

CENTRO UNIVERSITÁRIO DA FEI

Vagner do Amaral

**Extração e Comparação de Características Locais e
Globais para o Reconhecimento Automático de
Imagens de Faces**

São Bernardo do Campo
2011

Vagner do Amaral

**Extração e Comparação de Características Locais e
Globais para o Reconhecimento Automático de
Imagens de Faces**

Dissertação de Mestrado apresentada ao Centro
Universitário da FEI para obtenção do título de
Mestre em Engenharia Elétrica, orientada pelo
Prof. Dr. Carlos Eduardo Thomaz.

Amaral, Vagner do.

Extração e comparação de características locais e globais para o reconhecimento automático de imagens de faces / Vagner do Amaral. São Bernardo do Campo, 2011.

78 f. : il.

Dissertação (Mestrado) - Centro Universitário da FEI.

Orientador: Prof. Dr. Carlos Eduardo Thomaz

1. Reconhecimento de faces. 2. Análise de componentes principais. 3. Padrão binário local . I. Thomaz, Carlos Eduardo, orient. II. Título.

CDU 681.32



Centro Universitário da **FEI**

APRESENTAÇÃO DE DISSERTAÇÃO ATA DA BANCA JULGADORA

PGE- 10

Programa de Mestrado de Engenharia Elétrica

Aluno: Vagner do Amaral

Matrícula: 1091172

Título do Trabalho: **Extração e comparação de características locais e globais para o reconhecimento automático de imagens de faces.**

Área de Concentração: Inteligência Artificial Aplicada à Automação

Orientador: Prof. Dr. Carlos Eduardo Thomaz

ORIGINAL ASSINADA

Data da realização da defesa: 28 / outubro / 2011

A Banca Julgadora abaixo-assinada atribuiu ao candidato o seguinte:

APROVADO

REPROVADO

São Bernardo do Campo, 28 / 10 / 2011.

MEMBROS DA BANCA JULGADORA

Prof. Dr. Carlos Eduardo Thomaz

Ass.: _____

Prof. Dr. Marko Ackermann

Ass.: _____

Prof.^a Dr.^a Gilka Jorge Figaro Gattas

Ass.: _____

VERSÃO FINAL DA DISSERTAÇÃO

**ENDOSSO DO ORIENTADOR APÓS A INCLUSÃO DAS
RECOMENDAÇÕES DA BANCA EXAMINADORA**

Aprovação do Coordenador do Programa de Pós-graduação

Prof. Dr. Carlos Eduardo Thomaz

A minha família pelo apoio incondicional, a
minha esposa pela paciência e compreensão e
aos meus amigos pela motivação e incentivo.

AGRADECIMENTOS

Gostaria de agradecer ao meu orientador, Prof^o Dr. Carlos Eduardo Thomaz, inicialmente pela confiança em mim depositada e pelos ensinamentos, conselhos, críticas e opiniões que foram fundamentais para a execução deste trabalho. Porém, devo agradecê-lo principalmente pela paciência que me permitiu evoluir não apenas no âmbito acadêmico mas também no pessoal.

Aos professores da graduação Eduardo Rosalém Marcelino, Luciano Gaspar, Adão Alves Rodrigues e Paulo Marcotti, por me incentivarem a ingressar no ambiente acadêmico e aos professores Flávio Tonindandel, Paulo Sérgio Rodriguez, Paulo Eduardo Santos e Reinaldo Bianchi por compartilharem seu conhecimento durante a realização dos créditos do mestrado. Aos colegas de turma e de laboratório pela colaboração e companheirismo que abrandaram as dificuldades ao longo do curso.

Aos membros da banca Prof^a Dra. Gilka Jorge Figaro Gattas e Prof^o Dr. Marko Ackermann, cujas sugestões, questionamentos e críticas durante a qualificação contribuíram muito para melhorar a qualidade deste trabalho.

À equipe do projeto Caminho de Volta pela oportunidade única de integrar e contribuir com um grupo multidisciplinar composto por profissionais dedicados a uma causa nobre, a busca de crianças e adolescentes desaparecidos.

À FAPESP pela possibilidade de trabalhar com pesquisa científica em um projeto de relevância social e ao Centro Universitário FEI por todas as oportunidades, acadêmicas e profissionais, oferecidas a mim. Aos funcionários da secretaria do mestrado e do CGI, que acompanharam e contribuíram, direta ou indiretamente, com a conclusão deste trabalho.

E finalmente agradeço à minha família pelo apoio incondicional e por todos os esforços realizados para que eu pudesse alcançar meus objetivos e a minha querida esposa pela paciência e compreensão nos momentos de ausência. A todos, o meu muito obrigado!

"Não sabendo que era impossível, foi lá e fez."

Jean Cocteau

RESUMO

Pesquisas relacionadas ao reconhecimento de faces sempre receberam ênfase devido a sua aplicabilidade e abrangência. Inicialmente, em razão das limitações tecnológicas, os trabalhos desenvolvidos nesta área empregaram técnicas baseadas em características geométricas (locais). No entanto, com o avanço da capacidade computacional e a exploração de métodos estatísticos, a utilização de informações holísticas (globais) se tornou mais popular e fomentou a criação de diversas metodologias. Conseqüentemente, surgiram estudos com a intenção de comparar o desempenho das técnicas propostas e os resultados preliminares favoreceram a implementação de modelos holísticos. No entanto, trabalhos subsequentes concluíram que ambas as abordagens, locais e globais, são úteis de acordo com o contexto da aplicação. Este trabalho propõe e implementa uma comparação entre as técnicas Local Binary Pattern (LBP) e Principal Components Analysis (PCA) com o propósito de destacar características particulares de cada método no processo de reconhecimento automático de imagens de faces, utilizando apenas um exemplo de treinamento por pessoa. Os resultados demonstram um bom desempenho da técnica LBP em situações específicas, onde a base de treinamento é formada por indivíduos em pose frontal com expressão facial neutra. Contudo, empregando-se uma base de treinamento com poses e expressões faciais heterogêneas, em um teste leave-one-out, o PCA é mais eficiente. Estes resultados enfatizam conclusões prévias sobre a aplicação de abordagens adequadas ao contexto do problema tratado e permitem uma análise acerca das vantagens e desvantagens de cada método quando utiliza-se imagens de faces pré-processadas e normalizadas espacialmente.

Palavras-chave: Reconhecimento de faces; Análise de Componentes Principais; Padrão Binário Local.

ABSTRACT

Research on face recognition has always received attention due to its intrinsic cognitive scope and applicability. In the beginning, owing to technological limitations, works in this area have used techniques based on geometric (or local) features. However, with the increasing computational power and the exploration of statistical methods, the use of holistic (or global) information has become more popular and stimulated the development of different methodologies. Consequently, new studies have emerged with the intention to comparing the performance of the proposed techniques and the experimental results favoured the implementation of holistic models. However, more recent studies have highlighted that both approaches, local and global, are useful and depend on the application context. This dissertation proposes and implements a comparison between the techniques Local Binary Pattern (LBP) and Principal Components Analysis (PCA), in order to understand and enhance particular characteristics of each popular method in the process of automatic recognition of face images using only one training example per person. Our results have showed a good performance of the LBP technique in specific situations where the training set is composed of individuals in frontal pose with neutral facial expression only. However, using a training set with distinct poses and facial expressions, using a leave-one-out testing protocol, the PCA is more efficient. These results emphasize some previous conclusions available in the literature based on the use of local or global approaches depending on the application, but has also provided a detailed pattern recognition analysis about the advantages and disadvantages of each method when using pre-processed and previously spatially normalized face images

Keywords: Face recognition; Principal Components Analysis; Local Binary Pattern.

LISTA DE TABELAS

1.1	Aplicações típicas de reconhecimento facial.	2
-----	------------------------------------------------------	---

LISTA DE ILUSTRAÇÕES

3.1	Etapas presentes em um processo iterativo de reconhecimento de padrões.	13
3.2	Normalização de imagens frontais de face.	14
3.3	Rotação de imagem de face com base na posição das pupilas.	15
3.4	Distância interpupilar original (a) e escala normalizada para 100 pixels (b).	16
3.5	Relação das medidas de recorte, onde $x = 100$ pixels e $\kappa = 1,618$	17
3.6	Detalhe da etapa de suavização de contorno.	18
3.7	Imagem com histograma original e equalizada.	19
3.8	Imagens médias obtidas em ambientes controlados (a) e não-controlados (b), antes e depois da normalização.	20
3.9	Abordagens de mapeamento local para extração de características.	21
3.10	Extração global de características: matriz bi-dimensional de intensidade de pixels em escala de cinza.	22
3.11	Representação do espaço multidimensional de faces.	23
3.12	Imagem de face média.	23
3.13	Distribuição bidimensional de amostras e navegação no eixo de maior variância no subespaço gerado pelo PCA	25
3.14	Operação básica do LBP, com valor binário resultante 10111000 e decimal 184.	26
3.15	Vizinhança quadrada (8, 1), circular (8, 1), quadrada (8, 2) e circular (16, 2). O valor dos pontos descentralizados é determinado pela interpolação bilinear dos pixels vizinhos.	27
3.16	Imagem de face processada pelo LBP, dividida em grades 3×3 , 5×5 e 7×7 e regiões específicas dos olhos, boca e nariz.	27
3.17	Descritores de textura independentes, concatenados para compôr o histograma global.	28
3.18	Filtragem dos resultados de acordo com o contexto da busca.	29
4.1	Imagens da base FEI, contendo variações de expressão, pose e iluminação.	32

4.2	Imagens da base FERET, contendo uma imagem neutra e uma sorrindo por pessoa.	32
4.3	Imagens da base JAFFE, contendo uma imagem neutra e seis variações de expressão por pessoa	32
4.4	Imagens normalizadas das bases FEI, FERET e JAFFE.	33
4.5	De cima para baixo, deslocamento da imagem média sobre o eixo de maior variância gerado pelo PCA sobre a base FEI, FERET e JAFFE.	34
4.6	Da esquerda para direita, grades 3x3, 7x7 e 7x7 ponderada. As regiões preta, cinza escuro, cinza claro e branco, receberam os pesos 0, 1, 2 e 4, respectivamente.	35
4.7	Padrões de regiões faciais empregadas no reconhecimento com LBP.	35
4.8	Imagens pré-processadas pelo LBP com raios de vizinhança 1 e 2.	36
4.9	Comparação de grades na base FEI homogênea.	37
4.10	Comparação de grades na base FEI heterogênea.	37
4.11	Comparação de grades na base FERET homogênea.	38
4.12	Comparação de grades na base FERET heterogênea.	38
4.13	Comparação de grades na base JAFFE homogênea.	39
4.14	Comparação de grades na base JAFFE heterogênea.	39
4.15	Comparação de regiões faciais na base FEI homogênea.	41
4.16	Comparação de regiões faciais na base FEI heterogênea.	41
4.17	Comparação de regiões faciais na base FERET homogênea.	42
4.18	Comparação de regiões faciais na base FERET heterogênea.	42
4.19	Comparação de regiões faciais na base JAFFE homogênea.	43
4.20	Comparação de regiões faciais na base JAFFE heterogênea.	43
4.21	Comparação de combinações das regiões faciais na base FEI homogênea.	44
4.22	Comparação de combinações das regiões faciais na base FEI heterogênea.	44
4.23	Comparação de combinações das regiões faciais na base FERET homogênea.	45
4.24	Comparação de combinações das regiões faciais na base FERET heterogênea.	45
4.25	Comparação de combinações das regiões faciais na base JAFFE homogênea.	46

4.26	Comparação de combinações das regiões faciais na base JAFFE heterogênea.	46
4.27	Comparação de métricas na base FEI homogênea.	48
4.28	Comparação de métricas na base FEI heterogênea.	48
4.29	Comparação de métricas na base FERET homogênea.	49
4.30	Comparação de métricas na base FERET heterogênea.	49
4.31	Comparação de métricas na base JAFFE homogênea.	50
4.32	Comparação de métricas na base JAFFE heterogênea.	50
4.33	Comparação de métricas na base pré-processada FEI homogênea.	51
4.34	Comparação de métricas na base pré-processada FEI heterogênea.	51
4.35	Comparação de métricas na base pré-processada FERET homogênea.	52
4.36	Comparação de métricas na base pré-processada FERET heterogênea.	52
4.37	Comparação de métricas na base pré-processada JAFFE homogênea.	53
4.38	Comparação de métricas na base pré-processada JAFFE heterogênea.	53
4.39	Comparação entre as técnicas LBP e PCA na base FEI homogênea.	54
4.40	Comparação entre as técnicas LBP e PCA na base FEI heterogênea.	54
4.41	Comparação entre as técnicas LBP e PCA na base FERET homogênea.	55
4.42	Comparação entre as técnicas LBP e PCA na base FERET heterogênea.	55
4.43	Comparação entre as técnicas LBP e PCA na base JAFFE homogênea.	56
4.44	Comparação entre as técnicas LBP e PCA na base JAFFE heterogênea.	56

SUMÁRIO

1	Introdução	1
1.1	Motivação	2
1.2	Objetivo	3
1.3	Organização do Trabalho	3
2	Revisão Bibliográfica	4
2.1	Considerações Complementares	9
2.2	Desafios atuais	10
3	Metodologia	12
3.1	Aquisição de amostras	12
3.2	Pré-processamento	12
3.2.1	Normalização	14
3.2.2	Extração de características	20
3.3	Classificação	27
3.4	Pós-processamento	29
3.5	Considerações Complementares	30
4	Experimentos e Resultados	31
4.1	Base de Imagens de Face	31
4.1.1	Normalização das amostras	33
4.2	Projeção de auto-faces	33
4.3	Reconhecimento automático de imagens de faces	34
4.3.1	Experimentos com LBP	35
4.3.2	Experimentos com PCA	40
4.3.3	Comparação entre LBP e PCA	47

5 Conclusão	57
5.1 Trabalhos Futuros	58
REFERÊNCIAS	60

1 INTRODUÇÃO

Pesquisas relacionadas ao reconhecimento de padrões sempre receberam ênfase devido a sua aplicabilidade e abrangência (KITANI, 2007). Os estudos realizados nesta linha de pesquisa obtiveram resultados bem sucedidos em áreas da ciência como engenharia, medicina e economia. Entretanto, aplicações envolvendo processamento de imagens e visão computacional ainda proporcionam vários desafios à comunidade científica. Neste contexto, surgiram trabalhos com o intuito de criar técnicas capazes de atuar na identificação de pessoas através de padrões biométricos e consequentemente desenvolveram-se sistemas robustos utilizando características específicas, tais como: voz, íris, geometria da mão e impressão digital (THOMAZ, 1999). A demanda por métodos menos invasivos motivou a utilização da face como parâmetro de reconhecimento. No entanto, esta atividade é extremamente complexa para sistemas digitais e tem atraído a atenção de diversos grupos de pesquisa em todo o mundo (KANADE, 1973; ZHAO et al., 2003).

Os primeiros experimentos relacionados ao reconhecimento de faces tiveram início na década de 70 e sua expansão ocorreu de forma paralela ao avanço da capacidade computacional (DELAC; GRGIC; BARTLETT, 2008). O trabalho pioneiro de Kanade (KANADE, 1973) promoveu discussões nesta área e desde então inúmeras propostas foram apresentadas através de abordagens distintas. Tendo em vista este cenário, Brunelli e Poggio (BRUNELLI; POGGIO, 1993) realizaram uma análise entre técnicas baseadas em informações locais e globais, obtendo bons resultados ao utilizar apenas imagens frontais de faces. Outras pesquisas também apresentaram resultados promissores em ambientes controlados (ZHAO et al., 2003; PANTIC; ROTHKRANTZ, 2003) e atualmente novas técnicas de processamento têm contribuído para a resolução de problemas comuns a aplicações de reconhecimento facial em ambientes reais (DENG et al., 2010; TAN et al., 2006; MATURANA; MERY; SOTO, 2009), como será descrito com mais detalhes no capítulo seguinte.

Entre os seres humanos, o reconhecimento de um indivíduo com base em uma imagem de face é uma tarefa simples e possui grande relevância nas relações interpessoais (DUDA; HART; STORK, 2001). Portanto, a compreensão dos processos cognitivos, aliada ao desenvolvimento de arcabouços computacionais, fornece ótimas perspectivas à resolução de problemas de identificação baseados em biometria facial.

1.1 Motivação

Há muitos anos os desafios presentes nesta área fomentam o surgimento de conferências, publicações e protocolos específicos, destinados a padronização de métricas para avaliação das técnicas de reconhecimento facial. Segundo Zhao et al. (ZHAO et al., 2003), existem pelo menos duas razões para esta tendência. A primeira é a disponibilidade de tecnologias viáveis após 40 anos de pesquisa, e a segunda é a vasta gama de aplicações, vide Tabela 1.1.

Áreas	Aplicações
Entretenimento	Jogos digitais, realidade virtual, simulações, treinamentos, interação homem-máquina.
Segurança da Informação	Bancos de dados, terminais bancários, criptografia de arquivos e controle de acesso.
Aplicações legais e fiscalização	Vigilância em vídeo, monitoração de furto, rastreamento de suspeitos e investigação.
Cartões Inteligentes	Habilitação, RG, CPF, passaporte, título de eleitor e controle de fraudes.
Controle de Acesso	Simplificação de acesso predial e veicular.

Tabela 1.1: Aplicações típicas de reconhecimento facial.

Como pode-se observar na Tabela 1.1, as técnicas de reconhecimento facial têm sido utilizadas em diversas aplicações comerciais, em áreas distintas. No entanto, poucos esforços tem sido empregados em atividades de grande relevância social como por exemplo o reconhecimento de crianças e adolescentes desaparecidos.

Segundo dados da Associação Brasileira de Busca e Defesa das Crianças Desaparecidas e da CPI das Crianças e Adolescentes Desaparecidos, cerca de 40 mil crianças e adolescentes desaparecem no Brasil todo ano (BRASIL, ; DESAPARECIDOS,). Contudo, é impossível afirmar a veracidade desses dados devido a falta de políticas públicas que viabilizem o acompanhamento das ocorrências em todo o país, de forma integrada.

Neste contexto, o desaparecimento infanto-juvenil merece um estudo metódico, pois o tempo de busca pode se prolongar e as alterações fisionômicas geradas pelo envelhecimento dificultam o reconhecimento facial. Portanto, algumas iniciativas, como o projeto Caminho de Volta (GATTÁS; FIGARO-GARCIA, 2007), promovem a coleta de material genético de familiares para eventuais comparações. Embora o DNA seja o método mais seguro de identificação, a divulgação de fotos ainda é o principal meio de localização. Além disso, é comum o deslocamento de pessoas desaparecidas para outros lugares e o processo de coleta, transporte e análise das amostras biológicas é muito mais complexo e demorado que as etapas envolvidas no reconhecimento de imagens de face.

A ausência de uma legislação que exija, desde a infância, o registro de indivíduos nascidos no Brasil, torna a identificação dependente das fotografias fornecidas pelos familiares. Porém, tais imagens frequentemente são geradas sem nenhum critério de aquisição e costumam apresentar diversas variações de pose, tamanho e iluminação. Em muitos casos esta contribuição se limita a apenas uma única imagem, aumentando ainda mais o grau de dificuldade para o desenvolvimento de arcabouços computacionais capazes de auxiliar no processo de reconhecimento. Portanto, a compreensão e aprimoramento das técnicas de reconhecimento de padrões podem auxiliar o processo de triagem de crianças e adolescentes em situação de filiação estabelecida ou não.

1.2 Objetivo

Este trabalho propõe uma comparação entre métodos de extração e comparação de características faciais locais e globais, predominantes na literatura afim. Mais especificamente, pretende-se avaliar as técnicas *Principal Components Analysis* (PCA) e *Local Binary Pattern* (LBP), respectivamente. Objetiva-se, assim, não apenas identificar de forma empírica o método mais eficaz, mas também, fornecer uma contribuição relevante a respeito dos prós e contras de ambas abordagens. Tal proposta implica nos seguintes objetivos secundários:

- a) Analisar as características faciais de maior relevância no processo de reconhecimento;
- b) Desenvolver um modelo de arcabouço computacional completo para o reconhecimento automático de faces;
- c) Indicar a melhor solução para o problema de uma imagem de treinamento por pessoa, utilizando amostras que apresentem desvios de natureza distinta.

1.3 Organização do Trabalho

Este trabalho é composto por 5 capítulos. No capítulo 2, apresenta-se a revisão bibliográfica, em ordem cronológica, seguida de uma discussão sobre a contribuição dos trabalhos analisados e os desafios presentes nesta área. No capítulo 3, descrevem-se as etapas envolvidas no reconhecimento de padrões e as técnicas empregados neste estudo, sob o contexto do problema de reconhecimento facial. Os experimentos e resultados obtidos são relatados no capítulo 4. Por fim, no último capítulo, apresentam-se as conclusões e sugestões para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Neste capítulo são apresentados, em ordem cronológica, os principais estudos sobre reconhecimento facial e em seguida pretende-se comparar e discutir as características específicas de cada abordagem analisada com o propósito de relacioná-las ao contexto deste trabalho.

Kanade (KANADE, 1973), em sua tese, apresentou um estudo pioneiro sobre processamento de fotografias e reconhecimento automático de imagens de faces, no qual analisaram-se os métodos de digitalização e armazenamento mais adequados aos recursos disponíveis na época. Neste trabalho desenvolveu-se um processo de reconhecimento em dois estágios: na etapa de extração de características utilizou-se um operador Laplaciano (GONZALES; WOODS, 2000), com limiar heurístico, para a detecção de bordas e conseqüentemente a localização de elementos faciais como olhos, nariz, boca e formato do rosto. Nesta fase analisaram-se 800 imagens de face, divididas em quatro categorias: face completa sem artefatos, face completa com óculos, face com desvios de rotação ou inclinação e face com barba ou outros artefatos. Os resultados da extração foram classificados através de inspeção humana e obtiveram-se as seguintes taxas de extração correta 90.7%, 2.6%, 79.7% e 0%, respectivamente. No experimento final, utilizou-se 40 imagens de face de 20 indivíduos, sendo 17 homens e 3 mulheres, sem óculos, barba ou bigode. As duas imagens de cada indivíduo foram capturadas em locais diferentes com intervalo de 1 mês. Nesta etapa definiram-se 26 parâmetros faciais para comparação através da distância euclidiana ponderada, a fim de mensurar a similaridade entre as faces. Visando analisar a influência do pré-processamento, tais parâmetros foram divididos em duas categorias: extraídos por processos manuais ou automáticos. Verificou-se, também, a capacidade de reconhecimento com um número reduzido de características e deste modo, obteve-se resultados favoráveis ao procedimento manual empregando apenas as características mais relevantes.

Nos anos de 1987 e 1990, Sirovich e Kirby (SIROVICH; KIRBY, 1987; KIRBY; SIROVICH, 1990) apresentaram trabalhos empregando uma técnica de estatística multivariada, conhecida como *Principal Components Analysis (PCA)*, para representar imagens de face em um sub-espço reduzido preservando o máximo de informação original. Para os experimentos iniciais as imagens foram recortadas mantendo somente a região entre a base do nariz e o topo da sobrancelha. Utilizaram-se 115 fotografias, geradas em ambiente controlado, contendo exclusivamente indivíduos caucasianos do sexo masculino. Os resultados preliminares demonstraram que são necessárias menos de 100 componentes para a reconstrução adequada de uma imagem proveniente de um sub-espço reduzido, ou seja, uma imagem de face pode ser representada por menos de 100

atributos (ou variáveis). No estudo subsequente analisou-se a face delimitada por uma região oval abrangendo os olhos, o nariz e a boca. Deste modo pode-se observar a extração das características faciais, com a presença de artefatos e variações de iluminação, através das componentes principais geradas pelo PCA. Por fim, obteve-se uma taxa de compressão de 100:1, permitindo representar imagens de 91x51 pixels com apenas 50 variáveis.

Dando continuidade ao estudo de técnicas estatísticas neste contexto, Turk e Pentland (TURK; PENTLAND, 1991a, 1991b) relataram a primeira aplicação do PCA para reconhecimento de faces e cunharam o termo autofaces, também conhecido como *eigenfaces*, para descrever os autovetores que determinam o espaço de características faciais gerado por este processo. Em seus estudos, utilizou-se uma base composta por 2500 amostras de 16 indivíduos, dispostos conjuntos de treinamento contendo apenas uma imagem de cada pessoa, todas geradas sob as mesmas condições de iluminação, pose e escala. Para cada conjunto de treinamento classificou-se todas as imagens da base como sendo um dos 16 indivíduos e obteve-se, em média, aproximadamente 96% de acerto para variações de iluminação, 85% para variações de pose e 64% para variações de escala. No experimento seguinte adotou-se o mesmo procedimento, no entanto, definiu-se um limiar para rejeição de faces desconhecidas. Objetivando alcançar uma classificação 100% precisa a taxa de rejeição foi impelida à 19% para variações de iluminação, 39% para pose e 60% para escala. Ao definir-se, arbitrariamente, o limiar de rejeição em 20% as taxas de reconhecimento resultaram em 100%, 94% e 74%, respectivamente. Estes resultados demonstraram uma relação inversamente proporcional entre a taxa de classificação e o limiar de rejeição e indicaram, também, baixa sensibilidade a variações de iluminação e alta sensibilidade a variações de escala devido à tênue correlação entre imagens de tamanhos distintos.

Visando organizar os conceitos relacionados aos métodos propostos na época, em 1993, Brunelli e Poggio (BRUNELLI; POGGIO, 1993) classificaram as técnicas de reconhecimento facial como locais ou globais e descreveram uma comparação entre estas abordagens através de dois algoritmos de autoria própria. O primeiro algoritmo baseou-se em características geométricas, como altura e largura do nariz, posição da boca e o formato do rosto, e o segundo algoritmo realizou uma comparação dos pixels em escala de cinza. Para ambas implementações, empregou-se uma base contendo 188 imagens frontais, composta por 47 indivíduos, sendo 26 homens e 21 mulheres, com 4 amostras por pessoa. Duas imagens de cada indivíduo foram geradas na mesma sessão e as outras imagens em um intervalo de duas a quatro semanas. Os experimentos apresentaram 90% de reconhecimento correto para abordagem geométrica e 100% para abordagem holística, promovendo a utilização de técnicas globais.

Posteriormente, Chellappa et al. (CHELLAPPA; WILSON; SIROHEY, 1995) apresentaram uma extensa revisão teórica relacionada ao reconhecimento de faces realizado por máquinas e seres hu-

manos. Neste trabalho, analisaram-se questões psicofísicas e neurofisiológicas almejando compreender a relativa facilidade dos humanos no processo de reconhecimento facial. Contudo, também discutiram-se diversas técnicas, métricas e avaliações desenvolvidas nesta área e suas aplicações. Através deste estudo concluiu-se que: o reconhecimento automático de faces é essencial para aplicações legais no combate ao crime; os estudos psicofísicos e neurofisiológicos são úteis, porém, não essenciais à modelagem de sistemas automáticos de reconhecimento facial; a segmentação da face é uma tarefa fundamental no ciclo de reconhecimento, contudo, esta etapa carece de estudos efetivos; ambas abordagens locais e globais são úteis e devem ser aplicadas de acordo com o contexto; o reconhecimento de faces em vídeo é provavelmente o maior desafio presente nesta área; as imagens utilizadas nos experimentos devem provir de situações operacionais, similares ao ambiente de implantação dos sistemas de reconhecimento.

Em 1996, Belhumeur et. al. (BELHUMEUR; HESPANHA; KRIEGMAN, 1996) definiram uma nova abordagem para reconhecimento empregando a técnica *Linear Discriminant Analysis* (LDA), derivada do PCA e proposta também por Swets et. al. (SWETS; WENG, 1996), para representação de imagens. Neste trabalho, realizou-se uma comparação com o método de autofaces (TURK; PENTLAND, 1991a, 1991b) e a técnica proposta ficou conhecida como *fisherfaces*, por utilizar o modelo discriminante de Fisher (FISHER, 1936). Para os experimentos, utilizaram-se duas bases de imagens, uma composta por 500 amostras e outra por 176 amostras. Empregou-se a primeira para análise sob variações de iluminação e a segunda para tratar diferentes condições como gênero, expressão facial e presença de artefatos. Para validação deste estudo empregou-se o método "leave-one-out" e os resultados obtidos demonstraram que a técnica proposta apresentou resultados superiores em situações onde há variações extremas de iluminação e expressão facial.

Mais recentemente, Phillips et. al. (PHILLIPS et al., 2000) elaboraram um protocolo, denominado *Facial Recognition Technology* (FERET), para avaliação de técnicas automáticas e semi-automáticas de reconhecimento facial e criaram uma base de imagens extensa e padronizada. Projetou-se este modelo de modo que seja possível avaliar diferentes conjuntos de treinamento em tarefas de identificação e verificação, visando mensurar e estimular melhorias nesta área. Realizaram-se diversos experimentos a fim de estabelecer uma referência para trabalhos futuros. No entanto, para verificação, não houve nenhum algoritmo que apresentou resultados ótimos em todas as situações e concluiu-se que o bom desempenho em uma tarefa não é necessariamente preditivo para situações distintas.

Em 2002, Bartlett et. al. (BARTLETT; MOVELLAN; SEJNOWSKI, 2002) aplicaram uma variação do PCA, denominado *Independent Component Analysis* (ICA). Neste trabalho pretendia-se encontrar um bom conjunto de amostras para representar uma base de faces. Nos experimentos, utilizaram-se fotografias de 425 indivíduos, procedentes da base FERET (PHILLIPS et al., 2000),

contendo variações de expressão e intervalos de aquisição. Determinou-se o reconhecimento facial através de um classificador de vizinho mais próximo. Por fim, os resultados demonstraram um desempenho superior do método proposto, em relação ao PCA, em todas as situações analisadas.

No mesmo ano, Wu e Zhou (WU; ZHOU, 2002) propuseram uma extensão do PCA, denominado $(PC)^2A$, aplicado ao problema de reconhecimento facial com apenas uma imagem de treinamento por pessoa. Neste estudo utilizaram-se 400 imagens frontais de 200 indivíduos, 129 homens e 71 mulheres, selecionados aleatoriamente da base FERET. Para cada pessoa adotaram-se duas imagens, com expressões faciais distintas, com as quais foram gerados os conjuntos de treinamento e teste. Os experimentos com o método proposto obtiveram uma taxa de reconhecimento de 3% a 5% superior a técnica tradicional, utilizando entre 10% e 15% menos autofaces.

Com a intenção de revisar e atualizar os tópicos analisados por Chellappa et. al. (CHELLAPPA; WILSON; SIROHEY, 1995), Zhao et. al. (ZHAO et al., 2003) elaboraram um trabalho metucioso acerca das etapas de segmentação, extração de característica e classificação. Empregaram-se, neste estudo, diversas técnicas aplicadas a problemas de reconhecimento em imagens estáticas e análise de vídeo. As conclusões apresentadas reforçam discussões anteriores (CHELLAPPA; WILSON; SIROHEY, 1995) e enfatizam as seguintes questões: implementação de métodos compostos utilizando informações da fala, iris e impressão digital; problemas de iluminação, pose e envelhecimento compõem os maiores desafios nesta área; a padronização dos métodos de avaliação para os algoritmos é uma etapa crucial para o progresso do reconhecimento facial.

Motivados pela falta de direção, detalhamento e consenso entre os estudos precedentes, Delac et. al. (DELAC; GRGIC; GRGIC, 2005) apresentaram uma análise comparativa entre as três técnicas estatísticas mais populares PCA, ICA e LDA. Neste trabalho, realizaram-se diversos ensaios sob condições idênticas aplicando as seguintes métricas: Manhattan, Euclidiana, Mahalanobis e outras. Visando consistir os resultados com trabalhos análogos empregou-se imagens oriundas da base FERET (PHILLIPS et al., 2000), devido a sua padronização. Nos experimentos utilizaram-se 3 amostras por pessoa de 255 indivíduos para constituir o conjunto de treinamento e classificaram-se 1196 amostras, subdividas por variações de expressão, iluminação e alterações temporais de curto e médio prazo. Por fim, implementaram-se 16 algoritmos a fim de atender os cenários de teste e demonstrou-se que as combinações de técnicas e métricas avaliadas são adequadas para situações específicas e nenhum algoritmo pode ser considerado uma solução definitiva.

Em 2004, Perlibakas (PERLIBAKAS, 2004) comparou o desempenho de 14 métricas aplicadas ao reconhecimento de faces utilizando PCA e propôs uma variação da distância por erro médio quadrático. Neste trabalho, utilizaram-se fotografias de 423 indivíduos provenientes das bases AR (OHIO, 1998; MARTÍNEZ; BENAVENTE, 1998), IAM (BERN, 2008), BioID (AG, 2001), Yale

(YALE, 1997), PEIPA (ESSEX, 2004), MIT (MIT, 1991; TURK; PENTLAND, 1991a, 1991b), ORL (AT&T, 1994), Umist (UMIST, 1998) e FERET (PHILLIPS et al., 2000; STANDARDS; TECHNOLOGY, 1996). Nos experimentos adotaram-se duas amostras por pessoa, uma para treinamento e outra para comparação, e submeteram-se as imagens a uma etapa de normalização. Os resultados obtidos apresentaram melhor desempenho através das seguintes distâncias: Mahalanobis simplificada, angular ponderada, variação do erro médio quadrático proposta e angular entre vetores normalizados. Concluiu-se, também, que a combinação de algoritmos proporciona resultados superiores a implementação de distâncias isoladas.

Nos últimos anos, Ahonen et. al. (AHONEN; HADID; PIETIKAINEN, 2004, 2006) aplicaram ao problema de reconhecimento facial uma técnica denominada *Local Binary Pattern* (LBP), cujo conceito foi introduzido em 1996 por Ojala et. al. (OJALA; PIETIKAINEN; HARWOOD, 1996), como descritor de textura. Com a finalidade de avaliar o desempenho do algoritmo proposto utilizou-se o sistema CSU, desenvolvido por Bolme et. al (S. et al., 2003) conforme os procedimentos definidos no modelo FERET (PHILLIPS et al., 2000), salvo pequenas alterações, e conseqüentemente empregaram-se as imagens determinadas neste protocolo. Neste trabalho analisaram-se diversos parâmetros da técnica LBP, a fim de identificar as configurações adequadas para comparação com os seguintes métodos: PCA, *Bayesian Intrapersonal/Extrapersonal Classifier* (BIC) e *Elastic Bunch Graph Matching* (EBGM). Por fim, demonstrou-se um desempenho superior do LBP em todos os ensaios realizados. Obteve-se 97% de acerto com alterações de expressão, enquanto os métodos concorrentes não excederam 90%. Sob variações de iluminação alcançou-se 79% de acerto, contra 65%, 37% e 42% para PCA, BIC e EBGM, respectivamente.

Com o propósito de compreender os desafios relacionadas à limitação amostral, em 2006, Tan et. al. (TAN et al., 2006) compilaram, avaliaram e classificaram diversos trabalhos desenvolvidos com o intuito de solucionar esta questão. Neste estudo discutiram-se as conseqüências deste problema e temas como aquisição de imagens e sistemas de avaliação. A fim de realizar uma comparação justa empregou-se a mesma base de imagens para todas abordagens, de acordo com o procedimento FERET (PHILLIPS et al., 2000). Através dos resultados observados concluiu-se que: a solução deste problema pode beneficiar outras áreas relacionadas; pode-se utilizar técnicas alternativas para prover informações adicionais a fim de compensar a limitação amostral; a evolução desta área provavelmente será conduzida pela exploração de informações geradas a partir de outras faces e por modelos tridimensionais; a solução final deste problema depende muito dos avanços em áreas afins, tais como reconhecimento de padrões, aprendizado de máquina e de visão computacional, pois estas disciplinas apresentam diversos resultados sólidos; são necessários mais estudos sobre como a capacidade de generalização de um algoritmo é influenciado pelo tamanho do conjunto de treinamento. Contudo, acredita-se que o problema de reconhecimento facial pode

ser solucionado no âmbito do quadro técnico atual, devidos aos bons resultados que vem sendo alcançados.

Em 2009, Maturana et al. (MATURANA; MERY; SOTO, 2009) propuseram dois algoritmos para tratar variações de pose e alinhamento. Incorporou-se, neste trabalho, uma etapa de normalização para corrigir problemas de iluminação e compararam-se as técnicas LBP original (AHONEN; HADID; PIETIKAINEN, 2004, 2006), *Spatial Pyramid Matching* (SPM), *Naive Bayes Nearest Neighbor* (NBNN), PCA e LDA. Nos experimentos, utilizaram-se quatro bases de imagens, ORL (AT&T, 1994), Yale (YALE, 1997), Georgia Tech (TECHNOLOGY, 1999) e Yale B (YALE, 2005), contendo variações de pose, iluminação e expressão. Os resultados indicaram que a técnica NBNN é superior às demais, principalmente em situações onde há variações de pose. No entanto, o custo computacional deste método é relativamente alto se comparado ao LBP original.

Recentemente, Deng et al. (DENG et al., 2010) apresentaram uma nova técnica estatística para extração de características, derivada do PCA. O método proposto, denominado *Uniform Pursuit* (UP), visa solucionar o problema de uma amostra por pessoa e reduzir erros causados por faces semelhantes. Durante os experimentos empregaram-se metodologias de aprendizado específico (TAN et al., 2006) e genérico (WANG et al., 2006), utilizando fotografias procedentes da base FERET (PHILLIPS et al., 2000), para comparação com as técnicas PCA, PCA normalizada e ICA. Por fim, o método UP obteve 96.7%, 99.0%, 88.9% e 87.6% de acerto, sob variações de expressão, iluminação, intervalo de aquisição curto e intervalo de aquisição longo, respectivamente, superando as demais técnicas analisadas. Observou-se, também, que as duas abordagens de aprendizado empregadas demonstraram resultados efetivos. Portanto, acredita-se que a combinação de ambas seja um caminho promissor para aumentar a precisão do reconhecimento automático de faces.

2.1 Considerações Complementares

Inicialmente, em razão das limitações tecnológicas contemporâneas, os trabalhos desenvolvidos nesta área empregaram técnicas baseadas em características geométricas (KANADE, 1973). Com o avanço da capacidade computacional e a exploração de métodos estatísticos, a utilização de informações holísticas se tornou mais popular e fomentou a criação de diversas metodologias (SIROVICH; KIRBY, 1987; KIRBY; SIROVICH, 1990; TURK; PENTLAND, 1991a, 1991b; BELHUMEUR; HESPANHA; KRIEGMAN, 1996; BARTLETT; MOVELLAN; SEJNOWSKI, 2002; WU; ZHOU, 2002; THOMAZ; GILLIES; FEITOSA, 2003, 2004; THOMAZ; GIRALDI, 2010). Consequentemente surgiram estudos com a intenção de comparar o desempenho das técnicas propostas e os resultados preliminares favoreceram a implementação de modelos holísticos (BRUNELLI; POGGIO, 1993). No entanto, trabalhos subsequentes concluíram que ambas abordagens, locais e globais, são úteis de acordo com

o contexto e em situações específicas a implementação de modelos híbridos pode ser mais efetiva (CHELLAPPA; WILSON; SIROHEY, 1995; ZHAO et al., 2003).

A aplicação do PCA descrita por Sirovich e Kirby (SIROVICH; KIRBY, 1987; KIRBY; SIROVICH, 1990) é particularmente relevante devido aos problemas de alta dimensionalidade e redundância, presentes nas amostras tratadas no contexto deste trabalho. Logo, o arcabouço desenvolvido por Turk e Pentland (TURK; PENTLAND, 1991a, 1991b) tornou-se uma das metodologias mais bem sucedidas e referenciadas nesta área, pois fornece uma solução prática, rápida e relativamente simples que permite aprender novas imagens de face de modo não supervisionado, ou seja, sem interação humana.

Certas extensões do PCA apresentaram resultados bem sucedidos (BELHUMEUR; HESPANHA; KRIEGMAN, 1996; BARTLETT; MOVELLAN; SEJNOWSKI, 2002; THOMAZ; GIRALDI, 2010). Contudo, estudos comparativos, motivados pela falta de consenso na literatura afim, concluíram que não existe uma solução definitiva (DELAC; GRGIC; GRGIC, 2005). Em algumas situações, tais derivações obtiveram taxas de reconhecimento inferiores ao método original ou simplesmente falharam, devido a restrições da base de treinamento (TAN et al., 2006).

Com o surgimento do protocolo FERET os estudos sobre reconhecimento facial puderam empregar um modelo de avaliação padronizado, com o propósito de facilitar a comparação entre abordagens distintas (BARTLETT; MOVELLAN; SEJNOWSKI, 2002; WU; ZHOU, 2002; DELAC; GRGIC; GRGIC, 2005; PERLIBAKAS, 2004; AHONEN; HADID; PIETIKAINEN, 2004, 2006; TAN et al., 2006; DENG et al., 2010). Recentemente, a técnica LBP apresentou resultados promissores através deste protocolo e tem atraído a atenção da comunidade científica devido à sua baixa complexidade computacional e robustez à variações de pose e iluminação. Entretanto, aplicações reais de reconhecimento automático de faces apresentam características particulares, ausentes no modelo FERET, que serão abordados no escopo deste trabalho.

2.2 Desafios atuais

De acordo com Delac et al. (DELAC; GRGIC; BARTLETT, 2008), a comparação de duas imagens geradas em ambientes controlados é uma questão dada como resolvida por muitos cientistas, pois desvios mínimos são problemas triviais para o padrão dos métodos atuais que oferecem uma taxa de reconhecimento acima de 90%, desempenho superior à capacidade humana de reconhecimento nas mesmas condições. Entretanto, estes métodos ainda apresentam resultados inferiores aos seres humanos quando são submetidos a variações de pose, expressão, iluminação, envelhecimento e oclusão parcial da face.

Em aplicações reais, o número reduzido de amostras de treinamento e as variações complexas presentes nas amostras de teste acarretam problemas aos sistemas de reconhecimento (LU, 2003). Embora esta condição não comprometa plenamente os métodos geométricos (TAN et al., 2006), muitos estudos baseados em abordagens holísticas relataram depender diretamente da quantidade e representatividade das amostras de treinamento para obter resultados satisfatórios (DENG et al., 2010; TAN et al., 2006). Dada apenas uma amostra por pessoa, a maioria dos procedimentos analisados apresenta queda de desempenho ou até mesmo deixa de funcionar. Esta questão, conhecida como *"one sample per person problem"* provê inúmeros desafios à comunidade científica.

Outro problema presente em sistemas de reconhecimento consiste na rejeição confiável de amostras que não possuem imagens correspondentes no conjunto de treinamento (BRUNELLI; POGGIO, 1993), com o propósito de mitigar a ocorrência de falsos positivos, pois os algoritmos de busca sempre apresentam o resultado mais próximo, mesmo que este não seja o indivíduo procurado. Os estudos realizados por Turk e Pentland (TURK; PENTLAND, 1991b) e posteriormente por Perlibakas (PERLIBAKAS, 2004), utilizaram um limiar baseado em métricas para rejeitar amostras desconhecidas. O limite empregado na classificação pode ser definido através de heurísticas ou aprendizado, considerando a natureza e o comportamento dos dados processados. A resolução desta questão é fundamental para situações críticas, como por exemplo a busca de pessoas desaparecidas, onde a validação precisa dos resultados minimiza o gasto desnecessário com análise de DNA e o transtorno emocional dos familiares.

3 METODOLOGIA

O processo natural de reconhecimento de padrões tem evoluído durante milhares de anos e sempre desempenhou um papel fundamental para a sobrevivência das espécies (DUDA; HART; STORK, 2001). Entretanto, a criação de um modelo computacional capaz de realizar esta atividade ainda é um desafio. Este problema foi amplamente analisado por Duda et al. (DUDA; HART; STORK, 2001), visando a definição de um protótipo para sistemas de classificação contendo as seguintes etapas: aquisição, pré-processamento, classificação e pós-processamento. A Figura 3.1 delinea as principais etapas presentes em um processo automático de reconhecimento de padrões. Cada uma das etapas será detalhada nas seções seguintes, contextualizando o problema de reconhecimento de faces no modelo computacional reconhecido pela literatura.

3.1 Aquisição de amostras

A etapa de aquisição ocorre por meio de um sensor que transforma cada amostra em um sinal digital (DUDA; HART; STORK, 2001). Em aplicações de processamento de imagens e reconhecimento facial as amostras frequentemente são convertidas em uma matriz de pixels através de dispositivos ópticos, do tipo CMOS ou CCD (LITWILLER, 2001), presentes em equipamentos de foto e filmagem ou scanner. No entanto, sensores térmicos, sonoros ou eletromagnéticos também podem ser empregados na captação do sinal, sendo os dois últimos amplamente utilizados em métodos de diagnóstico assistido por computador (LEÃO, 2009). As imagens empregadas na divulgação de crianças e adolescentes desaparecidos, por exemplo, são geradas em ambientes não-controlados sem qualquer critério de aquisição.

3.2 Pré-processamento

Diversos métodos foram empregados para tratar problemas específicos em reconhecimento de padrões. Contudo, uma prática comum na maioria dos casos analisados é a inclusão de uma etapa de pré-processamento (BRUNELLI; POGGIO, 1993; DENG et al., 2010; KITANI, 2007; THOMAZ, 1999), na qual podem ser empregadas técnicas de normalização, segmentação e extração de características. Pretende-se assim, padronizar as amostras e facilitar o processamento dos dados.

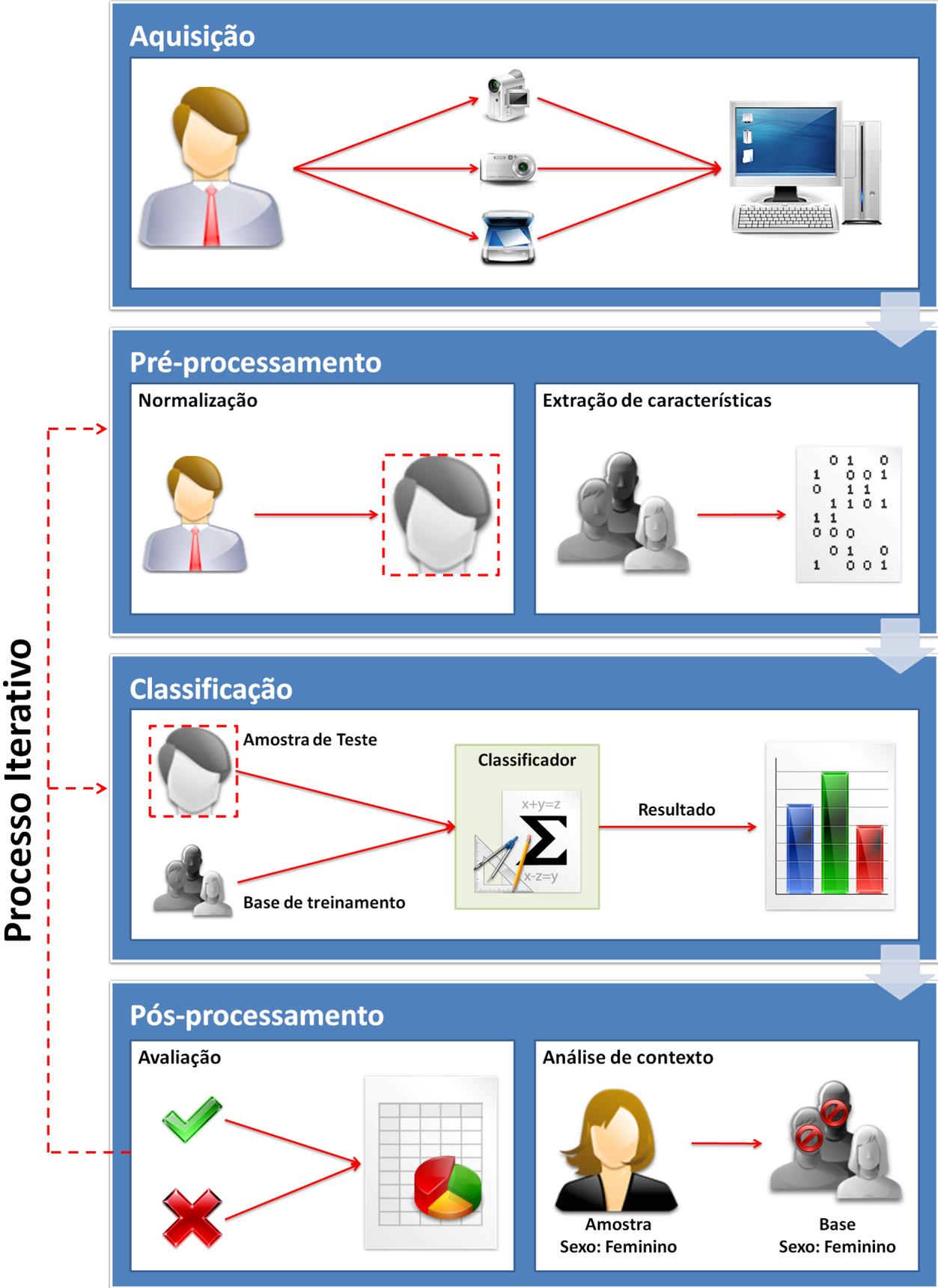


Figura 3.1: Etapas presentes em um processo iterativo de reconhecimento de padrões.

3.2.1 Normalização

As técnicas de normalização visam reduzir a variabilidade das amostras e remover informações indesejadas, geradas pelo processo impreciso de aquisição ou por limitações dos instrumentos utilizados (AMARAL et al., 2009; DUDA; HART; STORK, 2001). O processo de normalização empregado neste trabalho segue o modelo sugerido por Amaral et al. (AMARAL et al., 2009). Em que, tratam-se basicamente variações de rotação, translação, escala e iluminação, conforme ilustrado na Figura 3.2.

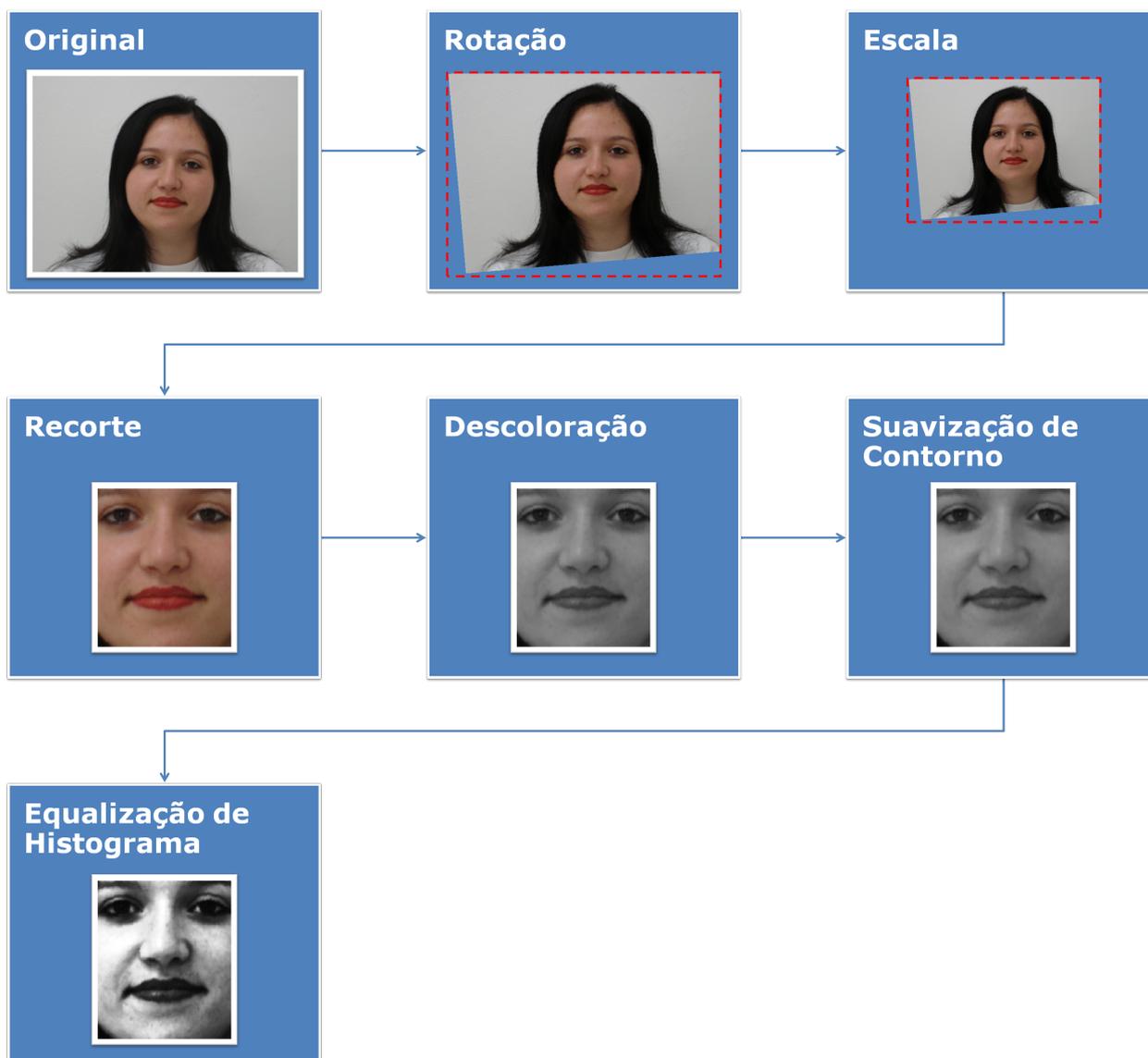


Figura 3.2: Normalização de imagens frontais de face.

Mais especificamente, este processo pode ser descrito da seguinte forma: através da identificação manual ou automática da posição das pupilas são definidos os pontos de referência. Em

seguida submete-se a imagem a uma etapa de rotação, cujo objetivo é alinhar esses pontos em relação ao eixo horizontal. Para isto utilizam-se os valores de seno e cosseno do ângulo θ , que descreve a inclinação da reta imaginária formada pelos pontos de referência. A Figura 3.3 ilustra esta etapa.

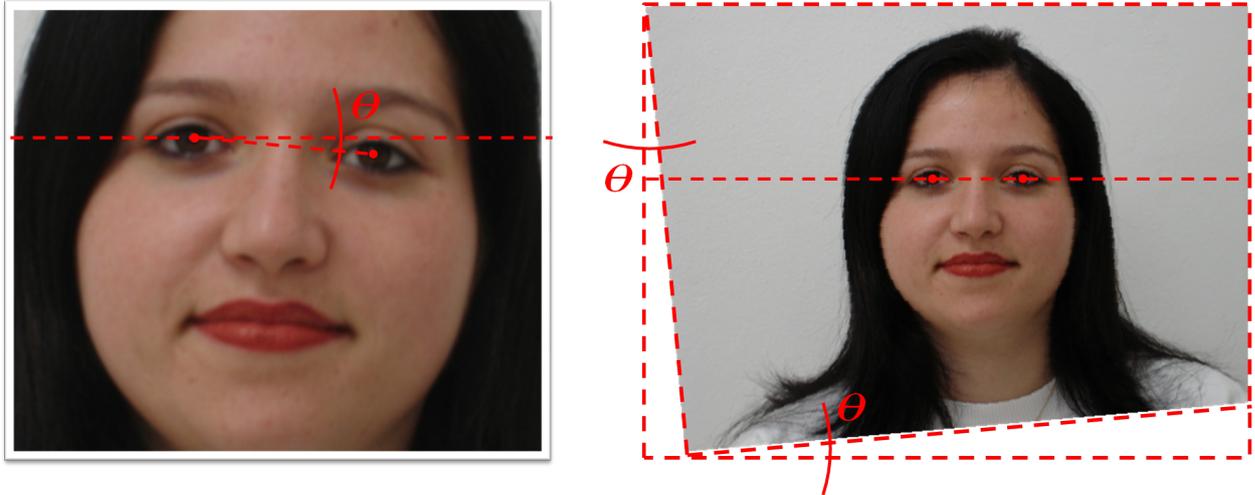


Figura 3.3: Rotação de imagem de face com base na posição das pupilas.

Sendo assim, cada pixel da imagem original pode ser mapeado em novas coordenadas pelas equações:

$$x_d = x_o \cos \theta + y_o \sin \theta \quad (3.1)$$

$$y_d = -x_o \sin \theta + y_o \cos \theta. \quad (3.2)$$

Contudo, este método gera sobreposição de pixels devido ao arredondamento dos resultados. Portanto, foi implementado um método inverso a partir da decomposição das equações (3.1) e (3.2). Calcularam-se os pontos de origem (x_o, y_o) correspondentes a cada ponto de destino (x_d, y_d) , utilizando uma técnica conhecida como rotação em duas passadas (CONCI; AZEVEDO; LETA, 2008). Deste modo, isolou-se inicialmente a posição de um dos eixos para compor o cálculo referente às coordenadas do segundo eixo, conforme equações (3.4) e (3.6), respectivamente,

$$x_d = x_o \cos \theta + y_o \sin \theta \quad (3.3)$$

$$x_o = \frac{x_d - y_o \sin \theta}{\cos \theta}, \quad (3.4)$$

onde $\cos \theta \neq 0$, e a fórmula resultante é aplicada ao cálculo da segunda coordenada:

$$y_d = - \left(\frac{x_d - y_o \text{sen}\theta}{\text{cos}\theta} \right) \text{sen}\theta + y_o \text{cos}\theta \quad (3.5)$$

$$y_o = \frac{-\text{sen}\theta \left(\frac{x_d}{\text{cos}\theta} \right) - y_d}{\text{sen}\theta \left(\frac{-\text{sen}\theta}{\text{cos}\theta} \right) - \text{cos}\theta}. \quad (3.6)$$

Finaliza-se esta etapa aplicando o valor obtido pela Equação (3.6) na variável correspondente da Equação (3.4). Dessa forma, a rotação é aplicada sem perda de informação da imagem original.

A seguir, realiza-se a normalização de escala. Neste processo, padroniza-se a distância interpupilar através da multiplicação de cada pixel da imagem destino pela razão entre a distância interpupilar original e um valor pré-definido, conforme exemplo apresentado nas figuras 3.4(a) e 3.4(b).

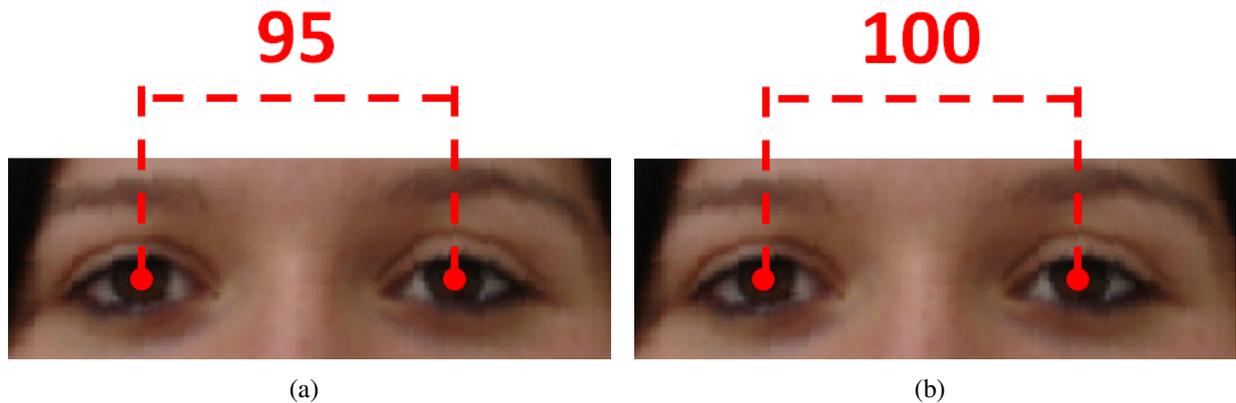


Figura 3.4: Distância interpupilar original (a) e escala normalizada para 100 pixels (b).

Após a normalização da escala é realizado o recorte. Nesta etapa eliminam-se artefatos desnecessários, visando preservar apenas as informações e características mais relevantes aos estudos sobre biometria de faces frontais. O trabalho de Amaral et al. (AMARAL et al., 2009) propôs como parâmetro delimitador a utilização da razão áurea, proporção presente em várias medidas faciais, denotada neste trabalho pela letra κ e com valor arredondado de 1,618, devido a sua relevância em estudos de proporção e harmonia facial (MORIHISHA, 2006). Deste modo, é possível representar essencialmente a face das pessoas, excluindo, por exemplo, a testa, o cabelo e o pescoço. A relação das medidas de recorte pode ser observada na Figura 3.5.

Na etapa subsequente, objetiva-se reduzir o esforço computacional através de um algoritmo de descoloração que substitua o sistema RGB (*Red, Green e Blue*), que trabalha com três canais de cores, por um sistema de camada única em escala de cinza. Esta transformação é gerada através da seguinte equação:

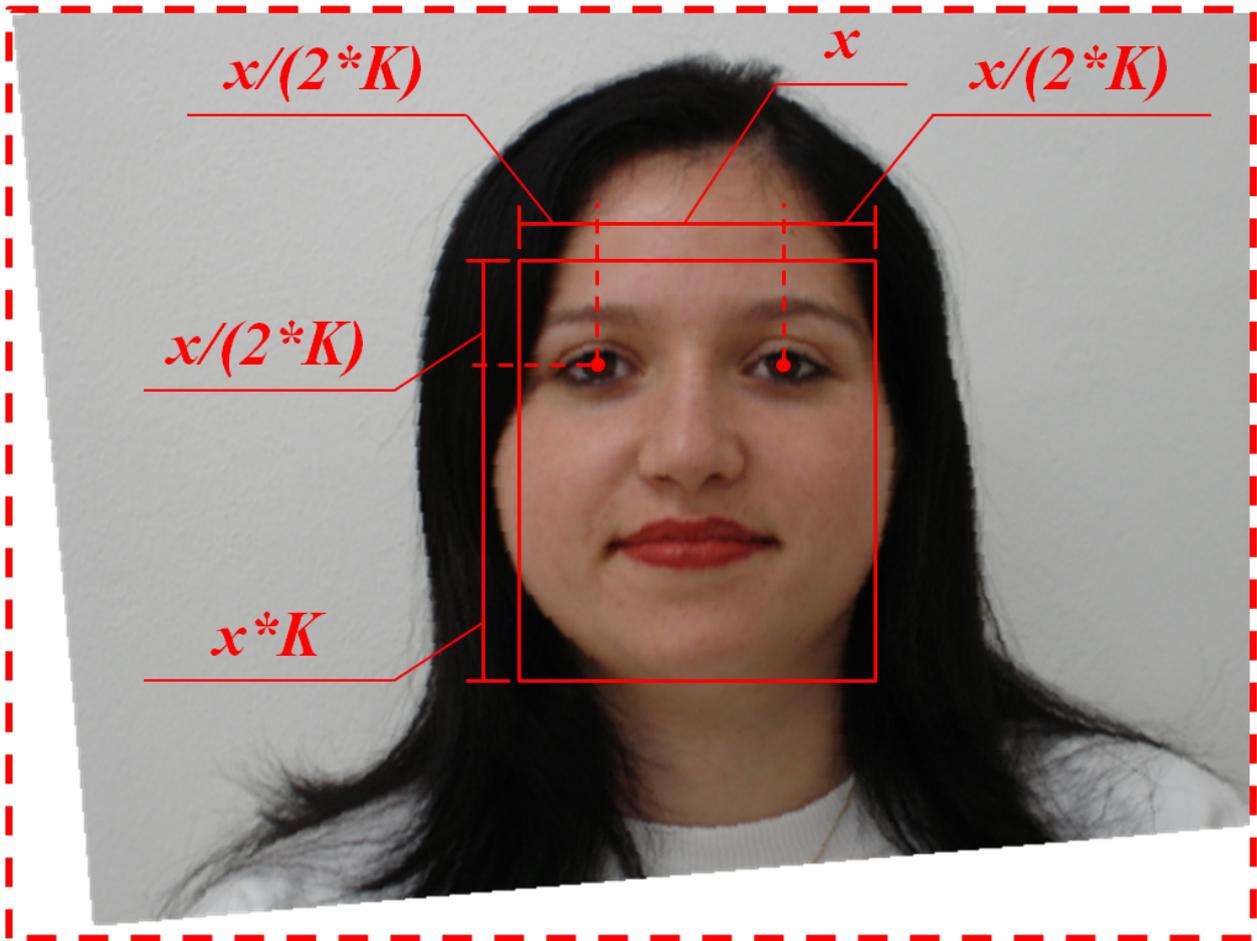


Figura 3.5: Relação das medidas de recorte, onde $x = 100$ pixels e $\kappa = 1,618$.

$$(x_d, y_d) = 0.299(x_o, y_o)_R + 0.587(x_o, y_o)_G + 0.114(x_o, y_o)_B. \quad (3.7)$$

As constantes 0.299, 0.587 e 0.114 correspondem ao modelo YIQ (*Luminance, Inphase e Quadrature*), projetado com base no percentual de sensibilidade da visão humana (GONZALES; WOODS, 2000). Portanto, a Equação (3.7) fornece um resultado mais adequado que a divisão proporcional, com 33% para cada cor do padrão RGB.

Em seguida aplica-se uma técnica de suavização de contorno visando minimizar o efeito de serrilhado gerado pela rotação. Utilizou-se uma matriz de 3x3, para mapear e substituir o valor do pixel central pela mediana dos pixels vizinhos, conforme exemplo:

$$\begin{pmatrix} 101 & 110 & 196 \\ 95 & 111 & 115 \\ 89 & 99 & 107 \end{pmatrix}. \quad (3.8)$$

Os valores contidos no kernel são ordenados e o valor central do conjunto se refere à mediana, conforme vetor abaixo:

$$[89;95;99;101;107;110;111;115;196;]. \quad (3.9)$$

No exemplo dado, o valor da mediana é o quinto elemento do conjunto ordenado e possui valor 107, que deverá substituir o antigo valor central 111 da matriz. O resultado pode ser observado na Figura 3.6:

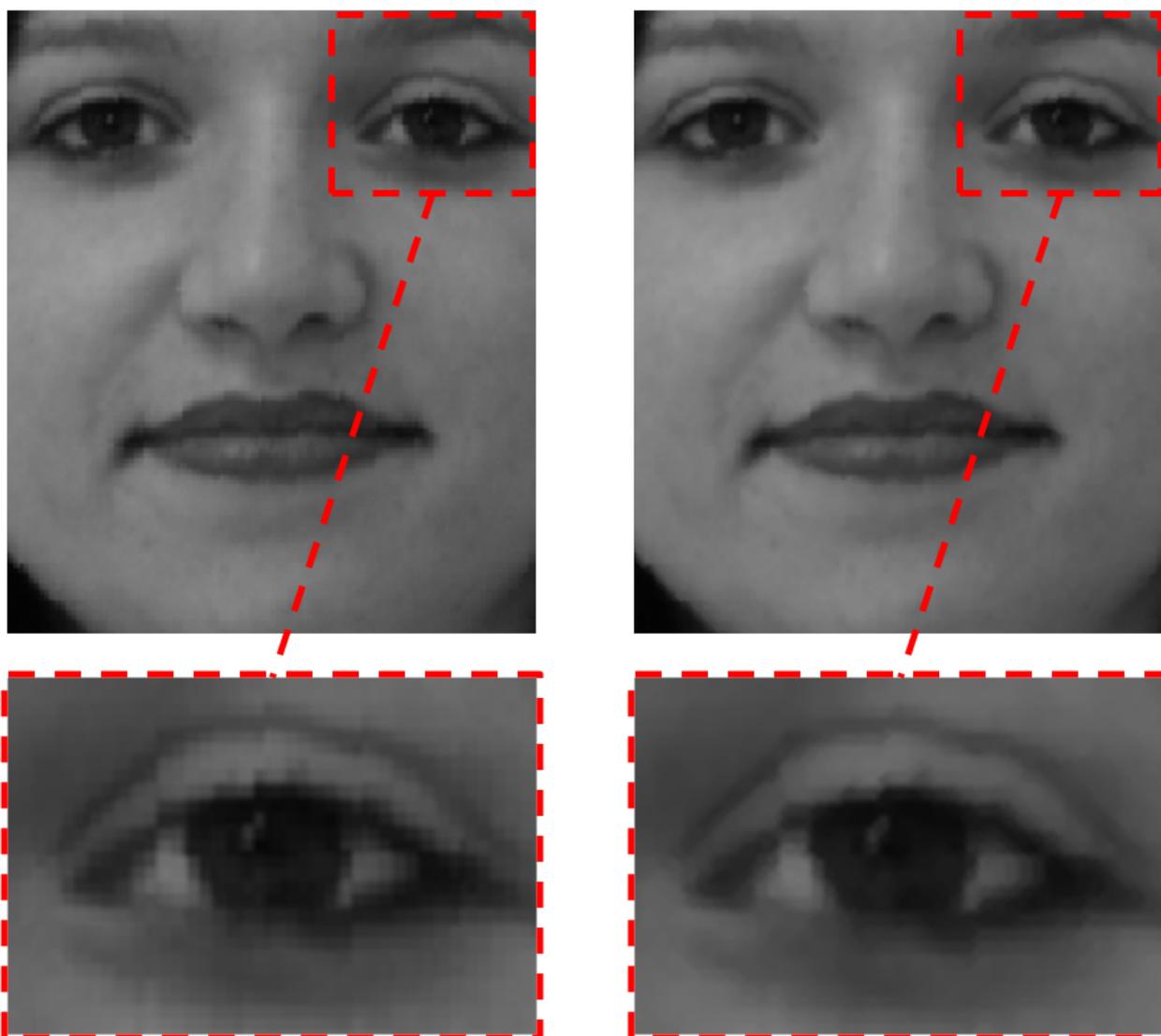


Figura 3.6: Detalhe da etapa de suavização de contorno.

A última etapa de normalização é a equalização de histograma. Segundo Conci, Azevedo e Leta (CONCI; AZEVEDO; LETA, 2008), este processo tem como objetivo melhorar o contraste das imagens através da representação dos níveis de cinza de maneira uniforme e distribuída. Portanto, ao final do processo de normalização acrescentou-se esta técnica visando otimizar imagens de face com baixo contraste, muito comuns em ambientes não-controlados onde a fonte de iluminação

é aleatória e muitas vezes precária. Outro benefício que esta técnica proporciona aos estudos biométricos se deve ao realce de rugas e outros detalhes faciais, aumentando assim a distinção entre faces de indivíduos diferentes. Para a aplicação desta técnica, calcula-se inicialmente um histograma, normalizado entre 0 e 1 (CONCI; AZEVEDO; LETA, 2008), que define a probabilidade original P de distribuição dos níveis de cinza presentes na imagem, através da seguinte equação:

$$P(k) = \frac{n_k}{l \times c}, \quad (3.10)$$

onde n_k é a quantidade de ocorrências do nível de cinza n_k , e l e c são os números de linhas e colunas da imagem, ou seja, $l \times c$ é igual ao número total n de pixels. Em seguida, para obter o histograma acumulado normalizado utiliza-se a seguinte função de distribuição:

$$T(k) = \sum_{i=0}^k \frac{n_i}{n}. \quad (3.11)$$

O resultado desta equação gera uma escala probabilística uniformemente distribuída entre 0 e 1. De acordo com Gonzales e Woods (GONZALES; WOODS, 2000), a conversão dos níveis de cinza para valores probabilísticos é realizada através da divisão de cada nível pelo número de valores possíveis. Portanto, a conversão da escala probabilística para níveis de cinza é obtida através da operação inversa e o nível de cinza equalizado de um pixel qualquer u da imagem original é dado por:

$$u_e = T(u_o * b), \quad (3.12)$$

onde u_e é o nível de cinza equalizado, u_o é o nível de cinza original, e o parâmetro b é o número de valores possíveis na escala de cinza. Para as imagens utilizadas neste trabalho, $b = 256$. O efeito deste procedimento pode ser verificado na imagem 3.7.



Figura 3.7: Imagem com histograma original e equalizada.

Contudo, a normalização de amostras geradas em ambientes controlados, frequentemente

utilizadas em estudos científicos, apresenta um grau de complexidade pequeno se comparado ao tratamento de amostras obtidas em ambientes reais. Logo, isto implica em um desafio ainda maior para aplicações práticas na área de Visão Computacional. Estudos realizados por Amaral et al. (AMARAL; THOMAZ, 2008; AMARAL et al., 2009) apresentaram resultados favoráveis à utilização deste procedimento aplicado à imagens frontais de face, oriundas de ambientes controlados e não-controlados. As imagens 3.8(a) e 3.8(b) apresentam os resultados obtidos nestes trabalhos.

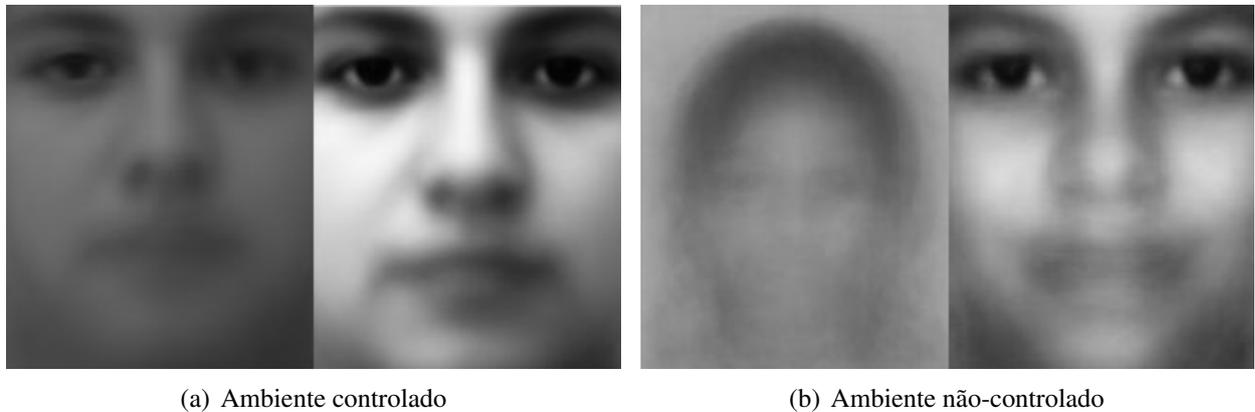


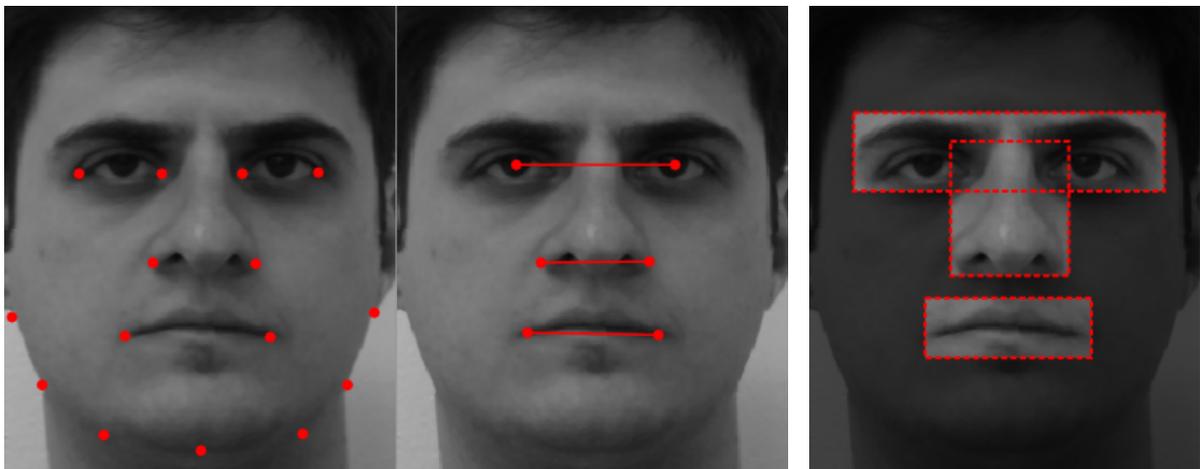
Figura 3.8: Imagens médias obtidas em ambientes controlados (a) e não-controlados (b), antes e depois da normalização.

3.2.2 Extração de características

A extração de características tem como objetivo reduzir a dimensionalidade dos dados eliminando informações redundantes (DUDA; HART; STORK, 2001). Imagens de face são frequentemente representadas por uma matriz de dimensão $l \times c$, contendo valores entre 0 e 255, referentes à intensidade de cada pixel em escala de cinza. O parâmetro l é o número de linhas e c o número de colunas (KITANI; THOMAZ, 2006), ou seja, a largura e a altura da imagem em pixels, respectivamente. Deste modo, as amostras processadas são frequentemente representadas por vetores contendo apenas as informações mais relevantes ao contexto da aplicação. Os métodos de extração empregados em reconhecimento facial costumam atender situações específicas e podem ser classificados em três categorias: locais, globais ou híbridos (GUTTA; WECHSLER, 1997; LIAU et al., 2008; TAN et al., 2006), conforme segue:

- a) Métodos locais: Estes métodos possuem duas abordagens principais. A primeira utiliza características geométricas, como distância e forma, para calcular medidas relativas entre componentes faciais particulares. Estas informações são representadas por um grafo que sintetiza a imagem de face e proporciona um processamento rápido (BRUNELLI; POGGIO, 1993). No entanto, os processos automáticos para demarcação das características são complexos e imprecisos, tornando-os dependentes de procedimentos manuais. A segunda abordagem baseia-se na

aparência de elementos independentes, ou seja, utiliza informações globais de regiões específicas da face, como, por exemplo, os olhos, nariz e boca. Permite-se, deste modo, explorar os fatores de relevância cognitiva de cada elemento facial (BRUNELLI; POGGIO, 1993; TAN et al., 2006). As figuras 3.9(a) e 3.9(b), ilustram as características frequentemente empregadas em ambas abordagens.



(a) Características geométricas.

(b) Aparência local.

Figura 3.9: Abordagens de mapeamento local para extração de características.

- b) Métodos globais: Métodos globais, também conhecidos como holísticos, dominam este campo de estudo desde 1990 (TAN et al., 2006) e representam cada amostra como uma matriz bidimensional (BRUNELLI; POGGIO, 1993). Esta estrutura permite o desenvolvimento de sistemas autônomos utilizando técnicas de aprendizagem. No entanto, estes métodos são altamente dependentes da quantidade e relevância das amostras de treinamento e exigem uma capacidade computacional maior para serem implementados, devido ao padrão de armazenamento dos dados. A Figura 3.10 ilustra a representação holística de uma imagem de face.
- c) Métodos híbridos: Tais métodos empregam informações locais e globais. Nesta abordagem, visa-se preservar as vantagens e reduzir as desvantagens de ambos. Entretanto, este processo de otimização requer uma análise precisa sobre quais características deverão ser combinadas (TAN et al., 2006).

3.2.2.1 Análise de Componentes Principais (PCA)

Também denominada transformada de Karhunen-Loève, ou em inglês de Principal Components Analysis, esta é uma das técnicas mais conhecidas e bem sucedidas no processo de reconhecimento facial (KITANI, 2007; TAN et al., 2006). Sua aplicação em imagens de faces foi proposta inicialmente por Sirovich e Kirby (SIROVICH; KIRBY, 1987), visando reduzir a dimensionalidade

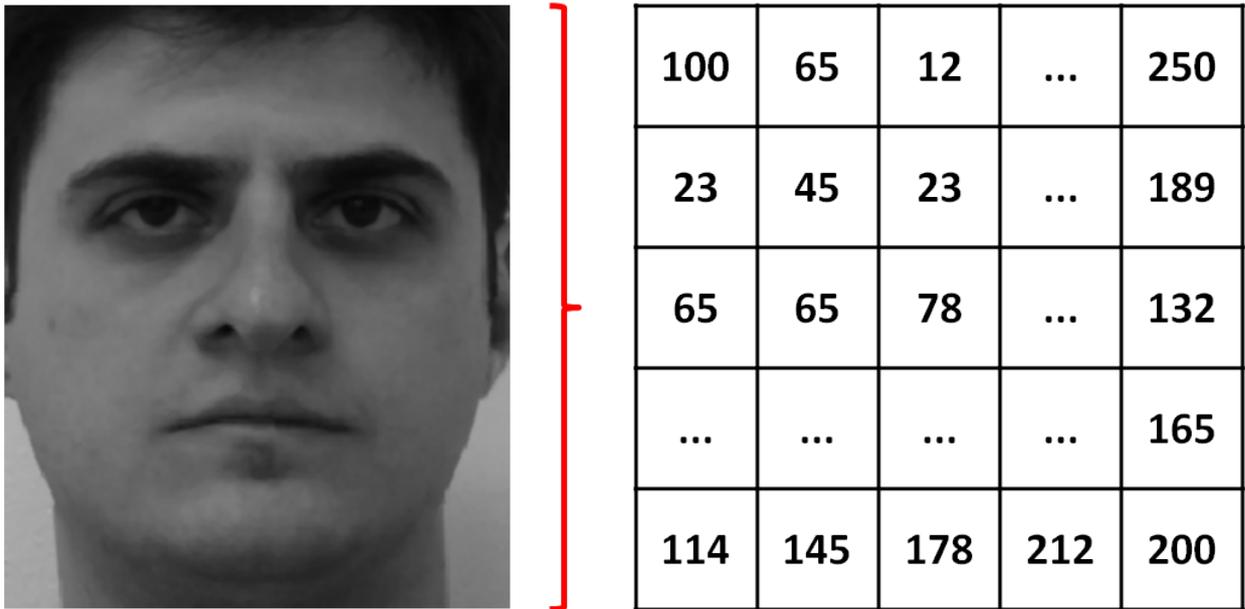


Figura 3.10: Extração global de características: matriz bi-dimensional de intensidade de pixels em escala de cinza.

dos dados preservando informações relevantes das amostras originais. Em 1991, Mathew Turk e Alex Pentland (TURK; PENTLAND, 1991a, 1991b) apresentaram um trabalho pioneiro, utilizando PCA aplicada ao reconhecimento de faces e difundiram o termo *eigenfaces*, também conhecido como auto-faces, em função dos autovetores utilizados para descrever a representação das faces no sub-espço gerado.

Inicialmente as imagens são representadas como vetores n -dimensionais gerados através da concatenação das linhas de cada imagem, onde $n = l * c$. Se considerarmos os valores do vetor x como coordenadas no espaço \mathcal{R}^n , as imagens de face serão pontos neste espaço. Sendo assim, um espaço multidimensional de faces contém todas as amostras de um determinado conjunto de treinamento. A Figura 3.11 ilustra a matriz de faces de dimensão $N \times n$, onde N é a quantidade de imagens de face e n é o produto $l * c$, ou seja, a quantidade total de pixels de cada imagem.

Segundo Kitani e Thomaz (KITANI; THOMAZ, 2006), o excesso de informação redundante permite que cada amostra seja representada pela sua variância em relação a média. Portanto, é plausível analisar esta variância em torno de uma média global. Dada a matriz do espaço de faces X , obtém-se o vetor média através da seguinte equação:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad (3.13)$$

onde N é a quantidade total de amostras do conjunto de treinamento. O vetor gerado é conhecido como "face média" (KITANI; THOMAZ, 2006) e representa as características comuns entre as faces

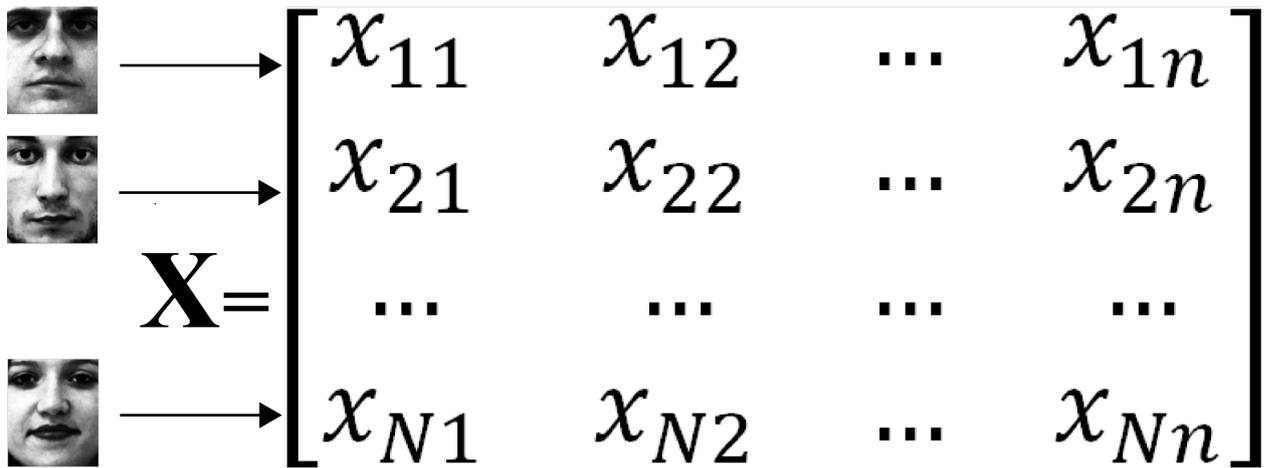


Figura 3.11: Representação do espaço multidimensional de faces.

que compõem o espaço multidimensional. Um exemplo de "face média" pode ser observado na figura 3.12.

Figura 3.12: Imagem de face média.

A equação 3.13 calcula a média aritmética de cada coluna da matriz X . Portanto, o desvio global de cada face deverá variar conforme:

$$z_i = x_i - \bar{x}, \quad (3.14)$$

onde $i = 1, 2, \dots, N$. Se então aplicarmos a equação 3.14 a todas as faces da matriz X , teremos uma matriz Z que contém apenas as variações de cada face em torno da média \bar{x} , conforme equação 3.15:

$$Z = [z_1, z_2, \dots, z_N]^T. \quad (3.15)$$

Em seguida, calculam-se os autovetores e autovalores, que descrevem as componentes principais, utilizando a matriz de covariância de Z . No entanto, é necessário um esforço computacional enorme para calcular a matriz de coeficientes $S_1 = Z^T Z$. Esta operação é inviável devido à alocação de memória limitada, visto que para uma imagem de 64x64 pixels, por exemplo, será gerada

uma matriz de 4096 x 4096 (KITANI; THOMAZ, 2006). Contudo, é possível calcular a matriz de covariância com esforço muito menor, gerando uma matriz $N \times N$, através da equação:

$$S_2 = \left(\frac{Z \cdot Z^T}{N-1} \right), \quad (3.16)$$

segundo Fukunaga (FUKUNAGA, 1990). Entretanto, os autovetores Φ da matriz S_2 calculada não são ortonormais. Portanto, Fukunaga (FUKUNAGA, 1990) propõe recalculá-los, conforme a equação abaixo:

$$\Phi^* = \frac{Z \cdot \Phi \cdot \Lambda^{-\frac{1}{2}}}{\sqrt{N-1}}, \quad (3.17)$$

onde Λ são os autovalores da matriz S_2 e Φ^* é a nova matriz de transformação, que descreve as faces do conjunto de treinamento nesse novo subespaço de menor dimensão, cujos autovetores são ortogonais e ortonormais. Portanto, qualquer análise sobre as amostras de treinamento deve ser executada a partir dessa matriz. A etapa de extração de características, com intuito de reduzir a dimensionalidade dos dados, se encerra aqui. Entretanto, tendo em vista a comparação e classificação das imagens em um processo de reconhecimento de padrões, as amostras de teste também devem ser projetadas neste mesmo subespaço, conforme a seguinte equação:

$$x_{pca} = (x_{teste} - \bar{x})^T \cdot \Phi^*, \quad (3.18)$$

onde x_{teste} é uma amostra de teste e x_{pca} é a representação dessa amostra no subespaço das componentes principais.

Outro aspecto relevante desta técnica é a capacidade de percorrer o subespaço gerado pelo PCA agregando à uma imagem as informações descritas por cada autovetor. Deste modo é possível visualizar e analisar as características mais variantes em cada conjunto de treinamento. O cálculo deste processo é realizado pela seguinte equação:

$$x_{nav} = x_i + \varphi_j^T b \sqrt{\lambda_j}, \quad (3.19)$$

onde φ_j^T é um autovetor, λ_j é um autovalor, b é o passo da navegação e x_{nav} é a representação gerada com as informações do autovetor percorrido. A Figura 3.13 ilustra, no subespaço gerado pelo PCA, a navegação sobre o eixo de maior variância e a distribuição bidimensional de um conjunto de amostras composto por imagens de faces com expressões neutra e sorrindo.

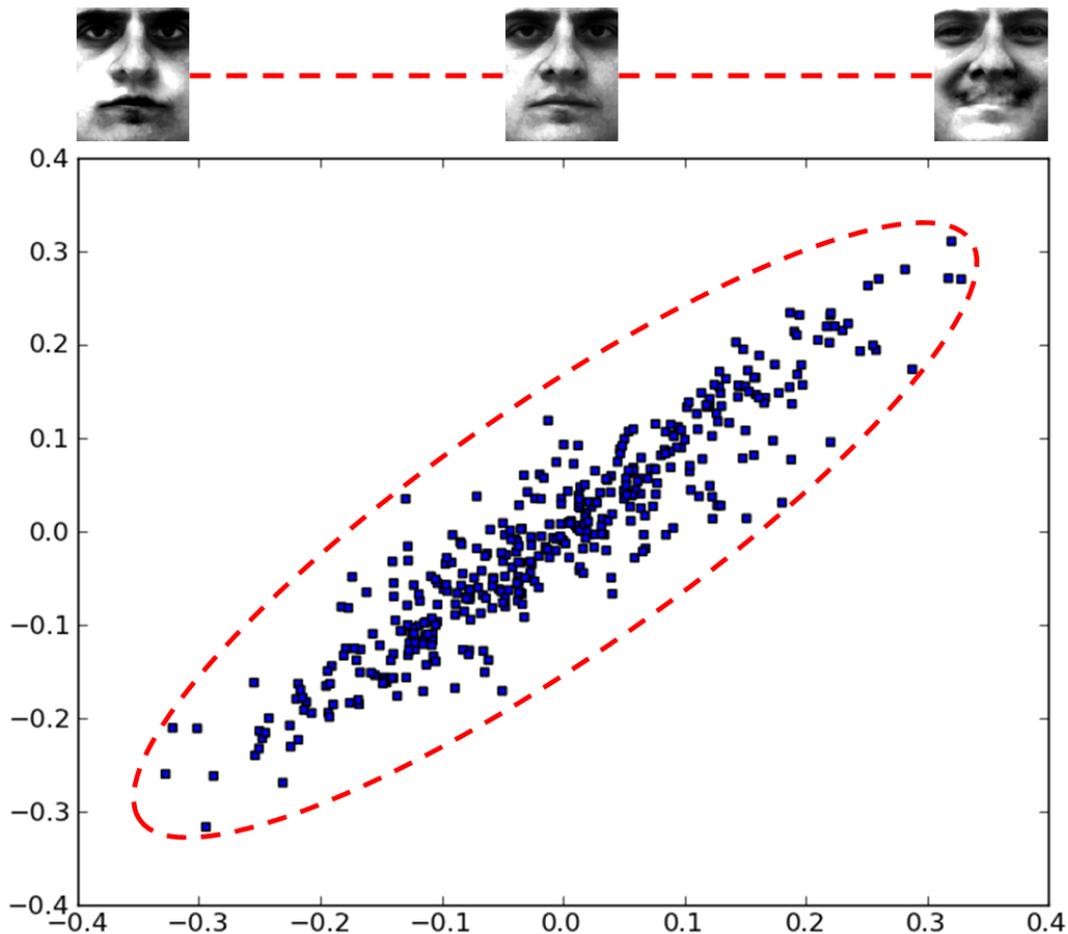


Figura 3.13: Distribuição bidimensional de amostras e navegação no eixo de maior variância no subespaço gerado pelo PCA .

3.2.2.2 Padrão Binário Local (LBP)

Inicialmente desenvolvido para atuar na análise de texturas (OJALA; PIETIKAINEN; HARWOOD, 1996), o LBP, do inglês Local Binary Pattern, tem sido empregado, com sucesso, no processo de reconhecimento de imagens de face. Isso ocorre pelo fato de que as faces podem ser vistas como uma composição de micro-padrões que são bem descritos por esta técnica (AHONEN; HADID; PIETIKAINEN, 2006). Neste método cada pixel de uma imagem é substituído por um valor binário. Este valor é determinado pela comparação de uma matriz quadrada contendo os pixels vizinhos, onde cada vizinho é comparado com o valor central, conforme a seguinte condição:

$$b_{ij} = \begin{cases} 0, & v_{ij} < v_c \\ 1, & v_{ij} \geq v_c \end{cases}, \quad (3.20)$$

onde v_{ij} é valor de um pixel na posição (i, j) e v_c é o valor central. Os valores obtidos para cada vizinho são concatenados e o número binário gerado é convertido na base decimal para substituir

o valor central v_c . A Figura 3.14 ilustra este processo, para uma matriz 3×3 de pixels vizinhos, e apresenta o resultado gerado por esta transformação:

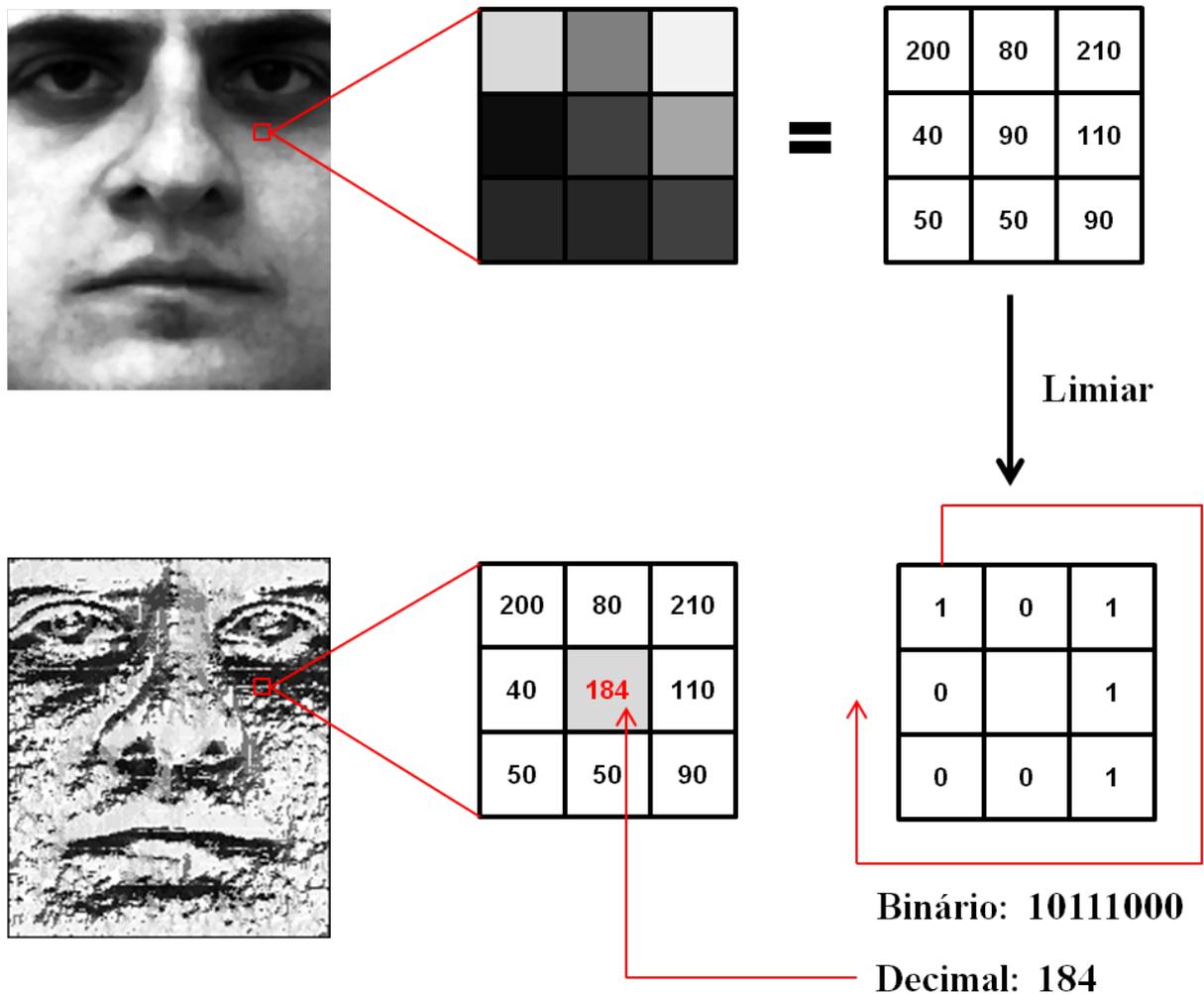


Figura 3.14: Operação básica do LBP, com valor binário resultante 10111000 e decimal 184.

O tamanho da vizinhança analisada pode variar e a seguinte notação (Q,R) é utilizada para determinar a configuração da matriz, onde Q é a quantidade de pontos comparados e R é o raio da vizinhança. Além disso, a vizinhança pode ser circular ou quadrada, de acordo com a Figura 3.15. Em seguida a imagem gerada é dividida em regiões, habitualmente organizadas em grades. Entretanto, é possível empregar outras formas com, ou sem, sobreposição. A Figura 3.16, apresenta exemplos dessas divisões.

Por fim os descritores de textura são extraídos de cada região isoladamente, calculando-se o histograma de intensidade dos pixels. Os vetores resultantes são concatenados em uma única matriz, denominada "*Spatially Enhanced Histogram*". A Figura 3.17 ilustra esta etapa.

A principal vantagem proporcionada pela técnica LBP origina-se da possibilidade de representação das imagens de face sem a necessidade de uma base de treinamento para auxiliar o

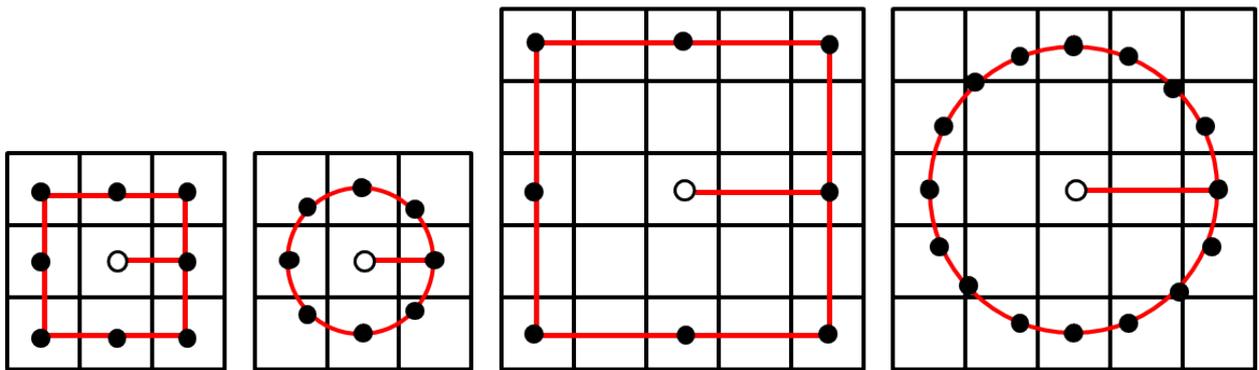


Figura 3.15: Vizinhança quadrada (8, 1), circular (8, 1), quadrada (8, 2) e circular (16, 2). O valor dos pontos descentralizados é determinado pela interpolação bilinear dos pixels vizinhos.

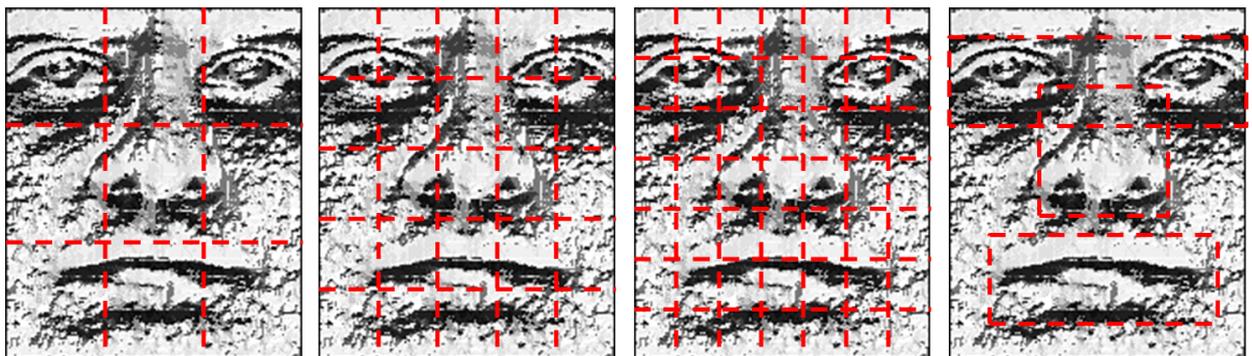


Figura 3.16: Imagem de face processada pelo LBP, dividida em grades 3×3 , 5×5 e 7×7 e regiões específicas dos olhos, boca e nariz.

processo de abstração das amostras.

3.3 Classificação

A classificação consiste em estabelecer o padrão correspondente a uma amostra através da criação de fronteiras de decisão baseadas em heurísticas ou geradas por técnicas de aprendizado (EICKE- LER; JABS; RIGOLL, 2000; YANG; JIN, 2006; CHEN et al., 2009). Segundo Duda et al. (DUDA; HART; STORK, 2001), a utilização de um classificador complexo pode prejudicar o processo de genera- lização, entretanto, a seleção de um limiar adequado é um dos principais problemas nesta área. Por outro lado, quando este processo visa determinar a semelhança entre amostras, empregam-se métricas que podem, ou não, levar em consideração a distribuição do conjunto de treinamento (PERLIBAKAS, 2004).

Em amostras processadas com o PCA habitualmente emprega-se um classificador de vi- zinho mais próximo, como por exemplo a distância Euclidiana. No entanto, a abstração obtida com o vetor de características permite o desenvolvimento de diversas teorias de classificação. Um estudo comparativo realizado por Perlibakas (PERLIBAKAS, 2004) propôs a implementação de uma

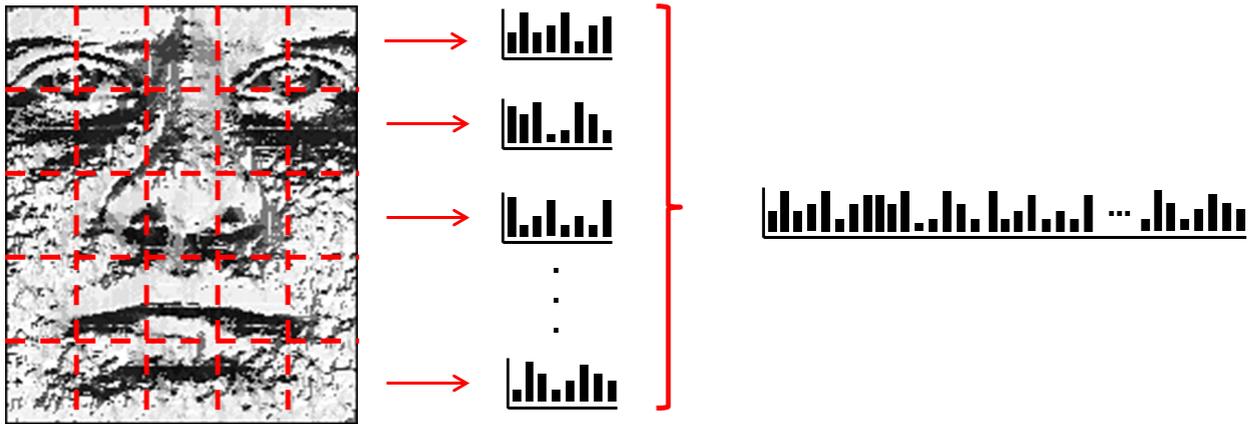


Figura 3.17: Descritores de textura independentes, concatenados para compôr o histograma global.

medida baseada no Erro Médio Quadrático e comparou 14 métricas e suas derivações, obtendo as melhores taxas de reconhecimento com a distância de Mahalanobis, ponderada pelos autovalores. No presente trabalho, adotaram-se as distâncias Euclidiana, Mahalanobis convencional, Mahalanobis ponderada e Erro Médio Quadrático modificado, descritas a seguir, considerando que $m \leq n$, x é a imagem de teste e y uma imagem da base de treinamento, ambas projetadas no subespaço gerado pelo PCA:

a) Euclidiana

$$d(x,y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2}; \quad (3.21)$$

b) Mahalanobis

$$d(x,y) = (x - y)S_2^{-1}(x - y)^T, \quad (3.22)$$

onde S_2 é a matriz de covariância da base de treinamento;

c) Mahalanobis ponderada

$$d(x,y) = - \sum_{i=1}^m x_i y_i \sqrt{1/\lambda_i}; \quad (3.23)$$

d) Erro médio quadrático modificado

$$d(x,y) = \frac{\sum_{i=1}^m \sqrt{1/\lambda_i} (x_i - y_i)^2}{\sqrt{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2}}. \quad (3.24)$$

Através do estudo realizado por Perlibakas (PERLIBAKAS, 2004), conclui-se que uma única abordagem é incapaz de resolver todos os problemas de classificação e a combinação de múltiplos classificadores, com base em aspectos distintos das amostras, pode melhorar a taxa de reconhecimento.

Para classificar as amostras geradas pelo LBP utilizam-se, basicamente, métricas que permitam explorar a relevância de algumas características faciais sobre outras, no processo de reconhecimento (ZHAO et al., 2003). Deste modo, é possível definir pesos específicos para cada região da grade de histogramas. Portanto, adotou-se neste trabalho a distância Qui-Quadrado, também empregada por Ahonen et. al. (AHONEN; HADID; PIETIKAINEN, 2004, 2006), definida abaixo:

$$\chi_w^2(x,y) = \sum_{i,j} w_j \frac{(x_{i,j} - y_{i,j})^2}{x_{i,j} + y_{i,j}}, \quad (3.25)$$

onde x e y são as matrizes de histogramas a serem comparadas, i é um elemento do histograma correspondente a j -ésima região e w_j é o peso pré-definido para a região j . Tal abordagem apresentou bons resultados nos estudos realizados por Ahonen et al. (AHONEN; HADID; PIETIKAINEN, 2004, 2006).

3.4 Pós-processamento

Esta etapa contempla as atividades de avaliação e análise de contexto, responsáveis por aprimorar o processo e adequar resultados imprecisos, respectivamente. A etapa de avaliação consiste na validação dos resultados e definição de ações paralelas. Supondo que uma amostra de teste seja identificada, a amostra de treinamento correspondente pode ser removida da base. Caso contrário, a amostra procurada pode ser incluída no conjunto de treinamento para futuras buscas. Na etapa de análise de contexto, o conhecimento *a priori* é utilizado como fator de decisão. Um resultado classificado como inválido pode ser substituído pelo resultado válido mais próximo, ou mais adequado às informações pré-estabelecidas (DUDA; HART; STORK, 2001). Considerando a hipótese de busca por uma pessoa do sexo feminino, qualquer pessoa do sexo masculino presente no resultado pode ser descartada, conforme exemplo apresentado na Figura 3.18:

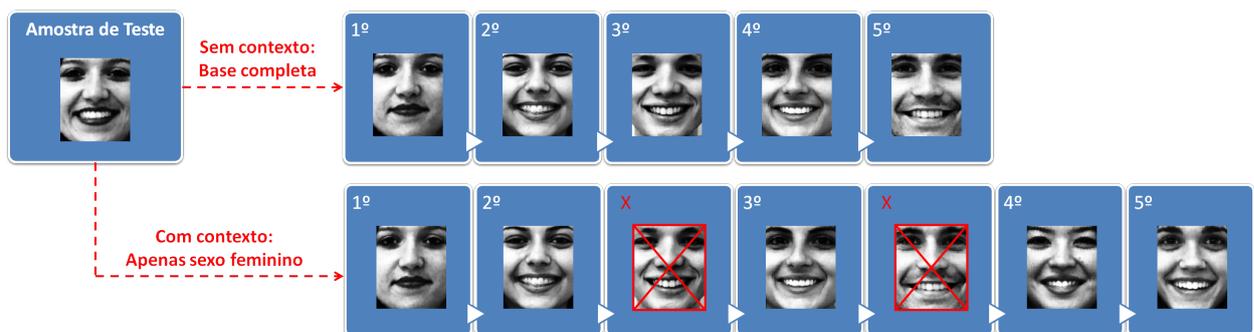


Figura 3.18: Filtragem dos resultados de acordo com o contexto da busca.

3.5 Considerações Complementares

Este capítulo descreveu o arcabouço principal de reconhecimento de padrões adotado na literatura afim para resolver problemas como o reconhecimento automático de faces, contexto de investigação deste estudo. Considerando a revisão de trabalhos realizada no capítulo anterior e a metodologia de normalização de (AMARAL; THOMAZ, 2008; AMARAL et al., 2009), fez-se uma descrição detalhada dos cálculos e métodos empregados para composição de todas as etapas do procedimento de reconhecimento automático de uma imagem de face, utilizando abordagens local e global de extração de características, baseadas respectivamente nas técnicas LBP e PCA. No próximo capítulo, descrevem-se os principais experimentos e resultados obtidos neste estudo utilizando bases públicas de imagens de faces.

4 EXPERIMENTOS E RESULTADOS

Neste capítulo são discutidos os resultados obtidos visando compreender os fenômenos envolvidos no processo de reconhecimento de imagens de faces. Com o propósito de analisar o desempenho de abordagens locais e globais, realizaram-se ensaios empíricos utilizando as técnicas LBP e PCA em três bases distintas. Todos experimentos realizados utilizaram scripts desenvolvidos em linguagem Python 2.7, utilizando as bibliotecas Numpy e Scipy.

4.1 Base de Imagens de Face

A definição das bases foi realizada levando-se em consideração as seguintes questões: permitir a comparação com outros estudos realizados nesta área; analisar o problema de restrição amostral; avaliar o processo de reconhecimento sob variações de expressão, pose e iluminação. Portanto, empregaram-se três conjuntos com características específicas, detalhadas a seguir:

- a) FEI Face Database: criada e mantida pelo Departamento de Engenharia Elétrica do Centro Universitário da FEI, esta base é composta por 14 imagens de 200 indivíduos, 100 homens e 100 mulheres, com variações de pose e iluminação, totalizando 2800 imagens, capturadas de acordo com o procedimento descrito por (JUNIOR; THOMAZ, 2006). Para este estudo empregaram-se 900 imagens de face, distribuídas em 5 grupos: 200 imagens frontais de faces com expressão neutra; 200 imagens frontais de faces com expressão sorrindo; 200 imagens de faces com variação de pose para a esquerda; 200 imagens de faces com variação de pose para a direita; e 100 imagens frontais de face com problemas de iluminação. Ilustrada na Figura 4.1.
- b) FERET: largamente empregada como parâmetro de avaliação e comparação em estudos de reconhecimento facial, esta base é composta por 14.126 imagens, de 1.199 indivíduos, com variações, de sexo, etnia, expressão facial e duplicidade com intervalos de captura. Neste estudo utilizaram-se 400 imagens, de 200 indivíduos, organizados em 2 grupos: 100 imagens frontais com expressão neutra e 100 imagens frontais com expressão sorrindo. Ilustrada na Figura 4.2.
- c) JAFFE: frequentemente aplicada a estudos sobre expressão facial, esta base é formada basicamente por mulheres japonesas e contém 213 imagens, de 10 pessoas distintas, apresentando 7 variações de expressão facial com duplicidades. Neste trabalho empregaram-se todas as imagens, com 1 expressão neutra e 6 variações para cada indivíduo. Ilustrada na Figura 4.3.



Figura 4.1: Imagens da base FEI, contendo variações de expressão, pose e iluminação.



Figura 4.2: Imagens da base FERET, contendo uma imagem neutra e uma sorrindo por pessoa.



Figura 4.3: Imagens da base JAFFE, contendo uma imagem neutra e seis variações de expressão por pessoa .

4.1.1 Normalização das amostras

As imagens empregadas nos experimentos deste trabalho foram normalizadas de acordo com a metodologia proposta por Amaral et al. (AMARAL et al., 2009), um exemplo dessas amostras pode ser observado através da Figura 4.4, onde as três bases utilizadas apresentam as mesmas dimensões e o mesmo aspecto visual.



Figura 4.4: Imagens normalizadas das bases FEI, FERET e JAFFE.

4.2 Projeção de auto-faces

Com o propósito de analisar a distribuição das amostras no subespaço gerado pelo PCA e identificar as características mais variantes em cada base, realizou-se o deslocamento da imagem média sobre o eixo de maior variância, para cada conjunto de processado, dentro dos limites $\pm 3\sqrt{\lambda_i}$. A Figura 4.5 ilustra esta análise.

Através da projeção no subespaço de características pode-se observar basicamente variações de expressão nas bases FEI e JAFFE e etnia na base FERET. De acordo com os resultados obtidos com as bases FEI e JAFFE conclui-se que a componente de maior variância descreve as características mais representativas de cada base. No entanto, a base FERET não é composta simetricamente por pessoas negras e caucasianas e mesmo assim as diferenças étnicas foram claramente mapeadas, isso ocorre pelo fato de que o PCA utiliza como variáveis de entrada o valor dos pixels de cada imagem em escala de cinza. Portanto, a tonalidade da pele afeta diretamente a distribuição das amostras. Além disso, cada autovetor não é responsável por apenas uma característica distinta



Figura 4.5: De cima para baixo, deslocamento da imagem média sobre o eixo de maior variância gerado pelo PCA sobre a base FEI, FERET e JAFFE.

e em uma análise mais detalhada identificam-se, também, variações de gênero nas bases FEI e FERET e tamanho da face na base JAFFE.

4.3 Reconhecimento automático de imagens de faces

Nos experimentos de reconhecimento automático de faces utilizaram-se duas metodologias para avaliação das técnicas empregadas. Com o intuito de comparar os resultados obtidos com a literatura afim adotou-se o protocolo FERET (PHILLIPS et al., 2000) , onde o conjunto de treinamento é gerado apenas com imagens frontais neutras e os testes são realizados com demais amostras presentes em cada base, neste trabalho está abordagem será denominada homogênea. No segundo modelo empregou-se a técnica de validação cruzada *leave-one-out*, a fim de utilizar um conjunto de treinamento heterogêneo, comum em ambientes reais. Para medir a precisão dos métodos analisados gerou-se uma taxa de classificação cumulativa, considerando o reconhecimento de uma amostra de teste ao longo das cinquenta amostras de treinamento mais semelhantes. Também é válido salientar que neste estudo a base FERET é a única que caracteriza o problema de uma amostra de treinamento por pessoa no ambiente heterogêneo. Portanto, quando necessário, analisou-se esta questão de modo particular.

4.3.1 Experimentos com LBP

Nestes experimentos avaliaram-se os parâmetros da técnica LBP e a taxa de reconhecimento utilizando regiões específicas da face. Empregaram-se, três grades distintas e sete padrões de regiões faciais para a extração dos histogramas, conforme apresentados nas figuras 4.6 e 4.7, respectivamente. Utilizaram-se, também, em todos os testes com LBP, raios de vizinhança 1 e 2, conforme ilustrado na Figura 4.8. Os resultados obtidos entre estes parâmetros de pré-processamento não apresentaram diferenças relevantes. No entanto, a definição de uma grade adequada se mostrou fundamental neste processo.

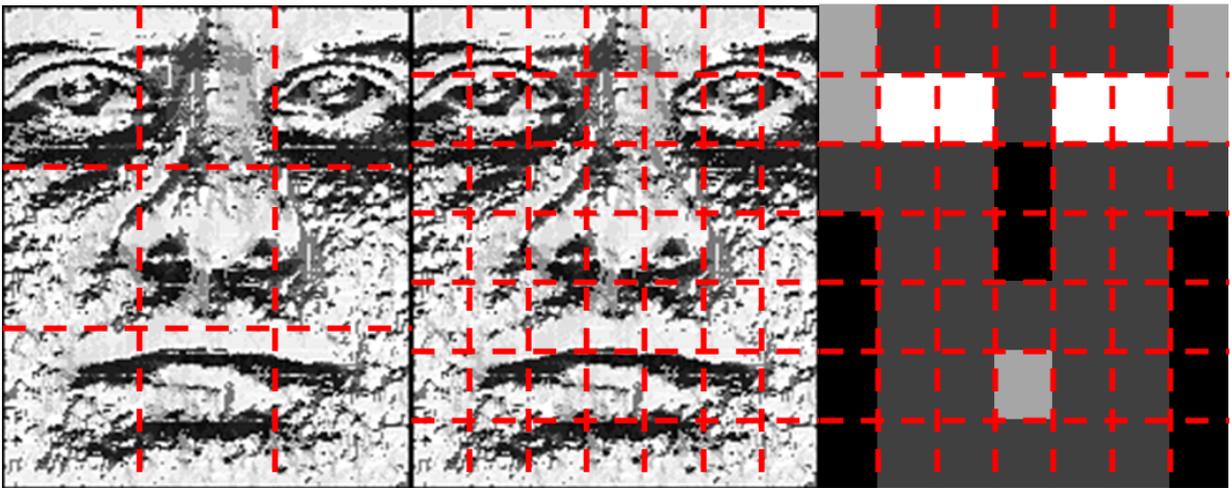


Figura 4.6: Da esquerda para direita, grades 3x3, 7x7 e 7x7 ponderada. As regiões preta, cinza escuro, cinza claro e branco, receberam os pesos 0, 1, 2 e 4, respectivamente.

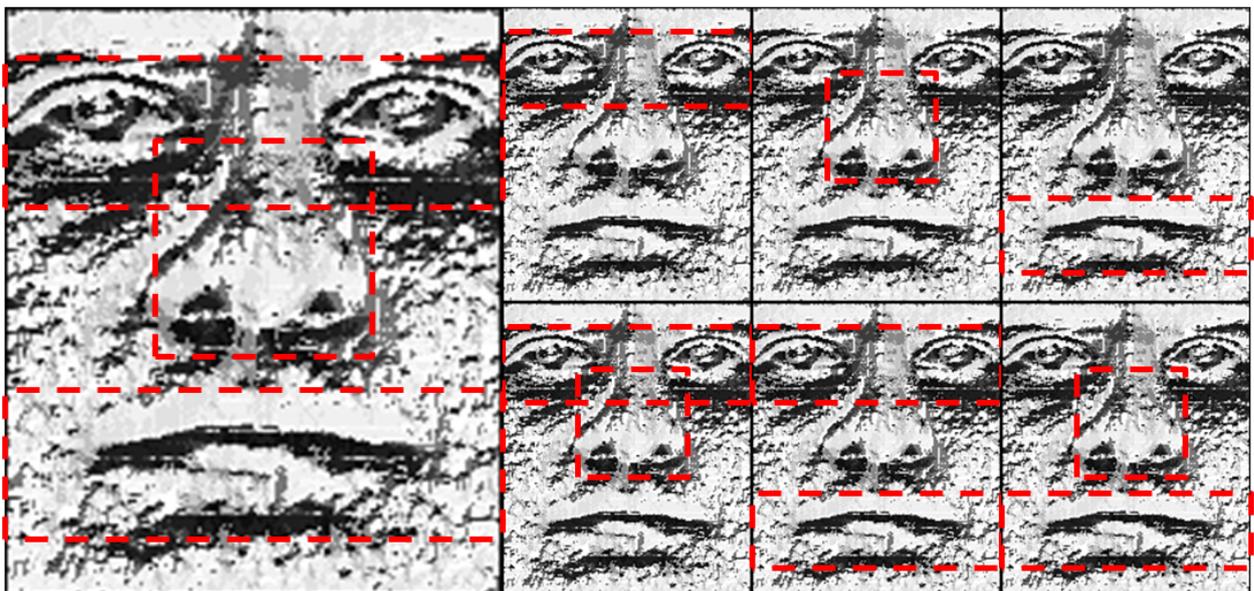


Figura 4.7: Padrões de regiões faciais empregadas no reconhecimento com LBP.

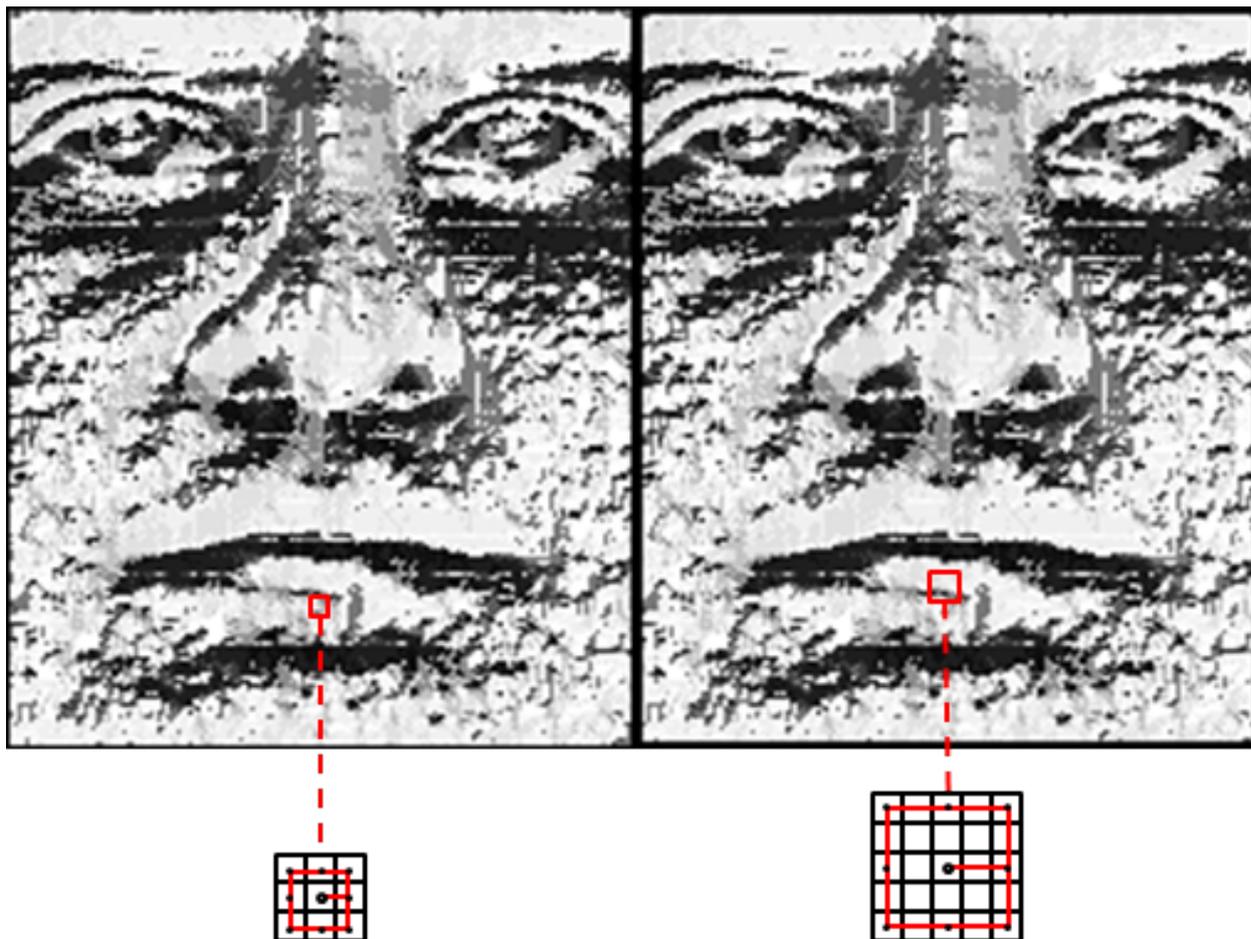


Figura 4.8: Imagens pré-processadas pelo LBP com raios de vizinhança 1 e 2.

No ensaio inicial, compararam-se as configurações para grades de extração dos histogramas locais utilizando os padrões 3x3, 7x7 e 7x7 ponderada. Em ambos modelos analisados, homogêneo e heterogêneo, observaram-se resultados superiores com as grades 7x7, com e sem ponderação. Contudo, no ambiente heterogêneo a grade ponderada foi superior ao mesmo layout sem ponderação, para as bases FEI e FERET, este resultado pode ser observado nas figuras 4.9, 4.10, 4.11 e 4.12. Isso ocorre por que tais ambientes possuem amostras de treinamento com variações de expressão, pose e iluminação. Esta característica torna ainda mais necessário o uso de pesos específicos para cada região facial, visando reduzir erros gerados por desvios nas imagens. Em relação à base JAFFE a taxa de reconhecimento foi altamente influenciada pela grande quantidade de amostras por pessoa. Portanto, os resultados obtidos não favorecem uma análise apurada das grades, conforme ilustrado nas figuras 4.13 e 4.14. Contudo, estudou-se este caso com o propósito de enfatizar os benefícios da utilização de múltiplas imagens de cada indivíduo em situações com grande variação amostral.

Em seguida verificou-se a taxa de classificação para regiões faciais específicas empregando-se os olhos, o nariz e a boca, individualmente. Os resultados gerados com as bases FEI e FERET,

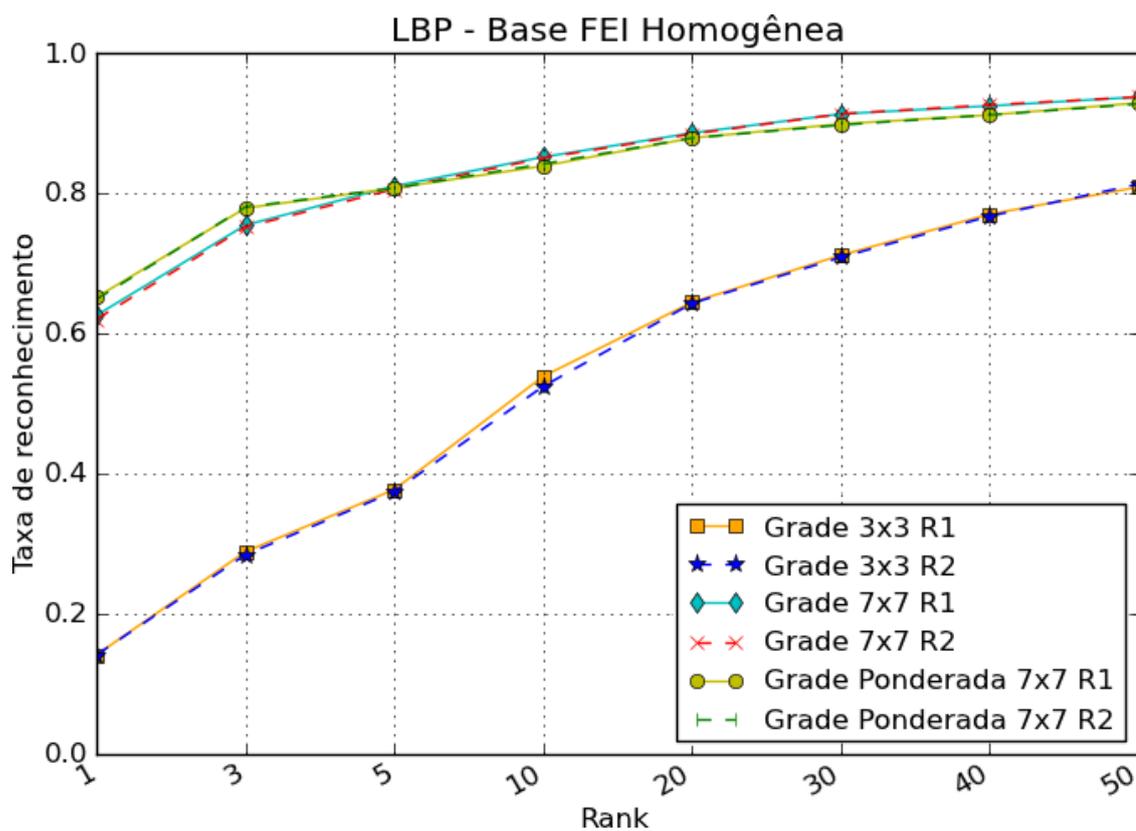


Figura 4.9: Comparação de grades na base FEI homogênea.

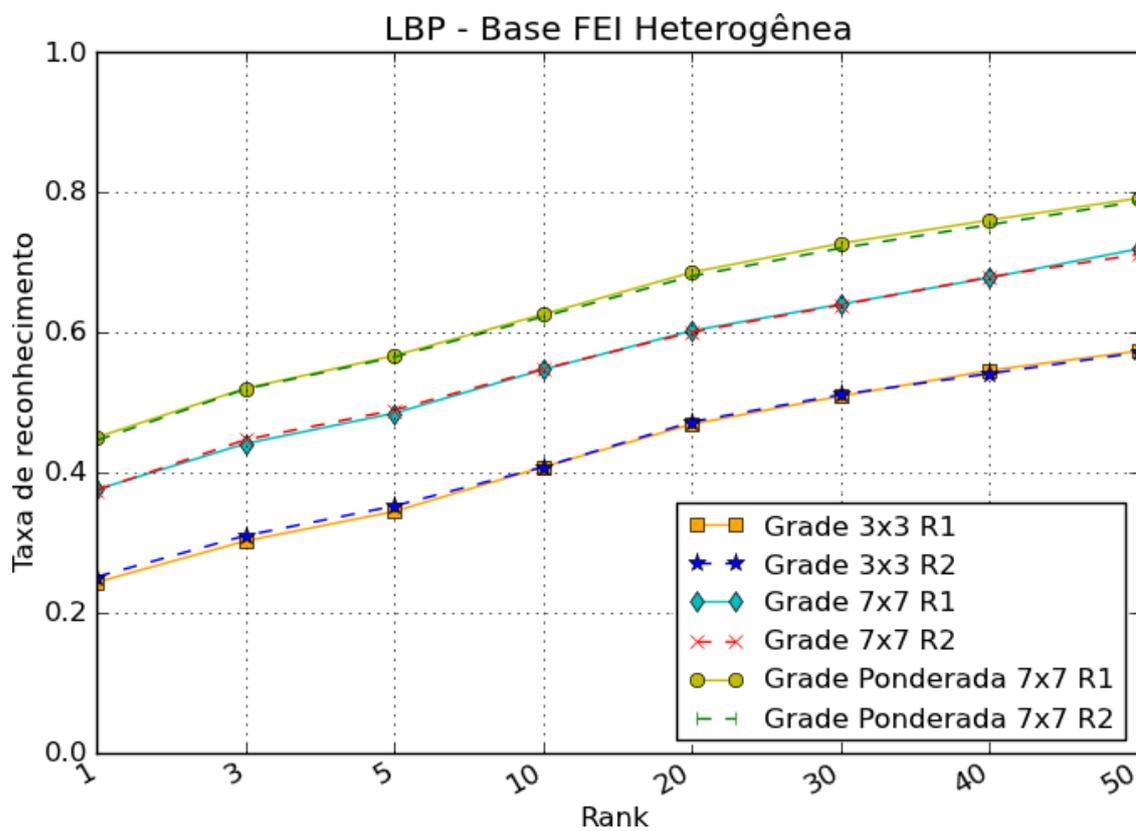


Figura 4.10: Comparação de grades na base FEI heterogênea.

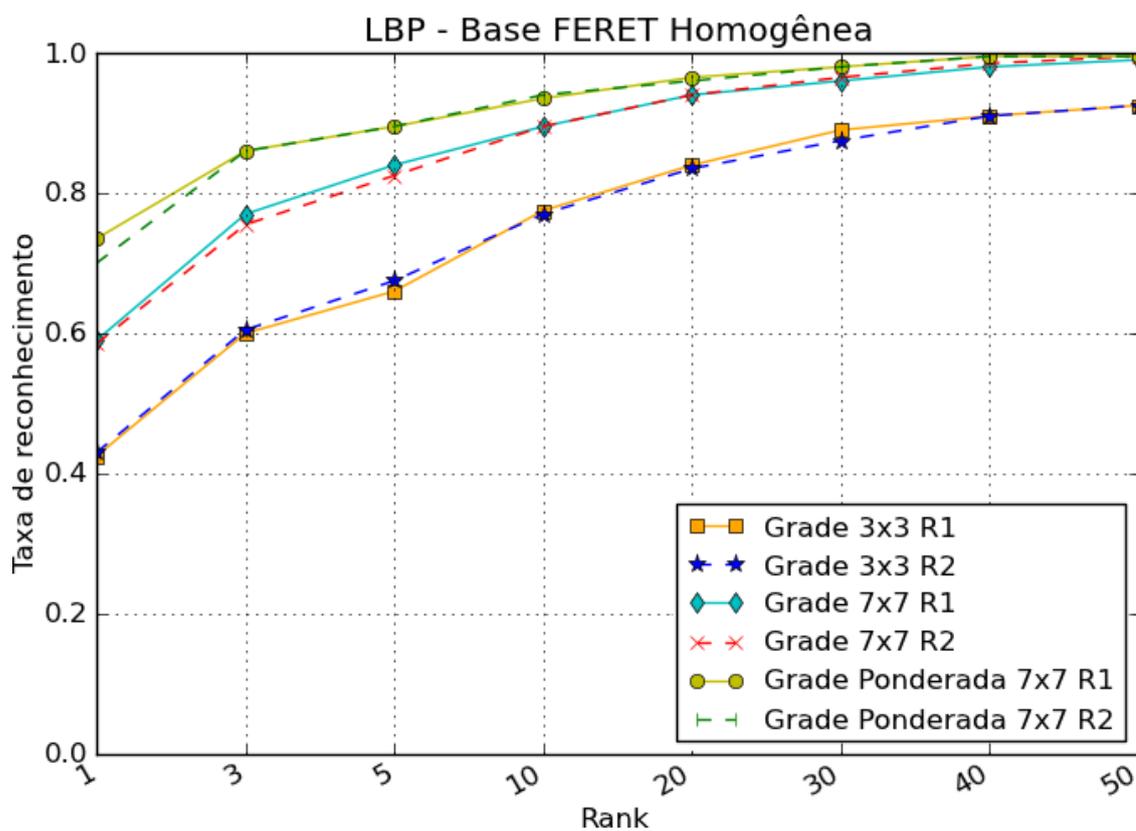


Figura 4.11: Comparação de grades na base FERET homogênea.

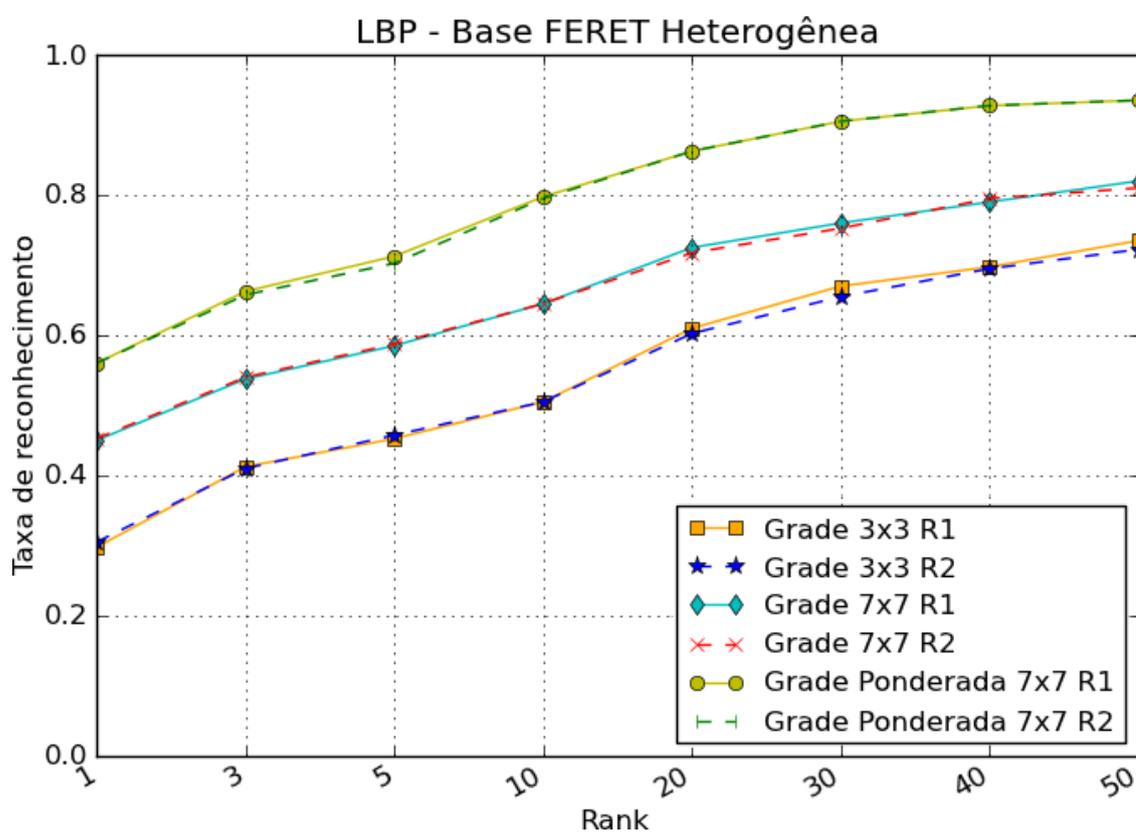


Figura 4.12: Comparação de grades na base FERET heterogênea.

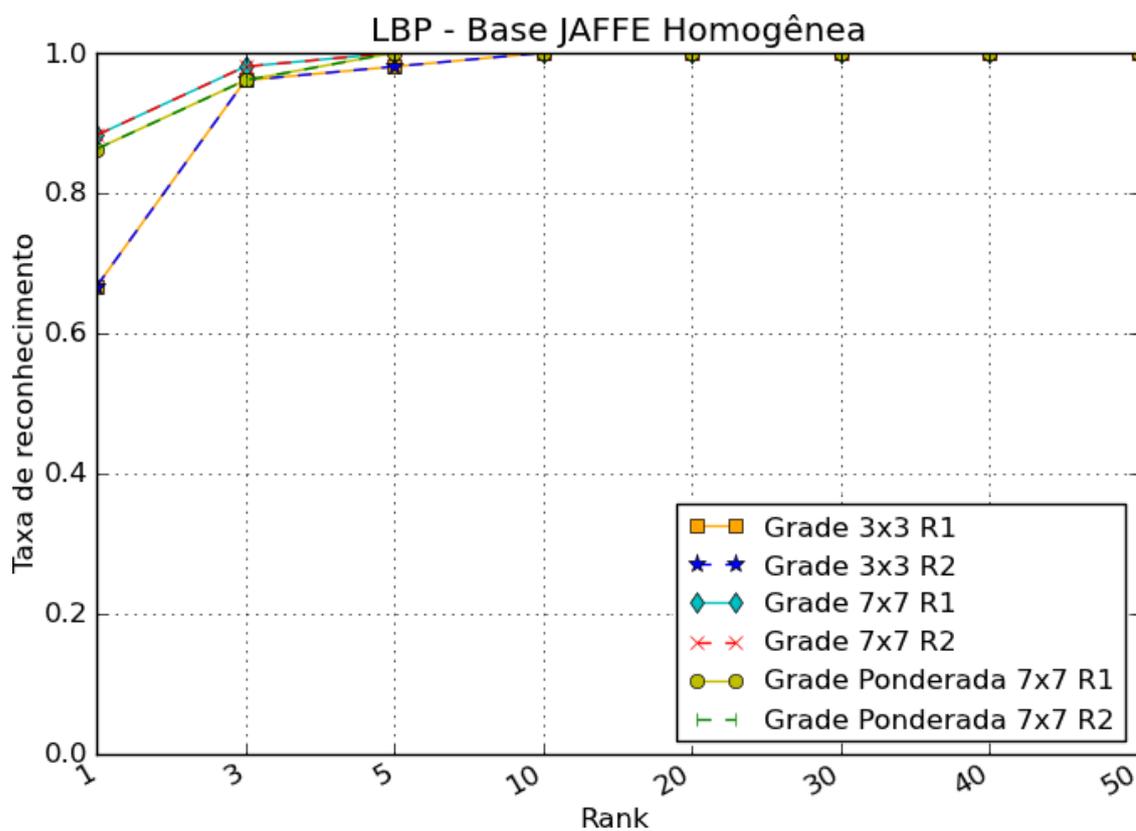


Figura 4.13: Comparação de grades na base JAFFE homogênea.

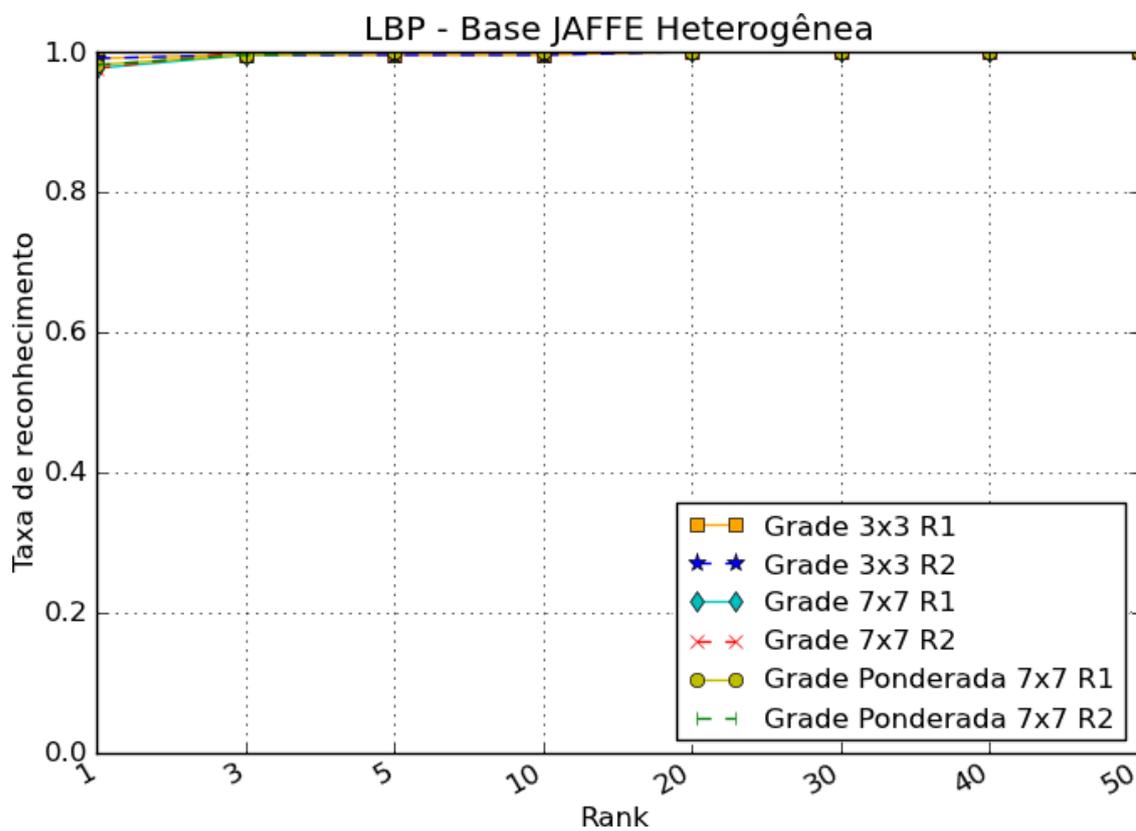


Figura 4.14: Comparação de grades na base JAFFE heterogênea.

destacaram a ineficiência da região bucal no processo de reconhecimento facial, conforme apresentado nas figuras 4.15, 4.16, 4.17 e 4.18. Pois esta área é a mais afetada pelas variações de expressão presentes nas amostras. Para as regiões dos olhos e do nariz, a base FEI apresentou resultados muito próximos em ambiente homogêneo e em ambiente heterogêneo a região do nariz foi superior, devido à estabilidade sob variações de expressão. Entretanto, para a base FERET obteve-se resultados superiores com a região dos olhos para os dois modelos analisados, homogêneo e heterogêneo. A base JAFFE apresentou um comportamento distinto, devido à sua estrutura baseada em variações de expressão intensas, no ambiente homogêneo apenas a região do nariz se destacou das demais, porém, no ambiente heterogêneo todas regiões tiveram um desempenho similar, de acordo com as figuras 4.19 e 4.20.

Analisou-se, também, a composição de padrões, utilizando as regiões especificadas no experimento anterior, através das seguintes combinações: olhos e nariz; olhos e boca; nariz e boca; olhos, nariz e boca. Com exceção da base JAFFE, todos testes demonstraram bom desempenho com a combinação das regiões do olhos e do nariz. Este resultado enfatiza a desvantagem de se empregar a região bucal no processo de reconhecimento automático de imagens de faces com variações de expressão, em ambientes homogêneos ou heterogêneos. Os resultados desta análise podem ser observados nas figuras 4.21, 4.22, 4.23, 4.24, 4.25 e 4.26.

Os experimentos desta seção permitiram definir os melhores parâmetros para utilização da técnica LBP no contexto analisado por este trabalho. Observou-se que os raios de vizinhança empregados apresentam uma diferença irrelevante para a aplicação deste estudo. Por outro lado, a redução das dimensões da grade melhora consideravelmente o desempenho do classificador, pois aumenta o grau de detalhamento do histograma gerado. Em relação à grade ponderada verificou-se um acréscimo na taxa de reconhecimento, principalmente em ambientes heterogêneos, devido à valorização de regiões mais discriminantes e a suavização de desvios amostrais. Contudo, os resultados obtidos no ambiente heterogêneo são muito baixos e ficam aquém das taxas geradas no ambiente homogêneo e apresentadas na literatura afim. Por fim, os ensaios acerca das regiões faciais específicas são favoráveis ao uso exclusivo dos olhos e do nariz como elemento de comparação no processo de reconhecimento automático de faces através da técnica LBP.

4.3.2 Experimentos com PCA

As análises realizadas com PCA foram conduzidas com o propósito de identificar as melhores métricas e a quantidade mínima de componentes necessárias para o reconhecimento de imagens de faces. Investigou-se, ainda, a capacidade de discriminação das amostras pré-processadas pelo operador binário do LBP através do subespaço gerado pelo PCA.

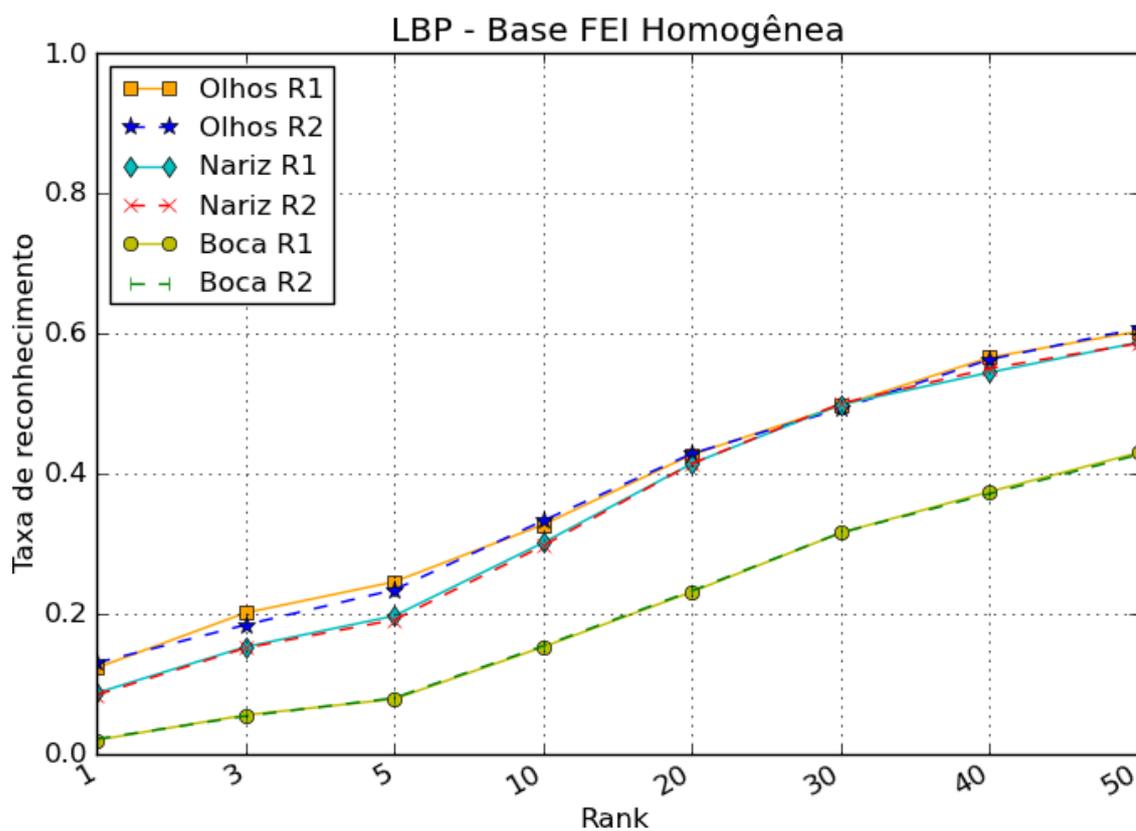


Figura 4.15: Comparação de regiões faciais na base FEI homogênea.

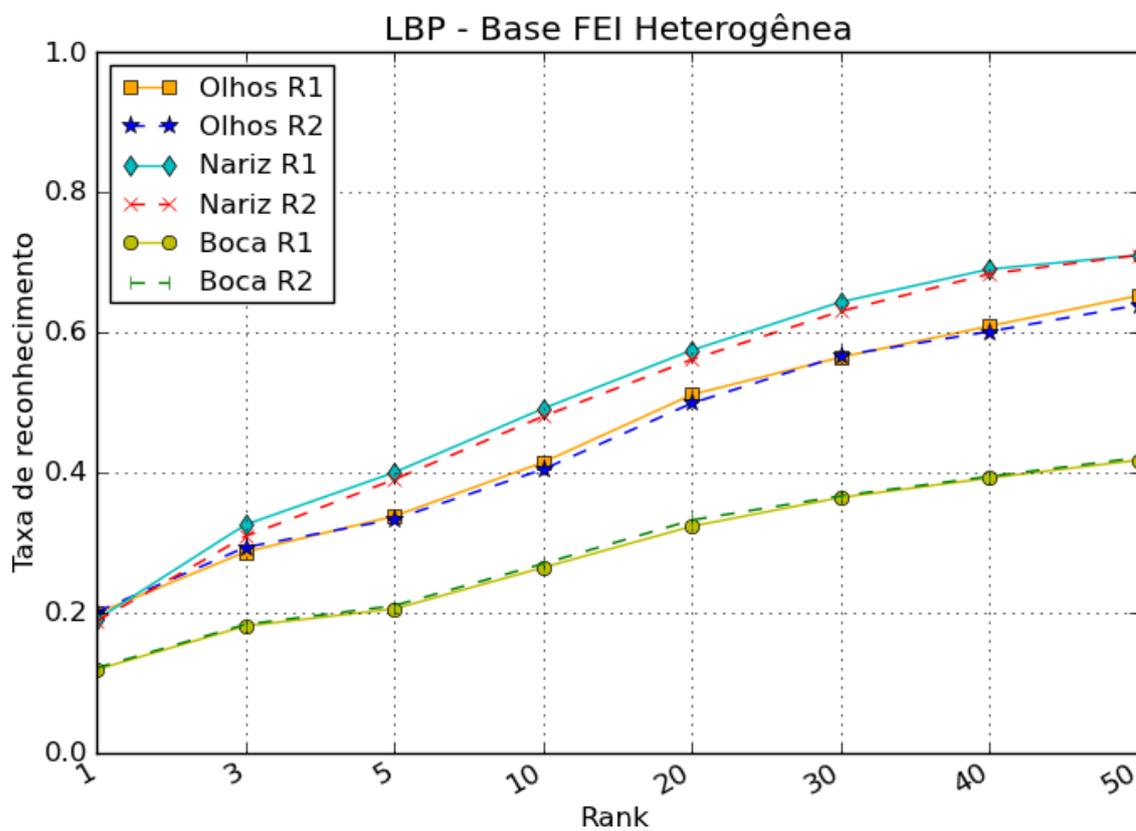


Figura 4.16: Comparação de regiões faciais na base FEI heterogênea.

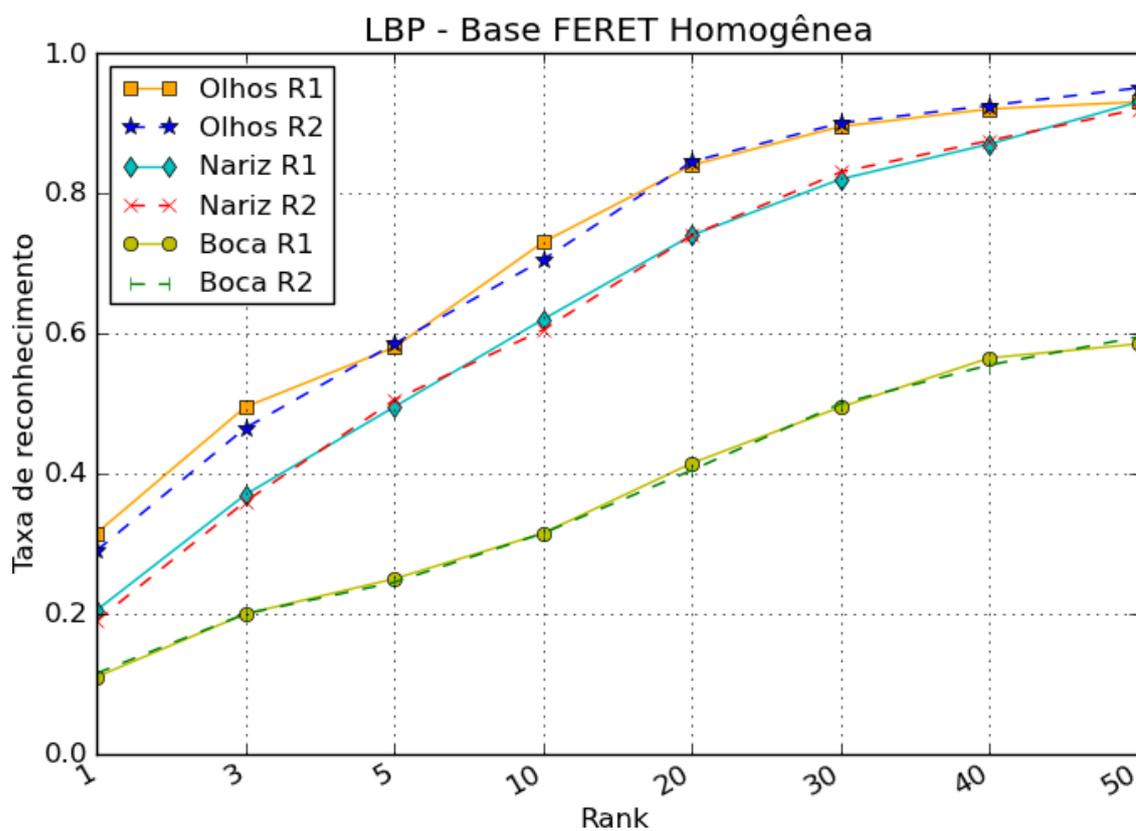


Figura 4.17: Comparação de regiões faciais na base FERET homogênea.

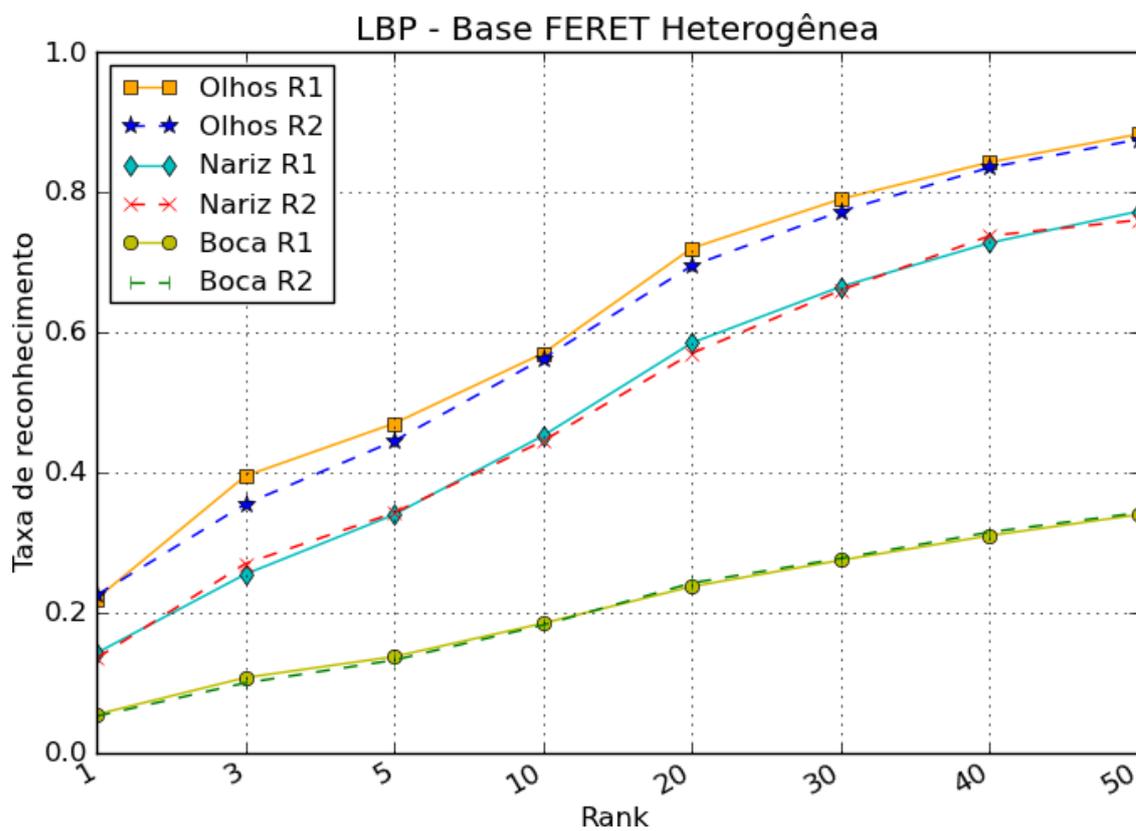


Figura 4.18: Comparação de regiões faciais na base FERET heterogênea.

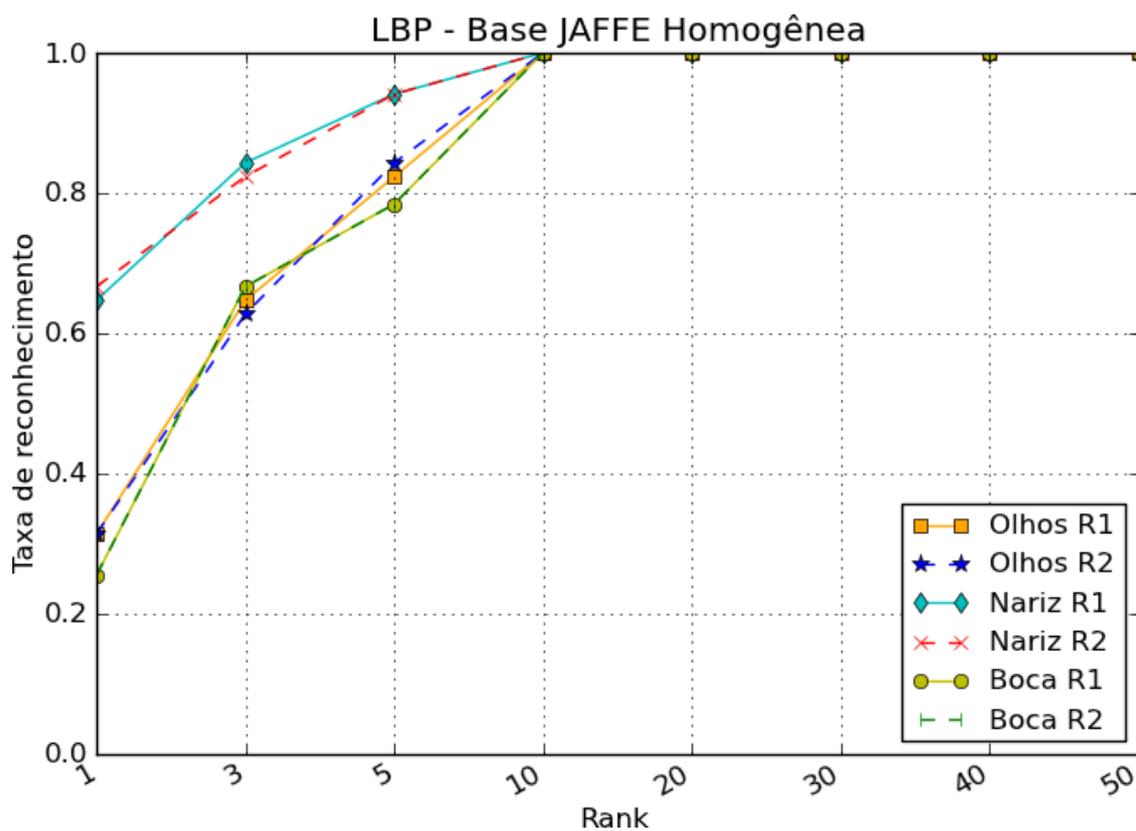


Figura 4.19: Comparação de regiões faciais na base JAFFE homogênea.

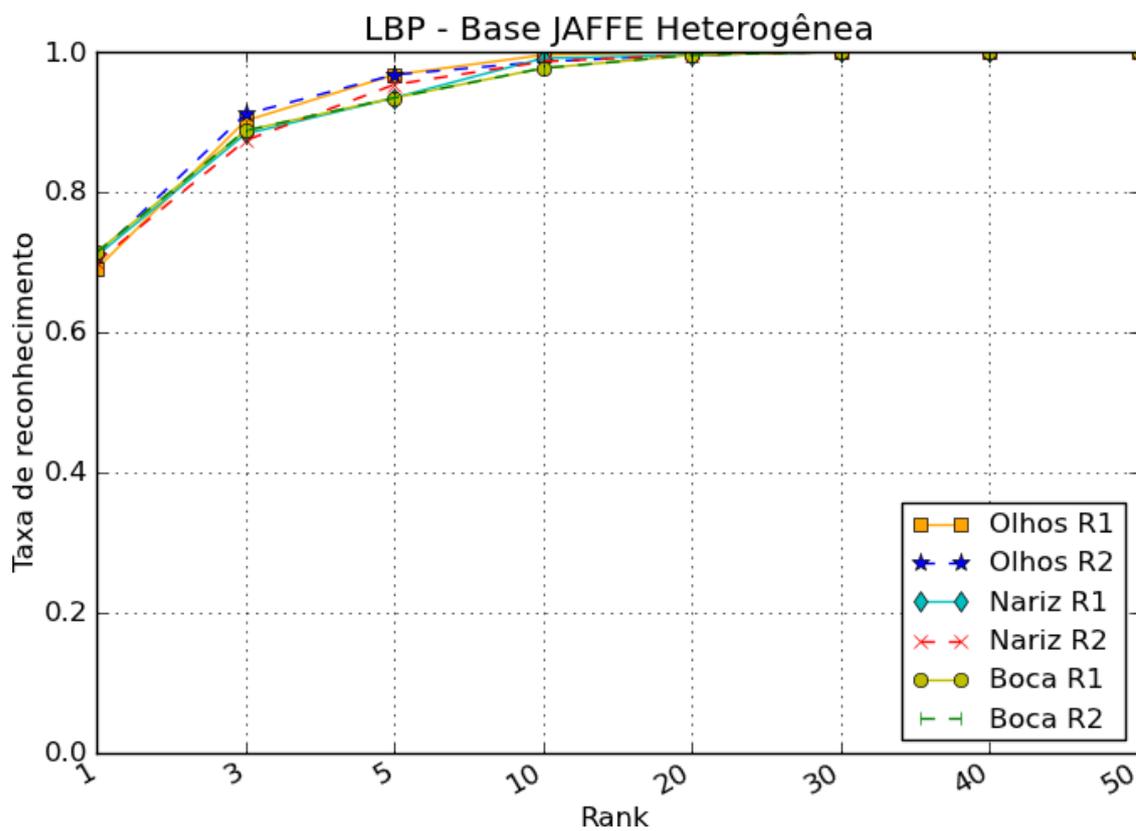


Figura 4.20: Comparação de regiões faciais na base JAFFE heterogênea.

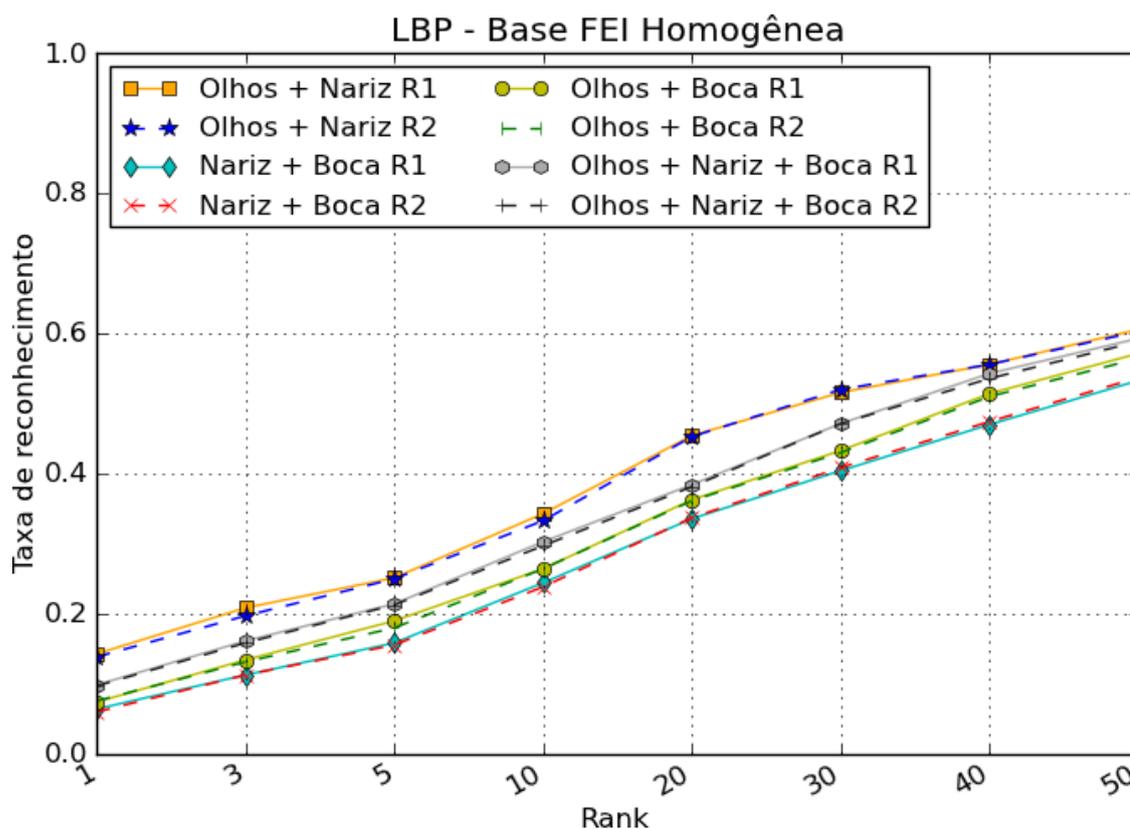


Figura 4.21: Comparação de combinações das regiões faciais na base FEI homogênea.

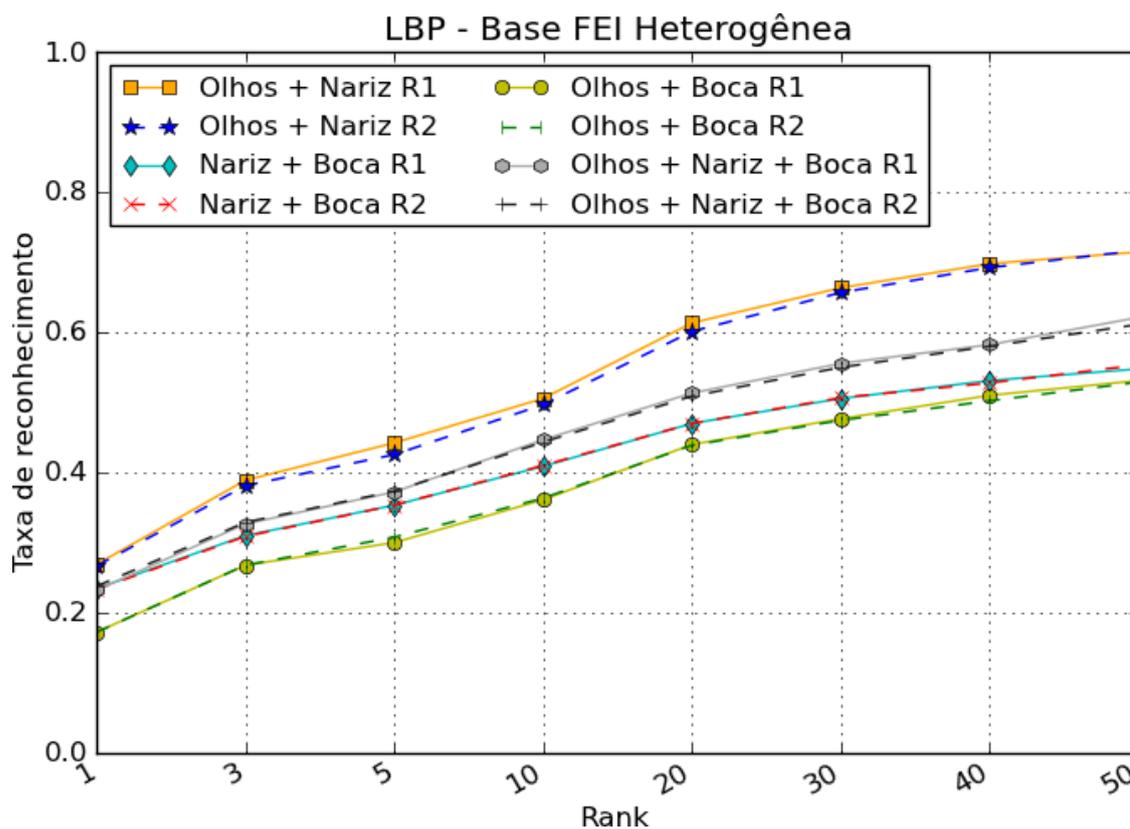


Figura 4.22: Comparação de combinações das regiões faciais na base FEI heterogênea.

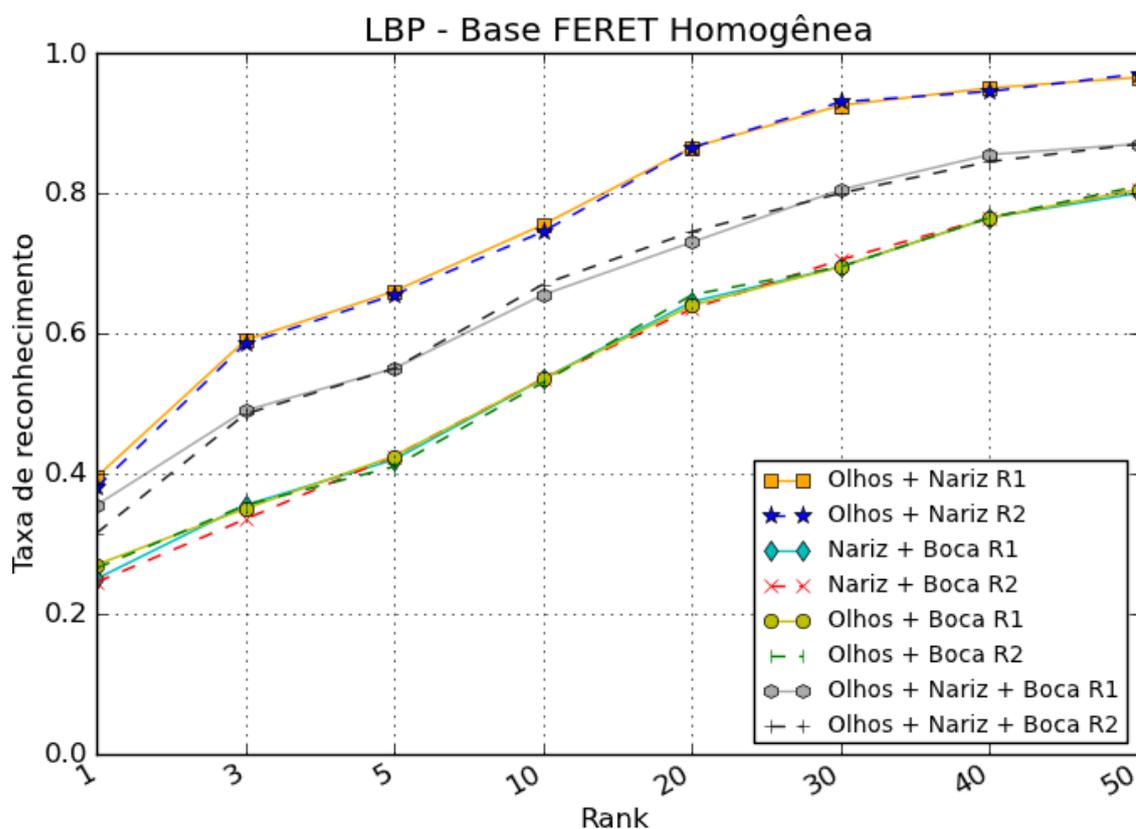


Figura 4.23: Comparação de combinações das regiões faciais na base FERET homogênea.

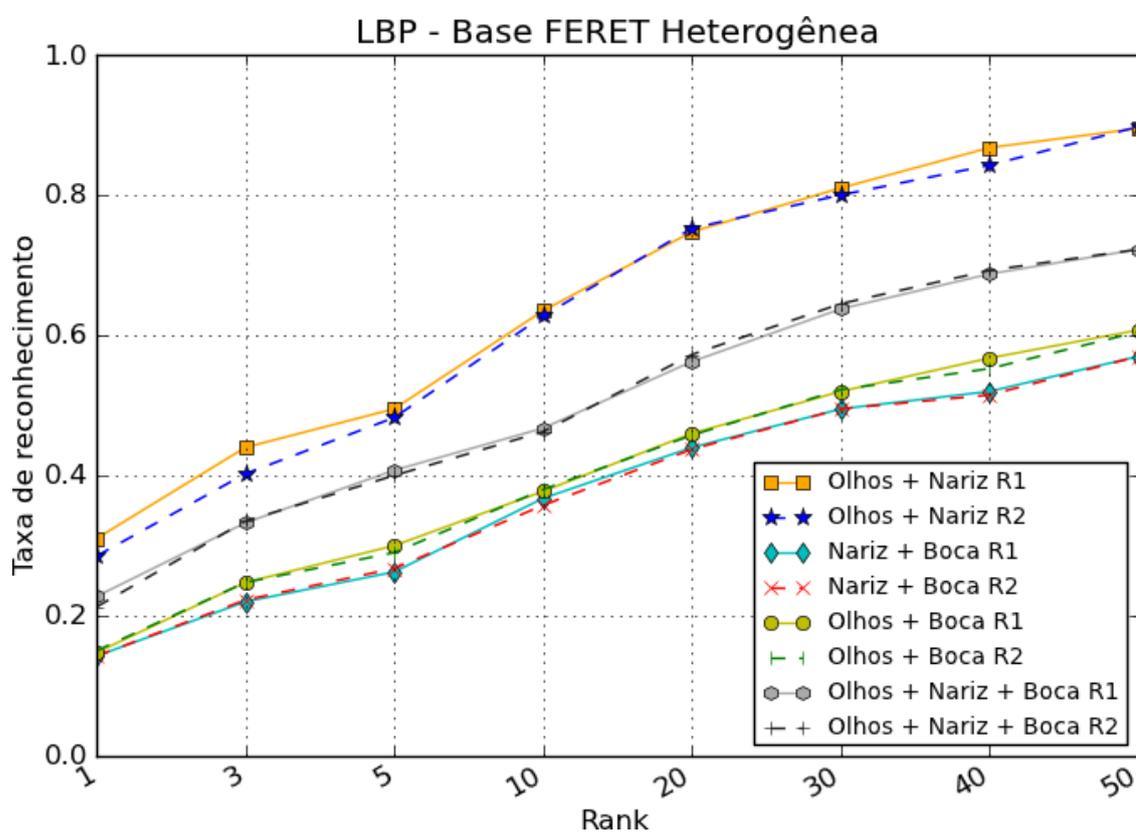


Figura 4.24: Comparação de combinações das regiões faciais na base FERET heterogênea.

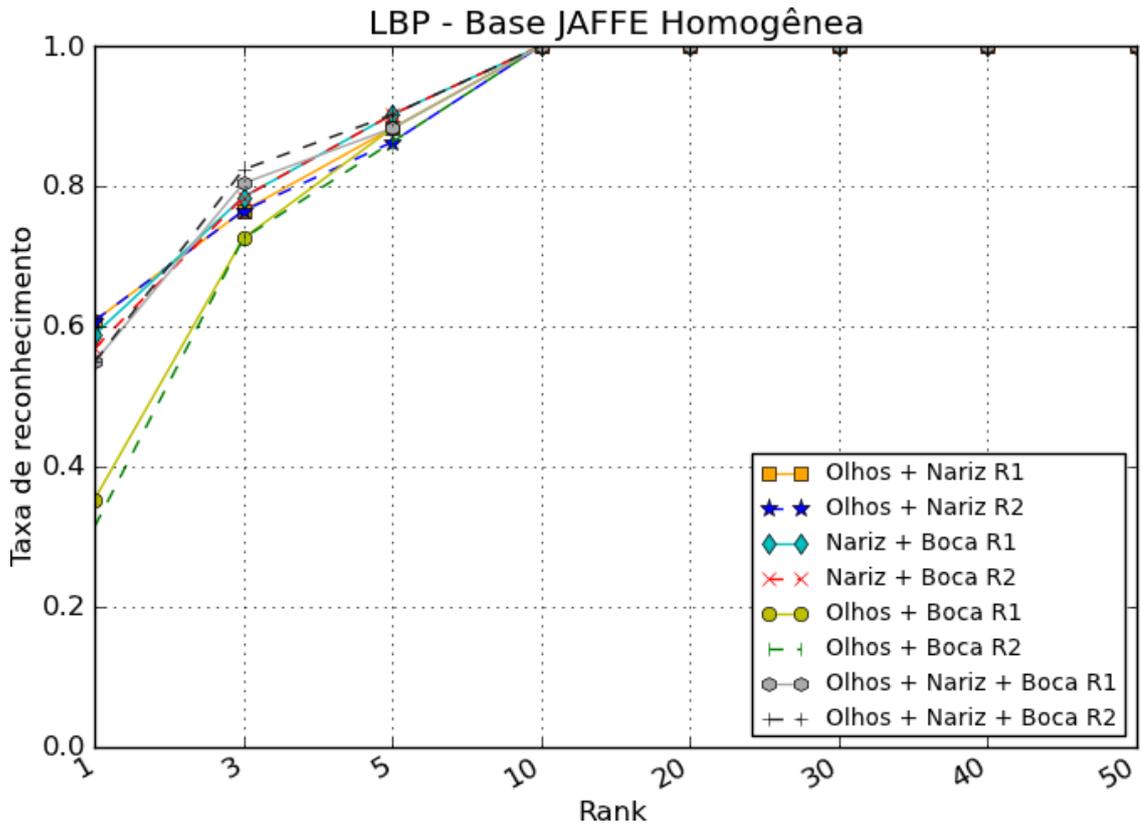


Figura 4.25: Comparação de combinações das regiões faciais na base JAFFE homogênea.

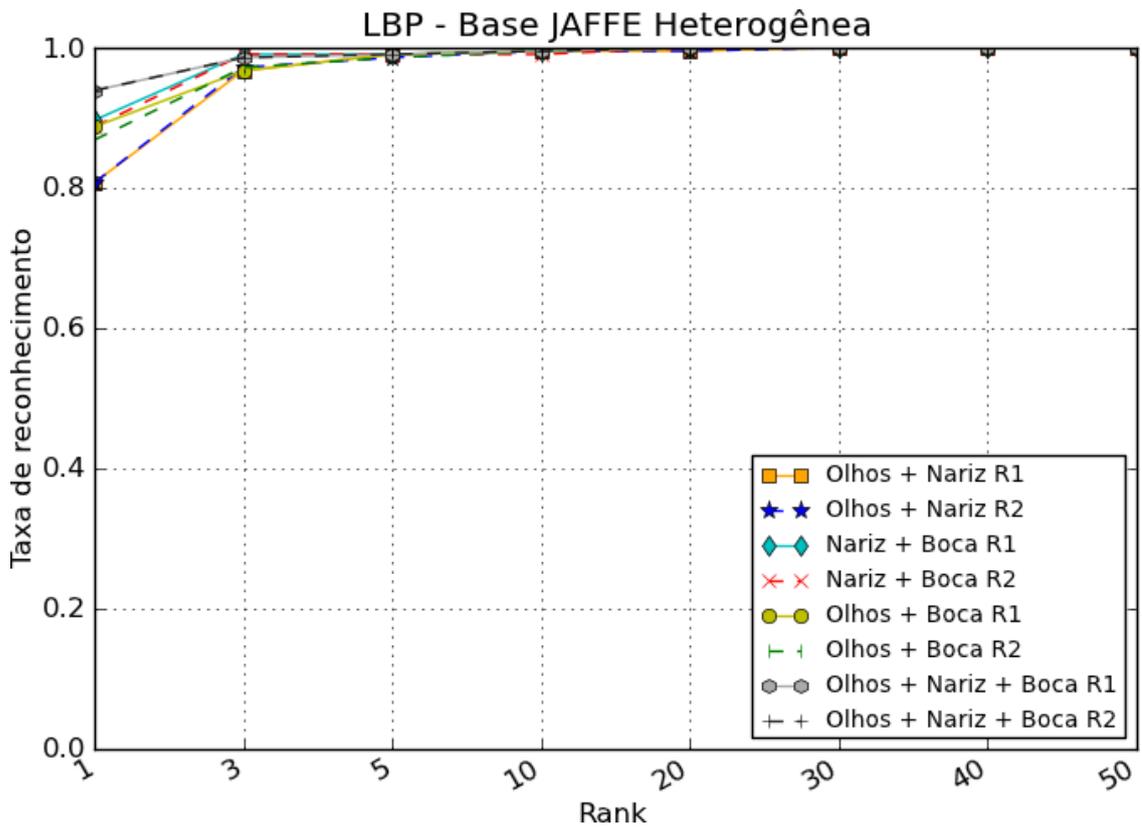


Figura 4.26: Comparação de combinações das regiões faciais na base JAFFE heterogênea.

Inicialmente, verificou-se a taxa de classificação das distâncias Euclidiana, Mahalanobis, Erro Médio Quadrático modificado e Mahalanobis ponderada, para quantidades distintas de componentes principais e precisão de acerto restrita à amostra mais semelhante. Em geral, para as três bases, obtiveram-se os melhores resultados com as métricas Mahalanobis e Mahalanobis ponderada, utilizando todas as componentes possíveis, conforme ilustrado nas figuras 4.27, 4.28, 4.29, 4.29, 4.31 e 4.32. Observou-se que, em ambientes homogêneos as métricas apresentaram crescimento contínuo com a adição de componentes. Entretanto, em ambientes heterogêneos as distâncias Mahalanobis e Mahalanobis ponderada demonstraram bom desempenho com um número reduzido de componentes para as bases FEI e FERET, respectivamente.

Conforme proposto, no experimento seguinte empregaram-se as amostras pré-processadas pelo operador binário do LBP como parâmetro de entrada para os cálculos do PCA, utilizando raios de vizinhança 1 e 2. Em geral, os resultados foram inferiores a abordagem original do PCA, para todas bases e ambientes e os raios de vizinhança novamente não apresentaram diferenças relevantes. Contudo, neste ensaio obtiveram-se as melhores taxas de reconhecimento para a distância de Erro Médio Quadrático modificado. As figuras 4.33, 4.34, 4.35, 4.36, 4.37 e 4.38 ilustram esta análise.

4.3.3 Comparação entre LBP e PCA

Utilizando como parâmetro os melhores resultados obtidos nos experimentos individuais pode-se comparar as técnicas LBP e PCA para bases e modelos de avaliação distintos. Portanto, empregaram-se as grades 7×7 e 7×7 ponderada para a técnica LBP e as métricas Mahalanobis e Mahalanobis ponderada, com todas componentes, para o PCA. Mantiveram-se as taxas para os raios de vizinhança 1 e 2 em carácter ilustrativo. Em geral, o desempenho do PCA foi melhor que o LBP. Considerando-se o índice cumulativo de precisão, em ambientes homogêneos os resultados de ambas técnicas foram semelhantes. No entanto, para ambientes heterogêneos o PCA apresentou taxas de reconhecimento superiores com as duas métricas propostas para esta análise. Estes resultados podem ser observados nas figuras 4.39, 4.40, 4.41, 4.41, 4.43 e 4.44.

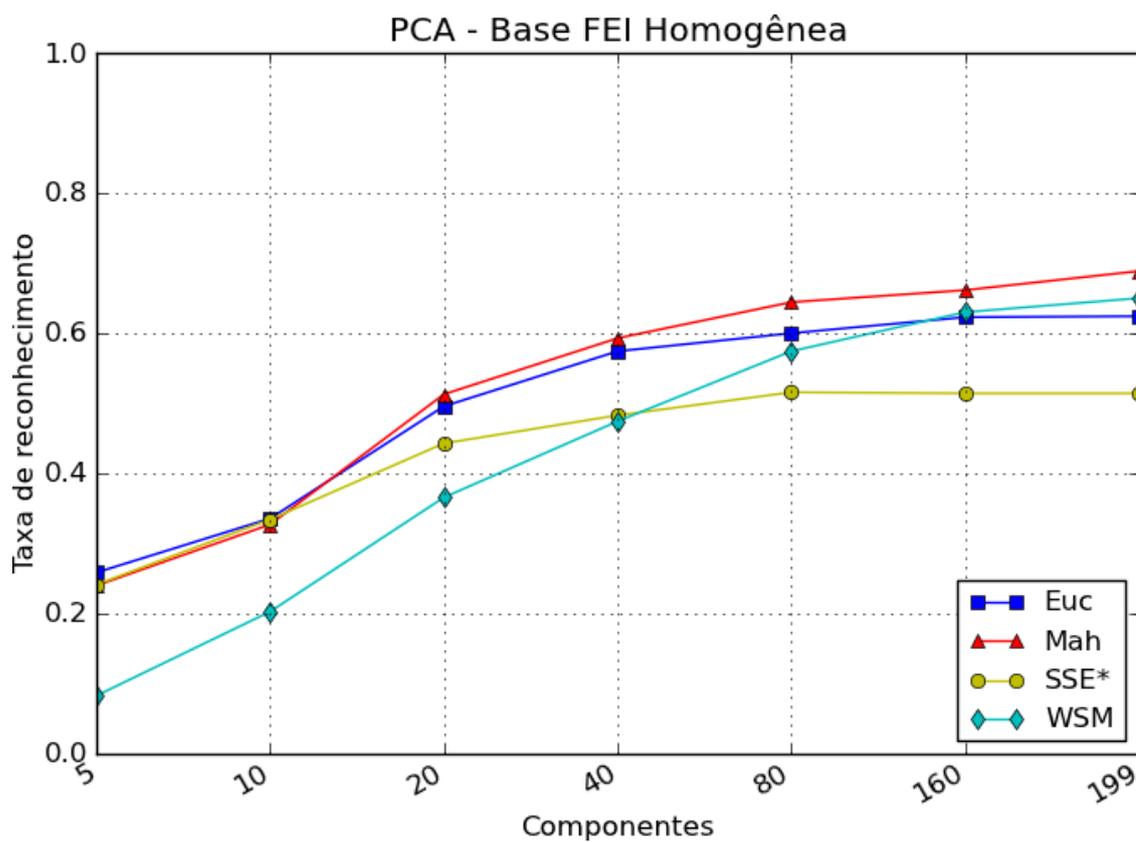


Figura 4.27: Comparação de métricas na base FEI homogênea.

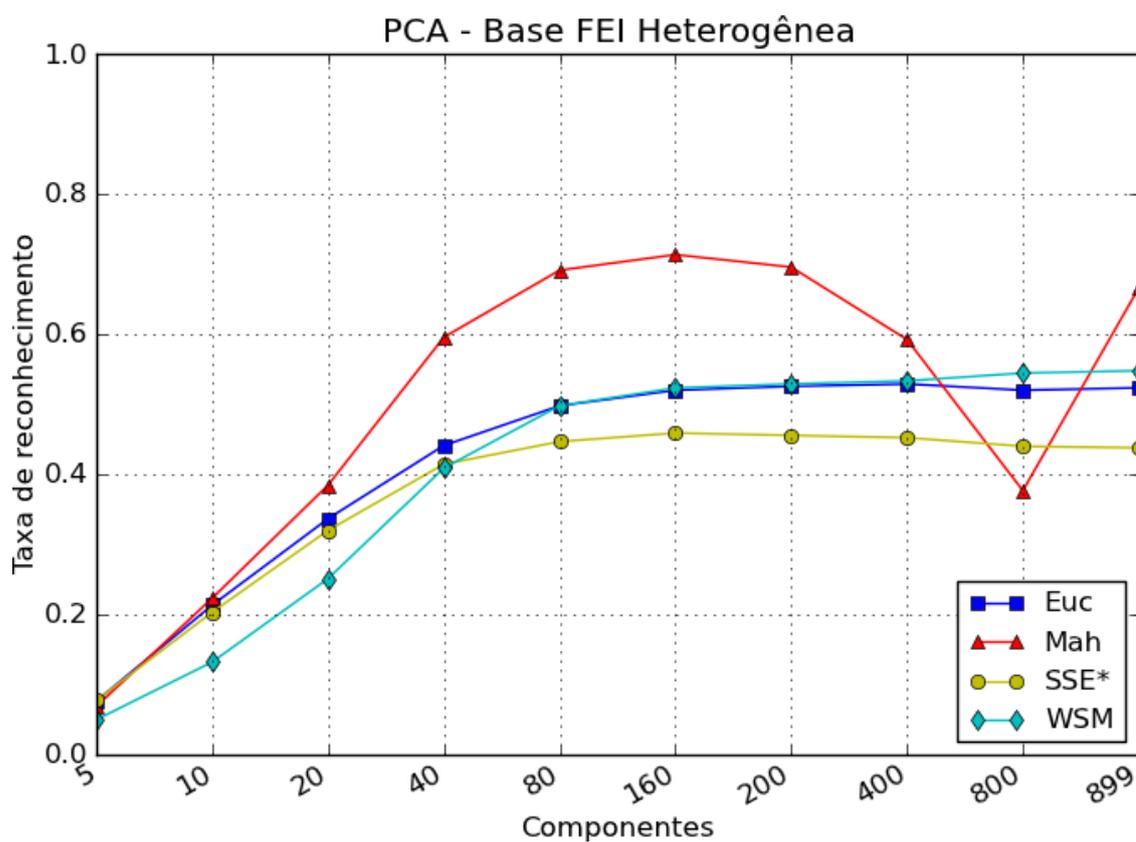


Figura 4.28: Comparação de métricas na base FEI heterogênea.

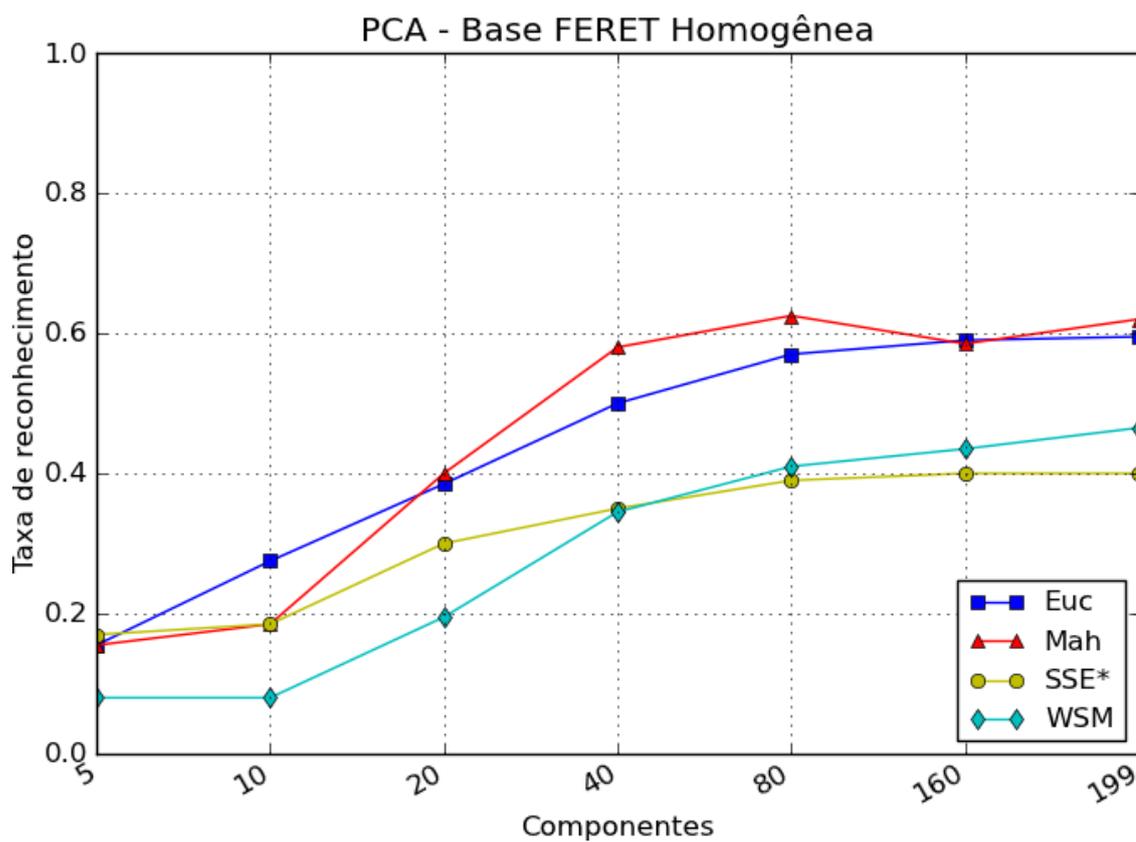


Figura 4.29: Comparação de métricas na base FERET homogênea.

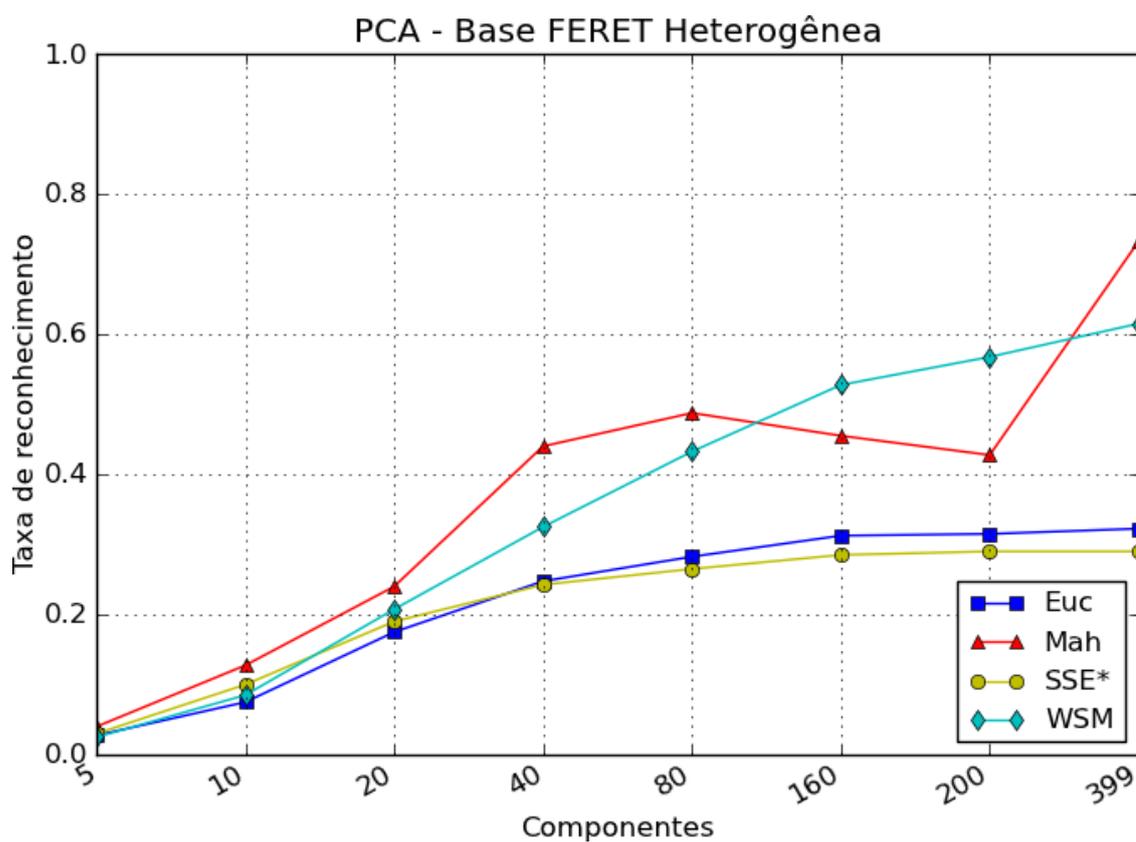


Figura 4.30: Comparação de métricas na base FERET heterogênea.

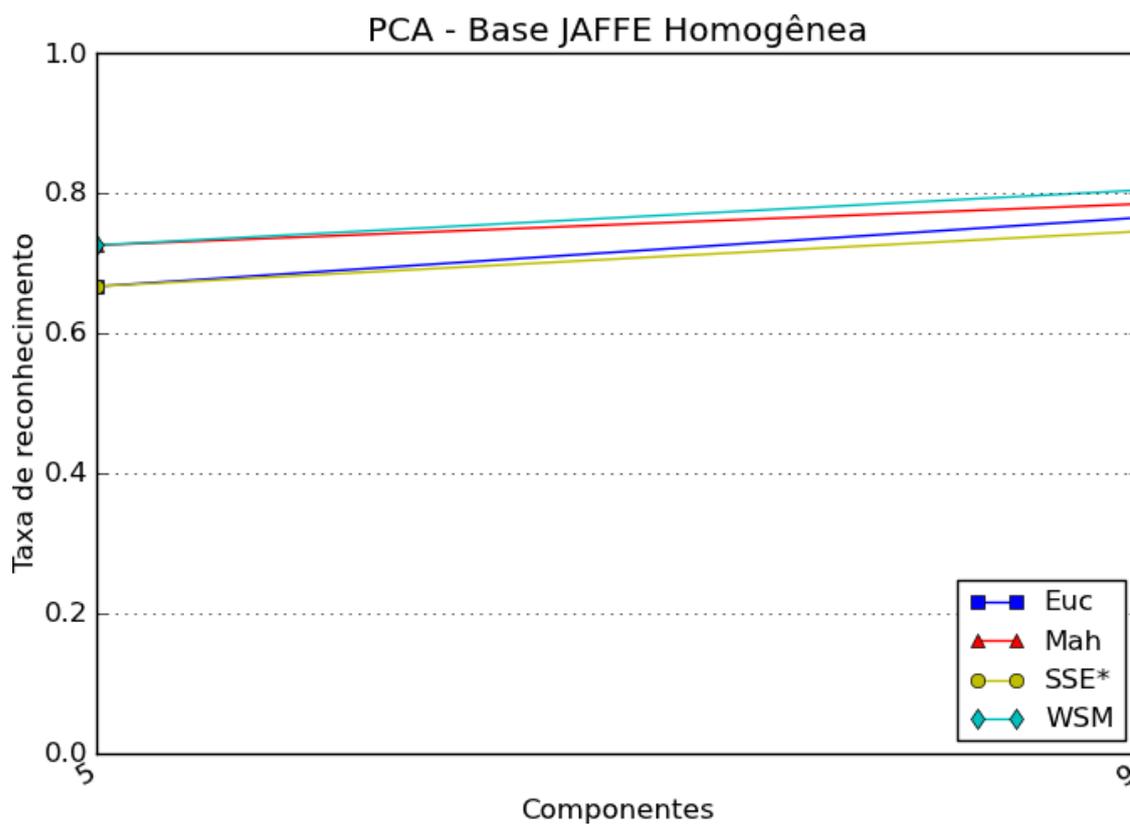


Figura 4.31: Comparação de métricas na base JAFFE homogênea.

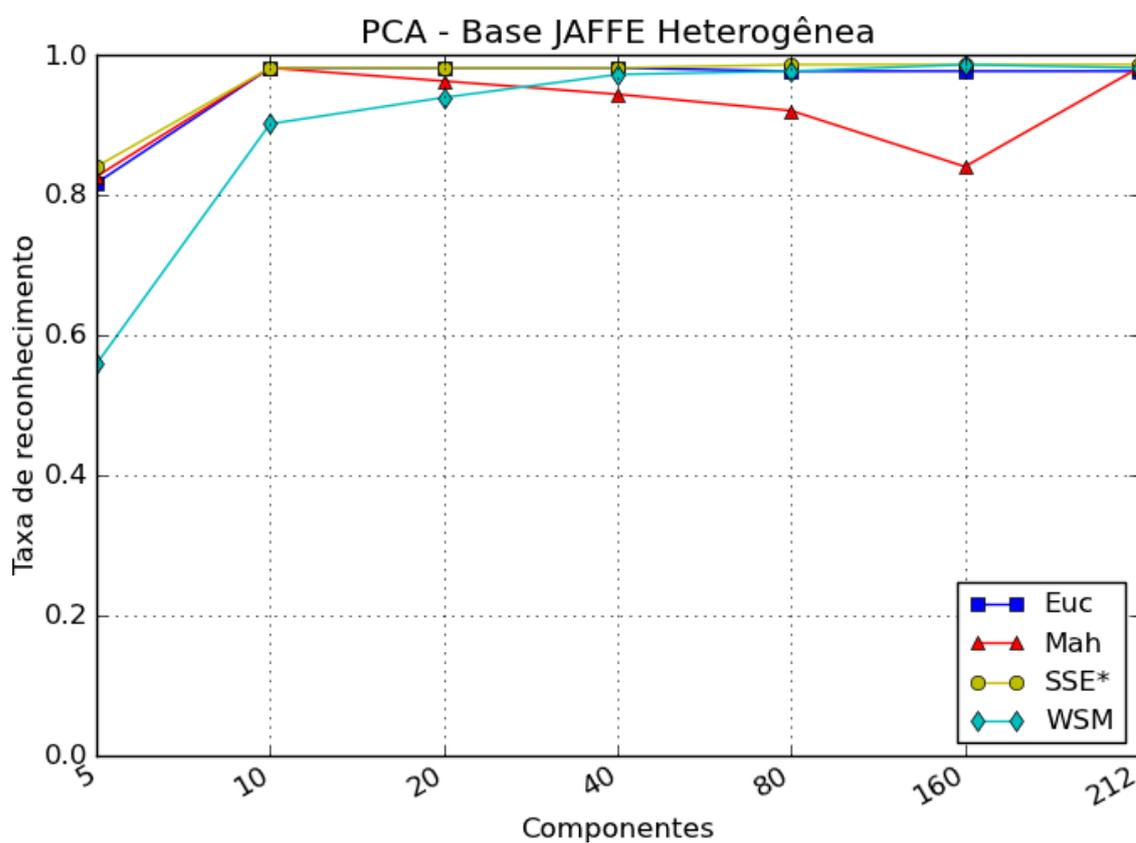


Figura 4.32: Comparação de métricas na base JAFFE heterogênea.

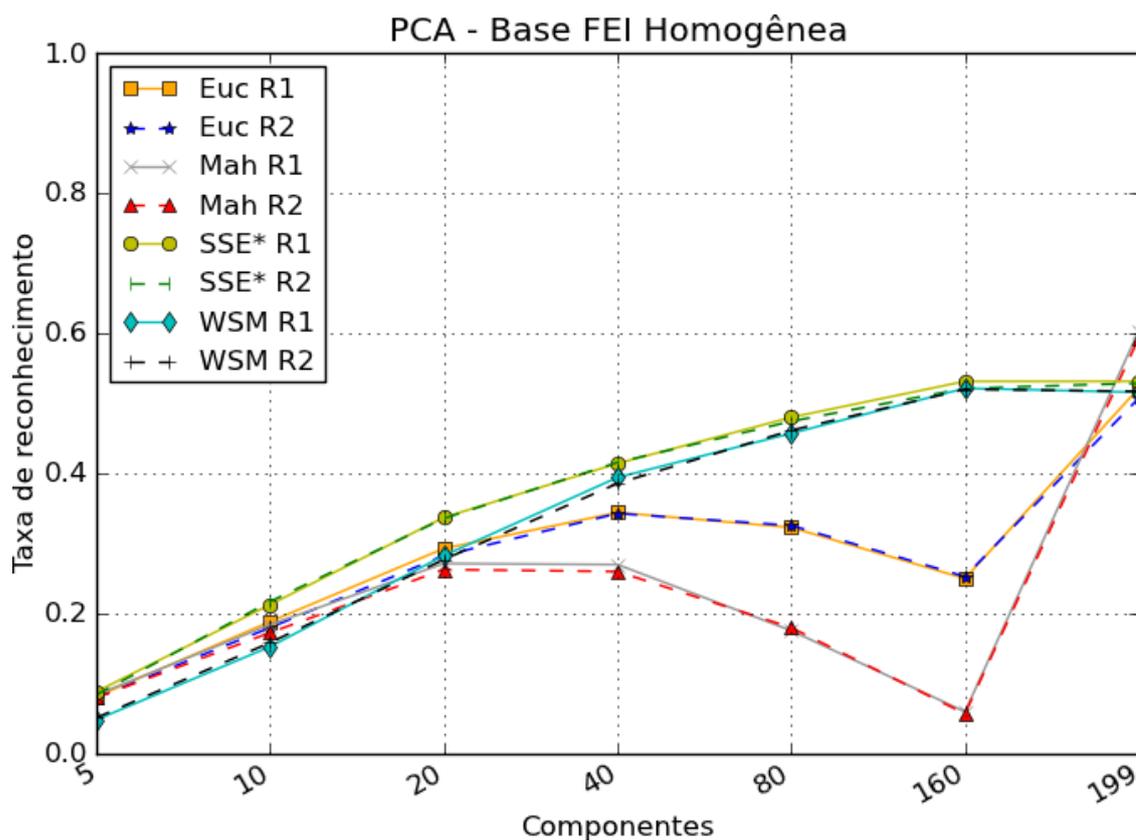


Figura 4.33: Comparação de métricas na base pré-processada FEI homogênea.

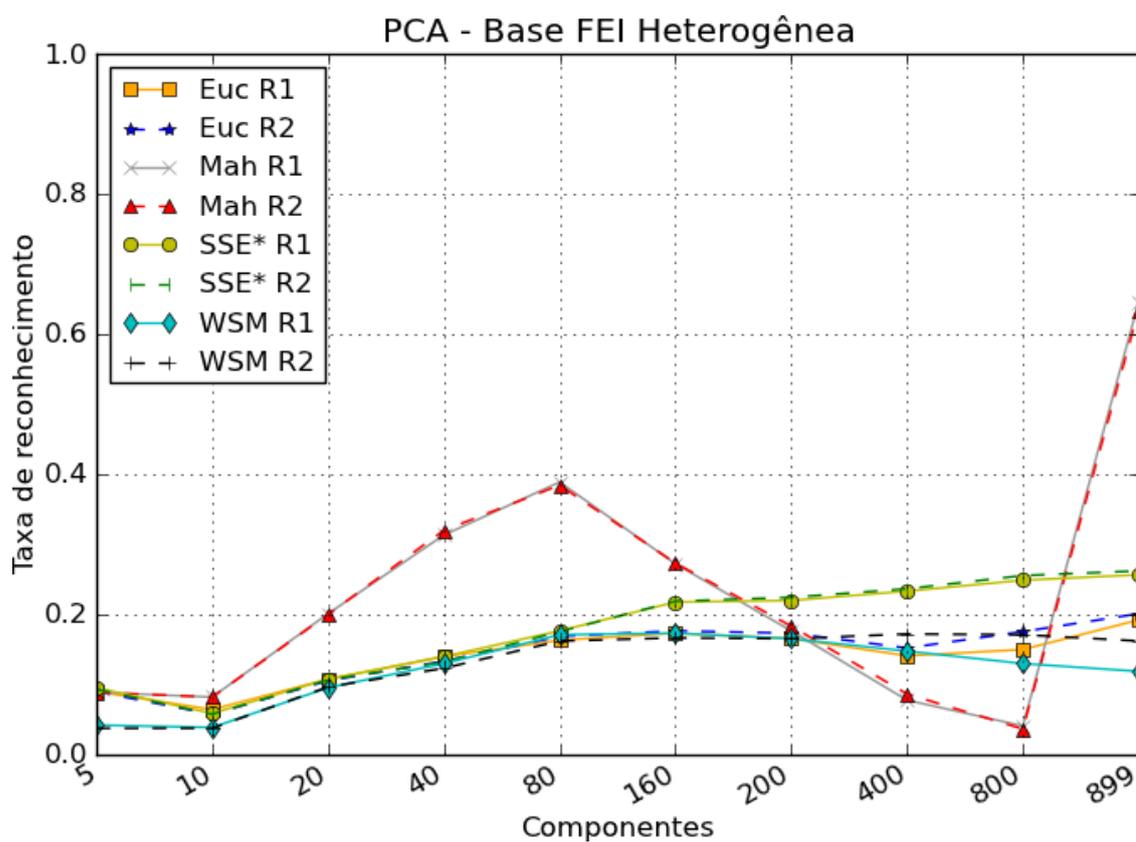


Figura 4.34: Comparação de métricas na base pré-processada FEI heterogênea.

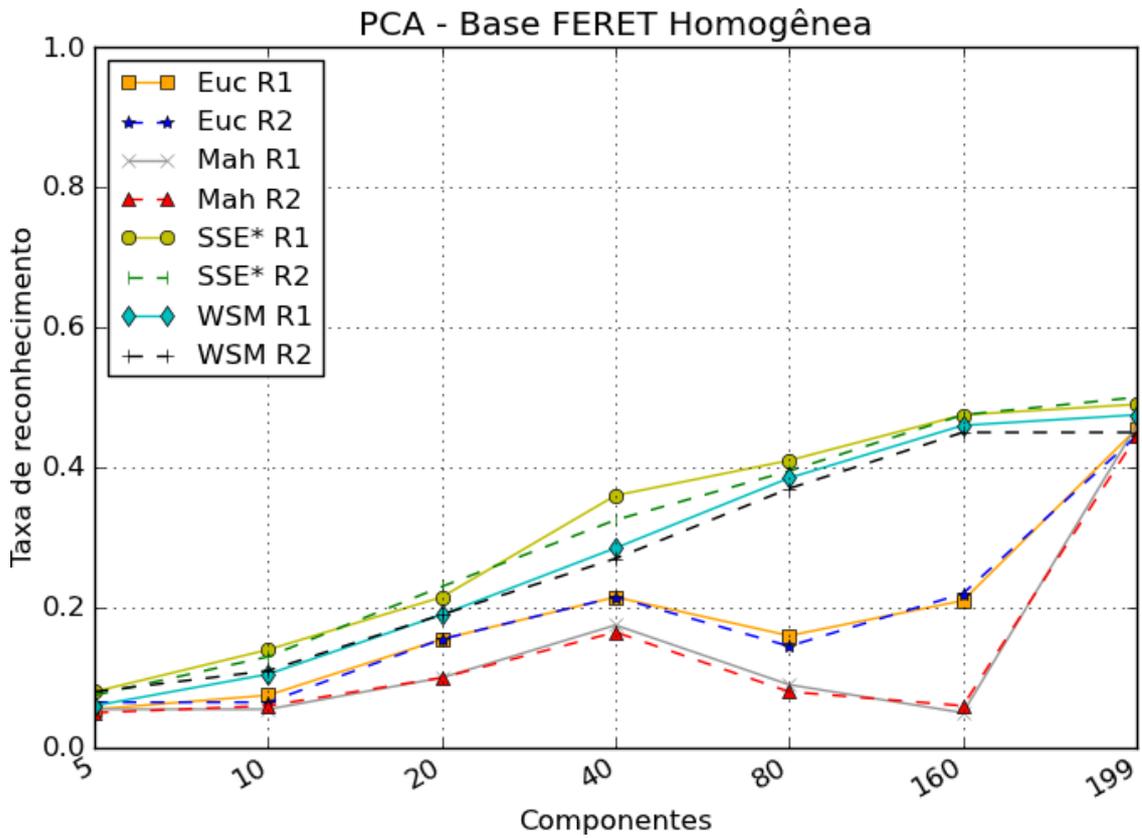


Figura 4.35: Comparação de métricas na base pré-processada FERET homogênea.

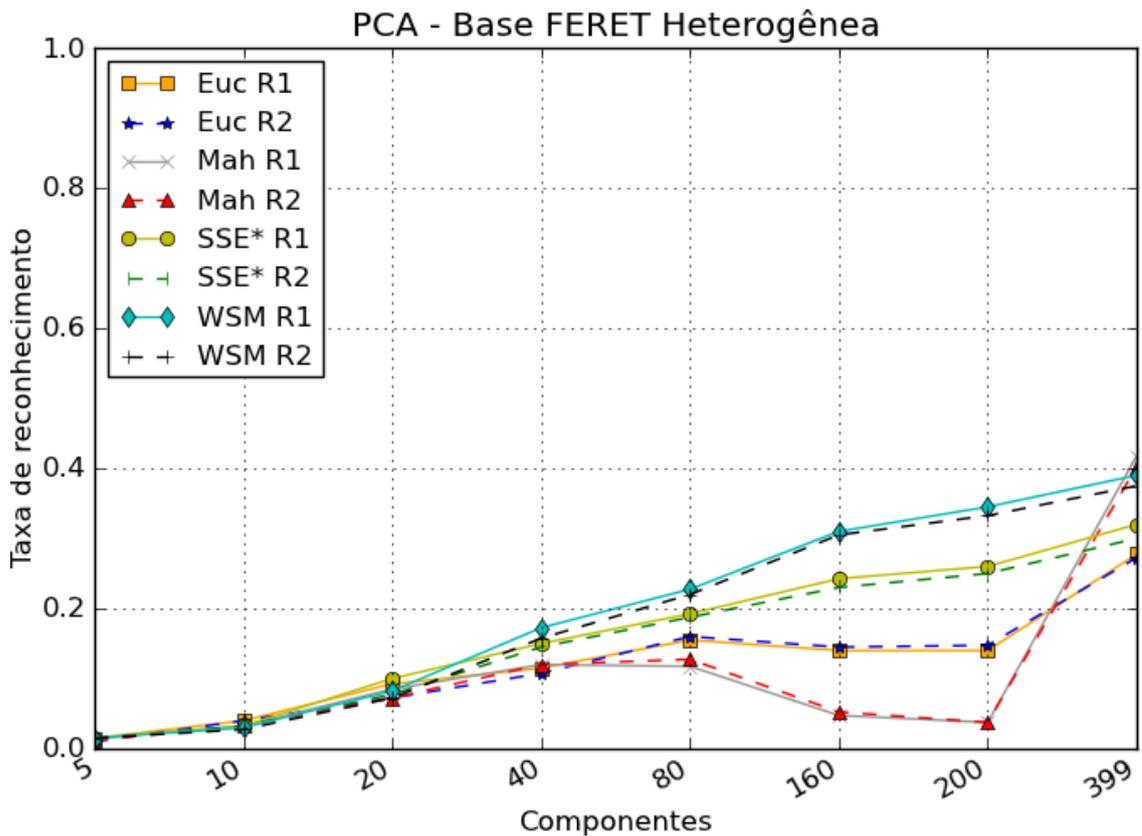


Figura 4.36: Comparação de métricas na base pré-processada FERET heterogênea.

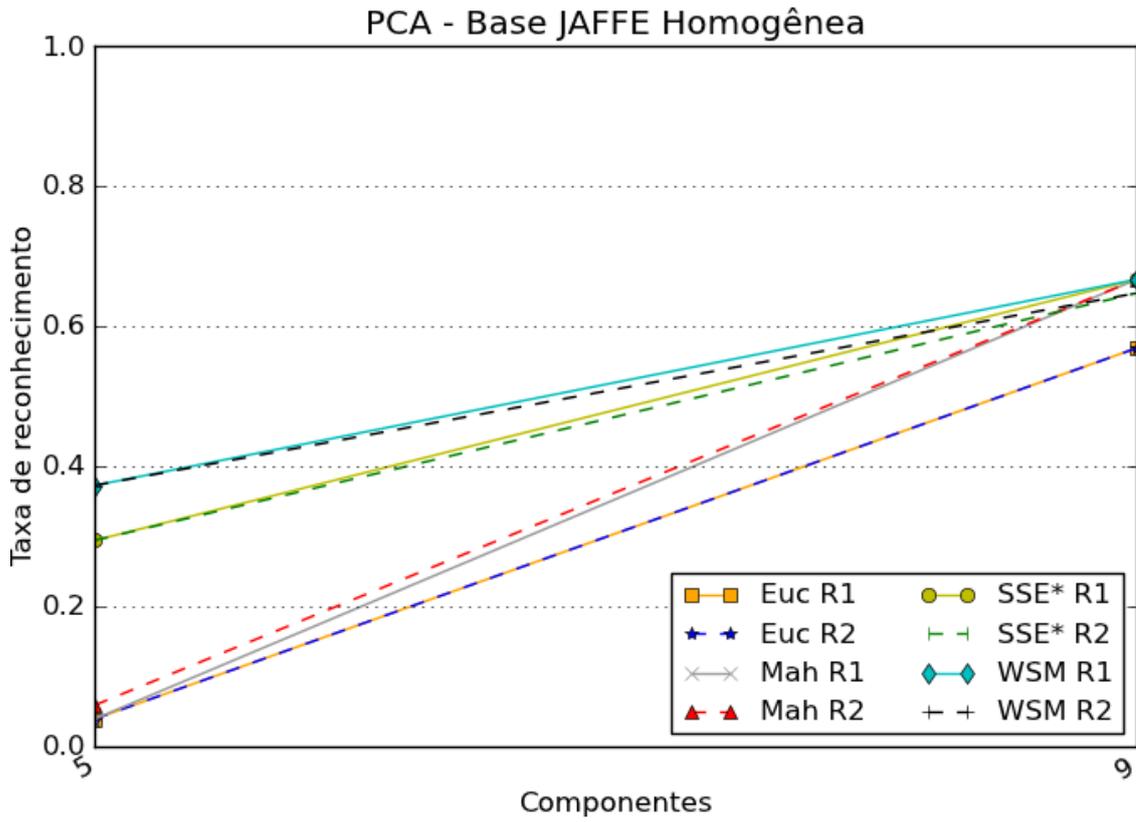


Figura 4.37: Comparação de métricas na base pré-processada JAFFE homogênea.

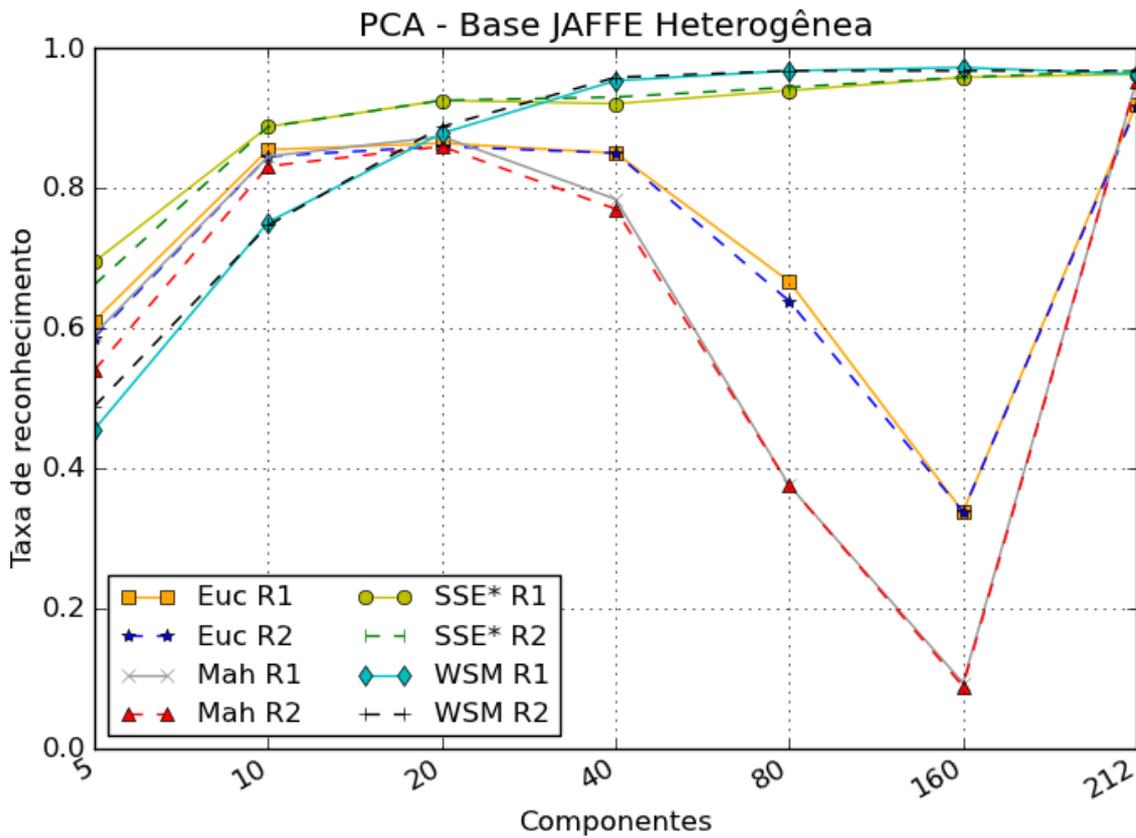


Figura 4.38: Comparação de métricas na base pré-processada JAFFE heterogênea.

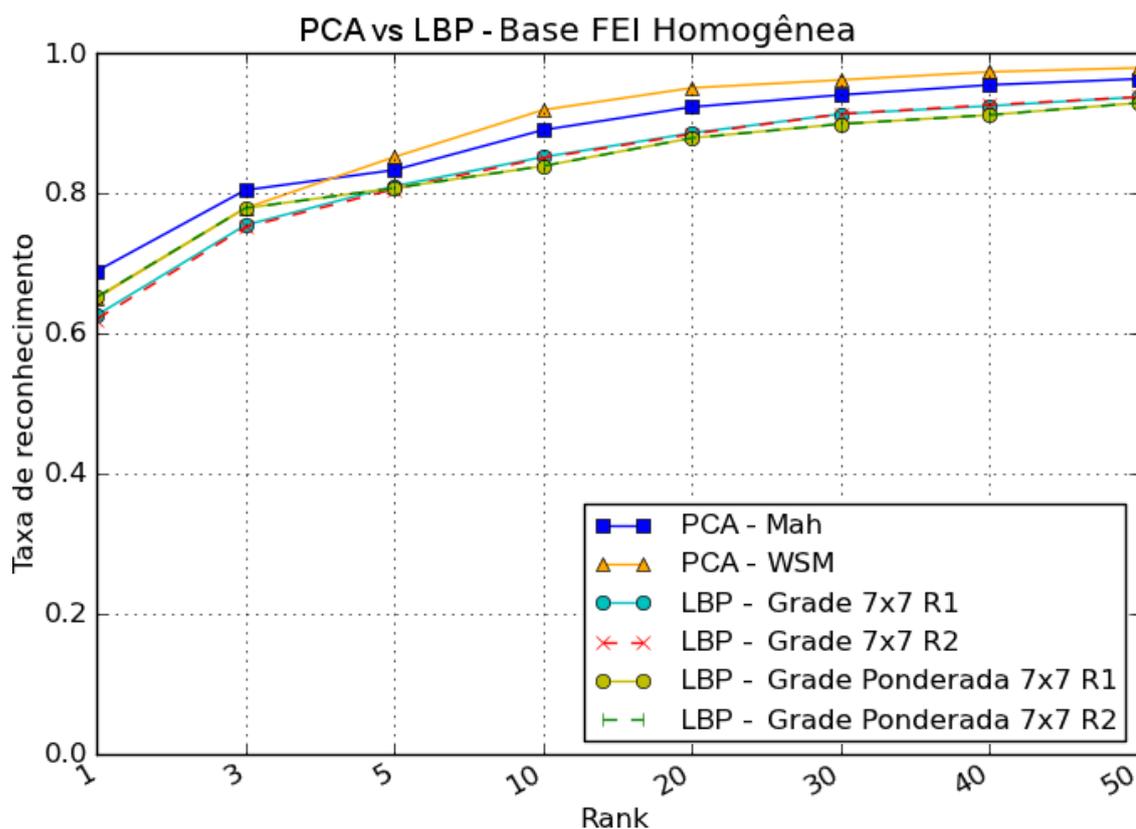


Figura 4.39: Comparação entre as técnicas LBP e PCA na base FEI homogênea.

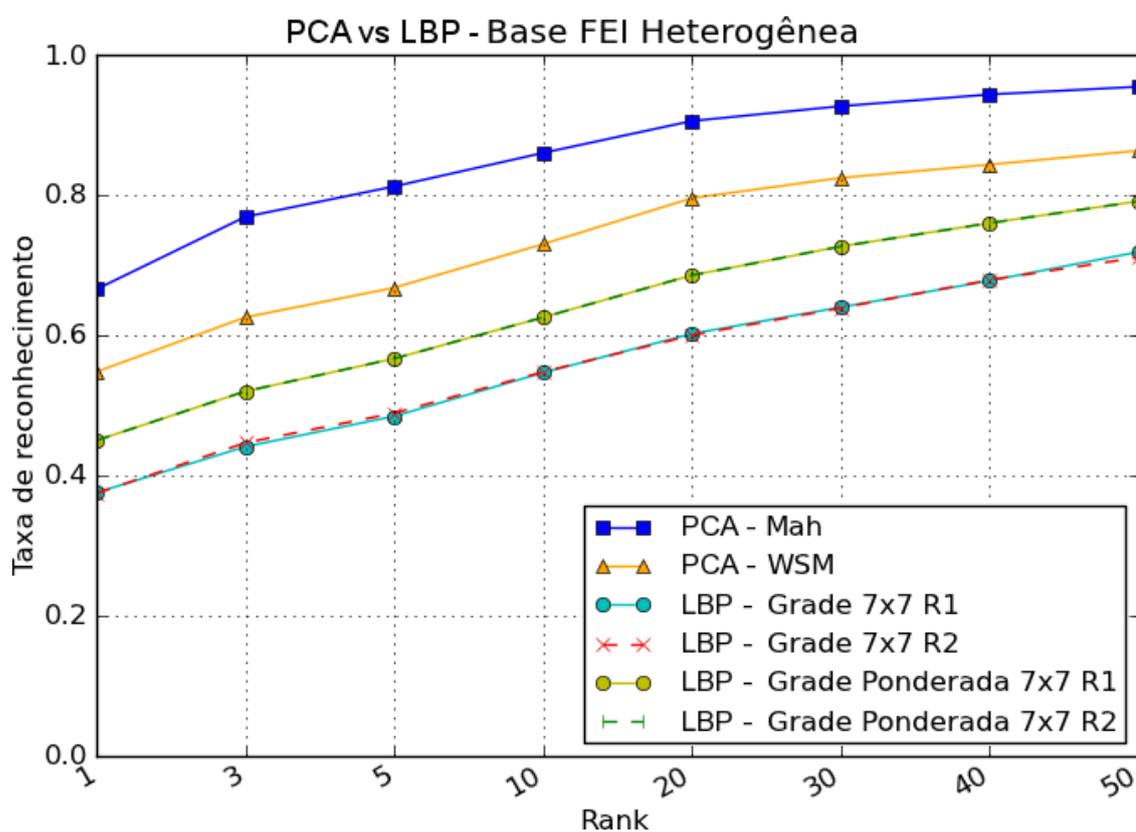


Figura 4.40: Comparação entre as técnicas LBP e PCA na base FEI heterogênea.

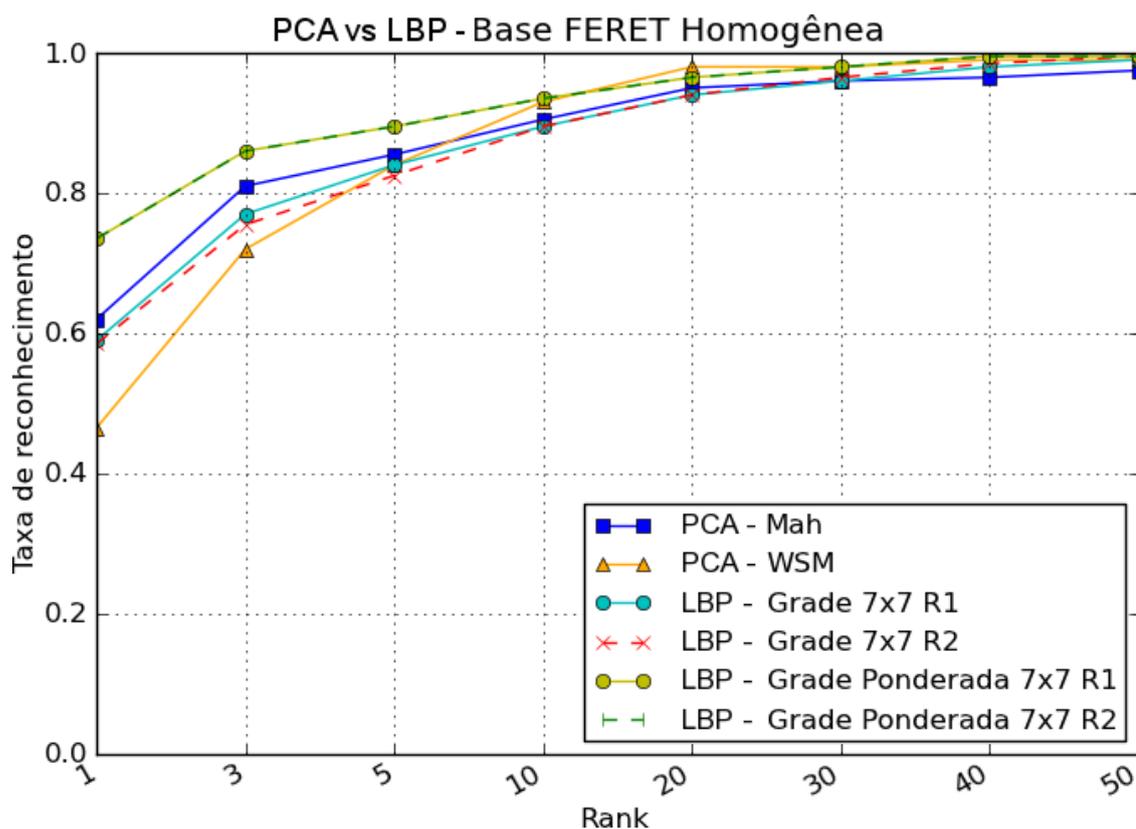


Figura 4.41: Comparação entre as técnicas LBP e PCA na base FERET homogênea.

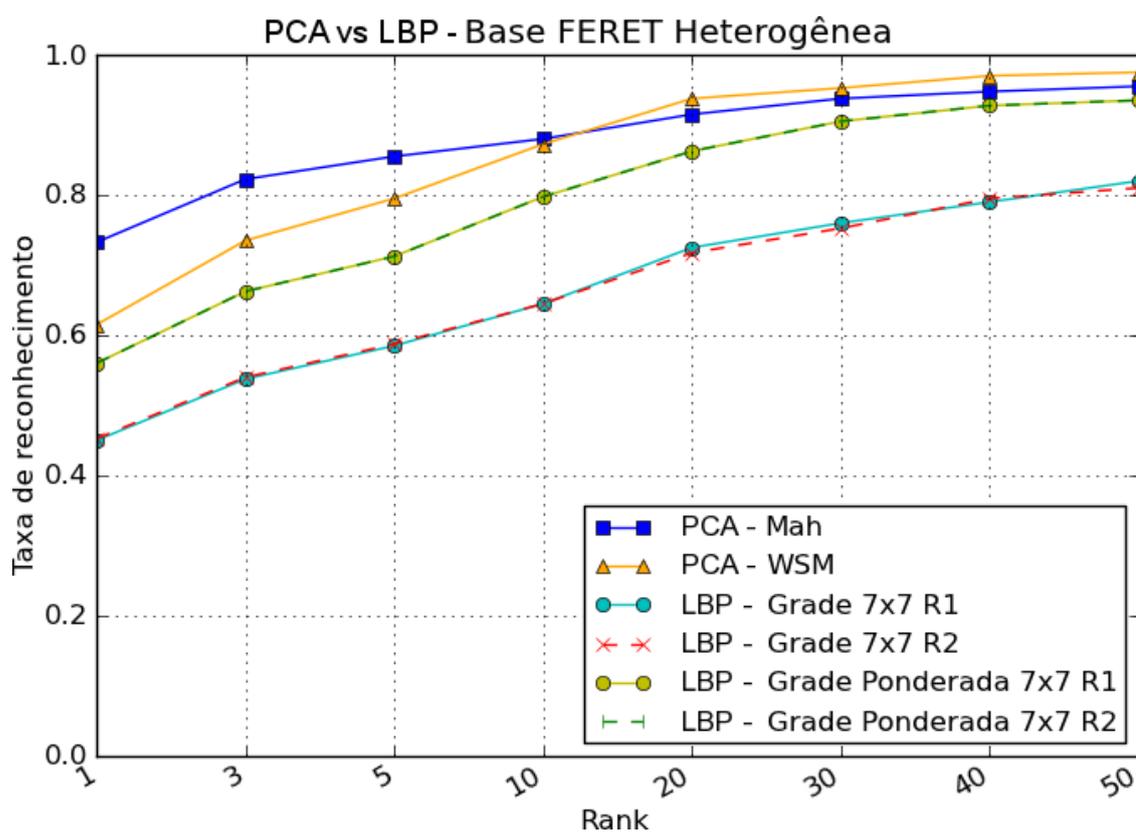


Figura 4.42: Comparação entre as técnicas LBP e PCA na base FERET heterogênea.

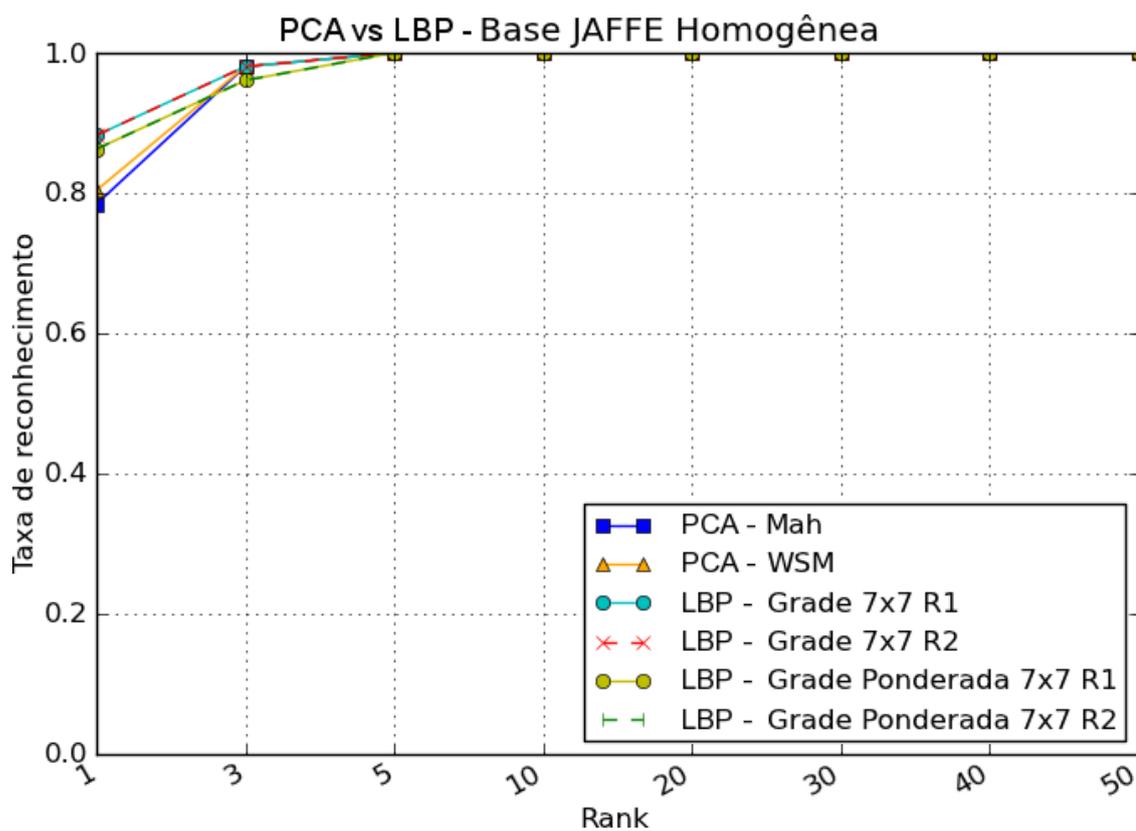


Figura 4.43: Comparação entre as técnicas LBP e PCA na base JAFFE homogênea.

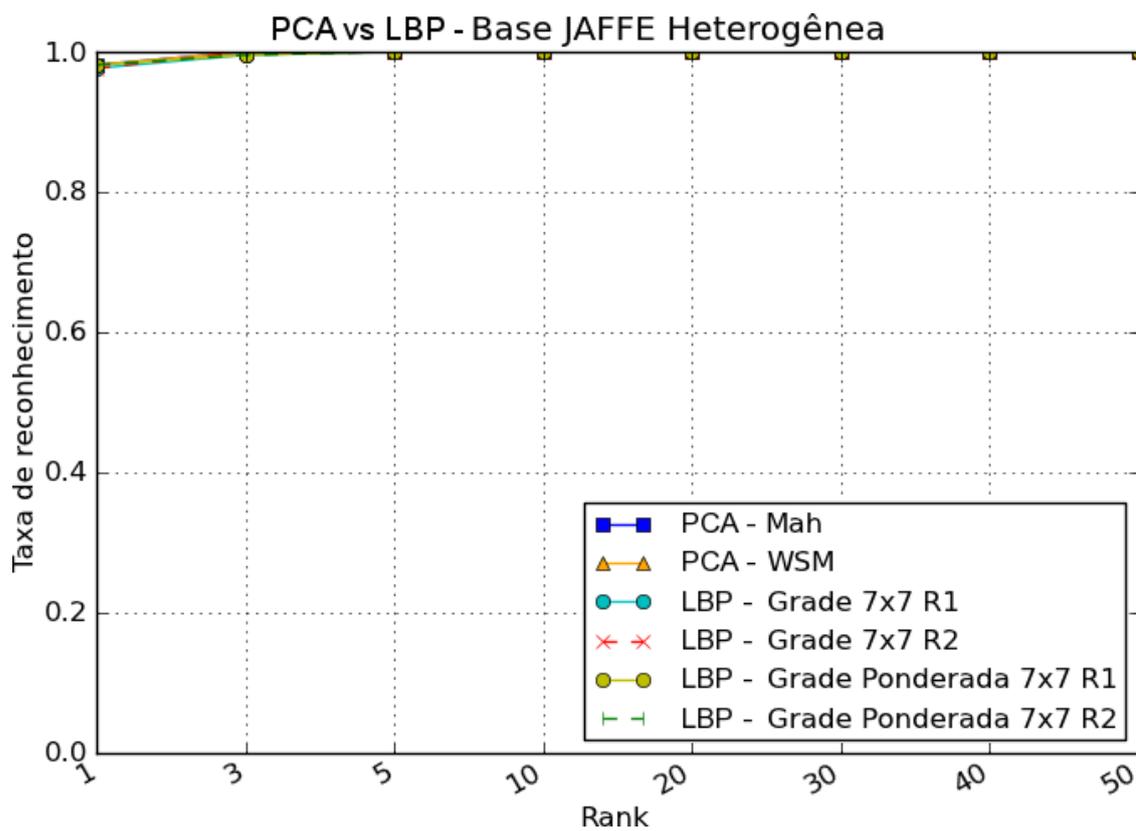


Figura 4.44: Comparação entre as técnicas LBP e PCA na base JAFFE heterogênea.

5 CONCLUSÃO

Neste trabalho realizou-se uma estudo comparativa entre as técnicas LBP e PCA, aplicadas ao reconhecimento automático de faces, com o propósito de identificar os parâmetros adequados para implementação de abordagens baseadas em características locais ou globais. Analisou-se a capacidade de classificação desses métodos sobre bases homogêneas e heterogêneas e em situações onde há apenas uma imagem de treinamento por pessoa, utilizando amostras com desvios de expressão, pose e iluminação. Para isto, desenvolveu-se um arcabouço computacional completo, composto pelas etapas de normalização, extração de características e classificação, presentes na literatura afim.

Como abordagem local investigou-se a técnica LBP, devido às boas taxas de reconhecimento apresentadas em estudos semelhantes e à possibilidade de utilizar regiões faciais específicas para a classificação das amostras. Identificou-se, no presente trabalho, que a variação dos raios de vizinhança, aplicados no pré-processamento, não proporciona melhorias relevantes. No entanto, resultados favoráveis puderam ser observados com a redução das áreas que compõem a grade de extração. Esta alteração melhora consideravelmente o desempenho do classificador, em detrimento do custo computacional, pois aumenta o grau de detalhamento do histograma gerado. A ponderação dessas áreas permite a valorização de regiões mais discriminantes e a suavização de regiões nocivas ao processo, melhorando ainda mais as taxas de reconhecimento.

Verificou-se também, que o LBP é robusto a variações de iluminação, escala e pose, por que extrai um padrão local relativo ao pixel central e utiliza comparação por histogramas. Além disso, este método não requer treinamento, portanto não precisa ser atualizado a cada adição de amostra na base de busca, e também é invariante ao problema limitação amostral. A utilização de múltiplas amostras por pessoa aumenta a probabilidade de reconhecimento. Contudo, observou-se uma fragilidade à variações na base de treinamento. Portanto, recomenda-se o uso de bases homogênea ou pesos que neutralizem as variações.

Em relação à análise de regiões faciais concluiu-se que os olhos são mais adequados para base homogênea e o nariz para base heterogênea, por que varia pouco sob desvios de expressão. Os experimentos com composições faciais apresentaram os melhores resultados com a combinação de olhos e nariz para ambos ambientes, homogêneos e heterogêneos, enfatizando ainda mais a relevância destas regiões para o reconhecimento de faces. Por outro lado, a região bucal foi pior em todos experimentos, devido a existência de variações de expressão.

Para análise de abordagens globais empregou-se o método PCA, uma das técnicas mais bem sucedidas no processo de reconhecimento facial. Verificou-se que para imagens pré-processadas, de acordo com a metodologia proposta por Amaral et al (AMARAL et al., 2009), deve-se utilizar todas componentes principais, pois a etapa de normalização elimina praticamente todos artefatos. Caso seja necessário limitar a quantidade de componentes deve-se empregar a distância Mahalano-bis ou Mahalanobis ponderada. E em situações onde há restrição amostral, ambientes heterogêneos são mais eficientes do que ambientes homogêneos. No entanto, reduzindo-se a quantidade de componentes obtiveram-se os melhores resultados com bases homogêneas, e para bases com múltiplas amostras por pessoa também é apropriado trabalhar neste ambiente.

Observou-se, também, que o uso da distância Euclidiana é adequado em bases homogêneas, que não contém desvios de expressão, pose e iluminação. E por este motivo é ruim em bases heterogêneas. A distância por Erro Médio Quadrático modificado obteve seu melhor desempenho com imagens pré-processadas pelo operador binário do LBP. Contudo, todos resultados obtidos pelo PCA com estas amostras são inferiores às outras abordagens.

Em geral, o PCA foi melhor que o LBP na maioria dos experimentos realizados. Analisando os resultados obtidos acredita-se que empregar um critério de pré-classificação considerando as 10 amostras mais semelhantes é razoável para verificação manual, pois possui na maioria dos casos analisados uma taxa de reconhecimento superior a 90%. Por fim, conclui-se, também, que ambientes homogêneos, compostos por imagens frontais neutras, são ideais para a implementação de sistemas de reconhecimento automático de faces. Este modelo é apropriado para situações onde há restrição amostral. No entanto, a utilização de múltiplas imagens por pessoa otimiza a classificação, caso todos indivíduos possuam as mesmas variações.

O reconhecimento automático de crianças e adolescentes desaparecidos proporciona a criação de um ambiente heterogêneo, no qual o PCA obteve resultados superiores, pois não há critério para aquisição das imagens de treinamento. Contudo, o LBP é mais adequado para tratar variações de captura e problemas de oclusão parcial da face, presentes nesta aplicação. Portanto, é necessário empregar ambas abordagens, locais e globais, para tratar características específicas de acordo com o contexto de cada busca.

5.1 Trabalhos Futuros

Com base na revisão bibliográfica realizada e nos resultados obtidos através dos experimentos deste estudo, algumas sugestões foram elaboradas com o propósito de contribuir com o desenvolvimento de pesquisas sobre reconhecimento automático de imagens de faces. Inicialmente, acredita-se que seja necessário o desenvolvimento de um novo protocolo de avaliação que con-

sidere ambientes heterogêneos e problemas de restrição amostral, a fim de, analisar problemas presentes em aplicações reais.

Quanto a melhoria dos métodos comparados, recomenda-se a implementação de técnicas de aprendizado para a otimização dos parâmetros do LBP e identificação das regiões mais discriminantes da face em um conjunto de treinamento específico. Sugere-se ainda, desenvolver abordagens híbridas combinando características das técnicas analisadas neste estudo, como por exemplo, processando os histogramas do LBP através do PCA para reduzir a redundância dos dados, ou também, empregar pesos faciais em uma abordagem global como o PCA.

REFERÊNCIAS

- AG, B. *BioID Face Database*. 2001. Acessado em: 03 Set. 2011. Disponível em: <<http://support.bioid.com/downloads/facedb/index.php>>.
- AHONEN, T.; HADID, A.; PIETIKAINEN, M. Face recognition with local binary patterns. *Computer Vision ECCV 2004*, Springer, v. 3021, p. 469–481, 2004.
- AHONEN, T.; HADID, A.; PIETIKAINEN, M. Face description with local binary patterns: Application to face recognition. *IEEE Transaction Pattern Analysis and Machine Intelligence*, IEEE Computer Society, v. 28, p. 2037–2041, 2006. ISSN 0162-8828.
- AMARAL, V. et al. Normalização espacial de imagens frontais de face em ambientes controlados e não-controlados. *FaSCi-Tech*, FATEC São Caetano, v. 1, n. 1, 10 2009.
- AMARAL, V.; THOMAZ, C. E. *Normalização espacial de imagens frontais de face*. [S.l.], 10 2008.
- AT&T. *ORL Face Database*. 1994. Acessado em: 03 Set. 2011. Disponível em: <<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>>.
- BARTLETT, M. S.; MOVELLAN, J. R.; SEJNOWSKI, T. J. Face recognition by independent component analysis. *IEEE Transactions on Neural Networks*, IEEE, v. 13, n. 6, p. 1450–1464, 2002.
- BELHUMEUR, P. N.; HESPANHA, J.; KRIEGMAN, D. J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. In: *Proceedings of the 4th European Conference on Computer Vision-Volume I - Volume I*. London, UK: Springer-Verlag, 1996. p. 45–58. ISBN 3-540-61122-3.
- BERN, U. de. *IAM Face Database*. 2008. Acessado em: 03 Set. 2011. Disponível em: <<http://www.iam.unibe.ch/fki/databases/iam-faces-database>>.
- BRASIL, D. do. *Cadastro Nacional para Crianças Desaparecidas Não Funciona*. Acessado em: 31 Jul. 2011. Disponível em: <<http://desaparecidosdobrasil.org/>>.
- BRUNELLI, R.; POGGIO, T. Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 15, n. 10, p. 805–821, 10 1993.
- CHELLAPPA, R.; WILSON, C.; SIROHEY, S. Human and machine recognition of faces: a survey. *Proceedings of the IEEE*, v. 83, n. 5, p. 705–741, may 1995. ISSN 0018-9219.
- CHEN, Y. et al. Similarity-based classification: Concepts and algorithms. *J. Mach. Learn. Res.*, Journal of Machine Learning Research, v. 10, p. 747–776, 2009. ISSN 1532-4435.
- CONCI, A.; AZEVEDO, E.; LETA, F. R. *Computação Gráfica*. 2. ed. Rio de Janeiro, RJ, BR: Editora Campus Elsevier, 2008. ISBN 978-85-35223-29-3.
- DELAC, K.; GRGIC, M.; BARTLETT, M. S. *Recent Advances in Face Recognition*. [S.l.]: In-teh, 2008.

- DELAC, K.; GRGIC, M.; GRGIC, S. Independent comparative study of pca, ica, and lda on the feret data set. *International Journal of Imaging Systems and Technology*, Citeseer, v. 15, n. 5, p. 252–260, 2005.
- DENG, W. et al. Robust, accurate and efficient face recognition from a single training image: A uniform pursuit approach. *Pattern Recognition*, Elsevier Science Inc., n. 5, p. 1748–1762, 2010. ISSN 0031-3203.
- DESAPARECIDOS, C. das Crianças e A. *Desaparecimento de Crianças e Adolescentes: CPI em ação*. Acessado em: 31 Jul. 2011. Disponível em: <<http://www.picriancasdesaparecidas.com.br/>>.
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification*. 2. ed. [S.l.]: Wiley-Interscience, 2001.
- EICKELER, S.; JABS, M.; RIGOLL, G. Comparison of confidence measures for face recognition. In: *FG '00: In proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000*. [S.l.]: IEEE Computer Society, 2000. p. 257.
- ESSEX, U. de. *PEIPA Face Database*. 2004. Acessado em: 03 Set. 2011. Disponível em: <<http://peipa.essex.ac.uk/index.html>>.
- FISHER, R. A. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, Wiley Online Library, v. 7, n. 2, p. 179–188, 1936.
- FUKUNAGA, K. *Introduction to statistical pattern recognition (2nd ed.)*. San Diego, CA, USA: Academic Press Professional, Inc., 1990. ISBN 0-12-269851-7.
- GATTÁS, G. J. F.; FIGARO-GARCIA, C. *Caminho de Volta: Tecnologia na busca de crianças e adolescentes desaparecidos no estado de São Paulo*. São Paulo, SP, BR: Secretaria Especial de Direitos Humanos: CONANDA - Conselho Nacional dos Direitos da Criança e do Adolescente, 2007. ISBN 978-85-89169-02-8.
- GONZALES, R. C.; WOODS, R. E. *Processamento de Imagens Digitais*. 1. ed. [S.l.]: Edgard Blücher, 2000. ISBN 8521202644.
- GUTTA, S.; WECHSLER, H. Face recognition using hybrid classifiers. *Pattern Recognition*, v. 30, n. 4, p. 539–553, 7 1997.
- JUNIOR, L. L. O.; THOMAZ, C. E. *Captura e alinhamento de imagens: Um banco de faces brasileiro*. [S.l.], 6 2006.
- KANADE, T. *Picture processing system by computer complex and recognition of human faces*. Tese (Tese de Doutorado) — Kyoto University, 1973.
- KIRBY, M.; SIROVICH, L. Application of the karhunen-loeve procedure for the characterization of human faces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 12, n. 1, p. 103 –108, jan 1990. ISSN 0162-8828.
- KITANI, E. C.; THOMAZ, C. E. *Um tutorial sobre análise de componentes principais para o reconhecimento automático de faces*. [S.l.], 5 2006.
- KITANI, E. K. *Análise de discriminantes lineares para modelagem e reconstrução de imagens de faces*. Dissertação (Dissertação de Mestrado) — Centro Universitário da FEI, 2007.

LEÃO, R. D. *Extração multilinear de informações discriminantes em imagens de ressonância magnética do cérebro humano*. Dissertação (Dissertação de Mestrado) — Centro Universitário da FEI, 2009.

LIAU, H. F. et al. New parallel models for face recognition. In: DELAC, K.; GRGIC, M.; BARTLETT, M. S. (Ed.). *Recent Advances in Face Recognition*. [S.l.]: In-teh, 2008. v. 1, p. 15–28.

LITWILLER, D. *CCD vs. CMOS: Facts and Fiction*. [S.l.], 2001.

LU, X. *Image analysis for face recognition - A brief survey*. 2003. Acessado em: 10 Mai. 2010. Disponível em: <<https://www.msu.edu/~lvxiaogu/publications/>>.

MARTÍNEZ, A.; BENAVENTE, R. *The AR Face Database*. Bellatera, Jun 1998.

MATURANA, D.; MERY, D.; SOTO, I. Face recognition with local binary patterns, spatial pyramid histograms and naive bayes nearest neighbor classification. In: *Proceedings of the 2009 International Conference of the Chilean Computer Science Society*. Washington, DC, USA: IEEE Computer Society, 2009. (SCCC '09), p. 125–132. ISBN 978-0-7695-4137-2.

MIT. *MIT Face Database*. 1991. Acessado em: 03 Set. 2011. Disponível em: <<http://vismod.media.mit.edu/vismod/>>.

MORIHISHA, O. *Avaliação comparativa entre agradabilidade facial, proporção áurea e padrão facial*. Dissertação (Dissertação de Mestrado) — Centro Universitário da FEI, 2006.

OHIO, U. de. *AR Face Database*. 1998. Acessado em: 03 Set. 2011. Disponível em: <<http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html>>.

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

PANTIC, M.; ROTHKRANTZ, L. J. M. Automatic analysis of facial expressions: The state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, v. 22, n. 12, p. 1424–1445, 12 2003.

PERLIBAKAS, V. Distance measures for pca-based face recognition. *Pattern Recognition*, Elsevier Science Inc, New York, NY, USA, v. 25, n. 6, p. 711–724, 2004. ISSN 0167-8655.

PHILLIPS, P. J. et al. The feret evaluation methodology for face-recognition algorithms. In: *IEEE Transaction on Pattern Analysis and Machine Intelligence*. Washington, DC, USA: IEEE Computer Society, 2000. v. 22, p. 1090–1104.

S., B. D. et al. The csu face identification evaluation system: its purpose, features, and structure. In: *Proceedings of the 3rd international conference on Computer vision systems*. Berlin, Heidelberg: Springer-Verlag, 2003. (ICVS'03), p. 304–313. ISBN 3-540-00921-3.

SIROVICH, L.; KIRBY, M. Low-dimensional procedure for the characterization of human faces. *Journal of Optical Society of America A*, OSA, v. 4, n. 3, p. 519–524, 1987.

STANDARDS, N. N. I. of; TECHNOLOGY. *FERET Face Database*. 1996. Acessado em: 03 Set. 2011. Disponível em: <<http://face.nist.gov/colorferet/>>.

- SWETS, D. L.; WENG, J. Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.*, IEEE Computer Society, Washington, DC, USA, v. 18, p. 831–836, August 1996. ISSN 0162-8828.
- TAN, X. et al. Face recognition from a single image per person: A survey. *Pattern Recognition*, Elsevier Science Inc., n. 9, p. 1725–1745, 2006. ISSN 0031-3203.
- TECHNOLOGY, G. I. of. *Georgia Tech Face Database*. 1999. Acessado em: 06 Set. 2011. Disponível em: <http://www.anefian.com/research/face_reco.htm>.
- THOMAZ, C. E. *Estudo de classificadores para o reconhecimento automático de faces*. Dissertação (Dissertação de Mestrado) — Pontifícia Universidade Católica do Rio de Janeiro, 1999.
- THOMAZ, C. E.; GILLIES, D. F.; FEITOSA, R. Q. Using mixture covariance matrices to improve face and facial expression recognitions. *Pattern Recogn. Lett.*, Elsevier Science Inc., New York, NY, USA, v. 24, p. 2159–2165, September 2003. ISSN 0167-8655.
- THOMAZ, C. E.; GILLIES, D. F.; FEITOSA, R. Q. A new covariance estimate for bayesian classifiers in biometric recognition. *IEEE Transactions on Circuits and Systems*, v. 14, n. 2, p. 214–223, 2004.
- THOMAZ, C. E.; GIRALDI, G. A. A new ranking method for principal components analysis and its application to face image analysis. *Image Vision Comput.*, Butterworth-Heinemann, Newton, MA, USA, v. 28, p. 902–913, June 2010. ISSN 0262-8856.
- TURK, M.; PENTLAND, A. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, MIT Press, n. 1, p. 71–86, 1991. ISSN 0898-929X.
- TURK, M.; PENTLAND, A. Face recognition using eigenfaces. *Computer Vision and Pattern Recognition*, IEEE Computer Society, p. 586–591, 1991.
- UMIST, U. de. *Umist Face Database*. 1998. Acessado em: 03 Set. 2011. Disponível em: <<http://www.sheffield.ac.uk/eee/research/iel/research/face>>.
- WANG, J. et al. On solving the face recognition problem with one training sample per subject. *Pattern Recogn.*, Elsevier Science Inc., New York, NY, USA, v. 39, p. 1746–1762, September 2006. ISSN 0031-3203.
- WU, J.; ZHOU, Z. Face recognition with one training image per person. *Pattern Recognition Letters*, Elsevier Science Inc., New York, NY, USA, v. 23, p. 1711–1719, December 2002. ISSN 0167-8655.
- YALE, U. de. *Yale Face Database*. 1997. Acessado em: 03 Set. 2011. Disponível em: <<http://cvc.yale.edu/projects/yalefaces/yalefaces.html>>.
- YALE, U. de. *Extended Yale Face Database B*. 2005. Acessado em: 06 Set. 2011. Disponível em: <<http://vision.ucsd.edu/leekc/ExtYaleDatabase/ExtYaleB.html>>.
- YANG, L.; JIN, R. *Distance Metric Learning: A Comprehensive Survey*. [S.l.], 5 2006.
- ZHAO, W. et al. Face recognition: A literature survey. *ACM Computing Surveys*, v. 35, n. 4, p. 399–458, 2003.