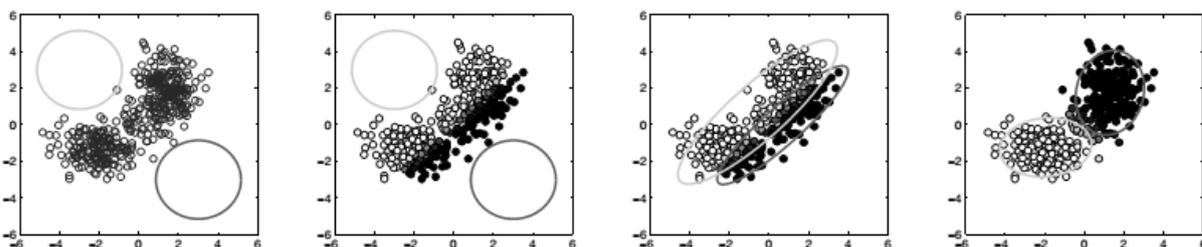


Figura 4 – Exemplo de execução do EM com dois componentes Gaussianos. (Fonte: Aggarwal e Reddy (2013))

### 2.2.3 Medidas de Proximidade

No processo de agrupamento deve-se ter como base alguma medida de similaridade ou dissimilaridade, que irá definir a alocação de objetos similares em um determinado

grupo. Essas medidas são de essencial importância nos algoritmos de agrupamento, pois fornecem critérios para definir se objetos próximos devem realmente compartilhar de um mesmo grupo. Existe uma grande variedade de medidas de similaridade e dissimilaridade. Dentre elas, podemos citar as medidas de correlação (Correlação de *Pearson* e Cosseno) e distância (Distância Euclidiana e *Mahalanobis*). Para desenvolvimento do presente trabalho, buscou-se focar em métricas mais conhecidas e abordadas no meio científico. Sendo assim, a distância euclidiana e a correlação de *Pearson* foram selecionadas, sendo descritas nas seções seguintes.

### 2.2.3.1 Correlação de *Pearson*

A Correlação de *Pearson* tem como finalidade medir a similaridade direcional entre dois pontos, não levando em consideração o tamanho dos atributos. O valor de saída do cálculo de correlação pode variar entre -1 e 1, onde: -1, são inversamente equivalentes; 0, os pontos não tem relação; 1, existe similaridade entre os pontos (Araújo, 2008). Denota-se a correlação de *Pearson* da seguinte forma:

$$Pearson(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}_e^T \mathbf{y}_e}{\|\mathbf{x}_e\| \|\mathbf{y}_e\|} \quad (2.1)$$

sendo  $\mathbf{x}_e = [x_1 - \bar{x}, \dots, x_d - \bar{x}]$  e  $\mathbf{y}_e = [y_1 - \bar{y}, \dots, y_d - \bar{y}]$  apresentam-se como vetores de diferença, e  $\bar{x} = \frac{1}{d} \sum_{i=1}^d x_i$  e  $\bar{y} = \frac{1}{d} \sum_{i=1}^d y_i$  (Araújo, 2008).

### 2.2.3.2 Distância Euclidiana

A Distância Euclidiana representa uma das medidas de distância mais populares no meio de pesquisa. Levando-se em consideração o agrupamento em si, esse tipo de distância funciona como medida de dissimilaridade. Resumidamente, podemos entender da seguinte forma: quanto mais próximos os pontos, mais similares eles são; quanto mais afastados, mais diferentes. Por sua vez, podemos defini-la de acordo com a seguinte expressão:

$$d(i, j) = \sqrt{\sum_{k=1}^r (X_{i_k} - X_{j_k})^2}$$

onde  $X_{i_k}$  e  $X_{j_k}$  representam as projeções dos pontos  $i$  e  $j$  na dimensão  $k$ ; ( $k = 1, 2, \dots, r$ ) (Green et al., 1989). Basicamente, na Distância Euclidiana podemos tomar  $d_{i,j}$  como distâncias físicas entre os pontos  $p$ -dimensionais  $\mathbf{x}'_i = (x_{i_1}, \dots, x_{i_p})$  e  $\mathbf{x}'_j = (x_{j_1}, \dots, x_{j_p})$  em um espaço Euclidiano (Everitt et al., 2011).

## 2.3 Processamento Digital de Imagens

A área de processamento de imagens se refere à manipulação de *pixels*, visando a melhoria na apresentação da imagem, o realce ou a eliminação de certas características, bem como a extração de informações (Gonzalez e Woods, 2000).

Entre as principais aplicações do Processamento de Imagens, temos a atenuação de ruídos, a segmentação de estruturas, o reconhecimento de padrões, dentre outras. Essa manipulação possibilita a obtenção das informações requeridas para, posteriormente, conduzir à uma tomada de decisão.

### 2.3.1 Representação de Imagens Digitais

Existem, basicamente, duas formas principais de se representar uma imagem digital: *Bitmap* e Vetorial. Os mesmos serão abordados de uma forma mais aprofundada a seguir.

- *Bitmap*: a composição de uma imagem é dada por um conglomerado de pontos, os *pixels*. Esses *pixels* são organizados em um “Mapa de Bits” (*Bitmap*), como uma matriz de pontos, disposta na tela da máquina de visualização em questão (computador, por exemplo). Cada elemento formador desse mapa traz consigo informações que irão definir a cor e as coordenadas daquele determinado ponto no plano geral da imagem. Além disso, a quantidade de *pixels* existentes vertical e horizontalmente formarão a resolução da imagem em questão.
- Vetorial: diferentemente da forma apresentada anteriormente, a representação de dados em suas características, que vão desde cor, tamanho e linhas/curvas vetorial, são definidas através de equações matemáticas. Essas singularidades permitem que a imagem seja modificada, em seu tamanho, forma e cor, sem que aconteça perda de resolução.

Quanto à codificação das cores das imagens, existem inúmeras vertentes, que abrangem a RGB, a CMYK e a HSV. Para realização dos testes abordaremos a codificação RGB, a mais comum entre todas as opções, e que se encaixa na proposição. Como definição, a codificação RGB (referente à “vermelho, verde e azul”), de acordo com Frery e Perciano (2013), representa um dos espaços de cores mais utilizados, e está voltado especialmente para imagens digitais de 8 *bits*. Esse modelo é baseado na teoria da visão de cores tricromática de *Young-Helmholtz* e no triângulo de *Maxwell*. Basicamente, as cores são formadas a partir da combinação de vermelho, verde e azul, em diferentes intensidades. Os *pixels* de uma imagem RGB tem, na forma mais simples, 8 *bits* de profundidade; esse valor resulta em um intervalo de [0, 255] de intensidade;

## 2.4 Histogram Oriented Gradients

O *Histogram Oriented Gradients* (HOG) caracteriza uma técnica do ramo da visão computacional ligada à extração de características de imagens. Sua preponderância está relacionada principalmente ao reconhecimento de objetos específicos em ambientes, bem como redução de dimensionalidade.

A definição de HOG pode ser encontrada em Dalal e Triggs (2005), o qual é caracterizado pela geração de histogramas locais da orientação dos gradientes de imagens em um denso *grid*. A ideia central de contribuição dessa estratégia de extração é que as características visíveis (aparência e forma) podem ser expressadas de maneira eficiente através da distribuição de gradientes de intensidade locais ou direções das bordas, mesmo que não se saiba previamente a posição destes (Dalal e Triggs, 2005). De forma sucinta, Gritti et al. (2008) define a essência do método como o uso da distribuição dos gradientes de intensidade dos *pixels*, bem como o direcionamento das bordas, na descrição da aparência e forma dos objetos presentes em uma imagem.

O funcionamento prático da extração das características HOG pode ser descrito pela divisão da imagem em pequenas regiões espaciais, denominadas “células”, onde cada célula acumula um histograma 1-D local das direções dos gradientes sobre os *pixels*. A combinação desses histogramas geram a representação da imagem. Para obtenção de resultados mais apurados, existe a possibilidade de realizar uma normalização de contraste, visando um melhor desempenho na variação de iluminação e sombreamento. Essa normalização é calculada através da medida de intensidade de regiões maiores, denominadas “blocos”, sendo essa medida aplicada às células presentes no bloco (Dalal e Triggs, 2005). Ainda segundo Dalal e Triggs (2005), os valores recomendados para parametrização são: máscara derivativa 1-D centrada  $[-1, 0, +1]$ , janela de detecção de  $64 \times 128$  *pixels*, células de proporção  $8 \times 8$ , e blocos de proporção  $16 \times 16$ .

A Figura 5 explica, de modo geral, o funcionamento procedural da técnica, realizado em cinco passos básicos, estendendo-se até a geração das características trabalhadas da imagem.

A Figura 6 ilustra o funcionamento dos procedimentos HOG de forma visual. Em (a) é apresentada a imagem original a ser trabalhada; no passo (b) é definida a magnitude do gradiente, por meio dos valores absolutos obtidos junto ao operador de diferenças melhorado; em (c) os *grids* são aplicados sobre as direções de bordas da imagem; finalizando, (d) expõe o agrupamento de células sobre os histogramas de gradientes.

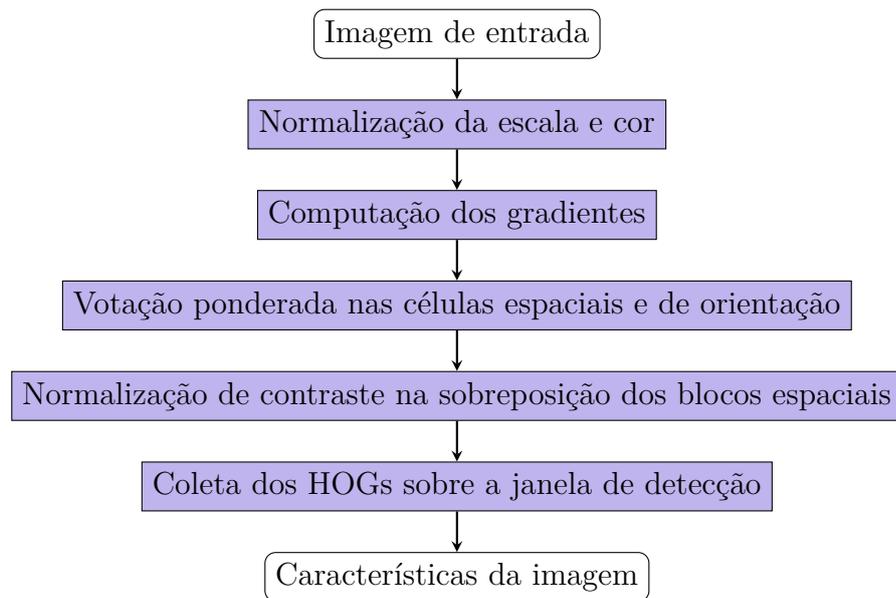


Figura 5 – Passos de execução da técnica HOG. (Adaptação: Dalal e Triggs (2005))

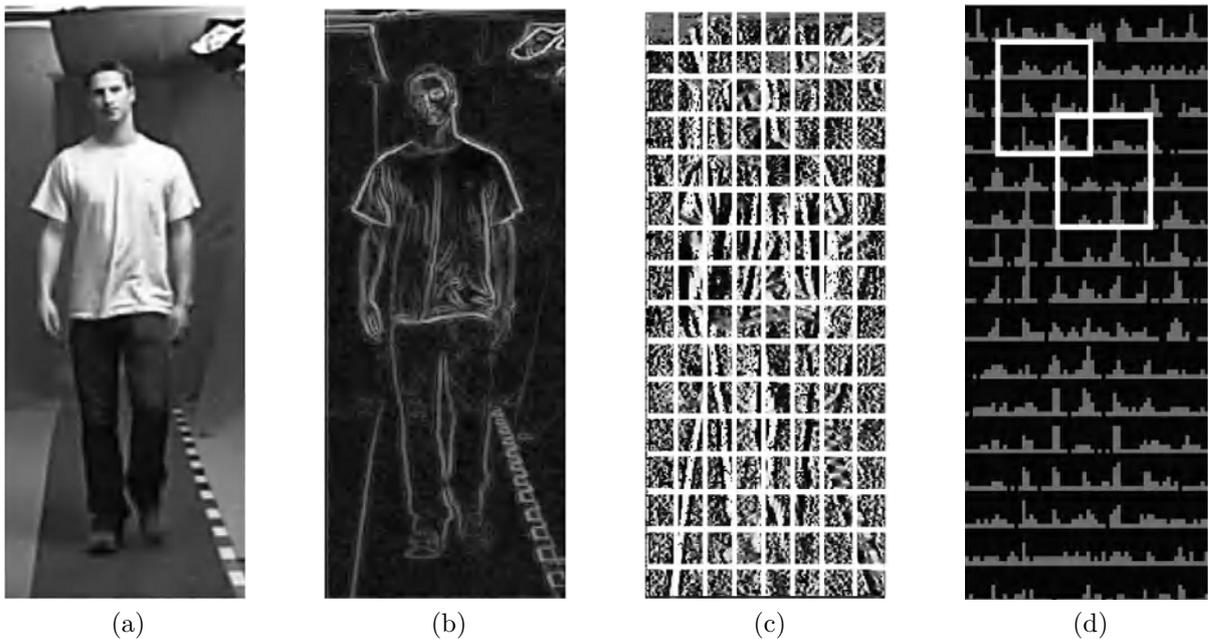


Figura 6 – Demonstração do processo HOG (Adaptação: Nixon e Aguado (2012))

### 2.4.1 Agrupamento de Imagens

O agrupamento de imagens representa uma importante vertente da visão computacional, visto que grandes bases de imagens são difíceis de se trabalhar, principalmente em processos de busca. Seu foco principal visa a descoberta não-supervisionada de aspectos visuais, como objetos ou propriedades da cena, que sejam comuns entre imagens em um aglomerado. Essas características visuais compartilhadas

possibilitam a formação de grupos mais refinados a partir de uma base de imagens (Faktor e Irani, 2012).

Em sua pesquisa, Goldberger et al. (2002) define o agrupamento de imagens como um meio de alto nível para descrição das figuras e seu conteúdo. Seu foco é classificar imagens em grupos, onde os mesmos sejam capazes de representar um resumo conciso, bem como visualização do conteúdo das mesmas. A principal aplicação dessas técnicas se encontra na pesquisa e recuperação eficientes de grandes bancos de dados de imagens. O processo de otimização acontece da seguinte forma: a busca das imagens aconteceria apenas no grupo de maior semelhança, evitando assim a busca exaustiva e economizando tempo computacional e de processamento.

Existem diversas aplicações possíveis para o agrupamento de imagens; entretanto, são poucas as pesquisas desenvolvidas na área. Dentre elas, podemos citar:

- Em sua pesquisa, Belhumeur et al. (2008) propõe uma forma automatizada de atribuição de espécies de plantas. Através de uma base de imagens, as plantas são classificadas por espécies, de acordo com características singulares das mesmas. Assim, pode-se otimizar o processo de identificação tradicional, que demanda uma tarefa manual de especialistas na área em questão;
- Pode-se verificar a utilização de técnicas de agrupamento de imagem em fotos digitais em Cooper et al. (2003), onde é proposto um método para reunião de figuras por meio da semelhança temporal.

## 2.5 Considerações Finais

Este capítulo trouxe os principais assuntos de referência a serem abordados no desenvolvimento geral do trabalho. Procurou-se tratar de forma mais concisa as características do agrupamentos e sua análise, descrevendo os algoritmos que irão receber o foco na realização do estudo, além de expor conceitos importantes no que tange ao processamento de imagens e visão computacional. Esses assuntos foram apresentados seguindo uma ordem de congruência para com o tema do trabalho.

O objetivo principal deste capítulo é a definição dos tópicos de preponderância no desenvolvimento do projeto, bem como a obtenção do conhecimento necessário acerca dos diversos assuntos técnicos que baseiam sua construção prática e teórica.

No capítulo seguinte, estarão presentes um apanhado considerável de trabalhos relacionados a técnicas de agrupamento de imagens, reconhecimento facial e otimização de

buscas. Sendo assim, poderemos observar as possíveis melhorias a serem implementadas, bem como aspectos técnicos complementares.

## 3 Revisão Bibliográfica

Neste capítulo, discorreu-se sobre um apanhado de trabalhos relacionados às técnicas de reconhecimento facial e agrupamento de imagens, principalmente sobre estudos que visam a melhoria do processo de extração de características de faces, técnicas de agrupamento de imagens, bem como otimização de busca e recuperação de itens.

Em Xu e Wunsch II (2007) é feito um amplo estudo comparando diferentes algoritmos de agrupamento em cinco problemas do mundo real: Problema do Caixeiro Viajante, Análise da Expressão Gênica de Dados, Agrupamento de DNA ou sequências proteicas, Reconhecimento de Expressões da Face Humana (Redução de Dimensionalidade) e Agrupamento de Documentos. Através dos resultados é possível notar que as técnicas de agrupamento são passíveis de aplicação em domínios distintos. Além disso, o trabalho torna claro que a superioridade de um algoritmo depende de fatores como os parâmetros utilizados para comparação e, acima de tudo, do domínio em questão.

Pautado pelo aumento da ineficiência dos motores de pesquisa web, a proliferação substancial de documentos presentes na Internet, e a necessidade e oportunidade de fornecer respostas de pesquisas mais refinadas, Jiang et al. (2000) propõe uma nova técnica de agrupamento de resultados de pesquisas *Web*. O sistema *Retriever*, como foi denominado, utiliza a ferramenta *Lycos* como motor de pesquisa, e os métodos de espaço vetorial e *N-Gram*, ambos baseados no algoritmo RFCMdd (*Robust Fuzzy C-Medoids*), na definição dos agrupamentos. Os teste propostos demonstraram bons resultados em *Web snippets*, tendo a técnica *N-Gram* uma melhor performance em relação a abordagem de espaço vetorial.

Baseando-se em pesquisas “web” e suas ineficiências, como a falta ou inadequação dos retornos com a informação previamente solicitada, Wang e Kitsuregawa (2001) demonstra uma abordagem de agrupamento baseado em *links* compartilhados pelas páginas. Seu objetivo principal visa agrupar resultados de pesquisas retornados por um motor de buscas *Web* em grupos semanticamente significativos, com abordagens estratégicas de co-citação e acoplamento. Para testes, foi desenvolvida uma variação do *K-Means* mais adaptada para lidar com ruídos. Os resultados advindos mostram que, em média, o agrupamento baseado em *link* apresentou melhores retornos que o agrupamento baseado em *snippet*.

Em um estudo relacionado à avaliação de conjuntos de recursos para fins de robustez no reconhecimento visual de objetos, Dalal e Triggs (2005) analisa descritores baseados em

gradiente e borda, bem como a influência de seus processos formadores no desempenho final. Seu projeto é baseado no desenvolvimento do reconhecimento facial aliado aos descritores HOG (*Histogram Oriented Gradients*). Na realização dos experimentos, visando a quantificação de performance, o método HOG desenvolvido foi comparado às abordagens *wavelets* de *Haar* generalizadas, PCA-SIFT, e técnicas de contexto de forma. Em seguida, as performances dos processos de normalização de cor/gama, computação de gradiente, orientação do encaixotamento, normalização e blocos de descritores, janela de detecção e contexto, e o classificador. Como resultados, pôde-se concluir que a utilização de histograma de gradientes orientados locais se mostrou eficiente na detecção de pessoas e diminuição de falsos positivos. Além disso, ficou comprovada a importância de uma parametrização mais cuidadosa dos descritores para obtenção de uma performance superior.

Visando uma melhor perspectiva relacionada a reconhecimento facial, Déniz et al. (2011) propõe o desenvolvimento de uma técnica HOG robusta. O trabalho está orientado em três objetivos gerais: construção de um descritor HOG usando um *grid* regular, para evitar problemas advindos de diferentes iluminações, poses e oclusões; combinar os descritores HOG em diferentes escalas; e promover uma redução de dimensionalidade linear, focando a resolução de ruídos. Para a localização das características, foram utilizadas *Active Appearance Models* (AAMs), além de PCAs (*Principal Component Analysis*) e LDAs (*Linear Discrimination Analysis*). Como resultado, obteve-se um aumento de quase 13% em performance de reconhecimento, em detrimento da performance de uma representação holística, sob a mesma parametrização.

Tendo como base o problema de lentidão de processamento enfrentado pela técnica HOG+SVM, Pang et al. (2011) apresenta duas soluções para enfrentar esse problema, prezando por agilizar o processo, sem que haja perda de qualidade da classificação. O primeiro passo consiste em reutilizar as características em bloco, aspirando a construção das especificidades HOG para as janelas de detecção de intersecção. A segunda forma tem como base a interpolação baseada em subcélula para cômputo eficiente das características HOG para cada bloco. Os resultados advindos dos testes apresentaram um aumento de até cinco vezes na detecção de humanos em imagens *top-view*.

Objetivando explorar a lacuna de trabalhos envolvendo a técnica HOG em reconhecimento facial, Shu et al. (2011) submeteu o HOG original e suas variações a testes, a fim de avaliar o efeito de diferentes fatores. Dentre os vários fatores experimentados, pode-se concluir que: na representação angular, o HOG em reconhecimento facial mostra-se mais eficiente devido a limitação imposta pela região da máscara facial, que exclui o plano de fundo; quanto mais orientações dos compartimentos, mais eficiente será a estrutura de padrão de face repassada pelo descritor; escala e tamanho da célula também influenciam no

desempenho da estrutura, onde sua diminuição positiva os resultados (todavia, a tolerância de variação intra-classe também diminui); a sobreposição impõe maior robustez para variações de pose ou expressões. Como resultados, pode-se comprovar que o uso de HOGs em reconhecimento facial possui praticamente o mesmo desempenho que descritores *Gabor*, e ainda melhor em detrimento de descritores LBP (*Local Binary Patterns*).

Ciente da importância do processo de agrupamento de imagens como meio não-supervisionado de criação de estruturas para formação de grupos complexos, Saffer (2011), em seu trabalho, compara os resultados da utilização dos algoritmos *K-Means*, *K-Median*, PSO e um híbrido de PSO e *K-Means*, aplicados em três imagens da massa cinzenta cerebral. Foi verificado então que uma imagem com uma maior quantidade de áreas favorece um melhor resultado, e o PSO e PSO híbrido obtiveram desempenhos melhores com imagens mais detalhadas; e medidas de compacticidade foram eficientes em todos os algoritmos.

Diante do crescimento das já enormes bases de dados de imagens digitais, desenvolver métodos de armazenamento e busca eficientes e otimizados tornou-se um desafio. Em sua pesquisa, Alattab e Kareem (2012) propõe um método de detecção de imagens faciais, além da recuperação dessas imagens nas bases de dados. No desenvolvimento do trabalho, foram abordadas ferramentas diversas, dentre elas: o descritor *Eigenfaces* foi selecionado com o intuito de realizar a extração das características faciais; o histograma de cor é usado como conteúdo visual (as representações de cores utilizadas foram RGB (*Red, Green and Blue*), HSV (*Hue, Saturation and Value*), e HSI (*Hue, Saturation and Intensity*)); para detecção e dimensionamento da face, o método *Viola-Jones* é empregado; na mensuração das distâncias, a medida escolhida foi a distância euclidiana; como instâncias, foi selecionada uma base de dados de 1500 imagens faciais de 150 pessoas diferentes (com expressões variadas). Como resultados, na utilização do método *Eigenfaces*, o espaço de cor HSV obteve maior acurácia dentre os demais; já na recuperação baseada em histograma de cor, o RGB demonstrou melhor performance.

O provimento dos *gadgets* trouxe consigo a dinamicidade e popularização da captura de imagens dos mais diversos tipos. Esse acontecimento acarretou no crescimento das bases de dados, bem como o posterior interesse científico na exploração desses *datasets*. Pautado na otimização dos ditos sistemas de Recuperação de Imagem Baseada em Conteúdo (RIBC), Obulesu et al. (2015) propõe um novo método de auxílio na extração de características faciais baseado em PBL (Padrões Binários Locais). O método consiste em prover um vetor de peculiaridades das características locais e regionais (que podem descrever eficientemente a face humana), usando PBL regional baseado em mediana. Além disso, visando correções de variação de iluminação e sombreamento local, são aplicadas técnicas de Correção Gama

(CG) e Diferença de Gaussiana (DG). Para experimentação, o método foi aplicado nas bases de dados FG-NET e *Google*, fazendo uso da medida de distância Qui-Quadrado; para cálculo de performance, a técnica de *Recall and Precision* foi escolhida. Os resultados comprovaram a superioridade do uso da mediana, em detrimento da mínima e máxima PBL, com apuração de 60% de ganho.

Vivemos em um mundo “digitalizado”, onde a quantidade dos mais variados dados cresce constantemente. Baseado nisso, o desenvolvimento de novas tecnologias que visam a recuperação eficiente de dados cresce. Como área importante de desenvolvimento, destaca-se a RIBC, provendo recuperação automática de imagens partindo de características mais preponderantes. Visto isso, Desai e Sonawane (2017) aborda, em seu trabalho, a utilização de três técnicas de extração de características faciais (*Gist*, HOG e DWT (Discret Wavelet Transform)) e seu impacto de performance no sistema RIBC. No desenvolvimento da pesquisa, foi utilizada a base de dados WLB (Weak Labeled Database), que contém 53448 imagens; para classificação do retorno da pesquisa, tornando assim a busca “não-trivial”, foi empregado o algoritmo kNN (*K-Nearest-Neighbor*), e a técnica LSH (Locality Sensitive Hashing), comumente abordada em espaços de alta dimensão. Como resultados, pode-se observar que, entre as três abordagens, o descritor HOG obteve a maior taxa de *recall* e *precision*, e a DWT, os piores resultados.

### 3.1 Considerações Finais

Este capítulo listou uma série de trabalhos com o foco relacionado a técnicas de reconhecimento facial, otimização de buscas ligadas a algoritmos de agrupamento, e estudo da abordagem de técnicas de escalonamento em diversos cenários diferentes.

Como objetivo principal, a seção presente trouxe a possibilidade de avaliar de forma mais abrangente, por meio do estado-da-arte, a realização do melhoramento de resultados em buscas de imagens faciais, seja pela proposição e desenvolvimento de técnicas de extração de características preponderantes, seja pela possibilidade de utilização do escalonamento das imagens como forma de otimização dessa pesquisa. Esse entendimento proporcionou o desenvolvimento de um projeto mais robusto e relevante.

No próximo capítulo será demonstrada a proposta de trabalho em si, expondo o problema a ser resolvido e sua conseqüente resolução, diante do conhecimento adquirido na pesquisa.

## 4 Recuperação de Imagens com Auxílio de Técnicas de Agrupamento

A análise de agrupamentos, dentro do contexto de imagens, constitui um importante segmento da Inteligência Computacional, com a presença de diversos trabalhos científicos que abordam a sua prática (Zhang et al. (2007), Goldberger et al. (2002), Petcu et al. (2011), Civicioglu et al. (2014), Faktor e Irani (2012), entre outros). A aglomeração de dados representa um método de organização de grandes conjuntos de informações com o intuito de promover, de uma forma mais simples e conveniente, sua interpretação e estudo. Com o crescimento exponencial dos dados e bases de dados, principalmente por meio da *World Wide Web*, a necessidade de sumarizar informações tornou-se imprescindível.

O presente capítulo versa sobre a descrição do problema enfrentado na busca de imagens faciais em grandes bases de dados. O processo usual, onde a imagem é pesquisada de forma “gulosa” (ou seja, realizando uma comparação entre a instância buscada e cada uma das imagens formadoras da base), nem sempre é o mais indicado devido ao custo de tempo computacional empregado, principalmente em se tratando de grandes conjuntos de imagens, já populares e cada vez mais presentes no meio acadêmico e social. Para encarar essa questão é proposta a realização de um pré-processamento por meio de agrupamentos, visando o encurtamento do caminho a ser percorrido na tarefa de retornar o resultado desejado. A seguir, são dispostos aprofundamentos dessas questões.

### 4.1 Descrição da Problemática

Como foi citado no Capítulo 1, a abrangência das imagens vêm alcançando as mais diversas áreas de estudo. A rapidez de proliferação e fortalecimento de conjuntos de dados em diversas vertentes, incluindo textos, áudios, geolocalizações e contextos associados, torna a diversificação dos dados ainda maior. É importante notar que essas características estão intimamente ligadas ao um novo contexto, muito citado na literatura, conhecido como *Big Data* (Wu et al., 2014).

A quantidade de imagens disponíveis e a pluralidade de aplicações traz consigo algumas implicações. Dentre elas está a dificuldade de recuperação desses dados. Realizar buscas em *datasets* de grandes proporções torna-se uma tarefa árdua, no que tange ao esforço computacional. Tendo em vista a orientação primordial de comparação de similaridade - as técnicas ditas “gulosas”, as quais demandam uma averiguação total da

base de dados para identificação do objeto idêntico ou mais semelhante - o processo em si, bem como os resultados posteriores, podem ser otimizados.

Esse cenário traz consigo a possibilidade do desenvolvimento de estratégias que viabilizem respostas mais rápidas. Para que isso aconteça são necessárias implementações que complementem a pesquisa em si, fugindo assim do procedimento usual. O uso de Inteligência Computacional, nesse sentido, é bastante viável. Os traços dessas aplicações podem ser abstraídas dos estudos de Jiang et al. (2000) e Wang e Kitsuregawa (2001). Se as imagens forem determinadas por uma série de características e agrupadas de acordo com as singularidades comuns entre elas, seria possível direcionar a busca para um determinado setor (baseado na semelhança) e diminuir o tempo de resposta na identificação de um rosto (caso ele exista na base), ou retornando a imagem mais semelhante. Sendo assim, as técnicas de agrupamento despontam como ferramentas de aperfeiçoamento dos procedimentos de busca.

Note que o agrupamento de imagens, enquanto processo de otimização, pode ser usado em qualquer trabalho relacionado a busca e inserção em bancos de imagens. Suponha, por exemplo, uma base policial com rostos de criminosos. Os registros são feitos (com mais de uma imagem, possivelmente) e todas elas são armazenadas indiscriminadamente em um banco de dados (mais de um, possivelmente). A busca e recuperação de um indivíduo, a partir da captura de uma imagem, pode inviabilizar uma ação efetiva das autoridades responsáveis, uma vez que a averiguação tem que obrigatoriamente verificar todos os registros.

Uma outra potencial forma de aplicação do processo de agrupamento em imagens está ligada ao reconhecimento facial, bem como ao procedimento de buscas em grandes bases de dados. De acordo com características extraídas dos elementos (rostos dos indivíduos), a implantação de técnicas de agrupamento poderia contribuir na dinamização e redução do tempo das buscas em grandes aglomerados de dados. Contextualizando, é possível utilizar os grupos formados, definidos pelo refinamento do algoritmo utilizado, aliado a seus diversos parâmetros, para delimitar um plano de busca mais otimizado.

É evidente a possibilidade de trabalho e prospecção relacionada a recuperação de imagens faciais em grandes bases de dados. Sendo assim, a implantação do pré-processamento dos dados norteados por técnicas de escalonamento torna-se uma alternativa viável. Entretanto, além do desenvolvimento da técnica em si, é de suma importância a comprovação da eficácia do método, por meio da avaliação do tempo computacional despendido, bem como a acurácia no retorno da imagem. Ou seja, além de tornar o processo de busca realmente mais ágil, deve-se prover resultados confiáveis, onde a imagem retornada seja realmente relevante e condizente com o item desejado, ou o

parâmetro previamente passado.

## 4.2 Solução Implementada

Diante da oportunidade de estudos correlatos às áreas de otimização de buscas e reconhecimento facial, como exemplificados nos trabalhos de Alattab e Kareem (2012), Obulesu et al. (2015) e Desai e Sonawane (2017), além da importância desse tema no âmbito científico/acadêmico, este trabalho se propõe a analisar o impacto da utilização de métodos de agrupamento no auxílio à otimização de buscas ligadas ao reconhecimento facial de imagens. A análise pretende, acima de tudo, a definição de eficiência relacionada ao tempo computacional dispendido no processo de comparação/recuperação, e eficácia, onde o resultado fornecido pela busca é realmente relevante e condizente com o esperado.

Para isso, partimos da suposição que as técnicas de agrupamento são ferramentas para organizar os dados de acordo com características semelhantes compartilhadas entre objetos. Dessa forma, podemos pensar na divisão dos dados em compartimentos, onde os objetos presentes em cada um deles sejam semelhantes de acordo com algum critério estabelecido.

Se esses conceitos foram levados para o contexto de imagens faciais de pessoas, a utilização de técnicas de agrupamento deve conseguir construir uma organização dos dados de forma que imagens de pessoas parecidas compartilhem um mesmo grupo, enquanto que imagens de pessoas com característica distintas sejam colocadas em grupos diferentes.

Esse princípio é fundamental para a ideia central deste trabalho. Considere, primeiro, a forma tradicional de se realizar busca em bases de dados com imagens de faces. Para identificar uma pessoa, é necessário ter uma amostra de imagem a ser testada. Normalmente, essa amostra é capturada através de algum sensor e, a partir disso, é feita uma busca em uma base de dados por uma imagem correspondente. Dessa forma, é possível obter alguma informação adicional dessa pessoa com base em registros armazenados separadamente. A Figura 7 ilustra a forma como isso é feito.

Diante do exposto, é possível notar que, a medida que uma base de dados cresce, o processo iterativo de uma busca trivial pode se tornar inviável, já que o princípio básico desse processo consiste em comparar a imagem teste com as imagens do banco de dados, até que seja encontrado um item que atenda os critérios estabelecidos para determinar que as imagens são da mesma pessoa.

É importante notar que o custo computacional para executar uma busca em uma grande base de dados pode impactar negativamente ou até inviabilizar algumas ações decorrentes da identificação do indivíduo presente na imagem. Como mencionado antes,

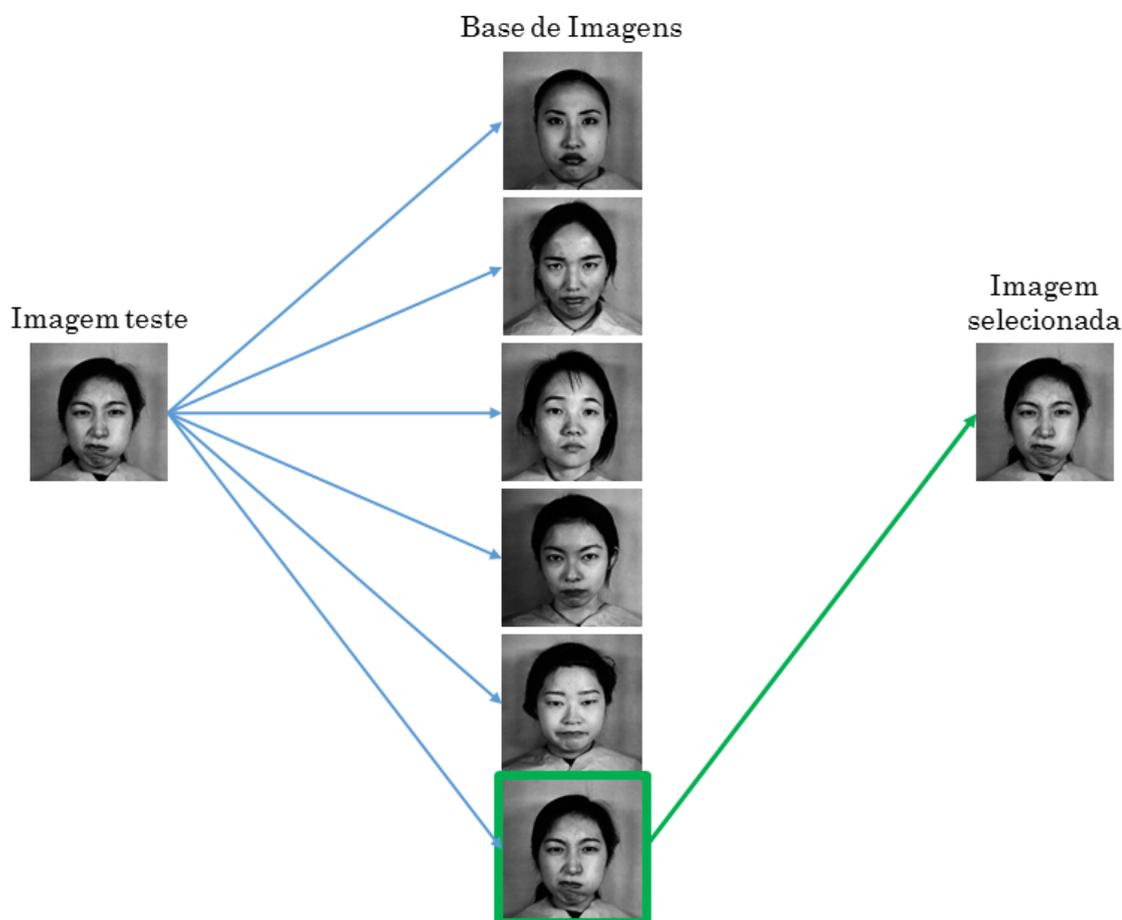


Figura 7 – Busca direta no conjunto de dados.

se uma ação policial, por exemplo, depende da identificação de uma pessoa que teve sua imagem capturada através de algum dispositivo de vídeo, o tempo gasto para obter as informações necessárias pode determinar o sucesso ou fracasso da ação.

Note que o tempo despendido para realizar esse procedimento deve ser grande porque o algoritmo de busca precisa comparar a imagem teste com todas as imagens do banco de dados, independentemente da relação direta, ou ausência de relação com o indivíduo em questão. Em algumas situações, são criados bancos de dados distintos para facilitar a busca, fornecendo uma “separação” entre itens, como a ação de segregar os seres entre sexo masculino e feminino, por exemplo. Entretanto, esse procedimento, que acontece de forma discricionária e depende diretamente da interpretação do agente manipulador da informação, nem sempre é confiável, o que pode acarretar no comprometimento da credibilidade do sistema, vistos casos de interpretações ambíguas e desconformes do contexto ideal ou até mesmo a disponibilidade do agente em si, para realização dessa triagem.

### 4.2.1 Sistema Proposto

Diante do explicitado anteriormente, se usarmos a ideia de dividir os dados em grupos, podemos reduzir o escopo a uma ou mais regiões de interesse no espaço de busca. A ideia de similaridade, fortemente presente nos conceitos da análise de agrupamentos, pode ser usada para determinar a região do espaço de busca onde a imagem desejada se encontra.

#### 4.2.1.1 Passo 1 - Agrupamento da base

O funcionamento geral desta abordagem começa com a divisão da base de dados em grupos. A Figura 8 ilustra uma pequena base de dados com seis imagens e uma possível divisão em grupos das imagens presentes na base.

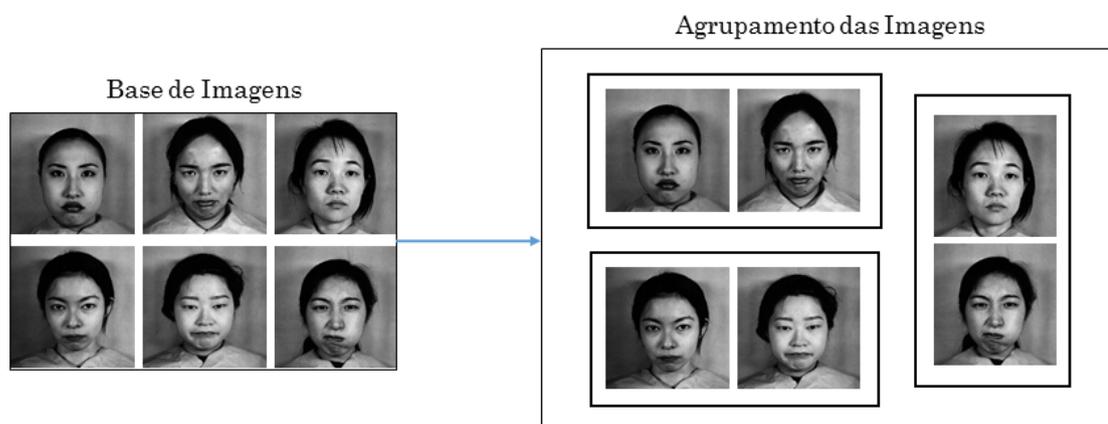


Figura 8 – Primeira etapa da busca com auxílio de agrupamento no conjunto de dados: divisão do conjunto de dados em grupos.

No caso deste exemplo, a base foi dividida em três grupos, contendo imagens de pessoas fisicamente semelhantes, de acordo com os critérios abordados. Para realização desse procedimento, é necessário o uso de uma medida de similaridade.

#### 4.2.1.2 Passo 2 - Verificação da semelhança entre imagem teste e grupos

Uma vez definida a organização dos dados, ao invés de prover uma busca “cega”, podemos direcioná-la para uma região onde a probabilidade de encontrar o que se procura é maior. Tal direcionamento é feito através do cálculo de similaridade entre a imagem teste e os grupos definidos no processo de agrupamento (representados pelos seus centróides).

Na implementação proposta neste trabalho, a imagem teste é comparada com todos os grupos que, na sequência, baseada na “semelhança” ou adequação entre ambos, são

ordenados. Dessa forma, a busca é iniciada no grupo cuja similaridade é maior, sendo passada para os demais grupos, em ordem de semelhança. Esse procedimento fornece uma organização prévia, sendo responsável por possibilitar a otimização da recuperação da imagem. A Figura 9 ilustra esse passo do procedimento, indicando em verde o primeiro grupo a ser buscado.

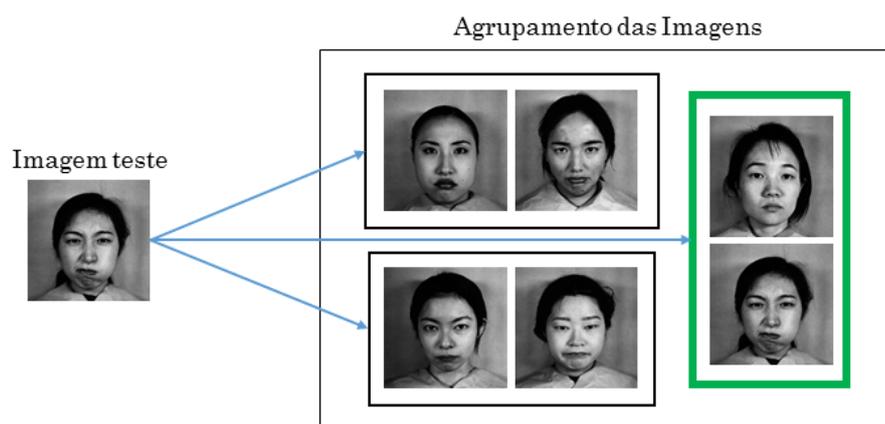


Figura 9 – Segunda etapa da busca com auxílio de agrupamento no conjunto de dados: determinação da proximidade dos grupos com a imagem teste.

### 4.2.2 Passo 3 - Checagem interna nos grupos

Dentro de cada grupo, as imagens são comparadas uma por vez com a amostra de teste. Para cada par, a similaridade é calculada e, caso o valor atinja um determinado limiar, o algoritmo considera que a aquela imagem corresponde ao item procurado. Essa etapa do procedimento pode ser visualizada na Figura 10.

A forma utilizada para determinar o limiar de identificação de uma imagem é definida a partir de uma análise estratificada dos dados. Esse processo visa decidir qual o valor exato de similaridade necessário para distinguir se duas imagens pertencem a uma mesma pessoa, ou se a pessoas diferentes. Mais detalhes sobre o cálculo desse limiar podem ser visualizados na Seção 6.2.2.

Vale salientar que, se a imagem de uma pessoa a ser buscada no banco de dados é exatamente igual àquela previamente armazenada, a definição do limiar é dispensada, bastando buscar por uma imagem cuja similaridade possui o valor máximo (ou mínimo), de acordo com a medida usada para calcular tal similaridade. Ou seja, a imagem teste e o item presente na base devem ser idênticos, apresentando assim um nível máximo de semelhança, uma situação facilmente tratada pelas características das medidas abordadas.

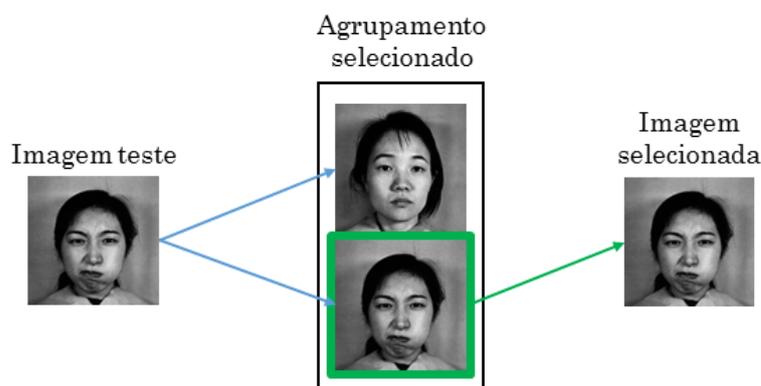


Figura 10 – Terceira etapa da busca com auxílio de agrupamento no conjunto de dados: busca local dentro dos grupos.

### 4.2.3 Descritores HOG

Um ponto que merece destaque é a forma como as imagens são representadas para os algoritmos de agrupamento e de busca. Conforme mencionado na Seção 2.2, os algoritmos de agrupamento organizam os dados com base nas suas características. Dessa forma, a maneira como uma imagem é representada determina o arranjo dos grupos executado pelos algoritmos.

A técnica escolhida neste trabalho, que intenta a extração das características das faces presentes nas imagens, é a de histogramas de gradientes orientados (Seção 2.4). Os histogramas de gradientes orientados constituem uma técnica comumente usada para identificar bordas e nuances em imagens. No contexto deste projeto, são capazes de identificar partes representativas do rosto de uma pessoa, como forma e textura, necessários para o reconhecimento facial em si.

Para garantir que as características sejam extraídas da melhor maneira possível, algumas transformações devem ser realizadas na imagem. A seguir, determinamos os passos seguidos na obtenção das características faciais de uma imagem original, por meio da abordagem HOG:

1. Inicialmente, são extraídos os dados RGB das imagens de teste e da base de dados. A Figura 11 representa uma das instâncias utilizadas na experimentação, em seu estado original;
2. Munido dos dados RGB, é realizado o processo de detecção e delimitação da região facial a ser trabalhada. A Figura 12 demonstra o resultado desse processamento;



Figura 11 – Imagem em seu estado original. (Fonte: Thomaz e Giraldi (2010))



Figura 12 – Seleção da face. (Adaptação: Thomaz e Giraldi (2010))

3. Na sequência, é definida uma escala de trabalho. No presente projeto, o fator de escala na execução dos experimentos foi reduzido devido ao custo computacional despendido, em detrimento da máquina utilizada no empreendimento dos testes (quanto maior a escala, maior a quantidade de “memória” requerida). Além disso, os dados RGB são transformados em escala de cinza, como presente na Figura 13;



Figura 13 – Mudança de proporção e transformação em escala de cinza (Adaptação: Thomaz e Giraldi (2010)).

4. Para melhor definição dos gradientes, o contraste da imagem é realçado através do histograma de equalização. O resultado pode ser visualizado na Figura 14;



Figura 14 – Melhoramento de contraste (Adaptação: Thomaz e Giraldi (2010)).

5. Para finalização do ensaio relacionado ao processamento da imagem, são extraídas as características HOG. Na Figura 15, é possível visualizar os traços obtidos diretamente, na imagem e na Figura 16, é demonstrado um gráfico dispondo os dados das características HOG extraídas, os quais são usados como parâmetros na realização dos agrupamentos e cálculo de distância entre as imagens. No escopo do presente trabalho, os dados finais obtidos representam um conjunto de 900 características, sensíveis à definição da escala (definida no item 3), variando de 0 a 0.38, aproximadamente.



Figura 15 – Visualização da características na imagem (Adaptação: Thomaz e Giraldi (2010)).

### 4.3 Considerações Finais

Este capítulo dispôs sobre a problemática envolvendo o crescimento massivo das bases de imagens, e a conseqüente dificuldade de realizar buscas nas mesmas. Adicionalmente, a possibilidade de exploração dessas bases é enorme, o que acarreta oportunidades de desenvolvimento de pesquisas relacionadas a diminuição do gargalo computacional que envolve buscas em bases de dados de grandes proporções.

O objetivo principal desta seção é expôr informações sobre como utilizar técnicas de agrupamento no auxílio da busca e recuperação de imagens em bancos de dados. Foram

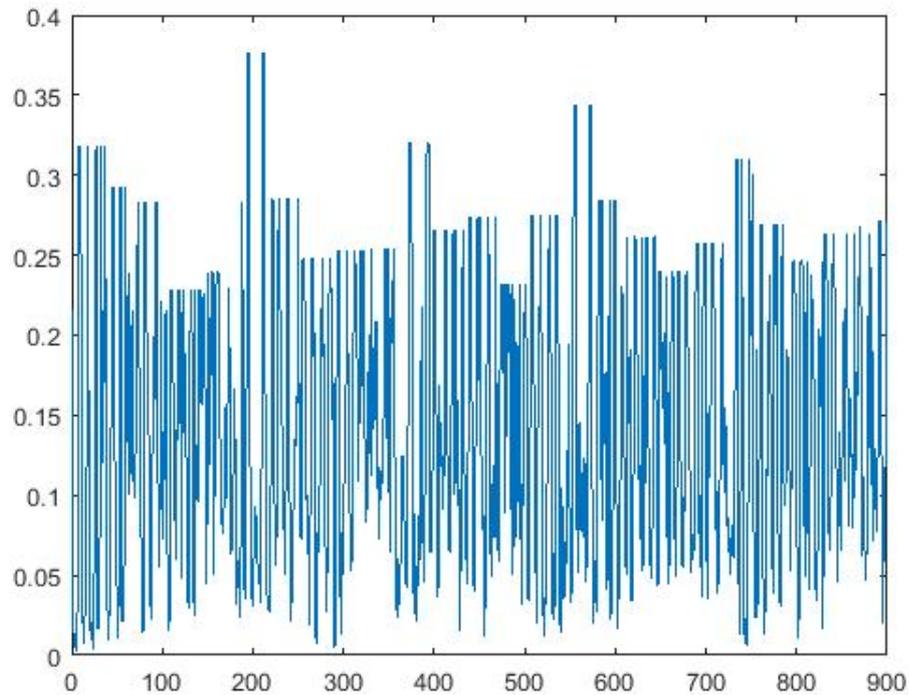


Figura 16 – Características HOG extraídas (Adaptação: Thomaz e Giraldi (2010)).

apresentados os conceitos ligados a representação de imagens e a forma como os algoritmos de agrupamento podem reduzir o custo computacional no procedimento de busca quando comparado a forma tradicional e exaustiva.

No capítulo subsequente, serão definidos os materiais e experimentos, que consistem nas ferramentas computacionais diretamente envolvidas na construção prática do estudo.

## 5 Materiais e Experimentos

Para realização do projeto foram definidas as ferramentas que caracterizam o estudo analítico. Dentre elas estão os algoritmos a serem explorados; as bases de dados escolhidas, que fornecerão as imagens necessárias para os testes; os parâmetros; e as métricas de comparação, decisivas na formulação do resultado final do estudo. Adicionalmente, as diversas métricas e características foram empregadas ao método proposto objetivando a colheita de resultados e posterior análise. Esses elementos são descritos nas seções seguintes.

### 5.1 Bases de Dados

Os dados utilizados no desenvolvimento dessa pesquisa são imagens reais retiradas de repositórios públicos. Existem alguns repositórios bastante utilizados em trabalhos relacionados à imagens. Como exemplo, o repositório *The Berkeley Segmentation Dataset and Benchmark* (Martin et al., 2001) reúne figuras de diferentes cenários em conjunto com designações humanas (usadas para referência) e automáticas de segmentação. O uso de repositórios desse tipo possui uma considerável importância, pois resulta na padronização dos grupo de imagens usadas pelos pesquisadores na área.

Três bases de imagens faciais são abordadas nesse trabalho. Esses *datasets* possuem imagens de pessoas reais, em posicionamento frontal, diferenciando-se em quantidade de instâncias, classes e distribuição de instâncias entre as classes. A Tabela 1 discrimina as características individuais das classes.

Base de dados	$n$	#C	Distribuição das classes
<i>faces94</i>	380	19	20 imagens para cada classe
FEI	400	200	2 imagens para cada classe
JAFFE	213	10	23, 22, 22, 20, 21, 21, 20, 21, 21, 22

Tabela 1 – Descrição das bases de imagens abordadas.

Discorreremos a seguir sobre as bases de dados “*Faces94*”, “FEI” e “JAFFE”, as quais são exploradas no decorrer das pesquisas e testes.

### 5.1.1 Faces94

Desenvolvida pelo Dr. Libor Spacek, por meio do grupo de Projetos de Pesquisas em Visão Computacional, da Universidade de Essex, no Reino Unido, essa base engloba imagens de 153 pessoas diferentes, sendo 20 fotos de cada pessoa, em poses/angulações e expressões diversas. As imagens possuem resolução de 180x200 *pixels*.

Dentre as demais informações adicionais relacionadas ao *dataset*, pode-se destacar: fotos com fundo verde; pequenas variações na posição da face; sem variação de luminosidade; mudanças de expressão facial consideráveis. Além disso, as fotografias foram retiradas em uma única sessão, e de modo sequencial (Hond e Spacek, 1997).

Para utilização dessa base de dados nos testes, foi definida uma escolha mais específica de imagens, sendo selecionadas apenas as fotografias femininas. A razão para essa delimitação foi a diminuição da quantidade de instâncias e classes envolvidas, buscando-se um maior assemelhamento entre as características de todas as bases. Essa “sub-base”, definida como *faces94-Woman* inclui 380 itens, sendo 19 mulheres diferentes (19 classes). Vale destacar que foi preservada a diversidade de indivíduos entre raças, estilos de corte de cabelo, além de itens acessórios, como a utilização de óculos.

### 5.1.2 FEI

A base de imagens Thomaz e Giraldi (2010) é provida pelo Laboratório de Inteligência Artificial da Fundação Educacional Inaciana, em São Bernardo do Campo/SP. O material fotográfico foi recolhido entre 2005 e 2006. O *dataset* é formado por fotos de 200 indivíduos (100 homens e 100 mulheres), sendo 14 retratos diferentes para cada integrante. A Figura 17 expõe alguns exemplos dos itens constantes na base.

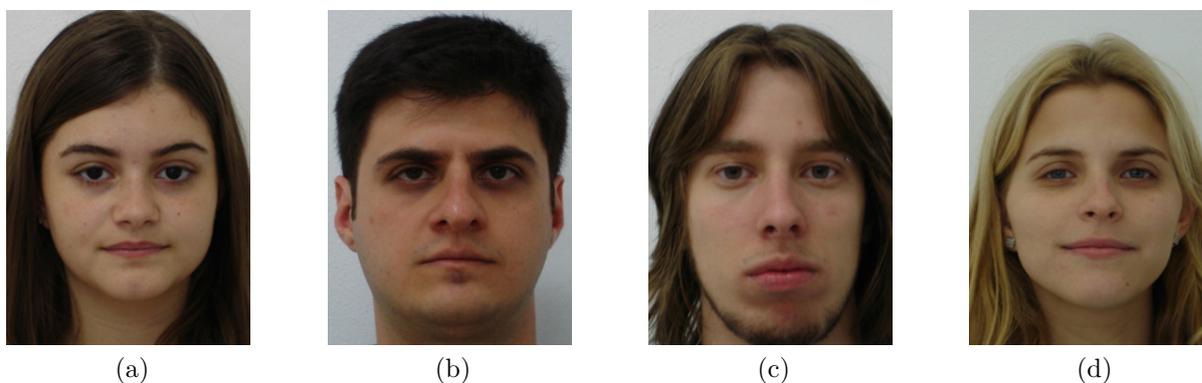


Figura 17 – Exemplos de imagens da base de dados FEI.

Todas as duas mil e oitocentas imagens são coloridas, com tamanho igual a 640x480 *pixels*, e escala variando em 10%. O plano de fundo é homogêneo, da cor branca. A

captura da cena segue o estilo frontal, com variação rotacional da cabeça do indivíduo em  $180^\circ$ . A idade dos envolvidos varia entre 19 e 40 anos, sendo preservadas a variedade de raça, aparência, cabelo e acessórios (Thomaz e Giraldi, 2010).

Visando o uso no processo de testes e validação, definiu-se a utilização apenas das capturas frontais, excluindo as com movimento rotacional. Portanto, são duas fotos para cada um dos 200 indivíduos, variando a expressão facial (“neutra” e “felicidade”), totalizando assim, 400 imagens e 200 classes.

### 5.1.3 JAFFE

A base de dados JAFFE, desenvolvida por Lyons et al. (1998), contém duzentas e treze imagens de dez modelos japonesas. Para cada pessoa foram colhidas diversas fotografias que representam expressões faciais diferentes, definidas como “neutra”, “felicidade”, “tristeza”, “medo”, “raiva”, “surpresa” e “nojo”. As figuras são em escala de cinza, de tamanho padrão, equivalendo a  $256 \times 256$  pixels. A Figura 18 demonstra exemplares das figuras contidas no *dataset* em questão.



Figura 18 – Exemplos de imagens da base de dados JAFFE.

Para realização dos testes experimentais, abordou-se a totalidade da base em questão, representando um quantidade razoável de objetos e classes (213 e 10, respectivamente).

## 5.2 Algoritmos e parâmetros

Para efetuação do escalonamento prévio das imagens, que representa um dos pontos cruciais da técnica proposta, são usados algoritmos de agrupamento. As estratégias abordadas foram *K-Means* e EM, técnicas reconhecidas e amplamente exploradas no meio científico. A seleção dos algoritmos levou em consideração a popularidade e reconhecido desempenho.

O algoritmo *K-Means*, bem como os processos de buscas abordados, usam um critério para medir a similaridade entre os dados. A maior parte deles trabalha com o contexto de distância para quantificar a proximidade (dissimilaridade) entre os pontos do conjunto de dados. Isso é perfeitamente plausível; já os dados são sempre de natureza numérica. O algoritmo EM, que trabalha internamente nas Misturas de Gaussianas Finitas, possui um modo de operação diferenciado na formação dos agrupamento, realizando essa tarefa por estimativas e probabilidades na definição dos parâmetros. Sendo assim, são abordadas duas técnicas de agrupamento com princípios diversificados. Além disso, a definição de algoritmos específicos se torna uma tarefa secundária, visto o pressuposto-base deste trabalho, que objetiva a análise da otimização das buscas e recuperações de imagens imposta pelo agrupamento em si, sendo desnecessária, nessa situação, uma análise inter-técnicas.

Além dos algoritmos definidos, as medidas de proximidade também são indispensáveis na realização da tarefa proposta. Dentre o contexto geral, podemos enumerar várias métricas de distância, como *Mahalanobis*, *Manhattan*, Euclidiana e *Minkowski* (Hair et al., 2005). Entretanto, como forma de tornar o escopo do trabalho limitado a um contexto mais específico, utilizou-se a distância Euclidiana como forma principal de aferição de dissimilaridade entre pontos no conjunto de dados. Isso está de acordo com a maior parte da literatura de aprendizado de máquina (Jain e Dubes, 1988). Além disso, visto a natureza dos dados a serem tratados (características faciais), a distância euclidiana desponta como um bom caminho de estudo. Alternativamente, explorou-se também o conceito das medidas de correlação, representadas no trabalho pela correlação de *Pearson*, para tratar similaridade entre os pontos. Tais medidas possuem a vantagem de encontrar padrões de proximidade sem levar em consideração a magnitude dos valores dos atributos, mas padrões de comportamento entre os exemplos do conjunto de dados (Jain e Dubes, 1988).

### 5.3 Descrição dos experimentos

Para verificação da relevância do método de busca por proximidade de agrupamento, foram realizadas experimentações que englobam ambos os cenários (com e sem a aplicação do método proposto). Também foi garantida a redundância da execução nas três bases (*faces94*, FEI e JAFFE) contendo imagens, tamanhos e números de classes diversificados.

Após a definição dos *datasets*, foram selecionadas dez imagens de cada base para formação de seus respectivos grupos de testes. Esse conjunto serve como parâmetro nas buscas dentro das bases de dados. Para entendimento da apresentação dos resultados, a Tabela 2 descreve as legendas usadas na apresentação dos resultados e as imagens

correspondentes às respectivas bases.

Legenda	<i>faces94</i>	FEI	JAFFE
Img 1	9336923.1	106a	KA.AN1.39
Img 2	9338535.1	11a	KL.AN1.167
Img 3	asewil.1	139a	KM.AN1.17
Img 4	ekavaz.1	13a	KR.AN1.83
Img 5	kaknig.1	147a	MK.AN1.125
Img 6	ksunth.1	185a	NA.AN1.211
Img 7	lfso.1	1a	NM.AN1.104
Img 8	sbains.1	34a	TM.AN1.190
Img 9	vstros.1	43a	UY.AN1.146
Img 10	yfhsie.1	83a	YM.AN1.61

Tabela 2 – Legendas e as imagens correspondentes nas três bases abordadas na pesquisa

A essência do trabalho consiste em realizar experimentos exaustivos nas tarefas de agrupamento de imagens e processamento da busca de similaridade entre os pares. O estudo envolve a aplicação dos algoritmos de agrupamento *K-Means* e EM, ambas estratégias “clássicas” e com boa exploração no meio acadêmico. Além disso, a distância euclidiana e correlação de *Pearson* são utilizadas como medidas de similaridade/dissimilaridade no cálculo das distâncias e, conseqüentemente, retorno do resultado mais adequado ao problema.

É importante ressaltar que o processo de implementação de técnicas e realização de testes foi desenvolvido por meio do *software Matlab*<sup>®</sup>. A máquina utilizada para execução dos códigos e instruções possui a seguinte configuração: processador Intel<sup>®</sup> Core<sup>™</sup>i5-2450M CPU @ 2.50GHz, com 4 GB de memória RAM.

Os experimentos seguem duas formas de organizações: as buscas diretas, e as buscas com auxílio de agrupamentos. Os dados extraídos das imagens, por meio dos descritores HOG, servem de parâmetros para ambas as execuções.

- No primeiro caso, testaram-se buscas diretas, sem nenhum método de arrançamento, visando encontrar a imagem igual ou similar. Esse procedimento envolve a comparação entre os dados da imagem teste e a base-alvo em questão; os itens são verificados sequencialmente, até que o resultado mais adequado seja retornado.
- No segundo caso, que trata da busca por proximidade de agrupamento, são aplicadas técnicas de agrupamento na otimização do retorno de pesquisa das imagens teste, corretas ou similares. Nesse caso, o procedimento inicial é o cálculo da distância da imagem teste e os centróides definidores dos grupos. Isso garante que a precedência

de buscas nos grupos irá seguir essa afinidade, sendo direcionada para o agrupamento mais próximo. Essa tratativa caracteriza a principal forma de “economia” de tempo computacional pois evita que imagens não condizentes sejam checadas sem necessidade (o que acontece na busca “gulosa”).

A totalidade dos testes, de acordo com a técnica utilizada e seus parâmetros, englobam a busca direta baseada em distância euclidiana e correlação de *Pearson*, busca com auxílio de agrupamento *K-Means* baseado em distância euclidiana e correlação de *Pearson*, e a busca com auxílio de agrupamento EM baseado em distância euclidiana e correlação de *Pearson*. Lembrando que todos esses ensaios foram submetidos à utilização das bases completas, onde a imagem teste está presente na mesma, bem como às bases onde a imagem teste inexistente.

Como parte final das experimentações, são feitas análises dos dados obtidos diante de todos os testes. Os tempos computacionais despendidos na realização das tarefas são calculados. O foco principal dessa colheita de tempos é verificar o quão mais ágil as técnicas de busca com auxílio de agrupamentos seria em detrimento das buscas diretas. Em paralelo, são verificados os acertos referentes às imagens retornadas nas buscas. Nesse caso, o objetivo é checar se a acurácia poderia interferir na qualidade do melhoramento advindo do projeto proposto.

## 5.4 Considerações finais

O presente capítulo discorreu acerca dos diversos experimentos desenvolvidos, no que tange ao seu funcionamento e aplicação na ideia proposta. Foram descritas as bases de imagens e os algoritmos, bem como a caracterização dos experimentos.

Esta seção teve como meta proporcionar o entendimento relativo ao funcionamento do método proposto, por meio da descrição de seu passo-a-passo, desde a obtenção dos dados até a análise dos resultados.

No capítulo seguinte são apresentados os resultados obtidos através da realização dos testes, além de suas respectivas avaliações, importantes para validação de preponderância e aplicabilidade das técnicas e metodologias propostas.

## 6 Resultados e Discussões

Este capítulo é dedicado a apresentação dos resultados obtidos utilizando a metodologia descrita no capítulo anterior. Os resultados estarão expostos em duas seções principais:

- busca de imagem existente na base de dados;
- busca de imagem inexistente na base de dados.

Nas duas abordagens as discussões serão realizadas com o objetivo de comparar o método tradicional de busca “gulosa” de uma imagem teste na base de dados (denominado de busca direta) com a abordagem baseada em técnicas de agrupamento utilizada neste trabalho (denominado como busca com auxílio de agrupamento), ambas baseadas em DE (Distância Euclidiana) e CP (Correlação de *Pearson*). A comparação é feita sempre usando dois aspectos fundamentais: tempo levado para a busca ser concluída (utilizando “segundos” como unidade de medida de tempo, em todos os casos), e a acurácia do sistema (capacidade de encontrar corretamente a imagem buscada, em porcentagem relativa). Dessa forma, a Seção 6.1 trata da busca de imagem existente na base de dados, e a Seção 6.2 exhibe os resultados da busca por imagem inexistente na base de dados.

### 6.1 Busca de Imagem Existente na Base de Dados

Como mencionado anteriormente, a análise comparativa entre as abordagens testadas será feita em dois aspectos distintos. O primeiro deles se refere a acurácia do sistema e será mostrado na Seção 6.1.1. O segundo aspecto analisado é o tempo de busca e os resultados podem ser vistos na Seção 6.1.2.

#### 6.1.1 Análise de Acurácia

Visando um melhor entendimento da qualidade dos resultados obtidos nos ensaios propostos, foram calculadas as acurácias ligadas a cada um dos diferentes testes. Esses cálculos se referem a situações de acerto na localização da imagem solicitada, em se tratando da busca por uma imagem presente na base de dados. A Tabela 3 expõe as apurações das porcentagens divididas entre os resultados dos experimentos individuais, bem como a acurácia porcentual geral de cada técnica.

Experimento	Base	Acurácia por experimento	Acurácia por técnica
Busca direta + DE	<i>faces94</i>	100%	100,00%
	FEI	100%	
	JAFFE	100%	
Busca direta + CP	<i>faces94</i>	100%	100,00%
	FEI	100%	
	JAFFE	100%	
K-Means + DE	<i>faces94</i>	100%	100,00%
	FEI	100%	
	JAFFE	100%	
K-Means + CP	<i>faces94</i>	100%	100,00%
	FEI	100%	
	JAFFE	100%	
EM + DE	<i>faces94</i>	100%	100,00%
	FEI	100%	
	JAFFE	100%	
EM + CP	<i>faces94</i>	100%	100,00%
	FEI	100%	
	JAFFE	100%	

Tabela 3 – Acurácia aferida nas experimentações

Baseado nos valores advindos dessa avaliação, pode-se concluir que todos os algoritmos foram capazes de encontrar a imagem buscada. Esse resultado é esperado, uma vez que a imagem teste é exatamente igual a uma das imagens presentes no conjunto de dados.

Esses testes foram conduzidos para garantir que a técnica proposta neste trabalho possui a real capacidade de encontrar uma imagem buscada e levar a análise principal para o contexto do tempo de busca, que será abordado na seção seguinte.

### 6.1.2 Análise do Tempo de Busca

As partes vindouras demonstram os dados referentes ao tempo gasto diante da busca de imagem previamente existente na base de dados. Essa busca se dá através de um processo comparativo de similaridade entre a imagem teste e a base de dados selecionada. Por finalidade, o objetivo desse método é encontrar uma imagem idêntica àquele objeto de análise, considerando a proximidade (através de distância igual a zero, ou correlação equivalente a um) entre a imagem teste e os demais itens formadores da base de dados.

A Tabela 4 mostra o resultado em termos de tempo de execução para as três bases de dados, utilizando o método de busca direta com distância euclidiana (DE) e

correlação de *Pearson* (CP) como medidas de proximidade. A tabela exhibe o tempo levado para encontrar cada imagem, bem como o tempo médio (e o desvio padrão) do método, considerando todas as imagens teste selecionadas.

<i>faces94</i>		
Imagem	Tempo mínimo	
	DE	CP
Img 1	0,92674	0,086495
Img 2	0,002258	0,008674
Img 3	0,0019	0,006769
Img 4	0,001869	0,006908
Img 5	0,002085	0,007263
Img 6	0,002228	0,006329
Img 7	0,0019	0,006677
Img 8	0,001881	0,007046
Img 9	0,001893	0,006567
Img 10	0,001866	0,020902
Tempo Médio	0,094462	0,016363
Desvio Padrão	0,292433	0,025033
FEI		
Imagem	Tempo mínimo	
	DE	CP
Img 1	0,062172	0,067042
Img 2	0,002303	0,008124
Img 3	0,001847	0,00759
Img 4	0,00187	0,006528
Img 5	0,0019	0,006618
Img 6	0,001789	0,0075
Img 7	0,002009	0,007059
Img 8	0,001812	0,006778
Img 9	0,00178	0,007434
Img 10	0,00203	0,007007
Tempo Médio	0,007951	0,013168
Desvio Padrão	0,019052	0,018936
JAFPE		
Imagem	Tempo mínimo	
	DE	CP
Img 1	0,163838	0,26185
Img 2	0,003836	0,013345
Img 3	0,002661	0,017972
Img 4	0,002567	0,013842
Img 5	0,000943	0,003887
Img 6	0,000964	0,003501
Img 7	0,000959	0,003171
Img 8	0,000964	0,003313
Img 9	0,001535	0,00377
Img 10	0,00094	0,003122
Tempo Médio	0,01792	0,032777
Desvio Padrão	0,05128	0,080682

Tabela 4 – Busca direta baseada em distância Euclidiana e correlação de *Pearson* nas bases completas.

Norteados por esses resultados, é possível notar que a busca na base *faces94* possui maior tempo médio de execução quando comparada aos outros dois conjuntos. Para todas as bases, nota-se uma variação entre o melhor e o pior tempo para encontrar a imagem

buscada (representada pelo desvio padrão), o que é normal para esse processo de busca gulosa, já que ela se dá de forma iterativa e a imagem pesquisada pode estar em qualquer lugar da base de dados.

É importante notar que, apesar de atingir bons resultados em algumas situações, a análise sobre esse tipo de busca deve sempre levar em consideração o tempo médio, já que, como mencionado anteriormente, a localização da imagem na base de dados influencia diretamente no processo de busca.

Por outro lado, a pesquisa com auxílio de agrupamentos visa organizar a base de dados de forma que a busca seja direcionada e, conseqüentemente, o tempo levado para recuperar a informação desejada seja reduzido. Esse procedimento caracteriza a ideia central do presente trabalho.

Para uma melhor apresentação, os resultados das buscas serão divididos em duas partes, de forma que o desempenho dos algoritmos utilizados possam ser analisados individualmente. Após isso, uma análise comparativa entre os dois algoritmos é realizada.

#### 6.1.2.1 Desempenho do Algoritmo *k-means*

As Tabelas 5, 6 e 7 mostram o resultados da busca por imagens existentes em bases previamente organizadas com o auxílio da técnica de agrupamento *k-means*, tendo como sustentação a distância euclidiana no cálculo de proximidade. Em cada tabela, são mostrados os tempos de busca para cada uma das imagens de teste (colunas) e a quantidade de grupos submetidos (linhas). As duas últimas colunas indicam o tempo médio e o desvio padrão do processo de busca, ambos considerando todas as imagens. Adicionalmente, foram destacados em cinza os melhores tempos para cada imagem (e o melhor tempo médio)<sup>1</sup>.

Visando validar os resultados obtidos relacionados ao tempo de execução das buscas com auxílio de agrupamento em detrimento da busca “gulosa”, os valores foram submetidos ao teste de *Wilcoxon* pareado, na expectativa de confirmar a hipótese de ganho de tempo relativo. No teste, a hipótese nula supõe que não existe diferença entre os desempenhos das abordagens que usam algoritmos de agrupamento e as que não fazem uso. Nos casos onde a hipótese nula foi rejeitada, os textos das células da tabela encontram-se sublinhados.

A partir dos resultados, é possível notar que os tempos de busca apresentam uma redução considerável, principalmente quando a base de dados se divide em um número de grupos maior que a metade possível. Esse comportamento foi notado em 9 buscas nas bases *faces94* e FEI, e 8 na base JAFFE. Se considerarmos o tempo médio do processo para

<sup>1</sup> Todas as tabelas que exibem resultados de buscas auxiliadas por técnicas de agrupamento seguirão este padrão.

Base faces94												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,059182	0,001107	0,002439	0,005685	0,002436	0,001291	0,002616	0,001461	0,001881	0,001907	0,008001	0,018030
3	0,001145	0,001351	0,000977	0,005711	0,001114	0,001691	0,002055	0,001012	0,001270	0,001064	0,001739	0,001436
4	0,002041	0,001046	0,000673	0,006216	0,001893	0,000989	0,001516	0,000872	0,001362	0,001376	0,001798	0,001612
5	0,001232	0,001259	0,001430	0,005708	0,001321	0,000679	0,001073	0,000886	0,001200	0,001221	0,001601	0,001460
6	0,000829	0,001381	0,001378	0,005849	0,000824	0,001129	0,001703	0,000767	0,001250	0,001372	0,001648	0,001506
7	0,001221	0,001216	0,000798	0,004208	0,000986	0,000977	0,001293	0,000756	0,000859	0,000915	0,001323	0,001031
8	0,000723	0,000932	0,000810	0,004363	0,000663	0,000792	0,001118	0,000870	0,000842	0,000752	0,001187	0,001123
9	0,000847	0,000908	0,001023	0,003299	0,001356	0,000897	0,001343	0,000872	0,000844	0,001009	0,001240	0,000748
10	0,000689	0,001143	0,003053	0,004535	0,000830	0,001118	0,001046	0,000780	0,000987	0,000767	0,001495	0,001271
11	0,000692	0,001030	0,002085	0,002280	0,001170	0,000661	0,000961	0,000682	0,000920	0,001167	0,001165	0,000570
12	0,000730	0,001205	0,002573	0,002592	0,000812	0,000701	0,000766	0,000646	0,000814	0,000648	0,001149	0,000772
13	0,000702	0,001721	0,002103	0,004095	0,000991	0,000961	0,001326	0,000764	0,001324	0,000653	0,001464	0,001035
14	0,000718	0,000934	0,003255	0,003332	0,000693	0,000984	0,000686	0,000667	0,000914	0,000684	0,001287	0,001064
15	0,000698	0,000823	0,003708	0,002124	0,001144	0,000938	0,000773	0,000655	0,000766	0,000652	0,001228	0,000976
16	0,001067	0,000719	0,002144	0,002512	0,000924	0,000658	0,001010	0,000657	0,001007	0,000652	0,001135	0,000655
17	0,001221	0,001032	0,002895	0,002811	0,000990	0,000641	0,001430	0,000680	0,000878	0,000766	0,001334	0,000836
18	0,000729	0,000718	0,002949	0,003744	0,000725	0,000670	0,001402	0,000646	0,000782	0,000822	0,001319	0,001106
19	0,000819	0,000988	0,002751	0,000705	0,000724	0,000674	0,001050	0,000659	0,000876	0,000666	0,000991	0,000634

Tabela 5 – Busca com auxílio de agrupamento usando o *K-Means* e distância euclidiana na base *faces94*.

Base FEI												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,055238	0,005528	0,001479	0,001485	0,001438	0,001927	0,006947	0,006678	0,001800	0,001475	0,008400	0,016616
3	0,001493	0,005093	0,001076	0,001246	0,001193	0,001199	0,004439	0,003131	0,001595	0,001244	0,002171	0,001499
4	0,001452	0,004429	0,001281	0,001207	0,001720	0,001168	0,003444	0,003999	0,001031	0,001530	0,002126	0,001300
5	0,001458	0,003718	0,001091	0,001632	0,001393	0,001149	0,005260	0,004732	0,001037	0,001135	0,002261	0,001646
6	0,000784	0,003869	0,001511	0,001108	0,000752	0,001121	0,003179	0,002794	0,001025	0,001117	0,001726	0,001122
7	0,001181	0,004686	0,000937	0,000981	0,001184	0,001061	0,004941	0,003840	0,000995	0,001052	0,002086	0,001682
8	0,001122	0,005291	0,000952	0,001609	0,000764	0,001057	0,004660	0,002588	0,001100	0,001051	0,002019	0,001646
9	0,001180	0,002996	0,000938	0,000980	0,001264	0,001638	0,003878	0,004153	0,000982	0,001436	0,001945	0,001247
10	0,001219	0,002950	0,000947	0,000986	0,001638	0,001126	0,002295	0,002712	0,000982	0,000903	0,001576	0,000787
11	0,001455	0,002650	0,001410	0,001408	0,000963	0,001025	0,002688	0,003838	0,000967	0,000932	0,001734	0,000988
12	0,001074	0,002677	0,000902	0,001267	0,000880	0,000932	0,003733	0,004019	0,001122	0,001362	0,001797	0,001216
13	0,000707	0,002873	0,000900	0,000907	0,000868	0,000954	0,002659	0,000942	0,001513	0,000928	0,001325	0,000789
14	0,001280	0,002651	0,000894	0,000928	0,001255	0,001441	0,002683	0,000884	0,001445	0,000871	0,001433	0,000688
15	0,002057	0,002663	0,000874	0,000770	0,000857	0,000971	0,002168	0,000916	0,000919	0,000963	0,001316	0,000696
16	0,002686	0,003516	0,000855	0,000823	0,000837	0,000956	0,002486	0,000918	0,001374	0,000980	0,001543	0,000981
17	0,002703	0,003977	0,000896	0,000951	0,000891	0,001344	0,002467	0,001323	0,000954	0,000913	0,001642	0,001057
18	0,003063	0,004530	0,000883	0,001113	0,000870	0,001055	0,002480	0,000934	0,000959	0,000985	0,001687	0,001258
19	0,001944	0,003535	0,000869	0,001342	0,000853	0,001059	0,002765	0,000906	0,000956	0,000889	0,001512	0,000943
20	0,003528	0,000679	0,000897	0,000915	0,000876	0,001537	0,002483	0,001311	0,001057	0,000934	0,001422	0,000903

Tabela 6 – Busca com auxílio de agrupamento usando o *K-Means* e distância euclidiana na base FEI.

Base JAFFE												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,046214	0,000863	0,000967	0,000969	0,000858	0,000868	0,001368	0,000859	0,000966	0,000885	0,005482	0,014313
3	0,000847	0,000871	0,000725	0,000803	0,000830	0,000816	0,000837	0,000880	0,000784	0,000864	0,000826	0,000047
4	0,000699	0,000811	0,000806	0,000668	0,000809	0,000874	0,000867	0,000849	0,000806	0,000804	0,000799	0,000067
5	0,000738	0,000918	0,000878	0,000735	0,000737	0,000730	0,000795	0,000771	0,000628	0,000739	0,000767	0,000082
6	0,000707	0,000693	0,000734	0,000700	0,000750	0,000701	0,000769	0,000746	0,000636	0,001140	0,000758	0,000140
7	0,000785	0,000638	0,000720	0,000703	0,000628	0,000731	0,000746	0,000724	0,000679	0,001092	0,000745	0,000131
8	0,000768	0,000744	0,000707	0,000640	0,000748	0,000741	0,000654	0,000858	0,000761	0,001084	0,000771	0,000126
9	0,000783	0,000625	0,000704	0,000638	0,000675	0,000657	0,000697	0,000673	0,000729	0,000739	0,000692	0,000049
10	0,000770	0,001448	0,000618	0,000640	0,000679	0,001159	0,000649	0,000713	0,000709	0,000676	0,000806	0,000274

Tabela 7 – Busca com auxílio de agrupamento usando o *K-Means* e distância euclidiana na base JAFFE.

todas as imagens, é fácil notar que à medida que o número de grupos cresce, o desempenho da busca melhora. Esse resultado é esperado, visto que com a divisão da base em grupos, cada aglomerado de informações fica mais homogêneo e o direcionamento na busca se torna mais preciso. Como os experimentos foram controlados para que o número de grupos máximo seja igual, de fato, ao número de faces de pessoas diferentes - com única exceção apresentada na base FEI -, é provável que cada grupo contenha em sua maioria imagens de uma única pessoa, garantindo um direcionamento eficiente no processo de busca.

Se compararmos a busca realizada com o auxílio do algoritmo *k-means* usando a distância euclidiana com a busca direta na base de dados, podemos notar que o tempo levado para encontrar a imagem correspondente tem uma grande redução. Esse comportamento é mais notório nas bases *faces94* e JAFFE. Especificamente, na base *faces94*, o tempo foi reduzido em uma escala de quase 100 vezes do tempo original, passando de 0,094462 para 0,000991 segundos.

A redução no tempo de busca também pode ser vista quando usamos o algoritmo *k-means* com a correlação de Pearson como medida de similaridade. As Tabelas 8, 9 e 10 trazem os valores do tempo de execução com a dita configuração.

Com o uso da correlação, os tempos de busca também são reduzidos; entretanto, em uma escala um pouco menor, quando comparados aos resultados usando a distância euclidiana. O tempo médio de busca com maior diferença entre as duas abordagens foi encontrado na base JAFFE, onde a essa diferença é quase 10 vezes menor que na busca direta (mudando de 0,032777 para 0,003928).

Outro comportamento similar ao *k-means*, baseado em distância euclidiana, refere-se a relação existente entre o número de grupos usados no escalonamento da base de dados e o ganho em desempenho advindo. Com o uso de correlação, notamos que a execução com número de agrupamentos maiores que a metade do total possível tende a entregar

Base faces94												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,061244	0,006373	0,002542	0,006812	0,006892	0,002882	0,006321	0,007274	0,023620	0,006025	0,012999	0,017943
3	0,002984	0,002615	0,002996	0,003519	0,003191	0,002686	0,003235	0,008172	0,010526	0,002845	0,004277	0,002743
4	0,003155	0,002417	0,003604	0,003554	0,002643	0,001640	0,003783	0,007364	0,008368	0,003044	0,003957	0,002169
5	0,002329	0,005028	0,001674	0,004172	0,002109	0,001661	0,004774	0,012598	0,012675	0,004554	0,005157	0,004150
6	0,002674	0,004245	0,001636	0,003668	0,002194	0,001678	0,004537	0,005616	0,011960	0,002271	0,004048	0,003080
7	0,003234	0,002396	0,001911	0,003446	0,002099	0,001996	0,003069	0,008639	0,007520	0,002787	0,003710	0,002377
8	0,001817	0,003078	0,001645	0,002400	0,002254	0,001724	0,001859	0,008763	0,008027	0,002611	0,003418	0,002666
9	0,002004	0,003261	0,001711	0,003209	0,001661	0,001663	0,003381	0,005887	0,010653	0,003128	0,003656	0,002771
10	0,004575	0,002778	0,001716	0,002970	0,002117	0,001590	0,002906	0,006846	0,009377	0,001978	0,003685	0,002555
11	0,001778	0,002951	0,002631	0,002608	0,002304	0,001780	0,002347	0,005810	0,006166	0,001788	0,003016	0,001618
12	0,002799	0,002434	0,001935	0,002356	0,002593	0,001784	0,002440	0,006024	0,006665	0,001670	0,003070	0,001769
13	0,001817	0,002307	0,001893	0,001766	0,003236	0,001757	0,003149	0,005506	0,005818	0,003414	0,003066	0,001513
14	0,001829	0,002035	0,001749	0,001797	0,002290	0,001667	0,002694	0,005802	0,005930	0,002018	0,002781	0,001654
15	0,001823	0,002524	0,001770	0,002998	0,002589	0,001786	0,002559	0,006212	0,006036	0,002335	0,003063	0,001663
16	0,001792	0,002274	0,001658	0,003041	0,001788	0,001783	0,002523	0,005862	0,006294	0,002428	0,002944	0,001710
17	0,001875	0,001806	0,001834	0,002996	0,002171	0,001710	0,002618	0,006700	0,002679	0,002076	0,002647	0,001489
18	0,001907	0,001744	0,002173	0,003180	0,002067	0,001749	0,002104	0,006901	0,002524	0,002361	0,002671	0,001545
19	0,001821	0,002429	0,002977	0,001767	0,001793	0,001822	0,001854	0,006975	0,001933	0,002459	0,002583	0,001595

Tabela 8 – Busca com auxílio de agrupamento usando o *K-Means* e correlação de *Pearson* na base *faces94*.

Base FEI												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,060927	0,013125	0,005070	0,004150	0,004541	0,004789	0,012311	0,016636	0,004548	0,005221	0,013132	0,017393
3	0,008362	0,010786	0,002855	0,003420	0,003484	0,003416	0,009710	0,011833	0,002635	0,003603	0,006010	0,003695
4	0,008209	0,008662	0,003110	0,003109	0,003143	0,004004	0,008809	0,009234	0,003917	0,002916	0,005511	0,002801
5	0,008633	0,008445	0,003667	0,003393	0,003049	0,003962	0,012143	0,002777	0,003251	0,003090	0,005241	0,003273
6	0,008531	0,007678	0,003430	0,003019	0,002537	0,003416	0,009380	0,002204	0,002688	0,002742	0,004563	0,002792
7	0,007755	0,003742	0,002623	0,002727	0,002170	0,002822	0,008291	0,002126	0,003309	0,002769	0,003833	0,002262
8	0,006100	0,003518	0,002116	0,002645	0,002164	0,002434	0,007904	0,002041	0,002496	0,002581	0,003400	0,001989
9	0,006017	0,002561	0,002381	0,002569	0,002103	0,002576	0,007624	0,002230	0,002539	0,003170	0,003377	0,001875
10	0,006125	0,002583	0,002035	0,002684	0,001888	0,003223	0,008300	0,002859	0,002495	0,002435	0,003463	0,002076
11	0,005688	0,002283	0,002182	0,002297	0,002153	0,002411	0,007649	0,002064	0,002243	0,003263	0,003223	0,001903
12	0,005574	0,002472	0,002413	0,002589	0,002023	0,002976	0,006359	0,002459	0,002503	0,002725	0,003209	0,001484
13	0,025307	0,002487	0,002154	0,002528	0,002157	0,002555	0,006322	0,001959	0,002595	0,002765	0,005083	0,007216
14	0,006292	0,002589	0,003189	0,002466	0,001991	0,007201	0,006064	0,002151	0,002931	0,002489	0,003736	0,001971
15	0,005436	0,002328	0,002348	0,002428	0,002010	0,007378	0,008018	0,002198	0,002311	0,002199	0,003665	0,002352
16	0,006103	0,002222	0,002109	0,002088	0,002149	0,006238	0,006913	0,002752	0,002111	0,002130	0,003482	0,002046
17	0,005608	0,001951	0,002285	0,002164	0,002111	0,006695	0,007069	0,002680	0,002262	0,002268	0,003509	0,002074
18	0,006095	0,002695	0,002602	0,002151	0,001913	0,006594	0,006844	0,002866	0,002609	0,002169	0,003654	0,002000
19	0,006177	0,002150	0,009883	0,002981	0,002117	0,006771	0,006876	0,002202	0,002464	0,002225	0,004385	0,002800
20	0,005674	0,002407	0,002549	0,002274	0,002222	0,006696	0,006142	0,002040	0,002739	0,002114	0,003486	0,001879

Tabela 9 – Busca com auxílio de agrupamento usando o *K-Means* e correlação de *Pearson* na base FEI.

Base JAFFE												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,073123	0,003751	0,003416	0,004515	0,003875	0,003851	0,010448	0,010924	0,011604	0,003954	0,012946	0,021411
3	0,003447	0,002911	0,003868	0,003465	0,002793	0,002925	0,007606	0,010234	0,012318	0,003643	0,005321	0,003466
4	0,002296	0,002980	0,002645	0,004114	0,003747	0,002974	0,008881	0,011600	0,013229	0,003867	0,005633	0,004042
5	0,002923	0,003003	0,002498	0,003404	0,003066	0,003058	0,010425	0,010213	0,012821	0,003317	0,005473	0,003986
6	0,002182	0,002347	0,002402	0,002866	0,002742	0,002485	0,007071	0,008956	0,011065	0,003101	0,004522	0,003262
7	0,003252	0,002021	0,002615	0,003690	0,003402	0,003396	0,007502	0,009803	0,008569	0,002168	0,004642	0,002853
8	0,002267	0,002168	0,002160	0,002766	0,002274	0,001975	0,007981	0,007365	0,027273	0,002283	0,005851	0,007859
9	0,002153	0,002111	0,002189	0,002131	0,002496	0,002251	0,006440	0,010646	0,006882	0,002537	0,003984	0,002975
10	0,002973	0,002005	0,002676	0,002097	0,002230	0,002010	0,006432	0,008433	0,008181	0,002240	0,003928	0,002658

Tabela 10 – Busca com auxílio de agrupamento usando o *K-Means* e correlação de Pearson na base JAFFE.

resultados mais satisfatórios.

#### 6.1.2.2 Desempenho do Algoritmo EM

A seguir são exibidos os resultados do processo de busca com auxílio do algoritmo de agrupamento EM. Seguindo a mesma metodologia da seção anterior, são inicialmente mostrados os resultados das buscas nas bases de dados usando a distância euclidiana como medida de dissimilaridade e, logo depois, os resultados da busca com a correlação como métrica de similaridade.

É importante salientar que a distância euclidiana e a correlação de *Pearson* são usadas, nesse caso, na aferição da proximidade entre imagem teste e as imagens presentes nos grupos. Durante o processo de construção do agrupamento, o algoritmo EM possui sua própria medida probabilística para quantificar e agrupar os dados.

Sendo assim, as Tabelas 11, 12 e 13 contém os resultados para a busca usando o algoritmo EM e a distância euclidiana. As Tabelas 14, 15 e 16, por sua vez, exibem os resultados do mesmo algoritmo, mas usando a correlação de *Pearson*.

O algoritmo EM, como pode ser visto nas tabelas, também gerou um ganho em desempenho no tempo de busca. O seu comportamento, quando usada a distância euclidiana, foi muito semelhante ao algoritmo *k-means* usando a correlação de *Pearson*. No entanto, um ponto importante a se ressaltar é que o algoritmo obteve seu melhor desempenho médio quando o número de grupos que divide a base de dados é mediano. Ou seja, os resultados mais otimizados ocorreram utilizando aproximadamente a metade da quantidade de grupos possíveis. O EM dividiu as bases *faces94*, FEI e JAFFE em 3, 5 e 4 grupos, respectivamente, para alcançar o menor tempo médio de execução.

Quando utilizamos a correlação para medir similaridade entre a imagem teste e a base de dados, organizada com o algoritmo EM, notamos o mesmo comportamento daquele

Base <i>faces94</i>												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,074676	0,003130	0,003777	0,002758	0,010136	0,003107	0,003165	0,002056	0,002396	0,002647	0,010785	0,022569
3	0,002722	0,002084	0,003288	0,001391	0,005871	0,001789	0,002108	0,002819	0,004104	0,001619	0,002780	0,001365
4	0,002403	0,002395	0,003162	0,001067	0,011305	0,003243	0,002987	0,003771	0,004185	0,003407	0,003793	0,002777
5	0,002834	0,002583	0,005401	0,002120	0,008040	0,002554	0,001047	0,002209	0,001431	0,002686	0,003091	0,002089
6	0,000990	0,001487	0,001058	0,000979	0,013137	0,001543	0,002940	0,004611	0,001485	0,004994	0,003322	0,003755
7	0,004247	0,001438	0,002996	0,000959	0,029795	0,006017	0,000957	0,003205	0,005194	0,003231	0,005804	0,008600
8	0,002346	0,003065	0,004203	0,003831	0,007098	0,001825	0,000938	0,003361	0,002679	0,003471	0,003282	0,001655
9	0,002180	0,003873	0,010260	0,002118	0,012245	0,001576	0,002225	0,000966	0,000938	0,003048	0,003943	0,003979
10	0,001875	0,002982	0,001286	0,001358	0,021210	0,007431	0,006194	0,002906	0,003992	0,007199	0,005643	0,005936
11	0,003551	0,004068	0,004675	0,001360	0,015447	0,001049	0,003544	0,003898	0,001357	0,000958	0,003991	0,004267
12	0,004224	0,004581	0,003326	0,002526	0,007791	0,002915	0,001901	0,003854	0,005980	0,005369	0,004247	0,001769
13	0,001689	0,002072	0,004532	0,003875	0,002645	0,002511	0,005377	0,003446	0,004970	0,002980	0,003410	0,001252
14	0,001303	0,002826	0,002524	0,005703	0,024281	0,001336	0,005322	0,000920	0,003946	0,003014	0,005118	0,006927
15	0,012004	0,005347	0,002237	0,004735	0,013811	0,000920	0,004411	0,003416	0,006700	0,004782	0,005836	0,004082
16	0,000934	0,001371	0,007712	0,004024	0,019006	0,001725	0,004653	0,006098	0,008889	0,005534	0,005995	0,005289
17	0,006938	0,002091	0,004025	0,003691	0,008282	0,002508	0,001295	0,006135	0,001620	0,006331	0,004292	0,002473
18	0,008332	0,001643	0,005337	0,019177	0,022570	0,001705	0,001344	0,005095	0,001984	0,005014	0,007220	0,007573
19	0,005422	0,005774	0,006516	0,005829	0,024916	0,011333	0,005940	0,006380	0,001292	0,003042	0,007644	0,006591

Tabela 11 – Busca com auxílio de agrupamento usando o EM e distância euclidiana na base *faces94*.

Base FEI												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,059497	0,003332	0,001983	0,008500	0,008087	0,004168	0,002196	0,001530	0,003611	0,009173	0,010208	0,017551
3	0,004101	0,003945	0,001439	0,006638	0,015801	0,001274	0,002124	0,001312	0,004265	0,009274	0,005017	0,004577
4	0,001928	0,002073	0,004585	0,012620	0,009363	0,002509	0,005059	0,001210	0,003204	0,003447	0,004600	0,003650
5	0,001330	0,004539	0,004329	0,005283	0,005384	0,003947	0,004345	0,001737	0,005400	0,003205	0,003950	0,001449
6	0,004344	0,002800	0,001073	0,004893	0,008880	0,005284	0,002887	0,002510	0,003767	0,005298	0,004174	0,002142
7	0,001756	0,005577	0,002391	0,012065	0,016256	0,004175	0,002118	0,001051	0,005531	0,003465	0,005439	0,004939
8	0,003999	0,004790	0,002757	0,010986	0,020969	0,004546	0,002582	0,006320	0,004297	0,002886	0,006413	0,005674
9	0,005043	0,002221	0,002127	0,006666	0,016635	0,004534	0,002572	0,003467	0,002772	0,009864	0,005590	0,004567
10	0,009019	0,001443	0,005082	0,052846	0,005894	0,004599	0,001679	0,001011	0,017840	0,012092	0,011151	0,015574
11	0,009091	0,003184	0,004213	0,029251	0,006396	0,008204	0,002888	0,003175	0,019150	0,050878	0,013643	0,015588
12	0,001866	0,003659	0,001060	0,022264	0,012116	0,006228	0,005782	0,005874	0,005462	0,044872	0,010918	0,013410
13	0,001582	0,005102	0,006433	0,021333	0,008257	0,005071	0,001721	0,002624	0,005240	0,044686	0,010205	0,013382
14	0,006912	0,006409	0,006554	0,004613	0,024550	0,004417	0,001856	0,006321	0,011868	0,022876	0,009638	0,007846
15	0,003685	0,004512	0,000971	0,018769	0,016490	0,001039	0,001987	0,002071	0,004029	0,007123	0,006068	0,006388
16	0,004433	0,002288	0,005615	0,008805	0,044811	0,006467	0,002884	0,003339	0,042765	0,008361	0,012977	0,016391
17	0,002691	0,004435	0,010660	0,008535	0,023226	0,003597	0,001449	0,007264	0,012373	0,008577	0,008281	0,006339
18	0,005270	0,007080	0,002491	0,009645	0,000997	0,005028	0,003877	0,003781	0,013324	0,008308	0,005980	0,003677
19	0,001387	0,003829	0,001774	0,002692	0,001709	0,005684	0,006002	0,010496	0,019153	0,012662	0,006539	0,005850
20	0,002649	0,002480	0,005630	0,011877	0,004077	0,003677	0,001958	0,006130	0,029837	0,022383	0,009070	0,009581

Tabela 12 – Busca com auxílio de agrupamento usando o EM e distância euclidiana na base FEI.

Base JAFFE												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,209299	0,001695	0,002078	0,001181	0,001751	0,000971	0,000950	0,001714	0,001523	0,002283	0,022345	0,065691
3	0,005328	0,002221	0,001365	0,002491	0,000890	0,001251	0,001747	0,001089	0,001222	0,002773	0,002038	0,001318
4	0,002517	0,004253	0,001601	0,001143	0,002283	0,000749	0,001183	0,001106	0,001545	0,001358	0,001774	0,001025
5	0,009507	0,001984	0,001191	0,003184	0,002093	0,001776	0,000735	0,002041	0,002278	0,002315	0,002710	0,002477
6	0,006719	0,003137	0,001620	0,002092	0,001089	0,000732	0,001936	0,001028	0,002491	0,001799	0,002264	0,001722
7	0,004602	0,002168	0,001026	0,001963	0,002536	0,001353	0,001442	0,000716	0,001156	0,001053	0,001802	0,001137
8	0,004443	0,003766	0,002615	0,001983	0,002204	0,002293	0,001618	0,001756	0,001112	0,001938	0,002373	0,001011
9	0,012013	0,003019	0,001697	0,002626	0,002444	0,001808	0,002055	0,002742	0,002953	0,001231	0,003259	0,003131
10	0,015712	0,003673	0,001551	0,002083	0,001973	0,002675	0,002347	0,003049	0,003857	0,001470	0,003839	0,004251

Tabela 13 – Busca com auxílio de agrupamento usando o EM e distância euclidiana na base JAFFE.

demonstrado quando é abordada a distância euclidiana. Nessa caso, apesar de semelhante, a diferença entre o tempo levado para executar a busca direta e o tempo despemido na busca com auxílio de agrupamento é menor. No melhor cenário, o algoritmo EM executou a busca com um tempo quase quatro vezes menor que a busca direta (na base JAFFE, a duração foi reduzida de 0,032777 para 0,008632).

Base faces94												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,112659	0,023613	0,003567	0,007133	0,008054	0,008208	0,007062	0,016147	0,004467	0,006879	0,019779	0,033185
3	0,006772	0,012658	0,005928	0,005906	0,010296	0,009196	0,006618	0,013896	0,003637	0,009291	0,008420	0,003235
4	0,006500	0,015630	0,009008	0,004250	0,008505	0,006855	0,003701	0,035700	0,009241	0,009043	0,010843	0,009337
5	0,005407	0,013911	0,010990	0,005072	0,006616	0,008561	0,008654	0,024820	0,007676	0,008610	0,010032	0,005811
6	0,014233	0,039279	0,012441	0,008376	0,013177	0,006745	0,033559	0,024156	0,010299	0,009233	0,017150	0,011301
7	0,007868	0,041561	0,005290	0,005520	0,004978	0,012828	0,018280	0,051970	0,004796	0,012393	0,016548	0,016700
8	0,014694	0,047060	0,013556	0,009932	0,006028	0,014021	0,008109	0,009896	0,013684	0,010168	0,014715	0,011714
9	0,012638	0,036425	0,002632	0,003101	0,019033	0,011253	0,012651	0,010561	0,009703	0,012544	0,013054	0,009492
10	0,004244	0,079209	0,006393	0,015409	0,010953	0,009326	0,012672	0,048762	0,004234	0,002402	0,019360	0,024925
11	0,013583	0,056187	0,007912	0,008973	0,012631	0,004768	0,023741	0,022026	0,006534	0,002636	0,015899	0,015767
12	0,008508	0,032123	0,012225	0,009536	0,018107	0,019400	0,007773	0,008325	0,009202	0,008338	0,013354	0,007814
13	0,016839	0,066482	0,007354	0,016463	0,012262	0,014025	0,050043	0,016505	0,002347	0,019074	0,022139	0,020056
14	0,018936	0,044960	0,018611	0,009622	0,009858	0,019460	0,022098	0,023729	0,015135	0,003593	0,018600	0,011198
15	0,004828	0,065295	0,003596	0,003548	0,014417	0,007284	0,034115	0,017569	0,002245	0,023621	0,017652	0,019700
16	0,022054	0,064443	0,015303	0,004625	0,008456	0,003804	0,014063	0,010465	0,002235	0,025471	0,017092	0,018320
17	0,003110	0,065717	0,004545	0,003244	0,013843	0,007519	0,032213	0,022735	0,019869	0,003167	0,017596	0,019638
18	0,053890	0,056515	0,016126	0,006118	0,006202	0,021770	0,051833	0,019872	0,006265	0,009128	0,024772	0,021017
19	0,020827	0,018022	0,021810	0,020783	0,015622	0,028877	0,084292	0,018570	0,019321	0,006781	0,025491	0,021385

Tabela 14 – Busca com auxílio de agrupamento usando o EM e correlação de *Pearson* na base *faces94*.

### 6.1.2.3 *k-means* vs EM

Quando comparamos a abordagem da busca com ou sem o auxílio de técnicas de agrupamento, é possível notar que organizar a base em grupos, promove, em todos os casos, um ganho em termos de tempo de execução. Ou seja, tanto o algoritmo *k-means* quanto o algoritmos EM agregaram benefícios, deixando as buscas mais rápidas. Se

Base FEI												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,067240	0,003381	0,003822	0,003932	0,006608	0,003267	0,019643	0,004758	0,004191	0,003592	0,012043	0,020015
3	0,003351	0,002884	0,003680	0,008241	0,006260	0,003450	0,025428	0,003668	0,009153	0,007848	0,007396	0,006753
4	0,009063	0,003448	0,002628	0,007953	0,007501	0,008287	0,015690	0,009075	0,005701	0,004256	0,007360	0,003750
5	0,009154	0,011292	0,010585	0,003841	0,005477	0,008874	0,033438	0,011192	0,007402	0,009953	0,011121	0,008215
6	0,019876	0,011654	0,010391	0,010213	0,011586	0,009984	0,024204	0,006875	0,013853	0,002955	0,012159	0,006077
7	0,018621	0,012438	0,011571	0,008429	0,005316	0,006242	0,020716	0,010055	0,014203	0,010591	0,011818	0,004956
8	0,041457	0,010335	0,004996	0,009216	0,006128	0,005294	0,021366	0,008535	0,011672	0,005167	0,012417	0,011311
9	0,015754	0,008261	0,002468	0,013675	0,006605	0,009369	0,008914	0,006970	0,012914	0,002828	0,008776	0,004392
10	0,035088	0,002430	0,013474	0,009403	0,004869	0,007632	0,024419	0,013068	0,015262	0,015095	0,014074	0,009638
11	0,012921	0,013389	0,012987	0,005125	0,007344	0,035069	0,056444	0,016534	0,006639	0,007270	0,017372	0,016204
12	0,055744	0,004488	0,015500	0,005566	0,008388	0,026668	0,056260	0,012284	0,020866	0,005151	0,021092	0,019748
13	0,007172	0,011977	0,012747	0,007096	0,013229	0,014259	0,076648	0,022572	0,014918	0,014933	0,019555	0,020526
14	0,031577	0,012429	0,015508	0,013995	0,004425	0,010358	0,058957	0,008735	0,017181	0,063183	0,023635	0,020998
15	0,014052	0,004695	0,016158	0,012195	0,013965	0,041760	0,093411	0,017463	0,020635	0,042385	0,027672	0,026170
16	0,010247	0,025276	0,008038	0,011755	0,010216	0,060858	0,033261	0,016711	0,023084	0,012537	0,021198	0,016113
17	0,049364	0,013956	0,017795	0,006176	0,014105	0,059886	0,004274	0,007269	0,013580	0,028088	0,021449	0,018901
18	0,007173	0,009169	0,007397	0,012834	0,018118	0,055983	0,007970	0,004994	0,014942	0,016471	0,015505	0,014893
19	0,048659	0,011294	0,003522	0,003304	0,018375	0,093207	0,015871	0,026589	0,014229	0,031911	0,026696	0,027073
20	0,053009	0,021681	0,013927	0,012974	0,013718	0,038994	0,035397	0,004786	0,023620	0,038819	0,025693	0,015241

Tabela 15 – Busca com auxílio de agrupamento usando o EM e correlação de *Pearson* na base FEI.

Base JAFFE												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,073781	0,005816	0,006081	0,005282	0,003275	0,005556	0,007481	0,015445	0,012154	0,019636	0,015451	0,021146
3	0,004542	0,002627	0,005481	0,005315	0,007876	0,002720	0,002931	0,022781	0,021114	0,010937	0,008632	0,007476
4	0,002467	0,005466	0,003881	0,003773	0,005305	0,003859	0,004620	0,026246	0,028961	0,017532	0,010211	0,010121
5	0,004747	0,005915	0,006003	0,008680	0,005969	0,004054	0,010799	0,015685	0,008499	0,017310	0,008766	0,004559
6	0,007770	0,010445	0,005839	0,007888	0,005776	0,009439	0,003882	0,025498	0,007855	0,007274	0,009167	0,006035
7	0,004811	0,006381	0,003492	0,008501	0,002182	0,006905	0,006011	0,018124	0,014610	0,020134	0,009115	0,006266
8	0,005976	0,009816	0,010495	0,003278	0,013098	0,011588	0,002672	0,024649	0,018023	0,033908	0,013350	0,009808
9	0,009043	0,003144	0,007263	0,002369	0,004358	0,010117	0,005561	0,054602	0,040882	0,027028	0,016437	0,018208
10	0,011492	0,006946	0,003223	0,010164	0,005362	0,004816	0,005455	0,008618	0,053246	0,046252	0,015557	0,018270

Tabela 16 – Busca com auxílio de agrupamento usando o EM e correlação de *Pearson* na base JAFFE.

compararmos os dois algoritmos de agrupamento, podemos perceber que o *k-means* obteve melhor desempenho para as duas medidas de similaridade. Uma síntese dos resultados pode ser vista na Tabela 17, além da disposição visual dos dados, presente na Figura 19.

Tendo como base os valores apresentados, principalmente pela validação do teste de *Wilcoxon* pareado, é possível perceber que a abordagem que utiliza o *k-means*, em geral, obteve uma diferença significativa do ponto de vista estatístico quando comparada ao método tradicional de busca “gulosa”. Nesse caso, quando a distância euclidiana é setada, apenas o resultado da base FEI não apresenta ganho comprovado; já no caso da correlação de *Pearson*, o ganho foi observado em todas as bases. A técnica EM mostrou-se estatisticamente insatisfatória, pois a otimização do tempo não foi comprovada em nenhum

Bases de dados	<i>faces94</i>		FEI		JAFFE	
	TM	DP	TM	DP	TM	DP
Busca direta	DE	0,094462 ± 0,292433	0,007951 ± 0,019052	0,017920 ± 0,051280	0,017920 ± 0,051280	0,032777 ± 0,080682
	CP	0,016363 ± 0,025033	0,013168 ± 0,018936	0,032777 ± 0,080682	0,032777 ± 0,080682	0,032777 ± 0,080682
<i>k-means</i>	DE	<u>0,000991 ± 0,000634</u>	0,001316 ± 0,000696	<u>0,000692 ± 0,000049</u>	<u>0,000692 ± 0,000049</u>	0,000692 ± 0,000049
	CP	<u>0,002583 ± 0,001595</u>	<u>0,003209 ± 0,001484</u>	<u>0,003928 ± 0,002658</u>	<u>0,003928 ± 0,002658</u>	0,003928 ± 0,002658
EM	DE	0,002780 ± 0,001365	0,003950 ± 0,001449	0,001774 ± 0,001025	0,001774 ± 0,001025	0,001774 ± 0,001025
	CP	0,008420 ± 0,003235	0,007360 ± 0,003750	0,008632 ± 0,007476	0,008632 ± 0,007476	0,008632 ± 0,007476

Tabela 17 – Resultados sintetizados para os métodos de busca.

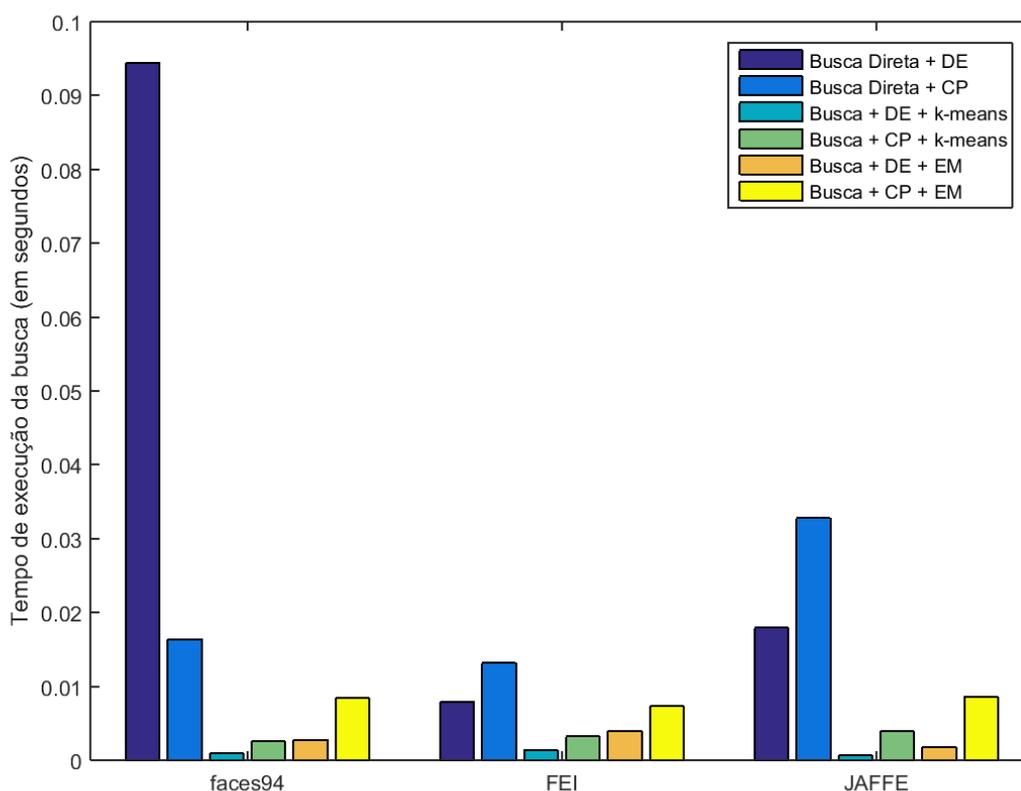


Figura 19 – Gráfico relacionado a totalidade dos resultados da buscas.

caso.

Pode-se observar, de acordo com os dados dos ganhos percentuais de tempo médio de execução, dispostos na Tabela 18, que abordam a busca com auxílio de agrupamento em detrimento da busca direta, um ganho razoável de desempenho. Nos melhores casos, esse incremento variou de 4,24% a 98,95%. Podemos verificar adicionalmente que o uso do *k-means* imprimiu a melhor taxa de ganho médio, com 98,95% na base *faces94*, 7,02% na base FEI, e 18,24% na JAFFE.

Outro ponto a ser notado nos resultados é que, em comparação com o *k-means*, o algoritmo EM sempre entregou melhores tempos com um número de grupos menor. O *k-means* é uma das técnicas de agrupamento que possui a tendência de dividir a base em grupos balanceados devido ao seu critério de agrupamento, que visa minimizar o erro quadrático médio (distância entre os pontos e os centros dos grupos aos quais pertencem) e cria conjuntos com tamanho e densidade semelhantes Jain e Dubes (1988). Devido a especificidade da presente proposta de otimização, a imagem teste é primeiramente comparada aos centros dos grupos (para determinar em qual deles a busca deve ser iniciada). Sendo assim, um número maior de grupos faria com que os centros se tornassem cada vez mais semelhantes à imagem teste.

<b>Base faces94</b>				
<b>k</b>	<b>K-Means + DE</b>	<b>K-Means + CP</b>	<b>EM + DE</b>	<b>EM + CP</b>
2	91,53%	3,56%	88,58%	-3,62%
3	98,16%	12,79%	97,06%	8,41%
4	98,10%	13,13%	95,99%	5,84%
5	98,31%	11,86%	96,73%	6,70%
6	98,26%	13,04%	96,48%	-0,83%
7	98,60%	13,40%	93,86%	-0,20%
8	98,74%	13,70%	96,53%	1,74%
9	98,69%	13,45%	95,83%	3,50%
10	98,42%	13,42%	94,03%	-3,17%
11	98,77%	14,13%	95,78%	0,49%
12	98,78%	14,07%	95,50%	3,19%
13	98,45%	14,08%	96,39%	-6,12%
14	98,64%	14,38%	94,58%	-2,37%
15	98,70%	14,08%	93,82%	-1,36%
16	98,80%	14,21%	93,65%	-0,77%
17	98,59%	14,52%	95,46%	-1,31%
18	98,60%	14,49%	92,36%	-8,90%
19	98,95%	14,59%	91,91%	-9,66%

<b>Base FEI</b>				
<b>k</b>	<b>K-Means + DE</b>	<b>K-Means + CP</b>	<b>EM + DE</b>	<b>EM + CP</b>
2	-0,47%	0,04%	-2,39%	1,19%
3	6,12%	7,58%	3,11%	6,11%
4	6,17%	8,11%	3,55%	6,15%
5	6,02%	8,39%	4,24%	2,17%
6	6,59%	9,11%	4,00%	1,07%
7	6,21%	9,88%	2,66%	1,43%
8	6,28%	10,34%	1,63%	0,80%
9	6,36%	10,37%	2,50%	4,65%
10	6,75%	10,27%	-3,39%	-0,96%
11	6,58%	10,53%	-6,03%	-4,45%
12	6,52%	10,54%	-3,14%	-8,39%
13	7,01%	8,56%	-2,39%	-6,76%
14	6,90%	9,98%	-1,79%	-11,08%
15	7,02%	10,06%	1,99%	-15,35%
16	6,78%	10,25%	-5,32%	-8,50%
17	6,68%	10,22%	-0,35%	-8,77%
18	6,63%	10,07%	2,09%	-2,47%
19	6,82%	9,30%	1,49%	-14,32%
20	6,91%	10,25%	-1,18%	-13,26%

<b>Base JAFFE</b>				
<b>k</b>	<b>K-Means + DE</b>	<b>K-Means + CP</b>	<b>EM + DE</b>	<b>EM + CP</b>
2	13,17%	20,99%	-4,68%	18,34%
3	18,10%	29,07%	16,81%	25,56%
4	18,12%	28,74%	17,09%	23,89%
5	18,16%	28,90%	16,10%	25,42%
6	18,17%	29,91%	16,57%	24,99%
7	18,18%	29,78%	17,06%	25,05%
8	18,15%	28,50%	16,46%	20,57%
9	18,24%	30,48%	15,52%	17,30%
10	18,12%	30,54%	14,91%	18,23%

Tabela 18 – Ganho percentual de tempo de execução em detrimento da busca “gulosa”.

Nos experimentos realizados neste trabalho, o número de grupos foi controlado pela quantidade de pessoas diferentes presentes na base de dados, com exceção da base FEI. Obviamente que em um cenário onde as bases de dados são bem maiores e não se sabe

ao certo o número de pessoas existentes, controlar o parâmetro referente ao número de grupos torna-se uma tarefa não trivial, exigindo uma análise mais profunda. Essa situação é simulada nos testes com a base de dados FEI, que originalmente possui 200 classes, com 2 imagens cada. Por outro lado, podemos ver que a simples divisão em grupos, em geral, leva a uma melhora de desempenho na busca.

## 6.2 Busca de Imagens Inexistentes na Base de Dados

Esta seção traz os resultados do processo de busca de uma imagem que não está inserida na base de dados. Da mesma forma que a seção anterior, os resultados estão divididos de acordo com o tipo de análise a ser investigada. A Seção 6.2.1 trata sobre a acurácia dos métodos quanto ao sucesso na localização das imagens corretas, e a Seção 6.2.2 aborda o tempo de busca para os algoritmos implementados.

### 6.2.1 Análise de Acurácia

Como procedido na seção anterior, apresentamos os dados relacionados ao acerto percentual na identificação da imagem mais adequada na base de dados alvo. As aferições são dispostas na Tabela 19, divididas entre os resultados dos experimentos individuais e a acurácia percentual geral de cada técnica.

Como podemos inferir nos resultados, o desempenho da busca com o auxílio de agrupamento de uma imagem desconhecida (não presente no conjunto de dados) apresentou resultados diferentes quando comparada à busca de imagens presentes no conjunto de dados. A queda na acurácia dos métodos é mais evidente nas bases FEI e JAFFE.

A redução no desempenho é mais notória quando se usa a correlação para medir a proximidade entre as imagens. Por outro lado, quando a distância euclidiana é empregada, o desempenho da busca usando agrupamento é melhor que a busca direta. Esse crescimento pode ser visto na base JAFFE, onde a acurácia passou de 90% com a busca direta, para 100% com o uso de algoritmos de agrupamento.

Dois dos motivos possíveis para a alteração de desempenho entre as abordagens, tanto para aqueles que pioraram quanto para os que melhoraram, seriam a escolha do limiar de similaridade usado na busca e a ordem com que as imagens são comparadas.

1. O limiar é decidido com base em um estudo que verifica quais as menores distâncias (ou maiores correlações) entre as imagens de pessoas diferentes e as maiores distâncias (menores correlações) entre imagens da mesma pessoa. Dessa forma, podemos definir um valor possível para determinar se duas imagens pertencem a mesma pessoa

Experimento	Base	Acurácia por experimento	Acurácia por técnica
Busca direta + DE	<i>faces94</i>	100%	90%
	FEI	80%	
	JAFFE	90%	
Busca direta + CP	<i>faces94</i>	100%	86,66%
	FEI	70%	
	JAFFE	90%	
K-Means + DE	<i>faces94</i>	100%	93,33%
	FEI	80%	
	JAFFE	100%	
K-Means + CP	<i>faces94</i>	100%	79,51%
	FEI	50,53%	
	JAFFE	88%	
EM + DE	<i>faces94</i>	100%	93,33%
	FEI	80%	
	JAFFE	100%	
EM + CP	<i>faces94</i>	100%	85,59%
	FEI	57,89%	
	JAFFE	98,89%	

Tabela 19 – Acurácia aferida nas experimentações.

ou não. No entanto, em algumas situações, a distância entre duas imagens de pessoas diferentes é menor que a distância entre duas imagens de uma mesma pessoa, principalmente devido em casos onde as expressões faciais são totalmente distintas (como na base FEI). Esse fator implica em uma sobreposição nas faixas determinantes de semelhança e, conseqüentemente, não é possível usar um valor único no discernimento de todas as imagens.

- Além disso, a ordem com que as imagens são comparadas influencia diretamente no desempenho das técnicas. Considere a base FEI, por exemplo, onde cada pessoa só é representada por duas imagens. Se a imagem teste é comparada a uma imagem de uma pessoa não correspondente, mas a distância entre elas é inferior ao limiar definido, o algoritmo termina a busca porque considera que a imagem correta foi encontrada. Por outro lado, se a imagem da pessoa correta fosse usada para comparação prévia, o valor da distância seria inferior ao de uma pessoa diferente (e ao limiar) e a busca terminaria com a resposta correta.

Esse tipo de comportamento desponta como um gargalo ao método, uma vez que foram determinados valores de distância e correlação absolutos para definição dos limiares.

A normalização ou outra transformação dos dados poderia ajudar a transpor esse empecilho e trazer resultados mais satisfatórios.

### 6.2.2 Análise do Tempo de Busca

Nesta seção serão discutidos os resultados para a segunda etapa dos experimentos, que consistem no processamento das buscas por imagens que não estão presentes nas bases de dados. Nesse caso, similarmente ao experimento anterior, a busca é feita com base na proximidade da imagem teste para os itens no conjunto de dados. No entanto, considerando que as imagens utilizadas representam dados de faces humanas, foi proposto um valor que visa determinar se uma imagem pode ser considerada da mesma pessoa.

Nesse sentido, foram realizados experimentos preliminares para determinar qual o valor limiar nas medidas de proximidade a ser usado pelos algoritmos de busca como “sinalização de parada”. Para cada base foi feito um estudo comparando um único item com todos os demais. A partir disso, foi possível determinar qual o melhor valor para os experimentos. Seguindo essa explanação, a Tabela 20 exibe os valores usados para cada base, considerando as duas medidas de proximidade.

<i>Base de Dados</i>	<i>Distância euclidiana</i>	<i>Correlação de Pearson</i>
<i>faces94</i>	1.9	0.8
FEI	1.7	0.78
JAFFE	1.7	0.82

Tabela 20 – Valores limiares utilizados para determinar que duas imagens pertencem a uma mesma pessoa.

Note que adotar valores diferentes para tais limiares se deve ao fato da natureza distinta das bases. Cada *dataset* foi construído sob diferentes metodologias e consequentes parâmetros/características diferenciados. Por exemplo, a base FEI possui somente duas imagens de cada pessoa (e expressões que demonstram um semblante sério e outra alegre). A base JAFFE, por sua vez, possui em torno de 20 imagens por pessoa, cada uma com uma expressão facial diferente. Além disso, as imagens da primeira base utilizam o esquema de cores RGB, enquanto que a última é composta por imagens em tons de cinza.

Vale salientar que a maioria dos sistemas de reconhecimento facial adota uma política controlada de formação da base de dados, com critérios no registro das figuras (posição no plano, variações nas expressões, etc). Sendo assim, em se tratando de um sistema homogêneo, esses valores podem ser extraídos diretamente do conjunto de dados existente.

A Tabela 21 mostra os resultados da busca direta no conjunto de dados, utilizando os valores limiares destacados anteriormente. De maneira similar aos resultados exibidos na Seção 6.1, temos uma grande variação em todas as bases de dados, indicando que a busca trivial pode ser custosa quando o processo tende a se repetir.

<i>faces94</i>		
Imagem	Tempo mínimo	
	DE	CP
Img 1	0,228588	0,238252
Img 2	0,00251	0,026805
Img 3	0,002002	0,024579
Img 4	0,005894	0,009083
Img 5	0,002072	0,016501
Img 6	0,003147	0,020984
Img 7	0,005173	0,024548
Img 8	0,002941	0,006628
Img 9	0,002012	0,007796
Img 10	0,003267	0,007586
Tempo Médio	0,025761	0,038276
Desvio Padrão	0,071279	0,070708
<b>FEI</b>		
Imagem	Tempo mínimo	
	DE	CP
Img 1	0,060715	0,073576
Img 2	0,012609	0,00956
Img 3	0,003352	0,006705
Img 4	0,002187	0,025976
Img 5	0,003855	0,008105
Img 6	0,002104	0,008958
Img 7	0,003429	0,008445
Img 8	0,005797	0,008947
Img 9	0,002225	0,032031
Img 10	0,002316	0,008619
Tempo Médio	0,009859	0,019092
Desvio Padrão	0,018148	0,021021
<b>JAFFE</b>		
Imagem	Tempo mínimo	
	DE	CP
Img 1	0,215873	0,074576
Img 2	0,001509	0,003722
Img 3	0,001312	0,002679
Img 4	0,003769	0,002698
Img 5	0,003724	0,003336
Img 6	0,001482	0,007861
Img 7	0,004335	0,003145
Img 8	0,001095	0,011049
Img 9	0,001058	0,008316
Img 10	0,000972	0,002714
Tempo Médio	0,023513	0,01201
Desvio Padrão	0,067601	0,022183

Tabela 21 – Busca direta de imagem não existente na base de dados baseada em distância Euclidiana e correlação *Pearson*.

6.2.2.1 Desempenho do Algoritmo *k-means*

As Tabelas 22, 23 e 24 exibem os resultados da busca das imagens nas bases, com a aplicação do método de organização do conjunto de dados em agrupamentos *k-means* usando a distância euclidiana. A organização das tabelas segue o mesmo formato adotado na seção anterior.

Base <i>faces94</i>												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,395673	0,002409	0,002066	0,0022	0,007446	0,005908	0,002521	0,001974	0,003049	0,002075	0,042532	0,124095
3	0,021589	0,002491	0,001948	0,00137	0,003384	0,002772	0,00162	0,001344	0,001729	0,001852	0,004010	0,006211
4	0,001135	0,001543	0,001118	0,001959	0,003216	0,00436	0,001609	0,001286	0,001228	0,001546	0,001900	0,001062
5	0,001603	0,000997	0,00186	0,002023	0,005203	0,003426	0,001505	0,002507	0,001394	0,001194	0,002171	0,001278
6	0,001186	0,004337	0,00176	0,002387	0,004646	0,001017	0,000975	0,001397	0,001342	0,001743	0,002079	0,001340
7	0,000837	0,003366	0,000982	0,003571	0,002542	0,00067	0,001409	0,001298	0,001536	0,001473	0,001768	0,001032
8	0,001191	0,004184	0,000698	0,001515	0,002155	0,000686	0,00084	0,000819	0,000697	0,003026	0,001581	0,001194
9	0,000982	0,005203	0,001432	0,001739	0,003255	0,00103	0,001413	0,00083	0,001517	0,003023	0,002042	0,001382
10	0,001108	0,004587	0,000725	0,001145	0,002997	0,001087	0,001281	0,001026	0,000924	0,002644	0,001752	0,001252
11	0,0014	0,004114	0,000976	0,00115	0,00276	0,001028	0,002837	0,001097	0,001018	0,003528	0,001991	0,001199
12	0,00158	0,004726	0,001324	0,002041	0,002761	0,00113	0,0028	0,000683	0,000963	0,004013	0,002202	0,001354
13	0,00397	0,004022	0,001015	0,000699	0,002284	0,001035	0,003302	0,000709	0,001208	0,003028	0,002127	0,001355
14	0,002249	0,003795	0,001055	0,000998	0,002335	0,000688	0,002384	0,000697	0,000687	0,002423	0,001731	0,001056
15	0,002847	0,003842	0,001064	0,000955	0,003087	0,00107	0,002101	0,000687	0,000652	0,002384	0,001869	0,001138
16	0,003028	0,002846	0,001106	0,001335	0,004545	0,0007	0,002421	0,000689	0,000679	0,004399	0,002175	0,001501
17	0,003811	0,003135	0,002388	0,000845	0,002693	0,001007	0,002201	0,002489	0,000682	0,003116	0,002237	0,001066
18	0,00228	0,002179	0,004199	0,001338	0,002348	0,00107	0,002177	0,002334	0,000688	0,00331	0,002192	0,001032
19	0,003191	0,002296	0,00323	0,000688	0,002426	0,000693	0,002512	0,002397	0,000913	0,003338	0,002168	0,001040

Tabela 22 – Busca com auxílio de agrupamento usando o *K-Means* e distância euclidiana na base *faces94*.

Base FEI												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,482212	0,008708	0,008708	0,004323	0,002698	0,003052	0,00393	0,004473	0,002818	0,002818	0,052374	0,151047
3	0,00295	0,010238	0,010238	0,003762	0,002712	0,003483	0,003383	0,001574	0,004155	0,004155	0,004665	0,003033
4	0,005014	0,003898	0,003898	0,004228	0,002416	0,002483	0,002803	0,001465	0,002566	0,002566	0,003134	0,001074
5	0,007234	0,001112	0,001112	0,003938	0,00747	0,00167	0,002588	0,001526	0,00347	0,00347	0,003359	0,002338
6	0,006529	0,001638	0,001638	0,004075	0,002335	0,002731	0,002395	0,001453	0,002208	0,002208	0,002721	0,001530
7	0,005832	0,001881	0,001881	0,002595	0,001288	0,002625	0,001664	0,001306	0,002285	0,002285	0,002364	0,001308
8	0,001996	0,002107	0,002107	0,0078	0,003211	0,003308	0,004105	0,004036	0,001691	0,001691	0,003200	0,001845
9	0,003998	0,002325	0,002325	0,00218	0,008372	0,002641	0,002872	0,003983	0,00172	0,00172	0,003214	0,001982
10	0,00694	0,00136	0,00136	0,004388	0,002003	0,00277	0,002804	0,00411	0,002656	0,002656	0,003105	0,001676
11	0,002062	0,002246	0,002246	0,004267	0,005509	0,002879	0,002811	0,003973	0,001962	0,001962	0,002992	0,001205
12	0,001991	0,001845	0,001845	0,00482	0,008341	0,002505	0,002262	0,002664	0,00148	0,00148	0,002923	0,002134
13	0,002301	0,006151	0,006151	0,002932	0,004135	0,002277	0,003913	0,00368	0,001957	0,001957	0,003545	0,001587
14	0,003684	0,002689	0,002689	0,00267	0,001862	0,002829	0,003363	0,00788	0,00209	0,00209	0,003185	0,001743
15	0,00549	0,002907	0,002907	0,003964	0,002015	0,003175	0,002877	0,008214	0,002569	0,002569	0,003669	0,001863
16	0,002191	0,001842	0,001842	0,004711	0,002061	0,004239	0,002159	0,009356	0,002713	0,002713	0,003383	0,002321
17	0,000854	0,003725	0,003725	0,00368	0,001777	0,00238	0,003018	0,010507	0,003244	0,003244	0,003615	0,002597
18	0,000592	0,002274	0,002274	0,009187	0,00143	0,004889	0,002488	0,002599	0,002832	0,002832	0,003140	0,002391
19	0,001151	0,002431	0,002431	0,013119	0,002447	0,002207	0,003073	0,009523	0,002578	0,002578	0,004154	0,003901
20	0,000828	0,001581	0,001581	0,01396	0,001769	0,001959	0,002376	0,006592	0,002619	0,002619	0,003588	0,003966

Tabela 23 – Busca com auxílio de agrupamento usando o *K-Means* e distância euclidiana na base FEI.

Base JAFFE												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,268493	0,001378	0,001398	0,001435	0,001128	0,004717	0,001051	0,001526	0,001625	0,000925	0,028368	0,084379
3	0,000857	0,001392	0,001315	0,001404	0,000993	0,002849	0,001308	0,00091	0,001266	0,000913	0,001321	0,000578
4	0,001348	0,002804	0,001153	0,001063	0,000799	0,002852	0,001224	0,001423	0,001027	0,000874	0,001457	0,000748
5	0,001003	0,002468	0,000904	0,000742	0,000803	0,003472	0,002269	0,001421	0,000803	0,000803	0,001469	0,000946
6	0,000671	0,003661	0,001069	0,000783	0,000877	0,003418	0,002419	0,001282	0,001287	0,000998	0,001647	0,001110
7	0,00079	0,002872	0,001115	0,00074	0,00076	0,00226	0,00234	0,001214	0,000744	0,000772	0,001361	0,000812
8	0,000799	0,003771	0,000802	0,0008	0,000748	0,00262	0,002104	0,000814	0,000793	0,000764	0,001398	0,001068
9	0,000821	0,002594	0,001049	0,000761	0,000689	0,00223	0,002079	0,001057	0,001046	0,000727	0,001305	0,000711
10	0,000698	0,00355	0,000789	0,000647	0,000703	0,0024	0,002033	0,000766	0,001101	0,000665	0,001335	0,000996

Tabela 24 – Busca com auxílio de agrupamento usando o *K-Means* e distância euclidiana na base JAFFE.

Quando analisamos o resultados da busca com o auxílio do *k-means* e distância euclidiana, podemos notar mais uma vez uma redução considerável no tempo de busca. Na totalidade das bases, o tempo foi compendiado em pelo menos 10 vezes se comparado a busca direta. Considerando esses resultados e também a tabela de acurácia vista anteriormente (Tabela 19), o uso do *k-means* em conjunto com a distância euclidiana conseguiu trazer, além do ganho temporal, benefícios nos aspectos de acurácia, fazendo com que mais imagens fossem corretamente identificadas.

A seguir, nas Tabelas 25, 26 e 27 são mostrados os resultados da busca com auxílio do *k-means* quando abordada a correlação para medir similaridade.

Base faces94												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,150079	0,012529	0,004139	0,009249	0,023297	0,008832	0,005409	0,003318	0,006658	0,027346	0,025086	0,044651
3	0,005021	0,004535	0,010751	0,005547	0,009997	0,018913	0,004034	0,002192	0,006011	0,012899	0,007990	0,005110
4	0,006296	0,009372	0,007018	0,005824	0,01044	0,007672	0,003047	0,003782	0,007443	0,015199	0,007609	0,003488
5	0,004708	0,005421	0,005318	0,004139	0,010232	0,009034	0,004724	0,005769	0,004475	0,014266	0,006809	0,003314
6	0,004862	0,007023	0,002379	0,004685	0,00939	0,009536	0,004721	0,003226	0,003231	0,011074	0,006013	0,003054
7	0,005486	0,004126	0,003352	0,006735	0,008578	0,007233	0,003946	0,00278	0,002787	0,018642	0,006367	0,004748
8	0,003965	0,006268	0,002133	0,003841	0,006032	0,009856	0,002814	0,003243	0,003101	0,01551	0,005676	0,004140
9	0,00282	0,004241	0,001891	0,00335	0,010133	0,008552	0,003613	0,002127	0,001843	0,012809	0,005138	0,003910
10	0,00292	0,0034	0,003338	0,003578	0,008475	0,009918	0,004003	0,002805	0,002153	0,012698	0,005329	0,003652
11	0,004894	0,004806	0,002363	0,003405	0,006779	0,006231	0,003621	0,002865	0,002986	0,011235	0,004919	0,002657
12	0,005666	0,002863	0,003245	0,002415	0,009981	0,009102	0,00475	0,00278	0,002847	0,006986	0,005064	0,002788
13	0,00399	0,002873	0,002077	0,003966	0,01028	0,009588	0,002891	0,002801	0,002931	0,010015	0,005141	0,003376
14	0,004945	0,002844	0,00284	0,002839	0,008862	0,002124	0,00312	0,002842	0,001858	0,008603	0,004088	0,002580
15	0,003871	0,002698	0,002122	0,010363	0,009057	0,002369	0,002227	0,003345	0,002221	0,011788	0,005006	0,003818
16	0,003407	0,003618	0,004071	0,013055	0,009105	0,002577	0,002141	0,002766	0,002907	0,009551	0,005320	0,003803
17	0,003048	0,00323	0,004927	0,009901	0,007995	0,002697	0,001916	0,002954	0,001959	0,011141	0,004977	0,003430
18	0,005418	0,002912	0,001621	0,010231	0,006871	0,002691	0,00274	0,003343	0,002313	0,011226	0,004937	0,003431
19	0,002472	0,002564	0,002079	0,013231	0,009582	0,001991	0,00258	0,002132	0,002724	0,007248	0,004660	0,003970

Tabela 25 – Busca com auxílio de agrupamento usando o *K-Means* e correlação de *Pearson* na base *faces94*.

O tempo de busca quando o agrupamento é gerado pelo algoritmo *k-means*,

Base FEI												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,068459	0,00904	0,005464	0,004418	0,027737	0,01831	0,00974	0,00974	0,003659	0,003707	0,016027	0,019918
3	0,00367	0,008889	0,003715	0,00303	0,026381	0,012301	0,008257	0,008257	0,002865	0,005692	0,008306	0,007070
4	0,005891	0,008989	0,003378	0,004465	0,024926	0,015598	0,006046	0,006046	0,003375	0,005964	0,008468	0,006786
5	0,005677	0,003297	0,002846	0,003497	0,023483	0,011093	0,007957	0,007957	0,00275	0,003577	0,007213	0,006365
6	0,005873	0,004794	0,004868	0,005141	0,022138	0,009671	0,006457	0,006457	0,002771	0,002988	0,007116	0,005626
7	0,00674	0,004882	0,003281	0,002897	0,021212	0,008159	0,005711	0,005711	0,002744	0,003172	0,006451	0,005494
8	0,003143	0,008045	0,004108	0,003354	0,021059	0,008408	0,005654	0,005654	0,002575	0,002426	0,006443	0,005561
9	0,006565	0,00472	0,003268	0,002807	0,022168	0,00814	0,008828	0,008828	0,002599	0,002439	0,007036	0,005917
10	0,002165	0,006746	0,002951	0,003879	0,028219	0,008281	0,007988	0,007988	0,002312	0,002711	0,007324	0,007766
11	0,001855	0,005004	0,005918	0,00283	0,025092	0,008061	0,009233	0,009233	0,002776	0,003643	0,007365	0,006792
12	0,004571	0,008477	0,004715	0,003325	0,004619	0,00774	0,009017	0,009017	0,002233	0,003967	0,005768	0,002537
13	0,002776	0,008824	0,002715	0,003392	0,004613	0,007988	0,006848	0,006848	0,002505	0,003446	0,004996	0,002402
14	0,001966	0,004989	0,004172	0,003939	0,006644	0,007245	0,008019	0,008019	0,002167	0,002881	0,005004	0,002345
15	0,006678	0,005241	0,002138	0,009703	0,00493	0,007894	0,011796	0,011796	0,002177	0,002069	0,006442	0,003796
16	0,001875	0,005165	0,003164	0,004101	0,004685	0,007615	0,012171	0,012171	0,00332	0,003245	0,005751	0,003712
17	0,001931	0,006751	0,003052	0,006331	0,005462	0,007661	0,007507	0,007507	0,00275	0,002463	0,005142	0,002338
18	0,001875	0,00594	0,003159	0,005876	0,014446	0,007002	0,007777	0,007777	0,002236	0,002579	0,005867	0,003793
19	0,002623	0,00429	0,00318	0,004274	0,009626	0,006788	0,004882	0,004882	0,002457	0,00346	0,004646	0,002166
20	0,005156	0,004268	0,003095	0,00552	0,008923	0,006684	0,009775	0,009775	0,002306	0,002434	0,005794	0,002903

Tabela 26 – Busca com auxílio de agrupamento usando o *K-Means* e correlação de *Pearson* na base FEI.

Base JAFFE												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,060729	0,003877	0,003749	0,004162	0,00262	0,012256	0,00348	0,003111	0,010726	0,002641	0,010735	0,017897
3	0,00263	0,002939	0,003265	0,005992	0,003306	0,013676	0,003755	0,002611	0,008318	0,003629	0,005012	0,003530
4	0,001867	0,00321	0,002992	0,003752	0,003382	0,011122	0,003258	0,003934	0,007409	0,002914	0,004384	0,002774
5	0,002317	0,002435	0,002557	0,002564	0,003476	0,009595	0,002831	0,002194	0,005638	0,003344	0,003695	0,002305
6	0,001881	0,00255	0,003003	0,001904	0,002602	0,00973	0,003479	0,003106	0,006744	0,002915	0,003791	0,002497
7	0,002041	0,002319	0,002686	0,001858	0,002996	0,0082	0,003229	0,007223	0,006419	0,002398	0,003937	0,002380
8	0,00184	0,002778	0,002448	0,001798	0,0026	0,002717	0,00208	0,008083	0,006436	0,002028	0,003281	0,002161
9	0,002096	0,002067	0,002044	0,001807	0,002733	0,002061	0,003117	0,006978	0,005709	0,002097	0,003071	0,001793
10	0,003402	0,002946	0,002013	0,001981	0,002815	0,002285	0,003092	0,010606	0,006077	0,003321	0,003854	0,002643

Tabela 27 – Busca com auxílio de agrupamento usando o *K-Means* e correlação de *Pearson* na base JAFFE.

utilizando-se a correlação para encontrar uma imagem teste na base de dados, é semelhante àquele encontrado quando se usa a distância euclidiana. De fato, os resultados apontam para um padrão de redução temporal nas duas medidas de similaridade. Entretanto, o uso da distância euclidiana (0,001581, 0,002364 e 0,001305 como melhores tempos médios nas bases *faces94*, FEI e JAFFE, respectivamente) ainda se mostra levemente superior à correlação (0,004088, 0,004646 e 0,003071, nas três respectivas bases).

Por outro lado, quando analisamos a acurácia do sistema, observamos uma queda nas bases FEI e JAFFE. Esse decréscimo no desempenho ocorre mais perceptivelmente na base FEI, onde a acurácia parte de 70% para 50,53%. Na base JAFFE, a queda é bem inferior, saindo de 90% na busca direta para 88% na busca com agrupamento.

Essa redução de desempenho deve ser considerada na avaliação sobre o custo-benefício, pois o uso da técnica reduz o tempo de busca, todavia deteriora a sua precisão, em alguns casos.

### 6.2.2.2 Desempenho do Algoritmo EM

As Tabelas 28, 29 e 30 exibem os resultados da busca das imagens nas bases, utilizando o método de organização do conjunto de dados em agrupamentos com auxílio do algoritmo EM usando a distância euclidiana. A organização das tabelas segue o mesmo formato adotado na seção anterior.

Base faces94												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,051014	0,009383	0,002201	0,00156	0,002156	0,007628	0,001519	0,002179	0,008675	0,001686	0,008800	0,015171
3	0,00126	0,001485	0,001151	0,003418	0,001861	0,004443	0,001757	0,001352	0,006639	0,001564	0,002493	0,001804
4	0,001216	0,001829	0,001589	0,001916	0,001402	0,004303	0,001573	0,001218	0,001611	0,000955	0,001761	0,000940
5	0,001599	0,004993	0,001484	0,001705	0,001504	0,004906	0,001569	0,001475	0,001007	0,001076	0,002132	0,001501
6	0,001134	0,002527	0,00153	0,000989	0,001457	0,004095	0,001315	0,001006	0,001505	0,001055	0,001661	0,000966
7	0,001254	0,001016	0,001325	0,001066	0,001422	0,002785	0,001427	0,001017	0,001402	0,001161	0,001388	0,000517
8	0,001264	0,002307	0,000996	0,000919	0,001344	0,005472	0,001381	0,000962	0,001387	0,001092	0,001712	0,001380
9	0,00122	0,000934	0,001471	0,001003	0,000966	0,002482	0,001406	0,001195	0,001535	0,001305	0,001352	0,000449
10	0,001346	0,000882	0,001316	0,002473	0,001308	0,003896	0,00147	0,001222	0,000949	0,001352	0,001621	0,000908
11	0,001262	0,000982	0,00137	0,001078	0,000917	0,002555	0,003684	0,000984	0,001329	0,00084	0,001500	0,000911
12	0,0013	0,000934	0,001348	0,000833	0,001377	0,002424	0,031352	0,00084	0,000969	0,000818	0,004220	0,009546
13	0,001319	0,001178	0,001265	0,000945	0,001357	0,003613	0,003602	0,000828	0,001347	0,000938	0,001639	0,001054
14	0,00123	0,000903	0,001375	0,000856	0,001516	0,001621	0,003839	0,001204	0,001318	0,001216	0,001508	0,000853
15	0,001286	0,001271	0,001296	0,001235	0,001349	0,002033	0,004716	0,000797	0,000961	0,001096	0,001604	0,001140
16	0,000899	0,000866	0,001222	0,001254	0,001333	0,003424	0,004498	0,000862	0,001089	0,000786	0,001623	0,001272
17	0,001333	0,000859	0,000908	0,000844	0,000951	0,001627	0,003761	0,000834	0,00132	0,000906	0,001334	0,000895
18	0,000898	0,004067	0,001307	0,000862	0,001367	0,001769	0,00391	0,000841	0,001276	0,002421	0,001872	0,001213
19	0,001361	0,000907	0,000895	0,000853	0,001559	0,001064	0,002654	0,000887	0,001286	0,003269	0,001474	0,000832

Tabela 28 – Busca com auxílio de agrupamento usando o algoritmo EM e distância euclidiana na base *faces94*.

Novamente, a busca com uso da distância euclidiana e auxílio de agrupamentos EM gera resultados semelhantes ao *k-means*. Nesse caso, a diferença entre os algoritmos só pode ser notada a partir da quarta casa decimal nos valores aferidos. Sendo assim, o algoritmo EM também traz benefícios relativos ao tempo de busca (reduzindo a duração em mais de dez vezes) quando usado em conjunto com a distância euclidiana. No quesito de acurácia, o EM também entrega resultados iguais ou melhores que a busca direta.

Esses resultados também são válidos para o EM quando usamos a correlação de *Pearson* para aferição de similaridade entre as imagens. Nesse caso, o desempenho para duas das três bases (*faces94* e *JAFFE*) foi ainda melhor que o *k-means*. No caso específico da base *faces94*, o tempo de busca foi reduzido em mais de setenta vezes em relação a busca direta. Os resultados do EM com correlação podem ser vistos nas Tabelas 31, 32 e 33.

Base FEI												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,064838	0,001548	0,001548	0,003389	0,001857	0,002768	0,002306	0,006922	0,009339	0,009339	0,010385	0,019381
3	0,002443	0,002284	0,002284	0,00315	0,001918	0,008529	0,001493	0,00872	0,010459	0,010459	0,005174	0,003831
4	0,003537	0,00117	0,00117	0,001824	0,001109	0,001461	0,0068	0,004827	0,010372	0,010372	0,004264	0,003724
5	0,001357	0,003313	0,003313	0,001237	0,001597	0,003942	0,003303	0,009858	0,015134	0,015134	0,005819	0,005487
6	0,004481	0,003961	0,003961	0,005171	0,000955	0,001547	0,003627	0,012602	0,016149	0,016149	0,006860	0,005816
7	0,001624	0,002257	0,002257	0,003371	0,002862	0,003021	0,002831	0,002972	0,00603	0,00603	0,003326	0,001509
8	0,003233	0,001172	0,001172	0,0032	0,002117	0,002489	0,007173	0,007179	0,006658	0,006658	0,004110	0,002520
9	0,004262	0,001853	0,001853	0,000962	0,002002	0,001529	0,004179	0,01508	0,002823	0,002823	0,003737	0,004130
10	0,001352	0,000936	0,000936	0,003745	0,001339	0,002619	0,005236	0,0209	0,024584	0,024584	0,008623	0,010305
11	0,001827	0,001752	0,001752	0,006265	0,002756	0,003214	0,012851	0,010122	0,019088	0,019088	0,007872	0,007019
12	0,00297	0,00089	0,00089	0,007864	0,00123	0,002503	0,016043	0,007847	0,008403	0,008403	0,005704	0,004885
13	0,001026	0,006676	0,006676	0,025052	0,012495	0,001638	0,003669	0,012564	0,003732	0,003732	0,007726	0,007294
14	0,002055	0,001968	0,001968	0,012169	0,008309	0,010477	0,013882	0,014079	0,007408	0,007408	0,007972	0,004761
15	0,005148	0,005428	0,005428	0,009693	0,007259	0,006245	0,018204	0,002395	0,00251	0,00251	0,006482	0,004722
16	0,002723	0,00229	0,00229	0,004894	0,027402	0,004772	0,018921	0,001106	0,022005	0,022005	0,010841	0,010371
17	0,002459	0,002054	0,002054	0,003818	0,007894	0,00148	0,009653	0,001723	0,005546	0,005546	0,004223	0,002846
18	0,002019	0,002927	0,002927	0,007073	0,005911	0,003295	0,005516	0,003326	0,008334	0,008334	0,004966	0,002376
19	0,001781	0,000879	0,000879	0,023795	0,029559	0,008688	0,01931	0,006215	0,001636	0,001636	0,009438	0,010782
20	0,000911	0,001446	0,001446	0,00783	0,004209	0,00142	0,002342	0,001252	0,000895	0,000895	0,002265	0,002197

Tabela 29 – Busca com auxílio de agrupamento usando o algoritmo EM e distância euclidiana na base FEI.

Base JAFFE												
k	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,054217	0,000944	0,003268	0,001061	0,000982	0,000964	0,000997	0,001592	0,001051	0,001115	0,006619	0,016739
3	0,001619	0,000915	0,003955	0,000816	0,001603	0,001065	0,000817	0,001287	0,000789	0,000837	0,001370	0,000963
4	0,000738	0,000802	0,002354	0,0009	0,001251	0,001105	0,000734	0,001276	0,000773	0,000794	0,001073	0,000496
5	0,002508	0,000783	0,002332	0,000961	0,001131	0,001336	0,000715	0,000806	0,000716	0,000746	0,001203	0,000673
6	0,000823	0,000767	0,002158	0,000942	0,000886	0,001099	0,000801	0,00131	0,002248	0,001047	0,001208	0,000549
7	0,004352	0,000736	0,002158	0,002259	0,000799	0,000957	0,000755	0,001123	0,002148	0,001173	0,001646	0,001132
8	0,000696	0,000756	0,002338	0,0011	0,000742	0,001022	0,000701	0,000782	0,002119	0,001024	0,001127	0,000601
9	0,001225	0,000778	0,003433	0,001062	0,001716	0,001448	0,000692	0,001099	0,002083	0,000651	0,001419	0,000842
10	0,000682	0,000676	0,002581	0,000821	0,000828	0,000744	0,000679	0,001199	0,002218	0,001021	0,001145	0,000687

Tabela 30 – Busca com auxílio de agrupamento usando o algoritmo EM e distância euclidiana na base JAFFE.

É importante observar que o uso de correlação trouxe prejuízos ao nível de acurácia na identificação de imagens. A base FEI teve sua precisão reduzida de 70% para 57,89%. Por outro lado, na base JAFFE houve um incremento nessa precisão, incrementando o valor da busca direta, anteriormente em 90%, para 98,89%. Sendo assim, o uso de correlação ainda solicita cautela, já que em alguns casos pode levar a uma diminuição na precisão da busca.

### 6.2.3 *k-means* vs EM

Com o objetivo de sintetizar a análise feita na comparação entre os métodos de busca, os desempenhos médios (e desvios padrões) de cada algoritmo foram compilados

Base faces94												
$k$	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,311399	0,003738	0,003912	0,010046	0,003935	0,006619	0,003785	0,005029	0,010628	0,02113	0,038022	0,096206
3	0,014218	0,002766	0,004373	0,005623	0,003297	0,00577	0,002831	0,0056	0,003168	0,033534	0,008118	0,009543
4	0,009796	0,002637	0,00316	0,004334	0,002681	0,008089	0,002736	0,003852	0,002641	0,021601	0,006152	0,005978
5	0,008698	0,009168	0,002742	0,004274	0,002572	0,009626	0,002752	0,004258	0,002635	0,002929	0,004965	0,002971
6	0,009731	0,011155	0,003485	0,006784	0,002735	0,010439	0,002563	0,002808	0,002767	0,002927	0,005539	0,003614
7	0,011048	0,007321	0,003286	0,003207	0,002727	0,02033	0,00267	0,006142	0,00255	0,002409	0,006169	0,005716
8	0,007049	0,00986	0,003328	0,002792	0,002331	0,014078	0,003219	0,00488	0,002352	0,002506	0,005239	0,003956
9	0,002434	0,010556	0,002345	0,002752	0,002936	0,013111	0,002503	0,004601	0,002436	0,003272	0,004694	0,003868
10	0,003953	0,006444	0,002337	0,003241	0,009764	0,01663	0,00225	0,003738	0,00214	0,002495	0,005299	0,004643
11	0,003252	0,006533	0,005139	0,003289	0,01009	0,011109	0,002114	0,004886	0,002767	0,002745	0,005192	0,003156
12	0,00316	0,007911	0,002303	0,00250	0,009986	0,009297	0,002048	0,002446	0,002148	0,002935	0,004473	0,003224
13	0,003331	0,010011	0,002346	0,002249	0,00886	0,011536	0,00225	0,005489	0,002191	0,0025	0,005076	0,003682
14	0,003251	0,010387	0,003159	0,005866	0,00917	0,008628	0,002168	0,005534	0,002162	0,002247	0,005257	0,003163
15	0,00283	0,015418	0,002852	0,003719	0,006265	0,011526	0,002159	0,004439	0,002251	0,004339	0,005579	0,004431
16	0,003692	0,005953	0,003249	0,003310	0,006387	0,008386	0,002191	0,008692	0,001983	0,002183	0,0046026	0,002558
17	0,002574	0,007105	0,00333	0,007769	0,007193	0,007846	0,002201	0,01133	0,002294	0,008562	0,006020	0,003183
18	0,002419	0,006574	0,003328	0,007805	0,006277	0,011707	0,002275	0,011577	0,002177	0,004064	0,005820	0,003640
19	0,002136	0,011764	0,002281	0,013973	0,006398	0,012178	0,002871	0,011617	0,002269	0,003385	0,006887	0,004923

Tabela 31 – Busca com auxílio de agrupamento usando o EM e correlação de *Pearson* na base *faces94*.

Base FEI												
$k$	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,149444	0,006194	0,005777	0,021868	0,02521	0,007309	0,004508	0,017914	0,00547	0,019094	0,026279	0,043973
3	0,004925	0,003953	0,00454	0,014003	0,035028	0,003269	0,003775	0,006798	0,00477	0,022723	0,010378	0,010628
4	0,008957	0,003381	0,010549	0,011751	0,012552	0,002932	0,011615	0,003643	0,004082	0,0182	0,008766	0,005108
5	0,002629	0,011232	0,009584	0,010557	0,015218	0,006443	0,010009	0,007194	0,003962	0,012831	0,008966	0,003914
6	0,008194	0,007293	0,010424	0,0326	0,008358	0,002828	0,010993	0,014619	0,00585	0,009733	0,011089	0,008188
7	0,002707	0,007208	0,003374	0,02135	0,022442	0,003319	0,003574	0,004096	0,003589	0,027621	0,009928	0,009780
8	0,0051	0,003003	0,004804	0,014318	0,017892	0,002528	0,007549	0,008089	0,00315	0,017856	0,008429	0,006058
9	0,007768	0,006698	0,014336	0,009282	0,017566	0,00284	0,004889	0,014275	0,003873	0,015844	0,009737	0,005366
10	0,002386	0,002714	0,003194	0,016257	0,011736	0,003779	0,006125	0,019995	0,003508	0,035642	0,010534	0,010801
11	0,003457	0,007122	0,003338	0,011063	0,024018	0,00214	0,014635	0,008582	0,002891	0,012667	0,008991	0,006879
12	0,021239	0,003097	0,005359	0,026619	0,011943	0,002281	0,021334	0,004979	0,002262	0,010391	0,010950	0,009066
13	0,002608	0,012199	0,00544	0,015075	0,021363	0,005047	0,004907	0,007779	0,003656	0,009298	0,008737	0,005919
14	0,002207	0,005504	0,002487	0,006914	0,028672	0,002304	0,015861	0,016893	0,003326	0,052051	0,013622	0,016050
15	0,007686	0,020625	0,004949	0,0365	0,023846	0,002221	0,019609	0,005507	0,003255	0,056963	0,018116	0,017648
16	0,005926	0,008482	0,004787	0,010434	0,055822	0,002144	0,026471	0,003194	0,002483	0,004538	0,012428	0,016841
17	0,00299	0,003293	0,014525	0,006417	0,01593	0,002903	0,03783	0,004996	0,002538	0,002594	0,009402	0,011153
18	0,003244	0,005692	0,011053	0,016854	0,006659	0,002508	0,01321	0,00705	0,0032	0,008611	0,007808	0,004699
19	0,003624	0,002371	0,002554	0,006646	0,039935	0,002143	0,070009	0,02254	0,008097	0,004359	0,016228	0,022433
20	0,002691	0,004006	0,004524	0,008193	0,002118	0,002151	0,027218	0,003154	0,002454	0,007653	0,006416	0,007628

Tabela 32 – Busca com auxílio de agrupamento usando o EM e correlação de *Pearson* na base FEI.

Base JAFFE												
$k$	Img 1	Img 2	Img 3	Img 4	Img 5	Img 6	Img 7	Img 8	Img 9	Img 10	TM	DP
2	0,066622	0,01379	0,003248	0,009776	0,022328	0,005074	0,012397	0,002825	0,003582	0,003215	0,0142857	0,019445
3	0,002504	0,012017	0,002439	0,008172	0,010945	0,00286	0,008693	0,002406	0,004399	0,002614	0,0057049	0,003851
4	0,00224	0,008569	0,002411	0,007748	0,010022	0,002729	0,010139	0,002388	0,003216	0,00236	0,0051822	0,003465
5	0,001925	0,007824	<u>0,002331</u>	0,006305	0,006802	0,002348	0,006841	0,002131	0,002243	0,002268	<u>0,0041018</u>	0,002476
6	0,00225	0,010709	0,007032	<u>0,006078</u>	0,009798	0,002103	0,007808	0,002016	<u>0,002138</u>	0,00307	0,0053002	0,003410
7	0,002142	<u>0,007148</u>	0,006263	0,010546	0,010361	<u>0,001897</u>	0,00712	0,002009	0,003023	0,001964	0,0052473	0,003487
8	0,001965	0,010224	0,006401	0,007527	<u>0,006626</u>	0,001917	0,007228	<u>0,001865</u>	0,002753	0,002748	0,0049254	0,003016
9	0,004013	0,010038	0,00717	0,006649	0,027429	0,002105	0,006871	0,001884	0,002879	0,003764	0,0072802	0,007546
10	<u>0,001776</u>	0,011727	0,007907	0,007577	0,007206	0,002064	<u>0,002042</u>	0,001908	0,003097	<u>0,001785</u>	0,0047089	0,003585

Tabela 33 – Busca com auxílio de agrupamento usando o EM e correlação de *Pearson* na base JAFFE.

na Tabela 34, e dispostos visualmente na Figura 20. Para cada base são exibidos os resultados obtidos durante os experimentos com a busca direta, e melhores tempos médios, na pesquisa com auxílio de agrupamento utilizando os algoritmos *k-means* e EM, visando a imagem teste em uma base onde a mesma inexistente, baseada em distância euclidiana (DE) e correlação de *Pearson* (CP).

Base de dados	<i>faces94</i>	FEI	JAFFE
	TM ± DP	TM ± DP	TM ± DP
Busca direta	DE 0,025761 ± 0,071279	0,009859 ± 0,018148	0,023513 ± 0,067601
	CP 0,038276 ± 0,070708	0,019092 ± 0,021021	0,012010 ± 0,022183
<i>k-means</i>	DE <u>0,001581</u> ± 0,001194	<u>0,002364</u> ± 0,001308	<u>0,001305</u> ± 0,000711
	CP <u>0,004088</u> ± 0,002580	<u>0,004646</u> ± 0,002166	<u>0,003071</u> ± 0,001793
EM	DE <u>0,001334</u> ± 0,000895	<u>0,002265</u> ± 0,002197	<u>0,001073</u> ± 0,000496
	CP <u>0,004473</u> ± 0,003224	<u>0,006416</u> ± 0,007628	<u>0,004101</u> ± 0,002476

Tabela 34 – Resultados sintetizados para os métodos de busca.

Os resultados sintetizados mostram que, quando comparada com a busca direta, a estratégia de organizar os dados em agrupamentos promove uma redução significativa no tempo despendido na produção busca. Esse fato é válido, de uma maneira geral, para todos os algoritmos de agrupamento e/ou medida de similaridade usados. Além de reduzir o valor médio do tempo de busca, o uso de agrupamento também reduz o desvio padrão, o que mostra uma maior robustez nas consultas e indica que o método é consistente no direcionamento das buscas.

Avaliando os pareceres do teste de *Wilcoxon* pareado, ressaltados nos valores sublinhados na Tabela 34, pode-se verificar a rejeição da hipótese nula próxima da totalidade dos resultados. Ou seja, a redução do tempo de pesquisa e retorno, confrontados os números apresentados na busca com auxílio de agrupamento e busca simplista, foi realmente validada, comprovando assim a eficiência do método. No contexto geral, apenas o uso de EM + correlação de *Pearson* na base JAFFE não conseguiu passar no teste.

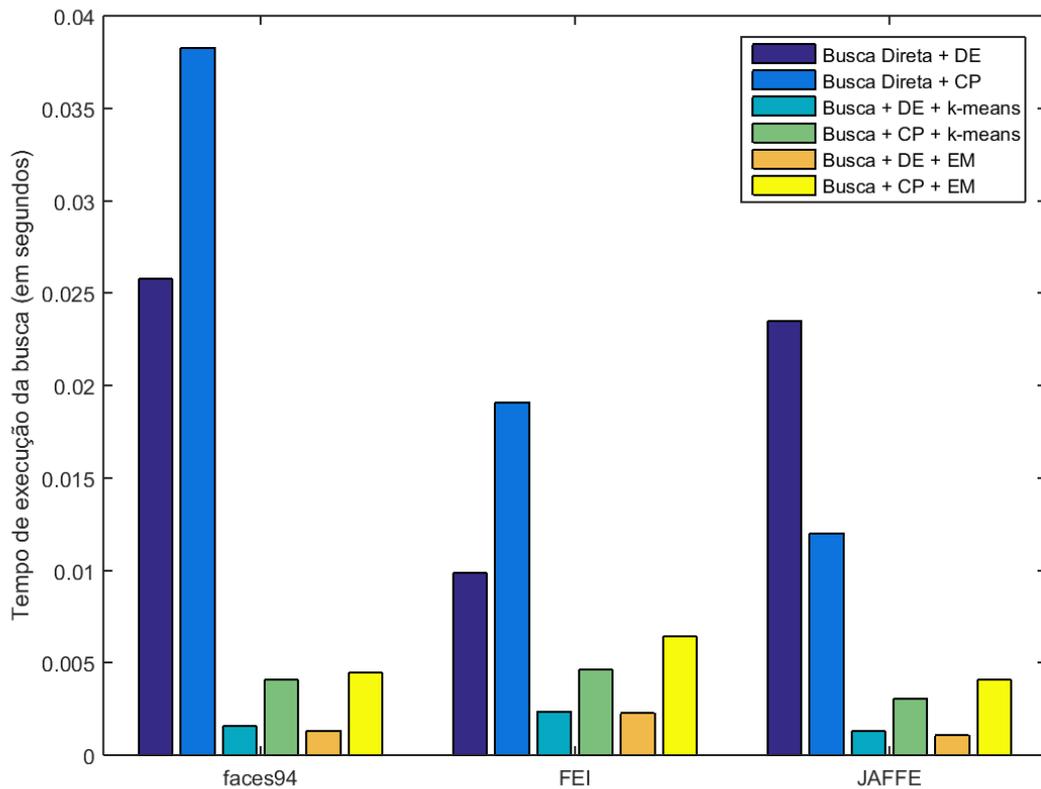


Figura 20 – Gráfico relacionado a totalidade dos resultados da buscas.

Observados os dados trazidos pela Tabela 35, que sintetiza o ganho percentual relativo obtido através da implementação da busca com auxílio de agrupamento como forma de otimização, o ganho de tempo médio no âmbito geral variou de 53,06% (usando *k-means* e correlação de *Pearson*) a 95,44% (usando EM e distância euclidiana), ambos na base JAFFE. Sendo assim, o incremento de desempenho é considerável.

<b>Base faces94</b>				
<b>k</b>	<b><i>k-means</i>+DE</b>	<b><i>k-means</i>+CP</b>	<b>EM+DE</b>	<b>EM+CP</b>
2	-65,10%	82,61%	65,84%	0,66%
3	84,43%	84,30%	90,32%	78,79%
4	92,62%	80,55%	93,16%	83,93%
5	91,57%	88,31%	91,72%	87,03%
6	91,93%	91,56%	93,55%	85,53%
7	93,14%	92,72%	94,61%	83,88%
8	93,86%	91,90%	93,35%	86,31%
9	92,07%	95,18%	94,75%	87,73%
10	93,20%	94,38%	93,71%	86,16%
11	92,27%	92,20%	94,18%	86,43%
12	91,45%	92,56%	83,62%	88,31%
13	91,74%	92,34%	93,64%	86,74%
14	93,28%	95,15%	94,15%	86,27%
15	92,75%	94,20%	93,77%	85,42%
16	91,56%	92,41%	93,70%	87,98%
17	91,32%	94,88%	94,82%	84,27%
18	91,49%	93,96%	92,73%	84,79%
19	91,58%	92,88%	94,28%	82,01%

<b>Base FEI</b>				
<b>k</b>	<b><i>k-means</i>+DE</b>	<b><i>k-means</i>+CP</b>	<b>EM+DE</b>	<b>EM+CP</b>
2	-431,23%	80,83%	-5,34%	-37,64%
3	52,68%	84,99%	47,52%	45,64%
4	68,21%	82,32%	56,75%	54,08%
5	65,93%	85,60%	40,98%	53,04%
6	72,40%	85,49%	30,42%	41,92%
7	76,02%	85,63%	66,27%	48,00%
8	67,54%	86,51%	58,31%	55,85%
9	67,40%	86,39%	62,10%	49,00%
10	68,51%	87,89%	12,54%	44,83%
11	69,66%	85,46%	20,16%	52,91%
12	70,35%	88,30%	42,14%	42,64%
13	64,04%	86,88%	21,64%	54,24%
14	67,70%	88,65%	19,14%	28,65%
15	62,79%	88,60%	34,25%	5,11%
16	65,69%	82,61%	-9,96%	34,90%
17	63,33%	85,60%	57,17%	50,76%
18	68,15%	88,29%	49,63%	59,10%
19	57,87%	87,13%	4,27%	15,00%
20	63,60%	87,92%	77,03%	66,39%

<b>Base JAFFE</b>				
<b>k</b>	<b><i>k-means</i>+DE</b>	<b><i>k-means</i>+CP</b>	<b>EM+DE</b>	<b>EM+CP</b>
2	-20,65%	10,69%	71,85%	-18,95%
3	94,38%	30,74%	94,17%	52,50%
4	93,80%	38,31%	95,44%	56,85%
5	93,75%	53,06%	94,88%	65,85%
6	93,00%	43,85%	94,86%	55,87%
7	94,21%	46,55%	93,00%	56,31%
8	94,05%	46,41%	95,21%	58,99%
9	94,45%	52,46%	93,97%	39,38%
10	94,32%	49,40%	95,13%	60,79%

Tabela 35 – Ganho percentual de tempo de execução em detrimento da busca “gulosa”.

### 6.3 Análise do Tempo de Execução do Processo de Agrupamento

A realização do agrupamento em si impõe um custo adicional no processo de busca. As tabelas seguintes exibem os dados de tempo de execução na realização da técnica de agrupamento do algoritmo *K-Means* baseado em distância Euclidiana e correlação de *Pearson* (Tabela 36). A Tabela 37 exhibe os valores para o algoritmo EM considerando as mesmas bases de dados. Nas tabelas, a coluna  $k$  representa a quantidade de grupos definidos; as demais demonstram as bases de imagens, diferenciando a distância euclidiana (DE) e correlação de *Pearson* (CR) e seus respectivos tempos despendidos na formação dos agrupamentos.

Nesse momento, é importante observar que os tempos exibidos aqui se referem somente ao processo de criação dos agrupamentos a partir da base de dados. A formação dos grupos é realizada somente uma vez, e norteia o modelo organizacional dos dados utilizado durante a busca.

k	<i>faces94</i>		FEI		JAFPE	
	DE	CP	DE	CP	DE	CP
2	1,292859	1,544066	4,573258	3,195139	1,342440	0,301482
3	4,629746	0,199839	7,105105	0,745093	5,290576	0,198617
4	2,160768	0,155611	12,280117	0,711610	8,263207	0,622065
5	4,951613	0,156749	14,359360	0,871255	5,415190	1,090455
6	5,110105	0,156711	15,830856	0,979041	6,659165	0,920586
7	2,699936	0,150026	16,029220	4,348705	6,857956	1,011610
8	5,822350	0,131282	17,847381	1,279930	7,354018	0,953375
9	6,393857	0,172201	18,597275	1,505665	6,045172	0,301689
10	6,396125	0,520879	31,180027	4,288163	6,529109	0,274484
11	6,714342	0,623538	26,993749	1,533293		
12	6,873097	0,896765	27,565557	3,757175		
13	8,366476	0,999207	30,763238	3,424708		
14	7,734583	0,254425	35,017127	4,298375		
15	7,122915	0,316776	32,958950	1,447478		
16	7,400137	0,447353	35,780406	2,714629		
17	14,342655	0,337435	32,045273	3,150790		
18	13,314018	0,251578	32,013598	1,634405		
19	13,050677	0,297956	42,575986	6,591505		
20			46,873027	1,777051		
Tempo Médio	6,909792	0,422911	24,084249	2,582053	5,972981	0,630485

Tabela 36 – Tempo de execução dos agrupamentos com o algoritmo *K-Means* usando a distância euclidiana e correlação de *Pearson* nas três bases de dados completas.

A partir dos resultados exibidos, é possível notar que existe um custo computacional elevado no desenvolvimento dos agrupamentos, principalmente se comparado ao tempo gasto na execução da busca. No entanto, conforme mencionado anteriormente, o processo de agrupamento é realizado somente uma vez, com o objetivo de organizar os dados de forma que a busca possa ser melhor direcionada e, conseqüentemente, o tempo despendido na localização de uma imagem seja reduzido.

$k$	<i>faces94</i>	FEI	JAFPE
2	0,880844	1,604709	0,739168
3	1,545754	5,606733	1,075957
4	2,319974	3,627020	1,504349
5	2,668018	2,912407	1,705102
6	2,729148	3,525533	1,765937
7	3,066885	3,339248	2,169498
8	3,291490	3,197086	2,459247
9	3,572033	4,773977	2,876696
10	4,440964	4,177334	3,092767
11	4,278310	3,964048	
12	5,201217	4,788640	
13	5,140904	6,003963	
14	5,551577	6,070752	
15	5,612253	6,967645	
16	6,215345	6,950417	
17	7,249656	7,849279	
18	8,155720	8,869243	
19	8,313453	7,305769	
20		9,437939	
Tempo Médio	4,457419	5,085211	1,932080

Tabela 37 – Tempo de execução dos agrupamentos com o algoritmo EM nas três bases completas.

Esse processo é semelhante ao da geração dos modelos das redes neurais artificiais, que possuem um alto custo para serem treinadas, mas uma vez feito isso, o tempo gasto para a classificação é baixo. Assim sendo, o benefício obtido pela organização das imagens em grupos é notório, já que o tempo de busca (operação maior frequência de realização) é drasticamente reduzido.

Os resultados exibidos são referentes ao agrupamento com toda as bases de dados completas. Os testes da busca de imagens inexistentes nas bases seguem o mesmo procedimento, com exceção do fato de que cada imagem teste é retirada da base antes da realização do agrupamento. Assim, o tempo de agrupamento entre as duas abordagens (busca de imagem existente e inexistente na base) é praticamente o mesmo.

Podemos ainda fazer algumas análises referentes a esse tempo de execução dos agrupamentos:

1. diferença entre algoritmos de agrupamento;
2. diferença entre medidas de proximidade;
3. diferença entre o número de grupos usado durante o agrupamento.

Quando comparamos as técnicas de agrupamento, notamos que o algoritmo EM tem desempenho semelhante ao *k-means*, na ocasião em que esse último usa a distância euclidiana no processamento do agrupamento. A exceção fica por conta da base FEI, onde

o tempo gasto pelo EM é bem inferior ao *k-means* (*k-means* com distância euclidiana obteve média de 24,084249 segundos na base FEI; já o EM, 5,085211 segundos). Por outro lado, quando o *k-means* usa a correlação como medida de proximidade, notamos uma grande diminuição no seu custo de execução. Sendo assim, se a medida escolhida para proximidade for correlação, o *k-means* tende a ser uma melhor opção no cenário dos experimentos realizados. O algoritmo EM surge como uma boa alternativa para o *k-means* quando a distância euclidiana precisa ser usada.

Um outro ponto importante a ser destacado é o fato de o agrupamento utilizando a correlação de *Pearson* ser bem mais eficiente, no que tange a otimização temporal, que o mesmo processo usando a distância euclidiana (no contexto deste trabalho, somente o algoritmo *k-means* foi usado com essa medida de similaridade). A partir disso, uma hipótese para essa diferença de desempenho seria a convergência mais rápida do algoritmo, provocada por um número menor de realocações dos grupos por parte das instâncias.

Como última análise a ser considerada, apresenta-se o número de grupos usados pelos algoritmos de agrupamento. É possível notar que, em geral, a medida que o número de grupos aumenta, o tempo levado para gerar o agrupamento também cresce. Esse comportamento é esperado, uma vez que a quantidade de comparações entre os objetos do conjunto de dados e os centros dos grupos se torna maior. Além disso, é possível que alguns grupos se tornem próximos, levando as instâncias a trocarem de conjuntos mais vezes durante a geração do modelo.

## 6.4 Considerações finais

O presente capítulo versou sobre os resultados obtidos através dos experimentos propostos em suas diferentes nuances. São apresentadas tabelas e aferições de cada um dos testes realizados, contendo valores relacionados ao tempo computacional despendidos nas execuções, bem como valores ligados a avaliações mais abrangentes de acurácia. O objetivo principal dessa seção é expôr os dados obtidos, de forma mais direta e explicativa, possibilitando assim um maior entendimento acerca do projeto, proporcionar a análise final.

Através dos resultados, foi possível concluir que o uso de algoritmos de agrupamento traz consigo benefícios em termos de tempo de busca de imagens faciais em bases de dados. Em todos os casos, esse tempo foi reduzido, seja em uma escala maior (quase 100 vezes mais rápido que o tempo da busca direta) ou menor (redução de três vezes do tempo). Além disso, a redução também no desvio padrão do tempo médio de busca mostra que o direcionamento da pesquisa é eficiente e impede que seja desperdiçado tempo em regiões

desinteressantes no espaço de busca.

Por fim, em alguns poucos casos, houve uma queda na acurácia da busca, provocando erros na identificação de pessoas. Esse comportamento foi notado para as versões que usaram a correlação como medida de similaridade entre imagens. O ocorrido pode indicar que tal medida é hipoteticamente ineficiente na captura de diferenças entre rostos distintos ou expressões distintas de um mesmo indivíduo.

O próximo capítulo apresenta as considerações finais, que englobam um resumo geral do projeto, bem como a finalização descritiva, trazendo consigo as possibilidades de aprimoramentos futuros.

## 7 Considerações Finais

Este trabalho apresentou um estudo relacionado ao uso de técnicas de agrupamento para auxiliar o processo de busca de imagens em bases de dados. A ideia central consiste em organizar os dados em grupos de forma que a busca por imagens possa ser direcionada e, conseqüentemente, obtenha ganhos computacionais em termos de tempo de execução.

O contexto escolhido para realizar os testes foi o de reconhecimento facial. Esse tema, além de trazer desafios relativos a complexidade da tarefa, possui alta relevância no desenvolvimento de aplicações inteligentes para os mais diversos fins. Entre as ações envolvendo o reconhecimento facial, merece destaque a área de segurança pública, cuja utilização se torna cada vez mais necessária, principalmente no âmbito brasileiro.

Com base nisso, foram realizados experimentos controlados objetivando a observação, diante da mudança de parâmetros e cenários diversos, do comportamento do processo de busca e recuperação de imagens com auxílio de agrupamentos, levando em consideração dois critérios principais: o tempo de busca e a acurácia do sistema. Em termos práticos, o tempo de processamento da busca representa um desafio, principalmente em relação a grandes bases de dados, enquanto que a precisão do sistema não pode ser prejudicada em detrimento de melhorias no custo computacional.

O estudo envolveu a aplicação dos algoritmos de agrupamento *k-means* e EM, ambas estratégias “clássicas” e com boa exploração no meio acadêmico. Vale lembrar que o desenvolvimento deste trabalho intenta a comprovação da otimização do uso de agrupamentos na busca em si, e não um *benchmarking* entre diversas técnicas para execução da mesma tarefa; sendo assim, a restrição do escopo a essas duas técnicas mostra-se satisfatória. Além disso, a distância euclidiana e correlação de *Pearson* são utilizadas como medidas de similaridade/dissimilaridade, mantendo a relação de medidas tradicionais para o estudo, além da diversidade de funcionamento que ambas possuem.

Com base nos resultados, foi possível concluir que o uso de técnicas de agrupamento traz vantagens nos dois aspectos analisados (acurácia e tempo de busca). O comportamento médio da busca quando os dados são organizados em grupos é mais eficiente, uma vez que é possível direcionar de forma sistemática o processo de comparação de similaridade. Quando confrontado com o procedimento tradicional de busca direta, onde a pesquisa é feita de maneira iterativa, o tempo médio para encontrar indivíduos na base se torna elevado, já que a localização da imagem a ser buscada na base de dados influencia diretamente no tempo. Por outro lado, quando organizamos os dados em grupos, temos duas vantagens claras: a

possibilidade de usar uma imagem (o ponto central do grupo) como o representante do grupo, e o fato de as imagens dentro do conjunto tenderem a ser similares. Esses dois fatores trazem subsídios para direcionar a busca e indicam, primeiramente, em que ponto pode ser iniciada (e continuada, de forma sistemática, de acordo com a proximidade) e, posteriormente, a busca local dentro do grupo é agilizada, uma vez que os itens do grupo são semelhantes.

Vale salientar que esse estudo abordou apenas três bases de dados. Cada uma dessas bases teve de ser estudada para que os parâmetros de busca fossem ajustados, na averiguação pelos melhores resultados possíveis. A análise desse processo em uma base maior pode se tornar mais complexa; todavia, na posse das informações necessárias, a obtenção/definição desses parâmetros é simplificada.

Além disso, é esperado que o aumento do número de imagens no conjunto de dados seja diretamente proporcional ao ganho de tempo na busca, se comparado ao processo trivial de buscas. Em alguns casos, como demonstrados através dos resultados, a busca com auxílio de agrupamento pode ser quase 100 vezes mais rápida que o processo tradicional. A validação dos resultados, provida pelo teste de *Wilcoxon* pareado, atesta que o algoritmo *k-means*, em destaque, realmente otimiza a busca e recuperação de imagens, em ambos os cenários (busca por imagem existente na base, e pela imagem inexistente na base).

O custo computacional inicial para se construir os agrupamentos é relativamente alto se comparado ao tempo de busca; entretanto, é compensatório, uma vez que o processo é realizado uma única vez (ou eventualmente, com a atualização massiva da base de dados) no início do processo. Esse tipo de custo benefício é o mesmo encontrado em técnicas de Aprendizado de Máquina, amplamente utilizadas na literatura, como as Redes Neurais Artificiais.

## 7.1 Trabalhos Futuros

Por se tratar de um estudo controlado, os resultados também são restritos ao escopo dos experimentos. Dessa forma, uma série de aprofundamentos podem ainda ser realizados com o objetivo de investigar o uso de agrupamentos em processos de busca de imagens. Alguns podem ser destacados a seguir:

- utilização de outros algoritmos de agrupamento para organização dos dados;
- utilização de bases de dados com mais imagens;
- utilização de outras codificações e métodos para extração de características de imagens;

- a possibilidade de combinar métodos de agrupamentos para produzir resultados mais robustos;
- uso de redução de dimensionalidade para diminuir a complexidade na extração de características;
- exploração outras codificações de cores para as imagens;
- utilização de diferentes tratamento e/ou transformações nas imagens.

Por fim, este trabalho deixa como contribuição a conclusão de que é possível otimizar o processo de buscas de imagens com o auxílio de técnicas de agrupamento, permeando opções para estudos seguintes, relativos a aspectos transversais na área de Aprendizado de Máquina e Visão Computacional.

## Referências

- Aggarwal, C. C. e Reddy, C. K. (2013). *Data Clustering: Algorithms and Applications*. Chapman & Hall/CRC, 1stedição. Citado 2 vezes nas páginas 10 e 24.
- Alattab, A. A. e Kareem, S. A. (2012). Efficient method of visual feature extraction for facial image detection and retrieval. In *2012 Fourth International Conference on Computational Intelligence, Modelling and Simulation*, páginas 220–225. Citado 3 vezes nas páginas 16, 33 e 37.
- Araújo, D. S. A. (2008). Algoritmos de agrupamento aplicados a dados de expressão gênica de câncer: Um estudo comparativo. Dissertação de Mestrado, UFRN. Citado na página 25.
- Belhumeur, P. N., Chen, D., Feiner, S., Jacobs, D. W., Kress, W. J., Ling, H., Lopez, I., Ramamoorthi, R., Sheorey, S., White, S., e Zhang, L. (2008). Searching the world's herbaria: A system for visual identification of plant species. *European Conference on Computer Vision*, página 116–129. Citado na página 29.
- Castiglione, A., Cattaneo, G., e Santis, A. D. (2011). A forensic analysis of images on online social networks. In *Intelligent Networking and Collaborative Systems (INCoS), 2011 Third International Conference on*, páginas 679–684. Citado na página 16.
- Chowdhry, D. A., Hussain, A., Rehman, M. Z. U., Ahmad, F., Ahmad, A., e Pervaiz, M. (2013). Smart security system for sensitive area using face recognition. In *2013 IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (CSUDET)*, páginas 11–14. Citado na página 17.
- Civicioglu, P., Atasever, U. H., Ozkan, C., Besdok, E., Karkinli, A. E., e Kesikoglu, A. (2014). Performance comparison of evolutionary algorithms for image clustering. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, páginas 71–74. Citado na página 35.
- Cooper, M., Foote, J., Girgensohn, A., e Wilcox, L. (2003). Temporal event clustering for digital photo collections. Citado na página 29.
- Dalal, N. e Triggs, B. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, páginas 886–893. Citado 4 vezes nas páginas 10, 27, 28 e 31.

- Dempster, A. P., Laird, N. M., e Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B*, 39(1):1–38. Citado na página 24.
- Desai, R. e Sonawane, B. (2017). *Gist, HOG, and DWT-Based Content-Based Image Retrieval for Facial Images*, páginas 297–307. Springer Singapore. Citado 3 vezes nas páginas 16, 34 e 37.
- Déniz, O., Bueno, G., Salido, J., e la Torre, F. D. (2011). Face recognition using histograms of oriented gradients. *Pattern Recognition Letters*, 32(12):1598 – 1603. Citado na página 32.
- Everitt, B. S., Landau, S., Leese, M., e Stahl, D. (2011). *Cluster Analysis*, volume 5. Wiley. Citado 3 vezes nas páginas 16, 20 e 25.
- Faceli, K., Carvalho, A., Lorena, A. C., e Gama, J. (2014). *Inteligência Artificial: uma abordagem de Aprendizado de Máquina*. LTC. Citado 2 vezes nas páginas 17 e 23.
- Faktor, A. e Irani, M. (2012). “clustering by composition” - unsupervised discovery of image categories. In *Computer Vision - ECCV 2012*, volume 7578 of *Lecture Notes in Computer Science*, páginas 474–487. Springer Berlin Heidelberg. Citado 2 vezes nas páginas 29 e 35.
- Frery, A. e Perciano, T. (2013). Image data formats and color representation. In *Introduction to Image Processing Using R*, SpringerBriefs in Computer Science, páginas 21–29. Springer London. Citado na página 26.
- Goldberger, J., Greenspan, H., e Gordon, S. (2002). Unsupervised image clustering using the information bottleneck method. Citado 2 vezes nas páginas 29 e 35.
- Gonzalez, R. e Woods, R. (2000). *Processamento de Imagens Digitais*. Edgard Blücher Ltda. Citado na página 26.
- Green, P. E., Carmone, F. J., e Smith, M. (1989). *Multidimensional Scaling*, chapter 5. Addison Wesley. Citado na página 25.
- Gritti, T., Shan, C., Jeanne, V., e Braspenning, R. (2008). Local features based facial expression recognition with face registration errors. In *Automatic Face Gesture Recognition, 2008. FG '08. 8th IEEE International Conference on*, páginas 1–8. Citado na página 27.
- Hair, J. F., Anderson, R. E., e Tatham, R. L. (2005). *Análise Multivariada de Dados*. Bookman. Citado na página 48.

- Hemmati, A. e Chung, K. S. K. (2014). Associations between personal social network properties and mental health in cancer care. In *Advances in Social Networks Analysis and Mining (ASONAM), 2014 IEEE/ACM International Conference on*, páginas 828–835. Citado na página 16.
- Hond, D. e Spacek, L. (1997). Distinctive descriptions for face processing. In *Proceedings of the 8th British Machine Vision Conference BMVC97*, páginas 320–329. Citado na página 46.
- Huang, W., Zhang, P., e Shen, M. (2014). Medical social media analytics via ranking and big learning: An image-based disease prediction study. In *Security, Pattern Analysis, and Cybernetics (SPAC), 2014 International Conference on*, páginas 389–394. Citado na página 16.
- Jain, A. K. e Dubes, R. C. (1988). *Algorithms for Clustering Data*. Prentice-Hall, Inc. Citado 2 vezes nas páginas 48 e 63.
- Jiang, Z., Joshi, A., Krishnapuram, R., e Yi, L. (2000). Retriever: Improving Web Search Engine Results Using Clustering. Relatório técnico, University of Maryland Baltimore County. Citado 2 vezes nas páginas 31 e 36.
- K, J. G. L. e Reshmi, K. (2015). Face recognition in surveillance system. In *Innovations in Information, Embedded and Communication Systems (ICIIECS), 2015 International Conference on*, páginas 1–5. Citado na página 17.
- Klimova, B., Poulouva, P., e Sucharda, O. (2014). Social networks and their use in education. In *Emerging eLearning Technologies and Applications (ICETA), 2014 IEEE 12th International Conference on*, páginas 233–237. Citado na página 16.
- Linden, R. (2009). Análise de agrupamento. In *Revista de Sistemas de Informação da FSMA*, number 4, páginas 18–36. Citado 2 vezes nas páginas 10 e 23.
- Lyons, M. J., Akemastu, S., Kamachi, M., e Gyoba, J. (1998). Coding facial expressions with gabor wavelets. In *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, páginas 200–205. Citado na página 47.
- Martin, D., Fowlkes, C., Tal, D., e Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int'l Conf. Computer Vision*, volume 2, páginas 416–423. Citado na página 45.
- Mooi, E. e Sarstedt, M. (2011). Cluster analysis. In *A Concise Guide to Market Research*, páginas 237–284. Springer Berlin Heidelberg. Citado na página 20.

- Nixon, M. e Aguado, A. S. (2012). *Feature Extraction Image Processing for Computer Vision*. Elsevier Science, 3ed.edição. Citado 2 vezes nas páginas 10 e 28.
- Obulesu, A., Kiran, J. S., e Kumar, V. V. (2015). Facial image retrieval based on local and regional features. In *2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, páginas 841–846. Citado 3 vezes nas páginas 16, 33 e 37.
- Pang, Y., Yuan, Y., Li, X., e Pan, J. (2011). Efficient {HOG} human detection. *Signal Processing*, 91(4):773 – 781. <http://www.sciencedirect.com/science/article/pii/S0165168410003476>. Citado na página 32.
- Petcu, D., Zaharie, D., Panica, S., Hussein, A. S., Sayed, A., e El-Shishiny, H. (2011). Fuzzy clustering of large satellite images using high performance computing. In *THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING*, volume 8183. Citado na página 35.
- Poulova, P. e Simonova, I. (2014). Social networks in higher education. In *Emerging eLearning Technologies and Applications (ICETA), 2014 IEEE 12th International Conference on*, páginas 443–446. Citado na página 16.
- Saffer, K. M. (2011). Comparison among image clustering algorithms. *Diyala Journal For Pure Science*, 7(3):242–258. Citado 2 vezes nas páginas 16 e 33.
- Sammut, C. e Webb, G. I. (2011). *Encyclopedia of Machine Learning*. Springer Publishing Company, Incorporated, 1stedição. Citado na página 24.
- Satari, B. S., Rahman, N. A. A., e Abidin, Z. M. Z. (2014). Face recognition for security efficiency in managing and monitoring visitors of an organization. In *Biometrics and Security Technologies (ISBAST), 2014 International Symposium on*, páginas 95–101. Citado na página 17.
- Shapiro, L. G. e Stockman, L. G. (2000). *Computer Vision*. Prentice Hall. Citado na página 16.
- Shen, H., Liang, X., e Wang, M. (2015). Emergency decision support architectures for bus hijacking based on massive image anomaly detection in social networks. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, páginas 864–869. Citado na página 16.

- Shu, C., Ding, X., e Fang, C. (2011). Histogram of the oriented gradient for face recognition. *Tsinghua Science & Technology*, 16(2):216 – 224. Acessado em: 28/07/2016. Citado na página 32.
- Thomaz, C. E. e Giraldi, G. A. (2010). A new ranking method for principal components analysis and its application to face image analysis. *Image and Vision Computing*, 28(6):902–913. Citado 6 vezes nas páginas 10, 42, 43, 44, 46 e 47.
- Wang, Y. e Kitsuregawa, M. (2001). Use link-based clustering to improve web search results. In *Web Information Systems Engineering, 2001. Proceedings of the Second International Conference on*, volume 1, páginas 115–124 vol.1. Citado 2 vezes nas páginas 31 e 36.
- Wu, X., Zhu, X., Wu, G. Q., e Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, 26(1):97–107. Citado na página 35.
- Xu, R. e Wunsch II, D. (2007). Computational intelligence in clustering algorithms, with applications. In *Algorithms for Approximation*, páginas 31–50. Citado 4 vezes nas páginas 10, 21, 22 e 31.
- Zhang, S., He, O., e Shi, Z. (2007). Efficient image clustering using a new image distance. In *Machine Learning and Cybernetics, 2007 International Conference on*, volume 3, páginas 1601–1605. Citado na página 35.