

ALEXANDRE FIENO DA SILVA

**Reconhecimento de Faces via PCA:
Análise de Desempenho**

Uberlândia, MG

Agosto/2006

ALEXANDRE FIENO DA SILVA

Reconhecimento de Faces via PCA: Análise de Desempenho

Dissertação apresentada como requisito parcial
à obtenção do grau de Mestre em Ciência da
Computação pelo Programa de Pós-Graduação
da Faculdade de Computação da Universidade
Federal de Uberlândia.

Área de Concentração: Banco de Dados

Orientador:

Prof^a. Dr^a. Celia Aparecida Zorzo Barcelos

Uberlândia, MG

Agosto/2006

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO

Os abaixo assinados, por meio deste, certificam que leram e recomendam para a Faculdade de Computação a aceitação da dissertação intitulada “Reconhecimento de Faces via PCA: Análise de Desempenho” por **Alexandre Fieno da Silva** como parte dos requisitos exigidos para a obtenção do título de **Mestre em Ciência da Computação**.

Uberlândia, 18 de Agosto de 2006

Orientador:

Prof^a. Dr^a. Celia Aparecida Zorzo Barcelos
Universidade Federal de Uberlândia UFU/MG

Banca Examinadora:

Prof. Dr. Maurílio Boaventura
Universidade Estadual Paulista UNESP/SP

Prof. Dr. Ilmério da Silva Reis
Universidade Federal de Uberlândia UFU/MG

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Data: Agosto, 2006

Autor: **Alexandre Fieno da Silva**
Título: **Reconhecimento de Faces via PCA: Análise de Desempenho**
Faculdade: **Faculdade de Computação**
Grau: **Mestrado**

Fica garantido à Universidade Federal de Uberlândia o direito de circulação e impressão de cópias deste documento para propósitos exclusivamente acadêmicos, desde que o autor seja devidamente informado.

*Dedico este trabalho primeiramente a Deus, pois sem Ele, nada seria possível
e não estaríamos aqui reunidos, desfrutando, juntos,
destes momentos que nos são tão importantes.*

*Aos meus pais Adão e Vera pelo esforço, dedicação e compreensão,
em todos os momentos desta e de outras caminhadas.*

*Em especial, à minha noiva Carolina pelo apoio,
amor e carinho a mim dedicados.*

*Aos meus amigos Ivan Lopes e Juliana Franciscani, que me ajudaram
no que puderam para que esta dissertação fosse concluída e, pelo
mútuo aprendizado de vida, durante nossa convivência,
no campo profissional e particular.*

Amigos, gratidão eterna!

Agradecimentos

À minha orientadora, Prof. Dra. Celia A. Zorzo Barcelos sou grato pela orientação.

Ao Prof. Dr. Ilmério Reis dos Santos pelas explicações e disponibilização de materiais usados na elaboração dessa dissertação.

Aos amigos Alexandre Grings, Vinicius (*smallville*), Vinícius Borges, Daniel, Sergio, Cristiane, Gustavo Carmo, Renê, Mylene, Marcio, Zé (cabelo), ao Marcos Aurélio e a todos aqueles que de alguma forma ajudaram com sugestões para esta dissertação.

Aos demais colegas de pós-graduação, pelas críticas e sugestões.

A minha família, a família de minha namorada, pelo apoio durante toda esta jornada.

Resumo

Com o aumento das atividades criminosas, como roubos a bancos, seqüestros, etc., a necessidade de identificar uma pessoa de forma mais segura e eficaz tornou-se imprescindível. Vários algoritmos são propostos a fim de resolver alguns desafios encontrados nos sistemas de recuperação, como efeitos de pose, iluminação, efeitos da idade e expressão facial.

Um método clássico e vastamente usado no reconhecimento de faces é a análise da componente principal (*Principal Component Analysis - PCA*) proposto por Turk e Pentland em 1991. O PCA tem como objetivo procurar as componentes principais da distribuição de faces ou seja os autovetores da matriz covariância do conjunto de faces, onde cada autovetor representa a variação entre as faces (autofaces).

Os estudos abordados neste trabalho são dedicados ao desempenho do PCA na recuperação de faces, considerando diferentes poses, expressões faciais e iluminação. O PCA é analisado quando a face consulta tem suas características modificadas (por exemplo, ruído, oclusões, recorte, etc.). É feita uma análise de como essas modificações podem alterar os resultados da busca, além da influência da iluminação e da pose na recuperação de face. Também é analisado o desempenho do PCA em relação à quantidade de autovalores que devem ser selecionados para se obter um bom resultado de recuperação e quais autovalores escolher.

No problema de iluminação um filtro de suavização é usado em um conjunto de faces e, posteriormente, o PCA é aplicado na parte da imagem que contém as texturas, os detalhes e os ruídos, oriundos da diferença da imagem original e da imagem suavizada. O objetivo é melhorar a recuperação de face sob as mais variadas condições de iluminação.

Palavras-chave: Análise da Componente Principal, Recuperação, Iluminação, Pose, Expressão facial.

Abstract

With the augment of criminal activities as bank robbery, kidnapping, etc., the needs of identifying a person in a safe manner became indispensable. Several algorithms have been proposed with the aim of resolving some of the challenges found in the retrieval systems as pose, illumination, age and facial expression effects.

A very well diffused method in face recognition is the PCA - Principal Component Analysis, proposed by Turk and Pentland in 1991. The PCA method has as the main goal to seek the principal components of faces distribution, which means, the eigenvectors of the covariance matrix of the face set, where each eigenvector represents the variation among the faces (auto faces) that can be seen as a feature set.

This work dedicates to the performance of the PCA method in face retrieval, taking into consideration the use of different poses, facial expression and illumination effects. The PCA is also evaluated when the query face has some of its features modified (i.e. noise, occlusion, cutting, etc). An analysis is done taking into consideration how these modified features could change the retrieval results, besides the illumination and poses influence. The performance of PCA method is also evaluated considering which ones and the quantity of eigenvalues to be selected by the algorithm, with the aim of reaching good retrieval results.

With the aim of verifying the illumination problem, a smooth filter is used in a set of faces and then the PCA method is applied in this set of images considering only those images pieces that have texture and noise presence, derived of the difference between the original and the smoothed image. The main goal is to improve the face retrieval under several illumination conditions.

Keywords: Principal Component Analysis, Retrieval, Illumination, Pose and facial expression.

Sumário

Lista de Figuras

Lista de Tabelas

Lista de Símbolos

1	Introdução	19
1.1	Introdução	19
1.2	Dificuldades no Reconhecimento de Face	22
1.2.1	A caricatura de uma face é fácil de ser reconhecida	22
1.2.2	A habilidade para tolerar a degradação aumenta com a familiaridade	23
1.2.3	As características faciais são processadas holisticamente	24
1.2.4	Sobrancelhas: uma característica muito importante para o reconheci- mento	24
1.3	Reconhecimento de face por computadores	25
1.4	Objetivos	26
1.5	Organização	27
2	Fundamentação Teórica	28
2.1	Reconhecimento de Padrões	28
2.1.1	Casamento de Modelos (<i>Template Matching</i>)	31

Sumário

2.1.2	Métodos Estatísticos	31
2.1.3	Análise estrutural-sintática	32
2.1.4	Redes Neurais	32
2.2	Reconhecimento de Faces	34
2.2.1	Detecção da Face	36
2.2.2	Normalização da face	37
2.2.3	Reconhecimento de face	37
2.2.3.1	Métodos Baseados na Aparência	38
2.2.3.2	Métodos Baseados em Características	39
2.3	Considerações Finais	39
3	Bases de dados	40
3.1	Introdução	40
3.2	Bases de dados de imagens de faces	40
3.2.1	Base de dados AR	41
3.2.2	Base de dados CMU-PIE	42
3.2.3	Base de dados FERET	43
3.3	Bases de dados usadas nesse trabalho	45
3.3.1	Base de dados ATT	45
3.3.2	Base de dados YaleA	46
3.3.3	Base de dados YaleB	47
3.4	Considerações Finais	47
4	Análise da Componente Principal	48
4.1	Introdução	48

Sumário

4.2	O PCA no Reconhecimento de Face	50
4.2.1	Método da Autoface	50
4.2.2	Projeção de uma face teste no espaço gerado pelas autofaces	54
4.3	Considerações Finais	55
5	Análise da Componente Independente	56
5.1	Introdução	56
5.2	Análise da Componente Independente	58
5.2.1	O modelo ICA	58
5.3	Arquiteturas ICA usadas no Reconhecimento de Faces	59
5.3.1	Arquitetura-I	60
5.3.2	Arquitetura-II	62
5.4	Resultados Experimentais	63
5.4.1	Comparação entre os métodos ICA e PCA	64
5.5	Considerações Finais	65
6	Análise Discriminante Linear	66
6.1	Introdução	66
6.2	Análise Discriminante Linear no Reconhecimento de Face	67
6.3	Resultados Experimentais	68
6.3.1	Reconhecimento com óculos	71
6.4	Considerações finais	72
7	Resultados Experimentais	73
7.1	Introdução	73

Sumário

7.2	Avaliação de um Sistema de Recuperação de Informação	73
7.2.1	Revocação e Precisão	74
7.3	Resultados Experimentais das autofaces	75
7.3.1	Seleção das autofaces	75
7.3.2	Etapa de geração das imagens consultas	76
7.3.3	Conjunto de treinamento e conjunto ideal	77
7.3.4	Experimento 1: Problemas envolvendo perturbações na face e mudan- ças da expressão facial	78
7.3.4.1	Resultados considerando o PCA construído com todas as im- agens do banco de dados	82
7.3.5	Experimento 2: Problema de diferentes condições de iluminação	93
7.3.6	Experimento 3: Problema de diferentes poses	109
7.4	Conclusões	121
8	O uso do PCA nas componentes de detalhes das imagens	122
8.1	Introdução	122
8.2	Descrição do Método Proposto	123
8.3	Resultados Experimentais	127
8.3.1	PCA (Detalhes) versus PCA (Padrão)	128
8.3.2	PCA (Detalhes) versus Outros Métodos	131
8.4	Considerações Finais e Trabalhos Futuros	133
9	Conclusão	134
	Referências	135

Lista de Figuras

1.1	Prosopagnosia.	19
1.2	Face montada a partir de duas faces famosas (J.F. Kennedy e Elvis Presley).	21
1.3	Caricaturas de pessoas famosas.	23
1.4	Seqüência de vídeo de segurança analisado por Burton et al.	23
1.5	Reconhecimento das características versus reconhecimento da face.	24
1.6	Amostra da contribuição da sobrancelha no reconhecimento de face.	25
2.1	Processos de sistema de reconhecimento de face.	34
2.2	Situação em que uma face é procurada em uma cena complexa e em uma cena não-complexa.	35
2.3	Verificação da face (autenticação) vs. Identificação da face (reconhecimento).	36
2.4	Imagem do perfil de uma face.	39
3.1	Base de dados AR.	42
3.2	Base de dados CMU-PIE (pose).	43
3.3	Base de dados CMU-PIE (iluminação e pose).	43
3.4	Base de dados FERET (Imagens frontais).	44
3.5	Base de dados FERET (Pose 1).	44
3.6	Base de dados FERET (Pose 2).	45
3.7	Base de dados ATT.	46
3.8	Base de dados YaleA.	46

Lista de Figuras

3.9	Base de dados YaleB.	47
4.1	Exemplo da concatenação de uma imagem 2D.	51
4.2	Exemplo das imagens de um conjunto de treinamento e a face média desse conjunto.	52
4.3	Exemplos das autofaces.	54
5.1	Arquiteturas usadas no ICA.	57
5.2	Arquitetura- <i>I</i>	60
5.3	Representação das imagens bases independentes.	60
5.4	Componentes independentes do conjunto de imagens obtidas pela arquitetura- <i>I</i>	62
5.5	Arquitetura- <i>II</i>	63
5.6	Imagens bases desconhecidas.	63
5.7	Imagens da representação ICA.	64
6.1	Exemplos de uma pessoa da base de dados YaleA e suas variações.	69
6.2	A variação no desempenho do método das autofaces.	69
6.3	Desempenho dos algoritmos quando aplicados à base de dados YaleA que contém variações na expressão facial e iluminação.	70
6.4	Comparação do PCA com o Fisherface em um problema de duas classes.	71
6.5	A imagem da esquerda é uma imagem da base de dados Yale de uma pessoa usando óculos. A imagem da direita é o Fisherface usado para determinar se uma pessoa está usando óculos.	72
7.1	Relação entre a revocação e a precisão.	75
7.2	Imagens usadas como consulta no processo de recuperação de faces.	77
7.3	Face montada a partir de duas faces de pessoas distintas.	77
7.4	Imagens usadas para formar o conjunto de treinamento.	78
7.5	Imagens usadas como conjunto ideal.	78

Lista de Figuras

7.6	Precisão: imagens com oclusão, ruído, com óculos escuros e sem um dos olhos e uma face montada a partir de duas faces.	79
7.7	Gráfico da imagem consulta com as características internas modificadas.	80
7.8	Gráfico da imagem consulta com metade da cabeça cortada, invertida, “estrita” e com “perucas”.	81
7.9	Imagens consultas usadas no processo de recuperação analisando a influência e a quantidade dos autovalores.	82
7.10	Gráficos da imagem consulta com oclusão (figura 7.9 – <i>a</i>).	83
7.11	Gráficos da imagem consulta sem metade da cabeça (figura 7.9 – <i>b</i>).	84
7.12	Gráficos das imagens consultas com ruído na face e nas características da face (figura 7.9 – <i>c</i> e <i>d</i>).	85
7.13	Gráficos da imagem consulta com a face invertida (figura 7.9 – <i>e</i>).	87
7.14	Gráficos da imagem consulta com a face “estrita” (figura 7.9 – <i>f</i>).	88
7.15	Gráficos da imagem consulta com a face com o cabelo cortado (figura 7.9 – <i>g</i>).	90
7.16	Gráficos da imagem consulta com as características internas modificadas (figura 7.9 – <i>h</i>).	91
7.17	Gráficos da imagem consulta de uma face pertencente à base de dados com as características internas modificadas (figura 7.9 – <i>i</i>).	93
7.18	Base de dados YaleB (iluminação).	94
7.19	Base de dados YaleB (faces usadas nas consultas).	94
7.20	Gráficos da precisão (Imagem consulta: figura 7.19- <i>a</i>).	96
7.21	Gráficos da precisão (Imagem consulta: figura 7.19- <i>b</i>).	97
7.22	Gráficos da precisão (Imagem consulta: figura 7.19- <i>c</i>).	99
7.23	Gráficos da precisão (Imagem consulta: figura 7.19- <i>d</i>).	100
7.24	Gráficos da precisão (Imagem consulta: figura 7.19- <i>e</i>).	102

Lista de Figuras

7.25	Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- <i>a</i>) em condições normais de iluminação.	103
7.26	Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- <i>b</i>) em condições normais de iluminação.	105
7.27	Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- <i>c</i>) em condições normais de iluminação.	106
7.28	Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- <i>d</i>) em condições normais de iluminação.	107
7.29	Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- <i>e</i>) em condições normais de iluminação.	108
7.30	Base de dados ATT (imagens usadas no cálculo das autofaces.)	109
7.31	Base de dados ATT (Imagens consultas).	109
7.32	Gráficos da precisão: Imagem consulta (figura 7.31- <i>a</i> .)	111
7.33	Gráficos da precisão: Imagem consulta (figura 7.31- <i>b</i> .)	112
7.34	Gráficos da precisão: Imagem consulta (figura 7.31- <i>c</i> .)	114
7.35	Gráficos da precisão: Imagem consulta (figura 7.31- <i>d</i> .)	115
7.36	Gráficos da precisão: Imagem consulta (figura 7.31- <i>e</i> .)	116
7.37	Gráficos da precisão: Imagem consulta (figura 7.31- <i>f</i> .)	117
7.38	Gráficos da precisão: Imagem consulta (figura 7.31- <i>g</i> .)	119
7.39	Gráficos da precisão: Imagem consulta (figura 7.31- <i>h</i> .)	120
8.1	Efeito da iluminação incidente na face.	122
8.2	Filtro de suavização aplicado em uma imagem de face.	126
8.3	Autofaces geradas a partir das imagens originais (<i>a</i>) e das imagens do conjunto $V = \{v_1, v_2, \dots, v_N\}$, (<i>b</i>).	126
8.4	Imagens da base de dados YaleB usadas para formar o conjunto de treinamento.	127

Lista de Figuras

8.5	Imagens da base de dados YaleB usadas na consulta.	128
8.6	Precisão vs Revocação da imagem consulta (<i>a</i>).	129
8.7	Precisão vs Revocação de outras faces de mesma intensidade de iluminação da face da figura 8.6.	129
8.8	Precisão vs Revocação: Gráfico usando a imagem consulta (<i>b</i>) : $A + 070E + 000$.	130
8.9	Precisão vs Revocação de outras faces de mesma intensidade de iluminação da face da figura 8.8.	130
8.10	Precisão vs Revocação: Gráfico usando a imagem consulta (<i>c</i>) : $A + 130E + 020$.	131
8.11	Precisão vs Revocação de outras faces de mesma intensidade de iluminação da face da figura 8.10.	131

Lista de Tabelas

1.1	Algumas empresas que desenvolvem sistema de reconhecimento de face.	22
2.1	Exemplos de aplicações no reconhecimento de padrões.	30
2.2	Alguns métodos de reconhecimento de faces de imagens estáticas.	37
3.1	Visão geral das bases de dados de faces.	41
3.2	Organização da base de dados AR.	42
3.3	Organização da base de dados CMU-PIE.	43
3.4	Organização da base de dados FERET.	45
3.5	Organização da base de dados ATT.	46
3.6	Organização da base de dados YaleA.	46
3.7	Organização da base de dados YaleB.	47
5.1	Taxa de reconhecimento da PCA, ICA- <i>I</i> e ICA- <i>II</i> na base de dados FERET.	65
6.1	Desempenho relativo dos algoritmos quando aplicados à base de dados YaleA.	70
6.2	Reconhecimento comparativo das taxas de erro para o reconhecimento de óculos/sem óculos com a base de dados YaleA [54].	72
8.1	Distribuição das imagens do conjunto de treinamento e de teste.	132
8.2	Comparação do PCA(Detalhes) com o trabalho de [74] usando a base de dados YaleB.	132

Lista de Símbolos

- M - Número de imagens da base de dados.
- N - Dimensão do vetor ($N = N_x N_y$), em que N_x é o número de linhas e N_y é o número de colunas.
- Γ_i - i -ésima imagem do conjunto de treinamento.
- Ψ - Face média.
- Φ_i - Diferença entre cada face do conjunto de treinamento com a face média.
- C - Matriz covariância ($N \times N$).
- L - Matriz covariância ($M \times M$).
- A - Vetor dos Φ 's calculados.
- P - Vetor das autofaces.
- ω_k - Pesos correspondentes à projeção de cada face no espaço das autofaces.
- Ω - Vetor que armazena todos os pesos calculados.
- Γ_{query} - Imagem usada como consulta.
- Φ_{query} - Diferença da imagem consulta com a face média.
- ω_{query} - Pesos correspondente à projeção da face consulta no espaço das autofaces.
- ϵ_k - Distância Euclidiana entre os k pesos encontrados.
- S - Conjunto desconhecido de imagens base estatisticamente independentes (Arquitetura ICA-I).
- A - Matriz de mistura (Arquitetura ICA-I).
- X - Faces originais (Arquiteturas ICA-I e ICA-II).
- W_I - Pesos aprendidos pelo ICA (Arquitetura ICA-I).
- U - Imagens base (componentes) independentes (Arquitetura ICA-I).
- R_m - Matriz dos coeficientes PCA usada na Arquitetura ICA-I.
- B - Coeficientes da combinação linear das componentes estatisticamente independentes com as imagens originais (Arquitetura ICA-I).
- X_{teste} - Imagem usada como consulta.

Lista de Símbolos

- S - Fontes desconhecidas (Arquitetura ICA-II).
- A - Imagens base desconhecidas (Arquitetura ICA-II).
- W_I - Pesos aprendidos pelo ICA (Arquitetura ICA-II).
- U - Coeficientes independentes (Arquitetura ICA-II).
- k - Quantidade de classes.
- N_j - Quantidade de faces pertencentes a cada classe k .
- Ψ_j - Face média de cada classe.
- S_w - Matriz dispersão intra-classe.
- S_b - Matriz dispersão entre-classes.
- S_T - Matriz dispersão total ($S_T = S_b + S_w$).
- P_{opt} - Matriz de projeção (critério de Fisher).
- P_{lda} - Matriz de projeção (Fisherfaces).
- P_{pca} - Matriz de projeção (autofaces).
- $I(x, y)$ - Imagem original.
- $u(x, y)$ - Imagem suavizada.
- $v(x, y)$ - Imagem contendo texturas, ruídos e detalhes da imagem original.

1 Introdução

1.1 Introdução

Nos últimos anos o reconhecimento de face despertou o interesse dos pesquisadores de diversas áreas, tais como, reconhecimento de padrões, visão computacional, processamento de imagens, neurociência, psicologia, etc., com propostas de diversos algoritmos [13,14,32,39,46].

Esse interesse foi motivado pela capacidade dos seres humanos em reconhecer pessoas e pelas diversas aplicações cotidianas, tais como, identificação criminal, verificação de cartões de crédito, segurança de sistemas, etc [39]. Em relação aos humanos há casos raros de certas deficiências do cérebro, em que essa capacidade é perdida, como por exemplo, a prosopagnosia. Um indivíduo com esta deficiência é incapaz de reconhecer uma pessoa por meio de sua face. Isso ocorre devido a algumas danificações no cérebro, fazendo com que a pessoa não enxergue com precisão algumas regiões da face, como os olhos, o nariz e a boca. A figura 1.1 ilustra uma face vista por uma pessoa sem essa deficiência e uma face vista por uma pessoa com prosopagnosia. Contudo, esses indivíduos são capazes de identificar pessoas usando outros centros do cérebro, como o centro responsável pelo reconhecimento de padrões, que a maioria das pessoas (sem essa deficiência) usa para reconhecer objetos [18].

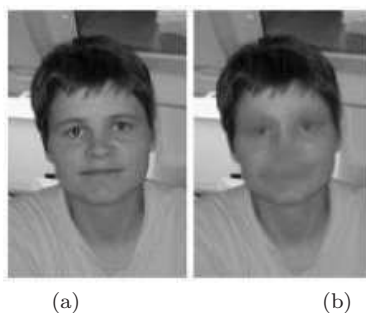


Figura 1.1: Prosopagnosia.
Fonte: Imagem obtida de [18].

Um outro fator que contribuiu para o interesse na área foi o aumento das atividades criminosas, como roubos a bancos, seqüestros, etc. A necessidade de identificar uma pessoa de forma mais segura e eficaz [9] tornou-se imprescindível. Uma das maneiras para se determinar essa identidade é o uso de suas características fisiológicas e comportamentais, denominadas biometria, única de pessoa para pessoa e muito difícil de ser roubada ou reproduzida. Geralmente, os sistemas biométricos incluem:

- reconhecimento de face;
- identificação por impressão digital;
- identificação pela geometria da mão;
- identificação da íris;
- reconhecimento por voz, e;
- identificação da retina.

A face, dentre as características biométricas, é a mais usada devido ao seu processo de obtenção não intrusivo, isto é, não há necessidade da cooperação da pessoa para que sua face seja capturada. Diferentemente de outros sistemas biométricos em que a participação do indivíduo é extremamente necessária, como por exemplo, na identificação por íris onde a pessoa precisa “se colocar” frente a um *scanner* para que a informação de sua íris seja capturada.

Entender o modo como os seres humanos reconhecem faces sob as mais diversas condições (iluminação, pose, dentre outros) [48] é fundamental para se desenvolver sistemas de reconhecimento de face. Contudo, desenvolver um sistema capaz de executar tais tarefas é muito difícil, devido às condições em que as faces são apresentadas, como por exemplo, imagem da face de uma pessoa na multidão ou uma imagem de face com pouca iluminação, além das diferenças na escala e na orientação ou por estar parcialmente tampada ou escondida.

Nos sistemas de reconhecimento de face por computador, um dos maiores problemas encontrados trata justamente dessas variações ambientais. Esses problemas são principalmente:

- **Mudança da expressão facial:** em que a face de uma pessoa sorrindo, chorando, com os olhos fechados, ou mesmo, uma pequena diferença na expressão facial que pode afetar significativamente os sistemas de reconhecimento de face.

- **Variação da iluminação:** a direção em que uma fonte de luz incide na face de uma pessoa pode prejudicar no seu reconhecimento. Um estudo dos efeitos da iluminação no reconhecimento de face [60] mostrou que iluminar a face de baixo para cima torna o reconhecimento de face ainda mais difícil.
- **Diferentes poses:** o ângulo, em relação ao eixo da câmera, pelo qual a imagem da face de uma pessoa foi obtida altera a precisão de um sistema de reconhecimento de face.

Durante anos houve intensa pesquisa científica no desempenho do reconhecimento de face procurando entender e caracterizar as representações e os processos envolvidos. A face, mais que qualquer outra parte do corpo, transmite individualidade, emoção, raça e idade (influência da idade), além de ser bastante útil para a análise de gênero (feminino ou masculino). Embora as pesquisas nas áreas da fisiologia, neurologia e psicologia tenham proporcionado um discernimento dos problemas de reconhecimento de face, os estudos ainda permanecem incompletos.

Mesmo sendo considerado muito superior aos sistemas computacionais, o sistema visual humano (SVH) apresenta algumas particularidades. Um bom exemplo a ser considerado pode ser a maneira como o SVH analisa e interpreta a face. De acordo com Sadr et al. [34], uma face montada à partir de duas faces distintas torna muito difícil o reconhecimento das duas identidades, pelo SVH. O contexto holístico¹ parece afetar o modo como as características individuais são processadas. Contudo, quando as duas metades da face são desalinhadas e a interpretação holística é interrompida, as duas identidades são facilmente reconhecidas (figura 1.2).

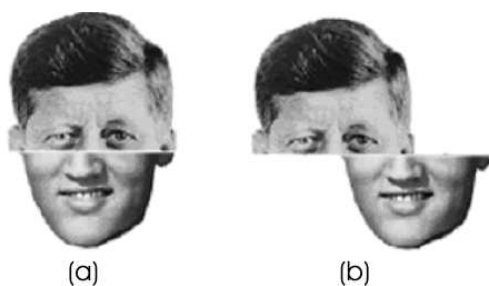


Figura 1.2: Face montada a partir de duas faces famosas (J.F. Kennedy e Elvis Presley).

Fonte: Imagem obtida de [65].

¹É a maneira como a face é analisada, isto é, seu contexto é considerada de uma só vez, pela análise do formato e das características internas.

Essa particularidade do SVH despertou o interesse dos pesquisadores de diversas áreas, tais como, reconhecimento de padrões, visão computacional, processamento de imagens, neurociência, psicologia, etc., a trabalharem no reconhecimento de face, propondo, assim, diversos algoritmos [13, 14, 32, 39, 46]. Apesar do desenvolvimento das pesquisas terem contribuído para a criação de sistemas de reconhecimento de face bastante eficientes, muitos em nível comercial (tabela 1.1), ainda não é possível encontrar sistemas que alcancem o mesmo desempenho do sistema visual humano.

Empresa	Site
AcSys Biometrics Corp	http://www.acsysbiometricscorp.com
Eyematic Interfaces Inc.	http://www.eyemetrics.com
Identix	http://www.identix.com
Viisage Technology	http://www.viisage.com

Tabela 1.1: Algumas empresas que desenvolvem sistema de reconhecimento de face.

Fonte: Tabela elaborada pelo autor.

1.2 Dificuldades no Reconhecimento de Face

Como citado anteriormente, o SVH possui algumas particularidades que são fundamentais no entendimento e desenvolvimento de sistemas de reconhecimento de face. Algumas dessas particularidades são comentadas a seguir.

1.2.1 A caricatura de uma face é fácil de ser reconhecida

Analisar as características faciais (por exemplo, sobrancelhas, olhos, nariz, boca, contornos da face e os contornos do cabelo), individualmente, ajuda a entender como essas características contribuem para o reconhecimento de face. Exemplos da exibição das características faciais podem ser encontrados nas faces representadas em caricaturas. O fato de que os seres humanos podem reconhecer faces conhecidas em caricaturas (figura 1.3) sem qualquer dificuldade, ainda não foi totalmente explorado nas pesquisas no reconhecimento de face [30, 50, 55, 69]. Mesmo com algumas características desenhadas em exageros ou baseadas em algumas características mais salientes, o reconhecimento não é afetado. A figura 1.3 ilustra caricaturas de pessoas famosas: (a) Vicent Van Gogh; (b) Jim Carrey; (c) Arnold Schwarzenegger; (d) Einstein; (e) G.W.Bush; e (f) Bill Gates.

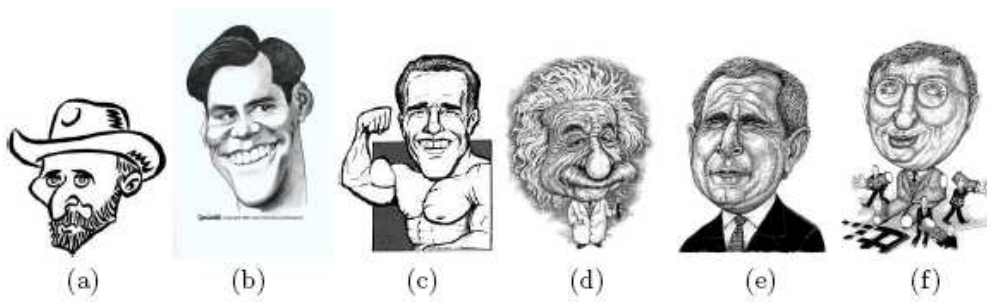


Figura 1.3: Caricaturas de pessoas famosas.
Fonte: Imagens obtidas de [1, 2]

1.2.2 A habilidade para tolerar a degradação aumenta com a familiaridade

A habilidade que o SVH tem em reconhecer a face aumenta dramaticamente com o grau de familiaridade da face visualizada. Burton [7] mostrou que o desempenho do reconhecimento de uma face por um observador, vista por câmeras de vídeo de baixa-resolução, é muito melhor quando as imagens dos indivíduos são familiares a esse observador. Analogamente, o desempenho do reconhecimento decai quando essa face não é muito familiar. A estrutura do corpo e o modo de andar são muito menos úteis para o reconhecimento do que as informações faciais (figura 1.4(a)), mesmo que a resolução naquela região esteja muito limitada. O desempenho do reconhecimento é levemente alterado quando o modo de andar ou a estrutura corporal é escondida (figura 1.4(b)), porém, quando a face é tampada o reconhecimento é dramaticamente afetado (figura 1.4(c)).

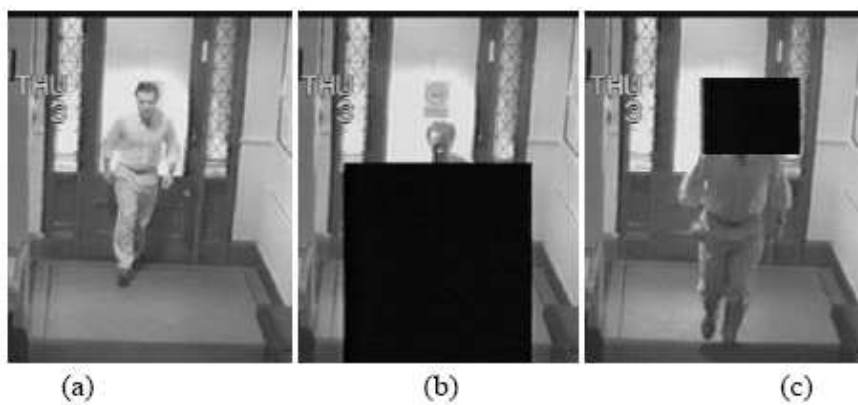


Figura 1.4: Seqüência de vídeo de segurança analisado por Burton et al.
Fonte: Imagem obtida de [65].

1.2.3 As características faciais são processadas holisticamente

De acordo com Sadr et al. [34] o contexto holístico (análise da face como um “todo”) parece afetar o modo como as características individuais são processadas. Esse resultado sugere que quando sozinha as características são algumas vezes suficientes para o reconhecimento facial. No contexto da face, contudo, a relação geométrica entre cada característica e o resto da face pode anular o diagnóstico daquela característica. Um exemplo pode ser visto na figura 1.5 em que é possível diferenciar as duas pessoas da figura (Bill Clinton e Al Gore) mesmo elas tendo as suas características internas (olhos, nariz e boca) idênticas.



Figura 1.5: Reconhecimento das características versus reconhecimento da face.
Fonte: Imagem adaptada de [56].

1.2.4 Sobrancelhas: uma característica muito importante para o reconhecimento

Nem todas as características faciais são igualmente analisadas no problema da identificação de uma face. Alguns resultados indicam a importância de analisar primeiro os olhos seguidos pela boca e então o nariz. Contudo, uma característica facial em particular tem recebido atenção dos pesquisadores, as sobrancelhas.

Sadr et.al [34] apresentou uma importante evidência sugerindo que as sobrancelhas poderiam não somente ser características importantes, mas que elas poderiam ser as mais importantes, superando até os olhos. Os experimentos foram realizados mostrando à expectadores imagens de faces de pessoas famosas com as sobrancelhas digitalmente apagadas. O desempenho do reconhecimento, quando as faces sem sobrancelhas foram apresentadas, foi

pior em relação às imagens originais. Esses resultados geraram algumas especulações sobre a importância da sobrancelha no reconhecimento de face. Primeiro, as sobrancelhas parecem ser muito importantes para transmitir emoções e outros sinais não verbais. Segundo, elas são úteis por ser uma característica facial muito “estável”.



Figura 1.6: Amostra da contribuição da sobrancelha no reconhecimento de face.
Fonte: Imagem obtida de [65].

1.3 Reconhecimento de face por computadores

Um sistema de reconhecimento de face é um sistema biométrico que identifica ou “analisa” seres humanos por meio da face. No geral, os sistemas de reconhecimento de face podem ser classificados em dois grupos, baseados no modo como a face é representada:

- *Métodos baseados na aparência*, que usam as características holísticas e são aplicados ou na face toda ou em regiões específicas da face (exemplos: PCA, LDA e ICA).
- *Métodos baseados em características*, que usam as características geométricas (boca, sobrancelhas, nariz, etc.) e a relação entre elas.

Dentro de cada grupo ainda é possível fazer mais classificações. A análise da componente principal (PCA) [51], a análise discriminante linear (LDA) [54] e a análise da componente independente (ICA) [44, 46] são os principais métodos baseados na aparência. O PCA [13, 51, 67] é uma técnica linear de projeção em um subespaço gerado pelos autovetores da matriz covariância das faces. Os vetores geradores definem as direções ao longo das

quais a variância das imagens originais é máxima. Enquanto o PCA lida com variâncias (estatísticas de segunda-ordem), o ICA [46] lida tanto com as estatísticas de segunda-ordem como as estatísticas de ordem superiores e projeta os dados de entrada em vetores bases estatisticamente independentes. Pode-se dizer que o ICA é uma generalização do PCA [44]. O LDA [54], diferente do PCA ou do ICA, usa as informações das classes e procura por um conjunto de vetores que maximize o critério discriminante de Fisher. Ele, simultaneamente, maximiza a matriz de dispersão entre-classes enquanto minimiza a matriz de dispersão intra-classe.

Os algoritmos baseados no PCA têm sido a base de numerosos projetos de pesquisas tanto na psicofísica quanto na visão computacional. Eles serviram e servem de comparação para novos algoritmos [32, 70], modelos computacionais na psicofísica [32] e de base para os algoritmos de reconhecimento de face [17, 21, 38, 51, 54].

Nos métodos baseados em características, a maioria usa o comprimento da cabeça, a distância entre os olhos e dos olhos com a boca, etc., ou as distâncias e ângulos entre os cantos dos olhos, os extremos da boca, nariz, etc. [73] no processo de reconhecimento. Um dos métodos baseados em características mais conhecido é o modelo Escondido de Markov [12, 27] (*Hidden Markov Model* - HMM).

1.4 **Objetivos**

O objetivo deste trabalho é analisar o desempenho do método PCA no reconhecimento de faces em diferentes condições (pose, iluminação e expressão facial). Especificamente, os objetivos são:

- Analisar o desempenho do PCA no reconhecimento de faces quando a face consulta é modificada por perturbações (por exemplo, ruído, oclusões, recorte, etc.) e pelas variações da iluminação e como essas modificações podem alterar os resultados da busca;
- Investigar a quantidade de autovalores e quais autovalores escolher para se obter um bom resultado de recuperação;
- Minimizar o efeito das variações de iluminação no reconhecimento de face com o uso de um filtro de suavização e a aplicação do PCA em imagens que contém texturas, detalhes e ruídos.

1.5 Organização

No Capítulo 2 apresenta-se os fundamentos dos sistemas de reconhecimento de padrões e dos sistemas de reconhecimento de face para computadores. No Capítulo 3 são apresentadas algumas bases de dados usadas no reconhecimento de face. Nos Capítulos 4, 5 e 6 são apresentados alguns dos principais métodos usados no reconhecimento da face: a análise da componente principal (PCA), a análise da componente independente (ICA) e a análise discriminante linear (LDA). No Capítulo 7 é apresentada uma análise experimental do método PCA sob variação da expressão facial e pose. No Capítulo 8 apresenta-se uma proposta para melhorar o desempenho do reconhecimento de face quando as faces estão sujeitas a diversas condições de iluminação. Finalmente, no Capítulo 9 apresenta-se conclusões relacionadas a esse trabalho.

2 Fundamentação Teórica

Nesta seção são apresentados os fundamentos dos sistemas de reconhecimento de padrões e de face.

2.1 Reconhecimento de Padrões

O reconhecimento de padrões é uma área de pesquisa que visa a classificação de objetos (padrões) em categorias ou classes. Por exemplo, dado um conjunto de k classes, c_1, \dots, c_k , e um padrão desconhecido x , um sistema de reconhecimento de padrões, auxiliado por pré-processamentos, extração e seleção de características, procura classificar x a uma classe c_i .

Segundo [10], nos últimos 50 anos de pesquisas, avanços foram obtidos possibilitando a evolução da pesquisa em aplicações mais complexas. Um exemplo é o reconhecimento de faces, que consiste em um problema de visão computacional que requer técnicas robustas a translação, a rotação, a alteração na escala e a deformações da face.

Há duas maneiras de se reconhecer ou classificar um padrão [10]:

- Classificação supervisionada (ex: análise discriminante): em que o padrão de entrada é identificado como um membro de uma classe pré-definida.
- Classificação não supervisionada (ex: clusterização): em que o padrão é determinado por uma classe, até o momento, desconhecida.

Observa-se, então, que o problema de reconhecimento de padrões consiste de uma tarefa de classificação ou categorização, em que as classes ou são definidas pelo projetista (classificação supervisionada) ou são “aprendidas” conforme a similaridade dos padrões (classificação não supervisionada).

O interesse na área de reconhecimento de padrões vem crescendo devido às aplicações, tanto desafiantes como computacionalmente mais exigentes (veja a tabela 2.1). Essas aplicações incluem mineração de dados, classificação de documentos, busca e organização em base de dados multimídia e biometria (identificação pessoal baseada nos atributos físicos, como a face e impressões digitais).

Com o avanço e disponibilidade de vários recursos computacionais tornou-se fácil o projeto e elaboração de métodos de análise e classificação de padrões. Em muitas aplicações, não há somente uma única abordagem para uma classificação “ideal”, de modo que a combinação de várias abordagens de classificadores é uma prática muito usada [10].

O projeto de um sistema de reconhecimento de padrões envolve, essencialmente, três aspectos:

- aquisição de dados (extração de características) e pré-processamento (seleção das características mais discriminantes);
- representação dos dados;
- tomada de decisões (construção de um classificador ou descritor).

A escolha de técnicas de pré-processamento, esquemas de representações e métodos para a tomada de decisões depende do domínio do problema. Um problema bem definido e restrito, em que se tem pequenas variações intra-classes e grandes variações entre-classes, indicará representações de padrões compactas (cujo a idéia é de que os dados estão bem organizados, isto é, dados similares ocupam regiões similares) e a uma estratégia de tomada de decisão simples (cujo objetivo é classificar corretamente indivíduos da mesma classe e separar os indivíduos de classes diferentes). O aprendizado, a partir de um conjunto de padrões (conjunto de treinamento), é um atributo importante e desejado na maioria dos sistemas de reconhecimento de padrões.

Um sistema de reconhecimento de padrões [10] pode ser classificado em:

1. casamento de modelos (*Template Matching*) [58, 59, 66];
2. método estatístico [70];
3. análise estrutural-sintática [22, 66];
4. redes neurais [66].

Domínio do Problema	Aplicação	Padrão de Entrada	Classes de Padrões
Bioinformática	Análise de seqüências	DNA/Seqüência de proteínas	Tipos conhecidos de genes/padrões
Mineração de dados	Busca por padrões significantes	Pontos em um espaço multi-dimensional	Compactar e separar grupos
Classificação de documentos	Busca na Internet	Documento texto	Categorias semânticas (ex: negócios, esportes, etc.)
Análise de documentos de imagem	Máquina de leitura para cegos	Documento imagem	Caracteres alfanuméricos, palavras
Automação industrial	Inspeção de circuitos impresso de placas	Intensidade ou alcance de imagem	Natureza do produto: defeituoso ou não
Recuperação de bases de dados multimídia	Busca na Internet	Vídeo clip	Gêneros de vídeo (ex: ação, diálogo, etc.)
Reconhecimento biométrico	Identificação pessoal	Face, íris, impressão digital	Usuários autorizados para controle de acesso
Sensoriamento remoto	Prognóstico da produção de colheita	Imagem multi-espectral	Categorias de aproveitamento de terra, desenvolvimento de padrões de colheita
Reconhecimento de voz	Inquérito por telefone sem assistência de operador	Voz em forma de onda	Palavras faladas

Tabela 2.1: Exemplos de aplicações no reconhecimento de padrões.

Fonte: Tabela obtida de [10].

Essas classificações não são necessariamente independentes, pois desde as pesquisas iniciais em reconhecimento de padrões, várias são as tentativas para os projetos de sistemas híbridos [10].

2.1.1 Casamento de Modelos (*Template Matching*)

Uma das primeiras e mais simples abordagens do reconhecimento de padrões é a técnica de casamento de modelos. O “casamento” é um processo usado para determinar a similaridade entre duas entidades (pontos, curvas ou forma) do mesmo tipo. O modelo é tipicamente uma forma $2D$ ou um protótipo [10].

O padrão a ser reconhecido é comparado com os modelos armazenados, levando-se em conta as principais variações possíveis (translação, rotação e mudanças de escala). A medida de similaridade é, geralmente, uma correlação ou uma função de distância. O casamento de modelos é computacionalmente exigente, mas a disponibilidade de recursos computacionais de hoje permite que essas abordagens sejam viabilizadas mais facilmente [10].

O casamento de modelo, mesmo eficaz em algumas aplicações, tem algumas desvantagens, como por exemplo, em casos em que os padrões são mal interpretados devido ao processo de visualização, mudanças no ponto de vista ou grandes variações intra-classes entre os padrões [58, 59, 66].

2.1.2 Métodos Estatísticos

No método estatístico cada padrão é representado por um conjunto m de características ou atributos chamado de vetor de característica m -dimensional. Padrões que possuem alguma relação ou particularidade em comum determinam uma classe. A efetividade do conjunto de características é determinada pela maneira como padrões de diferentes classes podem ser separados (limites de decisões).

O processo de decisão pode ser organizado da seguinte forma: dado um padrão, o objetivo é estabelecer limites de decisões, baseado no espaço de características, de modo a separar os padrões pertencentes a classes diferentes. Nos métodos estatísticos, os limites de decisões são determinados pela função probabilidade condicionada a uma classe de padrões [10, 70].

2.1.3 Análise estrutural-sintática

Em muitos problemas de reconhecimento envolvendo padrões complexos é mais apropriado adotar uma perspectiva hierárquica, em que um padrão é visto como uma composição de simples sub-padrões que são construídos de sub-padrões ainda mais simples [10].

Os métodos estruturais-sintáticos usam a estrutura interna do padrão como um elemento de análise. Eles baseiam-se no fato de que um padrão, como por exemplo, um objeto, pode ser escrito a partir de formas simples (primitivas), por meio de sua estrutura. A análise é realizada comparando cada uma das estruturas próprias do objeto com classes referências, ou determinando se o exemplo pertence ou não a famílias de modelos referências, gerados a partir de cada classe. Nessas técnicas, a representação do padrão não somente faz possível sua discriminação, como também sua reconstrução [22, 66].

Contornos e segmentos de contornos são as primitivas de componentes de padrão mais comuns. Além de contornos, regiões são também amplamente usadas. Elas podem ser quantitativamente caracterizadas por vários parâmetros, referindo-se ao tamanho e forma. Além disso, os histogramas de níveis de cinza baseados em parâmetros de regiões são bem conhecidos.

A representação do padrão na abordagem estrutural é baseada em elementos primitivos e seus relacionamentos. Essa informação é simbólica na sua natureza. Então, ao invés de um vetor de característica, como na abordagem estatística, são usadas outras estruturas de dados, freqüentemente de grande complexidade, como gráficos, etc.

Ao se projetar um sistema de reconhecimento de padrões sintático, é muito importante a questão da seleção de padrões primitivos. Na maioria das vezes o processo de seleção de primitivas é guiado por intuição e heurísticas.

2.1.4 Redes Neurais

A conectividade de uma rede neural determina sua estrutura. Grupos de neurônios podem ser localmente interconectados para formarem “aglomerados” que são conectados a outros aglomerados de forma imprecisa ou indireta. Alternativamente, neurônios podem ser organizados em grupos ou camadas que são (direcionalmente) conectadas a outras camadas. Assim sendo, as implementações de abordagem neural para reconhecimento de padrões requerem uma avaliação inicial de arquiteturas de redes neurais. As possibilidades são:

- projetar uma aplicação dependente da estrutura da rede que execute algumas computações desejadas;
- selecionar uma estrutura pré-existente “comumente usada” em que algoritmos de treinamento estejam disponíveis;
- adaptar a estrutura pré-existente para satisfazer uma aplicação específica.

De acordo com [10], modelos de redes neurais usam alguns princípios organizacionais, tais como aprendizado, generalização, adaptabilidade, tolerância a falhas e computação distribuída. Uma característica importante de redes neurais reside em sua habilidade de adaptação e aprendizagem, isto é, os modelos de redes neurais podem lidar com dados imprecisos e situações não totalmente definidas. Isso significa que ela é imune a pequenas falhas e que a rede poderá identificar um determinado padrão embora existam distorções em suas entradas [62].

Uma das redes neurais mais usadas para as tarefas de reconhecimento de padrão é o Mapa Auto-Organizável (*Self-Organizing Map* - SOM) ou rede de Kohonen. Esta rede exemplifica implementações neurais do aprendizado não-supervisionado, de forma a aglomerar ou auto-organizar padrões de entrada dentro de classes ou aglomerados baseada em alguma forma de similaridade. O processo de aprendizado envolve atualizações na arquitetura da rede e nos pesos das conexões, de forma que ela possa eficientemente executar uma tarefa específica de classificação/aglomeração.

A popularidade do uso de modelos de redes neurais para problemas de reconhecimento de padrões é devido à sua baixa dependência de conhecimento de domínio específico (relativo às abordagens baseadas em modelos e regras) e devido à disponibilidade de algoritmos eficientes de aprendizado.

Redes neurais provêm vários algoritmos não lineares para extração de características (usando camadas intermediárias) e classificação. Além do mais, existem algoritmos de extração de características e classificação que podem também ser mapeados em arquiteturas de redes neurais para implementação eficiente (*hardware*). Contrariando a aparente diferença de princípios de base, a maioria dos modelos de redes neurais bem conhecidos são implicitamente equivalentes ou similares a métodos estatísticos clássicos de reconhecimento de padrões.

2.2 Reconhecimento de Faces

O reconhecimento de face é uma área do reconhecimento de padrão executada especificamente em faces. Durante anos o estudo em faces vem atraindo mais e mais pesquisadores de diversas áreas, tais como, psicologia, ciência da computação, neurociência, etc.

Embora os seres humanos reconheçam faces em cenas desordenadas com relativa facilidade (capítulo 1), o reconhecimento automático é uma questão muito mais complicada.

De forma geral o problema pode ser formulado da seguinte maneira: dada uma imagem de uma cena estática ou de vídeo, o objetivo é identificar uma pessoa, nessa cena, com o uso de uma base de dados de faces. As informações secundárias, tais como raça, idade e sexo devem ser usadas para limitar as pesquisas. A solução do problema envolve a detecção da face a partir de uma cena com muitas pessoas (detecção), processamento da imagem (por exemplo, equalização de histograma, recorte, etc.) e a identificação (reconhecimento). A figura 2.1 ilustra esse sistema.

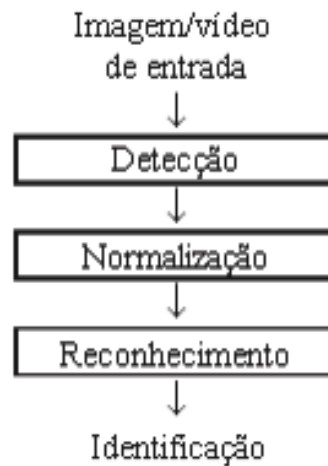


Figura 2.1: Processos de sistema de reconhecimento de face.

Fonte: Imagem obtida de [60].

Como pode ser visto, o processo de detecção é o primeiro passo de um sistema de reconhecimento de face: a detecção é realizada a partir de uma cena complexa, isto é, em que há muitas pessoas além de outros objetos na imagem ou a partir de uma cena não-complexa, em que as faces aparecem mais claramente. Exemplos de cenas complexas e não-complexas são mostrados na figura 2.2.

Posteriormente, a face detectada passa por um processo de normalização, com o objetivo



(a)



(b)

Figura 2.2: Situação em que uma face é procurada em uma cena complexa e em uma cena não-complexa.

Fonte: Imagem obtida de [60].

de tornar a face detectada ideal para o reconhecimento. O reconhecimento de face é executado como o último passo para identificar cada indivíduo comparando a face de entrada com uma base de dados de face existente. Se a face de entrada for similar a uma das imagens na base de dados, então ela será considerada “reconhecida”.

Geralmente, a etapa do reconhecimento é abordada de duas formas diferentes: (1) verificação da face (ou autenticação) e (2) identificação da face (ou reconhecimento). Como ilustrado na figura 2.3, a verificação da face envolve uma comparação “um-para-um” cuja identidade é solicitada, enquanto que a identificação da face envolve comparações “um-para-muitos” em que uma face (consulta) é comparada com todas as imagens de uma base de dados de modo a determinar sua identidade.

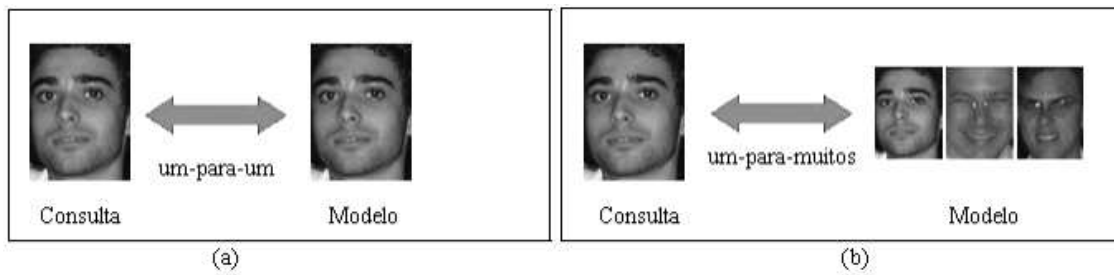


Figura 2.3: Verificação da face (autenticação) vs. Identificação da face (reconhecimento).
 Fonte: Imagem gerada pelo autor.

2.2.1 Detecção da Face

A detecção da face é realizada com o intuito de descobrir a presença de uma face em uma determinada imagem. Detectar a face antes de descobrir cada característica em particular poupa muito trabalho, uma vez que a maioria dos algoritmos baseiam-se na procura por tais elementos em toda a imagem. A vantagem de se detectar a face, em um primeiro momento, é que após esta fase a procura pelas características fica limitada apenas em uma determinada região da imagem. Basicamente, os métodos de detecção de faces podem ser divididos em três categorias:

1. métodos usando características invariantes;
2. métodos de casamento de modelos (*Templates*);
3. métodos baseados na aparência.

O objetivo dos métodos pertencentes à primeira categoria é encontrar as características geométricas ou estruturais da face, que existem mesmo quando a posição da pessoa, ponto de vista da câmera ou condições de iluminação variam.

Os métodos na segunda categoria visam detectar uma face procurando por ela em uma imagem e testando se ela corresponde a um modelo prévio da sua forma. Uma das formas mais comuns de modelar a forma de uma face é descrevê-la por meio de suas características geométricas básicas, como por exemplo, círculos, quadrados ou triângulos. Esta técnica é denominada *template*. A detecção do objeto, portanto, consistirá em achar a melhor correspondência entre a face presente na imagem e o seu molde (*template*).

Os métodos que compõem a abordagem baseada na aparência recebem este nome devido ao fato de não utilizarem nenhum conhecimento a priori sobre a face ou as características a serem

detectadas. Nesta classe de algoritmos surgem os conceitos de aprendizado e treinamento, uma vez que as informações necessárias para realizar a tarefa de detecção são retiradas do próprio conjunto de imagens sem intervenção externa.

2.2.2 Normalização da face

As imagens passam, geralmente, por um processo de normalização logo após a detecção. Esse processo pode, basicamente, ser dividido em dois outros: normalização geométrica e normalização fotométrica.

No processo de normalização geométrica, as imagens são normalizadas mantendo somente dados relacionados à face, como por exemplo, olhos, nariz e boca, e de modo que as imagens normalizadas, no final do processo, possuam a mesma posição e escala.

No processo de normalização fotométrico as condições da imagem são melhoradas de modo que ruídos, sombras e variações de contraste ou de brilho (causadas, principalmente, pelas mudanças de iluminação) não afetem, significativamente, o processo de reconhecimento.

2.2.3 Reconhecimento de face

Como comentado no capítulo 1, os métodos de reconhecimento de face podem ser divididos, basicamente, em dois grupos: os métodos baseados na aparência e os métodos baseados em características, que podem ainda, ser divididos em outras classificações (tabela 2.2).

Métodos de Reconhecimento de Face	
Métodos	Autores (trabalhos)
Métodos baseados na aparência	
Autofaces	Aplicação direta do PCA [49, 51]
Fisherface/sub-espaco LDA	FLD nas autoface [25, 54, 70]
SVM	Problema de 2-classes baseadas no SVM [73]
ICA	Análise de características baseada no ICA [44]
LDA/FLD	LDA/FLD em imagem natural [38]
PDBNN	Decisão probabilística baseada em redes neurais [64]
Métodos baseados em características	
Métodos geométricos puros	Métodos recentes [16]
Modelo de Markov Escondido	Métodos HMM [12, 27, 28]

Tabela 2.2: Alguns métodos de reconhecimento de faces de imagens estáticas.

Fonte: Tabela obtida de [73].

Os métodos baseados na aparência consideram todas ou algumas regiões da face no processo de reconhecimento. Dentre os métodos baseados na aparência, encontra-se o métodos das autofaces, proposto por Turk e Pentland [51].

Os métodos baseados em características usam as medidas de distância e ângulos entre pontos característicos da face como pontos extremos dos olhos, nariz, centro da boca, orelhas, etc. e combinações dessas medidas [73] (como por exemplo, a distância dos olhos com a boca) no processo de reconhecimento.

2.2.3.1 Métodos Baseados na Aparência

Os métodos baseados na aparência, também chamados de métodos holísticos, consideram todos os *pixels* da imagem ou de regiões características da face no processo de reconhecimento. Dentre os métodos baseados na aparência, os mais usados são a análise da componente principal (PCA) [13,51], a análise discriminante linear (LDA) [54] e a análise da componente independente (ICA) [44,46].

O método das autofaces (proposto por Turk e Pentland [13,51]) é baseado no PCA, e motivado pelo trabalho desenvolvido por Kirby e Sirovich [49]. O PCA é usado com o objetivo de caracterizar uma face mediante um conjunto de imagens de faces. A face é caracterizada por um pequeno conjunto de pesos correspondente à sua projeção no espaço das autofaces composta dos autovetores obtidos da variância das imagens de faces. O termo “autoface” é devido a semelhança que possuem com as imagens de faces.

O reconhecimento de face com o uso da análise discriminante linear (LDA) ou análise discriminante de Fisher é utilizado para encontrar uma transformação linear que maximize a dispersão entre-classe e minimize a dispersão intra-classe. Em outras palavras, o LDA agrupa as imagens de faces mais similares enquanto separa imagens de faces diferentes.

O método da análise da componente independente (ICA) proposto por Barlett et.al [44] codifica as imagens de faces com variáveis estatisticamente independentes. Enquanto o PCA procura direções no espaço de características que representam melhor as faces, o ICA procura direções que são mais independentes. Existem duas arquiteturas para o ICA [46]: a primeira é usada para encontrar um conjunto de imagens estatisticamente independentes que podem ser vistas como características independentes de um dado conjunto de imagens [6] e a segunda arquitetura é usada para encontrar filtros da imagem que produzem saídas independentes estatisticamente. Em ambas arquiteturas o PCA é usado para reduzir a dimensionalidade.

Os métodos acima citados serão abordados em mais detalhes nos capítulos 4, 5 e 6.

2.2.3.2 Métodos Baseados em Características

Os métodos de reconhecimento de faces baseados em características, fundamentam-se na construção de um vetor de características a partir das medidas de distância e ângulos entre os pontos característicos da face, tais como, pontos extremos dos olhos, nariz, orelhas, sobrancelhas, etc. e combinações dessas medidas [73].

Além da aplicação em imagens 2-D, esses métodos podem ser aplicados em imagens de faces de perfil, em que os pontos característicos são, por exemplo, a ponta do nariz, a sobrancelha, o queixo, etc (figura 2.4). A vantagem do reconhecimento a partir de perfis está no uso de informações que não estão disponíveis nas imagens frontais (por exemplo, o tamanho do nariz), além do fato de não ser difícil extrair esses pontos quando o fundo é uniforme.

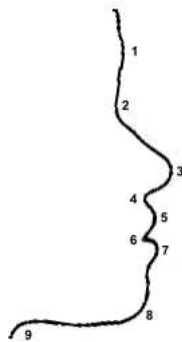


Figura 2.4: Imagem do perfil de uma face.

2.3 Considerações Finais

Neste capítulo foi apresentado os fundamentos do reconhecimento de padrões e do reconhecimento de face. Os principais métodos foram sucintamente descritos em ambas as abordagens. Em se tratando do reconhecimento de faces, os três principais métodos baseados na aparência, o PCA, o ICA e o LDA, são descritos. Uma abordagem mais detalhada é mostrada nos capítulos 4, 5 e 6, respectivamente.

3 Bases de dados

3.1 Introdução

O reconhecimento de face é uma das áreas de pesquisa mais populares da visão computacional. Com o desenvolvimento dos sistemas de reconhecimento de face, um grande número de bases de dados de faces têm sido criadas. Contudo, muitas dessas bases são feitas exclusivamente para determinados tópicos, como por exemplo, para avaliar o desempenho de sistemas de reconhecimento de face sob variação de iluminação, pose, expressão facial, etc. Nessa seção são descritas, resumidamente, algumas bases de dados usadas no reconhecimento de face e disponíveis publicamente.

3.2 Bases de dados de imagens de faces

A seguir é fornecida uma tabela resumindo os pontos-chaves da base de dados, incluindo o número de pessoas, as condições em que as imagens foram obtidas, a resolução da imagem, o número total de imagens e o endereço em que esta base de dados está disponível. A tabela 3.1 apresenta uma visão geral de uma boa parte das bases de dados usadas nos testes de reconhecimento de face. Casos em que o número exato das condições (pose, iluminação e expressão facial) não é determinado (ou por ter aumentado o número de imagens ou as condições não foram controladas durante a gravação) são marcadas com “++”.

Bases de Dados	Núm. de Faces	Pose	Iluminação	Expressão Facial
AR	126	1	4	4
ATT	400	1	++	++
BANCA	208	1	++	1
CAS-PEAL	1040	27	15	11
CMU PIE	68	13	43	3
FERET	1199	9 – 20	2	2
Harvard RL	10	1	77 – 84	1
KFDB	1000	7	16	5
MIT	16	3	3	1
NIST MID	1573	2	1	++
UMIST	20	++	1	++
U. Texas	284	++	1	++
U. Oulu	125	1	16	1
XM2VTS	295	++	1	++
YaleA	165	1	3	++
YaleB	640	9	64	1

Tabela 3.1: Visão geral das bases de dados de faces.

Fonte: Tabela obtida de [29].

3.2.1 Base de dados AR

A base de dados AR [11] foi coletada no Centro de Visão Computacional em Barcelona, Espanha, em 1998. Ela contém imagens de 126 pessoas (70 homens e 56 mulheres). As condições de gravação foram cuidadosamente controladas (parâmetros da câmera, iluminação, distância da câmera) para assegurar que as configurações fossem idênticas entre as pessoas. As imagens são coloridas (padrão RGB e de tamanho 768×576 pixels) e foram gravadas 2 vezes em um intervalo de 2 semanas. Durante cada sessão, as imagens de cada pessoa foram obtidas em 13 situações de variação da expressão facial, iluminação e oclusão. A figura 3.1 ilustra um exemplo de cada situação: (1) neutro, (2) sorrindo, (3) raiva, (4) gritando, (5) luz na esquerda, (6) luz na direita, (7) luz de ambos os lados, (8) com óculos de sol, (9) com óculos de sol e luz na esquerda, (10) com óculos de sol e luz na direita, (11) lenço no pescoço, (12) lenço no pescoço e luz na esquerda, (13) lenço no pescoço e luz na direita.

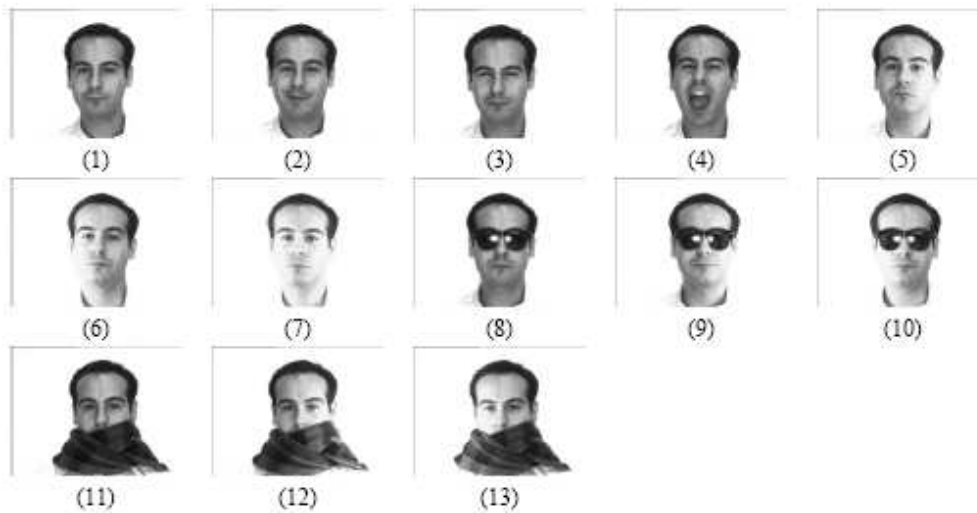


Figura 3.1: Base de dados AR.

Fonte: Imagens obtidas de [11].

No. de pessoas	Condições		Resolução da imagem	No. de imagens
126	Oclusão	6	768 × 576	3276
	Iluminação	3		
	Expressão facial	4		
<i>http : //rvl1.ecn.purdue.edu/~aleix/aleix_face_DB.html</i>				

Tabela 3.2: Organização da base de dados AR.

Fonte: Dados obtidos de [11].

3.2.2 Base de dados CMU-PIE

A base de dados CMU-PIE fornece amostras de um grande número de imagens em condições de iluminação e pose com expressões faciais diferentes. Essa base de dados contém mais de 40000 imagens de 68 pessoas¹. As fotos foram tiradas em 13 poses diferentes, com 43 condições de iluminação e com 4 de expressão facial. As imagens resultantes (coloridas e no padrão RGB) tem tamanho 640 × 480 pixels. As figuras 3.2 e 3.3 mostram exemplos das imagens nas condições diferentes de pose e iluminação.

¹Foram tiradas 600 imagens de cada pessoa em diferentes poses, iluminação e expressão Facial. Essa base de dados tem 40Gb e pode ser solicitada pelo email simonb@cs.cmu.edu.



Figura 3.2: Base de dados CMU-PIE (pose).

Fonte: Imagens obtidas de [68].

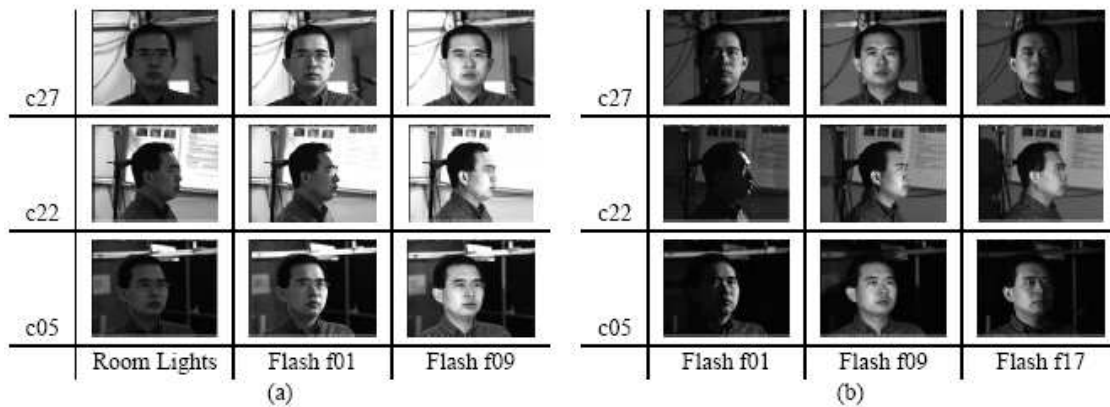


Figura 3.3: Base de dados CMU-PIE (iluminação e pose).

Fonte: Imagens obtidas de [68].

No. de pessoas	Condições		Resolução da imagem	No. de imagens
68	Pose	13	640 × 480	Mais de 40000
	Iluminação	43		
	Expressão facial	4		
http://www.ri.cmu.edu/projects/project_418.html				

Tabela 3.3: Organização da base de dados CMU-PIE.

Fonte: Dados obtidos de [68].

3.2.3 Base de dados FERET

A base de dados de tecnologia em reconhecimento de face (FERET) foi coletada na Universidade Mason George e patrocinada pelo Programa de Desenvolvimento de Tecnologia do Departamento de Defesa Anti-Drogas dos Estados Unidos (*US Department of Defense Counterdrug Technology Development Program*) [57] entre 1993 e 1997.

A base de dados FERET lista 24 categorias de imagens de face². Como os equipamentos usados nas gravações tiveram que ser reagrupados em cada sessão, pequenas variações são apresentadas entre as sessões. As imagens resultantes tem tamanho 256×384 pixels. A figura 3.4 mostra cinco grupos de imagens frontais.

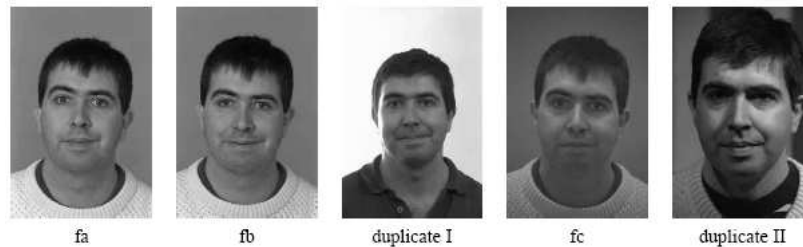


Figura 3.4: Base de dados FERET (Imagens frontais).

Fonte: Imagens obtidas de [57].

As imagens *fa* e *fb* foram obtidas em um espaço de tempo pequeno. Foi pedido para que as pessoas fizessem diferentes expressões faciais para as imagens *fb*. As mudanças resultantes na expressão facial são sutis, geralmente alternando entre “neutra” e “sorrindo”. As imagens na categoria *fc* foram gravadas por diferentes câmeras e em diferentes condições de iluminação. Uma quantidade de pessoas retornaram em uma data posterior à última sessão e foram retratadas novamente. Para as imagens no conjunto *Duplicate I*, passaram 1031 dias entre as sessões gravadas. Um subconjunto dessas imagens forma o conjunto *Duplicate II*, em que no mínimo 18 meses separaram cada sessão.

As categorias de imagens restantes abrangem um vasto conjunto de variações de pose. A figura 3.5 mostra as categorias *ba* até *bi*. As imagens de cada pessoa foram gravadas com seu corpo e sua cabeça giradas. O ângulo das poses variam de $+60^\circ$ até -60° .

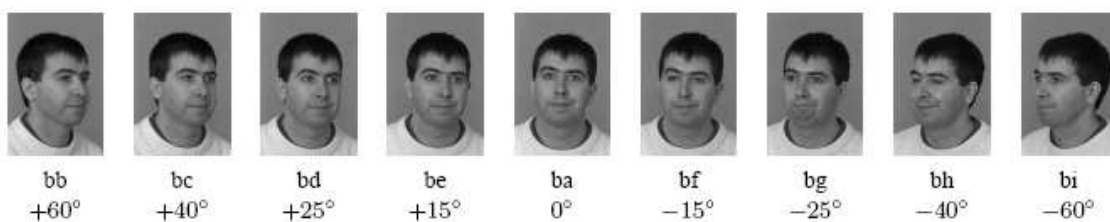


Figura 3.5: Base de dados FERET (Pose 1).

Fonte: Imagens obtidas de [57].

²Em http://www.itl.nist.gov/iad/humanid/feret/feret_master.html.

Um conjunto diferente das imagens de poses é mostrado na figura 3.6. Essas imagens foram obtidas a partir dos seguintes aspectos: perfil da direita e esquerda (chamadas de *pr* e *pl*), um quarto do perfil esquerdo e do direito (*ql* e *qr*) e metade do perfil direito e do esquerdo (*hr* e *hl*)³.

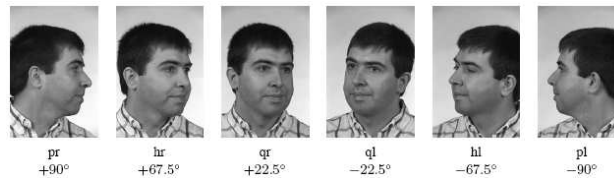


Figura 3.6: Base de dados FERET (Pose 2).
Fonte: Imagens obtidas de [57].

No. de pessoas	Condições		Resolução da imagem	No. de imagens
1199	Pose	9 – 20	256 × 384	14051
	Iluminação	2		
	Expressão facial	2		
	Espaço de tempo	2		
http://www.nist.gov/humanid/feret				

Tabela 3.4: Organização da base de dados FERET.
Fonte: Dados obtidos de [57].

Nesta seção foram apresentadas algumas bases de dados usadas nas pesquisas de face. Outras informações sobre as bases de dados da tabela 3.1 podem ser obtidas em [63].

3.3 Bases de dados usadas nesse trabalho

Nesta seção é descrita as bases de dados YaleA [3], YaleB [4] e ATT [5] usadas em nossos experimentos.

3.3.1 Base de dados ATT

A base de dados ATT foi coletada entre 1992 e 1994 [27]. Ela contém pequenas variações na iluminação, expressão facial (olhos abertos e fechados, sorrindo e sério) e detalhes faciais (com e sem óculos). A tabela 3.5 resume a base de dados ATT. Exemplos de algumas imagens podem ser vistos na figura 3.7.

³Mais informações sobre a base de dados FERET são encontradas em <http://www.nist.gov/humanid/colorferet>.



Figura 3.7: Base de dados ATT.

Fonte: Imagens obtidas de [5].

No. de pessoas	Resolução da imagem	No. de imagens
10	92×110	400
http://www.uk.research.att.com/facedatabase.html		

Tabela 3.5: Organização da base de dados ATT.

Fonte: Dados obtidos de [5].

3.3.2 Base de dados YaleA

A base de dados YaleA [3] contém imagens de 15 pessoas em 11 expressões faciais diferentes. Cada face é caracterizada por diferentes expressões faciais ou aspectos: feliz, normal, triste, sonolento, surpreso e piscando os olhos; com e sem óculos; e, por três direções de iluminação: luz central, luz lateral esquerda e luz lateral direita (figura 3.8). A tabela 3.6 resume a base de dados YaleA.



Figura 3.8: Base de dados YaleA.

Fonte: Imagens obtidas de [3].

No. de pessoas	Condições		Resolução da imagem	No. de imagens
15	Com e sem óculos	2	320×243	165
	Iluminação	3		
	Expressão facial	6		
http://cvc.yale.edu/projects/yalefaces/yalefaces.html				

Tabela 3.6: Organização da base de dados YaleA.

Fonte: Dados obtidos de [3].

3.3.3 Base de dados YaleB

A base de dados YaleB [4] foi montada para permitir testes dos métodos de reconhecimento de face sujeitas a grandes variações na iluminação e na pose. Imagens de 10 pessoas sob 64 condições de iluminação e 9 poses diferentes (uma frontal, cinco poses em 12° e três poses em 24°) a partir do eixo da câmera. A base de dados é dividida em quatro sub-conjuntos de acordo com o ângulo entre a fonte de luz e o eixo da câmera (12° , 25° , 50° , 77°). Exemplos de imagens da base de dados são mostrados na figura 3.9. A tabela 3.7 resume a base de dados YaleB.

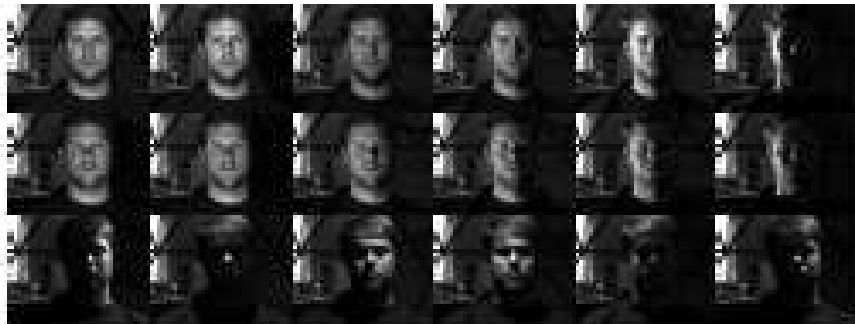


Figura 3.9: Base de dados YaleB.

Fonte: Imagens obtidas de [4].

No. de pessoas	Condições		Resolução da imagem	No. de imagens
10	Pose	9	640×480	5760
	Iluminação	64		
http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html				

Tabela 3.7: Organização da base de dados YaleB.

Fonte: Imagens obtidas de [4].

3.4 Considerações Finais

Cada base de dados descrita neste capítulo, além das citadas na tabela 3.1 são de extrema importância na análise do desempenho dos sistemas de reconhecimento de faces. As três bases utilizadas neste trabalho foram selecionadas por tratarem, especificamente, dos problemas como expressão facial, pose e iluminação.

4 Análise da Componente Principal

4.1 Introdução

A análise da componente principal (*Principal Component Analysis* - PCA) é uma técnica matemática que descreve um conjunto de dados usando “componentes principais”, escrita como combinações lineares dos dados originais. As componentes principais são determinadas em ordem decrescente de importância. A primeira componente possui mais informação do que a segunda, e assim por diante. O PCA tenta construir um pequeno conjunto de componentes que resumem os dados originais, reduzindo a dimensionalidade dos mesmos.

Como sendo uma das técnicas mais antigas e conhecida na análise multivariada e mineração de dados, o PCA foi apresentado por Pearson [19] que a usou em um contexto biológico e, posteriormente, foi desenvolvido por Hotelling [19] em trabalhos feitos na psicometria. Também foi desenvolvido independentemente por Karhunen no contexto da teoria da probabilidade e, generalizada por Loeve [19].

Dentre os objetivos do PCA, temos:

1. Redução da dimensionalidade;
2. Seleção de características: escolha dos componentes mais significantes.

O PCA é vastamente estudado e usado nas áreas de processamento de sinais e reconhecimento de padrões. Também é usado em técnicas de mineração e redução de dados. No processamento de imagens e visão computacional, as representações do PCA são usadas para solucionar problemas da face e de objetos, tais como reconhecimento, detecção, determinação de formas, aparência e gestos, dentre outros.

As componentes principais consideradas as mais importantes, possuem as direções em que os dados de entrada têm as maiores variâncias e, as outras componentes (componentes

com as menores variâncias, chamadas de componentes menores) são consideradas menos importantes ou associadas a ruídos.

De um modo geral, o PCA é relacionado e motivado por dois problemas:

1. Dado um vetor $x \in \mathbb{R}^n$, o objetivo é encontrar um subespaço linear m -dimensional ($m < n$) que minimize a distância de x a esse subespaço. Esse problema surgiu na área de compressão de dados, cujo objetivo é representar todos os dados por um número reduzido de parâmetros.
2. Dado um vetor $x \in \mathbb{R}^n$, a idéia é encontrar um subespaço linear m -dimensional que armazene a máxima variância de x . Esse problema é associado à extração de características em que o objetivo é reduzir a dimensão dos dados preservando o máximo de informações.

Ambos os problemas são baseados nos autovetores e autovalores da matriz covariância dos dados. O objetivo é encontrar um conjunto de vetores ortonormais v_i que melhor descreve a distribuição dos dados de entrada. Seja C a matriz covariância estimada

$$C = XX^T = V\Lambda V^T \in \mathbb{R}^{n \times n} \quad (4.1)$$

em que $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ é a matriz diagonal contendo os n autovalores, v_i é o autovetor associado a λ_i e $V = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$ é a matriz dos autovetores, v_i , da matriz covariância de

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (4.2)$$

é a matriz de dados original, em que $x_{ij} \in \mathbb{R}$, $i = 1, \dots, n$ e $j = 1, \dots, m$.

O PCA determina uma transformação linear dos elementos de X tal que

$$y = V^T X \quad (4.3)$$

onde $y = [y_1, y_2, \dots, y_n]$ é o vetor das componentes principais.

4.2 O PCA no Reconhecimento de Face

Alguns dos métodos de reconhecimento de face concentram-se em determinar quais características da face são mais relevantes, tais como, olhos, nariz, orelhas e boca. O método das autofaces busca um conjunto de características que não depende das formas geométricas da face, utilizando toda a informação da representação facial.

Baseadas na teoria da informação, as autofaces buscam identificar um pequeno número de características que são relevantes para diferenciar uma face de outras faces. Essas características podem ser analisadas apenas com a variação dos valores assumidos pelos pixels, em um conjunto de imagens de faces.

Em linguagem matemática, esse processo pode ser obtido via componentes principais da distribuição de faces, ou seja, através dos autovetores da matriz de covariância de um conjunto de imagens de faces. Esses autovetores podem ser vistos como um conjunto de características que descrevem a variação entre tais imagens. Cada autovetor descreve a variação dos pixels em uma região diferente das imagens [51], ou seja, significa que cada um deles descreve a variação dos pixels associados a diferentes características faciais. Por serem autovetores associados a um conjunto de faces, as componentes principais são chamadas de autofaces.

4.2.1 Método da Autoface

O método das autofaces, proposto em [51], trata-se de uma abordagem holística, que considera o reconhecimento de face como um problema de reconhecimento $2-D$. As imagens de faces são projetadas em um espaço de características (“espaço de faces”) que melhor representa a variação entre faces conhecidas. Este espaço é definido pelas “autofaces”, que são os autovetores do conjunto de faces. O reconhecimento de uma face é realizado pela sua projeção no subespaço gerado pelas autofaces e então pela comparação da posição obtida com a posição de indivíduos conhecidos.

Basicamente, o método consiste em:

- Formar um espaço, P , descrito pelos principais autovetores da matriz de covariância de um determinado conjunto de faces.
- Projetar uma imagem teste (Γ_q) neste espaço, gerando um vetor de pesos w ,

tal que $w = P^T \Gamma_q$.

- Comparar o vetor de pesos w obtido da projeção, com os vetores obtidos da projeção de faces conhecidas.

Consideremos a imagem de face I como sendo uma matriz de dimensão $N_x \times N_y$, em que N_x é o número de linhas e N_y é o número de colunas. É possível representar essa matriz como um vetor de dimensão N , em que $N = N_x N_y$, e que na realidade representa a matriz I concatenada. A figura 4.1 ilustra essa concatenação.

$$\begin{array}{c}
 \text{Imagem} \\
 = \\
 \begin{bmatrix}
 x_{11} & x_{12} & \dots & x_{1N_y} \\
 x_{21} & x_{22} & \dots & x_{2N_y} \\
 \vdots & \vdots & \ddots & \vdots \\
 x_{N_x 1} & x_{N_x 2} & \dots & x_{N_x N_y}
 \end{bmatrix}
 \cong
 \begin{bmatrix}
 x_{11} \\
 x_{12} \\
 \vdots \\
 x_{1N_y} \\
 x_{21} \\
 x_{22} \\
 \vdots \\
 x_{2N_y} \\
 \vdots \\
 x_{N_x 1} \\
 x_{N_x 2} \\
 \vdots \\
 x_{N_x N_y}
 \end{bmatrix}
 \end{array}$$

Figura 4.1: Exemplo da concatenação de uma imagem 2D.

Fonte: Imagem elaborada pelo autor.

Se considerarmos que a imagem de uma face pode ser descrita como um vetor no espaço \mathbb{R}^N , então essa face será um ponto nesse espaço [51]. A análise da componente principal é usada para encontrar os vetores que melhor caracterizem a distribuição dos pontos nesse espaço. Esses vetores definem o subespaço de faces, chamado de espaço de faces. Cada vetor tem tamanho N e descreve uma imagem de tamanho $N_x \times N_y$, sendo uma combinação linear das faces originais. Dado que estes vetores são os autovetores da matriz covariância das imagens originais, e que, a visualização desses vetores têm a aparência semelhante a uma face, eles são chamados de “autofaces”.

Dessa forma, seja $\Gamma = [\Gamma_1, \Gamma_2, \dots, \Gamma_M]$ o conjunto das imagens de treinamento (imagens selecionadas a partir de um banco de dados podendo ser todo o banco ou apenas algumas imagens) e M a quantidade de imagens desse conjunto.

De acordo com [23], há muita informação redundante na face e ela poderia ser representada somente pela variância que ocorre em torno de uma média, isto é, uma imagem de face poderia ser representada com poucas componentes principais devido à sua redundância.

Assim, torna-se relevante estudar essas variâncias considerando uma média geral. Essa média geral é calculada pela média aritmética dos vetores do conjunto de treinamento e tem a mesma dimensão das imagens de treinamento. A face média (figura 4.2) é

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (4.4)$$

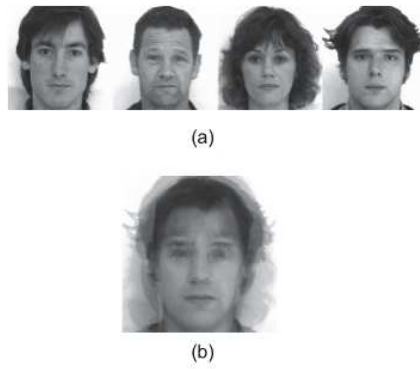


Figura 4.2: Exemplo das imagens de um conjunto de treinamento e a face média desse conjunto.

Cada face do conjunto de treinamento varia em relação à face média de acordo com

$$\Phi_i = \Gamma_i - \Psi \quad (4.5)$$

em que $i = 1, 2, \dots, M$.

O vetor Φ_i contém todas as variações de uma determinada face Γ_i em relação à face média Ψ . Dessa forma, se considerarmos que todas as faces serão subtraídas da face média, teremos uma nova matriz A que contém somente as variações de cada face em relação à face média [26].

$$A = [\Phi_1, \Phi_2, \dots, \Phi_M] \quad (4.6)$$

e, o cálculo da matriz covariância é dado por

$$C = AA^T. \quad (4.7)$$

O cálculo dos autovalores de C podem ser feitos baseados no SVD (*Singular Values Decomposition*). Caso o número de imagens $M \ll N$ então podemos usar o seguinte artifício algébrico para reduzir os cálculos ou mesmo, viabilizar o método.

Ao invés de resolver o sistema $(A^T A)x = b$ resolvemos $(AA^T)x = c$. Vejamos: seja λ_k e u_k os autovalores e autovetores de $C = AA^T$. Daí,

$$(AA^T)u_k = \lambda_k u_k. \quad (4.8)$$

Para determinar esses autovalores, λ_k e os correspondentes autovetores, u_k , podemos usar do seguinte artifício: seja μ_i e v_i , respectivamente, os autovalores e autovetores de $A^T A$, isto é

$$(A^T A)v_i = \mu_i v_i. \quad (4.9)$$

Multiplicando ambos os lados da equação 4.9 por A , tem-se

$$A(A^T A)v_i = A\mu_i v_i, \quad (4.10)$$

daí,

$$(AA^T)(Av_i) = \mu_i(Av_i). \quad (4.11)$$

Desta forma, observa-se da equação 4.8 e da equação 4.11, que os autovetores u_k de AA^T satisfazem $Av_i = \mu_i v_i$, ou seja, Av_i são os autovetores de C . Logo, conhecendo v_i conhecemos μ_i .

Desta forma, uma matriz de dimensão $M \times M$ é utilizada nos cálculos ao invés da matriz de dimensão $N \times N$, reduzindo substancialmente e de forma eficiente os cálculos computacionais. Uma vez que AA^T e $A^T A$ possuem os mesmos autovalores não-nulos [31], e que eles satisfazem o mesmo sistema de equações, é possível observar que, dentre os autovetores v_i , $(M - 1)$ correspondem aos autovetores, u_k , associados aos autovalores, λ_k , não-nulos.

Encontrados os autovetores que representam as imagens do conjunto de treinamento, cada autoface, p_i , será escrita como combinação linear desses autovetores com as faces do conjunto de imagens e podem ser pensadas como um conjunto de características que juntas identificam as variações entre as faces, e são escritas por

$$p_i = \sum_{j=1}^M v_{ji} \Phi_i. \quad (4.12)$$

O número máximo de autofaces que pode ser usada no processo de reconhecimento é igual ao número de imagens do conjunto de faces. Contudo, é possível usar um número menor de autofaces - aquelas que têm os maiores autovalores. Ao selecionar esse conjunto de autofaces, escolhemos aquelas que contribuem para a máxima variância dos dados. Uma ilustração das autofaces pode ser vista na figura 4.3.



Figura 4.3: Exemplos das autofaces.
Fonte: Elaboradas pelo autor.

Após determinar as autofaces, as imagens do conjunto de treinamento são projetadas no espaço de faces, efetuando-se a operação

$$\omega_k = p_k^T (\Gamma_k - \Psi), \quad (4.13)$$

para $k = 1, 2, \dots, M'$, em que M' é o número de autovetores associados aos autovalores mais significativos. A contribuição de cada autoface em relação às imagens do conjunto de treinamento é representada pelos pesos ω_k , estes organizados em um vetor Ω , tal que $\Omega = [\omega_1, \omega_2, \dots, \omega_{M'}]$.

4.2.2 Projeção de uma face teste no espaço gerado pelas autofaces

Os pesos encontrados pela equação 4.13 representam a contribuição de cada autoface em relação às imagens do conjunto de treinamento. Contudo, quando uma face que não pertence ao conjunto de treinamento for solicitada como consulta é preciso verificar a contribuição de cada autoface em relação à essa nova imagem.

Dessa forma, essa nova imagem (consulta) é subtraída da face média e projetada nas autofaces. A partir daí, os pesos resultantes da equação 4.13 são usados para procurar quais faces do conjunto de treinamento são similares, se houver alguma, e qual melhor descreve a nova imagem usada como consulta. Para determinar essa similaridade, o método PCA utiliza o cálculo da distância euclidiana.

Assim, seja Γ_{query} a face teste de dimensão $N = N_x N_y$. Calcula-se a diferença entre essa face e a face média, por

$$\Phi_{query} = \Gamma_{query} - \Psi, \quad (4.14)$$

em que Φ_{query} tem dimensão N . A projeção nas autofaces é dada por

$$\omega_{query} = p_k^T \Phi_{query}, \quad (4.15)$$

para $k = 1, 2, \dots, M'$.

O processo de similaridade é realizado por meio do cálculo da distância euclidiana entre os pesos da equação (4.13 e 4.15) pela equação

$$\varepsilon_k = \|(\omega_{query} - \omega_k)\|, \quad (4.16)$$

em que ω_{query} é o vetor de pesos resultantes da projeção da face consulta no espaço de face e ω_k é um vetor que representa a k -ésima face.

4.3 Considerações Finais

Neste capítulo foram apresentadas as bases teóricas para a compreensão da técnica estatística da análise das componentes principais (PCA), que é muito utilizada nas aplicações de caracterização e reconhecimento de faces, e as bases matemáticas do PCA.

5 Análise da Componente Independente

5.1 Introdução

Encontrar um meio de representar apropriadamente os dados originais é um problema importante e tem merecido a atenção de pesquisadores das mais diversas áreas. Pesquisas em redes neurais, estatística e processamento de sinais são alguns exemplos em que este problema tem tido destaque. Tal representação, geralmente, pode ser obtida como uma transformação linear dos dados originais. Assim, dada uma variável x (n -dimensional), o objetivo é encontrar uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definida por

$$f(x) = u, \quad (5.1)$$

que represente apropriadamente os dados. Na maioria dos casos, procura-se determinar essa representação por meio da transformação linear das variáveis observadas, x , isto é,

$$u = Wx, \quad (5.2)$$

em que W é a matriz da transformação f .

Vários métodos foram propostos a fim de encontrar transformações lineares apropriadas. Dentre esses métodos, temos por exemplo, a análise da componente principal, o fator de análise, a análise discriminante linear, a análise da componente independente, etc.

Recentemente, a análise da componente independente (ICA) vem chamando a atenção dos pesquisadores. O ICA tem como principal objetivo encontrar uma transformação em que as componentes u_i sejam estatisticamente tão independentes quanto possível uma das outras.

O ICA pode ser usado em um vasto campo de aplicações e entre elas podemos citar os sinais da atividade cerebral por meio de um eletro-encefalograma (EEG). O EEG consiste em

registros de potenciais elétricos em locais diferentes do escalpo cerebral, que são gerados a partir de componentes básicos da atividade cerebral. Neste caso, o ICA é usado para separar as componentes originais da atividade cerebral, em que apenas são conhecidas as misturas dessas componentes.

Atualmente, pesquisadores vêm adotando o ICA no reconhecimento de face, em duas arquiteturas diferentes: a arquitetura-I, que fornece um conjunto de imagens estatisticamente independentes, que podem ser vistas como um conjunto de imagens de características faciais independentes, e a arquitetura-II, cujo objetivo é encontrar coeficientes estatisticamente independentes para as imagens originais. Uma representação das arquiteturas ICA pode ser vista na figura 5.1.

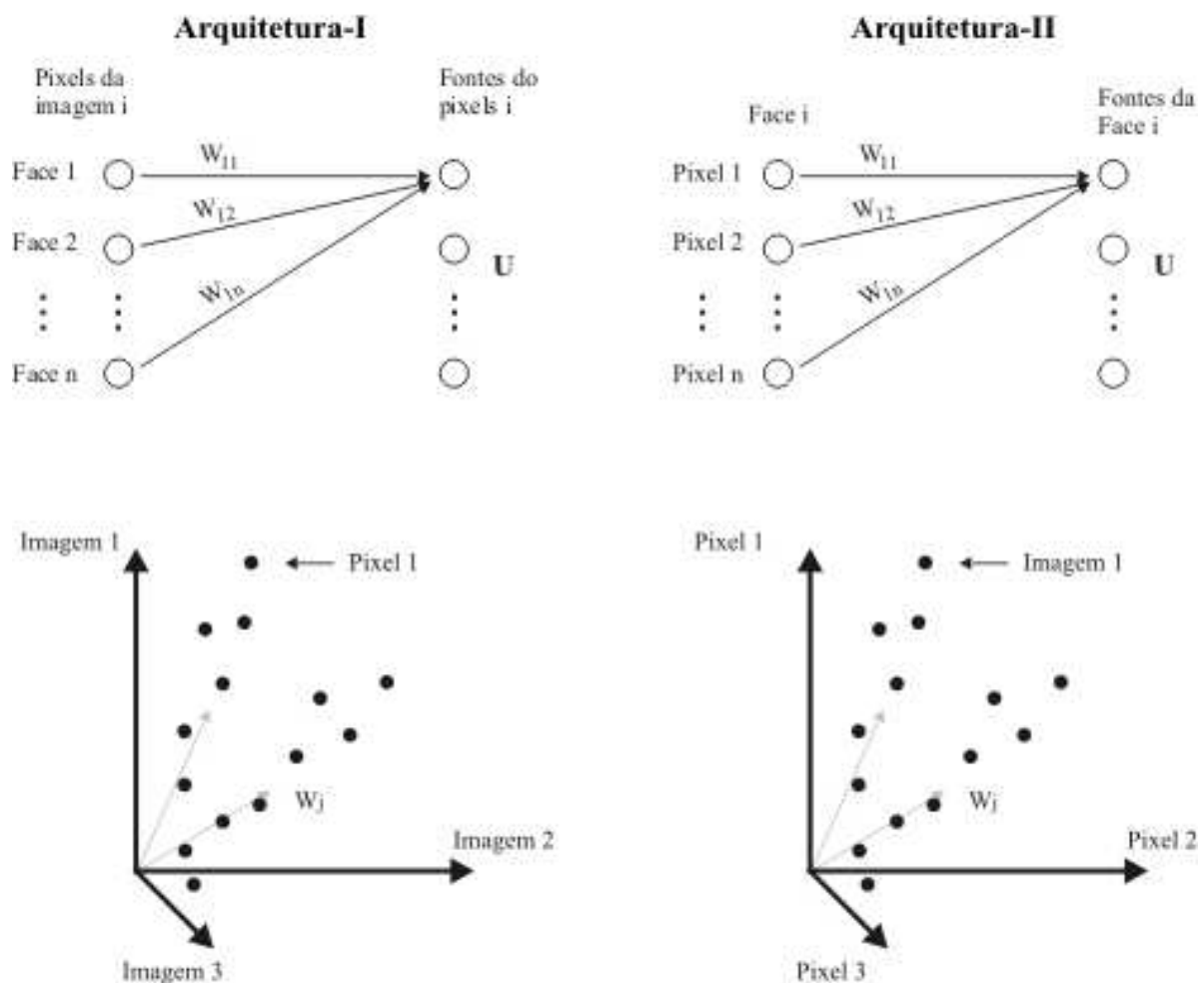


Figura 5.1: Arquiteturas usadas no ICA.
Fonte: Imagem obtida de [46].

5.2 Análise da Componente Independente

Alguns sistemas usados no reconhecimento de face são baseados nos mecanismos de aprendizado que são insensíveis às correlações das imagens de face. O método das autofaces [51] que é baseado na análise da componente principal (PCA), é um exemplo.

Métodos como o PCA são baseados nas estatísticas de segunda-ordem, geralmente usados, devido à sua simplicidade. Esses métodos assumem que todas as informações dos dados estão contidas na matriz covariância. Contudo, alguns conjuntos de dados, como por exemplo as faces, são diferenciadas por estatísticas de ordem superior. É importante investigar quais generalizações do PCA são sensíveis a essas estatísticas de ordem superior.

5.2.1 O modelo ICA

O ICA assume que um conjunto de variáveis observadas pode ser expresso como combinações lineares de variáveis desconhecidas e linearmente independentes. De alguma forma, o ICA separa os sinais gravados, em um conjunto de variáveis estatisticamente independentes.

Matematicamente, assume-se que m variáveis x_1, x_2, \dots, x_m são observadas, cada uma sendo uma combinação linear de m variáveis desconhecidas, u_1, u_2, \dots, u_m , de forma que

$$x_j = \sum_{i=1}^m a_{ji} u_i, \quad (5.3)$$

ou, na forma matricial, $x = Au$.

A equação 5.3 descreve como o dado gravado x é gerado pelas fontes u . As fontes u_i , que são chamadas de componentes independentes, não podem ser observadas diretamente. A matriz dos coeficientes A , chamada de matriz de misturas, também é desconhecida. O ICA fornece um modo de estimar tanto A quanto u . Para estimar a matriz A , o ICA possui como vantagem o fato de seus componentes serem estatisticamente independentes. A explicação de estimar A é a não-gaussianidade e, uma vez conhecida a matriz A e supondo que a mesma seja inversível, a sua inversa W pode ser aplicada para obter os componentes independentes: $u = Wx$.

Contudo, antes da aplicação do ICA, os dados precisam passar por uma fase de pré-processamento que consiste de dois passos: *sphering* e *whitening*. O passo *sphering* desloca os dados em torno de sua média, de tal forma que os dados resultantes tenham média zero.

No *whitening* o conjunto “*sphered*” de dados observados é transformado em um conjunto de dados não-correlacionados. A matriz de transformação ICA, aprendida com os dados *sphered* pode ser dada como $W_{ica} = WW_z$, em que W é a matriz aprendido calculada pelo ICA [44] e W_z a matriz *whitening*.

O PCA pode ser usado para executar essa transformação. Depois da fase de pré-processamento, a equação $x = Au$ assume a forma

$$x = E\{x\} + PAu \quad (5.4)$$

em que $E\{x\}$ é a esperança de x , P é uma matriz $n \times m$ obtida da aplicação do PCA ao dado “*sphered*” e m é o número de componentes principais.

A principal diferença entre o ICA e o PCA é que, o PCA assume que os componentes são não-correlacionados e o ICA assume que os componentes são independentes¹. O objetivo do PCA é encontrar os componentes de maior variância dos dados, enquanto que no ICA, os componentes independentes produzidos, fornecem um mecanismo de separação entre os dados que são assumidas serem independentes.

5.3 Arquiteturas ICA usadas no Reconhecimento de Faces

No contexto do reconhecimento de face, o uso do ICA foi proposto, primeiramente, por Bartlett et al. [44] e, posteriormente, por várias outras contribuições [15, 20, 37, 40, 53]. Em Bartlett et.al [44] duas arquiteturas diferentes são apresentadas para o reconhecimento de face. Na primeira, chamada de arquitetura-*I*, as imagens de faces são descritas como combinação linear de um conjunto desconhecido de componentes estatisticamente independentes e os componentes obtidos nessa arquitetura podem ser vistos como um conjunto de imagens de características faciais independentes (figura 5.4). Na segunda, chamada de arquitetura-*II*, as representações dos coeficientes (pesos) são estatisticamente independentes e, os componentes obtidos podem ser vistos como um conjunto de imagens com a aparência de uma face, assim como no PCA (figura 5.7).

Em contraste ao método PCA, em que a seleção das componentes principais é baseada no critério de energia (autovalores que representam as maiores variâncias), no ICA a seleção das componentes independentes não é evidente, uma vez que a energia dos componentes independentes não pode ser determinada.

¹Componentes independentes são sempre não-correlacionados, mas o inverso não é verdade [46].

5.3.1 Arquitetura-I

Na arquitetura-I, as faces originais, X , são descritas como combinação linear de um conjunto desconhecido de imagens base estatisticamente independentes, S , com os coeficientes dados por uma matriz de mistura A . A matriz de pesos W_I aprendida pelo ICA é usada para recuperar um conjunto de imagens base (componentes) independentes nas linhas de $U = [u_1, u_2, \dots, u_N]$ (figura 5.2).

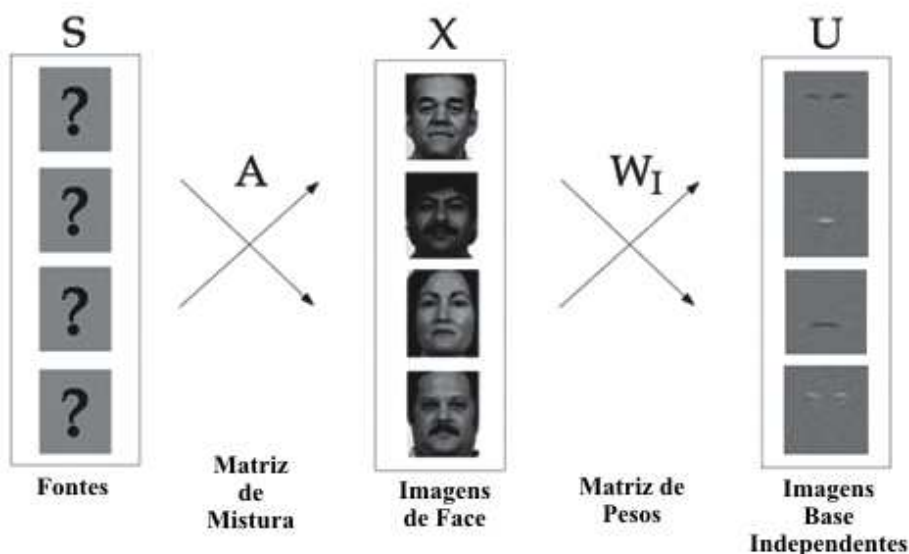


Figura 5.2: Arquitetura-I.
Fonte: Imagem obtida de [46].

Nesta arquitetura a matriz de entrada, X , é organizada de modo que as linhas representem imagens diferentes e as colunas representem pixels diferentes (figura 5.1 à esquerda). A projeção das imagens originais, X , no vetor de pesos, W_I aprendido pelo ICA, produz as imagens base independentes, U , tal que $U = W_I X$. A figura 5.3 mostra as imagens base independentes produzidas nessa arquitetura.

Figura 5.3: Representação das imagens bases independentes.
Fonte: Imagem obtida de [46].

Em Bartlett et.al [44] o PCA é aplicado primeiramente para projetar os dados em um sub-espço de dimensão m de modo a controlar o número de componentes independentes produzidos pelo ICA, permitindo-lhe criar sub-espços m -dimensionais. Em [20] também argumenta-se que a pré-aplicação do PCA melhora o desempenho do ICA [45, 47].

Seja P_m uma matriz de autovetores de um conjunto de faces obidos do PCA. Como as linhas da matriz de entrada são variáveis e as colunas são observações, o ICA é representado em P_m^T . Os m componentes independentes nas linhas de U são calculados como

$$U = W_I P_m^T. \quad (5.5)$$

Em seguida, é calculada a matriz dos coeficientes ICA, B , como combinação linear das componentes independentes, U , como descrito a seguir:

Considere R_m a matriz dos coeficientes PCA, tal que

$$R_m = X P_m \quad e \quad X = R_m P_m^T. \quad (5.6)$$

A partir da equação 5.5 e, supondo que W_I é inversível, obtém-se

$$P_m^T = W_I^{-1} U. \quad (5.7)$$

Então,

$$X = (R_m W_I^{-1}) U \quad (5.8)$$

$$= B U. \quad (5.9)$$

Cada linha de B contém os coeficientes da combinação linear das componentes estatisticamente independentes com as imagens de faces contidas nas linhas de X . Além disso, X é a reconstrução dos dados originais em relação ao erro quadrado-mínimo, como no PCA.

Quando uma nova imagem (imagem consulta), X_{teste} é solicitada, ela é projetada no sub-espço de face, $R_{test} = X_{test} P_m$, que é multiplicada por W_I produzindo um vetor de características ICA $B_{test} = R_{test} W_I^{-1}$. Esse vetor é comparado com os vetores características de cada indivíduo na base de dados de faces por meio da distância cosseno.

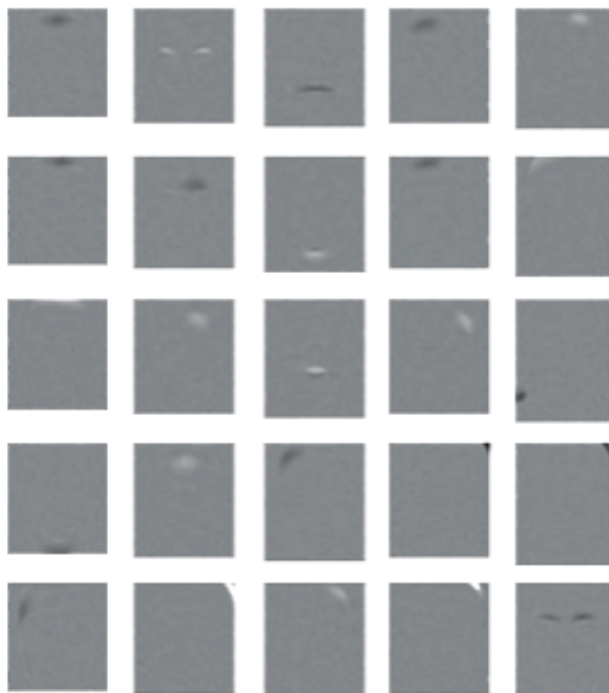


Figura 5.4: Componentes independentes do conjunto de imagens obtidas pela arquitetura-I.
 Fonte: Imagem obtida de [46].

5.3.2 Arquitetura-II

Enquanto as imagens base (componentes), u_i obtidas na arquitetura-I são estatisticamente independentes, os coeficientes, b_i que representam cada face não necessariamente são independentes [46]. O objetivo do ICA na arquitetura-II é encontrar uma representação em que os coeficientes usados para representar as faces sejam estatisticamente independentes (figura 5.5).

Nesta arquitetura a matriz de entrada, X , é organizada de modo que as linhas representem pixels diferentes e as colunas representem imagens diferentes (figura 5.1 à direita). As colunas da matriz $U = W_I X$ fornecem uma representação das imagens da matriz de entrada X . Cada coluna de U contém os coeficientes das imagens base em A , para reconstruir cada imagem de X (figura 5.6).

A representação de uma imagem teste, X_{teste} , é dada por $U_{teste} = W_I X_{teste}$.

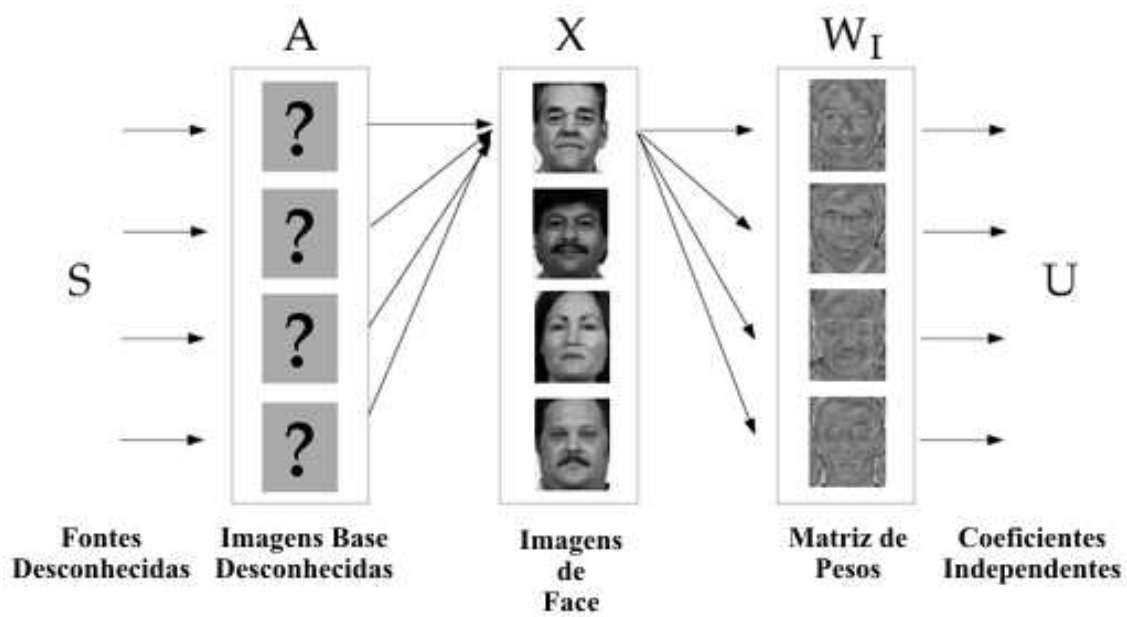


Figura 5.5: Arquitetura-II.

Fonte: Imagem obtida de [46].

Figura 5.6: Imagens bases desconhecidas.

Fonte: Imagem obtida de [46].

5.4 Resultados Experimentais

Nesta seção são descritos os resultados da avaliação das arquiteturas *I* e *II* do método da análise da componente independente, segundo [15]. As imagens usadas para avaliar o desempenho do ICA no reconhecimento de face, usada em [15], é da base de dados FERET [57], descrita no capítulo 3 seção 3.2.3. Os resultados obtidos foram comparados com o método PCA [15].

A base de dados FERET foi dividida em dois conjuntos: o conjunto consulta (*probe set*) e o conjunto de imagens (*gallery set*). Para cada imagem consulta, há uma única imagem da mesma pessoa na galeria. Se a primeira imagem recuperada a partir da galeria é a mesma pessoa usada na consulta, então a recuperação foi bem sucedida. Caso contrário, a recuperação terá fracassado. As imagens da base de dados FERET foram reduzidas para



Figura 5.7: Imagens da representação ICA.
Fonte: Imagem obtida de [46].

60×50 pixels para minimizar os custos computacionais.

5.4.1 Comparação entre os métodos ICA e PCA

Para comparar o desempenho da ICA com a PCA, um conjunto de treinamento foi montado usando 500 imagens da base de dados FERET. O desempenho dos métodos de reconhecimento de face foram avaliados com os vetores coeficientes b tendo o cosseno como medida de similaridade. Assim, os testes são executados com as distâncias cosseno e euclidiana como medida de similaridade. A tabela 5.1 ilustra a taxa de reconhecimento da PCA, ICA-I e ICA-II.

De acordo com a tabela 5.1, vemos que a arquitetura-II com a distância cosseno teve taxa de reconhecimento mais alta. Já com a distância euclidiana, observa-se que a arquitetura-I e a PCA tiveram resultados bem próximos. Podemos considerar que a arquitetura-II, ao usar a distância cosseno, é melhor que a PCA nos casos de reconhecimento de face.

Conjunto Consulta	ICA-I(%)		ICA-II(%)		PCA(%)	
	euclidiana	cosseno	euclidiana	cosseno	euclidiana	cosseno
<i>fc</i> (1195)	75.90	73.72	74.90	82.26	72.80	70.71
<i>fb</i> (194)	5.15	5.67	35.57	51.03	4.64	4.64
<i>duplicate I</i> (722)	32.96	36.15	37.81	48.48	33.24	35.32
<i>duplicate II</i> (234)	14.53	14.53	25.64	32.48	14.53	15.38
TOTAL (2345)	50.70	50.62	55.31	64.31	49.17	48.83

Tabela 5.1: Taxa de reconhecimento da PCA, ICA-I e ICA-II na base de dados FERET.
 Fonte: Imagem obtida de [15].

5.5 Considerações Finais

Segundo [15] comparar os métodos PCA e ICA não é uma tarefa fácil. Devem ser levados em conta as diferenças nas arquiteturas, as medidas de distância usadas, o tipo de problema enfrentado, como por exemplo, iluminação, expressão facial, dentre outros. No ICA sabe-se que a arquitetura-II dá uma taxa de reconhecimento maior quando usada a similaridade cosseno, enquanto que a distância euclidiana fornece próxima às taxas de reconhecimento da arquitetura-I e do PCA.

6 Análise Discriminante Linear

6.1 Introdução

A análise discriminante linear (LDA) ou análise discriminante de Fisher é uma técnica estatística que permite discriminar e classificar observações pertencentes a dois ou mais grupos mutuamente exclusivos definidos a priori, com base em um número de variáveis observáveis independentes. Essas variáveis observáveis são chamadas de “variáveis discriminantes”. Para isso é calculada uma “função discriminante”, que é uma função composta por índices, com pesos específicos.

A análise discriminante linear trata dos problemas relacionados com a separação de conjuntos distintos de observações (discriminação) e a alocação de novos objetos (classificação) em conjuntos previamente definidos. Essa técnica está inserida em um contexto mais amplo, a do reconhecimento de padrões, uma ciência que trata da classificação e descrição dos objetos.

A análise discriminante e a análise de classificação são técnicas multivariadas¹ interessadas, respectivamente, na separação de uma coleção de observações distintas e na alocação de novas observações em grupos previamente definidos. Apesar de estarem claramente interligadas, não devem ser confundidas. A análise discriminante se refere aos métodos de atribuição de classes a determinados conjuntos de dados. Já a classificação se refere à alocação de novos dados nos seus devidos grupos correspondentes.

As discriminações podem ser feitas através dos processos supervisionados que são utilizados quando se conhece o padrão (dados para treinamento) ou através dos processos não supervisionados, sendo estes recomendados quando não se tem um padrão conhecido.

O objetivo da análise discriminante é determinar um conjunto de coeficientes discrimi-

¹A análise multivariada tem como propósito medir, explicar e prever o grau de relação entre observações.

minantes para um conjunto de observações capaz de extrair a maior quantidade possível de informação quanto à classificação das observações nos grupos. Ela visa maximizar a variância entre-grupos em relação à variância intra-grupos, considerando que as observações estão previamente classificadas nos diversos grupos.

6.2 Análise Discriminante Linear no Reconhecimento de Face

A análise discriminante linear é aplicada no reconhecimento de face com o objetivo de maximizar a dispersão entre-classes, enquanto minimiza a dispersão intra-classe.

Considere uma imagem Γ_{ji} de tamanho N pixels² como a i -ésima imagem de uma classe de imagem C_j .

A face média da classe C_j , denotada por Ψ_j é obtida de

$$\Psi_j = \frac{1}{c_j} \sum_{i=1}^{c_j} \Gamma_{ji}, \quad (6.1)$$

em que c_j é a quantidade de faces pertencentes a cada uma das k classes, C_1, C_2, \dots, C_k , onde $C_j = \Gamma_{j1}, \dots, \Gamma_{jc_j}$.

A matriz dispersão intra-classe S_w , dada pela equação

$$S_w = \sum_{j=1}^k \sum_{i=1}^{c_j} (\Gamma_{ji} - \Psi_j)(\Gamma_{ji} - \Psi_j)^T, \quad (6.2)$$

representa a dispersão das imagens de face Γ_{ji} em relação à face média de sua classe, Ψ_j .

A matriz dispersão entre-classe S_b , definida por

$$S_b = \sum_{j=1}^k c_j (\Psi_j - \Psi)(\Psi_j - \Psi)^T, \quad (6.3)$$

representa a dispersão das faces médias de cada classe em relação à média total, Ψ dada por

$$\Psi = \frac{1}{\sum_{j=1}^k c_j} \sum_{j=1}^k \sum_{i=1}^{c_j} \Gamma_{ji}. \quad (6.4)$$

²Composta de maneira equivalente à descrita no capítulo 4 figura 4.1

O LDA tenta encontrar uma matriz de projeção P_{opt} que maximize a razão do determinante da matriz dispersão intra-classe com o determinante da matriz dispersão entre-classes (critério de Fisher) dada por

$$P_{opt} = \arg \max_P \frac{|P^T S_b P|}{|P^T S_w P|}. \quad (6.5)$$

De acordo com [54] S_w poderia ser singular. De modo a superar esse fato, diferentes métodos foram propostos [25, 35, 41, 54, 61, 72]. Dentre esses métodos, o mais usado é o Fisherface [25, 54, 72] que combina o PCA com o LDA.

O Fisherface projeta o conjunto de imagens em um espaço dimensional menor tal que a matriz dispersão intra-classe S_w resultante é não-singular. Isso é feito com o uso do PCA para reduzir a dimensionalidade projetando as faces originais em um espaço dimensional menor. Então, o LDA é aplicado para procurar a melhor característica discriminante linear naquele subespaço PCA. Mais formalmente, a matriz de projeção P_{opt} pode ser calculada como

$$P_{opt} = P_{lda} P_{pca}, \quad (6.6)$$

onde

$$P_{pca} = \arg \max_P |P^T S_T P|, \quad (6.7)$$

para $S_T = S_w + S_b$ e P_{lda} é dada por

$$P_{lda} = \arg \max_P \frac{|P^T P_{pca}^T S_b P_{pca} P|}{|P^T P_{pca}^T S_w P_{pca} P|} \quad (6.8)$$

6.3 Resultados Experimentais

Os resultados provenientes dos estudos de Belhumeur et.al. [54] que desenvolveram testes para determinar o desempenho do LDA sob diferentes condições de iluminação e expressão facial são apresentados nesta seção. A figura 6.1 ilustra imagens de uma pessoa. A primeira imagem foi obtida sob condição de iluminação ambiente, com expressão facial neutra e com a pessoa usando óculos, na segunda imagem, os óculos foram removidos. As imagens na posição de 3 a 5 foram obtidas sob três condições de iluminação diferentes e expressão facial neutra. As últimas 5 imagens foram obtidas sob iluminação ambiente com diferentes expressões faciais (feliz, triste, careta, sonolento e surpreso). As imagens foram manualmente centralizadas e cortadas em duas escalas diferentes: as imagens maiores incluem a face toda

e parte do fundo enquanto que as imagens cortadas incluem estruturas internas, tais como, testa, sobrancelhas, nariz, boca e queixo.



Figura 6.1: Exemplos de uma pessoa da base de dados YaleA e suas variações.
Fonte: Imagens obtidas de [3].

Nesse teste, as taxas de erros foram determinadas pela estratégia “*leaving one-out*” [54]: para classificar uma imagem de uma pessoa, aquela imagem foi retirada do conjunto de dados e a matriz P foi calculada. Todas as imagens na base de dados, excluindo a imagem teste, foram então projetadas no espaço reduzido e usadas na classificação. Para esse teste, cada pessoa no conjunto de treinamento é representada pela projeção de dez imagens, exceto pela imagem teste que é representada por somente nove.

Em geral, o desempenho do método das autofaces varia com o número de componentes principais. Assim, antes de comparar o método LDA com o método PCA, primeiramente é executado um experimento para determinar o número de componentes principais que resultam nas menores taxas de erros. A figura 6.2 ilustra a taxa de erro em relação à quantidade de componentes principais selecionadas para o conjunto com faces cortado, quando as três componentes principais iniciais foram retidas e quando elas foram colocadas.

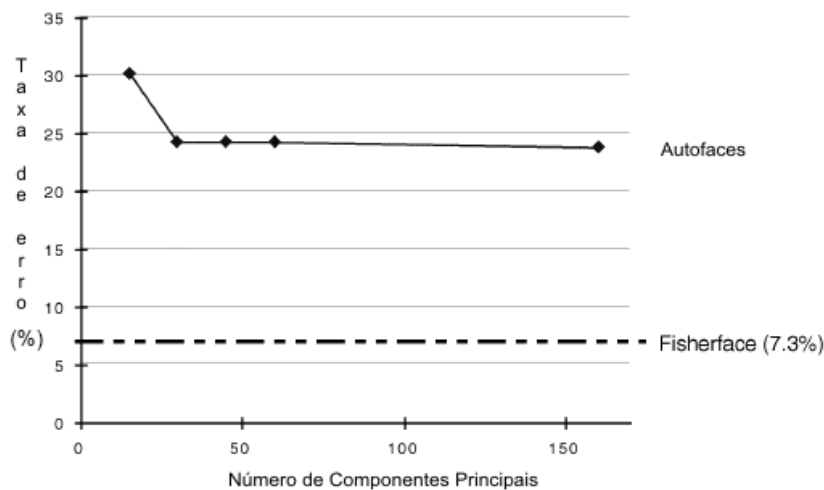


Figura 6.2: A variação no desempenho do método das autofaces.
Fonte: Imagem obtida de [54].

O desempenho dos algoritmos é evidente na figura 6.3. O método Fisherface tem taxas de erros que foram melhores que os outros métodos. Parece que o método Fisherface seleciona o conjunto de projeções que executa apropriadamente sobre variações de iluminação, expressão facial e presença de óculos.

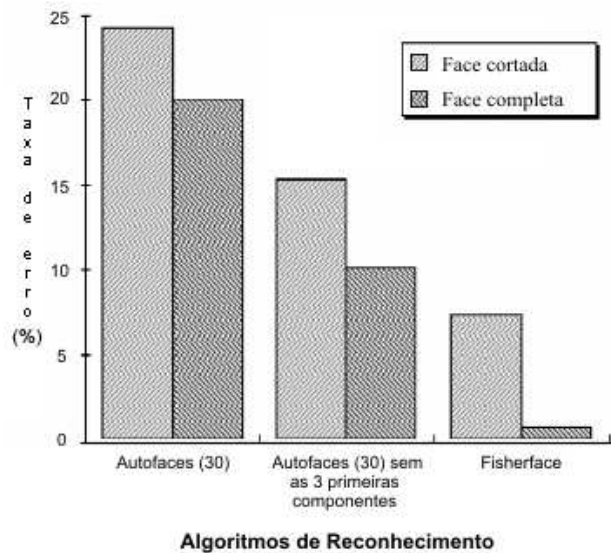


Figura 6.3: Desempenho dos algoritmos quando aplicados à base de dados YaleA que contém variações na expressão facial e iluminação.

Fonte: Imagem obtida de [54].

<i>Leaving-One-Out</i> da Base de dados YaleA			
Métodos	Espaço Reduzido	Taxa de erro (%)	
		Face cortada	Face Completa
Autofaces	30	24.4	19.4
Autofaces sem os 3 primeiros componentes	30	15.3	10.8
Fisherface	15	7.3	0.6

Tabela 6.1: Desempenho relativo dos algoritmos quando aplicados à base de dados YaleA.

Fonte: Dados obtidos de [54].

Todos os algoritmos foram melhores em imagens da face completa. Quando o método é treinado na face inteira, os pixels correspondentes aos contornos da face, que são perdidos quando a face é recortada, são selecionados como boas características na discriminação dos indivíduos, isto é, a forma global da face é uma característica poderosa na sua identificação. Contudo, esperou-se que as taxas de reconhecimento fossem muito menores para as imagens

de face toda se o fundo ou o estilo do cabelo fossem mudados, sendo que foi pior até mesmo que as imagens recortadas.

6.3.1 Reconhecimento com óculos

Nesse experimento, o método Fisherface foi comparado ao método das autofaces com imagens de faces que foram divididas em duas classes: usando óculos e não usando óculos. Com somente duas classes, as imagens podem ser projetadas para uma linha (figura 6.4) usando o método Fisherface. Com o PCA, a escolha das autofaces é independente da definição da classe.

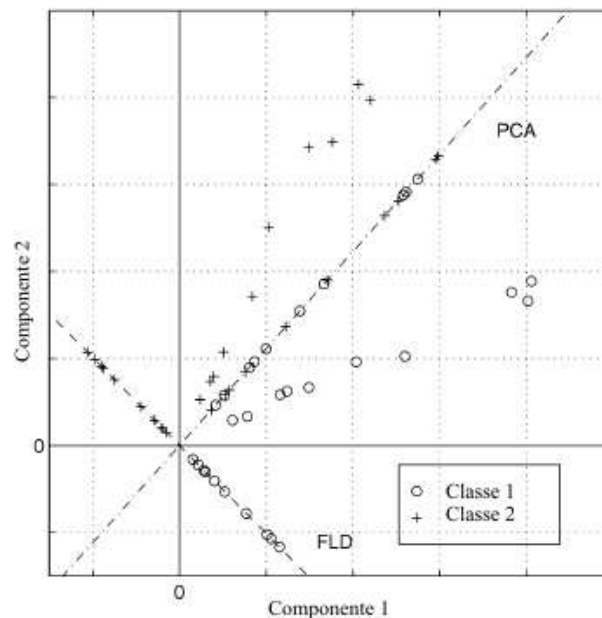


Figura 6.4: Comparação do PCA com o Fisherface em um problema de duas classes.
Fonte: Imagem obtida em [54].

As taxas de reconhecimento foram obtidas por classificar as imagens de cada pessoa. Todas as imagens daquela pessoa foram removidas da base de dados antes da matriz de projeção P ser calculada. A tabela 6.2 apresenta as taxas de erro para os dois métodos.

A taxa de reconhecimento pelo PCA já era prevista, pois, como na maioria dos casos, ela classificou ambas as imagens (com e sem óculos) na mesma classe. Por outro lado, o método Fisherface pode ser visto por derivar um modelo que é apropriado para encontrar óculos e ignorar outras características da face. Essa suposição é sustentada por observar o

Reconhecimento de óculos		
Métodos	Espaço Reduzido	Taxa de erro (%)
PCA	10	52.6
Fisherface	1	5.3

Tabela 6.2: Reconhecimento comparativo das taxas de erro para o reconhecimento de óculos/sem óculos com a base de dados YaleA [54].

Fonte: Dados obtidos de [54].

Fisherface na figura 6.5 conforme a matriz de projeção P . Naturalmente, espera-se que a mesma técnica poderia ser aplicada para identificar expressões faciais em que o conjunto de imagens de treinamento é dividido em classes baseadas na expressão facial [54].



Figura 6.5: A imagem da esquerda é uma imagem da base de dados Yale de uma pessoa usando óculos. A imagem da direita é o Fisherface usado para determinar se uma pessoa está usando óculos.

Fonte: Imagem obtida em [54].

6.4 Considerações finais

Neste capítulo encontra-se descrito o método LDA e sua utilização no reconhecimento de face. Também foi apresentada uma análise do desempenho comparando o LDA com o PCA, com base nos experimentos realizados por [54].

Os resultados sugerem que: (1) todos os métodos são bem executados quando uma imagem no conjunto teste é similar à uma imagem no conjunto de treinamento; (2) a remoção das três primeiras componentes principais não melhora o desempenho do PCA em relação às variações na iluminação; (3) os métodos parecem executar melhor o reconhecimento quando a face completa é usada; e (4) nos casos em que as faces são divididas em duas classes (com e sem óculos) o método Fisherface foi consideravelmente melhor que o PCA.

7 Resultados Experimentais

7.1 Introdução

Neste capítulo são apresentados os resultados experimentais do método das autofaces, considerando as variações na iluminação, expressão facial e pose, considerando o PCA com diferentes quantidades de elementos.

- **Experimento 1:** Mudanças da expressão facial e perturbações na face (por exemplo, faces com ruído, recortadas, com oclusões, etc.);
- **Experimento 2:** Variações da iluminação;
- **Experimento 3:** Variações da pose.

Estes experimentos foram motivados pelo fato de que é possível encontrar tais situações na vida cotidiana. Por exemplo, uma face adquirida de um sistema de segurança pode conter qualquer um dos problemas citados acima. Cada experimento é abordado e analisado, individualmente, na seção 7.3. Também é descrita a metodologia da avaliação de sistemas de recuperação de informação usado na análise dos resultados. Os experimentos foram divididos em três partes, de acordo com o tipo de problema abordado.

7.2 Avaliação de um Sistema de Recuperação de Informação

O desenvolvimento de técnicas de recuperação de informações realmente eficientes tem sido o centro das atenções dos estudos na área de Recuperação de Informação. Ao longo dos anos, várias métricas foram propostas a fim de testar e validar a eficiência destas técnicas. A revocação (Recall) e a Precisão (Precision) são as medidas de eficiência mais conhecidas e utilizadas na área [33].

O objetivo de se usar medidas de eficiência é indicar se o sistema realmente consegue recuperar informações relevantes para o usuário (uma boa revocação), ao mesmo tempo em que consegue excluir os itens irrelevantes (uma boa precisão).

7.2.1 Revocação e Precisão

A revocação (*Rev*) serve para indicar a proporção de documentos relevantes, recuperados em resposta a uma consulta do usuário, medida pela razão entre o total de documentos relevantes recuperados, N_{rr} , e o número total de documentos relevantes existentes, N_{re} , isto é,

$$Rev = \frac{N_{rr}}{N_{re}} \quad (7.1)$$

A precisão, (*Pre*), indica a habilidade do sistema em recuperar os documentos que são relevantes ao usuário. O cálculo da precisão é encontrado pela razão do total de documentos relevantes recuperados, N_{rr} , e o total de documentos recuperados na consulta, N_{tr} , isto é,

$$Pre = \frac{N_{rr}}{N_{tr}} \quad (7.2)$$

A figura 7.1 apresenta as possíveis situações entre a revocação e a precisão, em que os documentos recuperados são aqueles que se encontram no interior dos círculos.

Em algumas avaliações, a precisão é medida com um número fixo de documentos recuperados: “precisão em 25”, isto é, mede a precisão de um documento na posição 25 da lista de documentos recuperados. Em outros casos, a precisão e a revocação são comparadas mutuamente e visualizadas graficamente, como a precisão “11pt”, calculada a cada 10% de revocação, em revocação 20%, e assim por diante até 100% de revocação, e assume-se que todos os documentos relevantes foram recuperados.

O cálculo da precisão e revocação é feito percorrendo o conjunto de documentos recuperados (rank) e ao encontrar um documento relevante são calculados tanto a precisão como a revocação referente à esse relevante encontrado. Desta forma é possível, depois de encontrados e calculados a precisão e revocação para cada documento relevante, o gráfico “11pt” é desenhado.

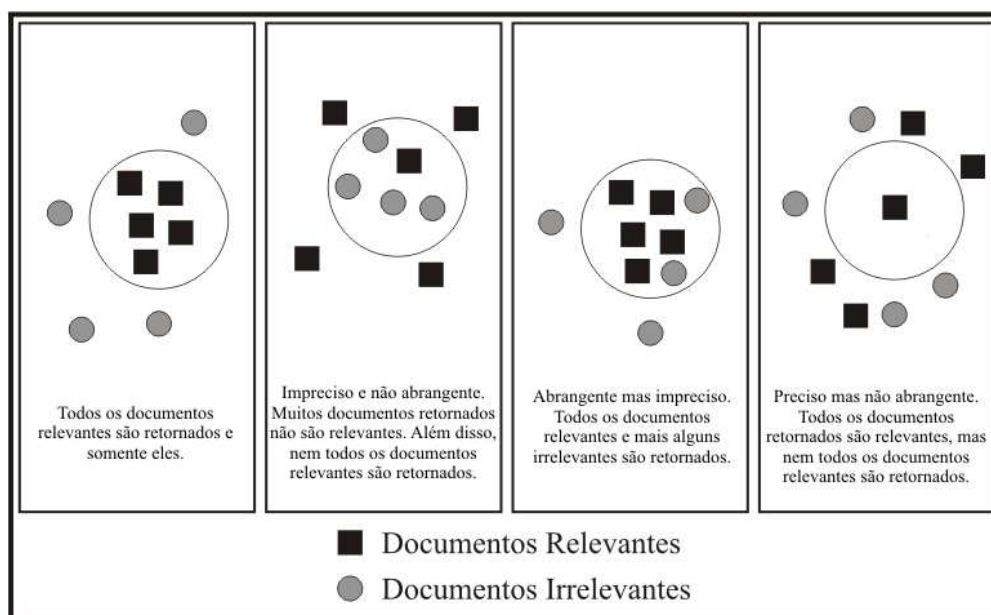


Figura 7.1: Relação entre a revocação e a precisão.
 Fonte: Imagem adaptada de [33].

7.3 Resultados Experimentais das autofaces

Esta seção apresenta os resultados obtidos usando o método das autofaces nos problemas envolvendo perturbações na face e mudanças da expressão facial, variações da iluminação e variações da pose. Tais problemas são enfocados ao explorarmos a quantidade de autofaces que deve ou pode ser selecionada bem como quais autofaces escolher, com o propósito de obter bons resultados de recuperação.

7.3.1 Seleção das autofaces

As autofaces foram geradas a partir de todas as imagens da base de dados YaleA [3] (descrita no capítulo 3), isto é, as 165 imagens de face. Dessa forma, 165 autofaces foram criadas e a escolha da quantidade e quais autofaces escolher foram organizadas em 4 grupos, de modo que:

- **1º grupo:** nesse grupo são consideradas 90 autofaces. É investigado o caso do sistema gerado pelas autofaces referentes aos 90 maiores autovalores e pelos 90 menores.
- **2º grupo:** neste 2º grupo é considerado o sistema gerado por 30 autofaces. Dois casos são também explorados na quantidade e quais autofaces escolher. O primeiro caso

refere-se aos 30 maiores autovalores e o segundo caso aos 30 autovalores ocupantes das posições: $1^a, 2^a, 3^a, 4^a, 5^a, 6^a, 7^a, 8^a, 9^a, 10^a, 15^a, 16^a, 17^a, 18^a, 19^a, 25^a, 26^a, 27^a, 28^a, 29^a, 50^a, 51^a, 52^a, 53^a$ e 104^a .

- **3º grupo:** nesse grupo são consideradas 12 autofaces. É investigado o caso do sistema gerado pelas autofaces referentes aos 12 maiores autovalores e pelos 12 autovalores ocupantes das posições: $1^a, 2^a, 3^a, 4^a, 5^a, 6^a, 40^a, 41^a, 50^a, 51^a, 90^a$ e 91^a .
- **4º grupo:** o 4º grupo considera o sistema gerado por 6 autofaces. Dois casos são explorados nesse grupo. O primeiro caso refere-se aos 6 maiores autovalores e o segundo caso aos 6 autovalores ocupantes das posições: $1^a, 2^a, 82^a, 83^a, 137^a$ e 138^a .

Em todos os casos de seleção das autofaces são escolhidas aquelas associadas aos autovalores de maior variância, os que encontram-se em posições “do meio” do conjunto ordenado de autovalores e aqueles de menor variância (encontram-se mais no final do conjunto ordenado dos autovalores).

7.3.2 Etapa de geração das imagens consultas

Nesta sub-seção é descrito como e quais faces foram usadas para gerar as faces consultas. A figura 7.2 ilustra as modificações realizadas nas imagens selecionadas. Com exceção da imagem (*n*) (figura 7.2), todas as outras faces pertencem à base de dados YaleA [3].

As imagens da figura 7.2 de (*a*) até (*i*) foram geradas adicionando ruídos e danificações (riscos, recortes, inversão, estreitamento da face, etc.) e, tem como objetivo verificar o desempenho do método das autofaces, utilizando essas imagens como consulta.

Nas imagens da figura 7.2 (*j*), (*l*) e (*m*) foram inseridos “disfarces” (perucas) visando verificar o desempenho do método das autofaces quando as informações relativas ao cabelo são modificadas.

A imagem da figura 7.2 (*n*) foi gerada usando uma imagem que não pertence à base de dados, em que foram inseridas as características (olhos, nariz e boca) de uma face pertencente à base de dados. Este experimento tem o objetivo de investigar o desempenho do PCA quando a consulta contém apenas as características internas da face (olho, nariz e boca) de alguma face da base de dados. Uma utilização disto é, por exemplo, se uma mulher se disfarçar com lenços e véus e apenas essas características são visíveis.

A figura 7.2 (o) é de uma face pertencente à base de dados cujas características internas (olhos, nariz e boca) são de outra face, também da base de dados.

Na figura 7.2 (p) duas faces de pessoas distintas presentes na base de dados foram usadas para gerar a imagem consulta, usando a metade superior e inferior de cada face. A figura 7.3 ilustra essa modificação.

Tanto a imagem da figura 7.2 (o) como a da figura 7.2 (p)) visam verificar o que mais influencia as autofaces na recuperação de faces, se as características faciais ou se o formato da face.

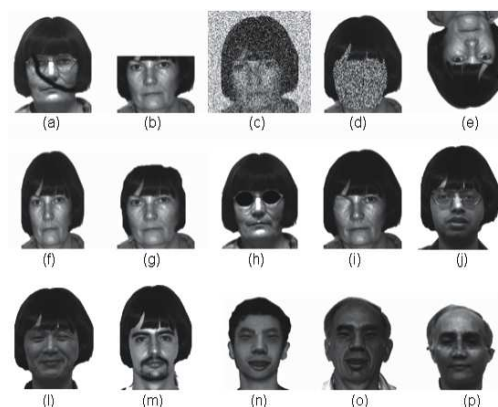


Figura 7.2: Imagens usadas como consulta no processo de recuperação de faces.

Fonte: Imagens elaboradas pelo autor.



Figura 7.3: Face montada a partir de duas faces de pessoas distintas.

Fonte: Imagem elaborada pelo autor.

7.3.3 Conjunto de treinamento e conjunto ideal

O conjunto de treinamento é formado por todas as imagens de uma base de dados que é utilizada no processo de obtenção das autofaces. Essas imagens podem ser selecionadas manualmente (como no caso deste trabalho) de acordo com a necessidade de quem está realizando os experimentos. O objetivo de usar apenas algumas imagens do banco de dados ao invés de todas está na minimização do custo computacional. Considere a base de dados YaleA para ilustrarmos um exemplo. Essa base possui 165 imagens de 15 pessoas em 11 condições

diferentes. Ao invés de usar todas as imagens são selecionadas apenas uma imagem de cada pessoa. Essas imagens formam o conjunto de treinamento. A ilustração desse conjunto pode ser visualizada na figura 7.4.

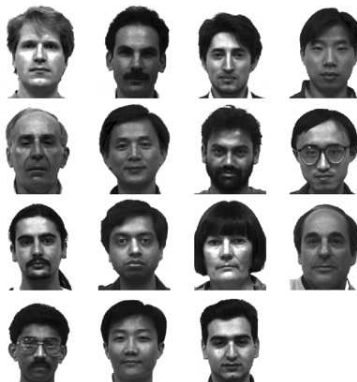


Figura 7.4: Imagens usadas para formar o conjunto de treinamento.

Fonte: Imagens elaboradas pelo autor.

O conjunto ideal é relacionado às imagens de uma determinada pessoa. Por exemplo, imagine a imagem de face de uma pessoa usada como consulta (imagem contornada em laranja na figura 7.5), em que existe, em um determinado banco de dados (neste caso as imagens da base de dados YaleA são usadas para ilustrar o exemplo), outras imagens desta mesma pessoa (porém em outras poses, condições de iluminação, etc.). O conjunto ideal nesse caso seria todas as imagens do banco de dados referentes à imagem de face da pessoa usada na consulta. A figura 7.5 ilustra uma imagem consulta e seu conjunto ideal.



Figura 7.5: Imagens usadas como conjunto ideal.

Fonte: Imagens elaboradas pelo autor.

7.3.4 Experimento 1: Problemas envolvendo perturbações na face e mudanças da expressão facial

Neste experimento o método das autofaces é avaliado quando a face consulta é modificada por meio de oclusões, disfarces ou até mesmo manipulações gráficas.

Os experimentos são divididos em dois grupos:

- **Primeiro grupo:** o PCA é executado em dois conjuntos de treinamento diferentes. O primeiro conjunto contém 1 imagem de face de cada pessoa da base e, são geradas 14 autofaces, que são usadas no reconhecimento; o segundo conjunto é formado por todas as 165 imagens de faces da base de dados e geradas 165 autofaces, que são selecionadas utilizando os 30 maiores autovalores.
- **Segundo grupo:** As autofaces são geradas sob as 165 imagens da base de dados e selecionadas de acordo com as configurações dos autovalores descritas na seção 7.3.1.

Os gráficos da precisão apresentam o desempenho do PCA em relação a cada face consulta da figura 7.2 com o uso de 14 autofaces e de 30 autofaces associadas aos maiores autovalores.

As faces consultas com oclusão (figura 7.2(a)), com ruído (figura 7.2(c, d)), com óculos escuros (figura 7.2(h)) e sem um dos olhos (figura 7.2(i)), recuperaram as imagens de faces referentes à pessoa pesquisada nas primeiras posições do conjunto resposta alcançando, portanto, precisão de 100%. O gráfico da figura 7.6-a representa a média para todas essas consultas. A precisão em relação a face consulta (figura 7.2(p)) é mostrada pelo gráfico da figura 7.6-b.

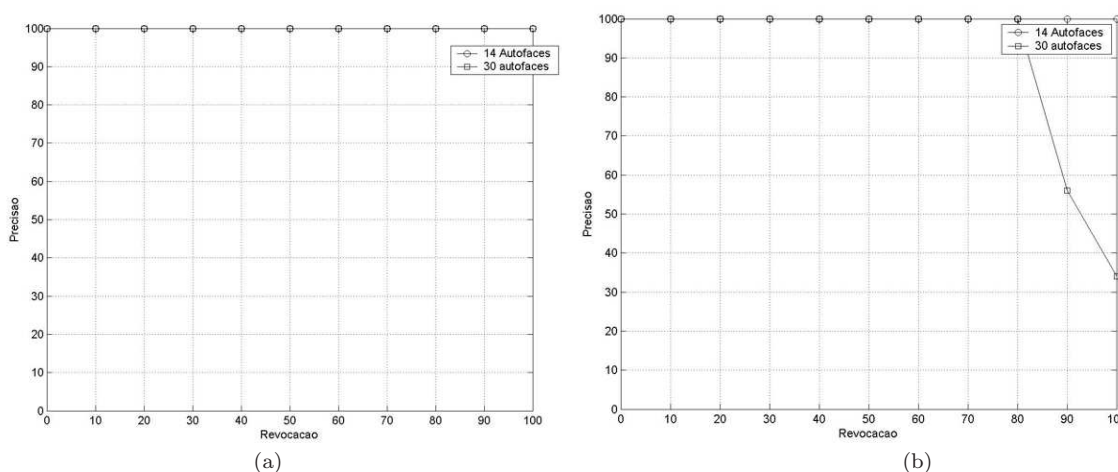


Figura 7.6: Precisão: imagens com oclusão, ruído, com óculos escuros e sem um dos olhos e uma face montada a partir de duas faces.

Fonte: Imagens elaboradas pelo autor.

No teste em que a face (o) (figura 7.2) é usada como consulta, o reconhecimento com o uso das 14 autofaces é menor do que com 30 autofaces (figura 7.7-a). O mesmo ocorreu

com a face consulta da figura 7.2 (n); contudo, a precisão aumentou quando o número de autofaces diminuiu (figura 7.7-b).

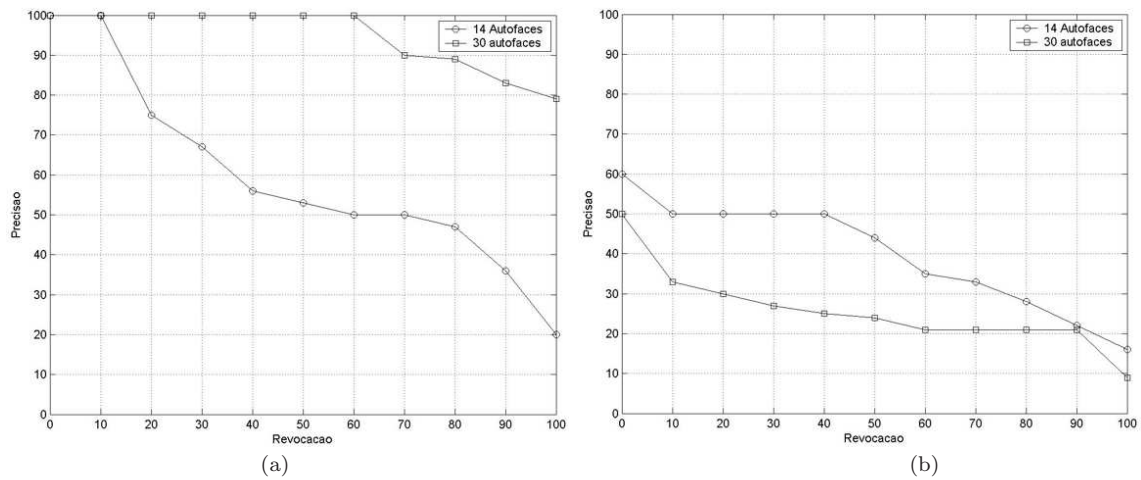
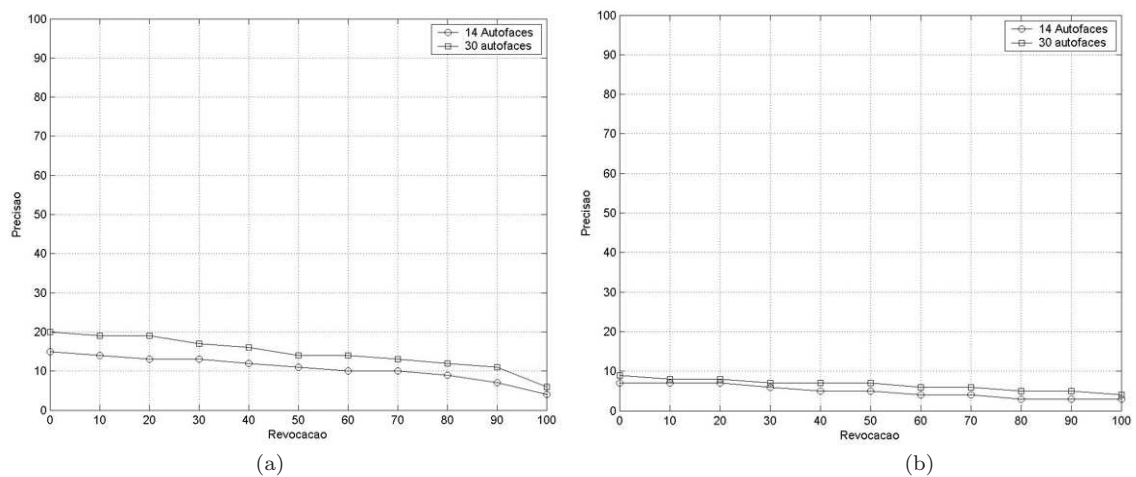


Figura 7.7: Gráfico da imagem consulta com as características internas modificadas.

Fonte: Imagens elaboradas pelo autor.

Contudo, verifica-se que algumas manipulações na imagem prejudicam o processo de reconhecimento. Os gráficos da figura 7.8 mostram a precisão em relação as imagens da figura 7.2(b, e, f, g, j, l, m), respectivamente.



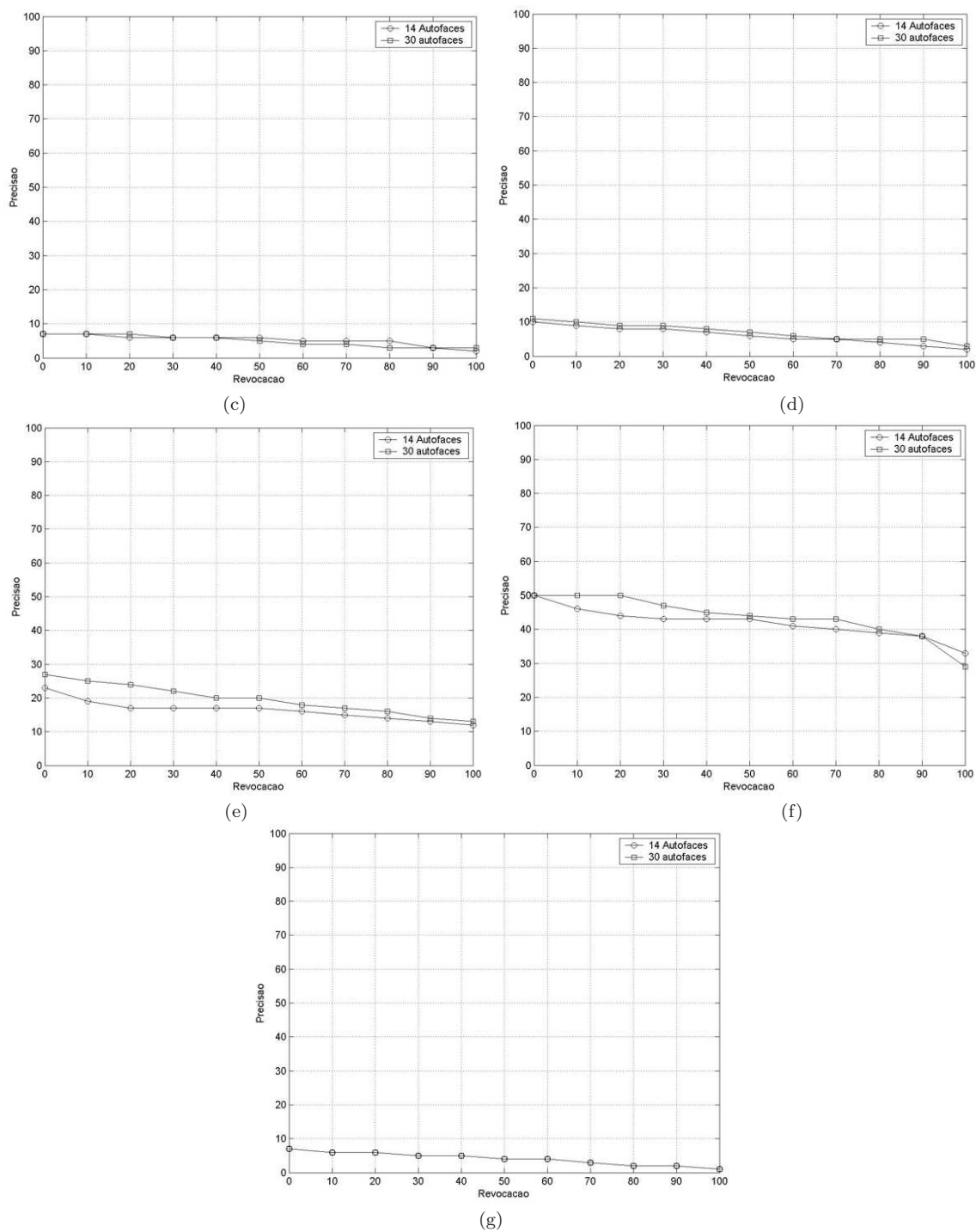


Figura 7.8: Gráfico da imagem consulta com metade da cabeça cortada, invertida, “estrita” e com “perucas”.
 Fonte: Gráficos elaborados pelo autor.

Com base nos resultados foi verificado que ambas as quantidades de autofaces (14 e 30) têm precisão muito similares. Isso leva-nos a crer que o conjunto de treinamento não necessariamente precisa conter todas as variações possíveis de uma mesma pessoa. O uso das imagens mais significativas (que pode variar de acordo com o idealizador do experimento) é suficiente para gerar autofaces que representem toda a base de dados, tendo como vantagem a redução dos cálculos computacionais e mais velocidade no processo de reconhecimento.

7.3.4.1 Resultados considerando o PCA construído com todas as imagens do banco de dados

Neste experimento é investigada a quantidade de autovalores que devem ser selecionadas para se obter um bom resultado de recuperação, bem como quais autovalores escolher (descrito na seção 7.3). Neste módulo, 9 das 15 imagens do experimento anterior são usadas na consulta (as imagens com perturbações na face e alteração das características internas). As imagens selecionadas são mostradas na figura 7.9.

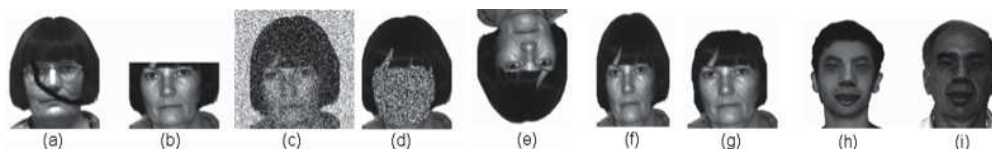


Figura 7.9: Imagens consultas usadas no processo de recuperação analisando a influência e a quantidade dos autovalores.

Fonte: Imagens elaboradas pelo autor.

O desempenho do PCA é analisado para cada imagem consulta por meio do gráfico da precisão e revocação. Para cada conjunto selecionado de autofaces um gráfico é mostrado. No caso em que a seleção das autofaces não modificam os resultados, apenas um gráfico é mostrado.

O gráfico 7.10 em (a) representa as consultas usando todos os conjuntos de autofaces com exceção das 90 menores autofaces, mostrado no gráfico 7.10 em (b).

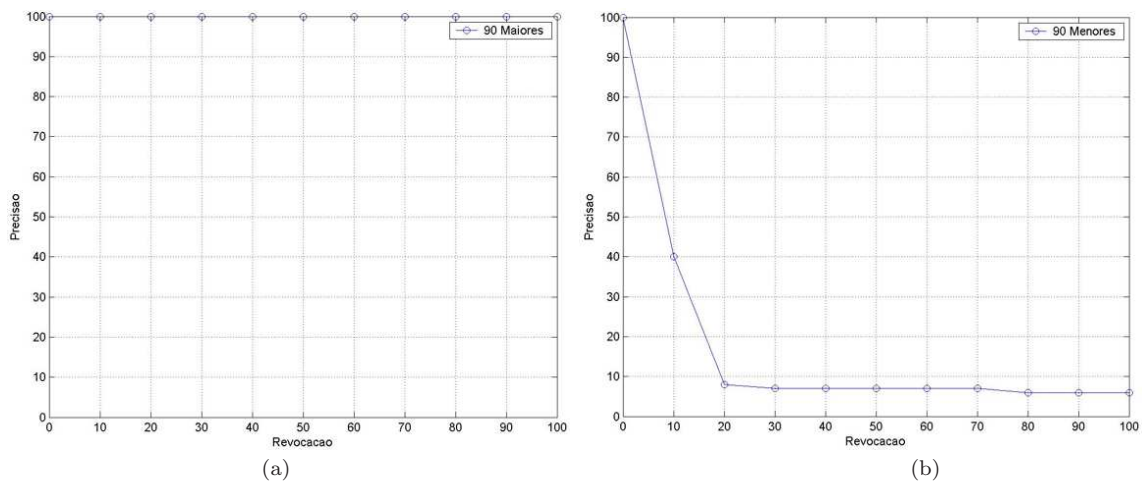
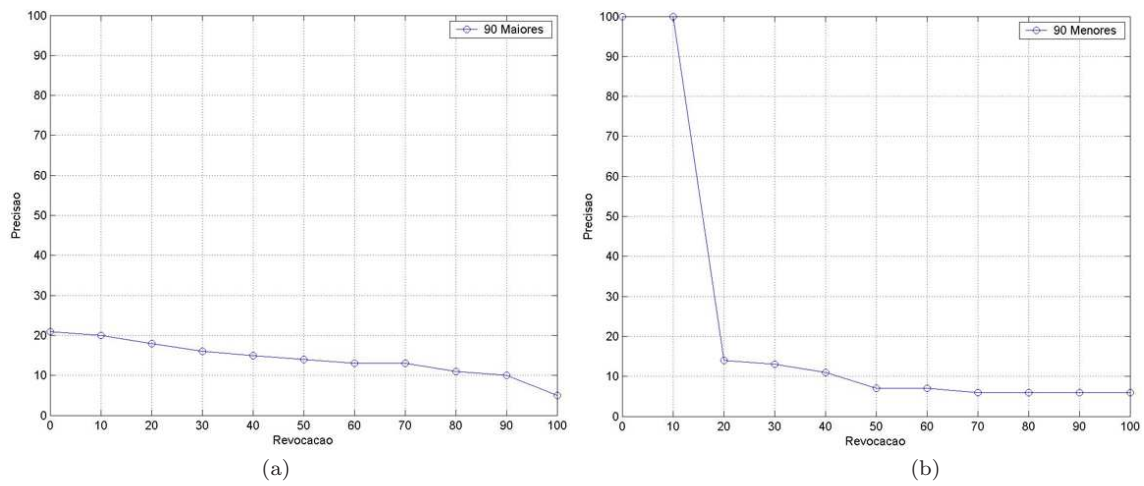


Figura 7.10: Gráficos da imagem consulta com oclusão (figura 7.9 – a).
Fonte: Gráficos elaborados pelo autor.

Os gráficos da figura 7.11 mostram o desempenho do PCA em relação à face consulta da figura 7.9 (b).



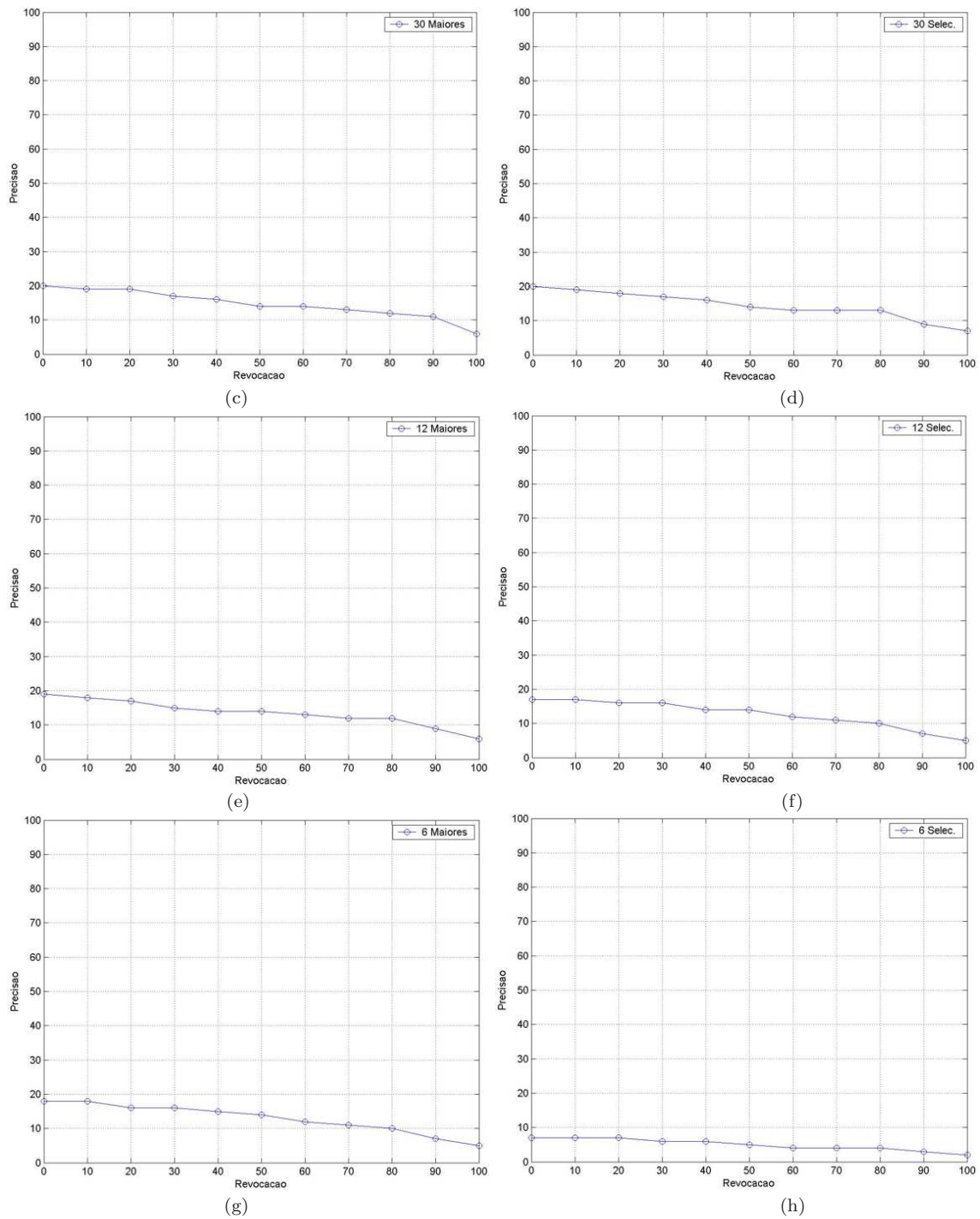


Figura 7.11: Gráficos da imagem consulta sem metade da cabeça (figura 7.9 – b).

Fonte: Gráficos elaborados pelo autor.

Os resultados da consulta com as imagens com ruído (figuras 7.9 (c) e (d)) são mostrados nos gráficos 7.12 (a, b, c, d). Os resultados com as 90 menores autofaces são mostrados nos

gráficos 7.12 (b) e (d)) e, os gráficos 7.12 (a) e (c), representam os conjuntos restantes de autôfaces. Os gráficos 7.12 (a) e (b) mostram a recuperação de face usando a face consulta da figura 7.9 (c) e os gráficos 7.12 (c) e (d) mostram o gráfico da recuperação de face usando a face consulta da figura 7.9 (d).

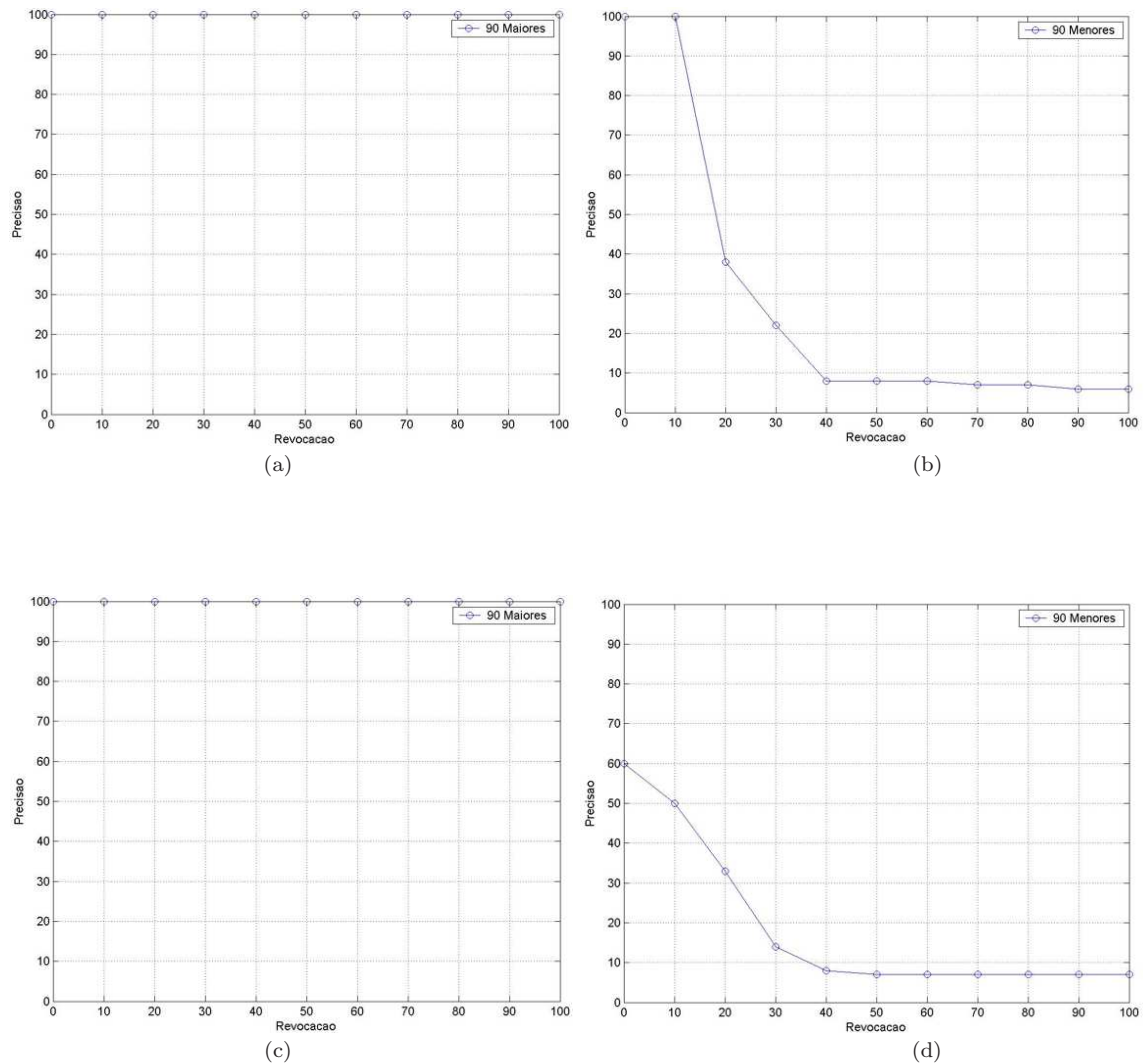
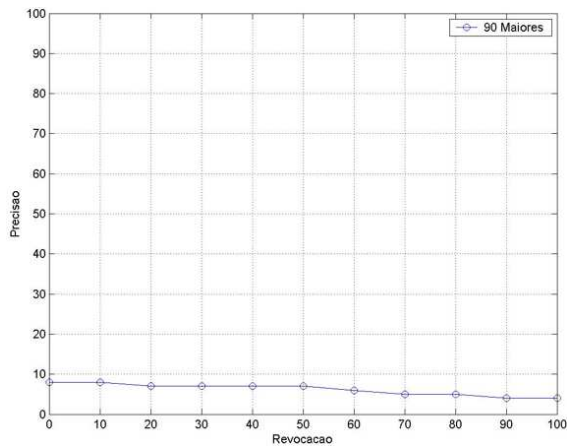


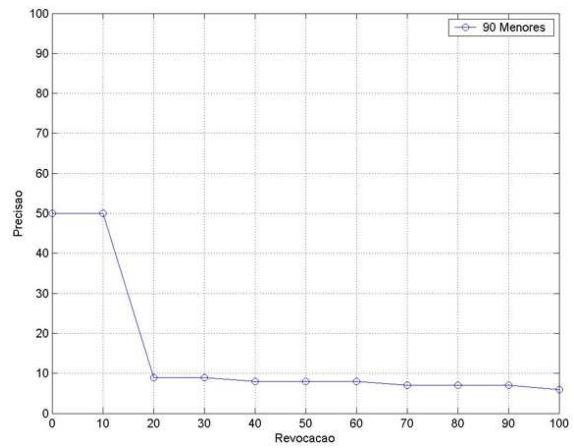
Figura 7.12: Gráficos das imagens consultadas com ruído na face e nas características da face (figura 7.9 – c e d).

Fonte: Gráficos elaborados pelo autor.

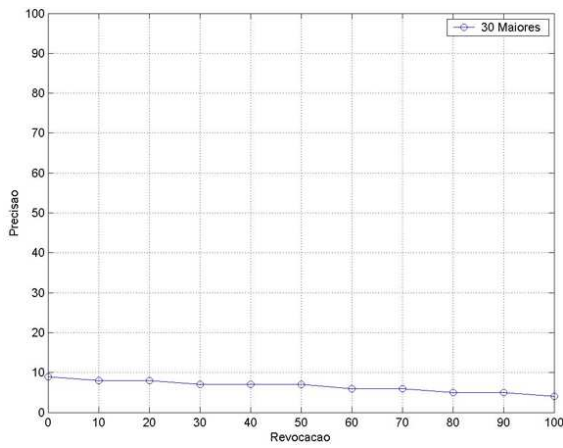
O resultado da consulta com a imagem invertida (figura 7.9 (e)) é mostrado nos gráficos 7.13 (a, b, c, d, e, f, g, h).



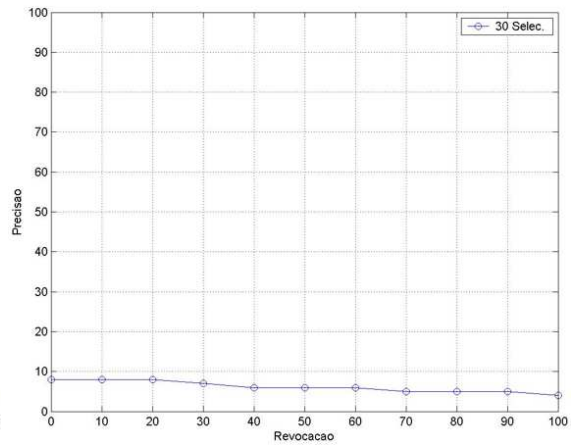
(a)



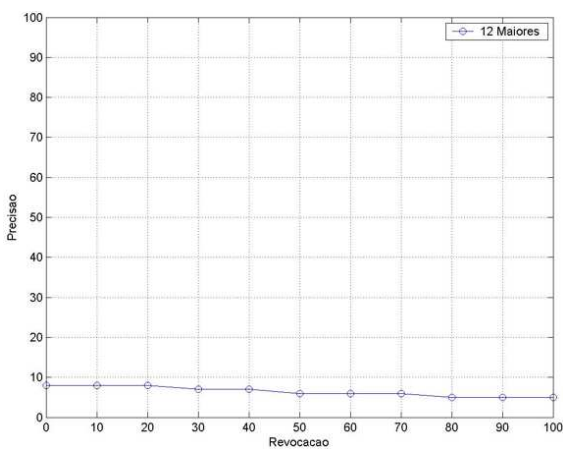
(b)



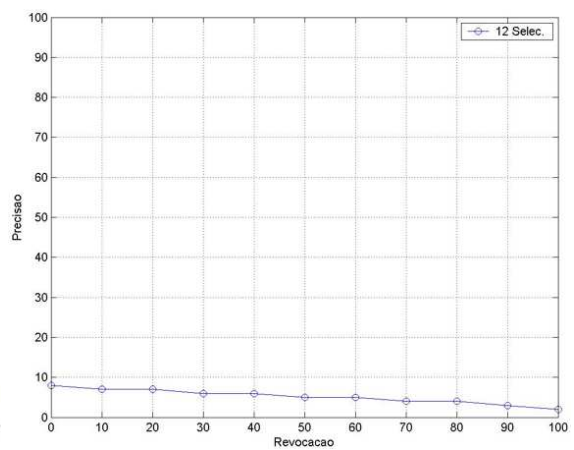
(c)



(d)



(e)



(f)

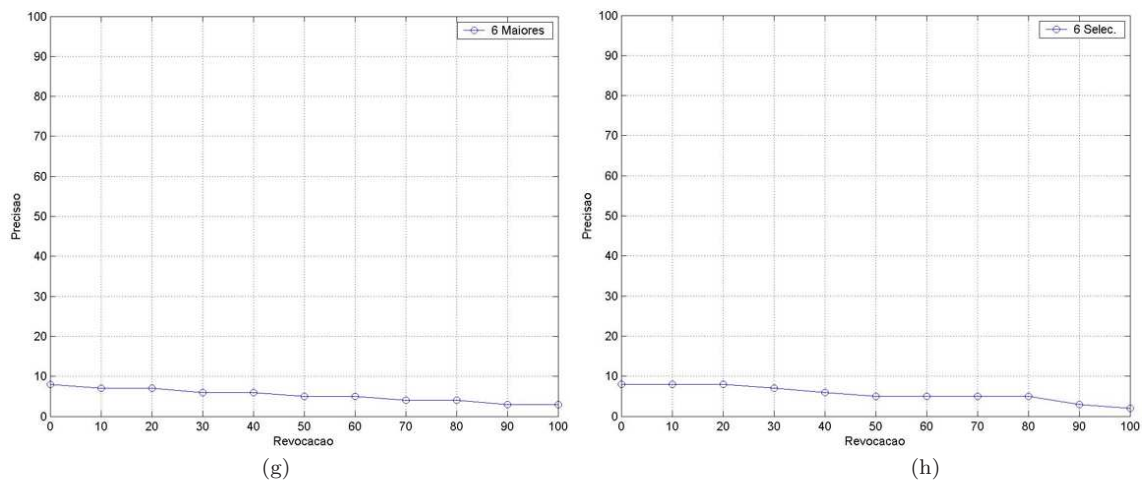
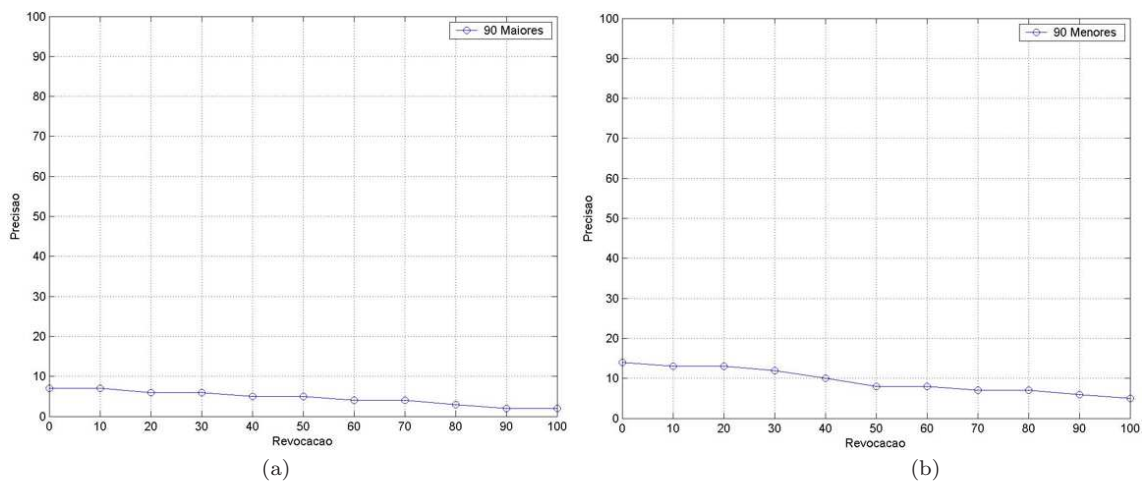


Figura 7.13: Gráficos da imagem consulta com a face invertida (figura 7.9 – e).
 Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a imagem “estreita” (figura 7.9 (f)) é mostrado nos gráficos 7.14 (a, b, c, d, e, f, g, h).



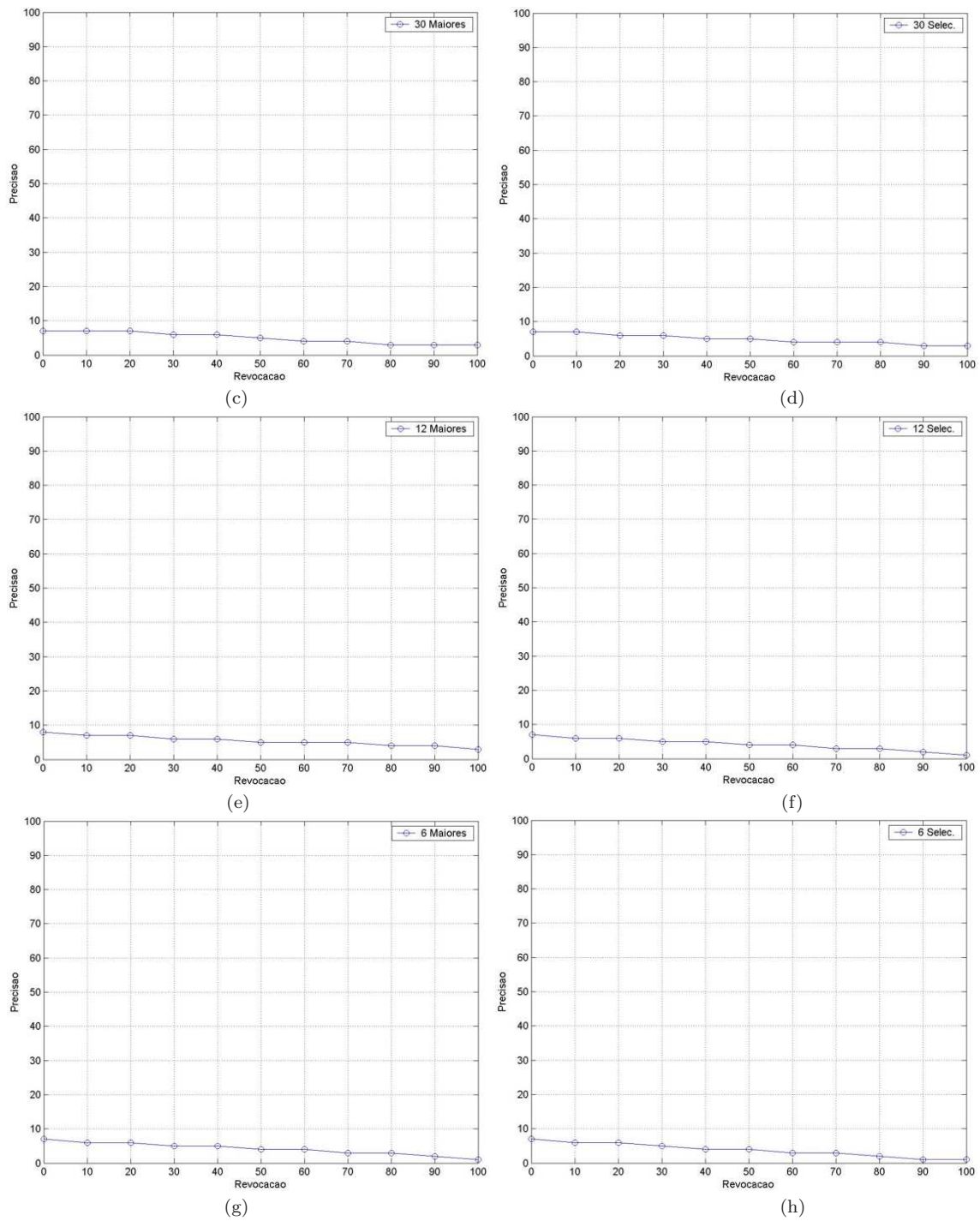
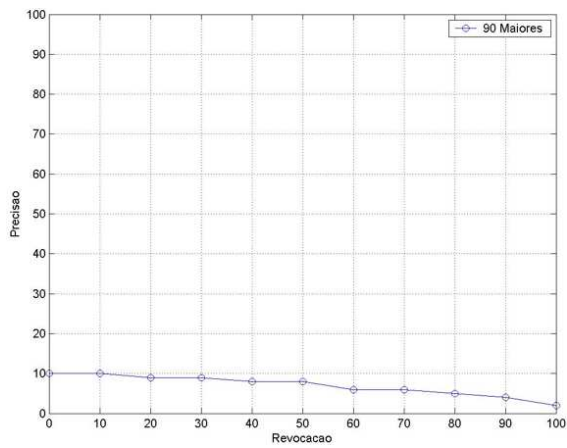


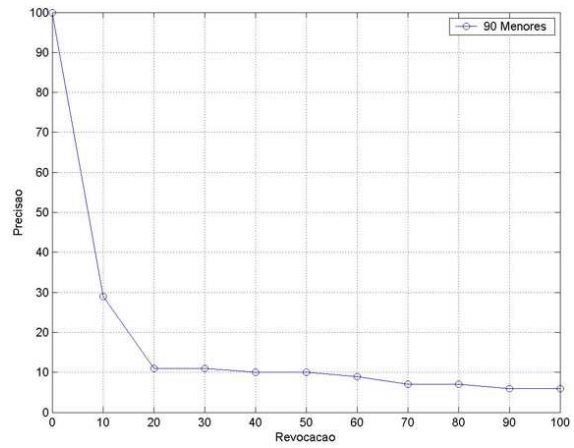
Figura 7.14: Gráficos da imagem consulta com a face “estrita” (figura 7.9 – f).

Fonte: Gráficos elaborados pelo autor.

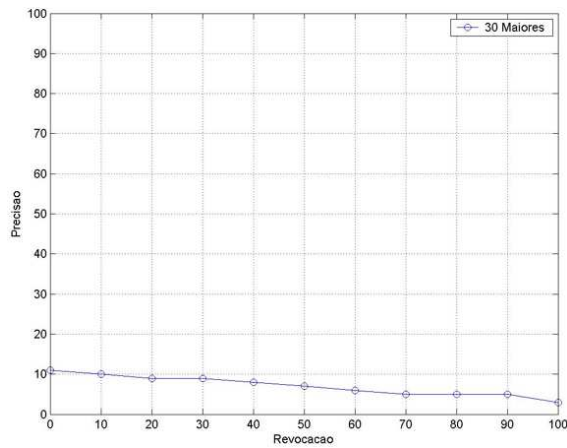
O resultado da consulta com a imagem com o cabelo cortado (figura 7.9 (g)) é mostrado nos gráficos 7.15 (a, b, c, d, e, f, g, h).



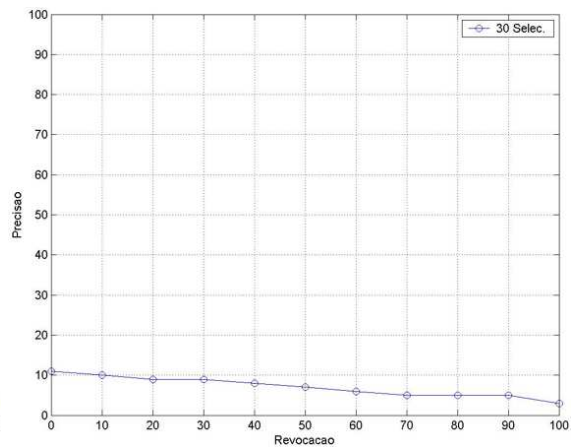
(a)



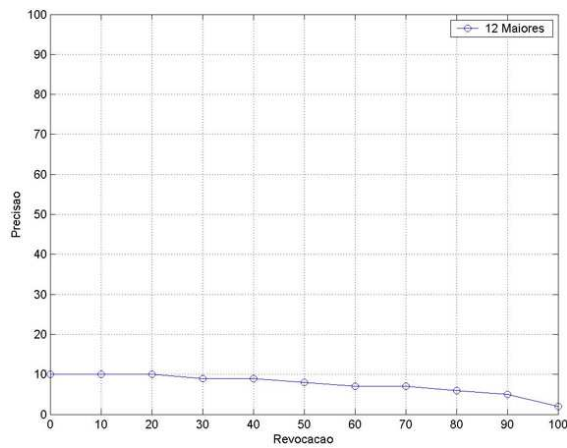
(b)



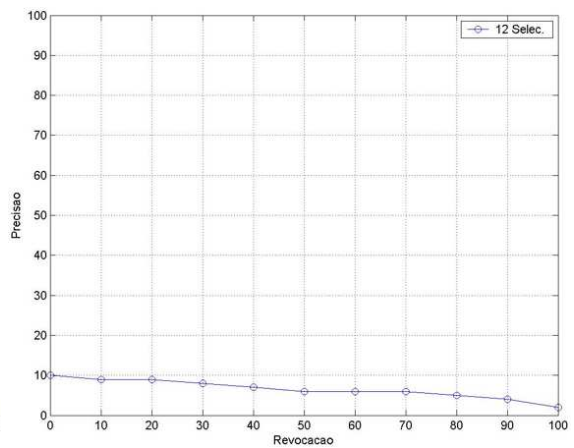
(c)



(d)



(e)



(f)

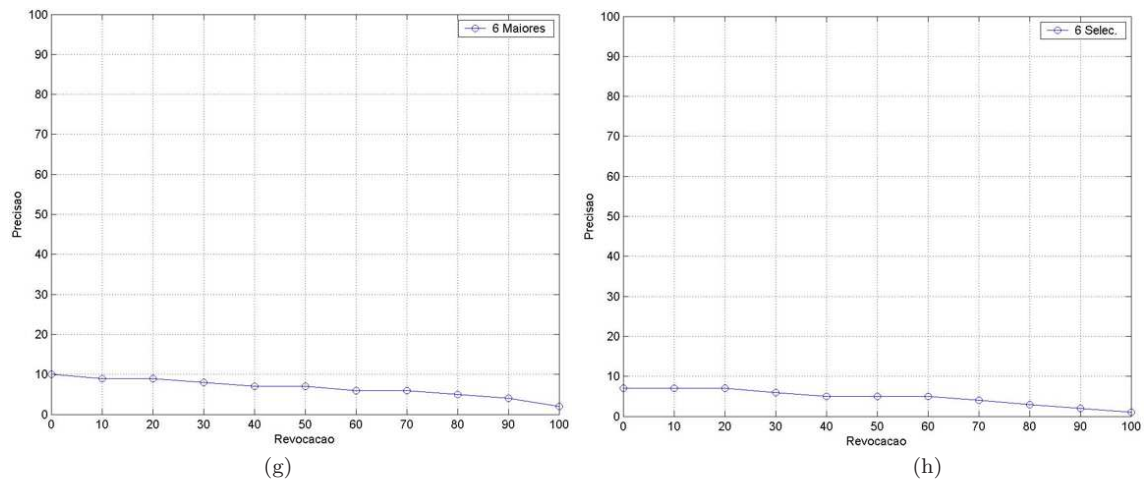
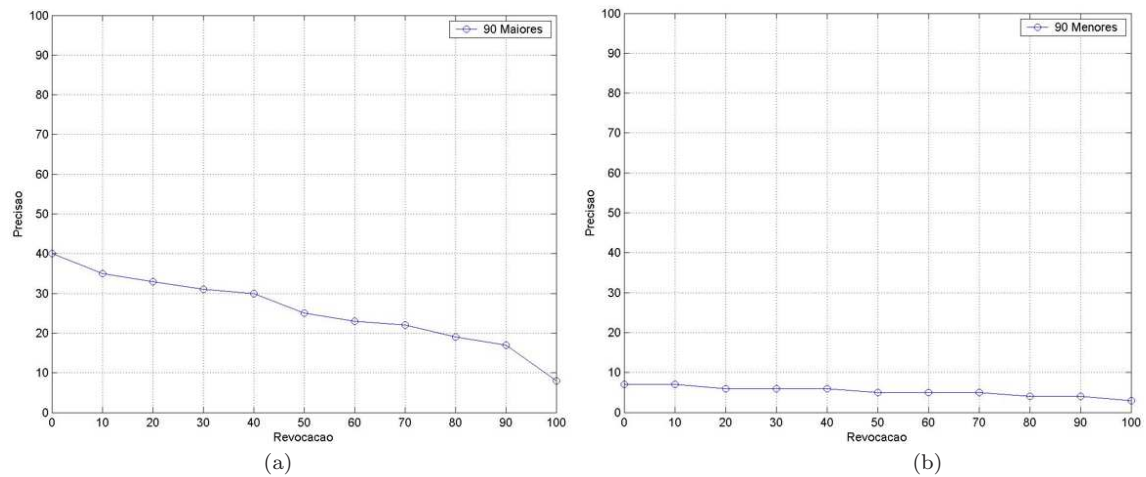


Figura 7.15: Gráficos da imagem consulta com a face com o cabelo cortado (figura 7.9 – *g*).
Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face de características internas modificadas (figura 7.9 (*h*)) é mostrado nos gráficos 7.16 (*a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*).



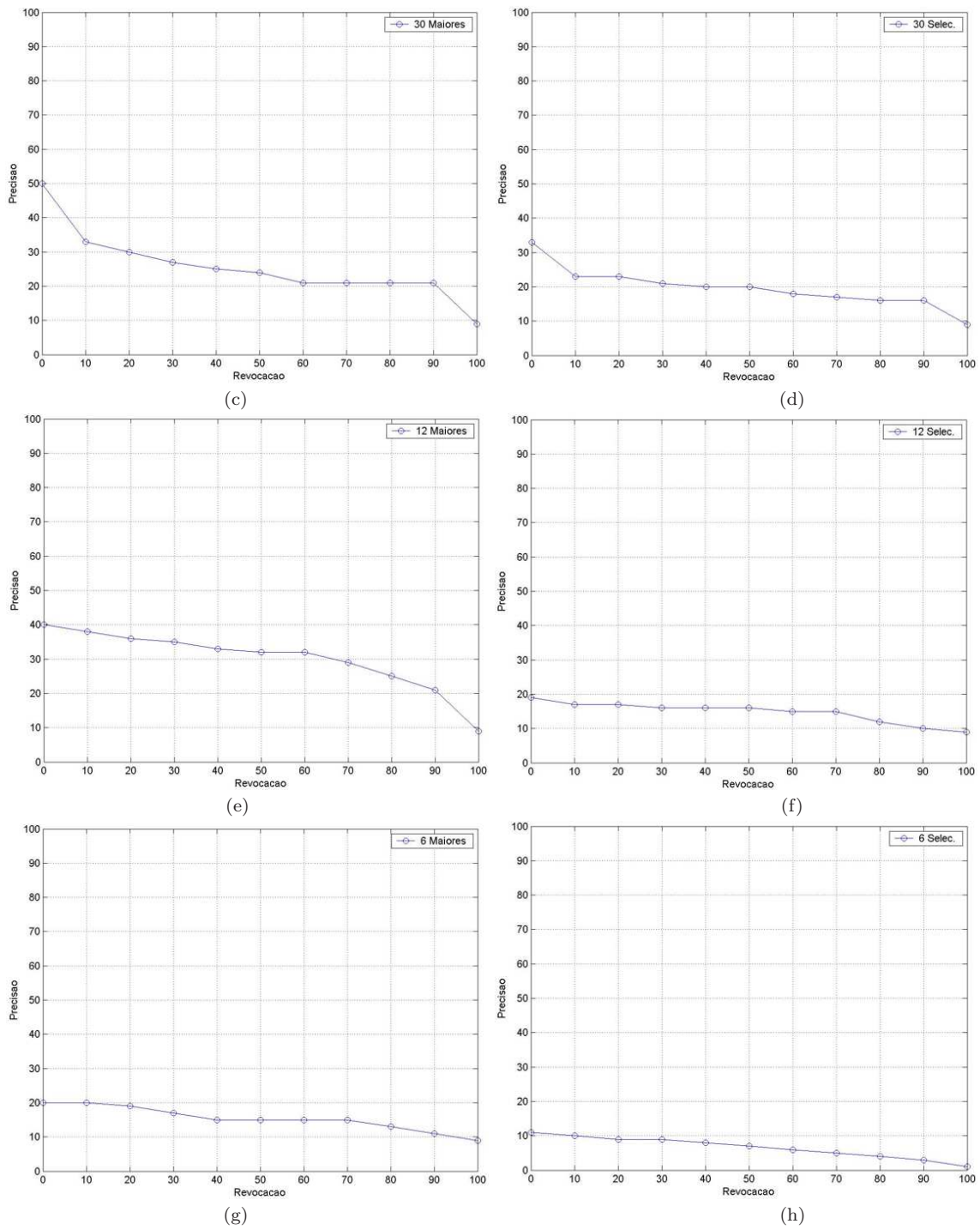
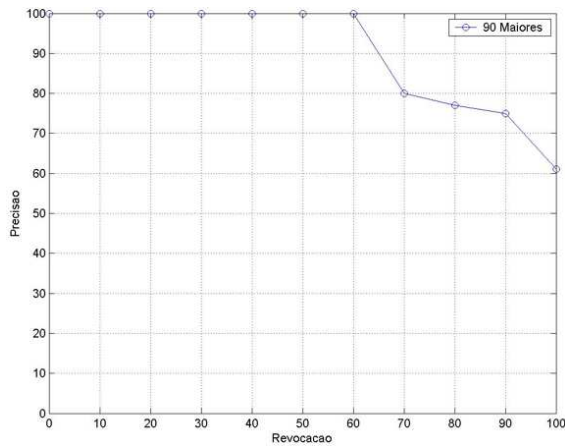


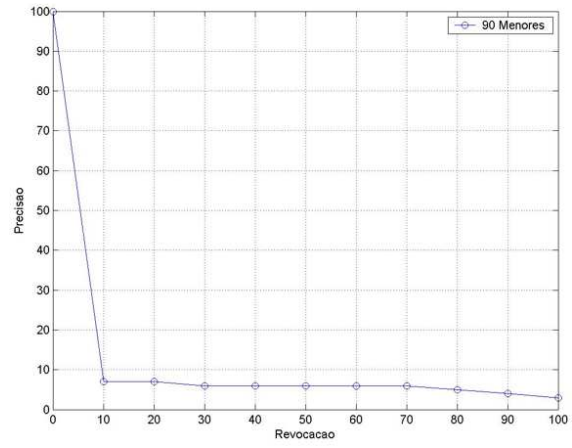
Figura 7.16: Gráficos da imagem consulta com as características internas modificadas (figura 7.9 – h).

Fonte: Gráficos elaborados pelo autor.

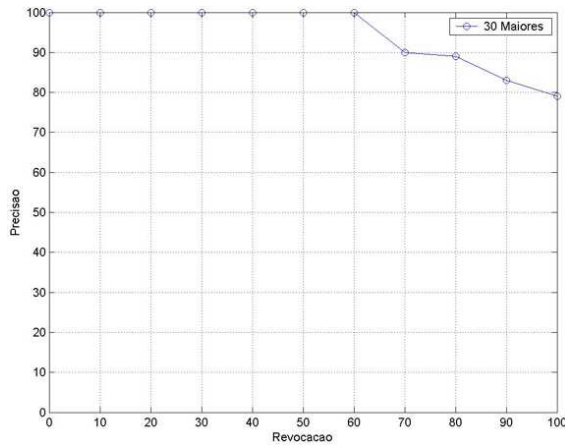
O resultado da consulta com a face pertencente à base de dados e de características internas modificadas (figura 7.9 (i)) é mostrado nos gráficos 7.17 (a, b, c, d, e, f, g, h).



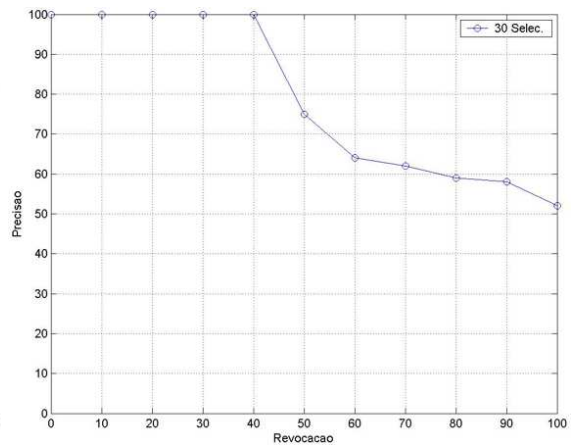
(a)



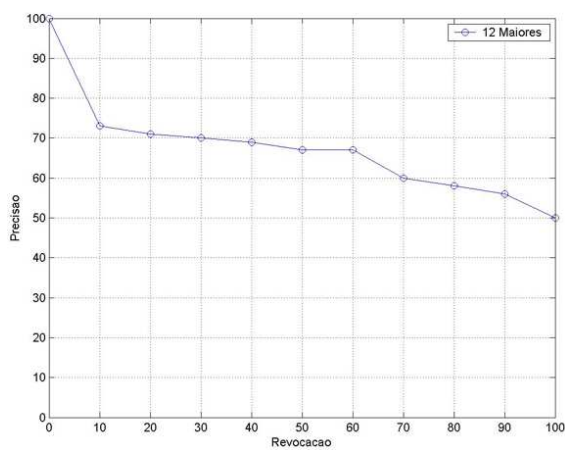
(b)



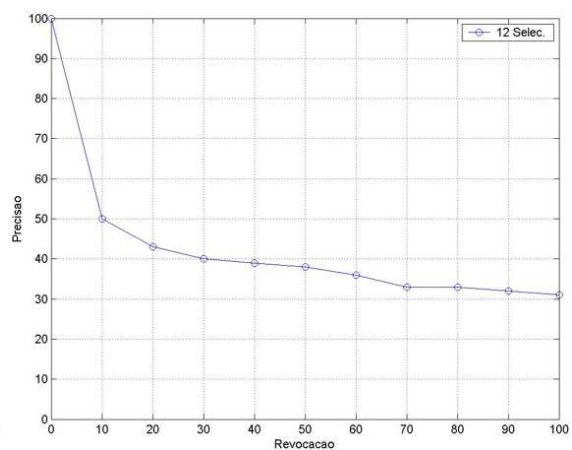
(c)



(d)



(e)



(f)

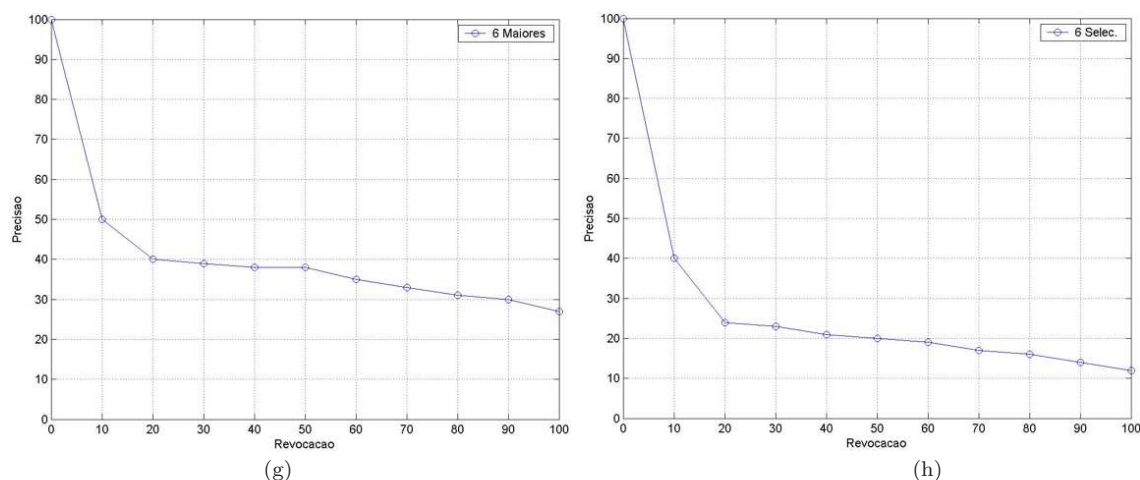


Figura 7.17: Gráficos da imagem consulta de uma face pertencente à base de dados com as características internas modificadas (figura 7.9 – *i*).

Fonte: Gráficos elaborados pelo autor.

As autofaces quando, aplicadas em situações em que a face consulta é alterada com perturbações (ruído, oclusões, etc.), mostraram-se robustas na maior parte dos casos. Nos casos em que havia oclusão e ruído na face, a seleção dos autovalores não alteraram o resultado da busca e todos os grupos de autovalores têm precisão 100%. Contudo, as autofaces selecionadas com base nos 90 menores autovalores, prejudicou substancialmente o reconhecimento. Provavelmente, os menores autovalores estão mais relacionados às características internas da face do que ao contorno. Como as imagens foram danificadas em sua maioria na parte interior da face, o desempenho do método das autofaces é prejudicado quando faz uso das autofaces relacionadas aos menores autovalores. Essa influência pode ser enfatizada com a análise dos resultados das imagens consultas da figura 7.9(*b, e, g*) em que as maiores taxas de precisão são as dos menores autovalores.

Já com as características internas da face alteradas, o uso dos maiores autovalores resultou em taxas de precisão mais altas. Mais uma vez, podemos supor que os maiores autovalores são mais influenciados pelo formato da face, justificando taxas mais altas de precisão do que as taxas das outras seleções de autovalores.

7.3.5 Experimento 2: Problema de diferentes condições de iluminação

Sendo a iluminação, um dos maiores desafios dos sistemas de reconhecimento de face, o objetivo desse teste foi verificar o comportamento do método das autofaces quando as luzes incidentes nas faces variam de intensidade e ângulo de incidência.

Neste experimento são usadas as imagens de faces da base de dados YaleB (capítulo 3 na seção 3.3.3). Essa base de dados possui imagens de muita e pouca incidência de luz, isto é, há faces em condições normais de iluminação e faces em que a iluminação é quase ou totalmente retirada. O conjunto de treinamento (usado para gerar as autofaces) é formado por 23 imagens de face (sob condições normais de iluminação) de cada pessoa da base de dados. Um exemplo das faces consultas e das diferentes iluminações aplicadas nas faces é mostrado na figura 7.18.



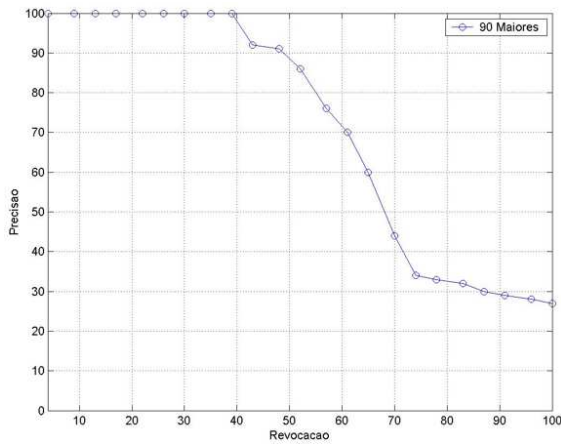
Figura 7.18: Base de dados YaleB (iluminação).
Fonte: Imagens obtidas de [4].

As faces usadas na consulta (vide figura 7.19) não pertencem ao conjunto de treinamento e estão sob condições de pouca iluminação. O desempenho das autofaces é analisado pela taxa de precisão de cada consulta considerando a configuração descrita na seção 7.3.1 para a seleção das autofaces.

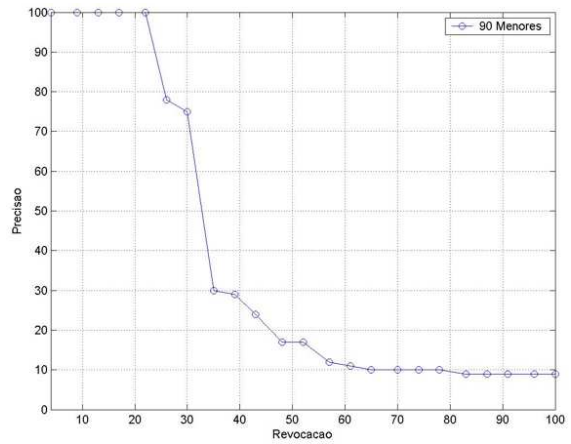


Figura 7.19: Base de dados YaleB (faces usadas nas consultas).
Fonte: Imagens obtidas de [4].

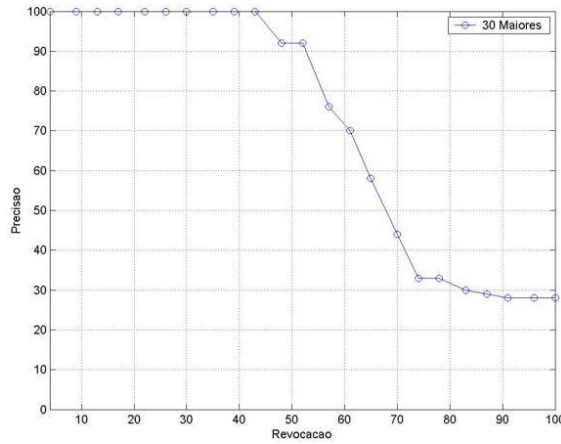
O resultado da consulta com a face da figura 7.19 (a) é mostrado nos gráficos 7.20 (a, b, c, d, e, f, g, h).



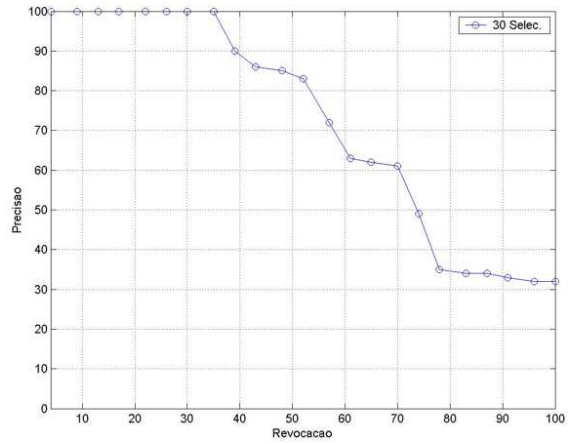
(a)



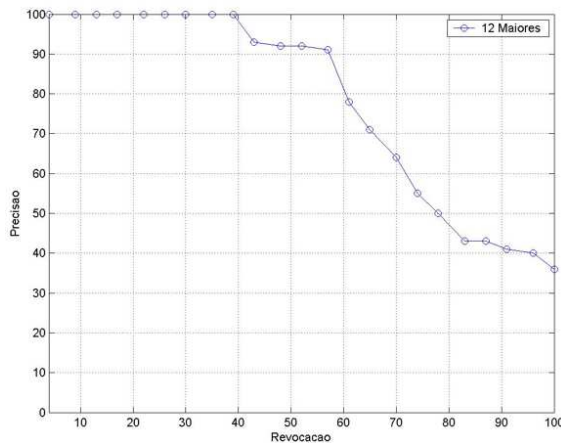
(b)



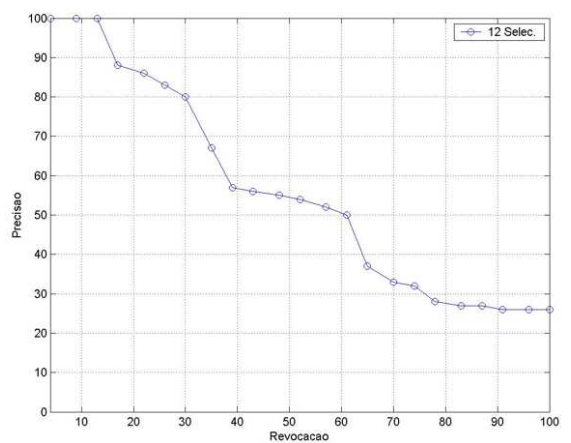
(c)



(d)



(e)



(f)

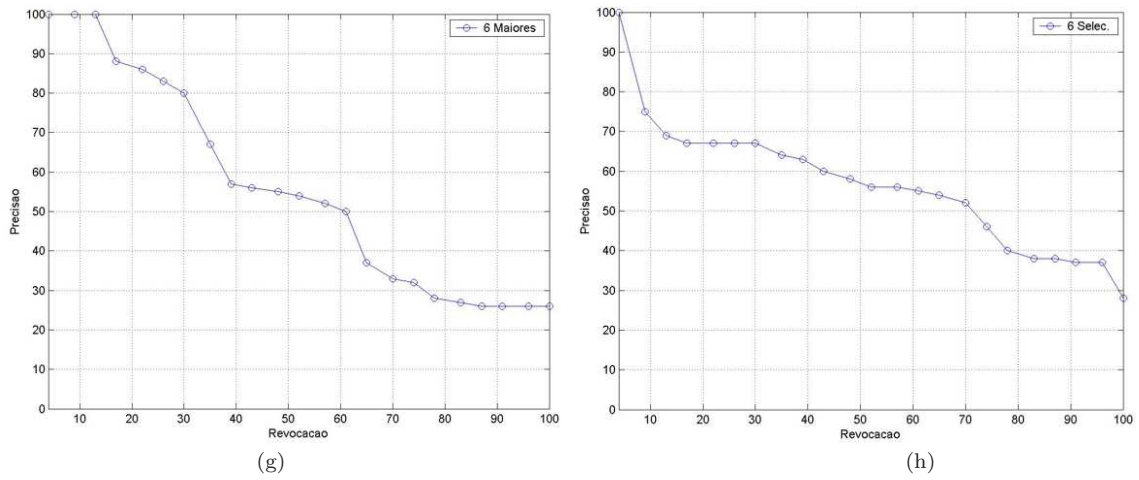
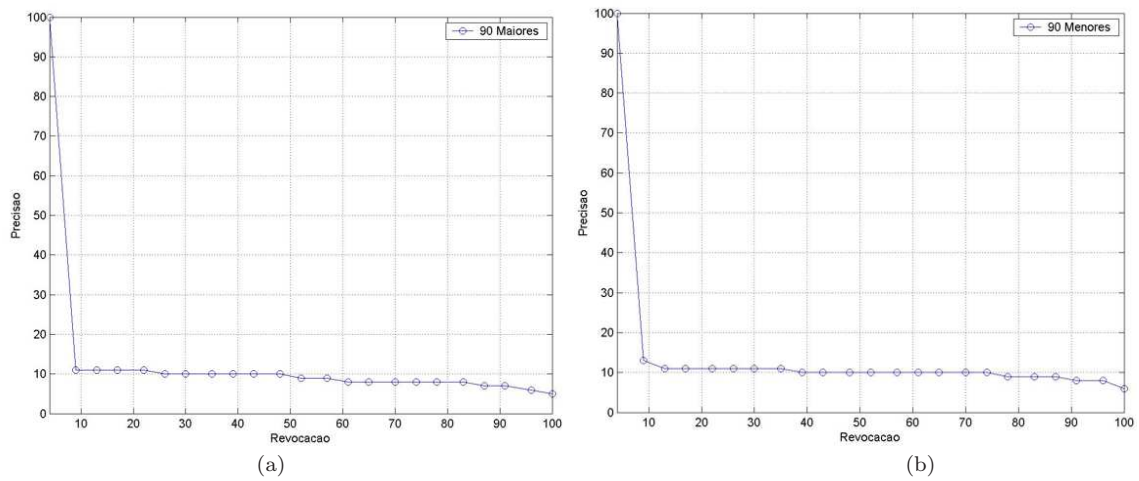


Figura 7.20: Gráficos da precisão (Imagem consulta: figura 7.19- a).
 Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face da figura 7.19 (b) é mostrado nos gráficos 7.21 (a, b, c, d, e, f, g, h).



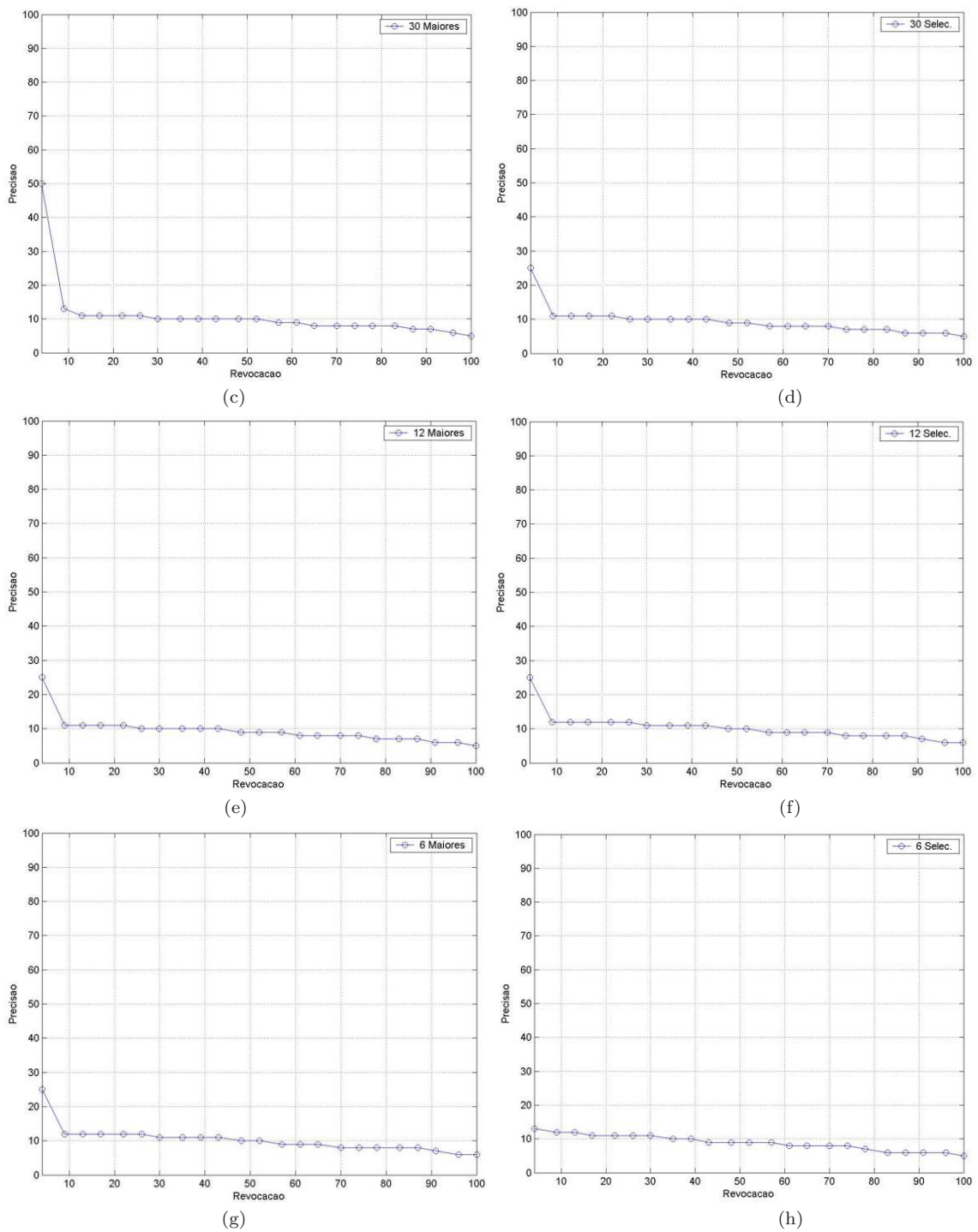
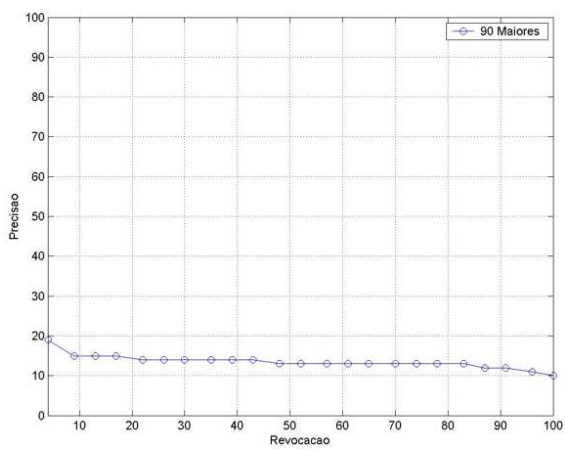


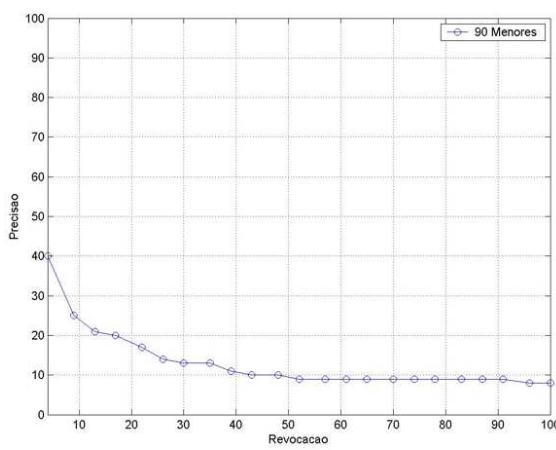
Figura 7.21: Gráficos da precisão (Imagem consulta: figura 7.19- b).

Fonte: Gráficos elaborados pelo autor.

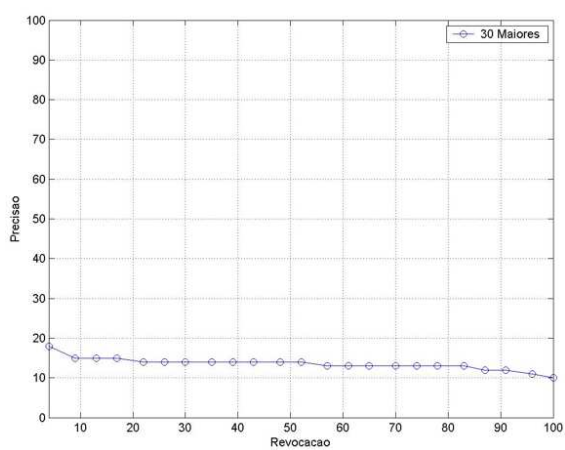
O resultado da consulta com a face da figura 7.19 (c) é mostrado nos gráficos 7.22 (a, b, c, d, e, f, g, h).



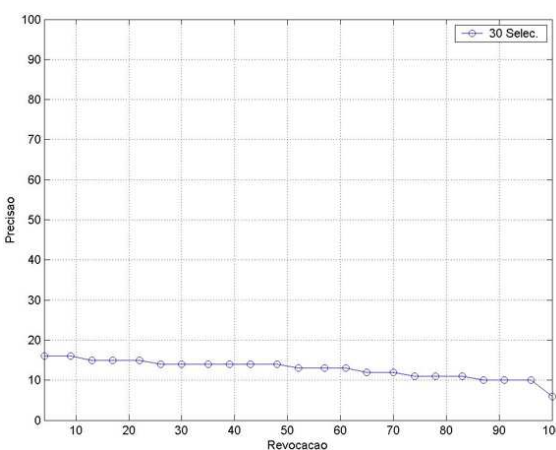
(a)



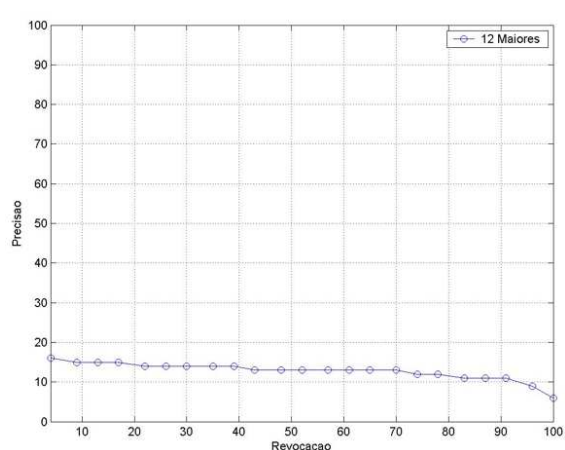
(b)



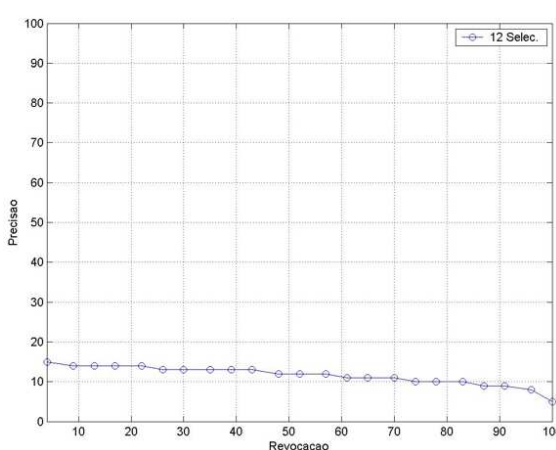
(c)



(d)



(e)



(f)

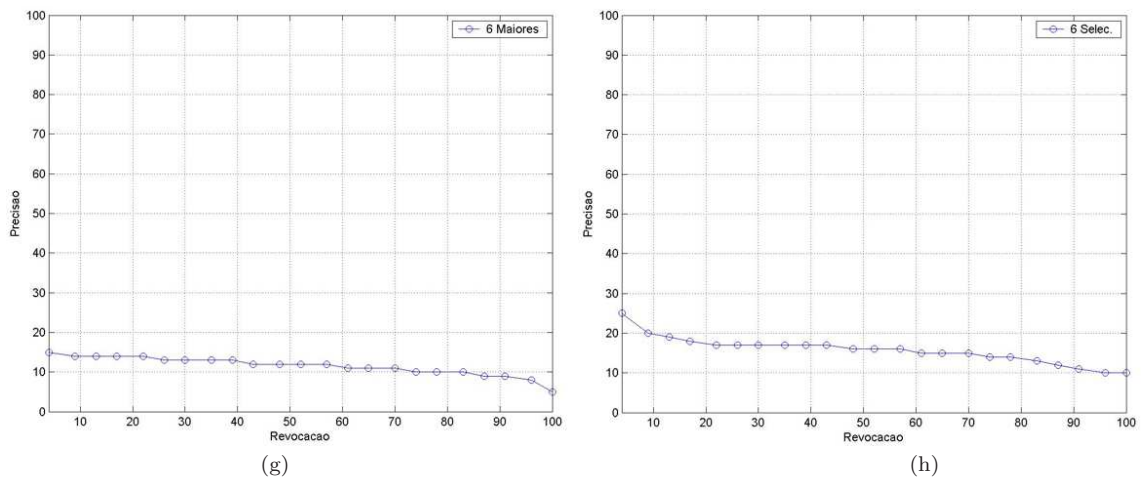
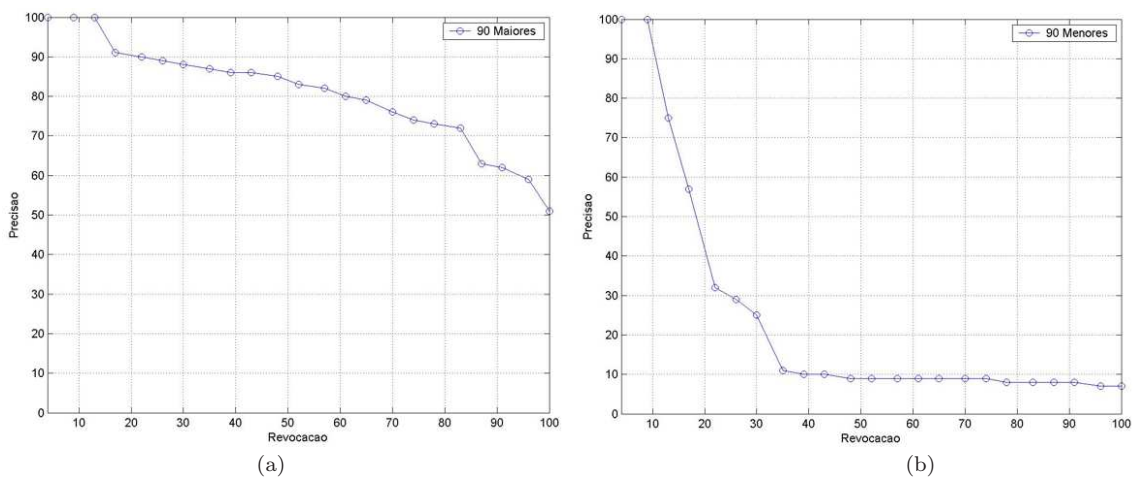


Figura 7.22: Gráficos da precisão (Imagem consulta: figura 7.19- c).
 Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face da figura 7.19 (d) é mostrado nos gráficos 7.23 (a, b, c, d, e, f, g, h).



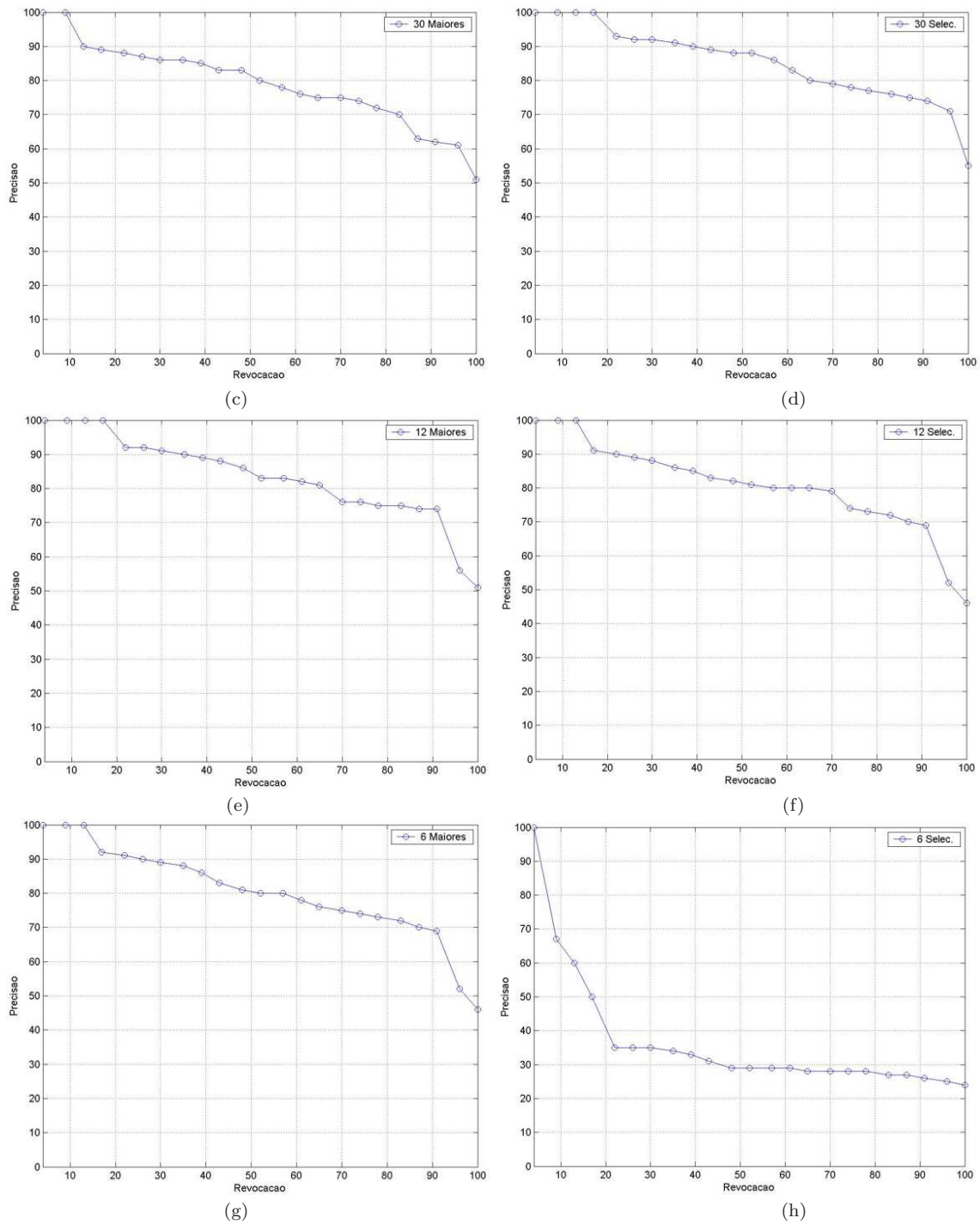
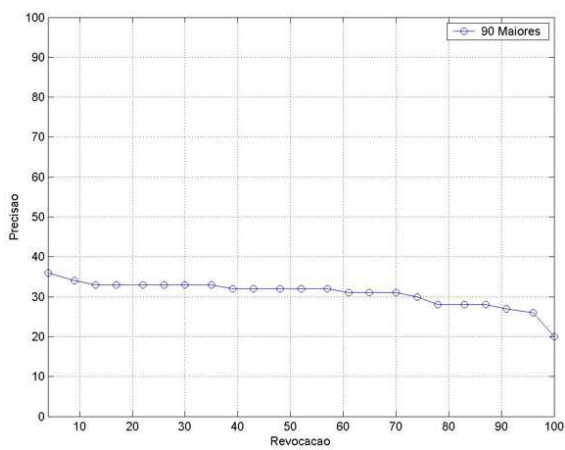


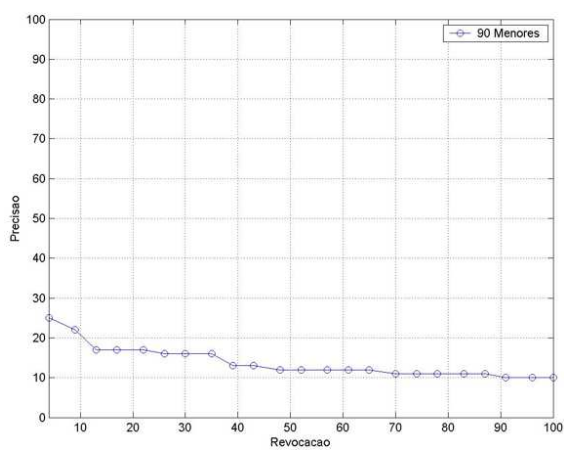
Figura 7.23: Gráficos da precisão (Imagem consulta: figura 7.19- d).

Fonte: Gráficos elaborados pelo autor.

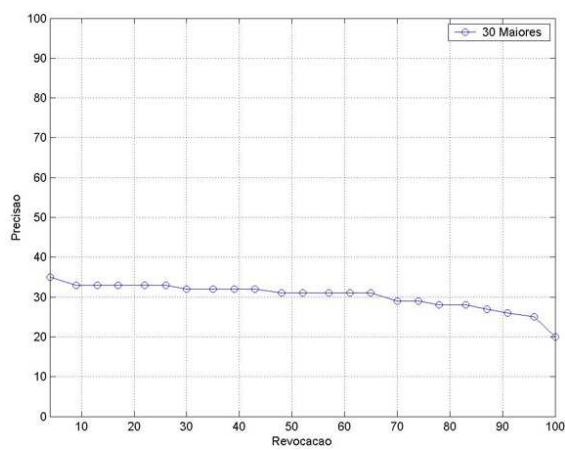
O resultado da consulta com a face da figura 7.19 (e) é mostrado nos gráficos 7.24 (a, b, c, d, e, f, g, h).



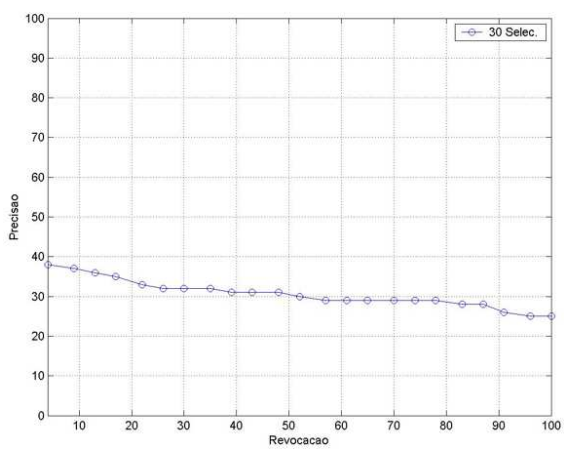
(a)



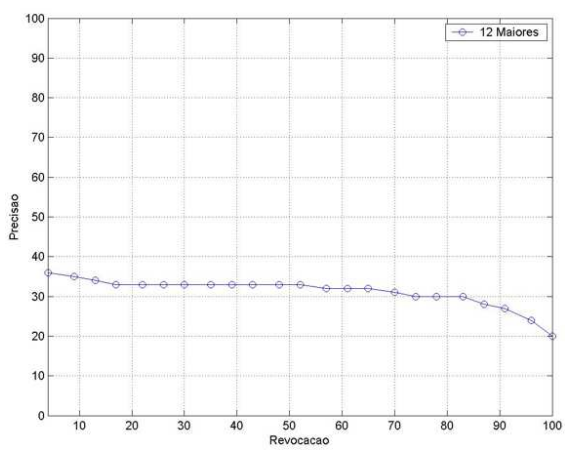
(b)



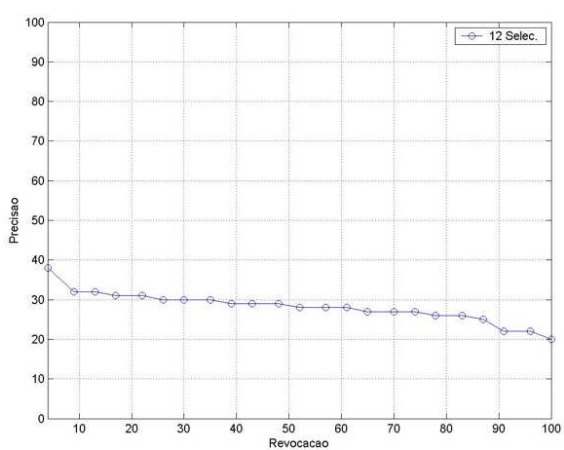
(c)



(d)



(e)



(f)

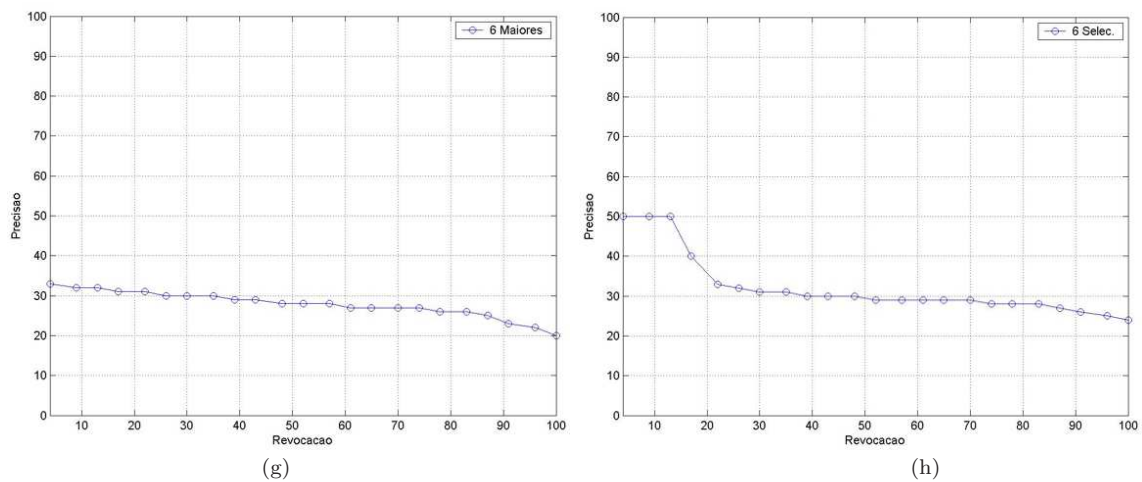
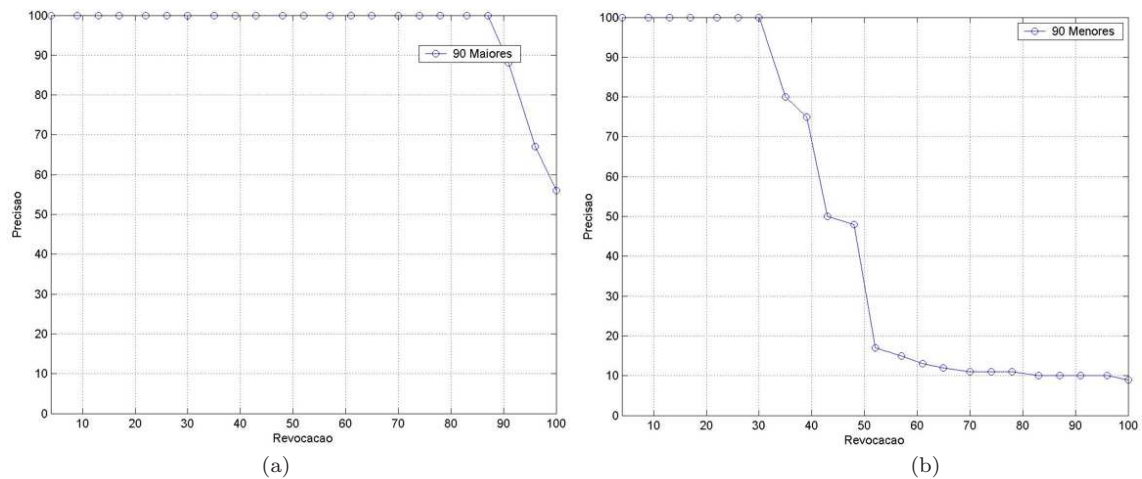


Figura 7.24: Gráficos da precisão (Imagem consulta: figura 7.19- e).

Fonte: Gráficos elaborados pelo autor.

Um novo teste é executado com as faces das mesmas pessoas usadas no teste anterior, porém, em condições normais de iluminação. Os resultados são avaliados em gráficos da precisão e revocação que mostram o desempenho das autofaces sob condições normais de iluminação. É possível notar, pelos gráficos, uma melhora significativa em relação aos gráficos das consultas em condições de iluminação variando.

Resultado da consulta com a face da figura 7.19 (a) em condições normais de iluminação é mostrado nos gráficos 7.25 (a, b, c, d, e, f, g, h).



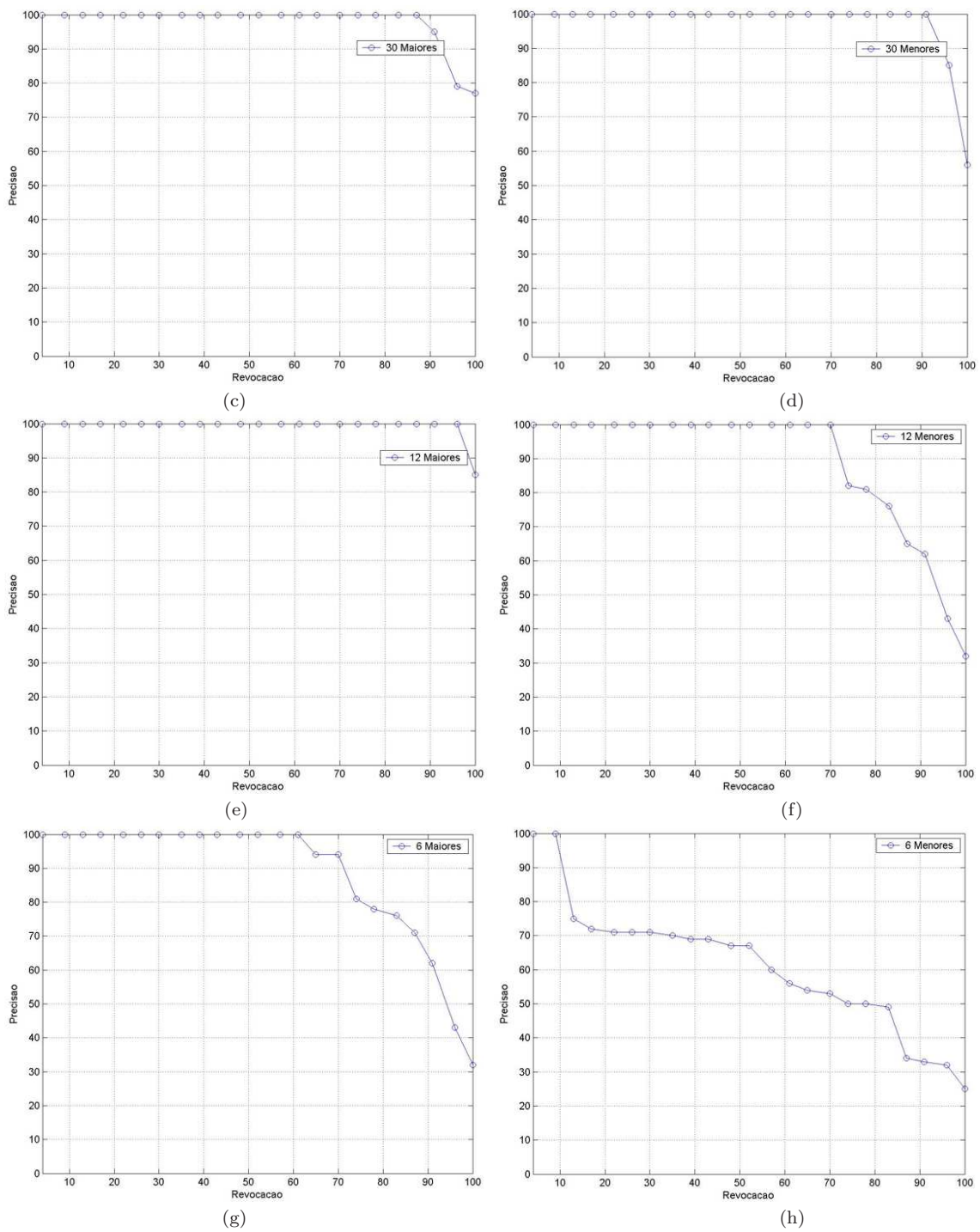
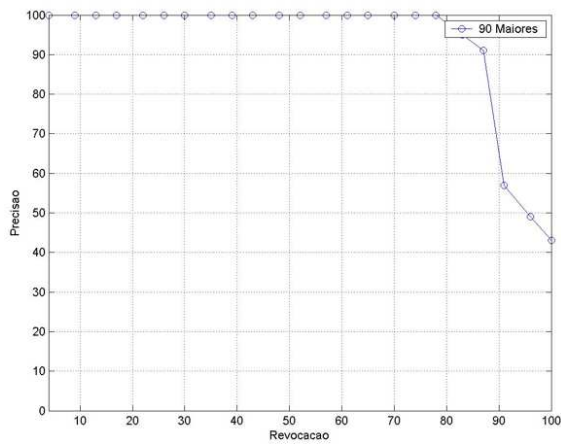


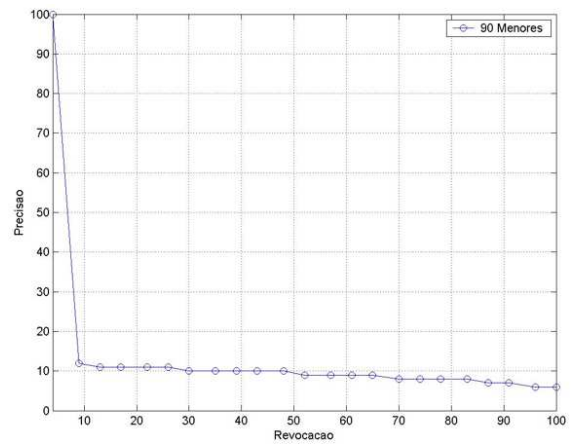
Figura 7.25: Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- a) em condições normais de iluminação.

Fonte: Gráficos elaborados pelo autor.

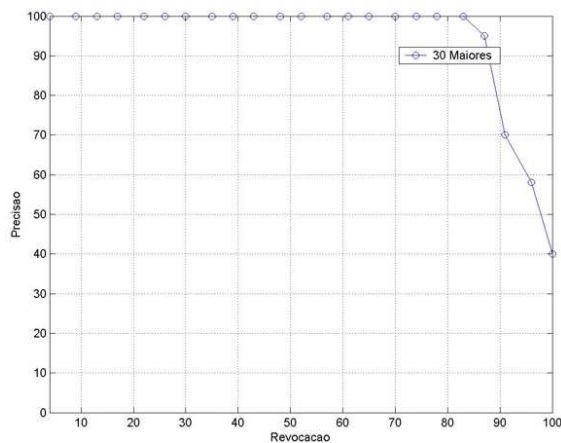
O resultado da consulta com a face da figura 7.19 (b) em condições normais de iluminação é mostrado nos gráficos 7.26 (a, b, c, d, e, f, g, h).



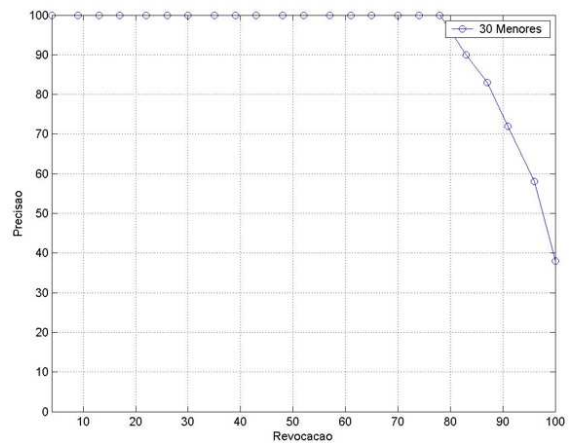
(a)



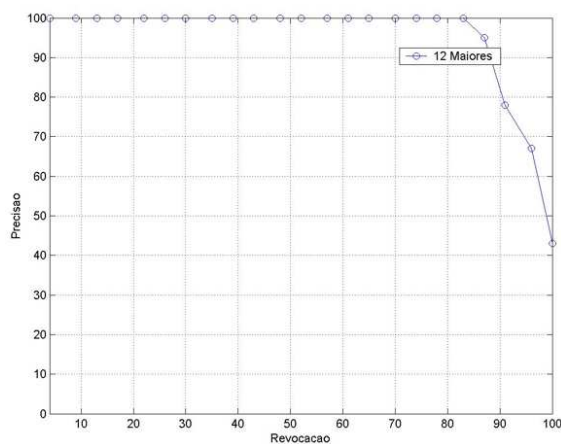
(b)



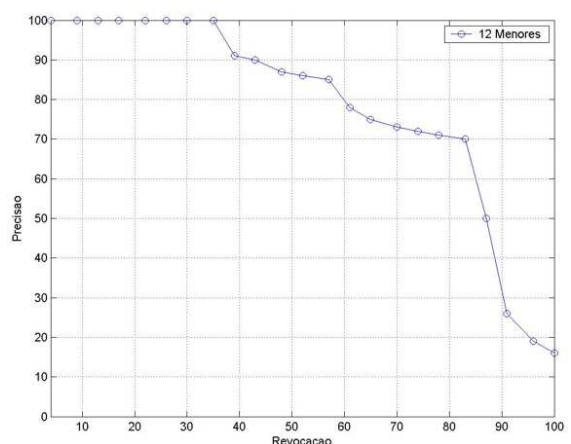
(c)



(d)



(e)



(f)

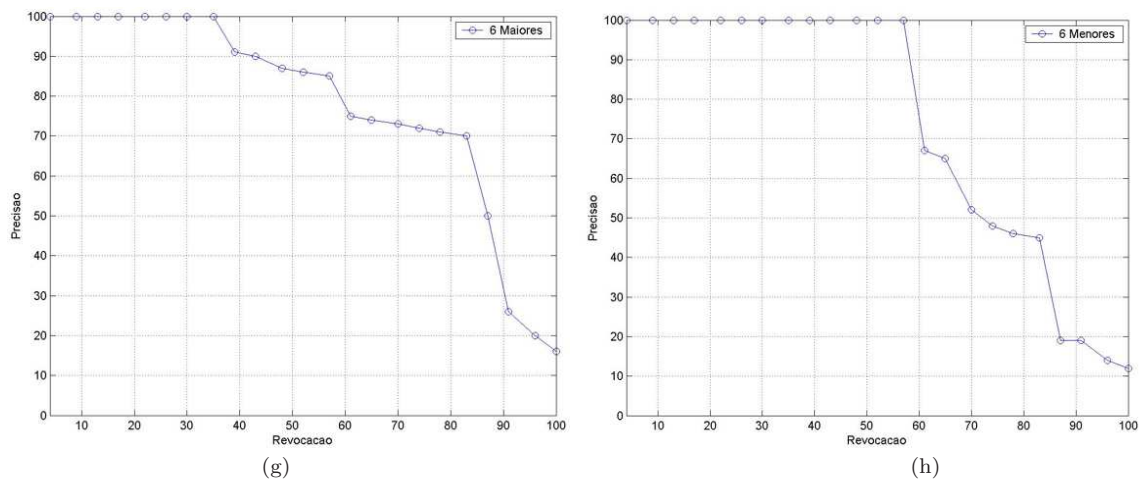
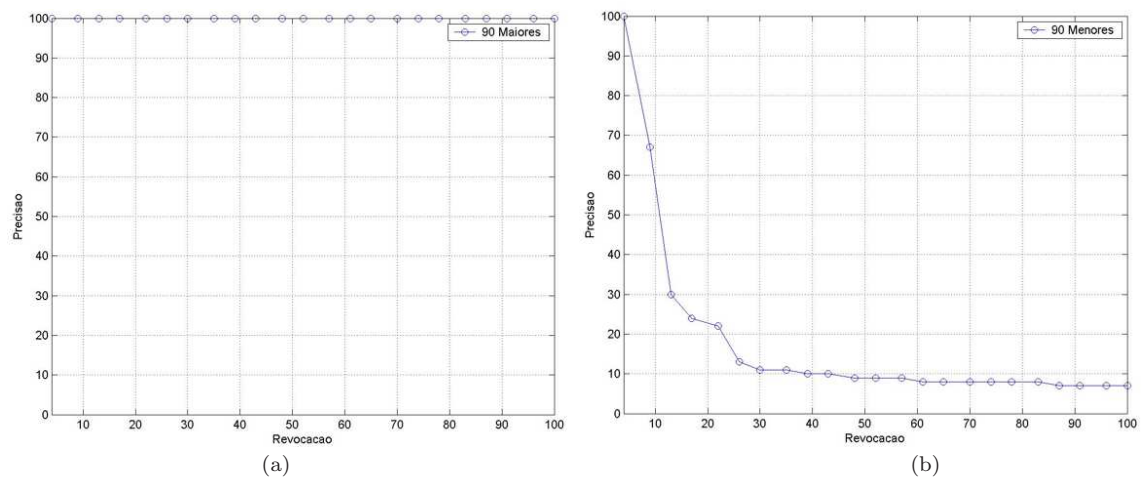


Figura 7.26: Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- b) em condições normais de iluminação.

Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face da figura 7.19 (c) em condições normais de iluminação é mostrado nos gráficos 7.27 (a, b, c, d). Os resultados gerados por essa consulta usando os 90 maiores autovalores, os 30 maiores e menores, os 12 e os 6 maiores autovalores geraram um gráfico igual ao da figura 7.27 (a).



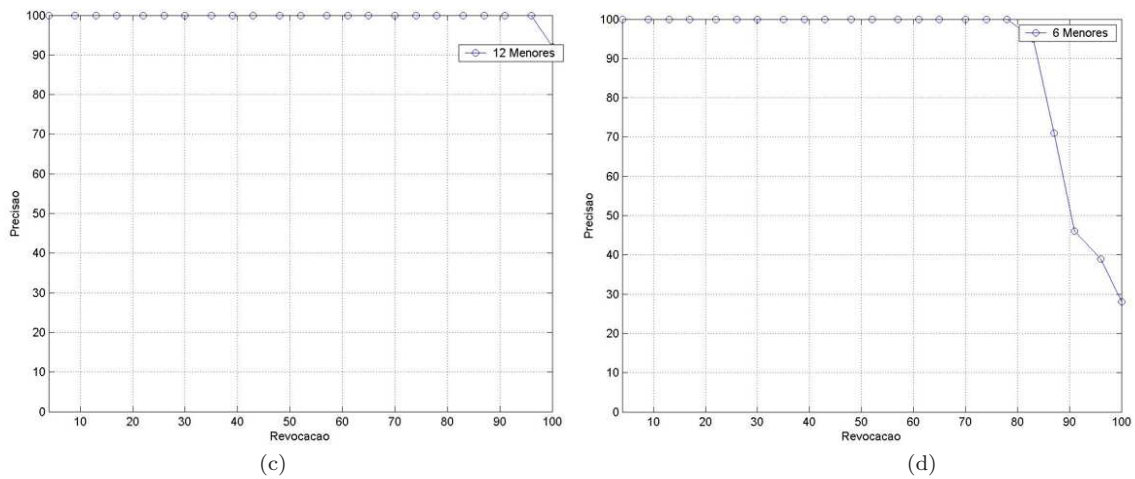
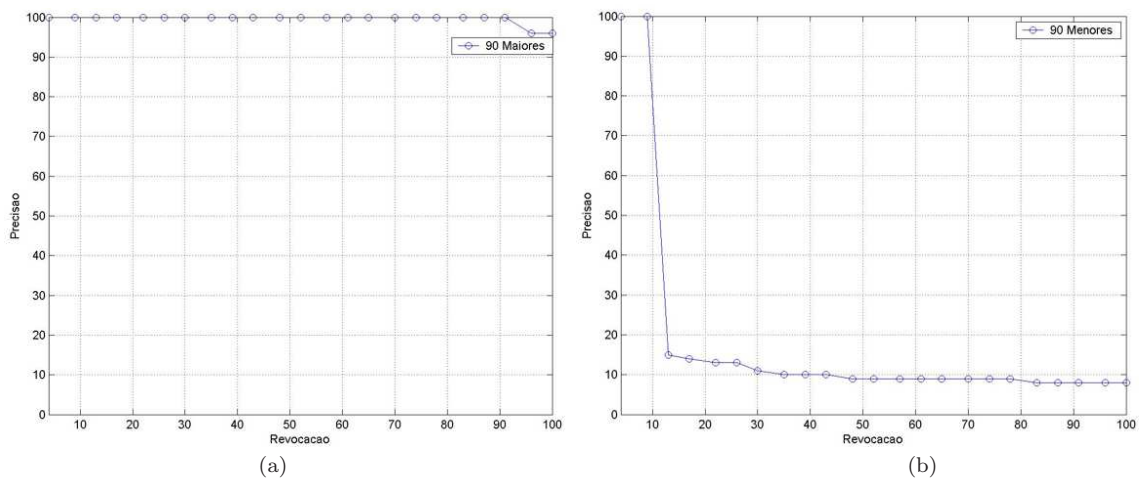


Figura 7.27: Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- c) em condições normais de iluminação.

Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face da figura 7.19 (d) em condições normais de iluminação é mostrado nos gráficos 7.28 (a, b, c, d, e, f, g, h).



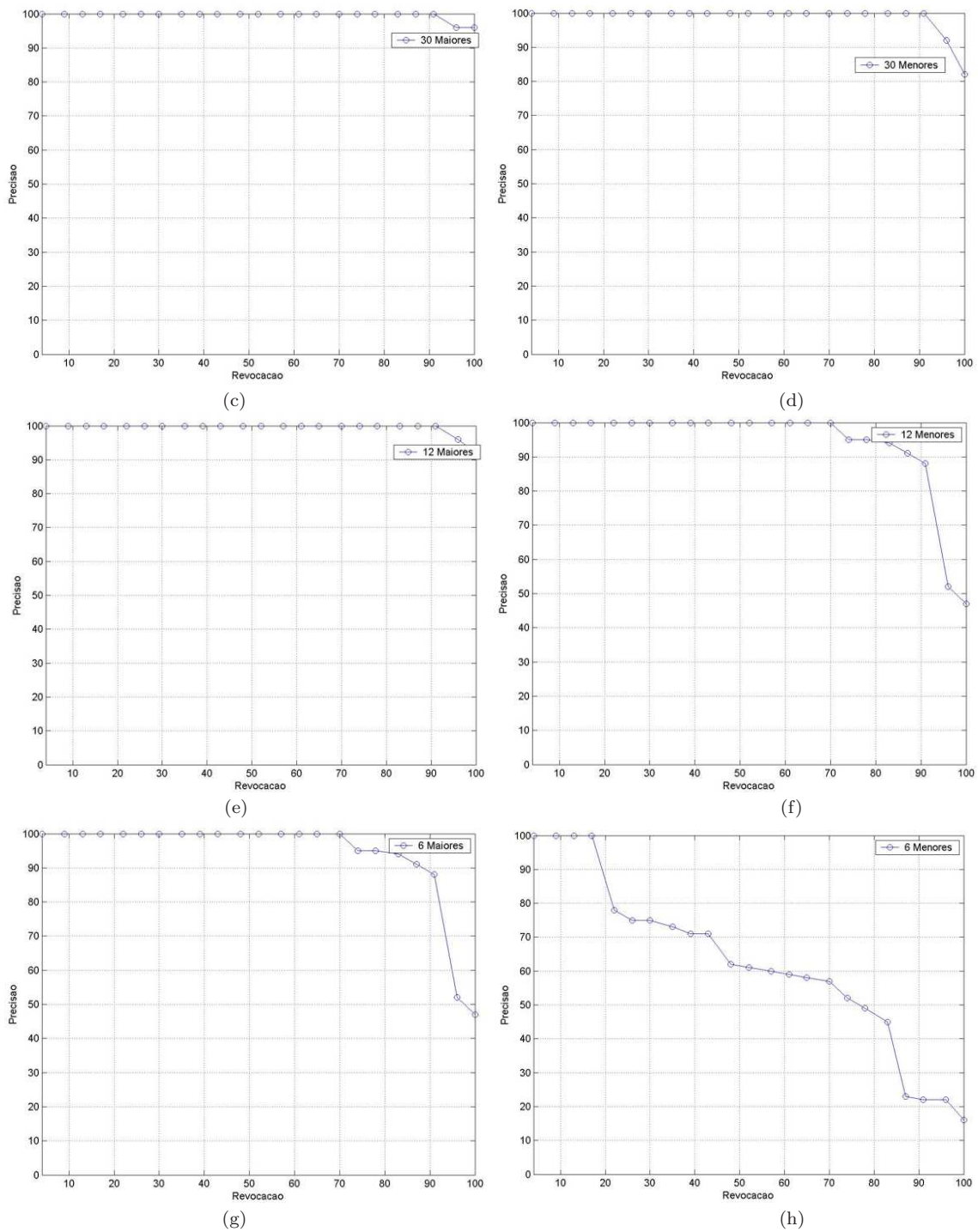


Figura 7.28: Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- d) em condições normais de iluminação.

Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face da figura 7.19 (e) em condições normais de iluminação é mostrado nos gráficos 7.29 (a, b, c, d). O resultado dos 30 maiores autovalores e selecionados é representado pelo gráfico da figura 7.29 (a), e o de 12 autovalores selecionados é representado pelo gráfico 7.29 (c).

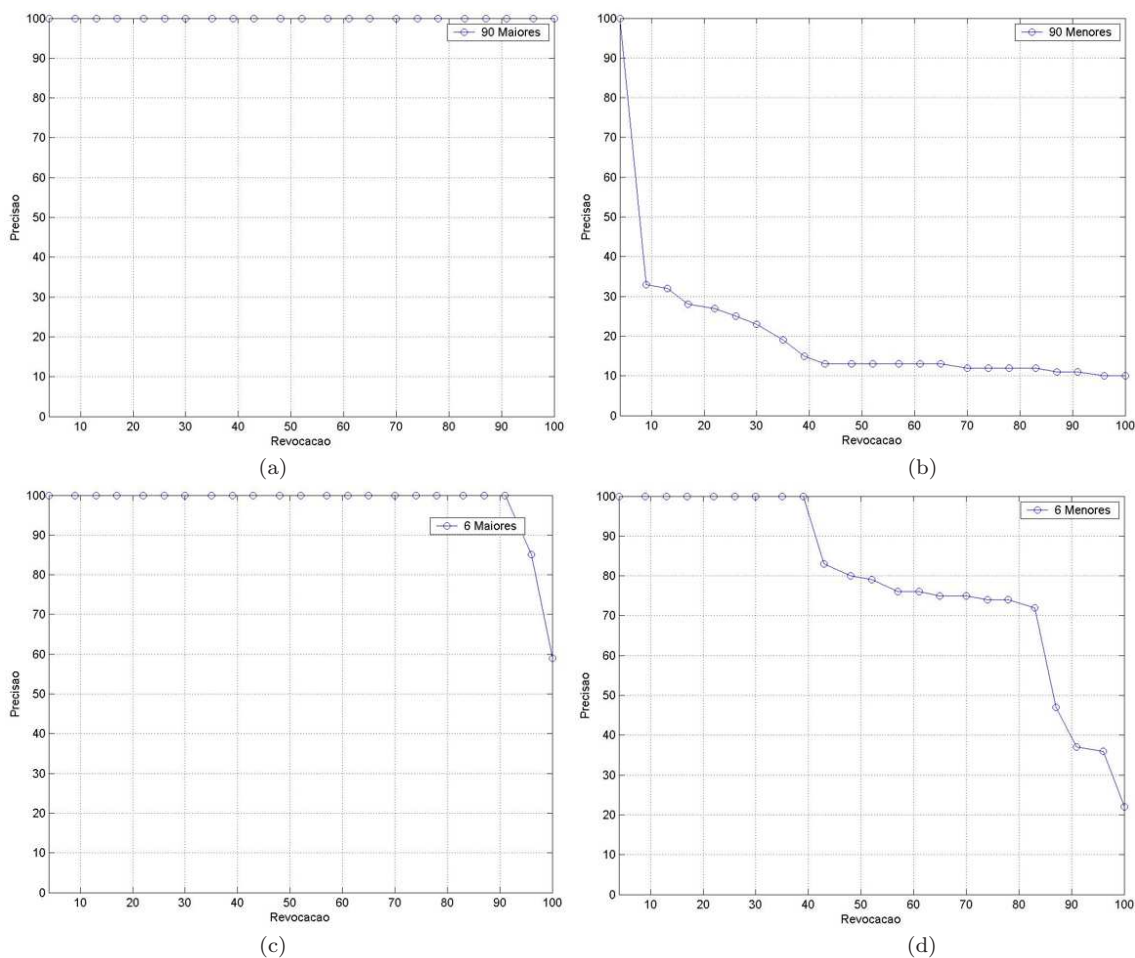


Figura 7.29: Gráficos da precisão em condições normais de iluminação (Imagem consulta: figura 7.19- e) em condições normais de iluminação.

Fonte: Gráficos elaborados pelo autor.

As autôfaces quando aplicadas em situações em que a face consulta possui poucas informações devido à falta de iluminação, mostraram-se sensíveis na maior parte dos casos. A seleção das autôfaces associadas aos autovalores não modificaram o resultado.

Os experimentos realizados nas faces (que tinham poucas informações devido à falta de iluminação) foram executados novamente, porém, essas faces estavam em condições normais de iluminação. É possível observar pelos gráficos uma melhora significativa com essas imagens

como consulta. A partir dos gráficos é observável que as autofaces associadas aos 30 maiores autovalores alcançaram taxas de precisão tão boas quanto usando as 90 maiores. De modo que o uso de menos autofaces “minimiza” os custos computacionais.

7.3.6 Experimento 3: Problema de diferentes poses

Assim como a iluminação, a pose¹ também é um dos desafios encontrados no reconhecimento de face.

Neste experimento visamos verificar o comportamento do método das autofaces em relação à pose. A base de dados ATT (descrita na seção 3.2) é usada nos experimentos. Essa base possui imagens de face de uma mesma pessoa em diferentes poses. Para formar o conjunto de treinamento são selecionadas 5 faces de cada pessoa da base de dados (figura 7.30) e, posteriormente, usadas para gerar as autofaces.



Figura 7.30: Base de dados ATT (imagens usadas no cálculo das autofaces.)
Fonte: Imagens obtidas de [5].

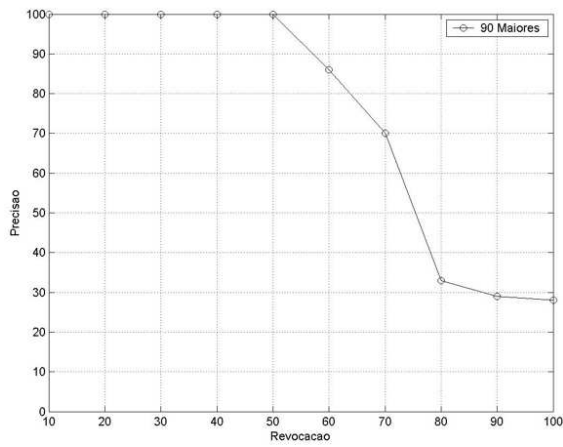
As consultas são realizadas com diferentes quantidades de autovalores e as imagens usadas na consulta (figura 7.31) não são usadas no cálculo das autofaces.



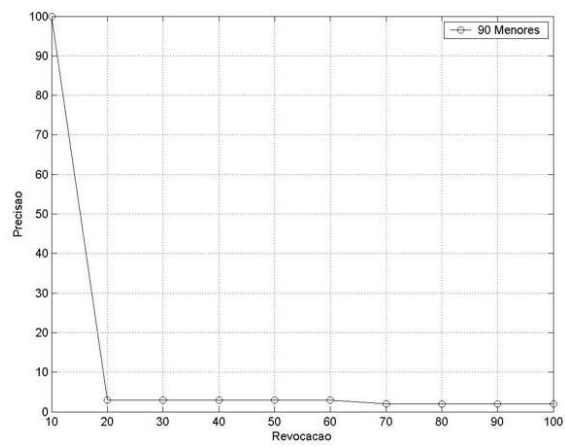
Figura 7.31: Base de dados ATT (Imagens consultas).
Fonte: Imagens obtidas de [5].

O gráfico da precisão e revocação é gerado para cada consulta. O resultado da consulta com a face da figura 7.31 (a) é mostrado nos gráficos 7.32 (a, b, c, d, e, f, g, h).

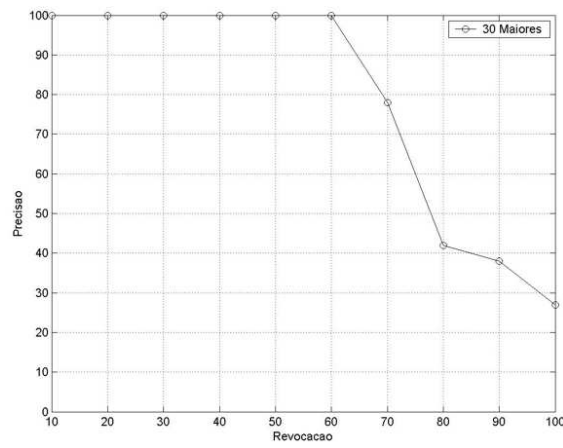
¹Pose é a posição que a face aparece na imagem (de perfil, levemente inclinada, etc).



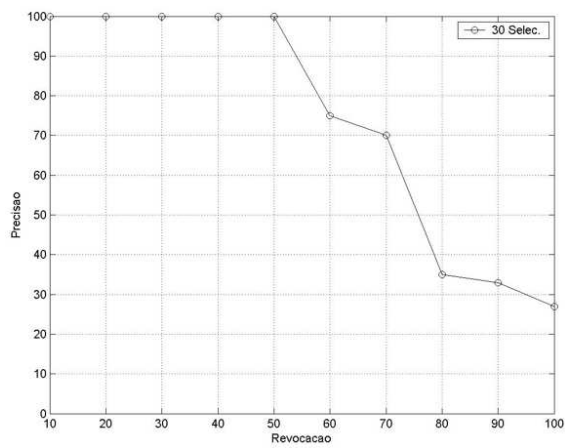
(a)



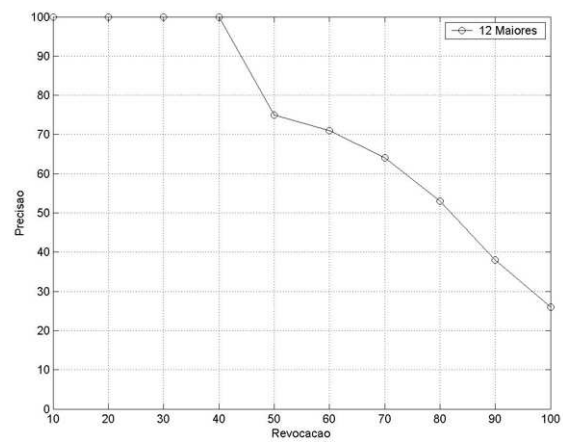
(b)



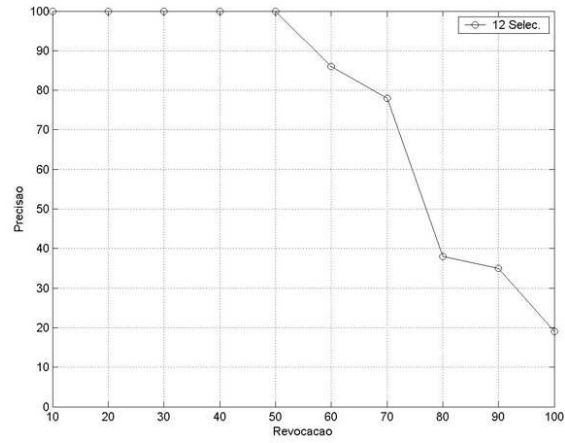
(c)



(d)



(e)



(f)

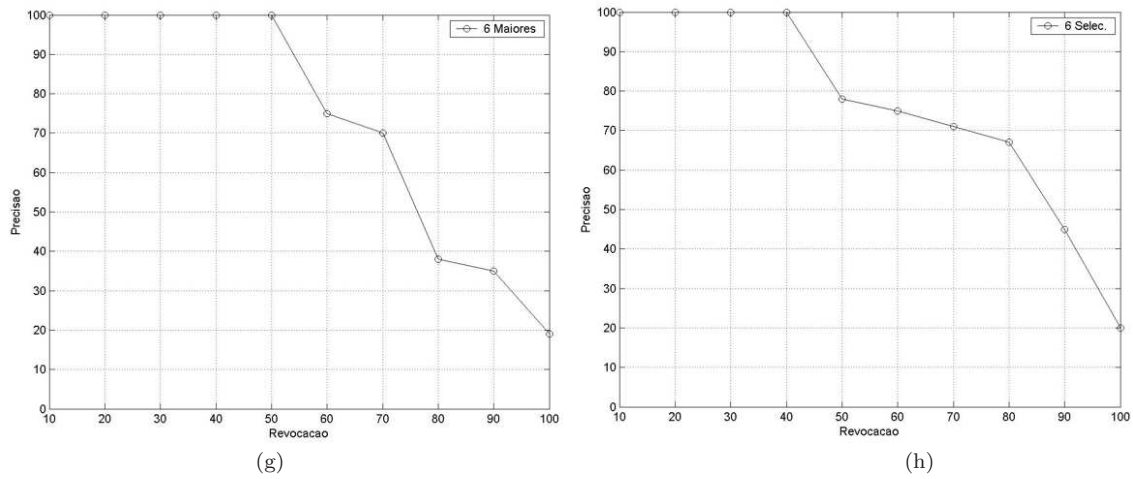
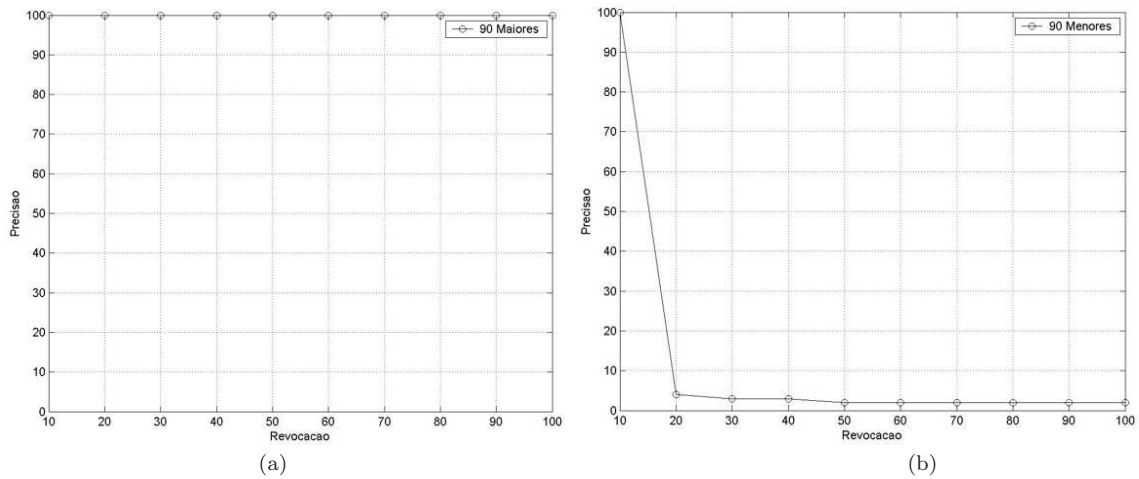
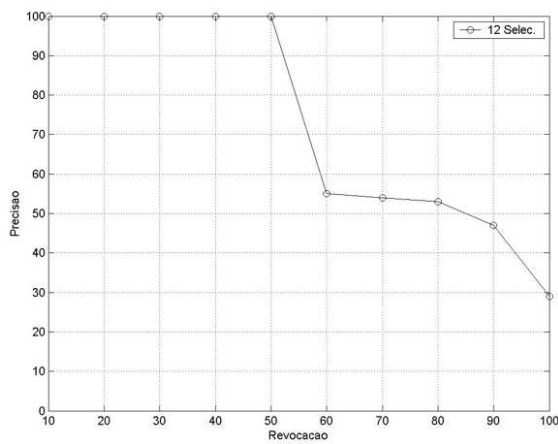


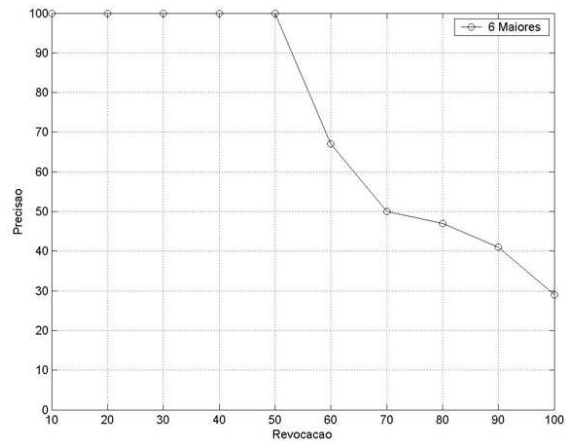
Figura 7.32: Gráficos da precisão: Imagem consulta (figura 7.31- a.)
 Fonte: Gráficos elaborados pelo autor.

O resultado com a face da figura 7.31 (b) é mostrado pelos gráficos 7.33 (a, b, c, d, e). O gráfico (a) representa as autofaces associadas aos 30 maiores e menores e, aos 12 maiores autovalores.

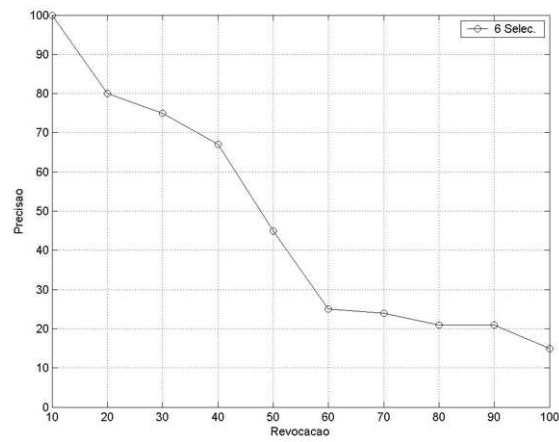




(c)



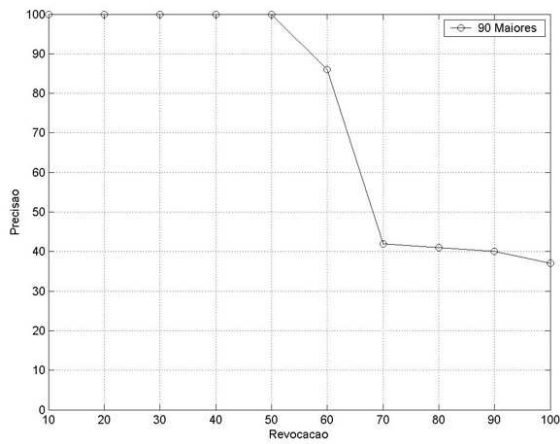
(d)



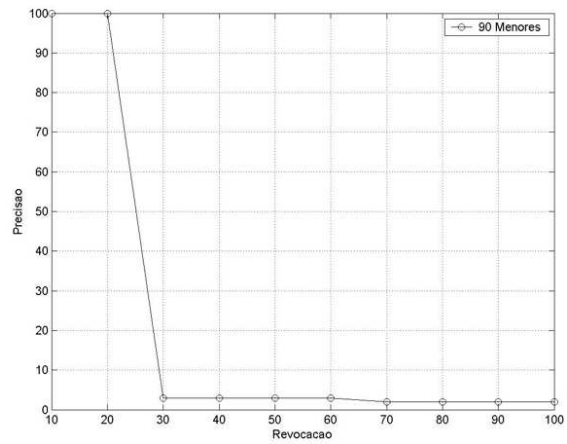
(e)

Figura 7.33: Gráficos da precisão: Imagem consulta (figura 7.31- b.)
Fonte: Gráficos elaborados pelo autor.

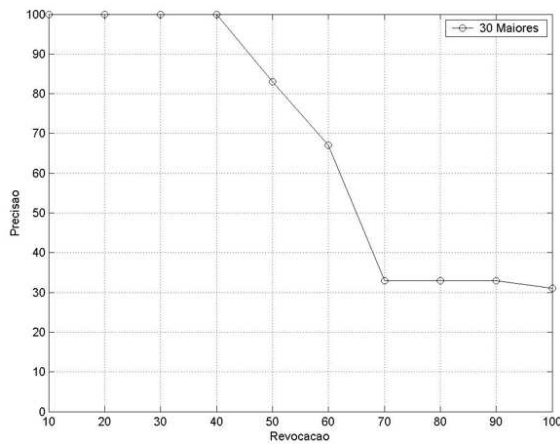
O resultado da consulta com a face da figura 7.31 (c) é mostrado nos gráficos 7.34 (a, b, c, d, e, f, g, h).



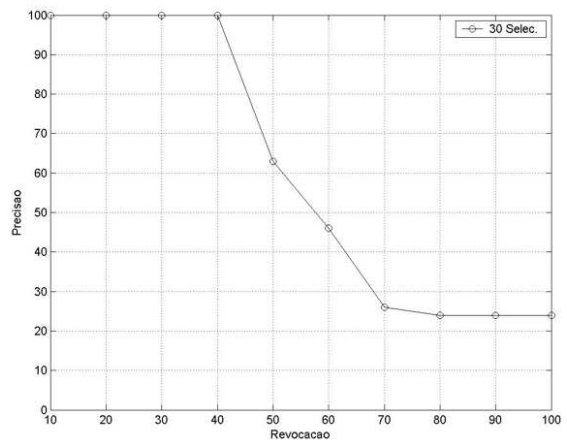
(a)



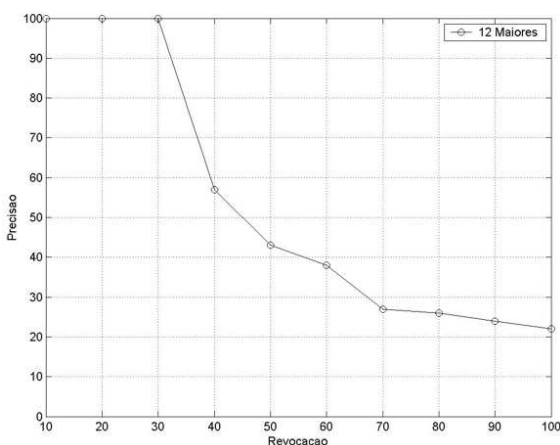
(b)



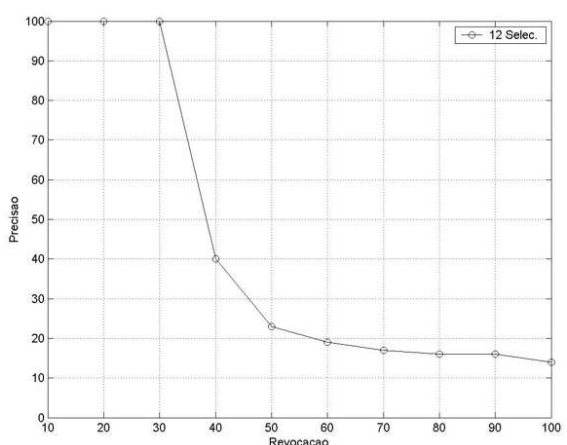
(c)



(d)



(e)



(f)

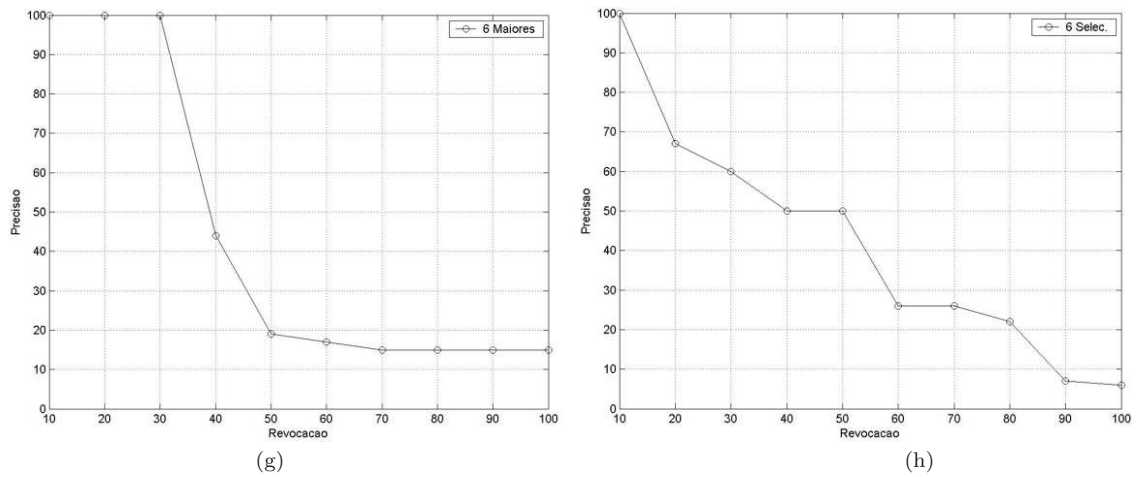
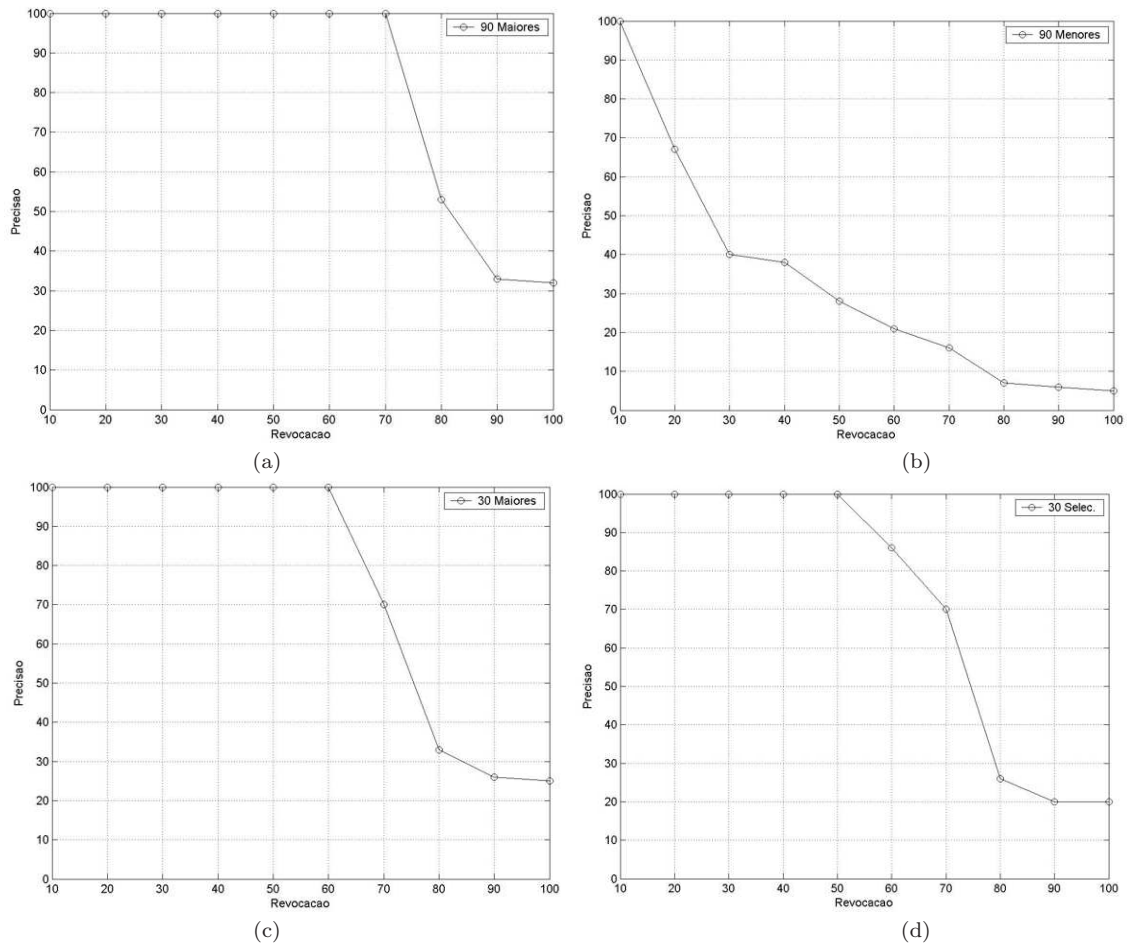


Figura 7.34: Gráficos da precisão: Imagem consulta (figura 7.31- c.)
 Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face da figura 7.31 (d) é mostrado nos gráficos 7.35 (a, b, c, d, e, f, g, h).



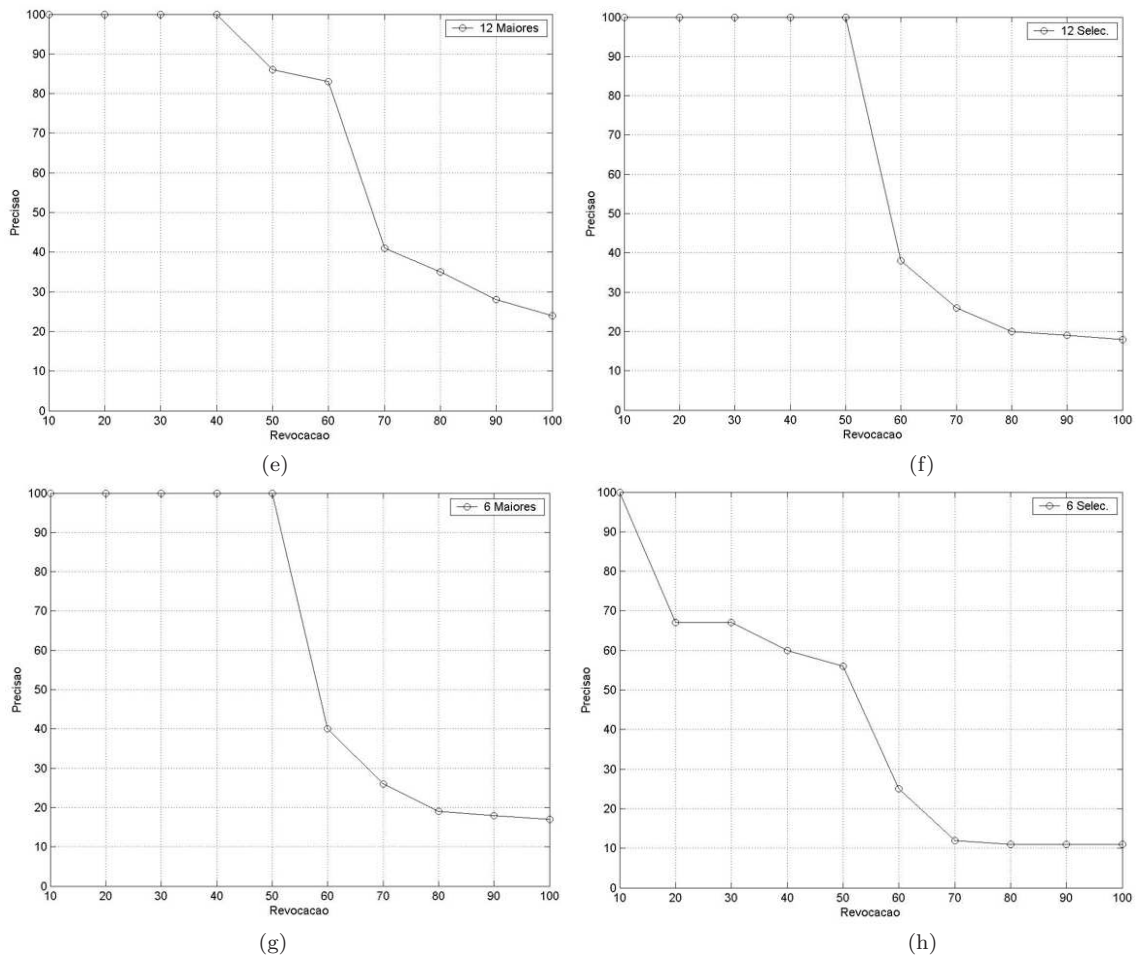


Figura 7.35: Gráficos da precisão: Imagem consultada (figura 7.31- d.)

Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face da figura 7.31 (e) é mostrado nos gráficos 7.36 (a, b, c). O gráfico (a) representa os resultados com os 30 autovalores (maiores e selecionados), com os 12 autovalores (maiores e selecionados) e com os 6 maiores autovalores.

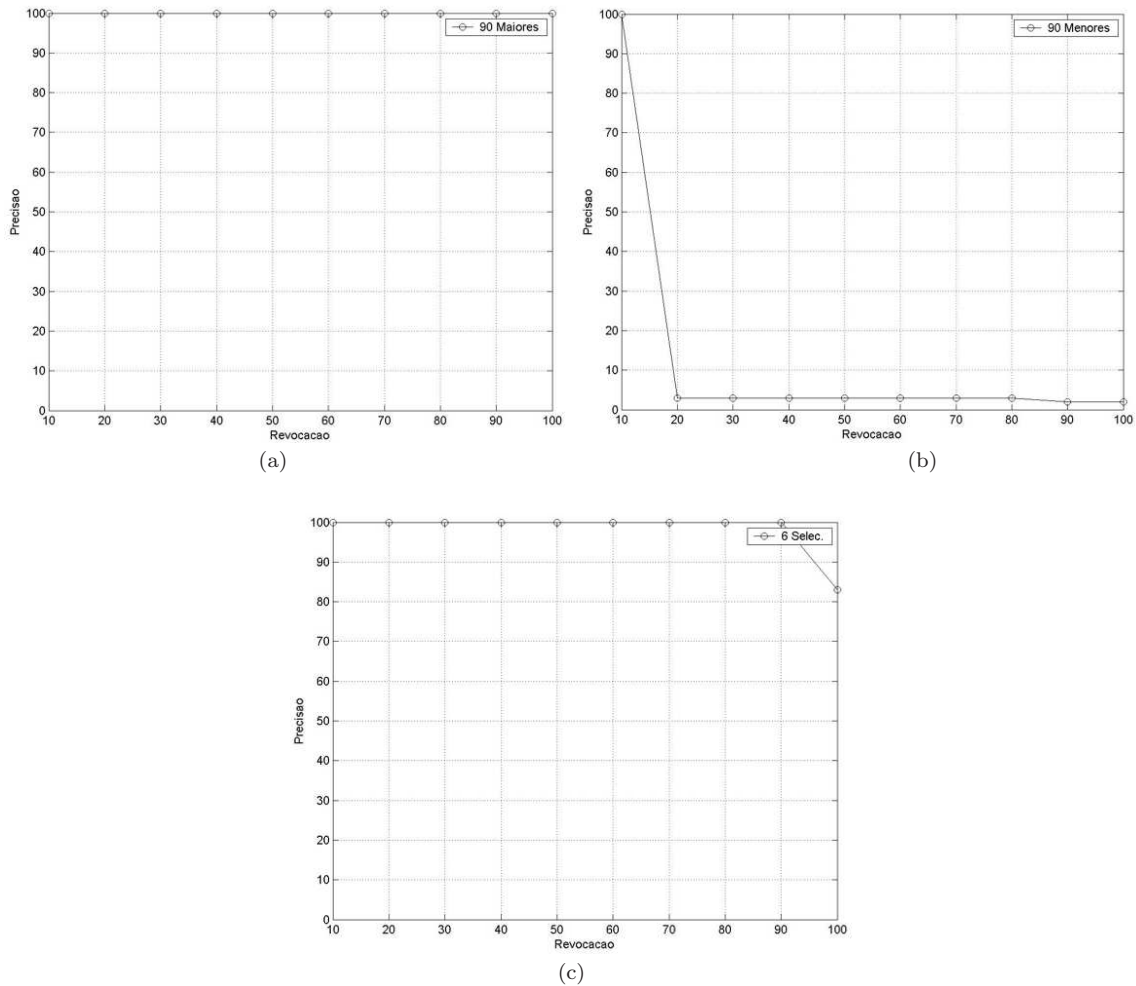


Figura 7.36: Gráficos da precisão: Imagem consulta (figura 7.31- e).
Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face da figura 7.31 (f) é mostrado nos gráficos 7.37 (a, b, c, d). O gráfico (a) representa os resultados com 30 e com 12 autovalores (maiores e selecionados).

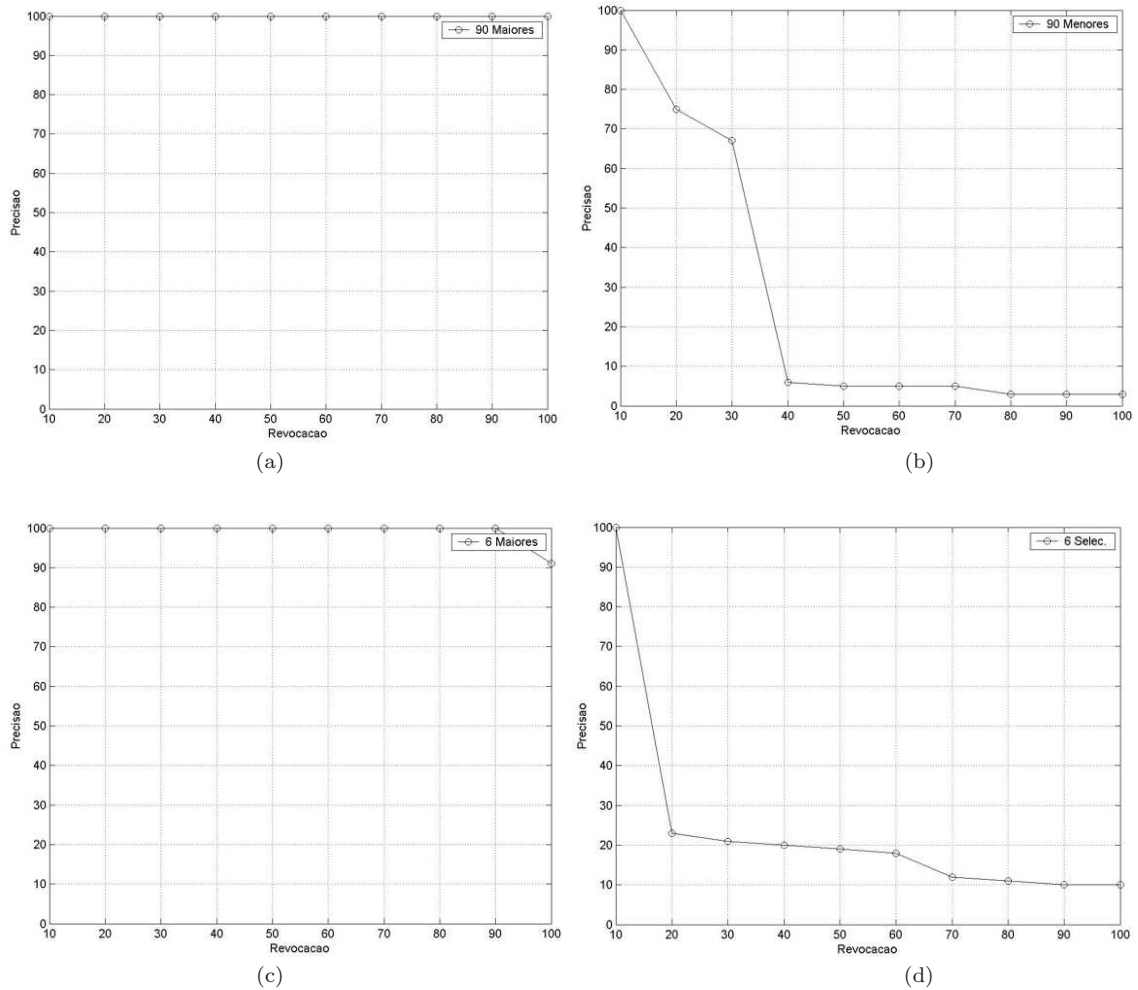
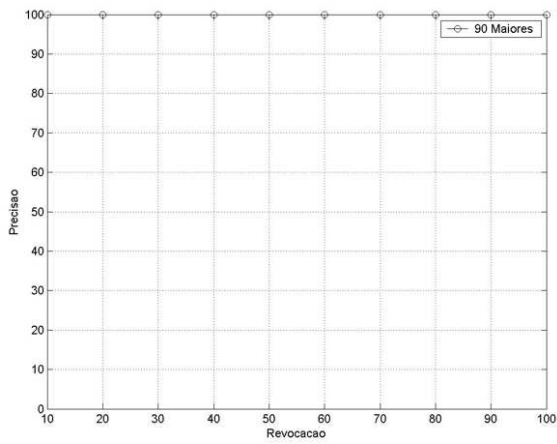
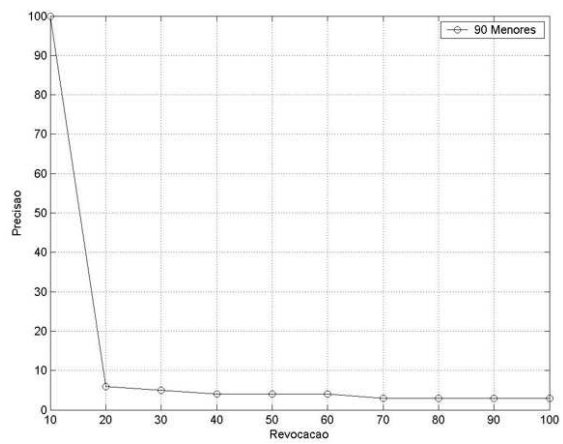


Figura 7.37: Gráficos da precisão: Imagem consulta (figura 7.31- *f*).
Fonte: Gráficos elaborados pelo autor.

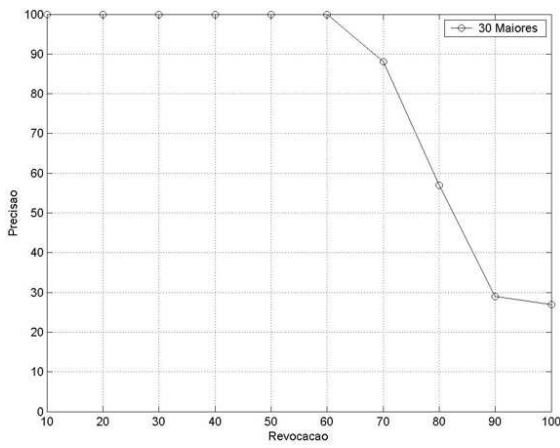
O resultado da consulta com a face da figura 7.31 (*g*) é mostrado nos gráficos 7.38 (*a, b, c, d, e, f, g, h*).



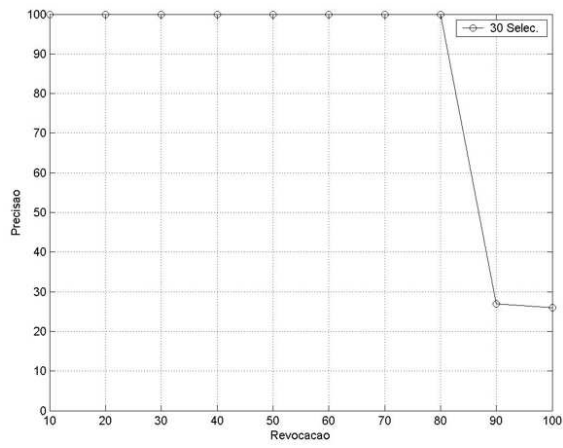
(a)



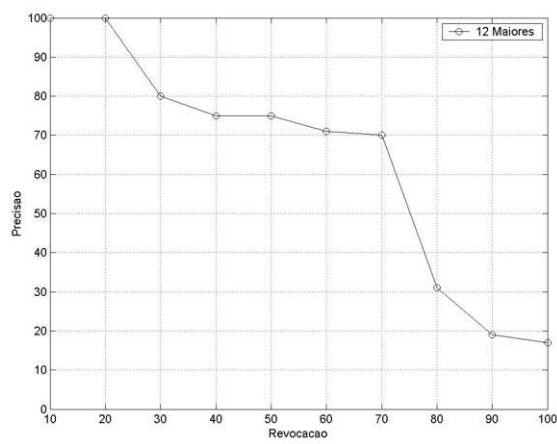
(b)



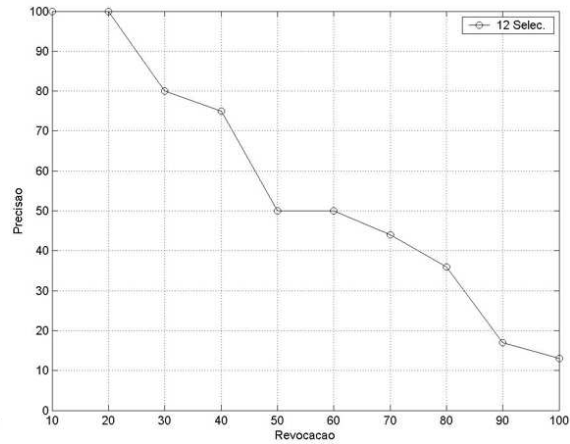
(c)



(d)



(e)



(f)

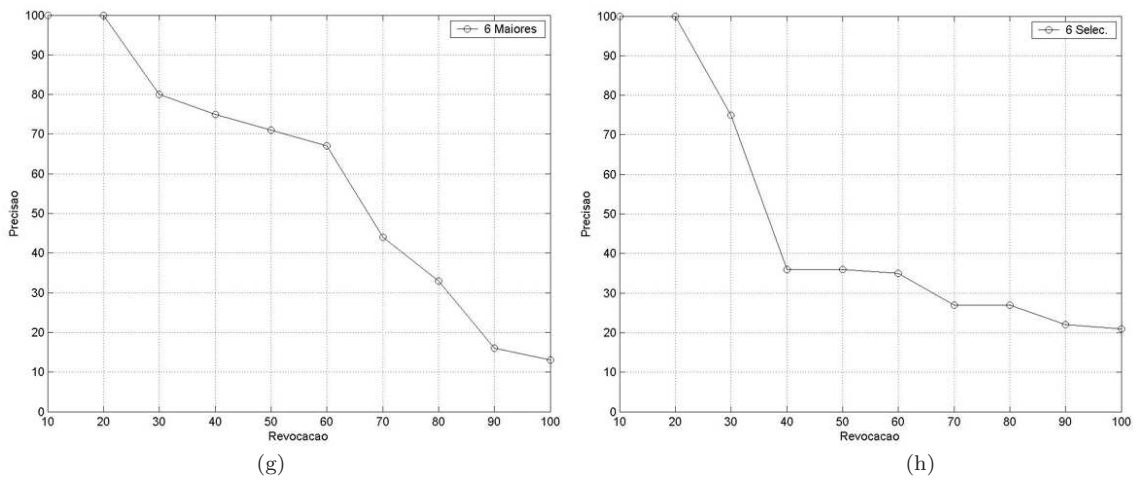
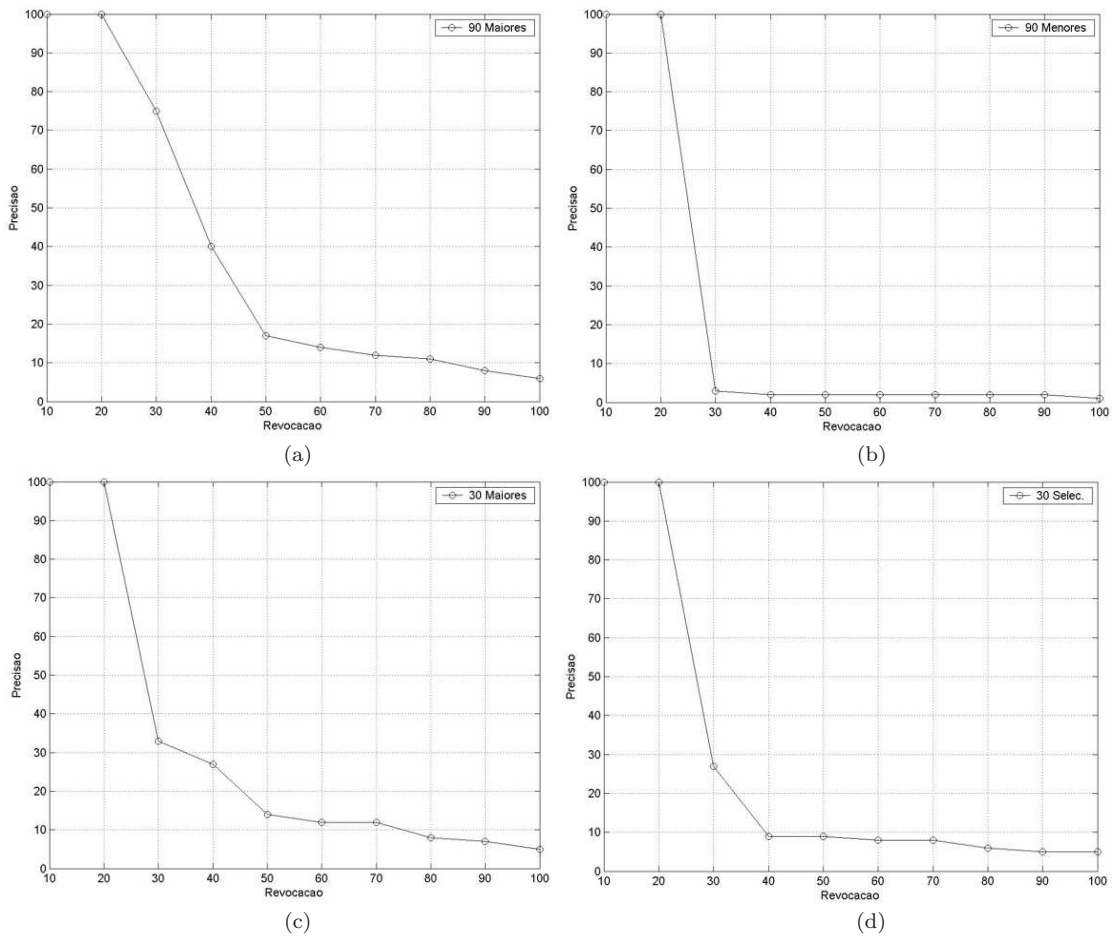


Figura 7.38: Gráficos da precisão: Imagem consultada (figura 7.31- g.)
 Fonte: Gráficos elaborados pelo autor.

O resultado da consulta com a face da figura 7.31 (h) é mostrado nos gráficos 7.39 (a, b, c, d, e, f, g, h).



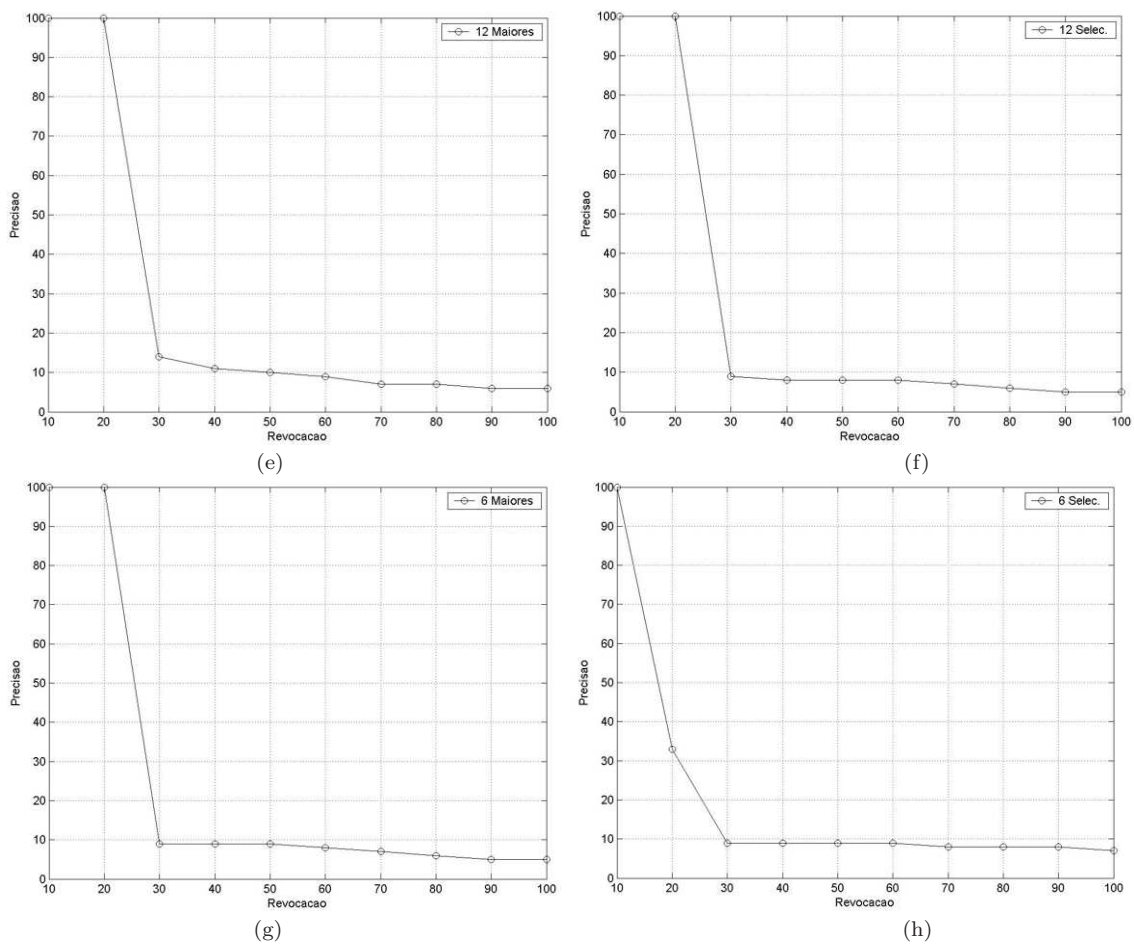


Figura 7.39: Gráficos da precisão: Imagem consulta (figura 7.31- *h*.)

Fonte: Gráficos elaborados pelo autor.

O método das autofaces, quando aplicado em poses, apresenta resultados razoáveis em relação a esse problema. Novamente verificamos que as autofaces relacionadas aos maiores autovalores representam melhores faces do que as autofaces relacionadas aos menores autovalores.

É importante lembrar que as faces usadas na consulta não foram usadas nos cálculos das autofaces e mesmo assim, em algumas situações as autofaces conseguiram recuperá-las nas primeiras posições do *rank* (imagens *b*, *e*, *f*, *g*), embora não tenha conseguido caracterizar a face (*h*).

7.4 Conclusões

Com a utilização das três bases de dados (YaleA, YaleB e ATT) cada problema encontrado pode ser verificado com mais eficácia (expressão facial, iluminação e pose). De acordo com os resultados obtidos é possível deduzir que as autofaces relacionadas aos maiores autovalores caracterizam melhor o formato da cabeça e as características exteriores, enquanto que as autofaces de menores autovalores possivelmente identificam as características internas. Isso pode ser visualizado pelos gráficos de precisão nos experimentos de expressão facial, em que os menores autovalores recuperaram faces com deformações nas características externas enquanto que foi prejudicado quando as deformações eram nas características internas.

No caso da iluminação o reconhecimento é prejudicado quando há pouca incidência de luz na face, como pode ser evidenciado pelo experimento 7.3.5. Contudo, isso já era de se esperar, visto que o processamento do PCA é holístico e as informações das faces são “escondidas” pelas sombras causadas pela falta de iluminação. Isso pode ser melhor compreendido pelos resultados mostrados pelos gráficos referentes às consultas que não eram prejudicadas pela falta de iluminação.

Em relação à pose, os resultados permaneceram próximos um dos outros, com exceção dos 90 menores autovalores que tiveram as menores taxas de recuperação. Assim como na iluminação é muito difícil determinar a quantidade ou quais autovalores devem ser selecionados para melhorar o reconhecimento de face.

8 O uso do PCA nas componentes de detalhes das imagens

8.1 Introdução

A imagem de uma mesma face pode apresentar diferentes aspectos devido à influência de fatores, como por exemplo: posição da face, variação da iluminação, variação da idade, etc. Estas influências geram grandes dificuldades no processo de reconhecimento e, ao longo dos anos, vêm sendo propostos diversos algoritmos para tratar tal problema [14, 39, 46, 51, 54]. Mesmo com o progresso obtido até hoje, as técnicas existentes ainda se mostram frágeis quando aplicadas a problemas gerais.

Dentre os problemas que prejudicam o desempenho dos sistemas de reconhecimento de face, a variação da iluminação é a que têm maior impacto. Um exemplo pode ser visto na figura 8.1, em que uma pessoa com a mesma expressão facial observada do mesmo ponto de vista parece diferente quando fontes de iluminação de mesma intensidade advinda de pontos distintos incidem sob a face.

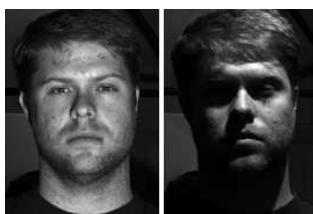


Figura 8.1: Efeito da iluminação incidente na face.
Fonte: Imagem obtida de [4].

Uma maneira de solucionar o problema da iluminação seria encontrar quais características da face são invariantes às mudanças da iluminação. Contudo, identificar essas características é um problema muito difícil. Uma alternativa proposta por vários pesquisadores é

construir sistemas que tentam modelar as variações de iluminação ao invés de tentar sistemas que sejam invariantes a elas [24, 74, 75].

Uma técnica usada na modelagem das variações de iluminação é aplicar o PCA não nas imagens originais, mas em características, ou seja, em imagens processadas via algum objetivo específico. Por exemplo, Zhao e Chellappa [71] desenvolveram um método cujo objetivo é obter uma imagem protótipo com iluminação normalizada utilizando o método *shape-from-shading* (SFS) em seguida, aplicar o PCA nessas imagens protótipos para realizar o reconhecimento de face. Xie e Lam [74] propuseram aplicar o PCA na imagem de face cujo fator de iluminação foi modificado em cada um dos triângulos da face triangularizada utilizando um método de normalização local (LN), outras abordagens com esse mesmo objetivo podem ser encontradas em [8, 36, 75].

Neste trabalho vamos aplicar o PCA nos detalhes de uma face. Cada imagem de face é decomposta em duas componentes: a componente de detalhes e a componente suavizada, e o PCA é então aplicado nas versões de detalhes de cada imagem do banco de dados. Para isto é efetuado uma etapa de pré-processamento, em cada imagem do banco de dados, que consiste em obter as duas componentes da imagem original. Basicamente, podemos considerar que cada imagem é uma soma de duas partes ou duas componentes, em que uma contém a parte suavizada (níveis de cinzas, sombras, etc.) e a outra contém as texturas, os ruídos e os detalhes. O objetivo é então aplicar o PCA nas componentes de detalhes que contém as informações que, nas imagens originais, estão “escondidas” pela falta de iluminação.

O método de suavização com preservação de bordas de Rudin, Osher e Fatemi ou modelo de ROF [42], foi o usado para pré-processar as imagens. Esse método extrai das imagens sua parte suavizada e, a parte de texturas, ruídos e detalhes é a componente que falta para compor a imagem original. O PCA foi então usado nas componentes de detalhes das imagens para gerar as autofaces. Outros métodos de suavização com preservação de bordas poderiam ser usados para obter a componente suavizada.

8.2 Descrição do Método Proposto

O método proposto tem como objetivo capturar as informações “perdidas” pela ausência de iluminação na face usando um modelo de preservação de bordas [42]. O método consiste em aplicar uma suavização com preservação de bordas em todas as imagens do banco de dados, f_i , obtendo imagens suavizadas, $u_i, i = 1, 2, \dots, N$. Métodos com preservação de bordas

têm a habilidade de calcular aproximações suavizadas u de f , preservando detalhes importantes como as bordas, as quinas, etc., enquanto ruídos ou pequenos padrões repetidos são removidos da imagem. A seguir, aplica-se o método PCA no conjunto $V = \{v_1, v_2, \dots, v_N\}$, em que $v_i = f_i - u_i$. A componente v_i contém os detalhes, ruídos e texturas.

Assim, dado uma imagem f , podemos decompô-la e obter outra imagem u , tal que u é uma simplificação de f . Em general, u é uma imagem formada por regiões homogêneas. Podemos assumir a seguinte relação entre f e u : $f = u + v$, em que a componente v representa as texturas, ruídos e detalhes da imagem original f . Em muitas aplicações, a componente v não é armazenada, pois ela contém os ruídos [42,43], enquanto que em outras esta componente v é importante, especialmente se ela representa textura. De acordo com Meyer [52] é possível extrair de f ambos os componentes u e v . A decomposição obtida pode ser útil para a segmentação e discriminação de textura, dentre diversas outras aplicações possíveis.

Neste trabalho usamos o modelo de preservação de bordas de ROF [42], obtido via minimização de energia e equações diferenciais parciais. Assim, considerando $f : R^2 \rightarrow R$ a imagem original, uma relação entre f e u pode ser expressa por um modelo linear, apresentando uma outra função v , tal que

$$f(x, y) = u(x, y) + v(x, y) \quad (8.1)$$

No modelo de Rudin et al. (equação 8.2), $v = f - u$ representa os detalhes, ruídos e textura enquanto u é uma imagem formada por regiões homogêneas. Dado f , tanto u como v são desconhecidos. Nessa abordagem, o problema da reconstrução de u a partir de f é dada pela minimização do funcional

$$F(u) = \int_{\Omega} |\nabla u| + \lambda \int_{\Omega} (f - u)^2 \, dx dy \quad (8.2)$$

em que ∇u é o gradiente de u , f a imagem original e λ o fator de ajuste ou termo de regularização.

O que leva à seguinte equação de Euler Lagrange:

$$2\lambda(u - f) = \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) \quad (8.3)$$

Então, a função que representa detalhes, ruídos e texturas é $v = f - u = -\frac{1}{2\lambda} \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right)$

mas isso não é calculado explicitamente no modelo de ROF, que calcula a imagem u . A componente v é, então, obtida via $f - u$.

A discretização do $\operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right)$ foi feita como em [43], ou seja, no ponto (ih, jh) o divergente foi aproximado por

$$\begin{aligned} \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) &\cong \left[\frac{u_{i+1,j} - u_{i,j}}{\sqrt{\left(\frac{u_{i+1,j} - u_{i,j}}{h}\right)^2 + \left(\frac{u_{i,j+1} - u_{i,j}}{2h}\right)^2}} - \frac{u_{i,j} - u_{i-1,j}}{\sqrt{\left(\frac{u_{i,j} - u_{i-1,j}}{h}\right)^2 + \left(\frac{u_{i-1,j+1} - u_{i-1,j}}{2h}\right)^2}} \right] \\ &+ \frac{1}{h^2} \left[\frac{u_{i,j+1} - u_{i,j}}{\sqrt{\left(\frac{u_{i+1,j} - u_{i-1,j}}{2h}\right)^2 + \left(\frac{u_{i,j+1} - u_{i,j}}{h}\right)^2}} - \frac{u_{i,j} - u_{i,j-1}}{\sqrt{\left(\frac{u_{i+1,j-1} - u_{i-1,j-1}}{2h}\right)^2 + \left(\frac{u_{i,j} - u_{i,j-1}}{h}\right)^2}} \right] \end{aligned} \quad (8.4)$$

Para resolver numericamente foi introduzido a seguinte equação linearizada (como em [43])

$$\begin{aligned} u_{i,j}^{n+1} &= f_{i,j} + \frac{1}{2\lambda h^2} \left[\frac{u_{i+1,j}^n - u_{i,j}^{n+1}}{\sqrt{\left(\frac{u_{i+1,j}^n - u_{i,j}^n}{h}\right)^2 + \left(\frac{u_{i,j+1}^n - u_{i,j}^n}{2h}\right)^2}} - \frac{u_{i,j}^{n+1} - u_{i-1,j}^n}{\sqrt{\left(\frac{u_{i,j}^n - u_{i-1,j}^n}{h}\right)^2 + \left(\frac{u_{i-1,j+1}^n - u_{i-1,j}^n}{2h}\right)^2}} \right] \\ &+ \frac{1}{2\lambda h^2} \left[\frac{u_{i,j+1}^n - u_{i,j}^{n+1}}{\sqrt{\left(\frac{u_{i+1,j}^n - u_{i-1,j}^n}{2h}\right)^2 + \left(\frac{u_{i,j+1}^n - u_{i,j}^n}{h}\right)^2}} - \frac{u_{i,j}^{n+1} - u_{i,j-1}^n}{\sqrt{\left(\frac{u_{i+1,j-1}^n - u_{i-1,j-1}^n}{2h}\right)^2 + \left(\frac{u_{i,j}^n - u_{i,j-1}^n}{h}\right)^2}} \right] \end{aligned} \quad (8.5)$$

As imagens suavizadas, u_i são subtraídas das imagens originais, f_i e, as imagens resultantes dessa subtração contém texturas, ruídos e os detalhes e são usadas para gerar as autofaces.

Um exemplo da aplicação do filtro de suavização na decomposição de face é mostrada na figura 8.2, em que são apresentadas duas imagens do banco de dados, uma com condições normais de iluminação e outra com pouca iluminação, bem como suas componentes u_i e v_i .

As imagens (a.1 e b.1) são as imagens originais; as imagens (a.2, e b.2) são as componentes, v_i 's, contendo os detalhes, texturas e ruídos; e, as imagens (a.3 e b.3) são as componentes suavizadas, u_i 's. Podemos observar que a imagem (b.2) apresenta detalhes que não são observados na imagem original (b.1), como por exemplo, alguns traços da boca, do nariz e do olho esquerdo, devido as poucas condições de iluminação¹.

Dado um banco de dados com N imagens, tomando M ($M \leq N$) imagens de treinamento

¹Os detalhes da boca, nariz e olho da imagem b.2 são mais visíveis na tela do computador, pois as imagens perdem qualidade na impressão.

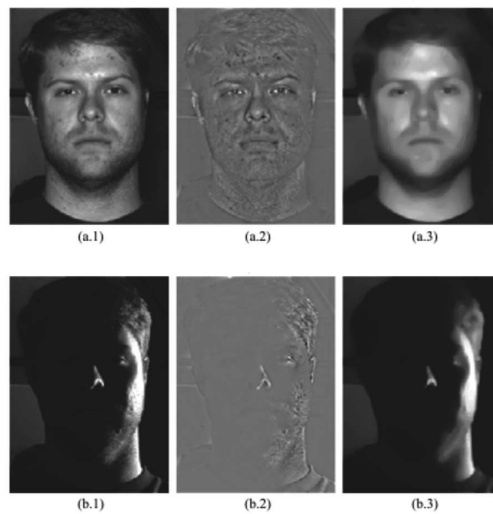


Figura 8.2: Filtro de suavização aplicado em uma imagem de face.
 Fonte: Imagens elaboradas pelo autor.

f_1, f_2, \dots, f_M , o PCA é, então, executado no conjunto $V = \{v_1, v_2, \dots, v_M\}$, que contém os detalhes, as texturas e os ruídos das imagens originais, f_i . Na figura 8.3 ilustramos as autofaces geradas pelo PCA quando aplicado nas imagens originais f_i (figura 8.3 (a)) e as autofaces geradas pelo PCA aplicado nas imagens com texturas, ruídos e detalhes v_i (figura 8.3 (b)).

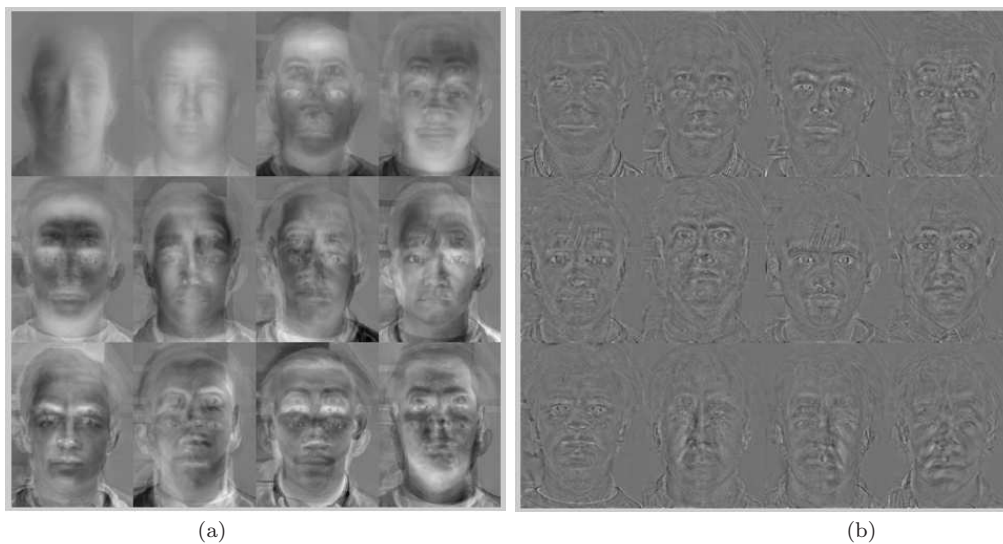


Figura 8.3: Autofaces geradas a partir das imagens originais (a) e das imagens do conjunto $V = \{v_1, v_2, \dots, v_N\}$, (b).
 Fonte: Imagem elaborada pelo autor.

Para ilustrar o desempenho do método proposto vamos comparar resultados obtidos quando o PCA é aplicado nas imagens originais, f_1, f_2, \dots, f_M , aqui denominado **PCA (Padrão)** com resultados obtidos usando o PCA aplicado no conjunto V aqui denominado **PCA (Detalhes)**. A principal diferença entre os métodos está nas imagens utilizadas para gerar as autofaces. Enquanto no PCA (padrão) as imagens usadas para gerar as autofaces são as imagens originais, f_i , no PCA (Detalhes) as imagens utilizadas são as do conjunto V .

8.3 Resultados Experimentais

Nesta seção, o PCA (Detalhes) é analisado com a utilização das imagens de face da base de dados YaleB [4] (capítulo 3 seções 3.3.3) e comparado com o PCA (Padrão). A base de dados YaleB possui $N = 640$ imagens de faces sob diversas condições de iluminação. As imagens são selecionadas de acordo com o azimute (representado pela letra A) e o ângulo de elevação da câmera (representado pela letra E). Azimute positivo significa foco luminoso na direita e elevação positiva significa foco luminoso acima da cabeça. Por exemplo, uma face em que o foco luminoso está na frente do seu rosto tem azimute e elevação 0° , isto é, $A + 000E + 000$.

Para os experimentos um conjunto de treinamento foi formado com 23 imagens de cada pessoa (num total de 64) sob condições normais de iluminação e expressão facial (com azimute de -5° à $+25^\circ$ e elevação de -1° à 0°) num total de $M = 230$ imagens (figura 8.4).

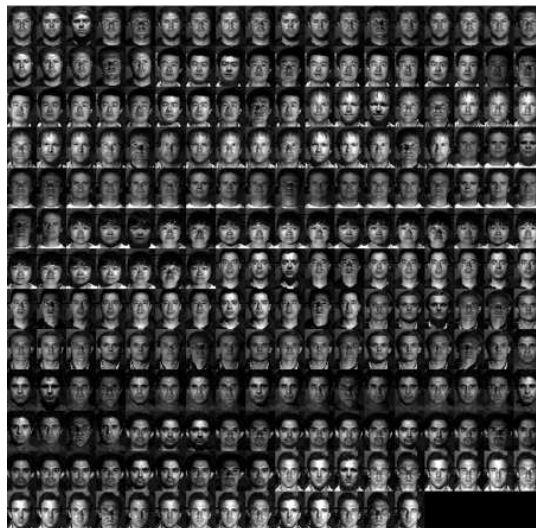


Figura 8.4: Imagens da base de dados YaleB usadas para formar o conjunto de treinamento.
Fonte: Imagem obtida de [4].

Três imagens pertencentes ao banco de dados e não no conjunto de treinamento foram selecionadas como consulta. Essas imagens selecionadas são apresentadas na figura 8.5 e possuem diferentes iluminações.

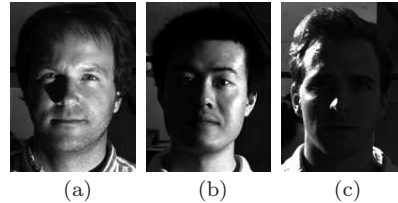


Figura 8.5: Imagens da base de dados YaleB usadas na consulta.
Fonte: Imagem obtidas em [4].

8.3.1 PCA (Detalhes) versus PCA (Padrão)

Para medir a precisão das consultas com o PCA (Detalhes) e com o PCA (Padrão) usamos o gráfico da precisão e revocação explicados no capítulo 7. A partir desse gráfico podemos comparar o desempenho de cada método. Os resultados serão apresentados em três experimentos de acordo com a intensidade da iluminação.

No primeiro experimento a consulta foi realizada com a figura 8.5(a) pouco prejudicada pela ausência de iluminação, onde todas as informações da face são visíveis com exceção da orelha esquerda. Neste caso, temos que os métodos, se analisarmos até 15% de revocação, apresentaram o mesmo desempenho (figura 8.6). Contudo, o desempenho do método proposto manteve-se constante até, aproximadamente, os 60% de revocação, enquanto que o nível de precisão e revocação do outro método apresentou uma queda. Como o PCA (Detalhes) é aplicado nas imagens de texturas, ruídos e detalhes, a quantidade de informações extraídas da face, lembrando que a face consulta tem pouca influência da iluminação, são maiores que as informações extraídas da imagem original. Mesmo tendo os métodos apresentados boas taxas de precisão, o PCA (Detalhes) conseguiu caracterizar melhor as faces, atingindo melhores resultados. Na figura 8.7 são apresentados mais resultados usando como consultas faces de outras pessoas.

No segundo experimento as informações da face começam a ser “perdidas” dificultando o seu reconhecimento (figura 8.5 (b)). Nesta imagem vemos que apenas parte do rosto está visível. Neste caso, temos que o nível de precisão caiu bastante em relação à consulta anterior. Fato este ocasionado pela ausência da iluminação incidida na face. O PCA (Detalhes), mesmo tendo uma queda de desempenho a partir de 45% de revocação) traz nas posições

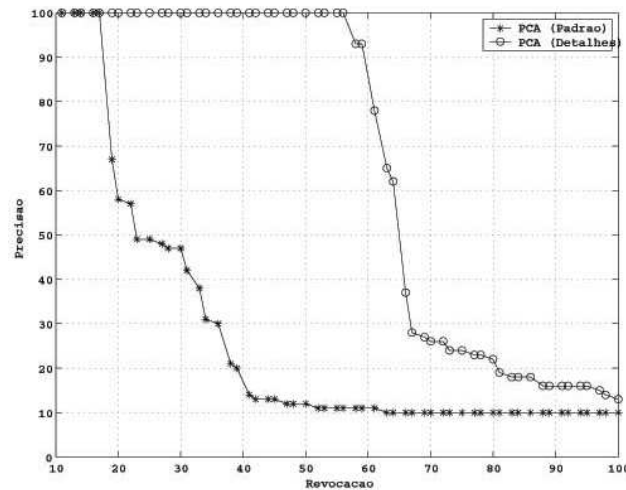


Figura 8.6: Precisão vs Revocação da imagem consulta (a).
Fonte: Gráfico elaborada pelo autor.

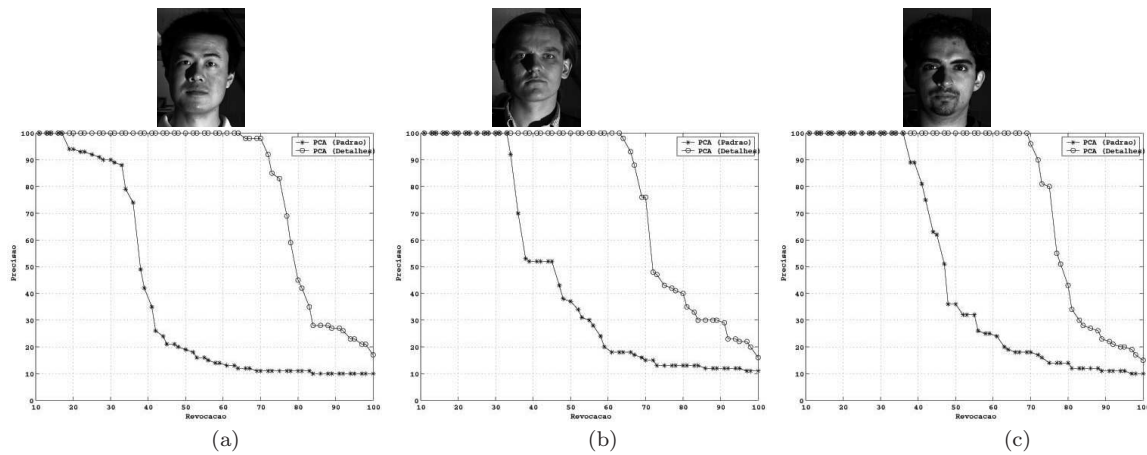


Figura 8.7: Precisão vs Revocação de outras faces de mesma intensidade de iluminação da face da figura 8.6.
Fonte: Imagem obtidas em [4].

mais relevantes (até 40% de revocação) maior quantidade de informações do que o PCA (Padrão)(figura 8.8). Na figura 8.9 são apresentados mais resultados usando faces de outras pessoas.

O terceiro experimento ilustra a situação onde não há muita informação disponível na face, devido à falta de iluminação, observa-se que o reconhecimento fica extremamente prejudicado. Mesmo assim o desempenho do PCA (Detalhes) é melhor que o do PCA (Padrão)(figura 8.10) até 50% de revocação tendo, a partir daí, praticamente o mesmo desempenho. Na figura 8.11 são apresentados mais resultados usando faces de outras pessoas.

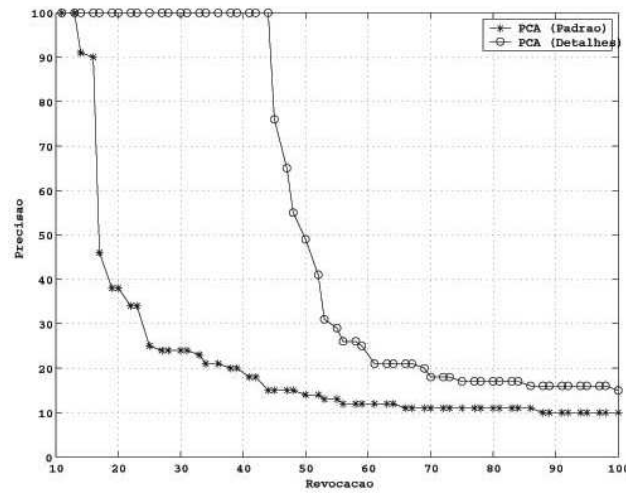


Figura 8.8: Precisão vs Revocação: Gráfico usando a imagem consulta (b) : $A + 070E + 000$.
Fonte: Gráfico elaborada pelo autor.

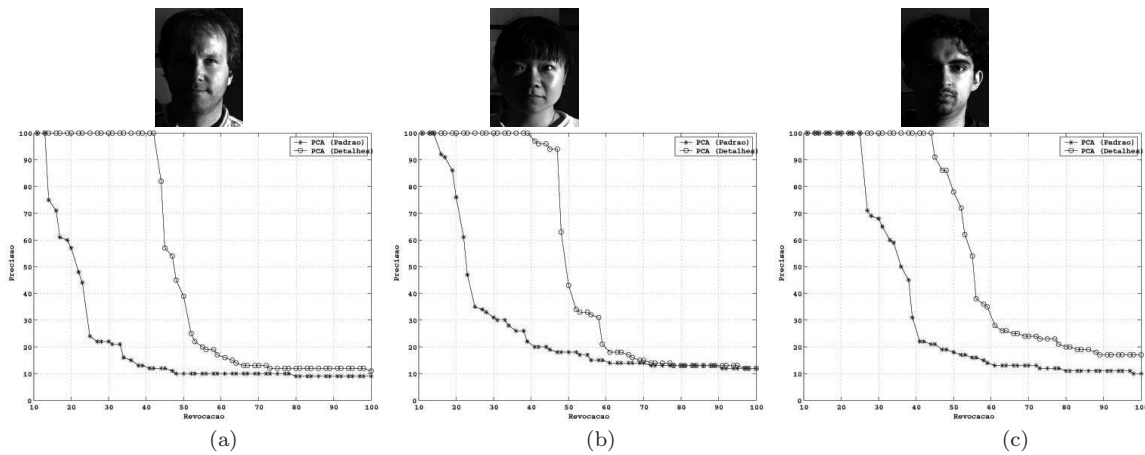


Figura 8.9: Precisão vs Revocação de outras faces de mesma intensidade de iluminação da face da figura 8.8.
Fonte: Imagem obtidas em [4].

Com os resultados mostrados é possível verificar que o PCA (Detalhes) obteve melhor desempenho que o PCA (Padrão). As informações provenientes da face, que antes estavam escondidas pela falta de iluminação, são melhor caracterizadas depois de aplicar o filtro de suavização com preservação de bordas [42] nas imagens de face. Se considerarmos até 30% de revocação para todas as consultas pode-se ver que o método proposto obteve quase que o dobro de precisão em relação ao PCA (Padrão).

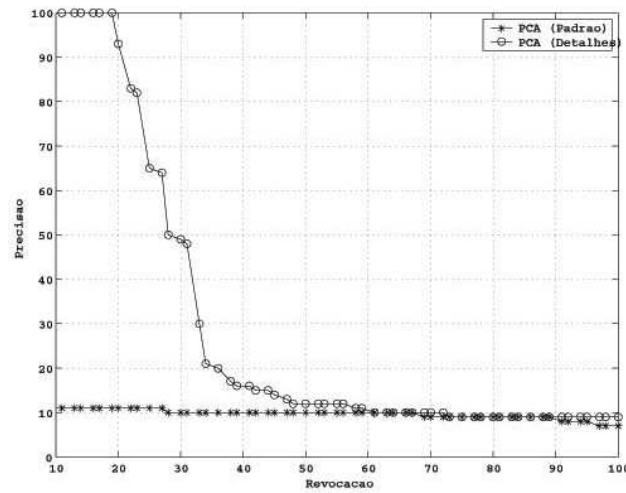


Figura 8.10: Precisão vs Revocação: Gráfico usando a imagem consulta
(c) : $A + 130E + 020$.

Fonte: Gráfico elaborada pelo autor.

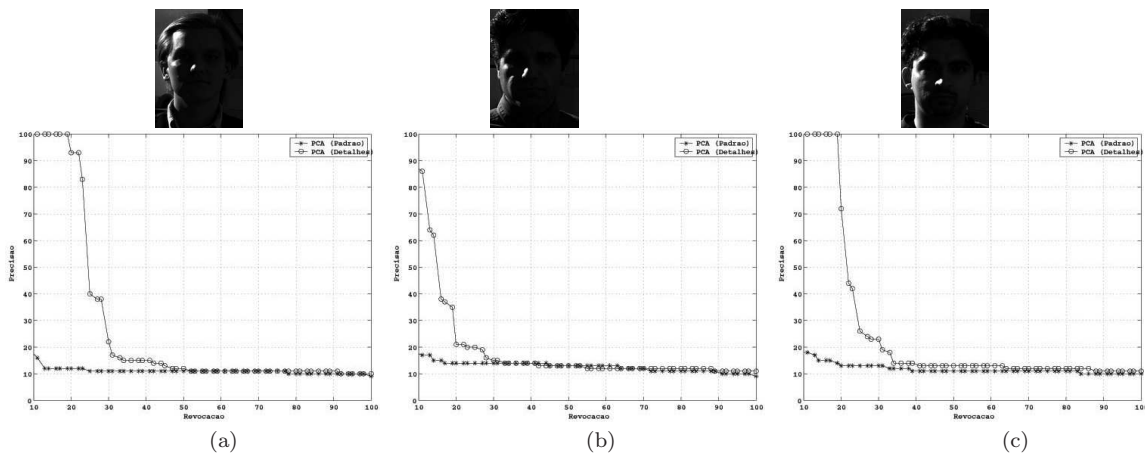


Figura 8.11: Precisão vs Revocação de outras faces de mesma intensidade de iluminação da face da figura 8.10.

Fonte: Imagem obtidas em [4].

8.3.2 PCA (Detalhes) versus Outros Métodos

Como vimos na seção anterior, o PCA (Detalhes) mostrou-se menos sensível às diferentes condições de iluminação do que o PCA (Padrão). Nesta seção o PCA (detalhes) é comparado com diversos métodos, tais como, o método PCA, equalização de histograma (HE) e o ICA, além da normalização local (LN) proposta por Xie e Lam [74]. No método de Xie e Lam as imagens de face passam por um processo de triangularização de modo que o efeito da iluminação em cada triângulo é modelado por um ruído multiplicativo e um ruído aditivo.

Em seguida, a técnica LN [74] é aplicada na imagem com objetivo de eliminar os efeitos da iluminação.

Para efeito de comparação os experimentos realizados utilizaram as mesmas condições do trabalho de Xie e Lam [74], isto é, foi utilizada a base de dados YaleB em que todas as imagens foram cortadas nas dimensões 64×64 pixels e alinhadas, manualmente, em relação aos olhos.

A quantidade de imagens usadas para formar o conjunto de treinamento e o número de imagens do conjunto teste também são as mesmas e estão apresentadas na tabela 8.1, obtida de [74].

	YaleB
Conjunto de Treinamento	10
Conjunto Teste	640

Tabela 8.1: Distribuição das imagens do conjunto de treinamento e de teste.
Fonte: Tabela obtida de [74].

Os resultados apresentados por [74] juntamente com os resultados obtidos pelo método proposto são apresentados na tabela 8.2.

Método	YaleB
PCA(Padrão)	60,3%
HE	63,3%
ICA	65,6%
PCA(Detalhes)	77,8%
LN	99,5%

Tabela 8.2: Comparação do PCA(Detalhes) com o trabalho de [74] usando a base de dados YaleB.

Fonte: Tabela obtida pelo autor.

A partir dos dados provenientes da tabela 8.2 vemos que o PCA (Detalhes) obteve um desempenho de aproximadamente 22% em relação ao PCA (Padrão) e de 16% em relação ao ICA. Contudo, seu desempenho foi inferior (aproximadamente 21%) à técnica de Xie e Lam [74, 75].

8.4 Considerações Finais e Trabalhos Futuros

Neste trabalho propomos um método cujo objetivo é caracterizar uma face que tem suas características “ocultas” pela falta de iluminação. Para isso usamos o método de suavização com preservação de bordas de ROF para a obtenção das componentes suavizadas, aplicando então o PCA nas componentes com detalhes, ruídos e texturas. Os resultados alcançados mostram que o método proposto é menos sensível a falta de iluminação que o método PCA (Padrão), que o HE e o ICA.

Um inconveniente do método LN é a definição do tamanho do bloco em que o processo de triangularização será feito que varia de acordo com o tamanho das imagens utilizadas e das características da base de dados. Porém, isso não invalida os bons resultados conseguidos com a aplicação do mesmo, o que justifica maiores investigações sendo tema de trabalhos futuros.

9 Conclusão

Neste trabalho foi analisado o desempenho do método das autofaces na recuperação de faces, a partir de um banco de dados de imagem, considerando diferentes condições, tais como, pose, iluminação e perturbações na face.

Em relação às perturbações na face, ruídos e oclusões não prejudicaram o desempenho das autofaces no processo de recuperação. O mesmo não acontece em outros tipos de manipulações, tais como, o “estreitamento” ou ainda quando a face é disposta de “cabeça-para-baixo”.

Também foi analisada a influência da quantidade e quais autovalores escolher, em que cada autovalor tem uma autoface associada, no processo de recuperação da face. Com respeito aos resultados, concluímos que: ao invés de usar todas as autofaces possíveis, a seleção das autofaces relacionadas aos maiores autovalores não afeta muito o desempenho; as autofaces associadas aos maiores autovalores parecem caracterizar o formato da cabeça, enquanto que as autofaces associadas aos menores autovalores parecem “identificar” as características internas (olhos, nariz, boca).

Também é explorado o desempenho do PCA aplicado na componente das imagens que contém as texturas, os detalhes e os ruídos, chamado PCA Detalhes. Para isto foi aplicado um filtro de suavização em toda a base de dados.

Observou-se que os resultados foram satisfatórios inclusive com uma melhora em relação ao problema de iluminação, ou seja, quando a face consulta tem pouca incidência de luz, o que prejudicou o resultado do PCA aplicado nas imagens originais.

Referências

- [1] Benjamin Caricature, <http://www.interchile.com/benjamin>.
- [2] Caricature Zone, <http://www.magixl.com/heads/view.html>.
- [3] YaleA - Face Database, 1997, <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.
- [4] YaleB - Face Database, <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.
- [5] ATT - Face Database, <http://www.uk.research.att.com/facedatabase.html>.
- [6] Bell A. and Sejnowski T. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7:1129–1159, 1995.
- [7] Burton A., Wilson S., Cowan M., and Bruce V. Face recognition in poor-quality video. *Psychological Science*, pages 243–248, 1999.
- [8] Chambolle A. and Lions P. Image recovery via total variation minimization and related problems. *Numer. Math.*, 76:167–188, 1997.
- [9] Jain A., Bolle R., and Pankanti S. *Biometrics: Personal Identification in Networked Society*. Kluwer Academic Publishers, Boston/Dordrecht/London, 1999.
- [10] Jain A., Duin R., and Mao J. Statistical pattern recognition: A review. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22:4–37, 2000.
- [11] Martinez A. and Benavente R. The ar face database. *CVC Technical Report*, 24, 1998.
- [12] Nefian A. and Hayes M. Hidden markov models for face recognition. *In Proceedings Intern. Conferenc. on Acoustics, Speech and Signal Process.*, pages 2721–2724, 1998.
- [13] Pentland A. and Turk M. Face recognition using eigenfaces. *In Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 586–591, 1991.
- [14] Schodl A., Haro A., and Essa I. Head tracking using a textured polygonal model. *In Workshop on Perceptual User Interfaces*, pages 43–48, 1998.
- [15] Draper B., Baek K., Bartlett M., and Beveridge J. Recognizing faces with pca and ica. *Computer Vision and Image Understanding (Special Issue on Face Recognition)*, 91(1-2):115–137, Julho 2003.

- [16] Manjunath B., Chellappa R., and Malsburg C. A feature based approach to face recognition. *In Proceedings, IEEE Conf. on Computer Vision and Patt. Recogn.*, pages 373–378, 1992.
- [17] Moghaddam B. and Pentland A. Beyond linear eigenspaces: Bayesian matching for face recognition. *Face Recognition: From theory to Applications*, pages 230–243, 1998.
- [18] Burman C. Prosopagnosia. <http://www.prosopagnosia.com>.
- [19] Cichocki C. and Amari S. Adaptive blind signal and image processing: Learning algorithms and application. *John Wiley Sons LTD*, 2002.
- [20] Liu C. and Wechsler H. Comparative assessment of independent component analysis (ica) for face recognition. *In: Second Internat. Conf. on Audio and Video Based Biometric Person Authentication*, 1999.
- [21] Liu C. and Wechsler H. A shape- and texture-based enhanced fisher classifier for face recognition. *IEEE Trans. Image Process.*, 10:598–608, 2001.
- [22] Morimoto C., Yacoob Y., and Davis L. Recognition of head gestures using hidden markov models. *In ICPR, Viena, Austria.*, 1996.
- [23] Thomaz C. Estudo de classificadores para reconhecimento automático de faces. *Dissertação de Mestrado-PUC RJ*, 1999.
- [24] Liu D., Lam M., and Shen L. Illumination invariant face recognition. *Pattern Recognition*, 38:1705–1716, 2005.
- [25] Swets D. and Weng J. Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18:831–836, 1996.
- [26] Kitami E. and Thomaz C. Um tutorial sobre a análise de componentes principais para o reconhecimento automático de faces. Relatório Técnico. Depto Eng. Elétrica-FEI.
- [27] Samaria F. Face recognition using hidden markov models. *PhD thesis, University of Cambridge, Cambridge, UK*, 1994.
- [28] Samaria F. and Yong S. Hmm-based architecture for face identification. *Image Vis. Comput.*, 12:537–583, 1994.
- [29] Ralhp G. Face database. *In S.Li and A.Jain, (ed). Handbook of Face Recognition. Springer-Verlag*, 2005.
- [30] Rhodes G. and Tremewan T. Understanding face recognition: Caricature effects, inversion, and the homogeneity problem. *Visual Cognition*, 1:257–311, 1994.
- [31] Guilherme H. Detecção facial: Autofaces versus antifaces. *Tese de Mestrado - Univers. Fed. Santa Catarina, Programa de Pós Graduação em Eng. Elétrica*, 2003.

-
- [32] Moon H. and Phillips P. Computational and performance aspects of pca-based face-recognition algorithms. *Perception*, 30:303–321, 2001.
- [33] Karlgren J. Stylistic experiments for information retrieval. *Phd Thesis, Suécia*, 2000.
- [34] Sadr J., Jarudi I., and Sinha P. The role of eyebrows in face recognition. *Perception*, 32:285–93, 2003.
- [35] Yang J., Yu H., and Kunz W. An efficient lda algorithm for face recognition. *In: Proceedings of the Sixth International Conference on Control, Automation, Robotics and Vision*, 2000.
- [36] Zhao J., Su Y., Wang D., and Luo S. Illumination ratio image: synthesizing and recognition with varying illuminations. *Pattern Recognition Letters*, 24:2703–2710, 2003.
- [37] Baek K., Draper B., Beveridge J., and She K. Pca vs. ica: A comparison on the feret data set. *In: Internat. Conf. on Computer Vision, Pattern Recognition and Image Processing, North Carolina*, 2001.
- [38] Etemad K. and Chellappa R. Discriminant analysis for recognition of human faces images. *Journal of the Optical Society of America*, 14:1724–1733, 1997.
- [39] Kukula and Elliott. Evaluation of a facial recognition algorithm across three illumination conditions. *IEEE Aerospace and Electronic System Magazine*, pages 19–23, 2004.
- [40] Choi C. Kwak N. and Ahuja N. Face recognition using feature extraction based on independent component analysis. *In: Internat. Conf. on Image Processing*, 2002.
- [41] Chen L., Liao H., Lin J., Ko M., and Yu G. A new lda-based face recognition system which can solve the small sample size problem. *Pattern Recognition*, 33:1713–1726, 2000.
- [42] Rudin L., Osher S., and Fatemi E. Nonlinear total variation based noise removal algorithms. *Physics D*, (60):259–268, 1992.
- [43] Vese L. and Osher S. Modeling textures with total variation minimization and oscillating patterns in image processing. *Journal of Scientific Computing*, 19, 2003.
- [44] Bartlett M. Face image analysis by unsupervised learning and redundancy reduction. *Phd. Thesis, UNIVERSITY OF CALIFORNIA, SAN DIEGO, Cognitive Science and Psychology*, 1998.
- [45] Bartlett M., Lades H, and Sejnowski T. Independent component representations for face recognition. *SPIE Symposium on Electronic Imaging: Science and Technology; Conference on Human Vision and Electronic Imaging III*, 1998.
- [46] Bartlett M., Movellan J., and Sejnowski T. Face recognition by independent component analysis. *IEEE Trans. on Neural Networks*, 13(6):1450–1464, Novembro 2002.

-
- [47] Bertalmio M., Luminata V., and Sapiro G. Simultaneous structure and texture image inpainting. *In IEEE Transactions on Image Processing*, 12, 2003.
- [48] Gian M. and Fabio R. Fusion of appearance-based face recognition algorithms. *Pattern Analysis Application*, 7:151–163, 2004.
- [49] Kirby M. and Sirovich L. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 1:103–108, 1990.
- [50] Lewis M. and Johnston R. Understanding caricatures of faces. *Journal of Experimental Psychology*, 51(2):321–346, 1998.
- [51] Turk M. and Pentland A. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.
- [52] Y. Meyer. Oscillating patterns in image processing and nonlinear evolution equations. *University Lecture Series*, 22, 2002.
- [53] Deniz O., Castrillon M., and Hernandez M. Face recognition using independent component analysis and support vector machines. *In: Third Internat. Conf. on Audio and Video Based Biometric Person Authentication*, 2001.
- [54] Belhumeur P., Hespanha J., and Kriegman D. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19, 1997.
- [55] Morris P. and Wickham V. Typicality and face recognition: A critical re-evaluation of the two factor theory. *Quarterly Journal of Experimental Psychology A*, 54(3):863–877, 2001.
- [56] Sinha P. and Poggio T. I think i know that face. *Nature*, 384:404, 1996.
- [57] Juang J. Phillips P., Wechsler H. and Rauss P. The feret database and evaluation procedure for face recognition algorithms. *Image and Vision Computing Journal*, 1996.
- [58] Feris R., Campos T., and Cesar R. Detection and tracking of facial features in video sequences. *In Lecture Notes in Artificial Intelligence*, 1973:129–137, 2000. Springer-Verlag Press.
- [59] Gonzalez R. and Woods R. Digital image processing. 1992. Addison-Wesley Publishing Company.
- [60] Hsu R. Face detection and modeling for recognition. *Phd. Thesis, UNIVERSITY OF MICHIGAN, Department of Computer Science and Engineering*, 2002.
- [61] Huang R., Liu Q., Lu H., and Ma S. Solving the small sample size problem of lda. *In: Proceedings of the 16th International Conference on Pattern Recognition*, 3:29–32, 2002.

-
- [62] Haykin S. *Redes neurais: Princípios e prática*. 2000. 2ed. São Paulo: Bookman.
- [63] Li S. and Jain A. *Handbook of face recognition*. Eds. Springer-Verlag, 2005.
- [64] Lin S., Kung S., and Lin L. Face recognition/detection by probabilistic decision based neural network. *IEEE Trans. Neural Netw.*, 8:114–132, 1997.
- [65] Pawan S., Benjamin B., Yuri O., and Richard R. Face recognition by humans: 20 results all computer vision researchers should know about. *Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology*, 2005.
- [66] Theodoridis S. and Koutroumbas K. *Pattern recognition*. Academic Press, USA, 1st edition, 1999.
- [67] Mathew T. *Eigenfaces and beyond*. Academic Press, 2005.
- [68] Sim T., Baker S., and Bsat M. The cmu pose, illumination and expression database. *IEEE Trans. on Pattern Anal. and Machine Intellig.*, 25, 2003.
- [69] Bruce V. and Burton M. Processing images of faces. Eds. Ablex Publishing, Norwood, NJ, 1992.
- [70] Zhao W., Krishnaswamy A., Chellappa R., Swets D., and Weng J. Discriminant analysis of principal components for face recognition. *Face Recognition: from Theory to Applications*, pages 73–85, 1998.
- [71] Zhao W. and Chellappa R. Illumination-insensitive face recognition using symmetric shape-from-shading. In: *Proc. IEEE Conf. CVPR, Hilton Head*, 2000.
- [72] Zhao W., Chellappa R., and Phillips P. Subspace linear discriminant analysis for face recognition. 1999. Technical Report.
- [73] Zhao W., Chellappa R., Phillips P., and Rosenfeld A. Face recognition: A literature survey. *ACM Computing Surveys*, 35:399–458, 2003.
- [74] Xie X. and Lam K. An efficient method for face recognition under varying illumination. In *Proc. Internat. Symposium on Circuits and Systems*, 2005.
- [75] Xie X. and Lam K. Face recognition under varying illumination based on a 2d face shape model. *Pattern Recognition*, 38:221–230, 2005.