

UNIVERSIDADE FEDERAL DO AMAZONAS  
FACULDADE DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

THIAGO AZEVEDO FALCÃO

COMPARAÇÃO DO DESEMPENHO DO CLASSIFICADOR DE  
NOVIDADES COM O CLASSIFICADOR DO VIZINHO MAIS PRÓXIMO  
NO RECONHECIMENTO FACIAL

MANAUS  
2014

UNIVERSIDADE FEDERAL DO AMAZONAS  
FACULDADE DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

THIAGO AZEVEDO FALCÃO

COMPARAÇÃO DO DESEMPENHO DO CLASSIFICADOR DE  
NOVIDADES COM O CLASSIFICADOR DO VIZINHO MAIS PRÓXIMO  
NO RECONHECIMENTO FACIAL

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade do Amazonas, como requisito parcial para a obtenção do título de Mestre em Engenharia Elétrica, área de concentração Controle e Automação de Sistemas.

Orientador: Prof. Dr. Cícero Ferreira Fernandes Costa Filho

Co-orientadora: Profa. Dra. Marly Guimarães Fernandes Costa

MANAUS  
2014

THIAGO AZEVEDO FALCÃO

COMPARAÇÃO DO DESEMPENHO DO CLASSIFICADOR DE  
NOVIDADES COM O CLASSIFICADOR DO VIZINHO MAIS PRÓXIMO  
NO RECONHECIMENTO FACIAL

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade do Amazonas, como requisito parcial para a obtenção do título de Mestre em Engenharia Elétrica, área de concentração Controle e Automação de Sistemas.

Aprovado em \_\_\_\_/\_\_\_\_/\_\_\_\_

BANCA EXAMINADORA

Prof. Dr. Cícero Ferreira Fernandes Costa Filho  
Universidade Federal do Amazonas

Profa. Dra. Marly Guimarães Fernandes Costa  
Universidade Federal do Amazonas

Jozias Parente de Oliveira  
Universidade do Estado do Amazonas

*Dedico este trabalho a Jesus Cristo,  
Monique, Mateus e aos meus pais.*

## AGRADECIMENTOS

Primeiramente a Deus, meu Criador, Salvador e Senhor. “Portanto, quer comais quer bebais, ou façais outra qualquer coisa, fazei tudo para glória de Deus”. (1 Coríntios 10:31).

À toda minha família pelo incentivo, principalmente meus pais, que sempre reuniram grandes esforços para meus estudos, à minha esposa e filho que me incentivaram na execução deste trabalho.

Ao Prof. Cícero Costa Filho, pela excelente orientação, aulas, apoio e acompanhamento durante o planejamento e desenvolvimento deste trabalho, e à Profa. Marly Guimarães Costa, pelo apoio e incentivo.

Ao Programa de Pós-Graduação em Engenharia Elétrica (PPGEE) da Universidade Federal do Amazonas e, em especial, ao Centro de Pesquisa e Desenvolvimento em Tecnologia Eletrônica e da Informação, CETELI, por fornecer a infraestrutura necessária para minha formação.

Ao Instituto Nokia de Tecnologia pelo incentivo à educação dado através de infraestrutura, tempo livre para estudo e apoio financeiro para publicação de artigo científico.

Aos meus colegas de trabalho que contrinuíram para meu crescimento na área de *Machine Learning*.

Enfim, a todos, que de alguma forma contribuíram para a realização deste trabalho.

## RESUMO

Este trabalho propõe a utilização do classificador de novidades para reconhecimento de faces, o qual é baseado no filtro de novidades, proposto por Kohonen. Para avaliar o desempenho do novo classificador é feita uma comparação com o classificador do vizinho mais próximo, usando a métrica da distância euclidiana. A base de dados utilizada para essa comparação foi a base ORL. A informação da face é extraída utilizando os métodos PCA, 2DPCA e  $(2D)^2$ PCA, sem usar qualquer tipo de pré-processamento (fotométrico ou geométrico). Os seguintes resultados são apresentados no modo de identificação: taxa de reconhecimento *rank 1* e as curvas CMC, no modo verificação: as taxas de correta aceitação (CAR), de erro equivalente (EER), as curvas ROC e área sob a curva ROC (AUC). Os resultados obtidos mostraram que o classificador proposto tem um desempenho melhor do que o desempenho do vizinho mais próximo e do que outros classificadores anteriormente publicados usando a mesma base, quando a estratégia de validação cruzada *10-fold* é usada, com essa estratégia a taxa de reconhecimento obtida foi de 100%.

Palavras-chave: sistemas biométricos, reconhecimento facial, classificador de novidades.

## *ABSTRACT*

This work proposes the new classifier for face recognition, novelty classifier, which is based on novelty filter proposed by Kohonen. In order to evaluate the new classifier performance, it is performed a comparison with nearest neighborhood classifier, which uses the Euclidian distance as distance metric. ORL face database was chosen to be used in this comparison. There was not any pre-processing (photometric or geometric) on face images. It was used the following feature extraction methods: PCA, 2DPCA and (2D)<sup>2</sup>PCA. Some results in identification mode are exposed through *rank 1* recognition rate and CMC curves. In verification mode, the results were presented by Correct Acceptance Rate (CAR), Equivalent Error Rate (EER), ROC curves and Area under the ROC curve (AUC). Results shown that the proposed classifier performs better than others previously published, when the *10-fold Cross Validation* method is employed as a test strategy. Recognition rate of 100% is achieved with this test methodology.

Keywords: Biometric system, face recognition, novelty classifier.

## LISTA DE FIGURAS

Figura 1: Cada linha contém 10 imagens pré-processadas de um indivíduo [Fonte: Sahoolizadeh e Ghassabeh, 2008] .....	22
Figura 2: Imagens resultantes do método. [Fonte: Liao e Isa, 2010] .....	23
Figura 3: Equalização do histograma de uma face. ....	29
Figura 4: Vetor de características de uma imagem em escala de cinza. [Fonte: Pinheiro, 2010] .....	30
Figura 5: Técnica PCA: (a) conjunto de pontos; (b) subespaço horizontal com distâncias do conjunto de pontos para as projeções não mínimas; (c) subespaço com distâncias do conjunto de pontos para as projeções mínimas. ....	31
Figura 6: Técnica PCA: (a) conjunto de pontos; (b) subespaço horizontal com projeções de variâncias não máximas; (c) subespaço com projeções de variâncias máximas. ....	31
Figura 7: 2DPCA Bidirecional ou $(2D)^2PCA$ .....	35
Figura 8: Variação do p vista graficamente. ....	36
Figura 9: Curvas dos pontos centrais de cada classe no espaço bidimensional. a) Distância Euclidiana; b) Distância de Mahalanobis. [Fonte: Theodoridis e Koutroumbas, 2009] .....	37
Figura 10: Nesse caso, $k=5$ , o ponto $x$ é classificado na classe dos pontos pretos. [Fonte: Duda et al, 2000] .....	38
Figura 11: Conceito do filtro novidade, mostrando o vetor novidade $\mathbf{x}$ e o vetor projetado $\mathbf{x}$ . [Fonte: Costa et al, 2012] .....	39
Figura 12: Classificador de Novidades.....	40
Figura 13: Distribuição de pontuação de clientes e impostores. [Fonte: Jain et al, 2011] .....	44
Figura 14: Impacto da variação do limiar ( <i>threshold</i> ) nos valores de FAR e FRR. ....	45
Figura 15: Exemplo de Curva ROC com 5 limiares (A, B, C, D e E).....	46
Figura 16: a) Exemplos de faces da base ORL; b) face média da base de faces ORL. ....	49
Figura 17: Vetor de características usando PCA .....	50
Figura 18: Matriz de características usando 2DPCA.....	50
Figura 19: Matriz de características usando 2DPCA bidirecional .....	51
Figura 20. Taxa de reconhecimento rank 1 versus número de autovetores usando PCA.....	55
Figura 21. Taxa de reconhecimento rank 1 versus número de autovetores usando 2DPCA....	56
Figura 22. Taxa de reconhecimento rank 1 versus número de autovetores usando $(2D)^2PCA$ .....	56
Figura 23. Curva CMC comparando os classificadores usando PCA .....	57
Figura 24. Curva CMC comparando os classificadores usando 2DPCA .....	57
Figura 25. Curva CMC comparando os classificadores usando $(2D)^2PCA$ .....	58
Figura 26. CAR (FAR=1%) versus número de autovetores usando PCA.....	59
Figura 27. CAR (FAR=1%)versus número de autovetores usando 2DPCA.....	60
Figura 28. CAR (FAR=1%)versus número de autovetores usando $(2D)^2PCA$ .....	60
Figura 29. CAR (FAR=0,1%) versus número de autovetores usando PCA.....	61
Figura 30. CAR (FAR=0,1%) versus número de autovetores usando 2DPCA.....	61
Figura 31. CAR (FAR=0,1%) versus número de autovetores usando $(2D)^2PCA$ .....	62
Figura 32. EER versus número de autovetores usando PCA .....	62
Figura 33. EER versus número de autovetores usando 2DPCA .....	63

Figura 34. EER versus número de autovetore usando $(2D)^2$ PCA.....	63
Figura 35. Curva ROC comparando os classificadores usando PCA.....	64
Figura 36. Curva ROC comparando os classificadores usando 2DPCA.....	64
Figura 37. Curva ROC comparando os classificadores usando $(2D)^2$ PCA.....	65

## LISTA DE TABELAS

Tabela 1: Resumo de artigos selecionados .....	18
Tabela 2: Banco de faces ORL .....	49
Tabela 3: Síntese do experimentos .....	53
Tabela 4: Melhores taxas de reconhecimento <i>rank 1</i> .....	58
Tabela 5: Melhores valores de CAR (FAR=1%) .....	65
Tabela 6: Melhores valores de CAR (FAR=0,1%) .....	66
Tabela 7: Melhores taxas de erro equivalente (EER) .....	66
Tabela 8: Melhores valores de Área sob a curva ROC (AUC) .....	67

## LISTA DE ABREVIATURAS E SIGLAS

<b>2DPCA</b>	Análise de Componentes Principais Bidimensional
<b>2D<sup>2</sup>PCA</b>	Análise de Componentes Principais Bidimensional e Bidirecional
<b>AUC</b>	<i>Area Under Curve</i>
<b>CAR</b>	<i>Correct Acceptance Rate</i>
<b>CMC</b>	<i>Cumulative Match Characteristic</i>
<b>CN</b>	Classificador de Novidades
<b>CRR</b>	<i>Correct Rejection Rate</i>
<b>DE</b>	Distância Euclidiana
<b>EER</b>	<i>Equivalent Error Rate</i>
<b>FAR</b>	<i>False Acceptance Rate</i>
<b>FERET</b>	<i>Facial Recognition Technology</i>
<b>FRR</b>	<i>False Rejection Rate</i>
<b>IEEE</b>	<i>Institute of Electrical and Electronics Engineers</i>
<b>kNN</b>	<i>k-nearest neighbors</i>
<b>LDA</b>	<i>Linear Discriminat Analysis</i>
<b>ORL</b>	<i>Olivetti Research Laboratory</i>
<b>ROC</b>	<i>Receiver Operating Characteristic</i>
<b>PCA</b>	<i>Principal Components Analysis</i>
<b>SVM</b>	<i>Support Vector Machine</i>

# SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>13</b>
<b>1.1 Objetivo .....</b>	<b>14</b>
1.1.1 Geral .....	14
1.1.2 Específico .....	14
<b>1.2 Organização do trabalho .....</b>	<b>14</b>
<b>2 REVISÃO BIBLIOGRÁFICA .....</b>	<b>16</b>
<b>2.1 Introdução .....</b>	<b>16</b>
<b>2.2 Pesquisa Bibliográfica .....</b>	<b>20</b>
2.2.1 Banco de Imagens.....	20
2.2.2 Pré-processamento.....	21
2.2.3 Extração de características.....	23
2.2.4 Classificação.....	25
<b>3 FUNDAMENTOS TEÓRICOS.....</b>	<b>28</b>
<b>3.1 Introdução .....</b>	<b>28</b>
<b>3.2 Pré-processamento.....</b>	<b>28</b>
3.2.1 Equalização de Histograma .....	28
<b>3.3 Extração de características .....</b>	<b>29</b>
3.3.1 Vetor de características.....	29
3.3.2 Análise de Componentes Principais - PCA .....	30
3.3.3 Análise de Componentes Principais Bidimensional - 2DPCA.....	33
3.3.4 2DPCA Bidirecional - (2D) <sup>2</sup> PCA.....	34
<b>3.4 Classificação .....</b>	<b>35</b>
3.4.1 Métrica ou Distância.....	35
3.4.2 Algoritmo do Vizinho mais Próximo .....	37
3.4.3 Filtro de Novidades .....	38
3.4.4 Classificador de Novidades .....	40
<b>3.5 Sistema Biométricos .....</b>	<b>40</b>
3.5.1 Identificação .....	41
3.5.2 Verificação.....	41

<b>3.6 Medidas de desempenho .....</b>	<b>42</b>
3.6.1 Identificação .....	42
3.6.2 Verificação.....	43
<b>4 MATERIAIS E MÉTODOS .....</b>	<b>48</b>
<b>4.1 Introdução .....</b>	<b>48</b>
<b>4.2 Banco de imagens .....</b>	<b>48</b>
4.2.1 Banco de Imagens ORL.....	48
<b>4.3 Pré-processamento das imagens.....</b>	<b>49</b>
<b>4.4 Extração de características .....</b>	<b>49</b>
<b>4.5 Técnicas de partição dos dados .....</b>	<b>51</b>
<b>4.6 Testes Experimentais.....</b>	<b>52</b>
<b>5 RESULTADOS E DISCUSSÕES .....</b>	<b>55</b>
<b>5.1 Introdução .....</b>	<b>55</b>
<b>5.2 Experimentos no modo identificação.....</b>	<b>55</b>
5.2.1 Impacto do número de autovetores versus taxa de reconhecimento <i>rank 1</i> .....	55
5.2.2 Curvas CMC .....	56
5.2.3 Taxa de Reconhecimento <i>rank 1</i> .....	58
<b>5.3 Experimentos no modo verificação .....</b>	<b>59</b>
5.3.1 Impacto do número de autovetores versus CAR .....	59
5.3.2 Impacto do número de autovetores versus EER.....	62
5.3.3 Curvas ROC.....	63
5.3.4 Resumo .....	65
<b>5.4 Discussão dos resultados .....</b>	<b>67</b>
<b>6 CONCLUSÃO.....</b>	<b>69</b>
<b>APÊNDICE .....</b>	<b>73</b>

# 1 INTRODUÇÃO

Um dos grandes desafios da Inteligência Computacional é entender como as pessoas se reconhecem através da face e desenvolver sistemas de reconhecimento faciais automáticos e confiáveis.

Estes desafios sustentam a Biometria, a ciência que autentica as pessoas através da medição de suas características físicas, tais como reconhecimento de íris e de impressão digital, ou características comportamentais ou psicológicas, tomando como exemplo a dinâmica de digitação, a análise de marcha (jeito de andar de uma pessoa) e reconhecimento por meio da voz. Essas características físicas e comportamentais não são necessariamente independentes, como no caso da face humana, que é uma mistura de características físicas e expressões emocionais. Essas características devem satisfazer os seguintes requisitos: universalidade (todos os indivíduos devem possuir a característica), unicidade ou distinção (dois indivíduos não podem ser o mesmo com relação à característica), permanência ou invariância (a característica deve ser invariante no tempo), e mensurabilidade (a característica deve ser mensurável). Na prática, ainda existem outros requisitos, como o desempenho (acurácia de identificação aceitável), aceitabilidade (aceitação do sistema biométrico pelos usuários) e resistência a fraude (LI e JAIN, 2011; WECHSLER, 2007) .

O Reconhecimento Facial tem se tornado uma das maiores tecnologias da Biometria, pois reconhecer indivíduos é o primeiro passo para construir máquinas inteligentes que possam trabalhar em ambientes humanos. A habilidade de reconhecer seres vivos em fotografias ou vídeos é uma inovação para uma vasta faixa de aplicações que incluem defesa, saúde, interação homem-computador, recuperação de imagem e mineração de dados, robótica industrial e pessoal, vigilância e segurança, e transporte (WECHSLER, 2007). O reconhecimento facial possui várias vantagens com relação a outros métodos biométricos, tais como íris e impressão digital, pois além de ser natural e não intrusivo, a face tem a possibilidade de ser captada à distância e de maneira secreta, sem consentimento do indivíduo (LI e JAIN, 2011).

Um sistema completo de reconhecimento facial pode ser resumido em duas partes: detecção de face e reconhecimento de face. A detecção facial limita a região da face na imagem de entrada. Uma localização precisa da face possibilita uma segmentação do contorno facial mais eficaz, aumentando a acurácia no processo de identificação e o reconhecimento de indivíduos (PADILLA, 2012). O Reconhecimento Facial tem como objetivo identificar a qual

pessoa determinada imagem de face pertence e pode ser dividido em três etapas: pré-processamento da imagem de face, extração de características e classificação. Geralmente, na avaliação do desempenho de sistemas de reconhecimento de indivíduos, as imagens de face são agrupadas em um conjunto de imagens para treinamento e outro para teste. No primeiro, o indivíduo correspondente a cada imagem de face de treinamento é previamente conhecido, no outro, as imagens de face de teste são submetidas ao processo de classificação para obtenção do desempenho do sistema (WECHSLER, 2007).

Apesar das pesquisas terem iniciado nos anos 60, sistemas de reconhecimento facial robustos e precisos ainda se constituem desafios para os pesquisadores de visão computacional e reconhecimento de padrões, especialmente sob ambientes não controlados.

## **1.1 Objetivo**

### **1.1.1 Geral**

O objetivo geral proposto nesta pesquisa é avaliar o reconhecimento facial usando o Classificador de Novidades.

### **1.1.2 Específico**

Os objetivos específicos consistem em:

- Avaliar o desempenho do classificador de novidades usando as técnicas de extração de características PCA, 2DPCA e  $(2D)^2$ PCA no reconhecimento facial.
- Comparar os classificadores de novidades e vizinho mais próximo no reconhecimento facial;

## **1.2 Organização do trabalho**

Este trabalho está dividido em seis capítulos descritos a seguir.

O Capítulo 1 apresenta uma introdução ao reconhecimento facial e seus problemas e desafios na área de reconhecimento de padrões, e também aborda os objetivos gerais e específicos propostos para a dissertação.

O Capítulo 2 contém trabalhos relevantes para a realização desta pesquisa e são relacionados com a metodologia de reconhecimento facial adotada, formando assim a revisão bibliográfica.

O Capítulo 3 engloba a fundamentação teórica, apresentando os conceitos básicos e matemáticos das técnicas empregadas no desenvolvimento deste trabalho.

O Capítulo 4 aborda os materiais e a metodologia que serão utilizados para alcançar os objetivos propostos.

O Capítulo 5 mostra resultados e discussões.

O Capítulo 6 são apresentadas as conclusões.

## 2 REVISÃO BIBLIOGRÁFICA

### 2.1 Introdução

O reconhecimento facial é um dos tópicos de pesquisa mais atuais na área de reconhecimento de padrões e processamento de imagens. É amplamente usado em medicina, controle de acesso, detecção forense, sistemas de vigilância e monitoração, robótica e interação homem máquina. Sendo assim, possui larga aplicação e valor comercial agregado. A maioria dos atuais algoritmos de extração de características para reconhecimento de faces é baseado em métodos estatísticos, tais como: *Eigenface* (PCA), 2DPCA, 2D<sup>2</sup>PCA, *Fisherface* (LDA), 2DLDA, 2D<sup>2</sup>LDA. Na fase de classificação, destacam-se a aplicação dos algoritmos Vizinho Mais Próximo (kNN), Redes Neurais e Máquina de Vetores de Suporte (SVM) (Oliveira et al, 2011).

Nesse trabalho, a principal base de dados empregada na revisão bibliográfica foi a IEEE (Institute of Electrical and Electronics Engineers) e a Science Direct.

Com base nos trabalhos lidos, conclui-se que um sistema completo de reconhecimento facial pode ser resumido em duas partes: detecção de face e reconhecimento de face. A detecção facial, primeira etapa de sistemas biométricos através da face, limita a região da face na imagem de entrada. Uma localização precisa da face possibilita uma segmentação do contorno facial mais eficaz, aumentando a acurácia no processo de identificação e reconhecimento de indivíduos (Padilla, 2012). Já o Reconhecimento Facial tem como objetivo identificar à qual pessoa determinada imagem de face pertence. Geralmente, as imagens de face são agrupadas em imagens para treinamento e para teste, o indivíduo correspondente a cada imagem de face de treinamento é previamente conhecido, já as imagens de face de teste são submetidas ao processo de classificação (Wechsler, 2007).

Este trabalho aborda somente a segunda parte, o reconhecimento facial, que pode ser dividida em três fases: pré-processamento, extração de característica e classificação.

Na Tabela 1 apresentam-se os principais artigos selecionados durante a pesquisa bibliográfica, organizados de acordo com as fases típicas do sistema de reconhecimento facial referidas acima. Primeiramente, serão apresentados os bancos de imagens de faces mais utilizados para teste e validação dos métodos propostos. Em seguidas serão discutidos os trabalhos apresentados na Tabela 1, sob a ótica das fases típicas de um sistema de reconhecimento facial.

Nessa revisão serão abordados também artigos relacionados com o classificador de novidades, método proposto nesse trabalho para a etapa de classificação no sistema de reconhecimento facial.

Tabela 1: Resumo de artigos selecionados

Ano	Autor(es)	Título	Materiais	Pré-processamento	Característica	Extração de características	Partição de dados	Classificador	Medida de Desempenho	Resultados
2004	Jian Yang, David Zhang, Alejandro F. Frangi e Jing-yu Yang	Two-dimensional PCA: a New Approach to Appearance-based Face Representation and Recognition	ORL, AR e Yale	Não	Intensidade do pixel	PCA; 2DPCA	Holdout (5 imagens de treinamento) e Leave-one-out	1NN (distância euclidiana)	Taxa de reconhecimento	PCA: 93,5% 2DPCA: 96% (ORL-Holdout); PCA: 97,5% 2DPCA: 98,3% (ORL-Leave-one-out)
2005	Daoqiang Zhang e Zhi-Hua Zhou	(2D)2PCA - Two-directional two-dimensional PCA for efficient face representation and recognition	ORL e FERET	FERET: imagens redimensionadas para 60x60 pixels	Intensidade do pixel	(2D) <sup>2</sup> PCA	Holdout (5 imagens de treinamento)	1NN (distância euclidiana)	Taxa de reconhecimento	ORL: 90,5% FERET: 85%
2005	Jian Yang, David Zhang, Xu Yong, Jing-yu Yang	Two-dimensional discriminant transform for face recognition	ORL	Não	Intensidade do pixel	LDA; 2DLDA	Holdout (5 imagens de treinamento)	1NN	Taxa de reconhecimento	LDA: 89,5% 2DLDA: 96,4%
2006	S. Noushath, G. Hemantha Kumar, P. Shivakumara	(2D)2LDA: An efficient approach for face recognition	ORL e Yale	Não	Intensidade do pixel	PCA; 2DPCA; (2D) <sup>2</sup> PCA; 2DLDA; (2D) <sup>2</sup> LDA	Holdout (5 imagens de treinamento)	1NN (distância euclidiana)	Taxa de reconhecimento	PCA: 95,5% 2DPCA: 97% (2D) <sup>2</sup> PCA: 97,75% 2DLDA: 98% (2D) <sup>2</sup> LDA: 98,5%
2011	C. MageshKumar, R. Thiyagarajan, S.P. Natarajan, S. Arulsevi, G.Sainarayanan	Gabor features and LDA based Face Recognition with ANN classifier	ORL	imagens redimensionadas para 128x128 para o filtro Gabor	Coeficientes da transformada Gabor wavelet	PCA; PCA+LDA	Holdout (5 imagens de treinamento)	1NN (distância euclidiana); Rede neural ( <i>Back propagation</i> )	Taxa de reconhecimento	1NN+PCA+LDA: 97,8%; ANN+PCA+LDA: 98,6%
2011	Shermina J.	Illumination invariant face recognition using Discrete Cosine Transform and Principal Component	Yale B (10 pessoas)	Normalização e Correção da iluminação usando DCT	Intensidade do pixel	PCA	Não menciona	Não menciona	Acurácia	94.2%

<b>Ano</b>	<b>Autor(es)</b>	<b>Título</b>	<b>Materiais</b>	<b>Pré-processamento</b>	<b>Característica</b>	<b>Extração de características</b>	<b>Partição de dados</b>	<b>Classificador</b>	<b>Medida de Desempenho</b>	<b>Resultados</b>
		Analysis								
2011	Oliveira, Luiz; Koerich, Alessandro; et al	2D Principal Component Analysis for Face and Facial-Expression Recognition	ORL, AR, Feret, Yale e Jaffe	Não	Intensidade do pixel	2DPCA + Seleção de características usando Algoritmos Genéticos	Particiona do em Treino, Validação, Busca e Teste.	kNN e SVM	Taxa de reconhecimento	ORL-kNN: 91,5% ORL-SVM: 93,1%
2011	Thai Hoang Le, Len Bui	Face Recognition Based on SVM and 2DPCA	ORL e Feret	imagens redimensionadas para 50x50 (FERET)	Intensidade do pixel	PCA; 2DPCA	Holdout	kNN e SVM	Acurácia	ORL PCA+SVM: 95,7% PCA+kNN: 95,2% 2DPCA+SVM: 97,3% 2DPCA+kNN: 96,2%

## 2.2 Pesquisa Bibliográfica

### 2.2.1 Banco de Imagens

Alguns grupos de pesquisadores e instituições disponibilizaram bases de faces humanas para a comunidade acadêmica. As bases mais usadas são: Yale, Yale B, ORL, AR, Feret e Jaffe, as quais serão discutidas a seguir.

A base de faces Yale possui 165 imagens em tons de cinza em formato GIF de tamanho 320x243 pixels de 15 indivíduos. São 11 imagens por pessoa, onde há variação de iluminação (luz proveniente do centro, da direita e da esquerda), de expressão facial (normal, feliz, triste, sonolento, surpreso e piscando os olhos) e indivíduos com e sem óculos. Em Yang et al (2004), a base de faces Yale é usada para avaliar a performance do sistema quando ambas as variações em iluminação e expressões faciais estão presentes. Yang *et al* (2004), Noushatha et al (2006) e Oliveira et al (2011) usaram esta base em seus experimentos. Esta base pode ser encontrada em: <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>

A base de faces Yale B possui 5.760 imagens faciais em tons de cinza de tamanho 640x480 pixels de 10 pessoas, cada uma vista sob 576 condições de visualização (9 poses x 64 condições de iluminação). Para cada indivíduo em uma específica pose, uma imagem com iluminação ambiente foi capturada, totalizando, de fato, em 5850 (5760 + 90) imagens. Shermina (2011) usou esta base para lidar com o problema da variação na iluminação. Esta base pode ser encontrada em: <http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html>.

A base de dados ORL (Laboratório de Pesquisa da Olivetti) contém 10 imagens diferentes de cada 40 indivíduos, totalizando 400 imagens de tamanho 92x112 pixels. Para alguns indivíduos, as imagens foram capturadas em diferentes épocas, variando em iluminação, expressões faciais (olhos abertos e fechados, sorrindo ou não) e com detalhes na face (com ou sem óculos). Todas as imagens foram capturadas com um fundo negro homogêneo. Yang *et al* (2004), Yang *et al* (2005), Zhang e Zhou (2005), Noushatha et al (2006), MageshKumar et al (2011), Oliveira et al (2011) e Le e Bui (2011) usaram esta base em seus experimentos. Esta base pode ser encontrada em: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>.

O banco de faces AR consiste de mais de 3.000 imagens faciais coloridas de tamanho 768 x 576 pixels de 126 indivíduos (70 homens e 56 mulheres), variando em iluminação, oclusão, idade e expressão facial (neutro, sorrindo, zangado, gritando). Este banco é muito

popular no que se refere a oclusões naturais, e é considerado bastante desafiante, visto que 50% das imagens possuem muitas oclusões (Martínez e Benavente, 1998; Jia e Martínez, 2008). Yang *et al* (2004) e Oliveira et al (2011) usaram esta base em seus experimentos.

O banco de faces FERET contém 14.126 imagens de 1.199 indivíduos. Alguns conjuntos de imagens foram tiradas em dias diferentes e outros ao longo de 2 anos, habilitando pesquisadores a estudar o impacto da variação no tempo. (Phillips et al, 2000), Zhang e Zhou (2005), Oliveira et al (2011) e Le e Bui (2011) usaram esta base em seus experimentos.

### 2.2.2 Pré-processamento

Nessa fase, a imagem é processada de forma a facilitar a extração das características relevantes para a identificação. Como o pixel da imagem é uma característica da face, logo, é importante e não pode ser descartado. As vezes, há a necessidade de um processamento, denominado de padronização ou normalização da imagem, o qual consiste em ajustar a imagem a um tamanho pré-definido, centralizar os olhos, ajustar a escala, enfim, buscar um meio de tornar possível a comparação entre imagens de diferentes indivíduos ou do mesmo indivíduo capturadas em diferentes condições.

Sahoolizadeh e Ghassabeh (2008) primeiramente redimensionaram as imagens de entrada para o tamanho de 40 x 40 pixels a fim de remover a informação do fundo e obter somente os detalhes da imagem de face. A seguir, equalizaram o histograma para espalhar a energia de todos os pixels e normalizaram as imagens para equalizar a energia relacionada com cada imagem de face. Um exemplo de resultado pode ser visto na Figura 1, a qual possui 10 imagens de face para cada indivíduo.



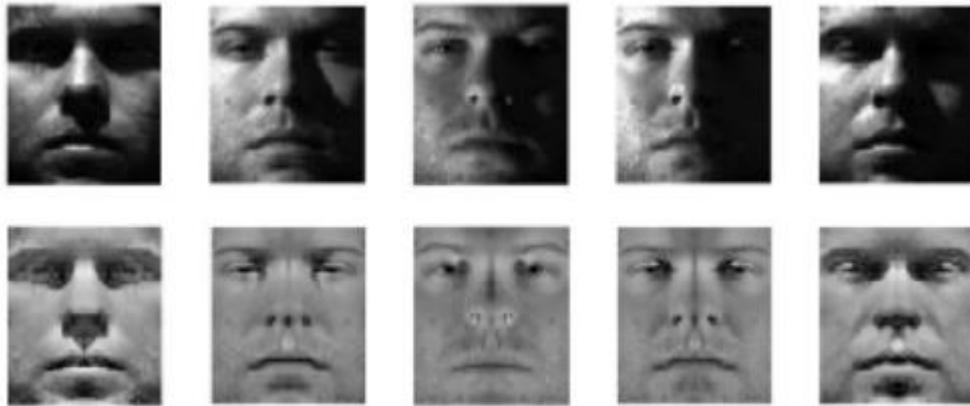
**Figura 1: Cada linha contém 10 imagens pré-processadas de um indivíduo [Fonte: Sahoolizadeh e Ghassabeh, 2008]**

Shermina (2011) e Liau e Isa (2010), com o intuito de corrigir a luminância não uniforme, fazem a normalização da iluminação através da construção de uma imagem com os componentes de baixa frequência da Transformada Discreta do Cosseno (DCT), encontra a média desta imagem ( $m$ ), aplica sobre cada pixel a seguinte equação:

$$c(x, y) = 0,5 * (m - \log L(x, y)) \quad (1)$$

Onde  $c(x, y)$  é o valor do novo pixel e  $L(x, y)$  a luminância. A variação da iluminação é corrigida usando a mesma banda de frequência da reflectância através das propriedades ímpar e par dos componentes da DCT na direção horizontal, gerando duas novas imagens:  $I_{\text{ímpar}}$  e  $I_{\text{par}}$ . Shermina (2011) considera que a face humana é simétrica, logo os pixels do lado esquerdo e direito de uma imagem de face são comparados. Se um pixel do lado direito é positivo e o pixel correspondente do lado esquerdo é negativo, ambos são ajustados através de um termo de compensação  $I_0$ , corrigindo assim a variação iluminação da imagem.

Este método considera que os componentes ímpares da DCT dizem respeito ao ruído e os componentes pares à informação desejada, sendo assim, a nova intensidade do pixel é definida como  $I(x, y) = I_0 + I_{\text{ímpar}}$ . No final, as imagens pré-processadas diferem-se das originais, pois um lado da face se baseia no outro. Este método atingiu uma acurácia de 94,2%, com uma taxa de falso positivo de 5,84% e taxa de falso negativo de 7,51%. A Figura 2 mostra um exemplo de imagens após este método de pré-processamento, onde é possível visualizar que partes da face (olhos e narizes, por exemplo) foram criados, pelo fato de que uma imagem de face não é exatamente simétrica.



**Figura 2: Imagens resultantes do método. [Fonte: Liau e Isa, 2010]**

No trabalho de MageshKumar et al (2011), as imagens foram apenas redimensionadas para 128x128 pixels para facilitar as representações da transformada Wavelet de Gabor, a qual tem como saída um vetor de características com uma dimensão alta. Esse vetor de características é passado, então, à etapa de redução de dimensionalidade, que utiliza a técnica PCA.

Yang *et al* (2004), Yang *et al* (2005), Zhang e Zhou (2005), Noushatha et al (2006), Oliveira et al (2011) e Le e Bui (2011) usam diretamente a intensidade do pixel em escala de cinza, sem a necessidade de um pré-processamento.

Como é visto no trabalho de Han et al (2013), o pré-processamento melhora o desempenho quando usado em imagens que possuem variação na iluminação, tais como as imagens de face da base Yale B.

### 2.2.3 Extração de características

O objetivo das técnicas de extração de características é encontrar um conjunto mínimo de características que, ainda assim, possam identificar de forma exclusiva uma face. É, também nesta etapa do processo que se concentra o maior esforço dos trabalhos acerca do reconhecimento facial. Vários métodos foram propostos ao longo dos anos a fim de se conseguir parâmetros que melhor caracterizassem a face de um indivíduo e assim viabilizar a tarefa de reconhecimento de indivíduos.

Várias formas de extração de características de imagens de face humana vêm sendo propostas, como as que usam PCA (Sirovich e Kirby, 1990; Turk e Pentland, 1991), 2DPCA (Yang *et al*, 2004),  $(2D)^2$ PCA (Zhang e Zhou (2005)), LDA (Belhumeur et al, 1997), 2DLDA (Yang et al, 2005),  $(2D)^2$ LDA (Noushath et al, 2006), entre outras.

Sirovich e Kirby (1990) foram os primeiros a usarem PCA eficientemente para representar imagens de faces humanas. Eles argumentaram que qualquer imagem de face poderia ser reconstruída aproximadamente como uma soma ponderada de uma pequena coleção de imagens que definiriam a base de faces (*eigenimages*), e a média da imagem da face. Dentro deste contexto, Turk e Pentland (1991) apresentaram o método *Eigenfaces* para reconhecimento facial, o qual já é bastante difundido atualmente. Desde então, PCA tem sido amplamente investigado e tornou-se uma das abordagens de maior sucesso no reconhecimento facial.

Seguindo os passos do PCA, o método de análise de componentes principais bidimensional (2DPCA) é proposto por Yang *et al* (2004), e difere do convencional (PCA) pelo fato de usar as matrizes de imagens bidimensionais (2D) diretamente no processo de obtenção da matriz de dispersão, sem a necessidade da transformação da matriz de imagem de um vetor unidimensional (1D). O algoritmo tem o mesmo objetivo do PCA, que é encontrar vetores de projeção ótimos, que compõem a matriz de projeção. A redução de dimensão é feita na direção horizontal da matriz da imagem. Este método é superior ao PCA em termos de eficiência computacional, pois calcula a matriz de covariância mais rápido. No entanto, o vetor de característica do 2DPCA é maior que o do PCA. Para solucionar este problema é recomendado aplicar PCA após o uso do 2DPCA para redução de dimensionalidade .

Como 2DPCA requer mais coeficientes para representar uma face do que PCA, Zhang e Zhou (2005) apresentaram o método bidirecional 2DPCA, também conhecido por  $(2D)^2PCA$ . Este método é uma extensão do método 2DPCA, que realiza redução em ambas as direções horizontal e vertical da matriz da imagem. Dessa forma, consegue representar uma face de forma eficiente, sem impactar na taxa de reconhecimento. O autor testou o método nas bases de faces ORL e FERET com classificador vizinho mais próximo, obtendo uma acurácia de 90,5% e 85%, respectivamente. Já Oliveira *et al* (2011) propuseram um método baseado em algoritmos genéticos multiobjetivos para encontrar os coeficientes mais discriminantes do 2DPCA, realizando os testes nas bases ORL, Yale, AR, Feret e Jaffe (*Japanese Female Facial Expression*).

O método LDA foi introduzido por Fisher (1936). No entanto, difundiu-se para a área de reconhecimento de faces a partir da publicação do artigo de Belhumeur *et al.* (1997), no qual os autores compararam PCA com LDA através do reconhecimento de imagens de pessoas com grandes variações de iluminação, e mostraram que o espaço de características criado pela transformação LDA proporcionou melhores resultados que o PCA.

Um outro método, denominado de Análise Discriminante de Fisher Bidimensional (2DLDA), foi aplicado no reconhecimento de faces usando o banco de faces ORL. O método é similar ao LDA convencional, mas ao invés de transformar a matriz de imagem em um vetor unidimensional (1D), as próprias matrizes bidimensionais (2D) são utilizadas diretamente no processo de obtenção da matriz de dispersão intraclasse e da matriz de dispersão interclasses, procurando encontrar um conjunto de vetores de projeção que tanto reduza a dimensão do espaço de imagens, como também maximize a separação (discriminação) entre as classes de imagens. Foi empregado o classificador vizinho mais próximo, com a métrica da distância Euclidiana, obtendo uma acurácia de 94%, contra 92,5% do 2DPCA (Yang et al, 2005).

Uma extensão do método 2DLDA é denominado de  $(2D)^2LDA$ . O 2DLDA realiza a redução da matriz original apenas na direção horizontal, enquanto que o  $(2D)^2LDA$  realiza a redução da matriz original nas direções horizontal e vertical. Com isso, a imagem projetada (matriz de característica) por este método é de dimensão muito menor do que a matriz de característica projetada pelo 2DLDA. Foi empregado o classificador vizinho mais próximo, com a métrica da distância Euclidiana, obtendo o melhor resultado de acurácia de 99,75% (base de faces ORL) e 99,33% (base de faces Yale) (Noushatha et al, 2006).

MageshKumar et al (2011) utilizaram como características os coeficientes da transformada Gabor Wavelet para formar um vetor de características de dimensão alta, ao qual foi aplicado PCA para redução de dimensionalidade e LDA para a extração de características. Após este processo, foram geradas 199 características que foram usadas como conjunto de treinamento para treinar a Rede Neural *Back Propagation*.

#### 2.2.4 Classificação

Esta última fase do processo de reconhecimento facial consiste em comparar as características extraídas de uma imagem de teste com as já previamente extraídas das imagens de treinamento e expressar essa comparação através de uma medida de similaridade, efetuando o reconhecimento ou não de um determinado indivíduo.

No trabalho de MageshKumar et al (2011), a fase de classificação usou uma rede neural *Back Propagation* com 199 neurônios na camada de entrada, pois há 199 características extraídas por PCA e LDA, e 30 na camada escondida. A rede neural foi treinada com 200 imagens de treinamento. Foram utilizadas 200 imagens de teste da base de faces ORL. Com a rede neural treinada, uma face desconhecida é submetido a teste, e então é

extraído o vetor característica da mesma, o qual é passado a entrada da rede para classificação. O melhor resultado foi uma acurácia de 98,6%.

Yang *et al* (2004) usaram o método Vizinho Mais Próximo (1NN) com a métrica da distância euclidiana em seus experimentos. Atingiram uma acurácia de 96% (estratégia metade treinamento, metade teste) e 98,3% (*Leave-one-out*) para testes relacionados à variação com o tempo, expressões faciais, detalhes faciais (presença de óculos), variação nas poses e escalas usando a base de faces ORL. Usando a base de faces AR, obtiveram acurácia de 67,6% para testes relacionados à diferentes épocas em que as imagens foram capturadas, 96,1% para testes relacionados à expressões faciais, e 89,8% para testes relacionados à variação de iluminação. Já na base de faces Yale foram: acurácia de 84,24% (estratégia *Leave-one-out*) para testes relacionados à expressões faciais e variação de iluminação. Foi observado que quanto maior o número de imagens de testes e de componentes principais, maior a acurácia.

Le e Bui (2011) avaliaram um sistema de reconhecimento facial usando 2DPCA (extração de características) e SVM (classificação), aplicando SVM através da decomposição *One-Against-All* para transformar um problema de várias classes em um problema de duas classes, usando as funções de kernel polinomial e de base radial (RBF). Os experimentos foram aplicados nas bases de face FERET e ORL. As faces das imagens da base FERET foram extraídas manualmente através do algoritmo de detecção de faces de Viola e Jones (2004). Logo em seguida as imagens foram redimensionadas para 50 x 50 pixels e aleatoriamente separou-se a base de faces em 3 subconjuntos, sob os quais a validação cruzada foi aplicada para obter os resultados. Este experimento resultou em uma acurácia de 97,3% (base ORL) e 95,1% (base FERET).

Oliveira *et al* (2011) realizaram testes com o classificador vizinho mais próximo e com o classificador SVM (kernel linear) nas bases ORL, Yale, AR, Feret e Jaffe, obtendo uma acurácia de 91,5%, 86,7%, 89,2%, 87,9% e 91%, com o classificador vizinho mais próximo, e com o classificador SVM obtiveram uma acurácia de 93,1%, 90%, 92,1%, 89,2% e 94%, respectivamente.

Todos os experimentos em Yang *et al* (2004), Yang *et al* (2005), Zhang e Zhou (2005), Noushatha et al (2006) utilizaram a distância euclidiana na classificação para avaliação dos métodos propostos: 2DPCA, 2DLDA, 2D<sup>2</sup>PCA e 2D<sup>2</sup>LDA.

O método de classificação que será empregado no reconhecimento facial é o classificador de novidades. Essa técnica foi empregada nos trabalhos de Costa et al (2012), para detecção de vazamento de gás em instalações petrolíferas, em Costa et al (2013), para reconhecimento de íris e em Costa e Moura (1995), para detecção de câncer de mama.

Em Costa et al (2012), o classificador de novidades foi aplicado com o objetivo de detectar vazamento de gás natural através das imagens capturadas por uma câmera CCD. Foi usada uma base de 60 imagens, com dois grupos de treinamento, um com 20 imagens e outro com 30 imagens. Em ambos os casos, o filtro novidade classificou corretamente todas as 60 imagens. Os melhores resultados atingiram uma sensibilidade e especificidade iguais à 1.

Em Costa et al (2013), os autores propõem um novo método de casamento (*matching*) no reconhecimento de íris, o qual usa o conceito do classificador de novidades. Este método é testado usando a base de imagens de íris UBIRIS em modo verificação e identificação usando 3 e 4 imagens para treinamento, pois a base UBIRIS possui apenas 5 imagens por indivíduo. Os melhores resultados obtidos foram uma acurácia de 99,1% (modo verificação) e 97,9% (modo identificação).

O classificador de novidades também foi usado para detecção de lesões em imagens de cintilografia de mama no trabalho de Costa e Moura (1995). No trabalho, propõe-se um método automatizado de diagnóstico, possibilitando a classificação de lesões como malignas ou benignas, caso sejam detectadas. Um conjunto de imagens cintilográficas de mamas saudáveis é utilizado para compor o conjunto de treinamento, e quando uma imagem de teste é submetida à entrada, tem-se na saída uma imagem cujos pixels referentes a regiões de lesões são marcados. Os melhores resultados atingiram uma sensibilidade de 90% e especificidade de 95% na detecção de tumores, uma sensibilidade de 80% e especificidade de 96% na detecção mamas com lesões e uma sensibilidade de 100% e uma especificidade de 91.4% na detecção de indivíduos normais.

Após esta revisão bibliográfica, fica evidente que a proposta de reconhecer indivíduos através de imagens faciais, utilizando o conceito de filtro de novidade para obtenção de uma medida de similaridade é uma proposta inovadora.

## 3 FUNDAMENTOS TEÓRICOS

### 3.1 Introdução

Neste capítulo descrevem-se os fundamentos matemáticos das técnicas empregadas no desenvolvimento deste trabalho.

Inicialmente descreve-se uma técnica de pré-processamento, a equalização de histograma, ela não será empregada neste trabalho, porém dá uma visão do que pode ser feito nessa fase. Após isso, são descritos os métodos utilizados para extração de características: PCA, 2DPCA e  $(2D)^2$ PCA.

Em seguida, é abordada a técnica proposta nesse capítulo para o reconhecimento de faces, o classificador de novidades e a técnica usada para comparação, o vizinho mais próximo com a métrica da distância euclidiana.

Por último, são apresentados sistemas biométricos, nos quais estão inseridas as curvas *Cumulative Match Characteristic* (CMC) e *Receiver Operating Characteristic* (ROC) utilizada para avaliação do desempenho de ambos os classificadores.

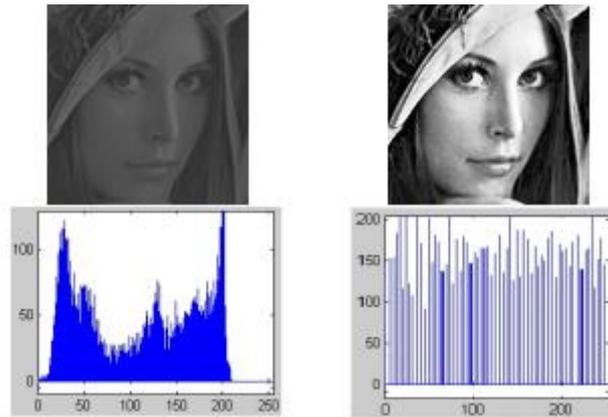
### 3.2 Pré-processamento

#### 3.2.1 Equalização de Histograma

A equalização de histograma é uma técnica a partir da qual se procura redistribuir os valores de tons de cinza dos pixels em uma imagem, de modo a obter um histograma uniforme, no qual o número (percentual) de pixels de qualquer nível de cinza é praticamente o mesmo (Filho e Neto, 1999).

A equalização do histograma é aplicada com o objetivo de melhorar o contraste das imagens, realçando a textura de regiões escuras ou sem nitidez e facilitando com isso as operações de extração de características.

Dois histogramas são mostrados na Figura 3, o da esquerda representa a face sem a aplicação da equalização e o da direita com equalização. Observa-se na figura da esquerda, que sofreu o processo de equalização, que mais detalhes da imagem são realçados.



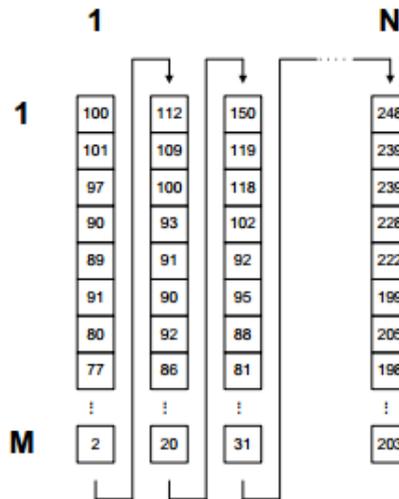
**Figura 3: Equalização do histograma de uma face.**

### 3.3 Extração de características

A extração de características consiste em transformar os dados de entrada em um conjunto de características. Essas características são escolhidas de tal maneira que contenham apenas informações relevantes ao invés de usar toda a informação crua inicial (Theodoridis e Koutroumbas, 2009; Duda et al, 2000). As informações de imagem são armazenadas em vetores de características.

#### 3.3.1 Vetor de características

Uma imagem é uma matriz de dimensão de  $m$  linhas por  $n$  colunas. O vetor de características com todas as informações iniciais da imagem, ou seja, um vetor cru, é obtido através da concatenação das  $n$  colunas da imagem, formando um vetor numérico de dimensão  $mxn$ , onde cada pixel da imagem é uma característica (Theodoridis e Koutroumbas, 2009; Duda et al, 2000). Na Figura 4 é mostrado um exemplo de vetor de características de uma imagem em escala de cinza, os valores de cada item do vetor correspondem aos níveis de cinza da imagem e variam de 0 à 255.



**Figura 4: Vetor de características de uma imagem em escala de cinza. [Fonte: Pinheiro, 2010]**

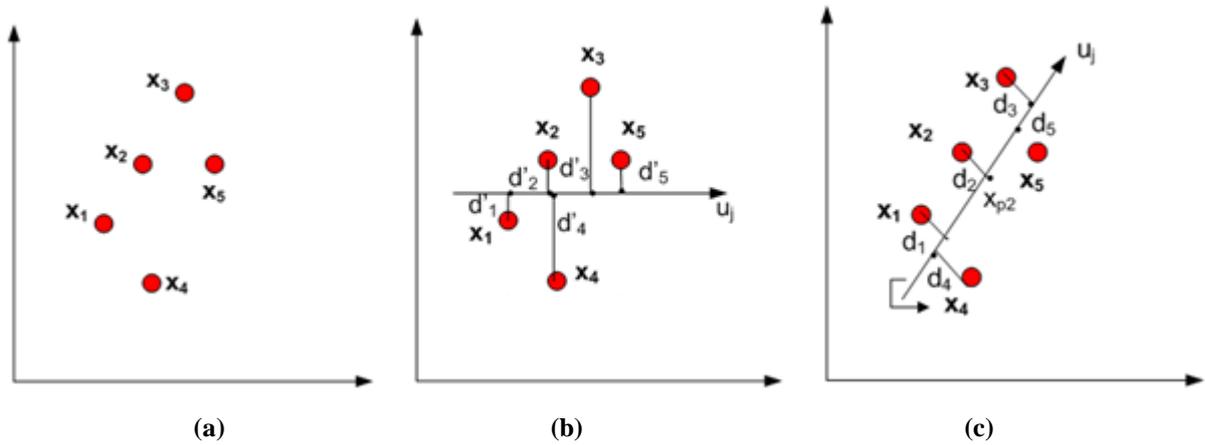
Como exemplo, as imagens da base de dados ORL, as quais possuem dimensão  $92 \times 112$  pixels, resultam em vetores crus de características com dimensão de  $10304 \times 1$ .

Este tipo de conversão de uma matriz bidimensional para um vetor unidimensional é necessário para o método de extração de características PCA. Já para os métodos 2DPCA e  $(2D)^2$ PCA, isto não é necessário, pois os mesmos trabalham diretamente com matrizes, ou seja, não é usado um vetor de características, mas sim uma matriz de características. Dito isso, pode-se abstrair que a dimensão dos vetores crus de características é geralmente muito grande.

Os métodos de extração de características acima citados são utilizados para a redução da dimensionalidade desses vetores. A seguir as três técnicas de extração de características citadas (PCA, 2DPCA e  $(2D)^2$ PCA), as quais foram utilizadas nesse trabalho para a extração de características das imagens da face, serão discutidas.

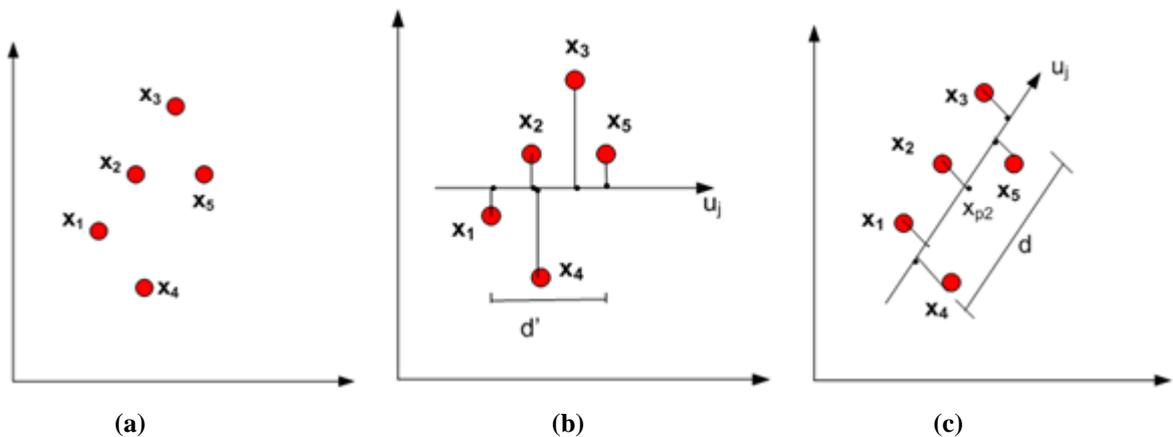
### 3.3.2 Análise de Componentes Principais - PCA

Karl Pearson (1901) definiu PCA como a tarefa de encontrar um subespaço de tal forma que uma função de custo definida como a soma das distâncias dos dados originais para suas projeções nesse subespaço seja minimizada.



**Figura 5: Técnica PCA: (a) conjunto de pontos; (b) subespaço horizontal com distâncias do conjunto de pontos para as projeções não mínimas; (c) subespaço com distâncias do conjunto de pontos para as projeções mínimas.**

Posteriormente, Harold Hotelling (1933) derivou as componentes principais a partir de um problema autovalores e autovetores, porém sem usar ainda a notação matricial e definiu PCA como a tarefa de encontrar um subespaço, de tal forma que a variância dos dados ortogonalmente projetados nesse sub-espaço seja maximizada.



**Figura 6: Técnica PCA: (a) conjunto de pontos; (b) subespaço horizontal com projeções de variâncias não máximas; (c) subespaço com projeções de variâncias máximas. .**

Sirovich e Kirby (1990) foram os primeiros a usarem PCA eficientemente para representar imagens de faces humanas. Eles argumentaram que qualquer imagem de face poderia ser reconstruída aproximadamente como uma soma ponderada de uma pequena coleção de imagens que definiriam como a base de faces (*eigenimages*) e a média da imagem da face. Dentro deste contexto, Turk e Pentland (1991) apresentaram o método *Eigenfaces* para reconhecimento facial, o qual já é bastante difundido atualmente. Desde então, PCA tem

sido amplamente investigado e tornou-se uma das abordagens de maior sucesso no reconhecimento facial.

A seguir é descrito a aplicação da técnica à imagens.

Dado  $N$  imagens de dimensão  $mxn$ , vistas como  $N$  vetores de dimensão  $(m*n)x1=d x1$ , em que  $d=m*n$ . Define-se a matriz de vetores de características crus como  $X = [x_1, x_2, \dots, x_N]$ , cuja dimensão é  $(m*n)xN$ . A matriz de diferença é definida como  $M = [m_1, m_2, \dots, m_N]$ , cuja dimensão é  $(m*n)xN$ , onde  $m_i$  é um vetor coluna de dimensão  $d x1$ , descrito pela equação ( 2 ), e  $\bar{x}$  é o vetor médio, descrito pela equação ( 3 ).

$$m_i = x_i - \bar{x} \quad (2)$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (3)$$

Pode-se mostrar que a variância dos dados ortogonalmente projetados de um conjunto de  $N$  vetores é mínima na direção do autovetor com maior autovalor da matriz de covariância, definida a seguir:

$$C = MM^T \quad (m*n)xN \times N x(m*n) \quad (4)$$

A matriz  $C$ , porém é de dimensão muito grande  $(m*n)x(m*n)$ . Determinar os autovetores e autovalores da mesma é uma tarefa muito trabalhosa, mesmo para imagens pequenas. Logo, é necessário um método computacionalmente adequado para encontrar tais autovetores. Se o número de amostras ( $N$ ) é menor que a dimensão do espaço ( $d=m*n$ ), haverá apenas  $N-1$ , ao invés de  $d$ , autovetores significantes (os autovalores restantes possuem autovalores correspondentes iguais a zero). Felizmente, é possível achar os autovetores de dimensão  $d$  calculando os autovetores de uma matriz bem menor de dimensão  $NxN$ . Se  $N < d$  (onde a maioria dos casos com poucas amostras se encontra), a matriz definida a seguir é usada.

$$S = M^T M \quad (5)$$

Isto é provado desta maneira:  $M^T M w_i = \lambda_i w_i$ , com autovalor  $\lambda_i$  e autovetor associado  $w_i$ . Multiplicando ambos os lados por  $M$ , temos  $MM^T M w_i = \lambda_i M w_i$ , onde verifica-se que  $v_i = M w_i$  é o autovetor de  $C = MM^T$ . Então podemos construir a matriz  $S = M^T M$  (matriz substituta da matriz  $C$ ). Dessa forma, os cálculos são substancialmente reduzidos da

ordem do número de pixels ( $d$ ) para a ordem do número do conjunto de amostras ( $N$ ) (Turk e Pentland, 1991).

A matriz  $U$  é formada pelos  $q$  autovetores que correspondem aos  $q$  maiores autovalores da matriz de covariância  $S$ , e é dada por  $U = [u_1, u_2, \dots, u_q]$  (dimensão  $dxq$ ).

A projeção da matriz  $X$  no subespaço principal é dado por:

$$X_p = U^T \cdot X \quad (6)$$

$U$  é um transformador linear de forma a representar  $X$  em um outro espaço cuja sua matriz de covariância seja diagonal.

A dimensão da matriz  $X_p$  é obtida a partir da equação (6):  $(qxd) (dxN) \rightarrow (qxN)$

A redução de dimensionalidade é obtida considerando que  $q < d$ . Quanto menos autovetores  $u_i$  forem utilizados na montagem da matriz  $U$ , maior a redução da dimensionalidade.

### 3.3.3 Análise de Componentes Principais Bidimensional - 2DPCA

A aplicação do método PCA em imagens exige que uma imagem  $A$  bidimensional seja convertida em um vetor unidimensional  $x$ . Como consequência, as dimensões desse vetor são muito grandes, o que torna a matriz de covariância  $C$  também muito grande. A tarefa de encontrar os autovetores de  $C$  é, então, computacionalmente, muito trabalhosa.

O método análise de componentes principais bidimensional (2DPCA) foi publicado no artigo de Yang *et al* (2004). Neste artigo, os autores aplicaram o método 2DPCA na extração de características para o reconhecimento de imagens de face. Este método difere do convencional (análise de componentes principais, PCA), pelo fato de usar as matrizes de imagens bidimensionais (2D) diretamente no processo de obtenção da matriz de dispersão, sem a necessidade da transformação da matriz de imagem de um vetor unidimensional (1D).

O algoritmo tem o mesmo objetivo do PCA, que é encontrar vetores de projeção ótimos, e é descrito a seguir.

A matriz de projeção  $U = [u_1, u_2, \dots, u_d]$  é formado pelos vetores  $u_1, u_2, \dots, u_d$ , os quais são obtidos através dos  $d$  autovetores associados aos  $d$  maiores autovalores da matriz de dispersão total, de dimensão  $n \times n$ :

$$G_H = \frac{1}{N} \sum_{i=1}^N (A_i - \bar{A})^T \cdot (A_i - \bar{A}) \quad (7)$$

Onde  $N$  é o número de imagens de treinamento,  $A_i$  é a  $i$ -ésima imagem da amostra de  $N$  imagens de dimensão  $mxn$ , e  $\bar{A}$  é a imagem média, descrita pela equação ( 8 ).

$$\bar{A} = \frac{1}{M} \sum_j A_j \quad (8)$$

Após a obtenção da matriz de projeção  $U$ , a imagem original  $A$  é projetada para, então se obter uma nova representação da imagem  $B = AU$  de dimensão  $mx d$ , com  $d < n$ . A redução de dimensão é feita na direção horizontal de  $A$ . Quanto menos autovetores  $u_i$  forem utilizados na montagem da matriz  $U$ , maior a redução da dimensionalidade. A redução da dimensionalidade nas duas dimensões é obtida através da técnica a seguir, 2DPCA Bidirecional.

### 3.3.4 2DPCA Bidirecional - (2D)<sup>2</sup>PCA

No artigo de Zhang e Zhou (2005), os autores apresentaram o método bidirecional 2DPCA, também conhecido por (2D)<sup>2</sup>PCA. Este método é uma extensão do método 2DPCA, que realiza redução nas direções horizontal e vertical da imagem original. Ele consiste em encontrar duas matrizes de projeções, definidas por  $U = [u_1, u_2, \dots, u_d]$  de dimensão  $mx d$  e  $V = [v_1, v_2, \dots, v_q]$  de dimensão  $qxn$ . A matriz  $U$  é obtida através do método 2DPCA e a matriz  $V$  pelo método denominado pelos autores de *alternative 2DPCA*. Este segundo método atua na direção vertical da matriz original, sendo implementado a partir da construção da matriz de dispersão total, de dimensão  $mxm$ , definida pela equação ( 9 ).

$$G_V = \frac{1}{N} \sum_{i=1}^N (A_i - \bar{A})(A_i - \bar{A})^T \quad (9)$$

Em seguida são determinados os  $q$  autovetores correspondentes ao  $q$  maiores autovalores de  $G_V$ , como é feito no método 2DPCA com a matriz  $G_H$ . A matriz de característica  $C$  de dimensão  $qxd$ , é obtida projetando-se a imagem  $A$  de dimensão  $mxn$  nas matrizes  $U$  e  $V$  simultaneamente, ou seja,

$$C = V^T A U \quad (10)$$

A redução de dimensionalidade é obtida de uma dimensão  $m \times n$  para  $q \times d$ , considerando que  $d < n$  e  $q < m$ . Quanto menos autovetores  $u_i$  e  $v_i$  forem utilizados na montagem da matriz  $U$  e  $V$ , maior a redução da dimensionalidade. Além disso, o tempo computacional também é reduzido e a acurácia de reconhecimento é geralmente melhor que a do método 2DPCA, como observado nos trabalhos de Zhang & Zhou, 2005.

A Figura 7 ilustra o método 2DPCA Bidirecional e suas reduções nas direções horizontal e vertical da imagem original  $A$ .

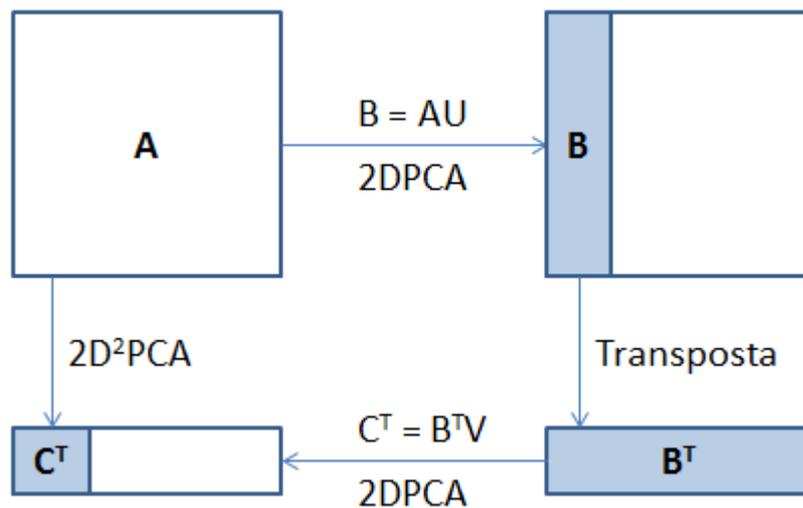


Figura 7: 2DPCA Bidirecional ou  $(2D)^2PCA$

### 3.4 Classificação

O processo de reconhecimento facial é feito a partir de um conjunto de treinamento, ou seja, uma base de faces pré-definida. As estruturas que representam as faces das pessoas (matrizes ou vetores de características) são armazenadas nessa base de faces. Para efetuar a classificação, a imagem da face de uma pessoa a ser reconhecida (amostra de teste) é transformada em um vetor ou matriz, para, então, ser comparada com todos os vetores ou matrizes armazenados nessa base de dados. A seguir aborda-se a teoria que suporta o reconhecimento de faces conforme entendido nesse trabalho.

#### 3.4.1 Métrica ou Distância

Uma métrica  $D(\cdot, \cdot)$  é meramente uma função que retorna um valor escalar entre dois padrões e possui quatro propriedades (Duda et al, 2000):

1. Não negatividade:  $D(a, b) \geq 0$

2. Reflexibilidade:  $D(a,b)=0$ , se e somente se  $a=b$
3. Simetria:  $D(a,b)= D(b,a)$
4. Desigualdade triangular:  $D(a,b)+D(b,c)\geq D(a,c)$

Sendo  $a, b$  e  $c$  vetores.

A Distância Euclidiana ou Métrica Euclidiana é uma métrica bastante usada em trabalhos acadêmicos. Ela deriva da métrica de Minkowski, também conhecida como norma  $L_p$ , a qual é dada por:

$$L_p(a, b) = \left( \sum_{i=1}^d |a_i - b_i|^p \right)^{1/p} \quad (11)$$

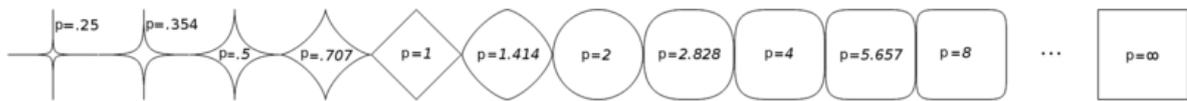
A Distância Euclidiana ( $L_2$ ), onde  $p=2$  e  $d$  é a dimensão, é dada por:

$$L_2(a, b) = \left( \sum_{i=1}^d |a_i - b_i|^2 \right)^{1/2} \quad (12)$$

Outra distância usada é a distância de Manhattan ou *City Block* ( $L_1$ ):

$$L_1(a, b) = \sum_{i=1}^d |a_i - b_i| \quad (13)$$

É possível variar o valor de  $p$ , gerando diversos tipos de distâncias ( $L_1, L_2$ , por exemplo) como foi observado acima. Esta variação do  $p$  pode ser vista graficamente na Figura 8. Vale ressaltar que quando  $p=2$  (Distância Euclidiana) é projetado um círculo.



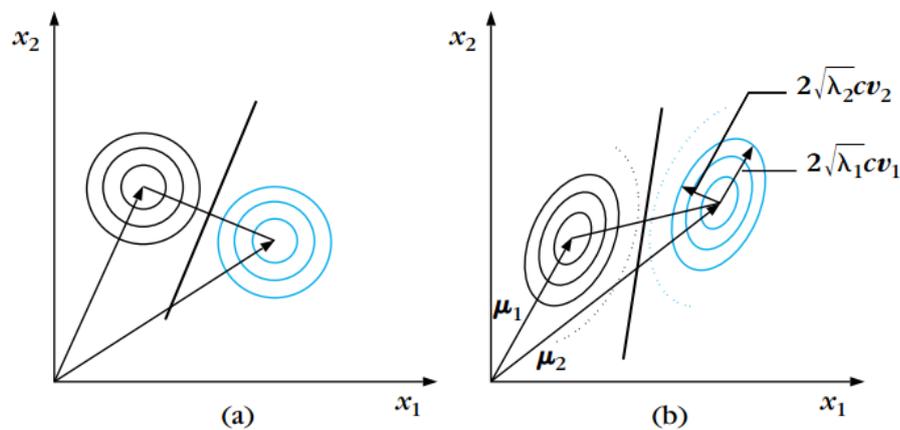
**Figura 8: Variação do  $p$  vista graficamente.**

A distância Euclidiana é extremamente sensível à variação na escala das variáveis envolvidas e não considera a correlação entre as variáveis. Para levar em consideração a covariância entre variáveis, a distância de Mahalanobis pode ser usada, eliminando assim, os problemas de escala e correlação. Ela é uma métrica bastante usada, onde dois vetores aleatórios são comparados, porém devem possuir a mesma distribuição, como também a mesma matriz de covariância  $S$ . A distância de Mahalanobis entre um vetor  $x$  de um grupo de

valores com média  $\mu$  pode ser definida como (Theodoridis e Koutroumbas, 2009; Duda et al, 2000):

$$D_m(x, \mu) = ((x - \mu)^T S^{-1}(x - \mu))^{1/2} \quad (14)$$

Na Figura 9b verifica-se que a Distância de Mahalanobis projeta elipses no espaço bidimensional, onde os eixos principais dessas elipses são dados pelos autovetores ( $v_1$  e  $v_2$ ) de  $S$ , e os autovalores ( $\lambda_1$  e  $\lambda_2$ ) determinam os tamanhos desses eixos (Theodoridis e Koutroumbas, 2009).



**Figura 9: Curvas dos pontos centrais de cada classe no espaço bidimensional. a) Distância Euclidiana; b) Distância de Mahalanobis. [Fonte: Theodoridis e Koutroumbas, 2009]**

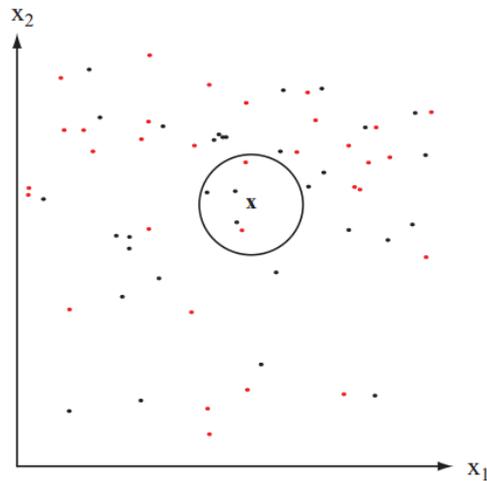
Quando a matriz de covariância  $S$  é identidade, a distância de Mahalanobis é igual à Distância Euclidiana, a qual projeta círculos, mostrados na Figura 9a. Se a matriz de covariância  $S$  é diagonal, a distância de Mahalanobis é igual à Distância Euclidiana Normalizada.

### 3.4.2 Algoritmo do Vizinho mais Próximo

O classificador k-Vizinhos Mais Próximos baseia-se em uma função cuja variável independente é a distância entre dois padrões. É um método que classifica um objeto desconhecido como pertencente à classe dos objetos mais próximos no espaço de características.

É um algoritmo que está entre os mais simples do aprendizado de máquina. Uma amostra de face é classificada pela maioria dos votos de suas vizinhas do conjunto de treinamento. Se  $k=1$ , o método é chamado de algoritmo do vizinho mais próximo, e a amostra

é simplesmente classificada como pertencente à classe de sua vizinha mais próxima (Theodoridis e Koutroumbas, 2009). A Figura 10 mostra um exemplo onde  $k=5$ . No exemplo mostrado, a amostra, ponto marcado com  $x$ , é classificada como pertencente à classe dos pontos pretos, pois dentro do círculo mostrado existem mais pontos pretos do que vermelhos.

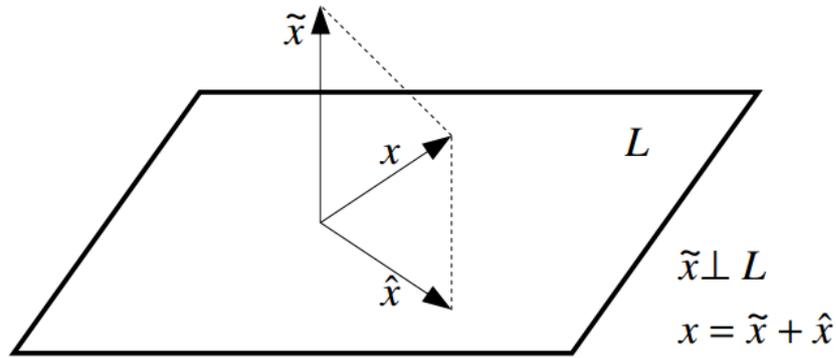


**Figura 10:** Nesse caso,  $k=5$ , o ponto  $x$  é classificado na classe dos pontos pretos. [Fonte: Duda et al, 2000]

### 3.4.3 Filtro de Novidades

O conceito de filtro de novidade foi descrito por Kohonen e Oja (1976) baseado no método de ortogonalização clássica de Gram-Schmidt (Kohonen, 1989; Costa e Moura, 1995).

Seja  $\{x_1, x_2, x_3, \dots, x_m\} \subset R^n$  um conjunto de vetores euclidianos de dimensão  $n$  que geram um subespaço de dimensão  $m$ ,  $L \subset R^n$ , com  $m < n$ . Considerando o subespaço  $L$ , um vetor arbitrário  $x \in R^n$  pode ser decomposto em duas componentes,  $\hat{x}$  e  $\tilde{x}$ , onde  $\hat{x} \in L$ , e  $\tilde{x} \perp L$  é a projeção de  $x$  perpendicular a  $L$ . O vetor  $\hat{x} \in L$  representa a componente de  $x$  que é “conhecida” pelo subespaço  $L$  e pode ser representada como uma combinação linear de  $\{x_1, x_2, x_3, \dots, x_m\}$ . O vetor  $\tilde{x} \perp L$  representa a informação “nova”, que é “desconhecida” pelo subespaço  $L$  e não pode ser representado como uma combinação linear de  $\{x_1, x_2, x_3, \dots, x_m\}$  (Costa e Moura, 1995). A Figura 11 ilustra estas duas componentes no espaço  $R^3$ .



**Figura 11: Conceito do filtro novidade, mostrando o vetor novidade  $\tilde{x}$  e o vetor projetado  $\hat{x}$ . [Fonte: Costa et al, 2012]**

Desta forma, considerando o subespaço  $L$ ,  $\tilde{x}$  é chamado de novidade e o sistema que extrai esta componente de  $x$  e a apresenta em sua saída pode ser chamado de Filtro de Novidade.

Para se determinar as componentes  $\hat{x}$  e  $\tilde{x}$  de um vetor  $x$ , é usado o processo de ortogonalização de Gram-Schmidt, descrito a seguir. A partir de um conjunto de  $m$  vetores independentes  $\{x_1, x_2, x_3, \dots, x_m\}$ , a base  $L$ , pode-se obter  $m$  vetores ortogonais  $\{v_1, v_2, v_3, \dots, v_m\}$  através do procedimento descrito em (15) e (16):

$$v_1 = x_1 \quad (15)$$

$$v_k = x_k - \sum_{i=1}^{k-1} \frac{(v_i, x_k)}{(v_i, v_i)} v_i, k = 2, 3, \dots, m \quad (16)$$

Onde  $\frac{(v_i, x_k)}{(v_i, v_i)} v_i$  é a projeção de  $x_k$  em  $v_i$ .

Dado um vetor de amostra  $x$ , sua componente “novidade” é calculada como o  $(m+1)$  passo do processo descrito em (16), ou  $\tilde{x} = v_{k+1}$ , conforme descrito na equação (17):

$$\tilde{x} = x - \sum_{i=1}^m \frac{(v_i, x)}{(v_i, v_i)} v_i \quad (17)$$

A norma do vetor  $\tilde{x}$ ,  $\|\tilde{x}\|$ , é usada como medida de desigualdade. Quanto menor a norma do vetor  $\tilde{x}$ , mais próximo ele estará da base  $\{v_1, v_2, \dots, v_m\} \in R^n$ .

Nesse trabalho, para classificar as faces utiliza-se o conceito de classificador de novidades, que é baseado no filtro de novidade.

### 3.4.4 Classificador de Novidades

O classificador de novidades utiliza vários filtros de novidades para realização de um classificador, conforme mostrado no diagrama em blocos da Figura 12. No primeiro nível do diagrama em blocos mostrado constroem-se tantos filtros de novidade (conforme descrito na seção anterior) quantas imagens de face existir no banco de dados. No segundo nível do diagrama em blocos mostrado calculam-se as normas dos vetores novidade. Quando uma amostra é aplicada na entrada do classificador, os vetores novidade são então calculados e extraídas as suas normas. No terceiro nível do diagrama em blocos mostrado faz-se uma comparação entre as normas dos vetores de novidade, para se determinar a menor delas. A amostra será classificada como pertencente a face para a qual o filtro de novidade apresentou o menor valor da norma. Assim, se dentre todas as normas a menor delas for  $|\tilde{x}_1|$ , então  $s_1 = 1, s_i = 0, i = 2, 3..N$ .

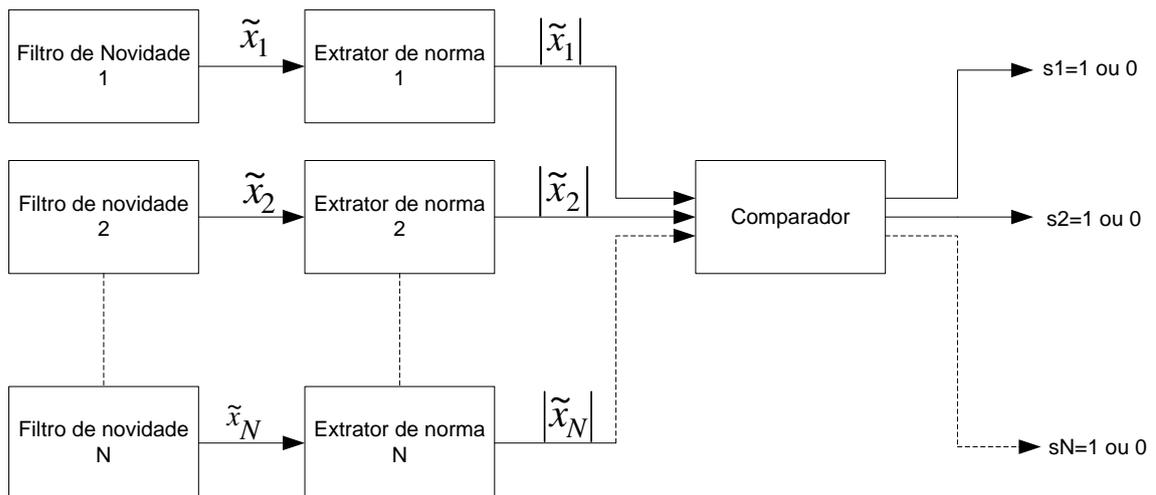


Figura 12: Classificador de Novidades.

## 3.5 Sistema Biométricos

Um sistema biométrico é essencialmente um sistema de reconhecimento de padrões que faz uma identificação pessoal estabelecendo a autenticidade de uma específica característica fisiológica ou comportamental de um usuário.

Identificação pessoal é o processo de associar um indivíduo a uma identidade. A identificação pode ser na forma de verificação (também chamada de autenticação), a qual implica autenticar uma identidade (“Eu sou quem eu afirmo ser?”), ou reconhecimento (também conhecido como identificação), o que implica determinar a identidade de uma

determinada pessoa a partir de um banco de dados de pessoas conhecidas ("Quem sou eu?") (JAIN et al, 2000).

Dessa forma, os sistemas biométricos podem ser empregados em dois tipos de aplicações: identificação e verificação (JAIN et al, 2000).

### 3.5.1 Identificação

No modo identificação, uma amostra biométrica do usuário é comparada com os modelos de todas as pessoas cadastradas na base de dados, possibilitando ao sistema emitir como saída:

- a identidade da pessoa, cujo modelo tem o mais alto grau de similaridade;
- ou uma decisão indicando que o usuário não está registrado na base de dados.

Formalmente, a identificação pode ser tratada da seguinte maneira: dada uma amostra de teste  $x$ , é necessário reconhecer a identidade  $I$  de um usuário, onde  $I \in \{I_1, I_2, \dots, I_N, I_{N+1}\}$ ,  $I_1, I_2, \dots, I_N$  correspondem às identidades dos  $N$  usuários cadastrados no sistema, e  $I_{N+1}$  indica o caso em que nenhuma identidade apropriada foi encontrada para a consulta. Se  $x_{I_n}$  é o modelo na base de dados correspondente à identidade  $I_n$ , e  $s_n$  é a pontuação de comparação (*match score*) entre  $x$  e  $x_{I_n}$ , onde  $n = 1, 2, \dots, N$ , a regra de decisão para identificação é:

$$x \in \begin{cases} I_{n_0}, & \text{se } n_0 = \arg \max_n s_n \text{ e } s_{n_0} \geq \eta \\ I_{N+1}, & \text{caso contrário} \end{cases} \quad (18)$$

onde  $\eta$  é um limiar pré-definido. A regra de decisão acima é usualmente conhecida como identificação em conjunto aberto, pois é possível retornar um resultado que indica que o usuário apresentou uma característica biométrica que não se encontra entre os  $N$  usuários cadastrados do sistema. Na prática, quase todos os sistemas de identificação biométrica usam identificação em conjunto aberto. É também possível forçar o sistema a retornar sempre uma identidade, independentemente do valor de  $s_{n_0}$ . Tal cenário é chamado de identificação em conjunto fechado, ou seja, qualquer indivíduo a ser testado deve ter um modelo previamente armazenado na base de dados.

### 3.5.2 Verificação

No cenário de verificação, uma pessoa afirma uma identidade particular e o sistema biométrico aceita ou rejeita a afirmação. A verificação é feita comparando a amostra adquirida no momento da afirmação com uma amostra previamente cadastrada (comparação um para um). Se as duas amostras, segundo algum critério, combinam suficientemente, a identidade afirmada é aceita como verdadeira e o usuário torna-se genuíno, caso contrário, a afirmação é rejeitada e o usuário é considerado um impostor. Quando a identidade é considerada como genuína, o usuário é autorizado a acessar os serviços providos pelo sistema (JAIN et al., 2000, 2011).

Formalmente, a verificação pode ser vista conforme o seguinte problema de classificação de duas categorias: dada uma identidade  $I$  e uma amostra de teste  $x$ , é necessário decidir se  $(I, x)$  pertence à classe genuíno ou impostor. Seja  $x_I$  o modelo cadastrado na base de dados correspondente à identidade  $I$ , onde  $x$  é comparado com  $x_I$ , gerando um uma pontuação de comparação (*match score*) entre eles. A regra de decisão é dada por:

$$(I, x) \in \begin{cases} \text{genuíno}, & \text{se } s \geq \eta \\ \text{impostor}, & \text{se } s < \eta \end{cases} \quad (19)$$

onde  $\eta$  é um limiar pré-definido. Se a distância for usada como pontuação de comparação no lugar da similaridade, os símbolos matemáticos mostrados na equação ( 19 ) devem ser invertidos (JAIN et al., 2011).

Neste trabalho, foi usada o valor da distância ao invés do grau de similaridade.

### 3.6 Medidas de desempenho

#### 3.6.1 Identificação

Uma maneira de avaliar o desempenho no modo identificação em conjunto fechado é construir a curva *Cumulative Match Characteristic* (CMC) e extrair a Taxa de Reconhecimento Rank 1.

Para traçar a curva CMC, é necessário dois conjuntos:

- Galeria (G): consiste em um conjunto formado pelos modelos biométricos de  $M$  indivíduo cadastrados no sistema. Um elemento deste conjunto é denotado por  $g_i \in G$ , onde  $i = 1, \dots, M$ .
- Conjunto de prova (Q): é formado pelas  $N$  amostras biométricas dos indivíduos a serem testados. Este conjunto não precisa ter amostras de todos os

indivíduos, no entanto, não pode possuir amostras de indivíduos cujos modelos não estão na galeria  $G$ . Um elemento deste conjunto é denotado  $b_j \in Q$ , onde  $j = 1, \dots, N$ .

Comparando uma  $b_j$  (do conjunto de prova) com um modelo biométrico  $g_i$  (da galeria), é gerada uma pontuação de similaridade  $s = s(b_j, g_i)$ . Uma pontuação de similaridade alta indica que dois elementos biométricos são similares. No caso, de se usar o valor da distância de dois elementos, a relação será inversamente proporcional, ou seja, quanto menor a distância, maior a similaridade entre os dois elementos biométricos. A comparação de todos os elementos de  $G$  e  $Q$  produz uma matriz de pontuação  $N \times M$ . Cada linha dessa matriz é representada por:

$$S_j = \{s(b_j, g_1), s(b_j, g_2), \dots, s(b_j, g_M)\}, j = 1, \dots, N. \quad (20)$$

Cada conjunto  $S_j$  é ordenado pela melhor similaridade:

$$s(b_j, g_{(1)}) \geq s(b_j, g_{(2)}) \geq \dots \geq s(b_j, g_{(M)}) \quad (21)$$

Nesse conjunto ordenado,  $s(b_j, g_{(1)})$  representa o modelo contido em  $G$  que resultou na melhor similaridade,  $s(b_j, g_{(2)})$  o modelo contido em  $G$  que resultou na segunda melhor similaridade e assim por diante. Assumindo um rank  $k$  (sendo  $1 \leq k \leq M$ ), uma identificação é bem sucedida quando o modelo referente ao mesmo indivíduo que possui a amostra  $b_j$  aparece entre os  $k$  primeiros elementos do conjunto ordenado (PINHEIRO, 2010).

A Taxa de Identificação para um rank  $k$  de um teste sobre o conjunto de prova  $Q$  de tamanho  $n$  é dado por:

$$p_k(Q) = \frac{c_k}{n}, \text{ onde } \sum_{b_j \in Q} b_k(b_j) \quad (22)$$

O termo  $b_k(b_j)$  é igual à 1 se houver um modelo do mesmo indivíduo de  $b_j$  entre os primeiros  $k$  elementos do conjunto ordenado, e zero caso contrário.

Assim, a curva CMC é construída traçando-se a taxa de identificação  $p$  em função do rank  $k$ .

### 3.6.2 Verificação

Na verificação, um usuário deve, primeiramente, afirmar que é um usuário cadastrado no sistema, então, o sistema biométrico aceitará ou rejeitará tal afirmação.

Dentro desse contexto, existem quatro tipos de ocorrências possíveis:

- Aceitação correta (*Correct Acceptance, CA*): ocorre quando o sistema aceita uma afirmação de identidade e a afirmação é verdadeira;
- Falsa aceitação (*False Acceptance, FA*): ocorre quando o sistema aceita uma afirmação de identidade, mas a afirmação é falsa;
- Rejeição correta (*Correct Rejection, CR*): ocorre quando o sistema rejeita uma afirmação de identidade e a afirmação é falsa;
- Falsa rejeição (*False Rejection, FR*): ocorre quando o sistema rejeita uma afirmação de identidade, mas a afirmação é verdadeira.

Considerando o valor de similaridade como uma variável aleatória, a qual possui distribuição de probabilidade para o caso em que clientes solicitam suas identidades, e uma distribuição de probabilidade para as solicitações de impostores, existe um limiar que define uma regra de classificação para o sistema, assim como é mostrado na Figura 13.

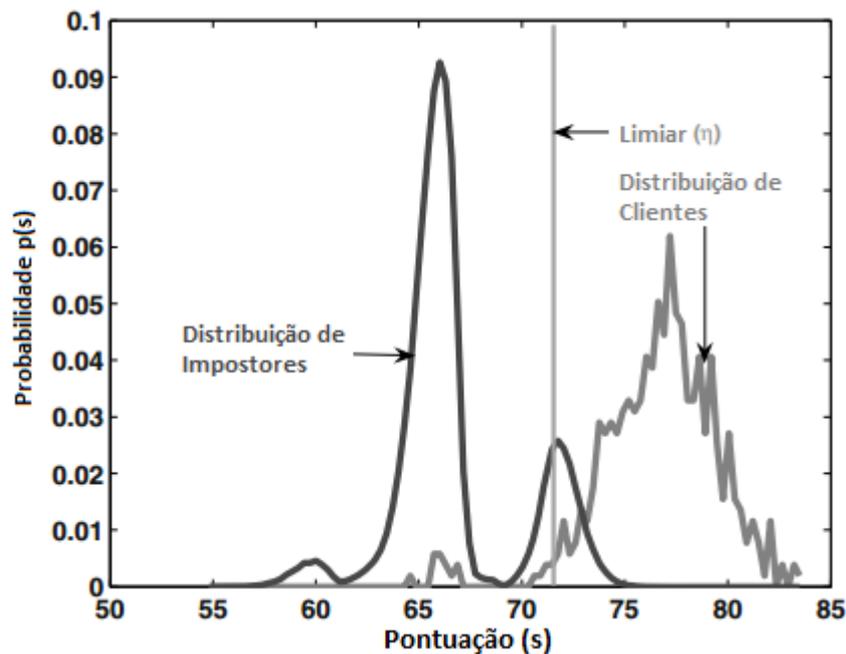


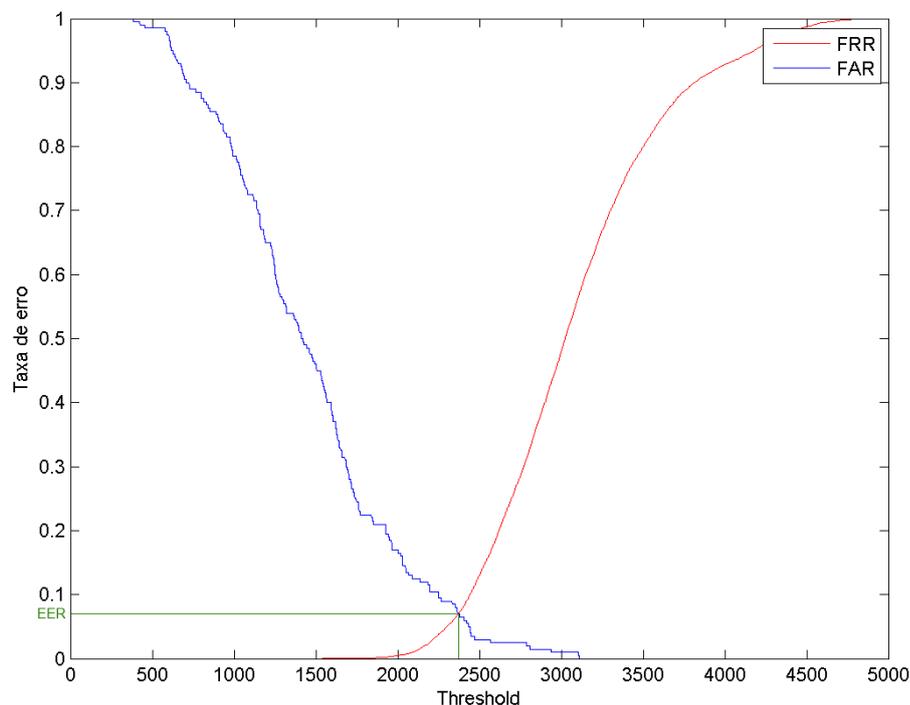
Figura 13: Distribuição de pontuação de clientes e impostores. [Fonte: Jain et al, 2011]

A partir deste dados, pode-se definir taxas métricas relacionadas ao modo de verificação:

- Taxa de Correta Aceitação ou *Correct Acceptance Rate (CAR)*: probabilidade de classificar um cliente como cliente;

- Taxa de Correta Rejeição ou *Correct Rejection Rate* (CRR): probabilidade de classificar um impostor como impostor;
- Taxa de Falsa Aceitação ou *False Acceptance Rate* (FAR): probabilidade de classificar um impostor como cliente;
- Taxa de Falsa Rejeição ou *False Rejection Rate* (FRR): probabilidade de classificar um cliente como impostor;
- Taxa de Erro Equivalente ou *Equal Error Rate* (EER): quando as distribuições de probabilidade para clientes e impostores se sobrepõem, a FAR e a FRR se cruzam num determinado ponto. Esse ponto é o EER, onde FRR é igual à FAR. Quanto menor este valor, melhor é o desempenho do sistema biométrico.

Fica claro que a FRR e a FAR são funções que dependem de um limiar  $\eta$ . Se o limiar  $\eta$  aumentar, FAR irá diminuir, porém FRR irá aumentar e vice-versa. Logo, em um sistema biométrico, não é possível diminuir essas duas taxas simultaneamente variando o limiar. Tais afirmações podem ser vistas na Figura 14.



**Figura 14: Impacto da variação do limiar (*threshold*) nos valores de FAR e FRR.**

A análise da curva *Receiver Operating Characteristic* (ROC) foi proposta por Metz (1978) para avaliar o desempenho de classificadores. Muitos trabalhos têm demonstrado a importância das curvas ROC na avaliação e comparação de classificadores, havendo um

aumento significativo no seu uso devido ao fato de que uma simples acurácia de classificação é geralmente uma medição de desempenho insuficiente (Fawcett, 2006).

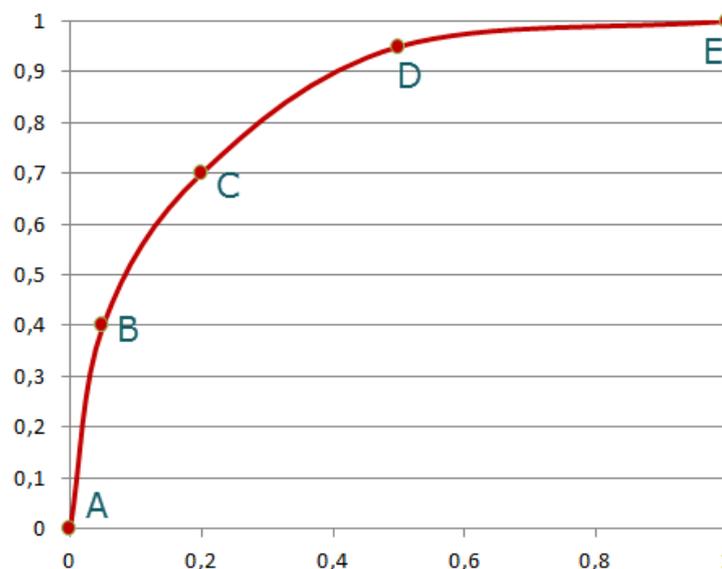
A Sensibilidade ( $S$ ), também conhecida como taxa de verdadeiro positivo ou revocação (*recall*), é definida como:

$$S = \frac{TP}{TP + FN} = \frac{TP}{P} \quad (23)$$

A especificidade ( $E$ ), também conhecida como taxa de verdadeiro negativo, é definida como:

$$E = \frac{TN}{TN + FP} = \frac{TN}{N} \quad (24)$$

A curva ROC trata-se de um gráfico bidimensional onde no eixo vertical temos a Taxa de Correta Aceitação (CAR) ou sensibilidade e no eixo horizontal temos a Taxa de Falsa Aceitação (FAR) ou 1 menos a especificidade. Para o traçado da curva de ROC varia-se um limiar de decisão e calcula-se diversos pares de sensibilidade e especificidade. A Figura 15 mostra um exemplo de Curva ROC obtida utilizando-se cinco limiares de decisão, correspondendo a cinco pares de sensibilidade e especificidade: A, B, C, D e E..



**Figura 15: Exemplo de Curva ROC com 5 limiares (A, B, C, D e E)**

Um classificador ideal apresentaria, para um dado valor de limiar de decisão, sensibilidade e especificidade iguais a um, isto é, encontra-se uma instância positiva sempre que a mesma for positiva e nunca afirma falsamente que uma instância positiva existe.

Para se comparar o desempenho de diferentes classificadores é utilizada uma medida de desempenho derivada da curva ROC: a Área Sobre a Curva ROC (*Area Under a Curve: AUC*).

Quando a curva é traçada através de um algoritmo de ajuste de curva, a área é calculada através de uma integração dessa curva:

$$AUC = \int_0^1 \frac{TP}{P} d\frac{FP}{N} \quad (25)$$

Quando se constrói a curva interligando uma grande número de pontos, a área pode ser calculada pela regra do trapézio:

$$AUC = \sum_{i=1}^{i=k-1} (FP_{i+1} - FP_i)(TP_i - TP_{i+1})/2 \quad (26)$$

O valor da AUC varia de 0,5 à 1, com 1 indicando que não houve erros.

Acurácia e a precisão são definidas a seguir.

$$Acurácia = \frac{TP + TN}{P + N} \quad (27)$$

$$Precisão = \frac{TP}{TP + FP} \quad (28)$$

## 4 MATERIAIS E MÉTODOS

### 4.1 Introdução

Nesse trabalho propõe-se o uso do classificador de novidades como ferramenta de reconhecimento de faces e comparar o seu desempenho com o desempenho do classificador vizinho mais próximo usando a métrica da distância euclidiana. Com esse intuito, nesse capítulo de materiais e métodos pretende-se: apresentar o banco de imagens que será utilizado para estabelecer essa comparação e apresentar a metodologia utilizada para realizar essa comparação. Como parte dessa metodologia serão abordados os métodos de partição de dados.

O método de reconhecimento facial pode ser dividido nas seguintes etapas: aquisição de imagem, pré-processamento, extração de características e classificação.

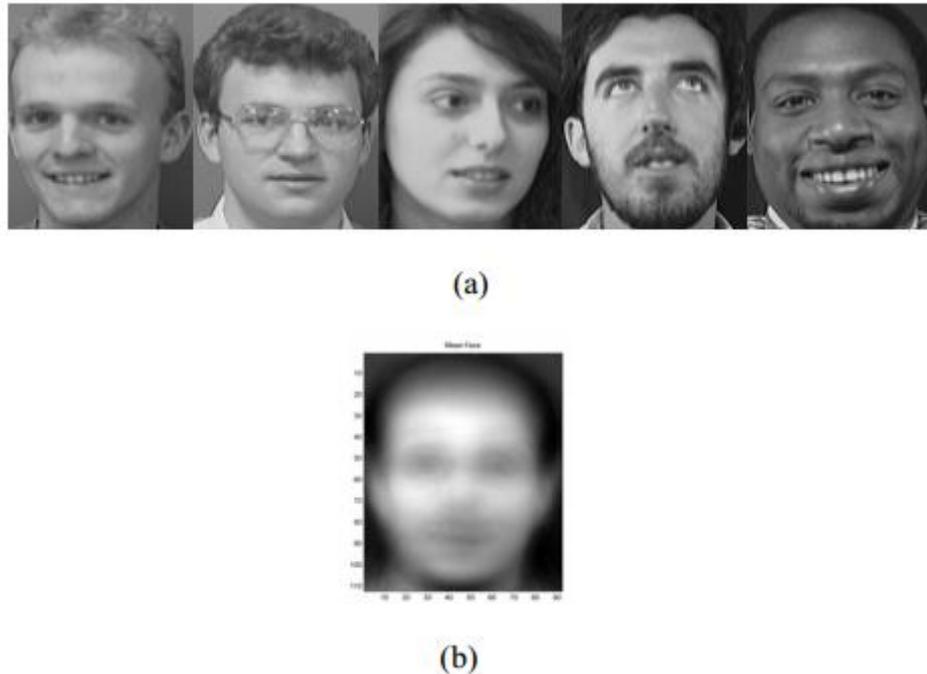
Nesse capítulo abordam-se as etapas Extração de Características e Classificação, visto que as imagens das faces já foram capturadas por outros grupos de pesquisas, e estão contidas no bancos de imagens de faces que será descrita a seguir.

### 4.2 Banco de imagens

Este trabalho usa a base de faces ORL pelo fato de ser uma das mais usadas pela comunidade científica, por possuir variação na escala e rotação, e por ser possível comparar o desempenho com outros classificadores que a utilizaram.

#### 4.2.1 Banco de Imagens ORL

A base de dados ORL (Laboratório de Pesquisa da Olivetti) foi coletada entre 1992 e 1994. Ela contém 10 imagens diferentes de 40 indivíduos (4 mulheres e 36 homens), totalizando 400 imagens de tamanho 92x112 pixels em escala de cinza. A idade dos indivíduos varia entre 18 à 81 anos, com a maioria sendo entre 20 à 35 anos. Para alguns indivíduos, as imagens foram capturadas em diferentes épocas, variando as condições de iluminação, expressões faciais (olhos abertos e fechados, sorrindo ou não) e contando ou não detalhes na face (óculos, por exemplo). Todas as imagens foram capturadas com um fundo negro homogêneo (Samaria e Harter, 1994; Gross, 2005). Esta base pode ser encontrada em: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>. A Figura 16 mostra algumas amostras e uma face média da base de faces ORL.



**Figura 16:** a) Exemplos de faces da base ORL; b) face média da base de faces ORL.

A Tabela 2 mostra um resumo da base de faces ORL.

**Tabela 2: Banco de faces ORL**

Número de indivíduos	Imagens por indivíduo	Resolução da imagem	Total de imagens
40	10	92x112	400

### 4.3 Pré-processamento das imagens

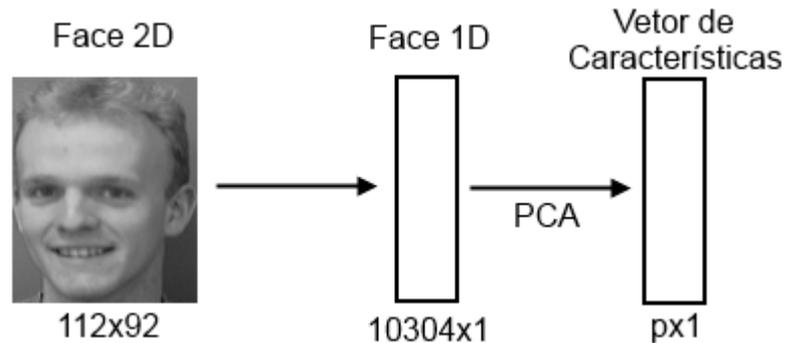
Foram usadas as imagens originais do banco de faces, ou seja, não houve pré-processamento fotométrico (correção de iluminação), nem geométrico (ajuste de escala, rotação, entre outros) nas imagens.

### 4.4 Extração de características

Cada imagem foi representada por um vetor de características, obtido através do método PCA, ou uma matriz de característica, obtida através do método 2DPCA ou  $(2D)^2PCA$ .

A Figura 17 mostra o processo de extração de características usando uma face da base de imagens ORL de 92x112 pixels. Primeiramente, a face é transformada em um vetor (10304x1), e então é passado pelo método PCA, o qual fará a redução de dimensionalidade, reduzindo a dimensão das características de 10304 para  $p$  elementos, onde  $p$  é o número de

autovetores escolhidos. Quanto menor o valor de  $p$ , maior é a redução de dimensionalidade. Geralmente  $p$  é igual  $N-1$  elementos, sendo  $N$  igual ao número de amostras de treinamento.

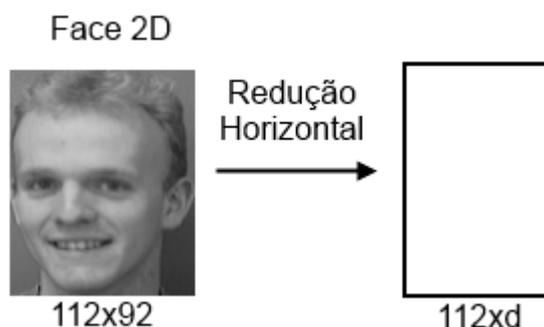


**Figura 17: Vetor de características usando PCA**

Foram usados os seguintes valores de  $p$  na estratégia metade-metade: 1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 112, 121, 144, 169, 199. Já na estratégia 10-fold cross validation foram usados os seguintes valores de  $p$ : 1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 112, 121, 144, 169, 199, 250, 300, 359.

A Figura 18 mostra o processo de extração de características usando o método 2DPCA e uma face da base de imagens ORL de 92x112 pixels. A matriz da face (112 linhas por 92 colunas) é usada diretamente pelo método 2DPCA, o qual fará a redução de dimensionalidade na direção horizontal, extraindo as características da ordem de 112xd elementos, onde  $d$  é o número de autovetores escolhidos. Quanto menor o valor de  $d$ , maior é a redução de dimensionalidade na direção horizontal.

Foram usados os seguintes valores de  $d$ : 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, gerando as seguintes matrizes: 112x1, 112x 2, 112x 3, 112x 4, 112x 5, 112x 6, 112x 7, 112x 8, 112x 9, 112x 10, 112x 15, 112x 20.

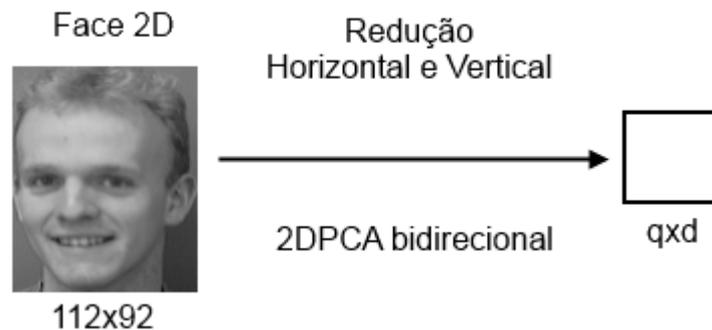


**Figura 18: Matriz de características usando 2DPCA**

A Figura 19 mostra o processo de extração de características usando uma face da base de imagens ORL de 92x112 pixels, pelos métodos 2DPCA bidirecional. A matriz da face

(112x92) é usada diretamente pelo método 2DPCA bidirecional, o qual fará a redução de dimensionalidade na direção horizontal e vertical ao mesmo tempo, extraíndo as características da ordem de 112x92 para  $q \times d$  elementos, considerando que  $d < n$  e  $q < m$ , onde  $d$  e  $q$  são o número de autovetores escolhidos. Quanto menor o valor de  $d$  e  $q$ , maior é a redução de dimensionalidade em ambas as direções.

Foram usados os seguintes valores de  $q$  e  $d$ : 2x2, 5x5, 10x10, 15x15, 20x20, 25x25, 30x30.



**Figura 19: Matriz de características usando 2DPCA bidirecional**

Todos estes métodos de extração de características já foram discutidos em detalhes no capítulo 3.3.

#### 4.5 Técnicas de partição dos dados

Existem várias técnicas para particionamento de conjuntos de treinamento e teste, dentre elas, destacam-se: *Holdout*, *Leave-one-out* e *Validação Cruzada (Cross Validation)*.

A técnica *Holdout* divide o conjunto de dados em dois subconjuntos mutuamente exclusivos: um para treinamento e outro para teste. Geralmente, o conjunto de dados é dividido ao meio. Assim, esse método tem a principal desvantagem de reduzir ambos os conjuntos de treinamento e teste (THEODORIDIS e KOUTROUMBAS, 2009; JAIN, DUIN e MAO, 2000).

O *Leave-One-Out*, como o próprio nome diz, deixa um fora, treina-se um classificador com  $N-1$  amostras e testa com a que restou. Essa técnica tem a principal desvantagem de um alto custo computacional, pois o experimento é executado  $N$  vezes (JAIN, DUIN e MAO, 2000).

O *n-Fold Cross Validation* é um meio termo entre *Holdout* e *Leave-One-Out*, pois divide o conjunto de dados em  $P$  subconjuntos disjuntos. Para treinar o classificador usamos

P-1 subconjuntos e o restante para teste. Isso é feito P vezes, variando o conjunto de teste utilizado e o de treinamento correspondente. Sua estimativa é menos tendenciosa que o *Holdout* e o custo computacional é menor do que o *Leave-One-Out* (JAIN, DUIN e MAO, 2000).

As possíveis e diferentes combinações de partições de treinamento e teste são executadas com o intuito de reduzir a variabilidade, a partir dos quais são calculados o desempenho médio e máximo.

Os experimentos utilizaram a base de faces ORL, a qual possui 10 imagens por indivíduo, sendo 40 indivíduos, totalizando 400 imagens. Foram utilizadas duas estratégias:

- *Metade-metade* ou *Holdout*: foram selecionadas cinco imagens por indivíduo para formar a base de treinamento e as cinco imagens restante são usadas como teste, ou seja, 50% treinamento e 50% teste. Não há sobreposição de imagens do conjunto de treinamento e teste. Buscando aumentar a confiança dos testes, o experimento é repetido 10 vezes, com diferentes amostras de treinamento e teste.
- *Validação Cruzada em 10 subconjuntos* ou *10-fold cross validation*: A base de faces ORL foi dividida em 10 subconjuntos, usando uma imagem de um indivíduo em cada subconjunto. Então, cada subconjunto possui 40 imagens, cada uma correspondendo a um indivíduo. Em cada experimento, 9 subconjuntos foram usados como conjunto de treinamento e o subconjunto restante para teste. Logo, os conjuntos de treinamento e teste possuem 360 e 40 imagens, respectivamente. O experimento é repetido 10 vezes, com diferentes subconjuntos de treinamento e teste.

#### **4.6 Testes Experimentais**

Para a realização dos experimentos de reconhecimento facial, foi utilizado um microcomputador com processador Intel Core i5-2540M, 2,6 Ghz e 8 GB de RAM, operando sob o sistema operacional Microsoft Windows 7 SP1 64 bits.

Todos os algoritmos dos experimentos foram escritos pelo autor deste trabalho e executados no programa MATLAB R2012a.

Foram implementadas as técnicas de extração de características PCA, 2DPCA e  $(2D)^2$ PCA, combinando-os com os classificadores vizinho mais próximo e classificador de novidades. A partir dessas combinações, pode-se obter uma comparação do desempenhos de ambos os classificadores.

A avaliação de desempenho utiliza duas estratégias de partição de dados: metade-metade e *10-fold cross validation* discutidas na seção 4.5. No caso do do classificador de novidades, o conjunto de treinamento forma uma base para o filtro de novidades, ou seja, a comparação será feita entre a amostra de teste e a base do filtro de novidades. Já no caso 1NN com a distância euclidiana, a amostra de teste será comparada com cada amostra de treinamento.

Os experimentos foram executados nos modos identificação (em conjunto fechado) e verificação, usando duas estratégias: *Holdout* (metade-metade) e *10-fold cross validation*, comparando com o 1NN, o qual utiliza a métrica da distância euclidiana.

A Tabela 3 resume os experimentos executados neste trabalho.

**Tabela 3: Síntese do experimentos**

Classificadores	Estratégias	Extração de características	Dimensões usadas
Classificador de Novidades e 1NN	Metade-metade	PCA	1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 112, 121, 144, 169,199
		2DPCA	112x1, 112x 2, 112x 3, 112x 4, 112x 5, 112x 6, 112x 7, 112x 8, 112x 9, 112x 10, 112x 15, 112x 20
		$(2D)^2$ PCA	2x2, 5x5, 10x10, 15x15, 20x20, 25x25, 30x30
	10-fold Cross Validation	PCA	1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 112, 121, 144, 169,199, 250, 300, 359
		2DPCA	112x1, 112x 2, 112x 3, 112x 4, 112x 5, 112x 6, 112x 7, 112x 8, 112x 9, 112x 10, 112x 15, 112x 20
		$(2D)^2$ PCA	2x2, 5x5, 10x10, 15x15, 20x20, 25x25, 30x30

A partir desta metodologia, no modo identificação, analisa-se o impacto do número de componentes principais (autovetores) no desempenho dos classificadores através da taxa de reconhecimento *rank 1*. Além disso, foram traçadas as curvas CMC comparando os classificadores e as estratégias de partição de dados.

Já no modo verificação foram executadas as seguintes análises:

- impacto do número de componentes principais (dimensão do vetor de características) no desempenho do classificador;
- sensibilidade ou taxa de correta aceitação (CAR) quando FAR=1%;
- taxa de erro equivalente (EER);
- curva ROC;
- e área sobre a curva ROC (AUC).

## 5 RESULTADOS E DISCUSSÕES

### 5.1 Introdução

Neste capítulo apresentam-se os resultados dos testes e as avaliações de desempenho do Classificador de Novidades (CN), usando métodos de extração de características PCA, 2DPCA e  $(2D)^2$ PCA na base de faces ORL.

A partir de agora será utilizada a nomenclatura CN para Classificador de Novidades, DE para Distância Euclidiana, 5-5 para a estratégia metade-metade e 9-1 para a estratégia *10-fold cross validation*.

### 5.2 Experimentos no modo identificação

#### 5.2.1 Impacto do número de autovetores versus taxa de reconhecimento *rank 1*

A Figura 20, Figura 21 e Figura 22 mostram curvas da taxa de reconhecimento *rank 1* versus o número de autovetores, para ambos os classificadores, utilizando os métodos PCA, 2DPCA e  $(2D)^2$ PCA, respectivamente.

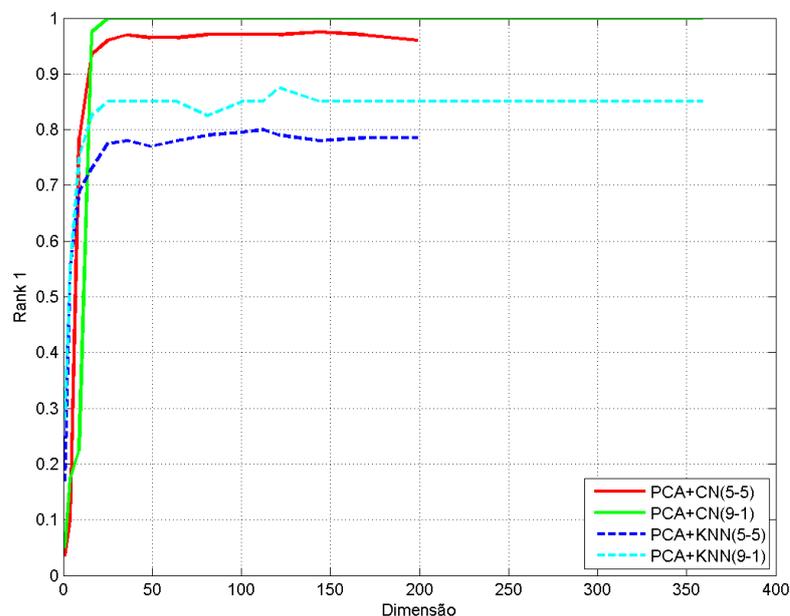


Figura 20. Taxa de reconhecimento rank 1 versus número de autovetores usando PCA

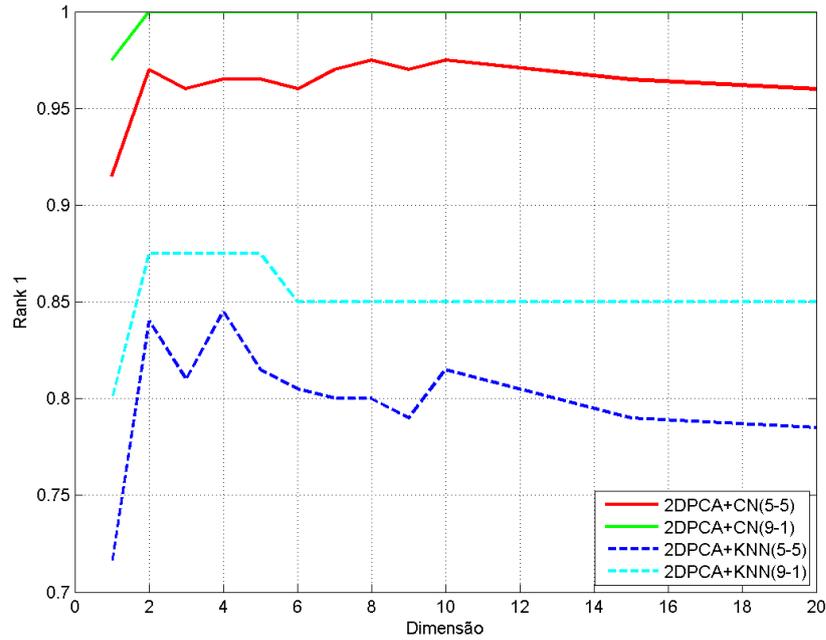


Figura 21. Taxa de reconhecimento rank 1 versus número de autovetores usando 2DPCA

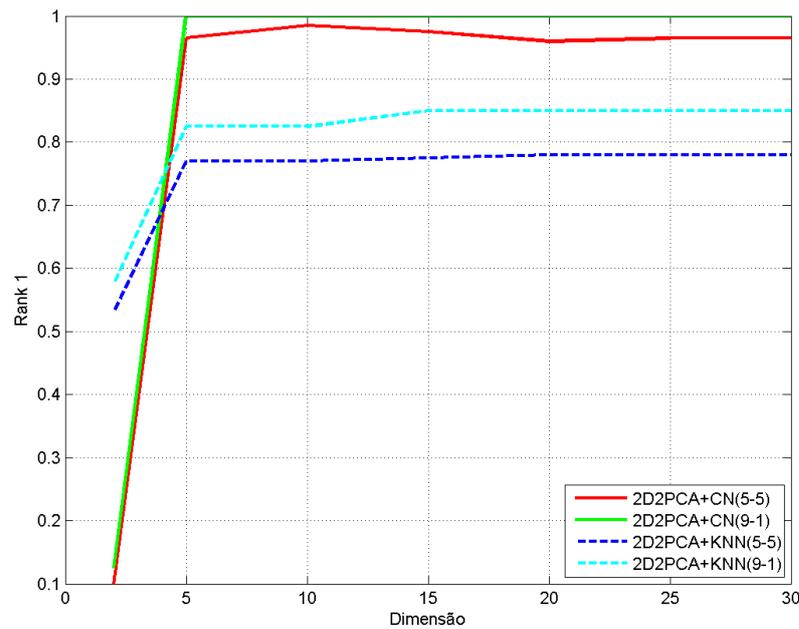


Figura 22. Taxa de reconhecimento rank 1 versus número de autovetores usando  $(2D)^2$ PCA

### 5.2.2 Curvas CMC

A Figura 23, Figura 24 e a Figura 25 mostram as curvas CMC geradas usando os métodos PCA, 2DPCA e  $(2D)^2$ PCA nas estratégias metade-metade (5-5) e *10-fold cross validation* (9-1).

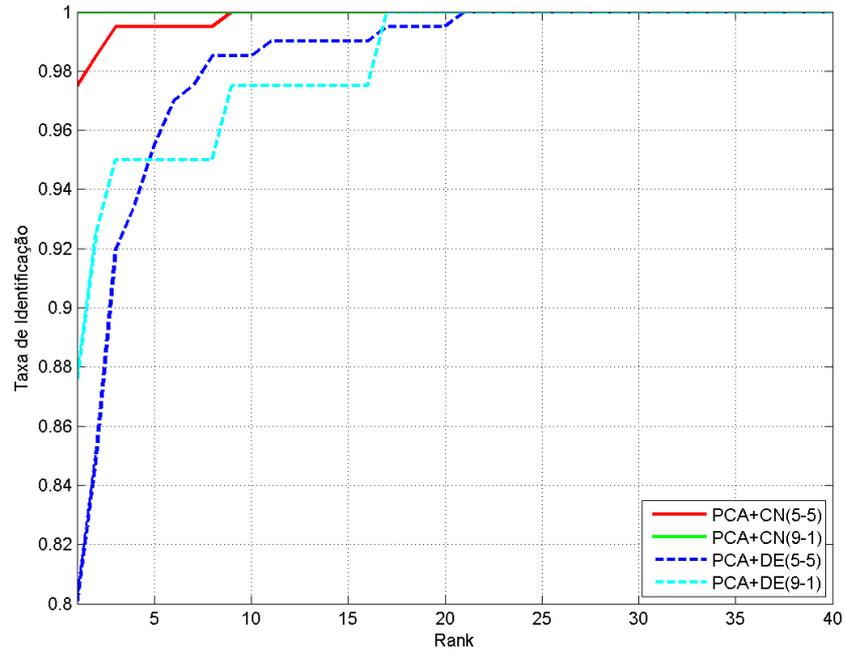


Figura 23. Curva CMC comparando os classificadores usando PCA

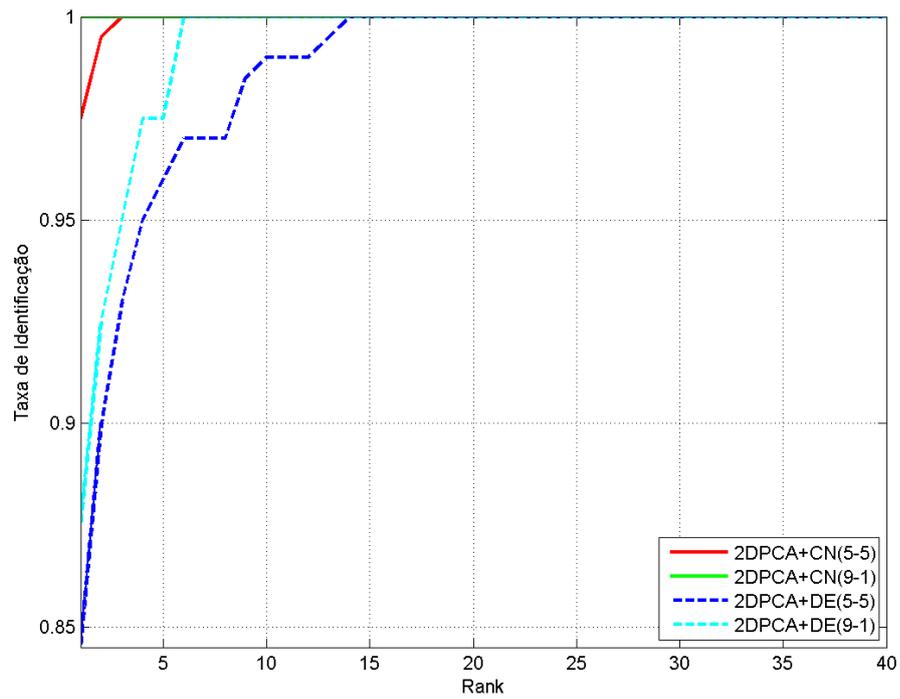


Figura 24. Curva CMC comparando os classificadores usando 2DPCA

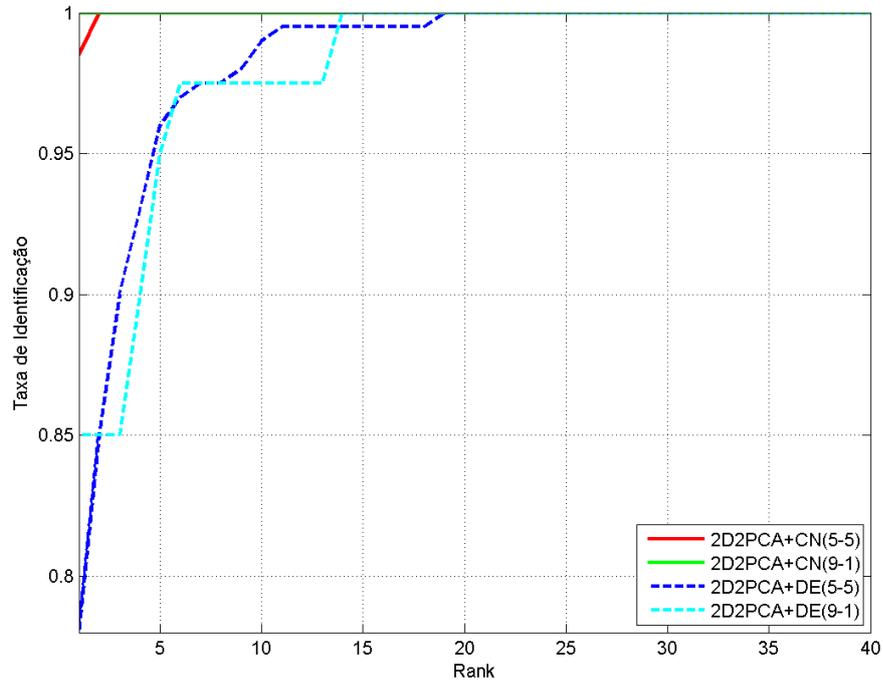


Figura 25. Curva CMC comparando os classificadores usando  $(2D)^2PCA$

### 5.2.3 Taxa de Reconhecimento *rank 1*

A Tabela 4 mostra o resumo geral do desempenho dos classificadores no modo identificação. Percebe-se que os classificadores apresentaram taxas de reconhecimento *rank 1* mais altas com a estratégia *10-fold Cross Validation*.

Tabela 4: Melhores taxas de reconhecimento *rank 1*

Classificador	Estratégia	Extração de características	Dimensão	Máxima
Classificador de Novidades	Metade-metade	PCA	144	97,5%
		2DPCA	112x8	97,5%
		$(2D)^2PCA$	10x10	98,5%
	10-fold cross validation	PCA	25	100%
		2DPCA	112x2	100%
		$(2D)^2PCA$	5x5	100%
Vizinho mais próximo	Metade-metade	PCA	112	80%
		2DPCA	112x4	84,5%
		$(2D)^2PCA$	20x20	78%
	10-fold cross validation	PCA	121	87,5%
		2DPCA	112x2	87,5%
		$(2D)^2PCA$	15x15	85%

### 5.3 Experimentos no modo verificação

#### 5.3.1 Impacto do número de autovetores versus CAR

A Figura 26, Figura 27 e a Figura 28 mostram valores de CAR (FAR=1%) versus o número de autovetores (dimensão do vetor de características), para ambos os classificadores, utilizando os métodos PCA, 2DPCA e  $(2D)^2$ PCA. Enquanto que a Figura 29, Figura 30 e a Figura 31 mostram valores de CAR (FAR=0,1%) versus o número de autovetores. Quanto menor a FAR, menor a probabilidade de permitir que um impostor seja aceito pelo sistema biométrico.

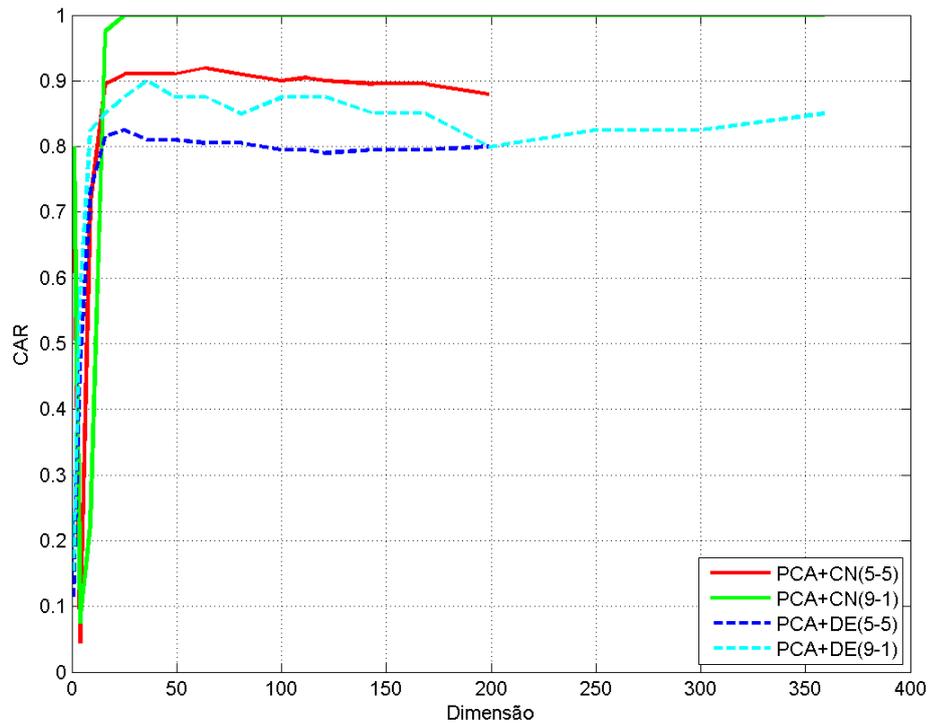


Figura 26. CAR (FAR=1%) versus número de autovetores usando PCA

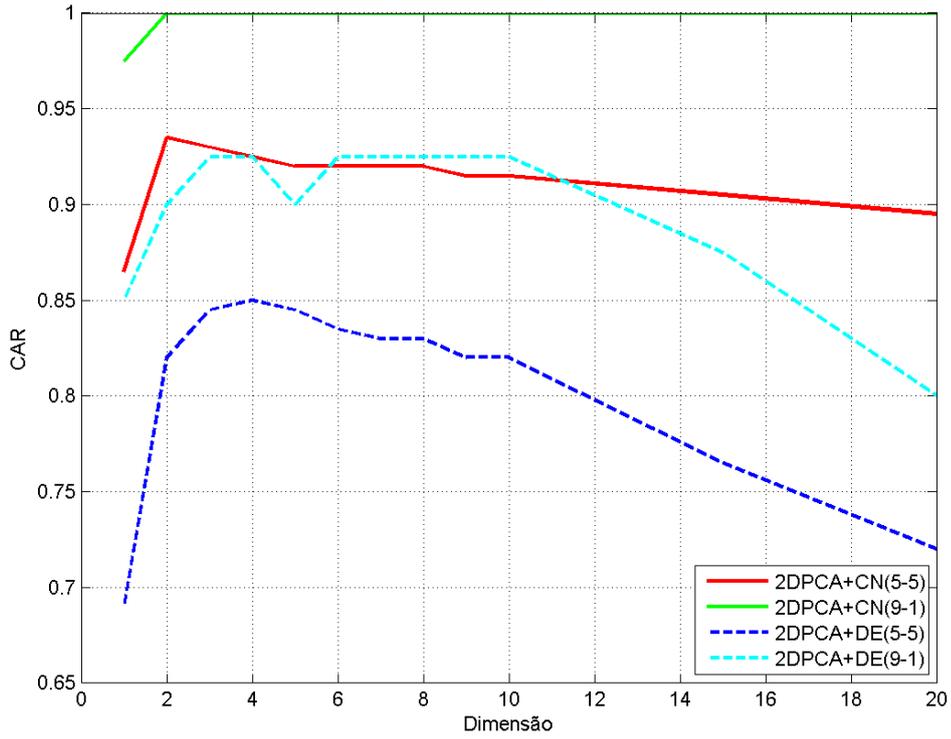


Figura 27. CAR (FAR=1%) versus número de autovetores usando 2DPCA

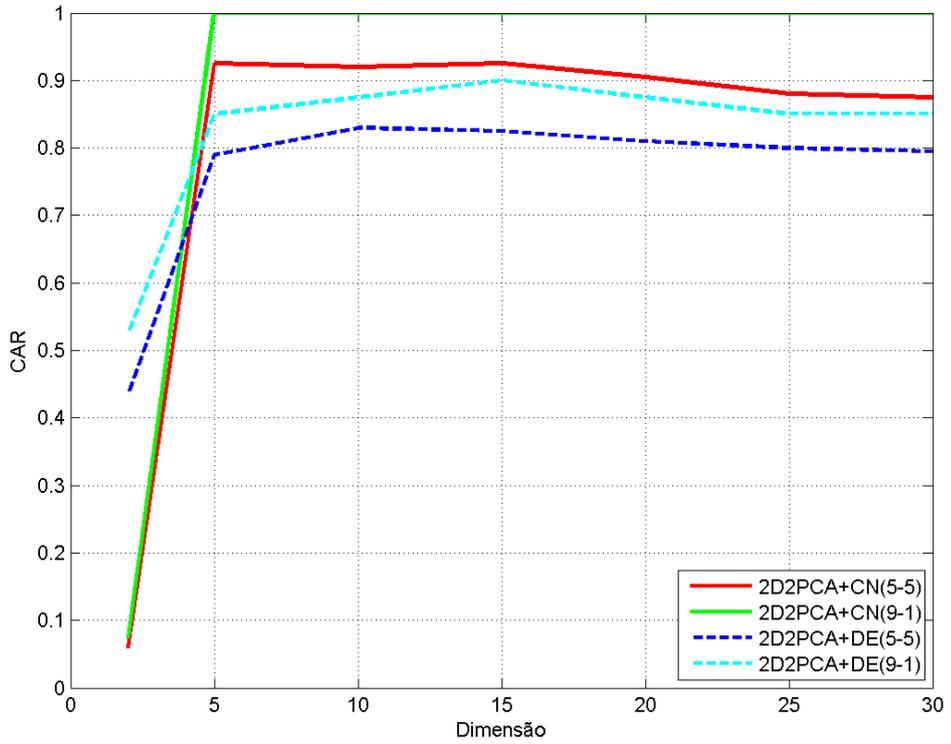


Figura 28. CAR (FAR=1%) versus número de autovetores usando (2D)<sup>2</sup>PCA

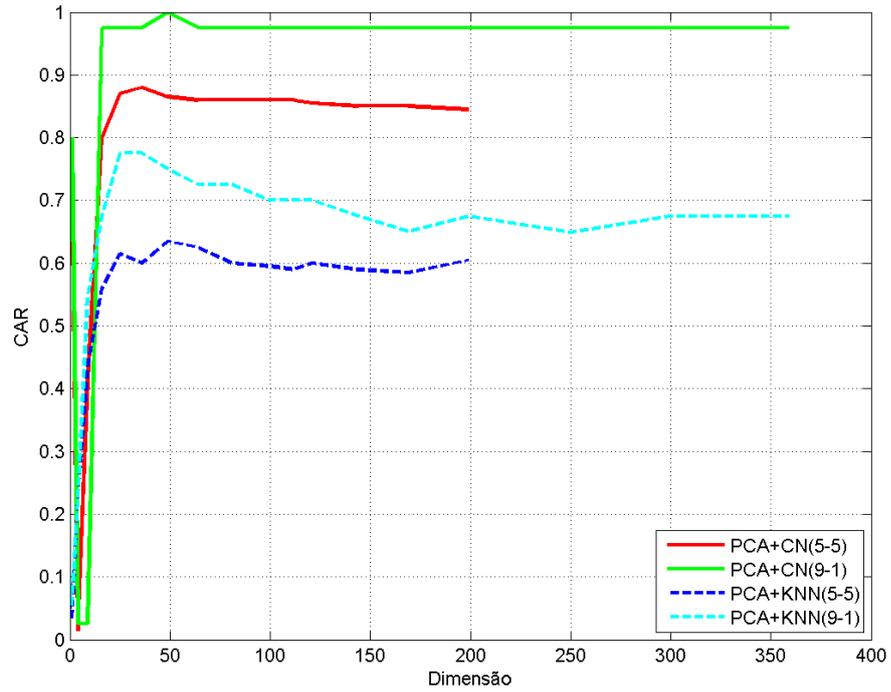


Figura 29. CAR (FAR=0,1%) versus número de autovetores usando PCA

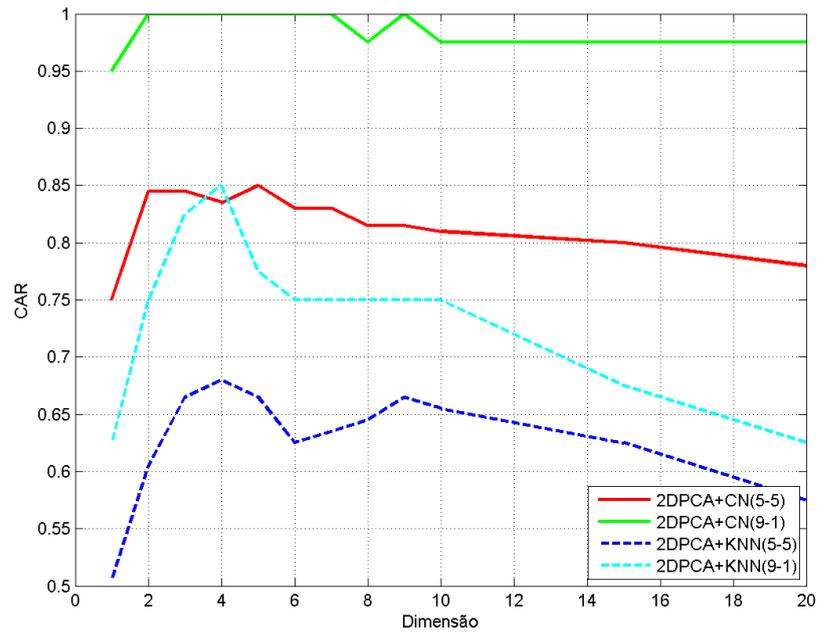


Figura 30. CAR (FAR=0,1%) versus número de autovetores usando 2DPCA

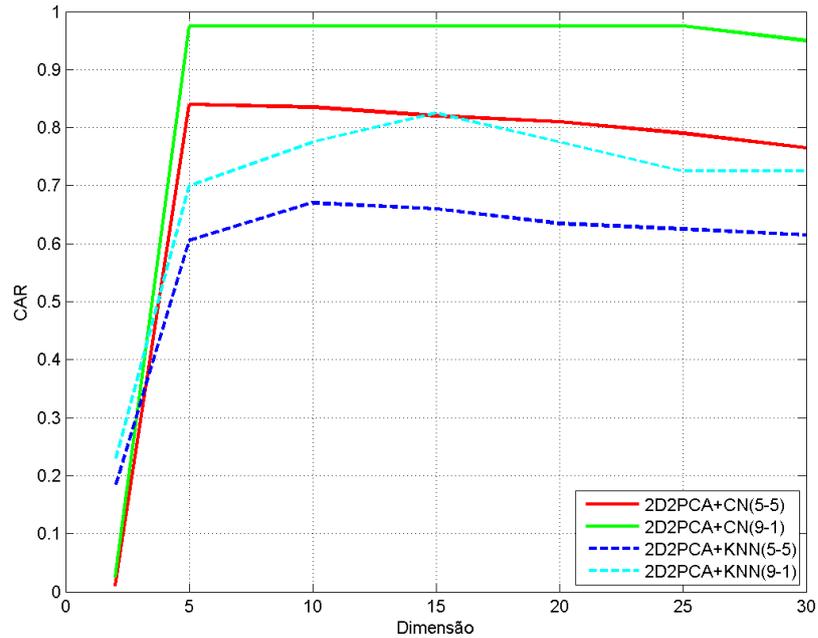


Figura 31. CAR (FAR=0,1%) versus número de autovetores usando  $(2D)^2PCA$

### 5.3.2 Impacto do número de autovetores versus EER

A Figura 32, Figura 33 e a Figura 34 mostra valores das taxas de erro equivalente (EER) versus o número de autovetores, para ambos os classificadores, utilizando os métodos PCA, 2DPCA e  $(2D)^2PCA$ . EER é o ponto onde FRR é igual à FAR.

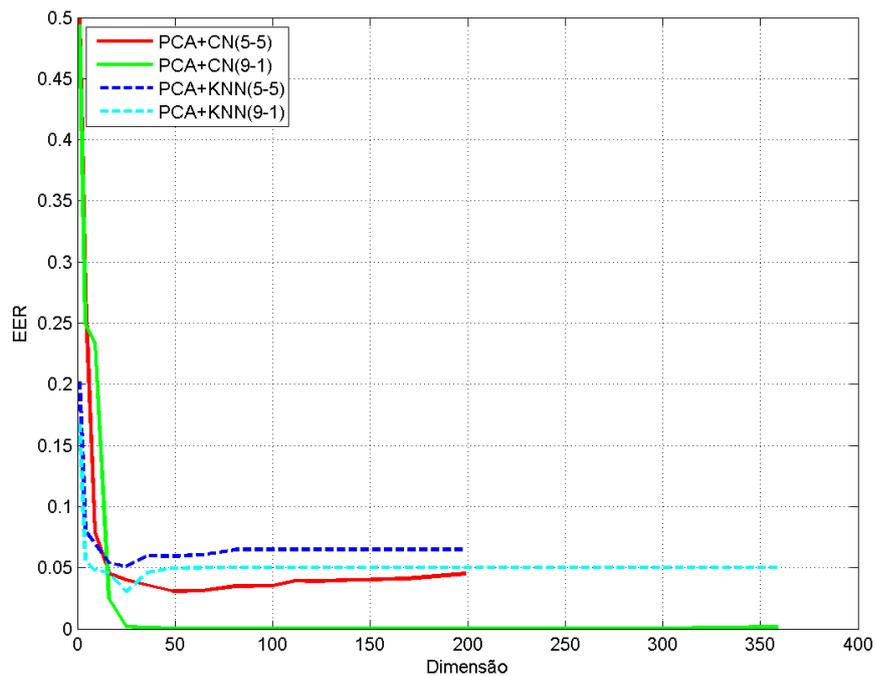


Figura 32. EER versus número de autovetores usando PCA

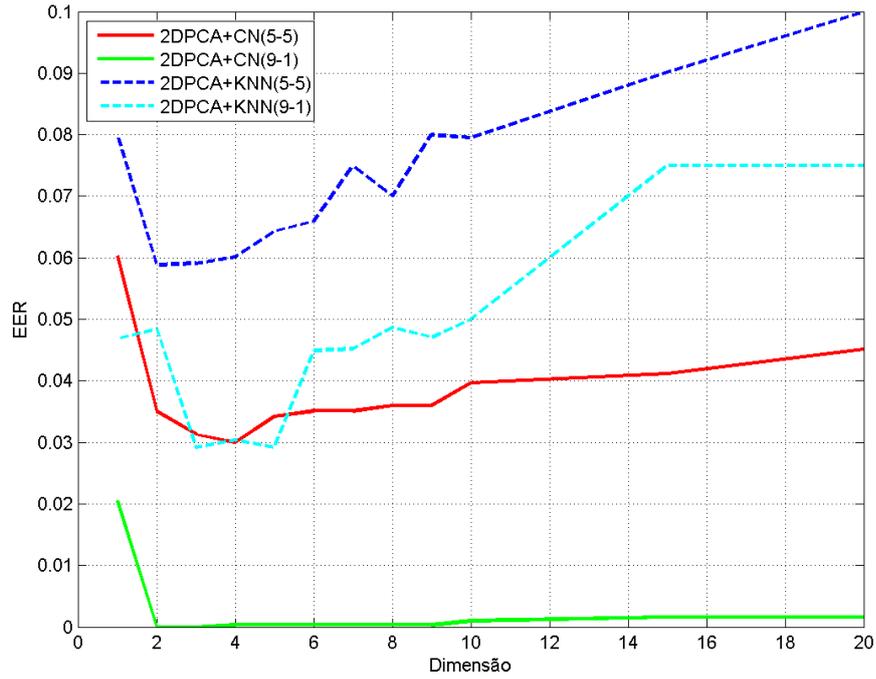


Figura 33. EER versus número de autovetores usando 2DPCA

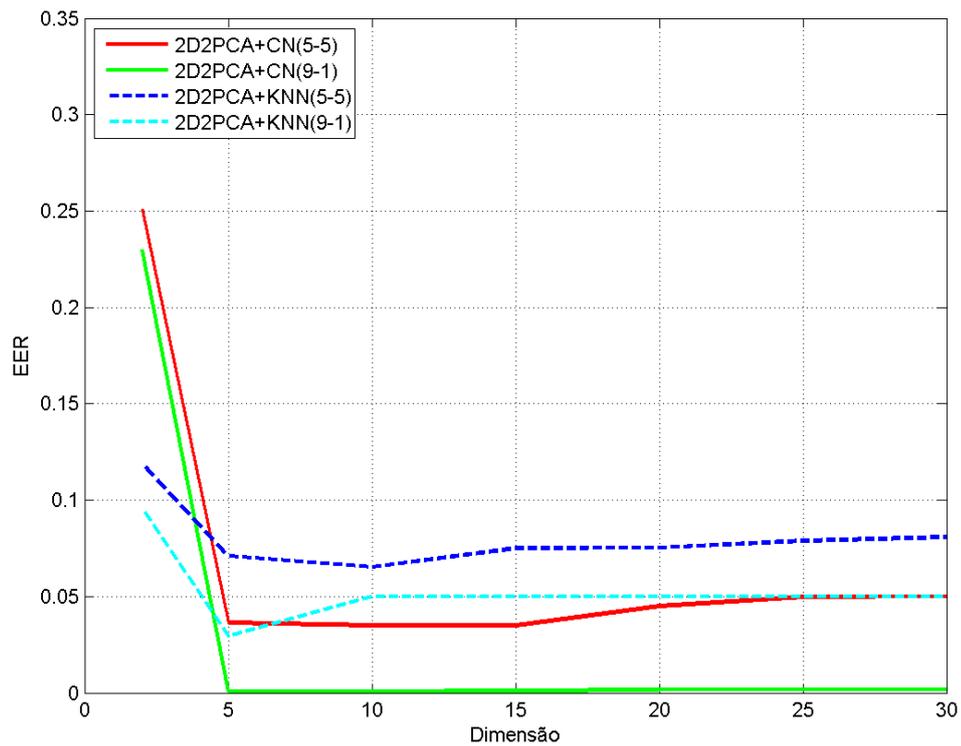


Figura 34. EER versus número de autovetore usando  $(2D)^2$ PCA

### 5.3.3 Curvas ROC

A Figura 35, Figura 36 e a Figura 37 mostram as curvas ROC.

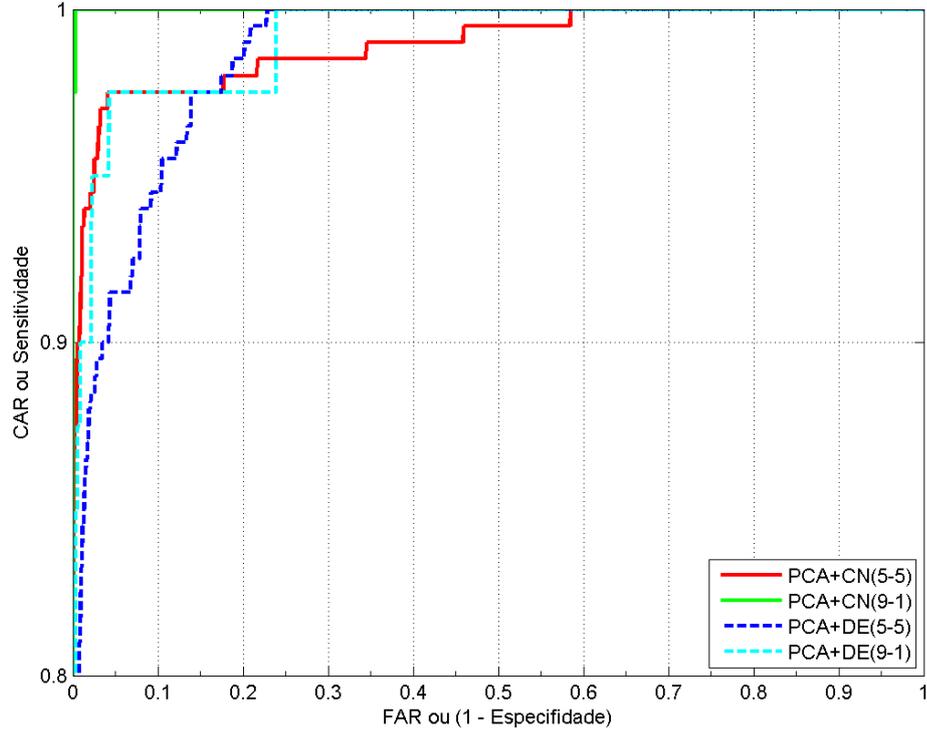


Figura 35. Curva ROC comparando os classificadores usando PCA

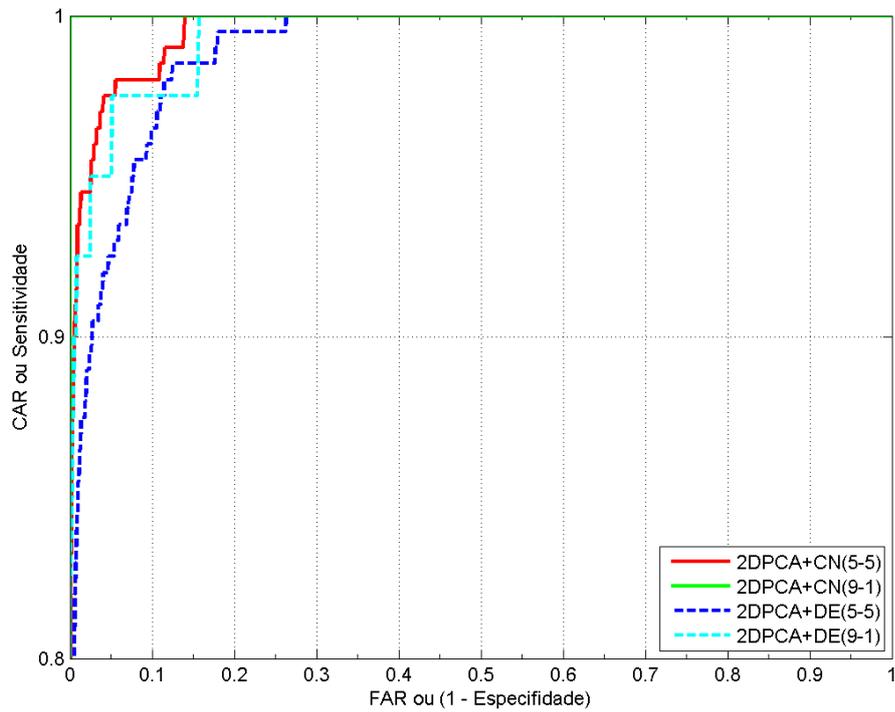
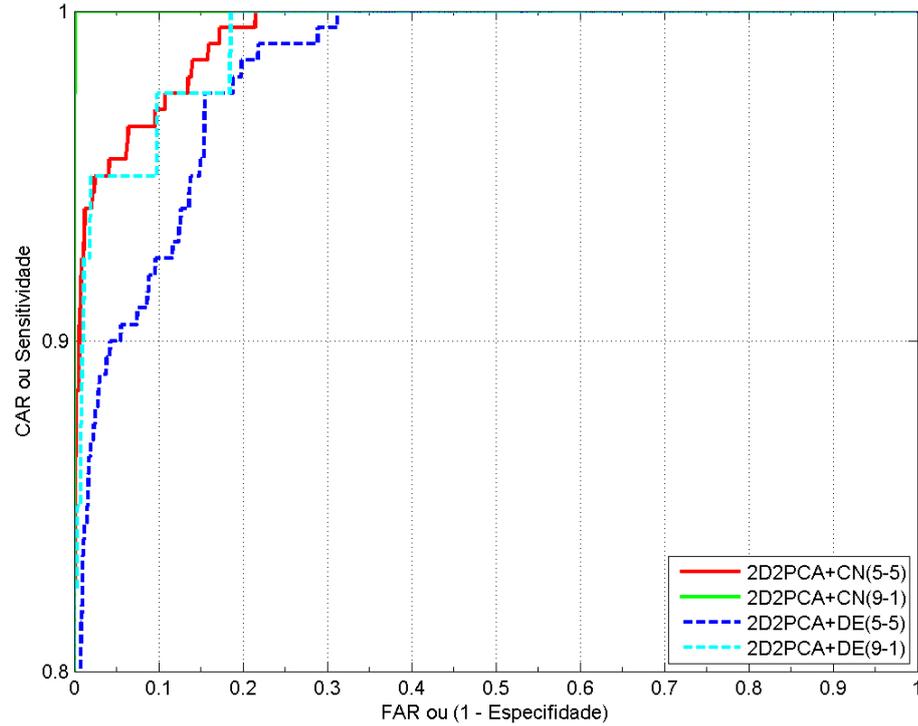


Figura 36. Curva ROC comparando os classificadores usando 2DPCA



**Figura 37. Curva ROC comparando os classificadores usando  $(2D)^2PCA$**

### 5.3.4 Resumo

A Tabela 5, Tabela 7 e a Tabela 8 mostram o desempenho dos classificadores no modo verificação, apresentando os melhores valores de CAR, EER e AUC.

**Tabela 5: Melhores valores de CAR (FAR=1%)**

Classificador	Estratégia	Extração de características	Dimensão	CAR (FAR=1%)
Classificador de Novidades	Metade-metade	PCA	64	92%
		2DPCA	112x2	93,5%
		$(2D)^2PCA$	5x5	92,5%
	10-fold Cross Validation	PCA	25	100%
		2DPCA	112x2	100%
		$(2D)^2PCA$	5x5	100%
Vizinho mais próximo	Metade-metade	PCA	25	82,5%
		2DPCA	112x4	85%
		$(2D)^2PCA$	10x10	83%
	10-fold Cross Validation	PCA	36	90%
		2DPCA	112x3	92,5%
		$(2D)^2PCA$	15x15	90%

**Tabela 6: Melhores valores de CAR (FAR=0,1%)**

Classificador	Estratégia	Extração de características	Dimensão	CAR (FAR=0,1%)
Classificador de Novidades	Metade-metade	PCA	36	88%
		2DPCA	112x5	85%
		(2D) <sup>2</sup> PCA	5x5	84%
	10-fold Cross Validation	PCA	25	100%
		2DPCA	112x2	100%
		(2D) <sup>2</sup> PCA	5x5	97,5%
Vizinho mais próximo	Metade-metade	PCA	49	63,5%
		2DPCA	112x4	68%
		(2D) <sup>2</sup> PCA	10x10	67%
	10-fold Cross Validation	PCA	49	77,5%
		2DPCA	112x4	85%
		(2D) <sup>2</sup> PCA	15x15	82,5%

Observando os resultados, o classificador de novidades alcançou as maiores taxas de Correta Aceitação (CAR) no PCA (92% e 100%), 2DPCA (93,5% e 100%) e (2D)<sup>2</sup>PCA (92,5% e 100%), nas estratégias metade-metade e *10-fold Cross Validation*, respectivamente.

**Tabela 7: Melhores taxas de erro equivalente (EER)**

Classificador	Estratégia	Extração de características	Dimensão	EER
Classificador de Novidades	Metade-metade	PCA	49	3,08%
		2DPCA	112x4	3,00%
		(2D) <sup>2</sup> PCA	15x15	3,49%
	10-fold Cross Validation	PCA	49	0,03%
		2DPCA	112x2	0,00%
		(2D) <sup>2</sup> PCA	5x5	0,06%
Vizinho mais próximo	Metade-metade	PCA	25	5,10%
		2DPCA	112x2	5,88%
		(2D) <sup>2</sup> PCA	10x10	6,52%
	10-fold Cross Validation	PCA	49	3,08%
		2DPCA	112x3	3,00%
		(2D) <sup>2</sup> PCA	5x5	2,92%

O classificador de novidades alcançou as menores taxas de erro equivalente no PCA (3,08% e 0,03%), 2DPCA (3,00% e 0%) e  $(2D)^2$ PCA (3,49% e 0,06%), nas estratégias *metade-metade* e *10-fold Cross Validation*, respectivamente.

**Tabela 8: Melhores valores de Área sob a curva ROC (AUC)**

Classificador	Estratégia	Extração de características	Dimensão	AUC
Classificador de Novidades	Metade-metade	PCA	36	0,9915
		2DPCA	112x2	0,9955
		$(2D)^2$ PCA	5x5	0,9949
	10-fold Cross Validation	PCA	25	0,9993
		2DPCA	112x4	1
		$(2D)^2$ PCA	15x15	1
Vizinho mais próximo	Metade-metade	PCA	25	0,9882
		2DPCA	112x4	0,9890
		$(2D)^2$ PCA	5x5	0,9851
	10-fold Cross Validation	PCA	25	0,9935
		2DPCA	112x2	0,9947
		$(2D)^2$ PCA	15x15	0,9909

Analisando os resultados, o classificador de novidades alcançou os maiores valores de AUC no PCA (0,9915 e 0,9993), 2DPCA (0,9955 e 1) e  $(2D)^2$ PCA (0,9949 e 1), nas estratégias *metade-metade* e *10-fold Cross Validation*, respectivamente.

#### 5.4 Discussão dos resultados

A maior taxa de reconhecimento *rank 1* atingida por Perlibakas (2004) foi de 85,34%, usando PCA na base ORL em conjunto com a distância do cosseno entre dois vetores, aplicando a transformada *whitening* descrita em Bishop (1995). O mesmo autor alcançou a menor taxa de erro equivalente (EER), 2,84% usando a distância de mahalanobis simplificada.

Chan et al (2010) obtiveram os melhores resultados na base ORL, usando PCA, estratégia *Leave-one-out* e kNN com a distância euclidiana. A taxa de reconhecimento e erro equivalente foram de 91,9% e 1,15%, respectivamente.

Yang et al (2004) usaram kNN com a distância euclidiana em seus experimentos. Atingiram uma acurácia de 96% (estratégia *Holdout*) e 98,3% (*Leave-one-out*).

Le e Bui (2011) usando 2DPCA e SVM na classificação atingiram 97,3% de taxa de reconhecimento.

MageshKumar et al (2011) usando PCA e redes neurais na classificação atingiram 98,6%.de taxa de reconhecimento usando estratégia *Holdout*.

Rouabhia e Tebbikh (2011) propõem e usam a distância matricial ponderada empregando no método 2DPCA, atingindo a melhor taxa de reconhecimento de 95,5%, utilizando a estratégia *Holdout*.

Através desses resultados apresentados verifica-se que:

- Ao comparar os resultados apresentados nesse trabalho em relação ao Vizinho mais próximo e ao Classificador de Novidades, verificamos que esse último apresentou resultados bem superiores, tanto no que diz respeito à taxa de reconhecimento como à taxa de erro equivalente;
- Ao comparar os resultados apresentados nesse trabalho para o Classificador de Novidades com os resultados da literatura anteriormente expostos nessa seção, verificamos que o mesmo apresenta os melhores resultados com a estratégia *10-fold Cross Validation*. Com respeito a estratégia metade-metade o Classificador de Novidades atingiu resultados inferiores aos publicados por MageshKumar et al (2011), porém superiores ao resultados apresentado por Rouabhia e Tebbikh (2011). Alguns autores não especificam a estratégia adotada, o que inviabiliza uma comparação rigorosa.

## 6 CONCLUSÃO

O Classificador de Novidades, anteriormente aplicado no reconhecimento de íris, detecção de vazamento de gás natural e detecção de lesões em imagens de cintilografia de mama, nunca tinha sido aplicado no reconhecimento facial, então neste trabalho, ele foi proposto como um novo método na classificação. Para testar e avaliar o desempenho do classificador proposto, o mesmo foi comparado o Vizinheiro mais próximo com a métrica da distância euclidiana, onde uma série de experimentos envolvendo identificação e verificação biométrica da face foram executados usando a base de faces ORL.

Os resultados mostraram que o classificador de novidades é eficaz no reconhecimento facial usando as conhecidas técnicas de extração de características: PCA, 2DPCA e  $(2D)^2PCA$ .

A grande vantagem do Classificador de Novidades é a utilização de algumas amostras para representar um modelo de um indivíduo (base do filtro de novidade).

Tendo em vista consolidar a utilização do classificador ora proposto, planeja-se como trabalhos futuros:

- o uso dos extratores de características LDA, 2DLDA e  $(2D)^2LDA$  na tentativa de melhorar a acurácia, visto que na literatura, geralmente, o método LDA supera o PCA;
- a utilização de outras bases de face;
- a utilização de algum pré-processamento fotométrico e geométrico;
- e uma comparação com outras distâncias, tais como distância de Mahalanobis, do cosseno, entre outras.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BELHUMEUR, P. N.; HESPANHA, J. P.; KRIEGMAN, D. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 19, n. 7, p. 711 - 720, Julho 1997.
- BOLLE, R. M. et al. **The Relation Between the ROC Curve and the CMC**. Fourth IEEE Workshop on Automatic Identification Advanced Technologies. Buffalo, NY: IEEE Computer Society. 2005. p. 15-20.
- COSTA, C. F. F. et al. Applying a novelty filter as a matching criterion to iris recognition for binary and real-valued feature vectors. **Signal, Image and Video Processing**, v. 7, n. 2, p. 287-296, Março 2013.
- COSTA, C.; MELO, R. O.; COSTA, M. G. F. Detecting Natural Gas Leaks Using Digital Images and Novelty Filters. In: KAMEL, M.; KARRAY, F.; HAGRAS, H. **Autonomous and Intelligent Systems**. Aveiro: Springer Berlin Heidelberg, v. 7326 of Lecture Notes in Computer Science, 2012. p. 242-249.
- COSTA, M.; MOURA, L. **Automatic assessment of scintmammographic images using a novelty filter**. Proceedings of the 19th Annual Symposium on Computer Applications in Medical Care. São Paulo: [s.n.]. 1995. p. 537-541.
- DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern Classification**. 2nd. ed. New York: Wiley-Interscience, 2000.
- FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, v. 27, n. 8, p. 861-874, Junho 2006.
- FISHER, R. A. The use of multiple measurements in taxonomic problems. **Annals of Eugenics**, v. 7, n. 2, p. 179-188, Setembro 1936.
- HAN, H. et al. A comparative study on illumination preprocessing in face recognition. **Pattern Recognition**, v. 46, n. 6, p. 1691-1699, Junho 2013.
- HOTELLING, H. Analysis of complex statistical variables into principal components. **Journal of Educational Psychology**, v. 24 (6), p. 417-441, 1933.
- JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical Pattern Recognition: A Review. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Washington, DC, v. 22, n. 1, p. 4-37, Janeiro 2000.
- JAIN, A. K.; ROSS, A. A.; NANDAKUMAR, K. **Introduction to Biometrics**. New York: Springer, 2011.
- JAIN, A.; HONG, L.; PANKANTI, S. Biometric identification. **Communications of the ACM**, New York, v. 43, n. 2, p. 90-98, Fevereiro 2000.
- KIRBY, M.; SIROVICH, L. Application of the Karhunen-Loeve procedure for the characterization of human faces. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 12, n. 1, p. 103,108, Janeiro 1990.

- KOHONEN, T.; OJA, E. Fast adaptive formation of orthogonalizing filters and associative memory in recurrent networks of neuron-like elements. **Biological Cybernetics**, Berlim, v. 21, n. 2, p. 85-95, 1976.
- KOHONEN, TEUVO, T. **Self-Organization and Associative Memory**. 3rd. ed. New York: Springer-Verlag New York, Inc., 1989.
- LE, T. H.; BUI, L. Face Recognition Based on SVM and 2DPCA. **International Journal of Signal Processing, Image Processing and Pattern Recognition**, v. 4, n. 3, Setembro 2011.
- LI, M.; YUAN, B. 2D-LDA: A statistical linear discriminant analysis for image matrix. **Pattern Recognition Letters**, v. 26, n. 5, p. 527–532, Abril 2005.
- LI, S. Z.; JAIN, A. K. **Handbook of Face Recognition**. 2nd. ed. London: Springer, 2011.
- LIAU, H.; ISA, D. New Illumination Compensation Method for Face Recognition. **International Journal of Computer and Network Security**, v. 2, n. 3, Março 2010.
- MAGESHKUMAR, C. et al. **Gabor features and LDA based face recognition with ANN classifier**. International Conference on Emerging Trends in Electrical and Computer Technology (ICETECT). Tamil Nadu: IEEE. 2011. p. 831 - 836.
- MARQUES FILHO, O.; VIEIRA NETO, H. **Processamento digital de imagens**. Rio de Janeiro: Brasport, 1999.
- METZ, C. E. Basic principles of ROC analysis. **Seminars in Nuclear Medicine**, Chicago, v. 8, n. 4, p. 283-298, Outubro 1978.
- NOUSHATH, S.; KUMAR, G. H.; SHIVAKUMARA, P. (2D)2LDA: An efficient approach for face recognition. **Pattern Recognition**, v. 39, n. 7, p. 1396–1400, Julho 2006.
- OLIVEIRA, L. et al. 2D Principal Component Analysis for Face and Facial-Expression Recognition. **Computing in Science & Engineering**, v. 13, n. 3, p. 9 - 13, Maio-Junho 2011.
- PADILLA, R. **Avaliação de Classificadores Haar Projetados para Detecção de Faces**. Universidade Federal do Amazonas. Manaus, p. 23. 2012.
- PEARSON, K. On lines and planes of closest fit to systems of points in space. **Philosophical Magazine Series 6**, v. 2, p. 559-572, 1901.
- PINHEIRO, C. **Reconhecimento de Indivíduos através da Íris usando Vetor Novidade**. Universidade Federal do Amazonas. Manaus. 2010.
- SAMARIA, F. S.; HARTER, A. C. **Parameterisation of a stochastic model for human face identification**. Proceedings of the Second IEEE Workshop on Applications of Computer Vision. Sarasota: IEEE. 1994. p. 138 - 142.
- SHERMINA, J. **Illumination invariant face recognition using Discrete Cosine Transform and Principal Component Analysis**. International Conference on Emerging Trends in Electrical and Computer Technology (ICETECT). Tamil Nadu: IEEE. 2011. p. 826 - 830.

- SIROVICH, L.; KIRBY, M. Low-Dimensional Procedure for Characterization of Human Faces. **J. Optical Soc. Am.**, v. 4, n. 3, p. 519-524, 1987.
- THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern Recognition**. 4th. ed. Oxford: Elsevier Academic Press, 2009.
- TURK, M.; PENTLAND, A. Eigenfaces for Recognition. **Journal of Cognitive Neuroscience**, v. 3, n. 1, p. 71-86, 1991.
- VIOLA, P.; JONES, M. J. Robust Real-Time Face Detection. **International Journal of Computer Vision**, v. 57, n. 2, p. 137-154, Maio 2004.
- WECHSLER, H. **Reliable Face Recognition Methods: System Design, Implementation and Evaluation**. New York: Springer, 2007.
- YANG, J. et al. Two-dimensional PCA: a new approach to appearance-based face representation and recognition. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 26, n. 1, p. 131 - 137, Janeiro 2004.
- YANG, J. et al. Two-dimensional discriminant transform for face recognition. **Pattern Recognition**, v. 38, n. 7, p. 1125–1129, Julho 2005.
- ZHANG, D.; ZHOU, Z.-H. (2D)2PCA: Two-directional two-dimensional PCA for efficient face representation and recognition. **Neurocomputing**, v. 69, n. 1-3, p. 224–231, Dezembro 2005.

## APÊNDICE

Este apêndice contém o artigo publicado resultante desta pesquisa:

1. **“Comparação do desempenho do Classificador de Novidades com o Classificador do Vizinho Mais Próximo no Reconhecimento Facial”**, publicado no *XI Simpósio Brasileiro de Automação Inteligente (SBAI)*, em outubro de 2013, Fortaleza, Ceará.
  - a. <http://www.sbai2013.ufc.br/pdfs/4767.pdf>