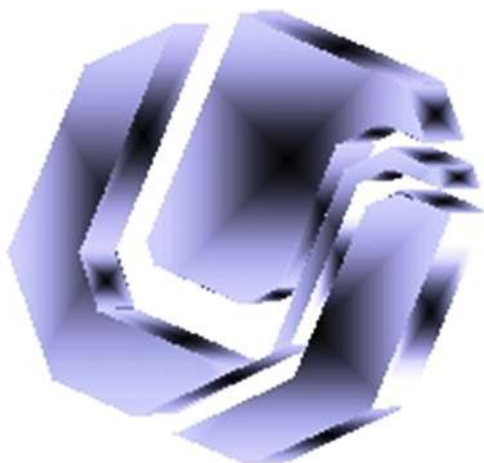


**UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**



**COMPARAÇÃO ENTRE ALGORITMO GENÉTICO, REDE
NEURAL ARTIFICIAL E ANÁLISE DE COMPONENTES
PRINCIPAIS NO RECONHECIMENTO DE FACES**

BENEDITO ALENCAR DE ARRUDA

2013

**UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA
PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

**COMPARAÇÃO ENTRE ALGORITMO GENÉTICO, REDE
NEURAL ARTIFICIAL E ANÁLISE DE COMPONENTES
PRINCIPAIS NO RECONHECIMENTO DE FACES**

Benedito Alencar De Arruda

Tese de doutorado apresentada por Benedito Alencar de Arruda à Universidade Federal de Uberlândia para obtenção do título de Doutor em Ciências.

Banca examinadora

Gilberto Arantes Carrijo, Pós-Dr – Orientador (UFU)

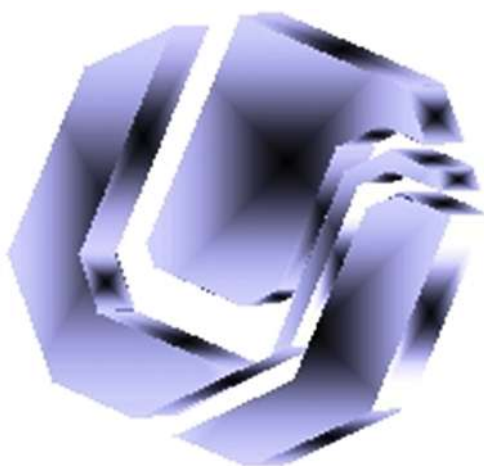
Antonio Cláudio Paschoarelli Veiga, Dr (UFU)

Edna Lúcia Flôres, Dra (UFU)

João Paulo Ignácio Ferreira Ribas, Dr (UFMT)

Ed´ Wilson Tavares Ferreira, Dr (IFMT)

**COMPARAÇÃO ENTRE ALGORITMO GENÉTICO, REDE
NEURAL ARTIFICIAL E ANÁLISE DE COMPONENTES
PRINCIPAIS NO RECONHECIMENTO DE FACES**



BENEDITO ALENCAR DE ARRUDA

Tese de Doutorado apresentada por Benedito Alencar de Arruda à Universidade Federal de Uberlândia como parte dos requisitos para obtenção do título de Doutor em Ciências.

Profº Gilberto Arantes Carrijo
Orientador

Profº Alexandre Cardoso
Coordenador do Curso de
Pós-Graduação

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG - Brasil

A779c Arruda, Benedito Alencar de, 1956-
2013 Comparação entre algoritmo genético, rede neural artificial e análise de componentes principais no reconhecimento de faces / Benedito Alencar de Arruda. - 2013.
 237 p. : il.

Orientador: Gilberto Arantes Carrijo.

Tese (doutorado) – Universidade Federal de Uberlândia, Programa de Pós-Graduação em Engenharia Elétrica.

Inclui bibliografia.

1. Engenharia elétrica - Teses. 2. Redes neurais (Computação) - Teses. 3. Reconhecimento de padrões - Teses. I. Carrijo, Gilberto Arantes, 1948- II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Engenharia Elétrica. III. Título.

CDU: 621.3

“ Todos saboreiam o fruto e ninguém vê a semente germinar; os que vêem a riqueza não vêem a economia; os que vêem a obra não vêem a labuta; os que vêem o saber não vêem o estudo; os que vêem a prosperidade não vêem o trabalho e os que vêem a recompensa não vêem as provações “.

Profº José Venâncio (1979)

AGRADECIMENTOS

A Deus em razão da crença que tenho pela existência de um ser supremo a quem atribuo a força recebida em muitos momentos difíceis ao longo desta caminhada.

A Universidade Federal de Uberlândia (UFU), instituição respeitada, na qual tenho convivido importantes anos de minha vida e que pude entender a fundo o lema escrito no muro da parte frontal do campus “um bem público a serviço do Brasil”.

Ao Professor Gilberto Arantes Carrijo, pesquisador de valor inestimável, majestoso na condução da orientação, imprescindível no estímulo nos momentos cruciais.

RESUMO

O reconhecimento de face tem-se mostrado uma técnica importante na identificação automática de pessoas. Ela é uma técnica biométrica de codificação de imagens caracterizada por explorar a auto-similaridade presente nas imagens digitais e cujo esforço computacional é considerável e tem exigido dos pesquisadores muita dedicação no sentido de cada vez mais aumentar a eficiência do processo. Nesta tese foi apresentado o estado da arte de sistemas de reconhecimento de faces citando-se diferentes formas de aplicações e trabalhos realizados por alguns pesquisadores. Foi mostrada também a metodologia empregada pelos classificadores Algoritmo Genético (AG), Rede Neural Artificial (RNA) e Análise de Componentes Principais (PCA). O objetivo deste trabalho foi realizar teste com esses classificadores para o reconhecimento de faces, analisando o percentual de acerto e o tempo de processamento. A razão de utilizar o AG é que a PCA, tradicionalmente empregada nesses casos, é muito lenta e de custo computacional elevado, tornando inviável em algumas aplicações, notadamente quando o banco de imagens é grande. Nos testes realizados neste trabalho foram utilizadas as bases de dados de dois arquivos de fotografia humana, a da Olivetti Research Laboratory, hoje base de dados de faces da Universidade Britânica de Cambridge (ORL) e a Face Recognition Data da Universidade de Essex, Reino Unido (UK). A taxa de acerto obtida utilizando AG, RNA e PCA foi maior do que aquelas obtidas pelos trabalhos pesquisados.

Palavras Chave: Reconhecimento de Faces, Rede Neural Artificial, Análise de Componentes Principais, Algoritmo Genético.

ABSTRACT

The face recognition has been shown to be an important technique in automatic identification of persons. It is a biometric image coding characterized by exploring self-similarity present in digital images and whose computational effort is significant and has required dedication of researchers in order to increasingly enhance the efficiency of the process. This thesis was presented the state of the art face recognition systems citing different forms of applications and work done by some researchers. It was also shown the methodology used by the classifiers Genetic Algorithm (GA), Artificial Neural Network (ANN) and Principal Component Analysis (PCA). The aim of this study was to test these classifiers for face recognition by analyzing the percentage of accuracy and processing time. The reason to use the GA is that the PCA, traditionally used in such cases, is very slow and the high computational cost, making it impractical in some applications, especially when the database is large images. In tests performed in this study were used the databases of two human photography files, the Olivetti Research Laboratory database today faces the British University of Cambridge (ORL) and Face Recognition Data, University of Essex, UK (UK). The accuracy rate obtained with GA, ANN and PCA were higher than those obtained by the studies reviewed.

Keywords: Face Recognition, Artificial Neural Networks, Principal Components Analysis, Genetic Algorithm.

COMPARAÇÃO ENTRE ALGORITMO GENÉTICO, REDE NEURAL ARTIFICIAL E ANÁLISE DE COMPONENTES PRINCIPAIS NO RECONHECIMENTO DE FACES

Sumário

CAPÍTULO I :	ESTADO DA ARTE.	1
1.1	Introdução	1
1.2	Reconhecimento de impressões digitais.	5
1.3	Reconhecimento de íris.	6
1.4	Reconhecimento de retina.	8
1.5	Reconhecimento facial.	8
1.5.1	Espaço-imagem e o espaço-face	11
1.5.2	Reconhecimento para autenticação e vigilância.	11
1.6	Aplicações do reconhecimento de faces e os seus desafios.	13
1.6.1	Identificação de faces.	13
1.6.2	Observações de pessoas e de lugares.	15
1.6.3	Cartões de identificação.	15
1.6.4	Investigação policial.	15
1.6.5	Indexação de banco de dados.	16
1.6.6	Identificação de imagem por método interativo.	16
1.6.7	Identificação de imagem por lembrança.	17
1.6.8	Utilizando multimídia.	17
1.6.9	Reconhecimento da cor da pele.	18
1.6.10	Reconhecimento de emoções.	18

1.6.11	Outras aplicações.	18
1.7	A limitação do sistema de reconhecimento de faces.	19
1.8	Trabalhos Realizados.	20
1.9	Considerações Finais Deste Capítulo.	20
 CAPÍTULO II: FUNDAMENTAÇÃO TEORICA DOS CLASIFICADORES: ALGO - RITMO GENÉTICO, REDE NEURAL ARTIFICIAL E ANÁLISE DE COMPONENTES PRINCIPAIS.		
2.1	Introdução.	22
2.2	Algoritmos Genéticos (AG)	23
2.2.1	Desenvolvimento do algoritmo genético.	24
2.2.2	Operadores genéticos.	25
2.2.2.1	Operador de cruzamento.	25
2.2.2.2	Operador de mutação..	26
2.2.3	Métodos de seleção.	26
2.2.4	Métodos de atualização da população.	27
2.2.5	Passos de um algoritmo genético..	27
2.3	Rede Neural Artificial.	28
2.3.1	Funcionamento do neurônio biológico.	28
2.3.2	Modelo do neurônio artificial..	30
2.3.3	Função de ativação.	32
2.3.3.1	Função limiar (degrau).	32
2.3.3.2	Função sigmóide..	33
2.3.3.3	Função signum.	34
2.3.3.4	Tangente hiperbólica.	35
2.3.4	Quanto ao treinamento da RNA.	35

2.3.5	Quanto ao Cálculo da RNA.	36
2.4	Análise de Componentes Principais (PCA)	39
2.4.1	Quanto ao treinamento da PCA.	42
2.4.2	Quanto ao reconhecimento pela PCA.	42
2.4.3	Cálculo pela PCA	42
2.5	Considerações Finais Deste Capítulo.	46
CAPÍTULO III: MATERIAIS E METODOLOGIA UTILIZADA PARA O RECO- NHECIMENTODE FACES.		47
3.1	Introdução.	47
3.2	Materiais	48
3.2.1	Quanto ao hardware e o software	49
3.2.2	Especificação da base de dados ORL.	49
3.2.3	Especificação da base de dados UK.	50
3.3	Metodologia Dos Testes Realizados	52
3.3.1	Metodologia do Algoritmo Genético para os arquivos ORL e UK...	53
3.3.1.1	Utilização do arquivo ORL pelo AG.	55
3.3.1.1.1	Utilização de toda imagem (pose) do arquivo ORL pelo AG.	55
3.3.1.1.2	Utilização das subposes do arquivo ORL pelo AG	56
3.3.1.2	Utilização do arquivo UK pelo AG	57
3.3.1.2.1	Utilização de toda imagem (pose) do arquivo UK pelo AG.	57
3.3.1.2.2	Utilização das subposes do arquivo UK pelo AG	58
3.3.2	Metodologia em relação à Rede Neural Artificial.	58
3.3.2.1	A rede neural artificial para o arquivo ORL.	59
3.3.2.2	A rede neural artificial para o arquivo UK.	59

3.3.3	Metodologia relacionada à Análise de Componentes Principais	59
3.3.3.1	A análise de componentes principais (PCA) para o arquivo ORL	60
3.3.3.2	A análise de componentes principais (PCA) para o arquivo UK	60
3.4	Sequência de utilização dos classificadores.	61
3.5	Apresentação dos Resultados.	61
3.6	Considerações Finais Deste Capítulo.	61
CAPÍTULO IV: UTILIZAÇÃO DO ALGORÍTMO GENÉTICO NO RECONHECI- MENTO DE FACES		63
4.1	Introdução.	63
4.2	A utilização do Algoritmo Genético utilizando o arquivo ORL.....	64
4.2.1	Quanto a utilização de toda a imagem (pose) ORL.....	65
4.2.2	Quanto a utilização de subimagens (subposes) ORL.....	70
4.2.2.1	Utilizando 154 coeficientes DCT, com um de cada subpose ORL.....	71
4.2.2.2	Utilizando 308 coeficientes DCT, com dois de cada subpose ORL.....	77
4.2.3.3	Utilizando 616 coeficientes DCT, com quatro de cada subpose ORL.....	84
4.3	A utilização do Algoritmo Genético usando o arquivo UK.	91
4.3.1	Quanto a utilização de toda a imagem (pose) UK.	92
4.3.2	Quanto a utilização de subimagens (subposes) UK.	98
4.3.2.1	Utilizando 550 coeficientes DCT, com um de cada subpose UK	99
4.3.2.2	Utilizando 1100 coeficientes DCT, com dois de cada subpose UK	105

4.3.2.3	Utilizando 2200 coeficientes DCT, com quatro de cada subpose UK	112
4.4	Considerações Finais Deste Capítulo.	120
CAPÍTULO V: UTILIZAÇÃO DA REDE NEURAL ARTIFICIAL NO RECONHECIMENTO DE FACES.		
122		
5.1	Introdução.	122
5.2	Experiência com a Rede Neural Artificial utilizando o arquivo ORL.	123
5.2.1	Experiência com 64 coeficientes DCT utilizando arquivo ORL.	123
5.2.2	Experiência com 100 coeficientes DCT utilizando arquivo ORL.	127
5.2.3	Experiência com 256 coeficientes DCT utilizando arquivo ORL.	131
5.3	Experiência com a Rede Neural Artificial utilizando o arquivo UK.	135
5.3.1	Experiência com 64 coeficientes DCT utilizando arquivo UK.	135
5.3.2	Experiência com 100 coeficientes DCT utilizando arquivo UK.	140
5.3.3	Experiência com 256 coeficientes DCT utilizando arquivo UK.	144
5.4	Considerações Finais Deste Capítulo.	148
CAPÍTULO VI: UTILIZAÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS NO RECONHECIMENTO DE FACES.		
150		
6.1	Introdução.	151
6.2	Experiência com a Análise de Componentes Principais (PCA) utilizando arquivo ORL.	152
6.3	Experiência com a Análise de Componentes Principais (PCA) utilizando arquivo UK.	154
6.4	Considerações Finais Deste Capítulo.	157
CAPÍTULO VII: COMPARAÇÃO DOS RESULTADOS OBTIDOS ENTRE O ALGORITMO GENÉTICO, A REDE NEURAL ARTIFICIAL E A ANÁLISE DE COMPONENTES PRINCIPAIS.		
159		

7.1	Introdução.	159
7.2	Resumo da experiência com o Algoritmo Genético.	159
7.2.1	Resumo da experiência com o Algoritmo Genético para o arquivo ORL	159
7.2.2	Resumo da experiência com Algoritmo Genético para o arquivo UK	162
7.3	Resumo da experiência com a Rede Neural.	165
7.3.1	Resumo da experiência com a Rede Neural Artificial para o arquivo ORL	165
7.3.2	Resumo da experiência com a Rede Neural Artificial para o arquivo UK	168
7.4	Resumo da experiência com a Análise de Componentes Principais.	171
7.4.1	Resumo da experiência com a Análise de Componentes Principais para o arquivo ORL... ..	171
7.4.2	Resumo da experiência com a Análise de Componentes Principais para o arquivo UK.	172
7.5	Comparação Dos Resultados Entre os Métodos do AG, RNA e PCA.	174
7.5.1	Comparação dos resultados entre os métodos do AG, RNA e PCA para o arquivo ORL.	174
7.5.2	Comparação dos resultados entre os métodos do AG, RNA e PCA para o arquivo UK.	178
7.6	Amstras de comparações de imagens reconhecidas erradas pelos algoritmos	181
7.7	Análises Dos Resultados.	182
7.8	Sobre O Desempenho E Aplicações Dos Classificadores.	183
7.9-	Considerações Finais Deste Capítulo.	184
CAPÍTULO VIII: CONCLUSÕES, CONTRIBUIÇÕES, SUGESTÃO DE TRABALHOS FUTUROS E PUBLICAÇÕES		186
8.1	As Implementações deste Trabalho	186

8.2	Conclusões e Contribuições.	187
8.3	Sugestões de trabalhos futuros.	188
8.4	Publicações.	191
REFERÊNCIAS BIBLIOGRÁFICAS.		192
APÊNDICE.		198
A	Introdução a Transformada Discreta do Cosseno.	198
A.1	Introdução.	198
A.2	Transformada Discreta do Cosseno.	199
B	Experiências Práticas: Banco de Imagens Familiar e Interação com o Meio Externo.	205
B.1	Introdução.	205
B.2	Experiência com fotografia familiar.	205
B.2.1	Resultados.	206
B.3	Experiência Com a Controladora Lógica Programável.	211

LISTA DE FIGURAS

Figura	Título	Página
1.1	Diagrama de um sistema biométrico.	4
1.2	Gráfico com o percentual do emprego biométrico.	5
1.3	Aspecto da impressão digital.	6
1.4	Olho humano destacando a íris	8
1.5	Aspecto do globo ocular humano com a retina.	9
1.6	Diagrama da sequencia para o reconhecimento de face	10
1.7	Disposição do reconhecimento de face. Leitora óptica-computador	15
2.1	Principais passos de um algoritmo de genético.	27
2.2	Neurônio biológico.	29
2.3	Neurônio de McCullogh & Pitts. Modelo geral do neurônio.	31
2.4	Gráfico da função de limiar.	33
2.5	Gráfico da função sigmóide.	34
2.6	Gráfico da função signun.	34
2.7	Gráfico da função tangente hiperbólica	35
2.8	Gráfico da distribuição de dados pela PCA.	40
2.9	Diagrama demonstrativo da PCA com imagens aleatórias.	41
3.1	Figura da árvore com vários rostos.	48
3.2	Conjunto de pessoas e suas respectivas poses do arquivo ORL.	50
3.3	Representação do arquivo ORL com dez poses por classe.	50

3.4	Conjunto de imagens de pessoas e suas respectivas poses extraídas do arquivo UK.	51
3.5	Representação do arquivo UK com 20 poses por classe.	52
3.6	Representação do procedimento de reconhecimento de faces.	52
3.7	Representação de um vetor imagem P, um vetor indivíduo I e um vetor máscara M para n coeficientes.	53
3.8	Localização dos coeficientes DCT utilizados para cada pose ORL.	55
3.9	Representação de um vetor imagem Px, um vetor indivíduo Ix e um vetor máscara Mx para 64 coeficientes.....	56
3.10	Representação de uma pose ORL de 92 colunas por 112 linhas dividida em 154 subposes X.....	56
3.11	Localização dos coeficientes DCT utilizados para cada pose UK.	57
3.12	Representação de um vetor imagem Py, um vetor indivíduo Iy e um vetor máscara My para 64 coeficientes.....	57
3.13	Representação de uma pose UK de 180 colunas por 200 linhas dividida em 550 subposes Y.....	58
4.1	Representação de uma classe por coeficientes DCT com 10 poses. Cada pose é um vetor com 64 elementos, arquivo ORL.....	65
4.2	Taxa de acerto do AG considerando 64 coeficientes, arquivo ORL.	67
4.3	Tempo de treinamento do AG considerando 64 coeficientes, arquivo ORL. ...	68
4.4	Tempo de reconhecimento do AG considerando 64 coeficientes por pose, arquivo ORL.....	69
4.5	Gráficos de erros do AG em função do número de gerações sendo 5 poses para o treinamento e 5 poses para o reconhecimento – usando 64 coeficientes DCT, para 40 classes, arquivo ORL.....	70
4.6	Representação da subpose X_1 , base para o vetor característica de 154 coeficientes DCT do arquivo ORL.....	71
4.7	Vetor representativo da primeira pose P através de 154 coeficientes DCT mais significativo de cada subpose - arquivo ORL.....	72
4.8	Conjunto de vetores com 154 coeficientes DCT representativo da classe C_1 - arquivo ORL.....	72
4.9	Taxa de acerto do AG considerando 154 coeficientes, arquivo ORL.	74

4.10	Tempo de treinamento do AG considerando 154 coeficientes, arquivo ORL...	75
4.11	Tempo de reconhecimento do AG considerando 154 coeficientes por pose, arquivo ORL.....	76
4.12	Gráficos de erros do AG em função do número de gerações sendo cinco poses para o treinamento e cinco para o reconhecimento – usando 154 coeficientes DCT para 40 classes, arquivo ORL	77
4.13	Representação da subpose X_1 , base para o vetor característica de 308 coeficientes DCT, arquivo ORL.....	78
4.14	Vetor representativo da primeira pose P através de 308 coeficientes DCT mais significativos de cada subpose - arquivo ORL.....	79
4.15	Conjunto de vetores com 308 coeficientes DCT representativo da classe C_1 - arquivo ORL.....	79
4.16	Taxa de acerto do AG considerando 308 coeficientes, arquivo ORL.	81
4.17	Tempo de treinamento do AG considerando 308 coeficientes, arquivo ORL. ..	82
4.18	Tempo de reconhecimento do AG considerando 308 coeficientes por pose, arquivo ORL.....	83
4.19	Gráficos de erros do AG em função do número de gerações sendo cinco poses para treinamento e cinco poses para o reconhecimento – usando 308 coeficientes DCT, para 40 classes, arquivo ORL. ...	84
4.20	Representação da subpose X_1 , base para o vetor característica de 616 coeficientes DCT, arquivo ORL.....	85
4.21	Vetor representativo da primeira pose P através de 616 coeficientes DCT mais significativos de cada subpose, arquivo ORL.	86
4.22	Conjunto de vetores com 616 coeficientes DCT representativo da classe C_1 - arquivo ORL.....	86
4.23	Taxa de acerto do AG considerando 616 coeficientes, arquivo ORL.	88
4.24	Tempo de treinamento do AG considerando 616 coeficientes, arquivo ORL. ..	89
4.25	Tempo de reconhecimento do AG considerando 616 coeficientes por pose, arquivo ORL.....	90
4.26	Gráficos de erros do AG em função do número de gerações sendo 5 cinco poses para o treinamento e 5 cinco poses para o reconhecimento – usando 616 coeficientes DCT, para 40 classes, arquivo ORL.	91

4.27	Representação de uma classe por coeficientes DCT com 20 poses. Cada pose é um vetor com 64 elementos, arquivo UK.	93
4.28	Taxa de acerto considerando 64 coeficientes por pose, arquivo UK.	95
4.29	Tempo de treinamento considerando 64 coeficientes por pose, arquivo UK....	96
4.30	Tempo de reconhecimento do AG considerando 64 coeficientes por pose, arquivo UK....	97
4.31	Gráficos de erros do AG em função do número de gerações sendo 10 poses para o treinamento e 10 poses para o reconhecimento – usando 64 coeficientes DCT, para 120 classes, arquivo UK... ..	98
4.32	Representação da subpose Y_1 , base para o vetor característico de 550 coeficientes DCT, arquivo UK....	99
4.33	Vetor representativo da primeira pose G com 550 coeficientes DCT mais significativos de cada subpose - arquivo UK....	100
4.34	Conjunto de vetores com 550 coeficientes DCT cada um, representativo da classe F_1 - arquivo UK....	100
4.35	Taxa de acerto considerando 550 coeficientes por pose, arquivo UK.	102
4.36	Tempo de treinamento considerando 550 coeficientes por pose, arquivo UK..	103
4.37	Tempo de reconhecimento do AG considerando 550 coeficientes por pose, arquivo UK....	104
4.38	Gráficos de erros do AG em função do número de gerações onde foram utilizadas 10 poses para o treinamento e 10 poses para o reconhecimento – usando 550 coeficientes DCT, para 120 classes, arquivo UK.	105
4.39	Representação da subpose Y_1 , base para o vetor característico de 1100 coeficientes DCT, arquivo UK....	106
4.40	Vetor representativo da primeira pose G através de 1100 coeficientes DCT mais significativos de cada subpose, arquivo UK....	107
4.41	Conjunto de vetores com 1100 coeficientes DCT cada um, representativo da classe F_1 - arquivo UK....	107
4.42	Taxa de acerto considerando 1100 coeficientes por pose, arquivo UK.	109
4.43	Tempo de treinamento considerando 1100 coeficientes por pose, arquivo UK	110
4.44	Tempo de reconhecimento do AG considerando 1100 coeficientes por pose, arquivo UK....	111

4.45	Gráficos de erros do AG em função do número de gerações onde foram utilizadas 10 poses para o treinamento e 10 poses para o reconhecimento – usando 1100 coeficientes DCT, para 120 classes, arquivo UK....	112
4.46	Representação da subpose Y_1 , base para o vetor característico de 2200 coeficientes DCT, arquivo UK....	113
4.47	Vetor representativo da primeira pose G através de 2200 coeficientes DCT mais significativos de cada subpose, arquivo UK....	114
4.48	Conjunto de vetores com 2200 coeficientes DCT cada um, representativo da classe F_1 , arquivo UK....	114
4.49	Taxa de acerto considerando 2200 coeficientes por pose, arquivo UK.	117
4.50	Tempo de treinamento considerando 2200 coeficientes por pose, arquivo UK	118
4.51	Tempo de reconhecimento do AG considerando 2200 coeficientes por pose, arquivo UK....	119
4.52	Gráficos de erros do AG em função do número de gerações onde foram utilizados 10 poses para o treinamento e 10 poses para o reconhecimento – usando 2200 coeficientes DCT, para 120 classes, arquivo UK.	120
5.1	Curvas das taxas de acertos da RNA em função do nº de classes usando de uma, três e cinco poses para o reconhecimento, 64 coeficientes- ORL.	125
5.2	Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, três e cinco poses, 64 coeficientes - ORL.	126
5.3	Curva do tempo de treinamento da RNA em função do nº de classes para 64 coeficientes por pose - ORL....	127
5.4	Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, três e cinco poses para o reconhecimento, 100 coeficientes- ORL.	129
5.5	Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, três e cinco poses, 100 coeficientes – ORL.	130
5.6	Curva do tempo de treinamento da RNA em função do nº de classes para 100 coeficientes por pose - ORL....	131
5.7	Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, três e cinco poses para o reconhecimento, 256 coeficientes- ORL.	133
5.8	Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, três e cinco poses, 256 coeficientes – ORL.	134

5.9	Curva do tempo de treinamento da RNA em função do nº de classes para 256 coeficientes por pose - ORL.....	135
5.10	Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, 64 coeficientes- UK.	138
5.11	Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, cinco e 10 poses, 64 coeficientes - UK.	139
5.12	Curva do tempo de treinamento da RNA em função do nº de classes para 64 coeficientes por pose - UK.	140
5.13	Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, 100 coeficientes- UK.	142
5.14	Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, cinco e 10 poses, 100 coeficientes - UK.	143
5.15	Curva do tempo de treinamento da RNA em função do nº de classes para 100 coeficientes por pose - UK.	144
5.16	Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, 256 coeficientes- UK.	146
5.17	Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma a cinco poses, 256 coeficientes - UK.	147
5.18	Curva do tempo de treinamento da RNA em função do nº de classes para 256 coeficientes por pose - UK.	148
6.1	Curvas das taxas de acertos da PCA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, arquivo ORL.....	153
6.2	Curvas dos tempos de reconhecimentos da PCA em função do nº de classes utilizando uma, cinco e 10 poses, arquivo ORL.	154
6.3	Curvas das taxas de acertos da PCA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, arquivo UK.	156
6.4	Curvas dos tempos de reconhecimentos da PCA em função do nº de classes utilizando uma, cinco e 10 poses, arquivo UK.	157
7.1	Comparação entre as taxas de acerto do AG considerando 64, 154, 308 e 616 coeficientes, arquivo ORL.	161
7.2	Comparação entre os tempos de reconhecimentos do AG considerando 64, 154, 308 e 616 coeficientes, arquivo ORL.	162
7.3	Comparação entre as taxas de acertos do AG considerando 64, 550, 1100 e 2200 coeficientes, arquivo UK.	164

7.4	Comparação entre os tempos de reconhecimentos do AG considerando 64, 550, 1100 e 2200 coeficientes, arquivo UK.	165
7.5	Comparação entre as taxas de acertos do RNA considerando 64, 100 e 256 coeficientes, arquivo ORL.	167
7.6	Comparação entre os tempos de reconhecimentos do RNA considerando 64, 100 e 256 coeficientes, arquivo ORL.	168
7.7	Comparação entre as taxas de acertos do RNA considerando 64, 100 e 256 coeficientes, arquivo UK.	170
7.8	Comparação entre os tempos de reconhecimentos do RNA considerando 64, 100 e 256 coeficientes, arquivo UK.	171
7.9	Comparação entre as taxas de acerto da PCA para os arquivos ORL e UK	173
7.10	Comparação entre os tempos de reconhecimentos da PCA para os arquivos ORL E UK.	174
7.11	Comparação das taxas de acertos entre o AG, a RNA e a PCA, arquivo ORL	176
7.12	Comparação dos tempos de reconhecimentos entre o AG, a RNA e a PCA, arquivos ORL. ...	177
7.13	Comparação do tempo de treinamento entre o AG, a RNA e a PCA – arquivo ORL.	178
7.14	Comparação das taxas de acertos entre o AG, a RNA e a PCA - arquivo UK.	179
7.15	Comparação dos tempos de reconhecimentos entre o AG, a RNA e a PCA – arquivos UK.	180
7.16	Comparação do tempo de treinamento entre o AG, a RNA e a PCA – arquivo UK.	181
7.17	Comparação de imagens de faces reconhecidas errada pelo AG. As duas primeiras faces são do arquivo ORL e as outras duas do arquivo UK... ..	181
7.18	Comparação de imagens de faces reconhecidas erradas pelo RNA. As duas primeiras faces são do arquivo ORL e as outras duas do arquivo UK... ..	182
7.19	Comparação de imagens de faces reconhecidas erradas pelo PCA. As duas primeiras faces são no arquivo ORL e as outras duas do arquivo UK.	182

LISTA DE TABELAS

Tabela	Título	Página
4.1	Resultados obtidos do reconhecimento de face utilizando 64 coeficientes DCT de toda pose, arquivo ORL.	66
4.2	Resultados obtidos do reconhecimento de face considerando 154 coeficientes DCT arquivo ORL.	73
4.3	Resultados obtidos do reconhecimento de face considerando 308 coeficientes DCT arquivo ORL..	80
4.4	Resultados obtidos do reconhecimento de face considerando 616 coeficientes DCT da subimagem, arquivo ORL.	87
4.5	Resultados obtidos do reconhecimento de face utilizando 64 coeficientes DCT da pose toda, arquivo UK.	99
4.6	Resultados obtidos do reconhecimento de face considerando 550 coeficientes DCT por pose, arquivo UK.	101
4.7	Resultados obtidos do reconhecimento de face considerando 1100 coeficientes DCT por pose, arquivo UK.	108
4.8	Resultados obtidos do reconhecimento de face considerando 2200 coeficientes DCT da subimagem, arquivo UK.	116
5.1	Resultados da experiência com a RNA utilizando cinco poses para o treinamento e uma a cinco poses para o reconhecimento para 64 coeficientes – ORL. . . .	124
5.2	Resultados da experiência com a RNA utilizando cinco poses para o treinamento e de uma a cinco poses para o reconhecimento para o nº de 100 coeficientes - ORL reconhecimento e de treinamento.	128
5.3	Resultados da experiência com a RNA utilizando cinco poses para o treinamento e de uma, três e cinco poses para o reconhecimento para o nº de 256 coeficientes - ORL.	132

5.4	Resultados da experiência com a RNA utilizando 10 poses para o treinamento e uma a 10 poses para o reconhecimento de duas a 40 classes, 64 coeficientes DCT – UK.	136
5.5	Resultados da experiência com a RNA utilizando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de 50 a 120 classes, 64 coeficientes DCT – UK.	137
5.6	Resultados da experiência com a RNA utilizando 10 poses para o treinamento e de uma a 10 poses (NP) para o reconhecimento de 2 duas a 40 classes, 100 coeficientes DCT – UK.	141
5.7	Resultados da experiência com a RNA utilizando 10 poses para o treinamento e de uma a 10 poses (NP) para o reconhecimento de duas a 120 classes, 100 coeficientes DCT – UK.	141
5.8	Resultados da experiência com a RNA utilizando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de duas a 40 classes, 256 coeficientes DCT – UK.	145
5.9	Resultados da experiência com a RNA utilizando 10 poses para o treinamento e de uma a 10 poses (NP) para o reconhecimento de duas a 120 classes, 256 coeficientes DCT – UK.	145
6.1	Resultado da experiência com a PCA utilizando de uma a 10 poses para o reconhecimento de faces – arquivo ORL.	152
6.2	Resultado da experiência com a PCA utilizando de uma a 10 poses para o reconhecimento de faces - arquivo UK.	155
7.1	Resultados obtidos do reconhecimento da face, AG, 5 poses para o reconhecimento, arquivo ORL para 2 a 40 classes.	160
7.2	Resultados obtidos do reconhecimento da face, AG, 10 poses para o reconhecimento, arquivo UK para 5 a 60 classes.	163
7.3	Resultados obtidos do reconhecimento da face, AG, 10 poses para o reconhecimento, arquivo UK para 70 a 120 classes.	163
7.4	Resultados obtidos do reconhecimento da face, RNA, 5 poses para o reconhecimento, arquivo ORL para 2 a 40 classes.	166
7.5	Resultados obtidos do reconhecimento da face, RNA, 10 poses para o reconhecimento, arquivo UK para 5 a 60 classes.	169
7.6	Resultados obtidos do reconhecimento da face, RNA, 10 poses para o reconhecimento, arquivo UK para 70 a 120 classes.	169
7.7	Resultados obtidos do reconhecimento de face, PCA, 10 poses para o reconhecimento, arquivo ORL para 2 a 40 classes.	172

7.8	Resultados obtidos do reconhecimento da face, PCA, 10 poses para o reconhecimento, arquivo UK para 2 a 40 classes.	172
7.9	Resultados obtidos entre os classificadores RNA, PCA e AG (100 gerações) considerando 5 poses para o reconhecimento, tempo em segundos - arquivo ORL.	175
7.10	Resultados obtidos entre os classificadores RNA, PCA e AG (100 gerações) considerando 10 poses para o reconhecimento, tempo em segundos - arquivo UK.	179

LISTA DE ABREVIATURAS

AFIS	Automated Fingerprint Identification Systems
AG	Algoritmo Genético (Genetic Algorithm – GA)
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
FT	Fourier Transform
JPEG	Joint Photograph Experts Group
KLT	Karhunen-Loève Transform
ORL	Olivetti Research Laboratory
PCA	Principal Component Analysis
PGM	Portable Gray Map

CAPÍTULO I

ESTADO DA ARTE

1.1- Introdução

Pode-se afirmar que na atualidade está cada vez mais atraente a utilização do processamento digital de imagens (PDI) na execução de pesquisas diversas. Os avanços são de tal ordem que não se imagina o desenvolvimento sem a presença marcante, indispensável dos recursos computacionais. Muitas áreas do conhecimento, tais como engenharia, medicina, astronomia, aplicações industriais, etc., utilizam aplicações baseadas em conceitos e técnicas de PDI, cujos resultados visam, de uma forma geral, à interpretação humana ou à utilização em visão computacional.

O Brasil adquiriu recentemente (novembro de 2010) um supercomputador que aparece entre os mais rápidos sistemas computacionais do mundo [1]. Chamado Tupã, este novo supercomputador do Instituto Nacional de Pesquisas Espaciais (Inpe), ocupa o 29º lugar que é a mais alta colocação já alcançada por uma máquina instalada no Brasil. O resultado também o classifica como o mais poderoso do hemisfério Sul. Com velocidade máxima de 258 TFlop/s, equivalente a 258 trilhões de cálculos por segundo, Tupã é hoje o terceiro mais poderoso entre os supercomputadores dedicados à previsão numérica operacional do tempo e do clima sazonal. Essa super máquina será utilizada para pesquisa e fornecimento de dados na área meteorológica para prevenção de catástrofe.

Diversos avanços têm-se presenciado nos últimos anos entre os quais se pode destacar o processamento de imagens para reconhecimento de faces. Com o surgimento de novos serviços de comunicação digital pela internet ou para comunicações multimídia, como teleconferências e televisão digital, o desenvolvimento de pesquisas

na área de reconhecimento de imagens de pessoas tem aumentado substancialmente, ao mesmo tempo que lança um desafio enorme para a utilização das técnicas no sentido de tornar célere suas respostas. Entre vários padrões e aplicações do PDI como geoprocessamento, sensoriamento remoto, cartografia e muitas outras, está a técnica de reconhecimento de pessoa pela imagem da face, da íris, da digital, do som da voz etc.

O estudo de reconhecimento de pessoas avança via PDI, porém, ainda apresenta algumas deficiências, sendo a principal delas o elevado tempo de processamento exigido na etapa de codificação. A partir disso, muitos modelos híbridos envolvendo a codificação têm sido propostos com o objetivo de desenvolver codificadores que mantêm a qualidade da imagem reconstruída e com um desempenho satisfatório.

A evolução acelerada da tecnologia vem respaldando em todos os aspectos as pesquisas nas modalidades de reconhecimento de pessoas. Até o mais simples cidadão vive hoje imerso em um mundo de informações que se torna impossível ignorar os acontecimentos. A tecnologia está disseminada e cada vez mais ao alcance das pessoas e, mesmo que estas não queiram adquirir equipamentos sofisticados, os avanços tecnológicos podem ser observados em sua volta. Para quem não sai de casa, as tais inovações podem ser vistas pela televisão, nas exposições de filmes contemporâneos que mostram as inúmeras formas de uso da tecnologia, como o salto de qualidade da própria televisão referindo-se notadamente ao sistema digital.

Existem muitas áreas para a aplicação das técnicas de reconhecimento de face que podem ser classificadas em duas vertentes: a autenticação e a vigilância, dependendo do elemento a ser avaliado. Na primeira vertente, os próprios usuários se apresentam para declarar sua legitimidade, enquanto na segunda existe um monitoramento constante de um espaço ou ambiente na espera de eventos que possam caracterizar violação das normas de acesso.

Na modalidade de autenticação, as melhorias das técnicas avançam significativamente no sentido de contribuir para o controle ou as restrições de acesso a locais estabelecidos, como áreas privativas que por questões de segurança ou de sigilo não podem ser acessadas por pessoas que não pertencem as equipes que nelas trabalham. Em muitos casos, a restrição se estende inclusive ao pessoal da própria organização, sendo o acesso permitido apenas a quem possui credenciais definidas por hierarquia ou por funcionalidade desenvolvida.

Esse controle pode ser realizado de várias formas. Uma delas seria colocar na entrada do recinto a ser observado, uma pessoa que conhecesse todos os funcionários e

que de modo visual controlaria o acesso. O problema desse tipo de tarefa é que se torna praticamente inviável quando se tratar de uma empresa com grande número de funcionários. Fica difícil confiar apenas na boa memória do responsável pela verificação.

Devido a isso, foi necessário a busca por sistemas de segurança que automatizassem os mecanismos de verificação de acesso ou de contribuição com tarefas relativas à proteção de patrimônios críticos.

Para a utilização de sistema de autenticação, aquele que identifica uma pessoa para um determinado fim, se faz necessário preparar uma base de dados contendo os critérios para se liberar o acesso e montar os dispositivos que executam os procedimentos necessários.

Para tal, hoje em dia se emprega muito os cartões de identificação que contêm ou senhas em código de barras ou microchips com circuitos inteligentes. Esses cartões são apresentados ao sistema de segurança que faz a comparação entre as informações lidas do cartão e aquelas contidas no banco de dados.

Esses tipos de identificação apresentados estão longe de serem as melhores opções. Não são totalmente seguras, uma vez que senhas são facilmente esquecidas ou obtidas sem muito esforço e os cartões podem ser utilizados por outras pessoas que desejam de alguma forma burlar o sistema de autenticação. Ou seja, para se ter uma identificação mais confiável, o dispositivo utilizado precisa conter características inerentes ao usuário.

O avanço das técnicas de processamento de imagem e a gradativa queda nos preços dos equipamentos estão estimulando o surgimento de novas técnicas que, por sua vez, ajudam as instituições públicas e particulares a zelarem pelos seus bens.

Essas novas tecnologias estão baseadas em características únicas de cada indivíduo, como a face, a íris, a retina, a impressão digital etc., fixando sua identificação perto da margem zero de erro. Por exemplo, a impressão digital é muito mais difícil de ser reproduzida do que um código de barras.

Portanto, entra-se, na chamada biometria que, segundo o dicionário Aurélio, “é a ciência da identificação baseada na medição precisa de traços biológicos”. A tecnologia biométrica é usada para a identificação de pessoas por meio dos traços físicos únicos e que são difíceis de serem reproduzidos. Esta é uma antiga ambição de cientistas e bastante explorada na ficção científica e no cinema. Nos filmes “Gattaca”, “Minority

Report” e “Adorável Mundo Novo”, esse tema é bastante abordado e mostra como essa tecnologia pode-se tornar parte do dia a dia das pessoas.

Além do reconhecimento pela impressão digital, íris, retina e face, existem outros métodos como o reconhecimento pelas ondas cerebrais, voz, palma da mão, assinatura e digitação. Embora todos esses métodos tenham qualidades, a identificação mais segura é aquela realizada pela íris, que, apesar do seu custo ainda ser alto, apresenta uma margem de erro de 0,05%.

No Brasil, estão cada vez mais disseminados os meios biométricos como senha para o acesso. Tanto as corporações privadas quanto as públicas empregam essa tecnologia que se devem transformar em uma alternativa de segurança e de controle de pessoas para um grande número de empresas. Os próprios ministros do Supremo Tribunal Federal acessam seus computadores por meio da identificação pelas impressões digitais.

Mas, nem tudo são flores. Essa mesma tecnologia pode ser utilizada por organização criminosa para, por exemplo, esconder informações sigilosas, burlar as impressões digitais, desqualificar o reconhecimento facial e até mesmo a retina, embora, neste caso, seja muito mais difícil.

Com o lema "seu corpo é sua senha", o sistema biométrico está com a aceitabilidade crescente na sociedade em função dos crimes de toda ordem que geram pânico na população e fazem das triagens de pessoas procedimentos cada vez mais imperativos.

Conforme o diagrama mostrado na Figura 1.1, a biometria pode designar mensurações fisiológicas ou características comportamentais. Neste sentido, qualquer característica humana pode servir como identificação de uma pessoa desde que satisfaça os requisitos da universalidade (todos os indivíduos possuem determinada característica), unicidade (fenômeno provado que não existem dois indivíduos com o mesmo padrão), imutabilidade (a característica estudada não muda com o tempo) e coletabilidade (característica que pode ser mensurada ou medida) [2].

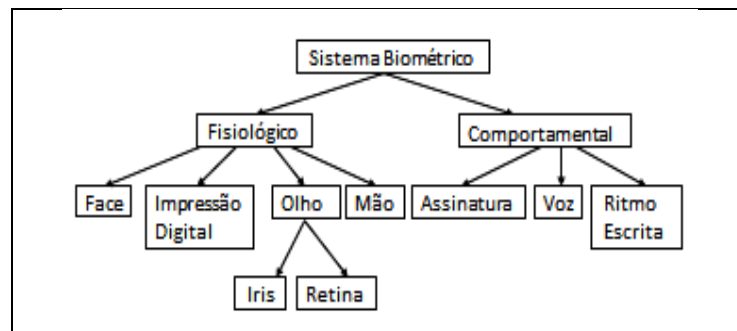


Figura 1.1: Diagrama de um sistema biométrico.

A Figura 1.2 mostra um gráfico com o percentual do emprego biométrico até o ano de 2007. Nota-se que a maior taxa de uso é o da impressão digital com 48%, enquanto o menos utilizado é o do emprego de assinaturas. Isto mostra o grande desafio técnico que vem pela frente, que é aumentar a taxa de uso biométrico dos sistemas utilizando a íris, a retina, a face etc.

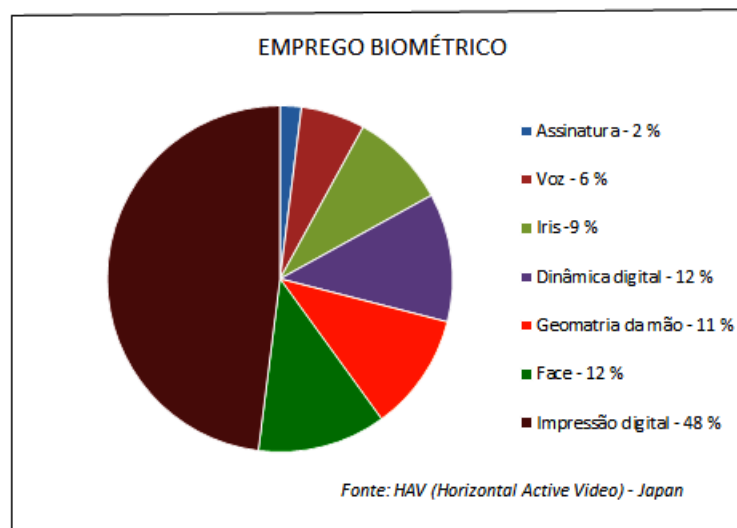


Figura 1.2: Gráfico com o percentual do emprego biométrico.

Nas seções 1.2, 1.3, 1.4 são mostrados os métodos biométricos por: impressão digital, íris e retina respectivamente, sendo que o método de reconhecimento de face ganha destaque a partir da seção 1.5.

1.2- Reconhecimento de impressões digitais

É um processo largamente utilizado por órgãos de investigação. Desenvolvida nos anos 60, era no começo bastante artesanal, o que exigia muita perícia por parte dos investigadores, no entanto, essa modalidade recebeu grande investimento em pesquisas e hoje o reconhecimento de impressões digitais é realizado automaticamente. Na Figura 1.3 tem-se o aspecto da impressão digital [3].

Esta forma é muito utilizada nas investigações forenses por meio de softwares chamados Automated Fingerprint Identification Systems (AFIS), que usam grandes arquivos previamente montados com dados e imagens completas das impressões digitais para serem comparadas com aquelas recolhidas geralmente em cenas de crimes. Outra forma é aquela destinada à autenticação que usa técnicas conhecidas como Finger Scan (FS). Tanto uma técnica como a outra fazem o reconhecimento de padrões das digitais e geram um modelo a partir da imagem original. Esta última forma é destinada à pesquisa rápida mesmo se tratando de grandes bases de dados. A escolha de qual técnica utilizar depende do grau de precisão desejado [4].

As principais tecnologias utilizadas para a captura de imagens de impressões digitais são:

- 1- Ótica – o usuário pressiona o dedo sobre a face ótica do leitor de impressão digital. Neste caso, os resíduos deixados por outros usuários podem comprometer a captura de novas impressões.
- 2- Chips de silício – utiliza sinais elétricos para formação da imagem. Como se trata de um dispositivo pequeno é muito utilizado em celulares e laptops [5].
- 3- Ultrassom - é o mais preciso de todos, porém, muito mais caro. Este tipo gera imagens em alta definição mesmo em situações adversas, como resíduos e sujeiras. A imagem é formada baseada em cálculos de distância, levando em consideração a impedância da pele, do ar e do próprio equipamento.



Figura 1.3: Aspecto da impressão digital.

1.3- Reconhecimento de Íris

A utilização da íris como elemento de identificação foi sugerida pela primeira vez no final dos anos 80, mas foi em 1993 que John Daugman [6] apresentou o primeiro e mais conhecido algoritmo para o reconhecimento automático de indivíduos utilizando a estrutura da íris. O algoritmo de extração de características aplica filtros complexos 2D Gabor Wavelets gerando uma *template* binário pela codificação das informações dos coeficientes produzidos. A comparação das íris é realizada utilizando-se a análise do valor de distância da Hamming (dH) calculada a partir de duas *templates*. Se a dH for menor que um determinado valor limiar, as *templates* são consideradas originais da íris dos mesmos indivíduos, caso contrário, elas são de indivíduos diferentes.

Muitos outros algoritmos foram propostos para o reconhecimento da íris como a sugestão feita por Wildes [7] que realiza um tipo de normalização em que a imagem da íris é deformada geometricamente por sucessivos deslocamentos e rotações. Muitos outros algoritmos estão disponíveis para o reconhecimento de íris e são extremamente eficientes na identificação de pessoas pela peculiaridade desse órgão do ser humano. Pela íris poderia, por exemplo, reconhecer uma pessoa em um banco de dados que contivesse toda a população do mundo com margens ínfimas de erros.

Essas técnicas se baseiam em características visíveis da íris que tem um tamanho de cerca de 11 mm e é capaz de fornecer cerca de 266 pontos únicos de identificação.

Como se pode ver na Figura 1.4 [8], a principal característica da íris é o tecido rugoso que aparenta dividi-la em forma radial.

A captura da imagem é realizada por câmera que utiliza tanto a luz visível quanto a luz infravermelha a uma distancia de até 1 metro da face do usuário. Os algoritmos são tão eficientes que podem fazer o tratamento da imagem mesmo com apenas 1/3 da imagem da íris captada. Esta é uma situação comum no momento da captura da imagem, pois a abertura do olho e a inclinação do ângulo entre o rosto e a câmera são variáveis.

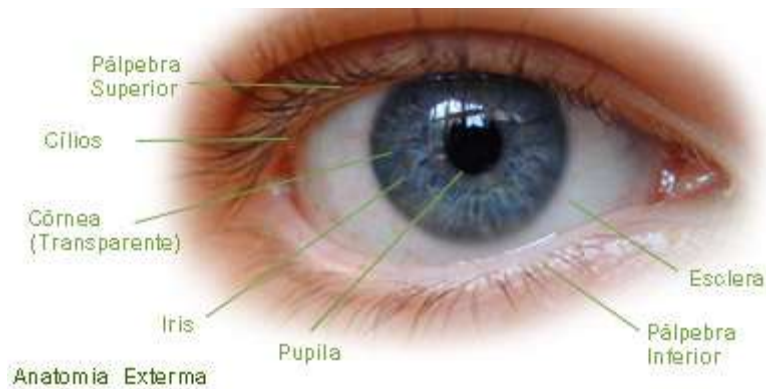


Figura 1.4: Olho humano destacando a íris.

1.4- Reconhecimento de Retina

Assim como o reconhecimento da íris, o da retina é muito eficiente na identificação individual das pessoas. No entanto, trata-se de um método intrusivo que requer proximidade da câmera ao olho para se obter a imagem precisa. Como se pode ver na Figura 1.5 [9], a retina é o nervo que se localiza no fundo do globo ocular. Desde 1930, cientistas estudam a retina e afirmam que os padrões dos vasos sanguíneos são particulares, exclusivos a cada indivíduo e podem ser utilizados como elemento identificador de pessoas. A tarefa, portanto, se baseia no reconhecimento desses padrões. Apesar de muito eficiente, o reconhecimento da retina é menos utilizado devido à necessidade de aproximar a câmera dos olhos. Os dispositivos atuais ficam a cerca de 1,5 cm da córnea, o que causa desconforto aos usuários e, por isso, a identificação por meio da retina só é recomendada quando a identificação é de grande relevância, não podendo ocorrer erro.

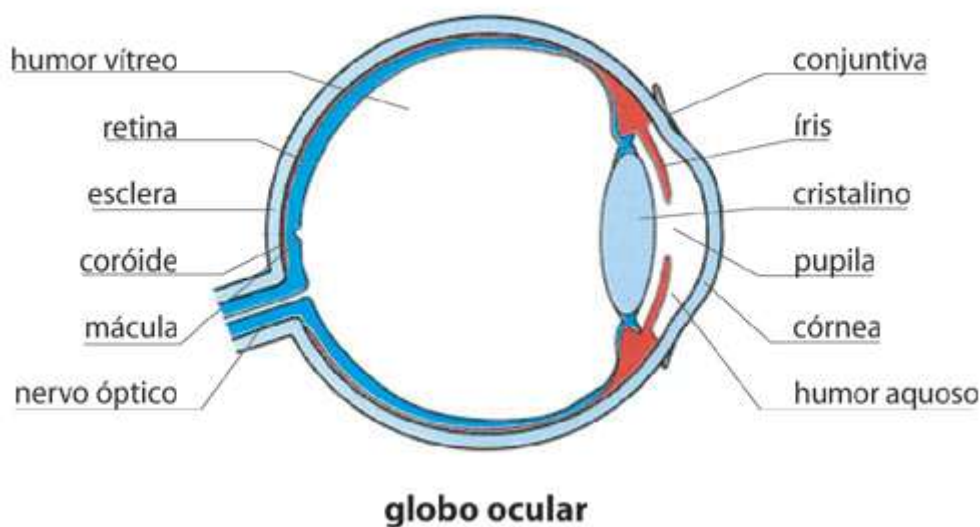


Figura 1.5: Aspecto do globo ocular humano com a retina.

1.5- Reconhecimento Facial

O reconhecimento de face tem sido objeto de dedicada atenção do ser humano no seu dia a dia. O simples ato de receber uma pessoa em uma residência faz com que o anfitrião desencadeie inconscientemente os dispositivos cerebrais visando ao reconhecimento da face do visitante. A automação deste procedimento está sendo acompanhada pela evolução dos sistemas computacionais que buscam automatizar as formas de reconhecimento faciais por meio de imagens digitais estáticas e aquelas dinâmicas em forma de vídeos, permitindo assim a autenticação biométrica e a interação dos seres humanos com as multimídias.

As pesquisas de reconhecimento de faces são motivadas pelo desafio que essa tarefa impõe considerando a multiplicidade de situações, notadamente aquelas que implicam nas alterações que uma face pode sofrer, rapidamente, de um instante para outro, tais como gestos, poses, adereços etc. Trata-se de um campo de aplicação muito vasto, pois tudo na vida é realizado por pessoas, mas nem todas as pessoas podem fazer tudo e por isso a necessidade da identificação dentro de cada finalidade. A constatação do aumento do número de pessoas nos centros urbanos seguido do aumento de empresas dos vários setores e das questões forenses faz da identificação individual de pessoas uma medida preponderante no ordenamento das respectivas atividades. O avanço

tecnológico é o principal parceiro neste trabalho à medida que disponibiliza novas câmeras digitais, acesso à internet e equipamentos de fácil deslocamento cuja portabilidade pode agilizar todo o processo.

O reconhecimento de faces tem várias vantagens sobre outras técnicas biométricas como o reconhecimento por meio da íris ou da retina. O sistema é preparado para identificar faces presentes em imagens e em vídeos, automaticamente, de maneira natural e não intrusiva.

Também para o reconhecimento de faces pode-se apontar duas linhas básicas de utilização: a autenticação e o reconhecimento.

A autenticação é do tipo um para muitos, ou seja, para o reconhecimento faz-se a captura da imagem da face e compara essa com todas as imagens existentes em um banco de dados ou a compara com uma lista de natureza pré-estabelecida como, por exemplo, uma lista de suspeitos dos quais obviamente deseja-se rápida identificação. Esta última lista é de tamanho bem menor do que um banco de dados geral, tornando-se muito mais rápidas a busca e a conclusão dos procedimentos.

O reconhecimento de face é um problema visual. A face está sujeita à variação de iluminação, pose, expressão e, mesmo sendo uma imagem tridimensional, a identificação é baseada em imagem bidimensional. Com tantas variáveis presentes, é preciso tomar cuidado para diminuir o máximo a taxa de erro e, por isso, o padrão de reconhecimento é composto de quatro fases distintas: detecção, alinhamento, extração de características e comparação, conforme o diagrama da Figura 1.6.

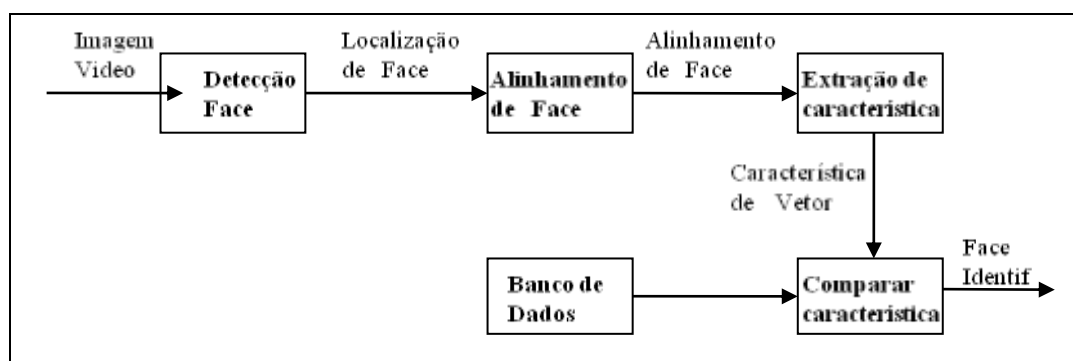


Figura 1.6: Diagrama da sequência para o reconhecimento de face.

A fase de detecção é a localização da face pela segmentação da área da imagem.

Em caso de vídeo, existe a necessidade de se utilizar o recurso chamado *face tracking* que processa a imagem em movimento.

A *fase de alinhamento* serve para destacar com mais exatidão a localização e fazer a normalização da área onde a face se situa, fornecendo a escala de cada face detectada. Nesta fase, os componentes faciais como olhos, nariz e boca e ainda a linha de contorno da face são definidos. Com base nos pontos de localização, a imagem é normalizada primeiramente de acordo com as propriedades geométricas, tais como tamanho e pose para depois ser normalizada de acordo com as propriedades fotométricas, como iluminação e escala de cinza, ou seja, a cor.

A *fase de extração* das características é executada após a face ser normalizada geométrica e fotograficamente. Nesta fase, obtêm-se as informações efetivas que são utilizadas para distinguir a face normalizada de todas aquelas de diferentes pessoas existentes na área da imagem.

A *fase de comparação* é aquela na qual o vetor de características extraído da face que está sendo manipulada é comparado com aqueles vetores derivados das imagens armazenadas em um banco de dados.

O resultado do reconhecimento de face depende altamente das características que são extraídas da imagem modelo (aquela que se deseja reconhecer) e dos métodos de classificação utilizados para distinção de faces e a extração efetiva de características.

1.5.1- Espaço-imagem e o espaço-face

O espaço-imagem é o ocupado plenamente pela imagem. O espaço-face é aquele ocupado pela face. Para especificar uma imagem aleatória no espaço-imagem, é preciso especificar todos os valores dos pixels. Assim, a dimensão nominal do espaço é expressa pela representação do número pixel contido nas M linhas e nas N colunas. Uma grande proporção de pontos no espaço-imagem não quer dizer que representa fisicamente uma face. Assim, o natural é que as imagens-faces sejam confinadas em um subespaço de um espaço-imagem.

É comum modelar o espaço-face como principal destaque no espaço-imagem. O propósito dessa modelagem é analisar o subespaço e determinar os modos principais para serem computados em função dos valores dos pixels.

1.5.2- Reconhecimento para autenticação e vigilância

O problema de detecção e reconhecimento facial tem vários níveis de aplicação e dificuldade. Por ser um órgão em três dimensões, a face está suscetível a uma variedade de modificações por parte da pessoa, sejam por situações intrínsecas ao corpo, como idade, posição, estado físico, estado emocional, doença, mímicas etc, sejam por fatores externos como iluminação, cores de ambientes, maquiagem, objetos diversos.

Dessa forma, o reconhecimento de face defronta-se com dois desafios: o primeiro é encontrar na imagem a face [10,11] e o segundo é extrair da face as características que possam gerar um modelo para o armazenamento destinado a comparação [11]. Os sistemas de identificação geralmente utilizam um conjunto de imagens em um determinado intervalo de tempo. Isso facilita a procura de características da face em uma sequência de imagem em movimento, uma vez que os diferentes ângulos e as expressões podem facilmente distinguir a face e seus detalhes do restante do ambiente. A partir dessas informações, cria-se um modelo que então é utilizado para a comparação de padrões com a face a ser autenticada ou reconhecida.

Os sistemas de reconhecimento de faces para autenticação, são relativamente simples de implementar, pois o usuário se apresenta para a identificação sabendo que a sua imagem será comparada com outras já cadastradas em um banco de dados. A dificuldade acentua quando se trata do reconhecimento de pessoas que se encontram em ambientes, que pretende-se a identificação e que nem sempre desejam ser reconhecidas. Nestes casos elas não se apresentam aos dispositivos destinados a esse fim, dificultando o processo.

Além da autenticação, a área de vigilância é outro grande emprego dos sistemas de reconhecimento de faces. Ou seja, existe a necessidade de se avaliar permanentemente um determinado ambiente que está sendo guardado. Por exemplo, a vigilância de um ambiente no qual não podem permanecer pessoas por alguma razão. Uma câmera apontada para esse espaço tenta detectar movimentação e, em caso positivo, dispara algum aviso. O sistema de detecção de movimento procura, em uma sequência de imagens, sinais que confirmem a existência de movimentação no ambiente monitorado, marcando os locais onde o movimento ocorreu.

Outra aplicação é o reconhecimento de rostos em uma multidão usando a técnica *tracking*, que habilita o sistema a acompanhar o movimento de alguma pessoa encontrada no campo de visualização. Essa técnica permite que o equipamento seja ligado só quando existe necessidade. Isto é importante porque é menor a possibilidade do operador, que acompanha o monitor visualmente, deixar passar detalhes. A aplicação

dessa modalidade é importante em caso de trabalho ininterrupto, pois, contribui para evitar equívocos.

Embora tanto a autenticação como a vigilância tratem do casamento de padrões, para a vigilância tem-se um complicador que é o fato do objeto alvo não fazer questão de se aproximar ou mesmo se esquivar, dos dispositivos de gravação para o reconhecimento. As variáveis são muitas, podendo-se citar algumas como a distância da câmera, o deslocamento horizontal e vertical, a iluminação etc.

No sistema de vigilância, a entrada de dados não é interrompida e também não se tem muito tempo para obter a resposta. Isto exige grande poder de processamento da máquina que roda o aplicativo de reconhecimento. Para contornar o problema do tempo de máquina, utilizam-se duas técnicas que se complementam. A primeira, de precisão questionável, baseia-se no movimento da imagem, é de rápida execução e restringe o espaço de busca das características, enquanto a segunda, baseada em estruturas complexas e de execução lenta, trabalha no espaço de busca restante [10].

Por isso, o processo de reconhecimento deve ocorrer a todo instante, levando em consideração todas as variáveis presentes quando do aparecimento do objeto e utilizar as modelagens que busquem tal finalidade [12].

1.6- Aplicações do reconhecimento de faces e os seus desafios

Como um dos muitos métodos biométricos não invasivos, a tecnologia de reconhecimento de face está sendo cada vez mais empregada para o reconhecimento de pessoas no dia a dia. A seguir, serão citadas algumas categorias de cenários para utilização da tecnologia de reconhecimento de faces em um futuro próximo.

1.6.1- Identificação de faces

Os sistemas de reconhecimento de faces identificam as pessoas pela suas imagens de faces [13]. Em contraste com os sistemas tradicionais de identificações, aqueles por cartões, por senhas etc., os sistemas de reconhecimento de faces partem do princípio de que existem conjuntos de pessoas que autorizam o reconhecimento de suas faces como forma de seleção ou triagem para entrar em determinado ambiente e ainda

podem fazer a identificação destas para que se tenha acesso a computadores, operações bancárias e outros.

Esse processo biométrico de identificação tem a vantagem de eliminar a utilização dos dispositivos tradicionais de identificação evitando o uso indevido desses, como na hipótese de cartões perdidos ou roubados serem utilizados causando transtornos aos titulares.

No futuro, serão utilizados para identificação em diversas situações na área comercial, nas escolas e universidades para o controle de estudantes, na emissão de documentos pessoais, na área industrial e na área de serviços.

Desde o ano 2000, a Organização Internacional de Aviação Civil adota o reconhecimento de face como o mais adequado dos métodos biométricos para o controle de passageiros nos embarques para viagens aéreas [14].

As aplicações em larga escala ainda impõem muitos desafios como a utilização sistêmica no reconhecimento de autenticidade das pessoas.

Alguns desafios da identificação de face são citados a seguir e exigem muito da engenharia computacional no sentido de disponibilizar um conjunto de hardware e software capaz de responder satisfatoriamente aos requisitos de alta velocidade com baixa taxa de erro [14].

São inúmeras as formas que esses sistemas podem contribuir para tornar mais céleres e eficazes os procedimentos burocráticos de controle e de gerenciamento tanto da área pública quanto da privada: o controle de fluxos imigratórios, a emissão de passaportes, a autenticidade da foto da carteira de motorista, do título de eleitor, da carteira de saúde, a identificação para permitir o acesso a veículos, a computadores, a programas e transações on-line, a sistema de educação a distancia, a provas e exames on-line, a sistemas de segurança de voos, a segurança de intranet e internet, a registros médicos, a vigilância de movimento em terminais, a vigilância em parque de diversão, a controle de acesso em tribunais, a possibilidade do reconhecimento de ladrões em lojas por meio de vídeos, a investigação e acompanhamento de suspeitos, a verificação de pessoas indesejáveis em determinado ambiente, o reconhecimento por meio da captura da imagem da face de pessoas criminosas etc.. A Figura 1.7 mostra uma das formas de captura da imagem de faces por leitora óptica cujo sinal é enviado ao computador para o devido processamento. Este exemplo foi feito com a foto do autor deste trabalho.



Figura 1.7: Disposição do reconhecimento de face. Leitora óptica-computador.

Um exemplo notável da utilização da tecnologia de reconhecimento de face é o emprego pelo departamento de migração dos Estados Unidos que seleciona aproximadamente 50.000 pessoas por ano, para receber visto de permanência, a chamada loteria do *Green card*. Os pedidos desses vistos, cada um, têm a foto facial do interessado. O sistema compara as fotos dos pedidos com um banco de dados para evitar que sejam concedidos vistos a quem já ganhou visto anteriormente. E, como as fotos são cadastradas no sistema, elas podem ser também utilizadas para outros processos de identificação, como, por exemplo, nas questões de controle migratório e de segurança.

1.6.2- Observações de pessoas e de lugares

Em uma atividade policial, a observação de pessoas ou lugares por meio de sistemas de reconhecimento não tem boa confiabilidade e por isso é uma operação que apresenta baixo nível de satisfação. O reconhecimento passa pelo desafio que representa o processamento de imagens submetidas a fatores de grande variabilidade como a clareza, as orientações de face e o movimento desordenado das pessoas.

Mesmo assim, as pesquisas continuam intensas visando à eficiência desses sistemas. Pelo mundo afora, centenas de câmeras são instaladas em prédios, monumentos e torres, estrategicamente posicionadas para observar pessoas e lugares.

1.6.3- Cartões de identificação

A aplicação da técnica de reconhecimento de face por meio de cartão, na essência, é a combinação de duas técnicas. A primeira é a gravação no cartão das características matemáticas da face durante a fase de registro. A segunda é a comparação dessas características com aquelas extraídas da face da pessoa, capturada por uma leitora, no momento da identificação. A precisão do cálculo da similaridade entre as características da face armazenada no cartão e a da face capturada depende do tempo transcorrido na comparação entre as duas imagens.

1.6.4- Investigação policial

O reconhecimento de face na investigação policial é realizada por um sistema eficiente que ajuda na rápida captura dos suspeitos mesmo com informações incompletas de suas identidades e até em alguns casos com as informações baseadas somente no esboço do retrato falado.

Os oficiais e detetives de posse das fotos, podem enviá-las pela internet e receber o retorno com o resultado do reconhecimento sem precisar deslocarem-se centenas de quilômetros para obterem as respostas da comparação da imagem de uma pessoa suspeita com aquelas dos arquivos forenses. Também a tecnologia do reconhecimento de face permite procurar suspeitos nos registros de maníacos sexuais permitindo que as testemunhas verifiquem vários conjuntos de fotos para ajudarem na elucidação de crimes desta natureza.

1.6.5- Indexação de banco de dados

Durante os anos 90, a recuperação das imagens a partir de textos-base contidos nos grandes bancos de dados ficou praticamente inviável devido à indexação manual [15]. Para resolver o problema, as palavras-chave passaram a ser indexadas pelas características visuais das próprias imagens. Esse procedimento foi um grande avanço na questão de rapidez, embora a indexação pelas características visuais das imagens tivessem grande limitação pois cada imagem apresenta diferente iluminação, textura, estado de preservação etc.

Recentemente, pesquisadores vêm experimentando a combinação tecnológica das análises de imagens com as de reconhecimento de faces para melhorar a qualidade e o grau de acerto das imagens recuperadas. Embora o reconhecimento de faces tenha utilizado principalmente as bases de dados destinados às atividades forenses, atualmente aquelas técnicas passaram a ser empregadas em outros tipos de base de dados, como álbuns de fotografias ou álbuns de outras imagens. O desempenho desse sistema de recuperação é ainda baixo em função do tamanho desses arquivos e das condições com as quais as imagens são capturadas.

1.6.6- Identificação de imagem por método interativo

Em um álbum pessoal de fotos digitais têm-se muitas imagens com ou sem faces humanas. Nessas fotos, estão misturados objetos, paisagens e seres humanos. O reconhecimento das faces nas imagens pode ser realizada com procedimentos que limitam o espaço observado para se encontrar mais rapidamente a imagem da face [16].

Quando o sistema detecta em uma foto faces de pessoas, ele tenta comparar estas faces com aquelas contidas em arquivos e informa o usuário sobre a semelhança. O usuário confirma se é aquela face a desejada. Se não confirmar, o sistema faz nova busca no arquivo. Uma vez confirmada a comparação, o nome é gravado em todas as fotos que têm aquela como original.

1.6.7- Identificação de imagem por lembrança

Uma das tarefas mais importante é reconhecer uma face utilizando somente a lembrança. A ideia é implementada usando mapas auto-organizados que alocam posições de imagens vizinhas [17]. O sistema pede ao usuário a imagem que ele considera semelhante àquela desejada e, em seguida, ele mostra as novas imagens mais parecidas com aquela selecionada pelo usuário. E, assim, o usuário interage com o sistema até encontrar a imagem desejada.

1.6.8- Utilizando multimídia

Faces humanas são frequentemente vistas em jornais, filmes, vídeos e outras multimídias e, assim, podem ser utilizadas para indexar os conteúdos destas mídias.

Uma das dificuldades da utilização direta do reconhecimento de face em multimídia é que geralmente o arquivo para comparação não está disponível. Desta forma, a identidade da pessoa é revelada por um método que pesquisa na *web* ou textos processados, por exemplo, nomes e imagens para identificar certa pessoa [18]. O trabalho inicial é relacionar os nomes com as faces encontradas nas várias mídias e então o detector de faces cria uma galeria de imagens identificadas [19]. A partir disso, ao deparar com uma imagem na mídia sem nome, o sistema compara as faces contidas na imagem com aquelas da galeria de fotos e retorna uma mensagem informando a identidade da face.

Outro sistema cria uma biblioteca de vídeo digital destinada a melhorar o reconhecimento das pessoas [20]. Milhares de horas de vídeo são indexadas e arquivadas para serem pesquisadas e recuperadas por usuários via rede de computadores.

Também foi desenvolvido um sistema para identificação em tempo real por meio de um conjunto de câmeras que acompanham as pessoas e sumarizam as faces humanas no domínio de compressão de vídeo [21]. O detector utiliza dados MPEG comprimidos para a detecção da face mesmo estando em movimento. Como a orientação das faces e a iluminação não são controladas, a exatidão do reconhecimento de face é relativamente baixa.

1.6.9- Reconhecimento da cor da pele

Embora o ponto ótimo de cada interface seja reconhecer e entender o movimento do corpo humano, obter este ponto é estar certo do local e do movimento da face. A análise da cor da pele oferece boa sugestão para a localização da face e, neste sentido, estudos de imagens em deslocamento estão sendo realizados baseados na segmentação da imagem de cor base cujo algoritmo implementa um sistema de cores não estacionárias [22]. Embora alguns trabalhos tenham alcançado relativo sucesso, esta modalidade de reconhecimento ainda precisa de muito estudo.

1.6.10- Reconhecimento de emoções

É a situação em que o computador está habilitado a entender o estado emocional das pessoas e assim interage com os usuários. A forma mais clara de uma pessoa mostrar seu estado emocional é por meio das expressões faciais. Um sistema criado para detectar as diferentes expressões faciais faz a classificação e aponta o estado emocional da pessoa por meio de um monitor de vídeo [23].

1.6.11- Outras aplicações

Alguns dos tipos de aplicações vistos acima requerem compatibilidade entre o sistema de reconhecimento de face e o domínio do espaço e o tempo no qual ela está inserida.

Uma aplicação que está desassociada com o tempo é o reconhecimento de fotos antigas na qual a idade e a qualidade das fotos passam a ser complicadores para se estabelecer o algoritmo de processamento, pois a esse têm que ser incorporadas outras variáveis como sinais de dobras, perda parcial da imagem, o aspecto de envelhecimento que altera o valor do cinza de cada pixel etc.

Outra área em que as pesquisas dos sistemas inteligentes avançam é o reconhecimento de um determinado ambiente.

Com essa percepção se pode alcançar um estágio significativo de interatividade homem-máquina que irá beneficiar o trabalho em diversos ramos das atividades humanas.

1.7- A limitação do sistema de reconhecimento de faces

Os sistemas de reconhecimento de faces não são ainda consistentes o bastante para proporcionar em qualquer situação alto índice de acerto.

Em utilização de pouca intensidade com arquivos pequenos e fotos bem definidas o índice de acerto chega bem próximo a 100%. A probabilidade de erro aumenta à medida que os arquivos para a verificação crescem de tamanho com o consequente aumento do tempo de máquina e ainda se as imagens originais estiverem em mau estado de conservação e/ou os gestos e ângulos de tomadas forem variados etc.

1.8- Trabalhos Realizados

Muitas pesquisas em reconhecimento de faces já foram e estão sendo realizadas nos mais variados métodos, sempre na perspectiva de encontrar um modelo mais eficiente. São citadas nesta seção alguns desses trabalhos.

Mendonça [24] fez uma comparação entre os métodos de normalização da iluminação utilizados para melhorar a taxa do reconhecimento facial. O objetivo dessa pesquisa foi o de avaliar os algoritmos de normalização da iluminação como: Principal Component Analysis(PCA), Vector Quantization(VQ), Wavelets e Rede Neural Artificial (RNA). Segundo o trabalho, os resultados obtidos apresentaram uma boa taxa do reconhecimento facial, exceto para o método de normalização wavelet .

Kim [25], fez uma experiência utilizando a Principal Component Analysis (PCA), em que o tempo de processamento foi excessivo para o reconhecimento e também ele afirmou que os autovetores que representam cada face são muito grandes.

Matos [26] propôs um método de reconhecimento de faces baseado na seleção de coeficientes DCT associados a distância. A taxa de acerto foi compatível com o estado da arte e mostraram-se eficazes em relação às alterações de iluminação e expressões faciais.

Amarapur [27] apresentou um trabalho em que a limitação do modelo é que o desempenho só é bom para faces com tomada frontal. Se a imagem é tomada em diferentes ângulos, a taxa de erro é muito grande.

Li [28] fez a redução do espaço de características com aumento da taxa de acerto, por meio de um modelo melhorado de algoritmo genético chamado Chaos.

Alves [29] aplicou o algoritmo genético para o reconhecimento de faces com a entrada sendo o número de segmentos áureos (razão áurea) encontrados na face. Essa entrada permitiu a identificação de indivíduos.

1.9- Considerações Finais Deste Capítulo

Neste capítulo foram mostradas referências a muitos sistemas de reconhecimento de faces em vários cenários de aplicação. Atualmente, a tecnologia do reconhecimento de faces avança pelos setores produtivos, comerciais e de serviços que passam a utilizá-la como uma necessidade de aumentar a qualidade dos resultados quando se trata de controle de acesso a ambientes específicos e de proteção patrimonial. Cada vez mais as

peessoas e os lugares estão sendo observados com a finalidade de proteger patrimônios, assegurar a convivência harmoniosa entre as pessoas ou identificar suspeitos de crimes. Por isso, a colaboração e a interação entre os pesquisadores das universidades e das indústrias podem tornar o sistema de reconhecimento de faces mais confiável, célere e robusto.

As pesquisas buscam incessantemente melhorar a exatidão do reconhecimento e, para tanto, elas trabalham com a combinação de métodos biométricos e de segurança e, nessa direção, além das digitais, da íris, da retina e da face, já se utiliza em muitas aplicações o som da voz. Um exemplo de métodos combinados é a segurança em aeroporto onde o sistema de reconhecimento de face pode também trabalhar junto com escaneador de bagagem, detector de metal, detector de componente químico e segurança de pontos de checagens.

A conclusão deste capítulo é a importância e a explícita maneira de como o reconhecedor de face pode ser utilizada no dia a dia do ser humano, em um futuro próximo. Ao se dirigir para o trabalho, o reconhecedor de face instalado no carro autoriza ou não a ignição do motor. Caso se utilize ônibus ou metrô, a foto do cartão pré-pago é comparada com aquela retirada por leitora no momento do embarque. As entradas de edifícios serão equipadas com reconhecedor de faces para permitir ou não o acesso às dependências do prédio. Em viagens, o reconhecedor pode fornecer segurança na utilização do cartão de crédito e passaporte. Ao chegar a casa, o sistema reconhece a face do proprietário e, então, abre a porta para sua entrada.

As limitações das tecnologias mostradas são mais um desafio aos pesquisadores que não devem cessar seus estudos pela enorme perspectiva futura da área.

O próximo capítulo apresenta os aspectos teóricos dos classificadores: algoritmo genético (AG), rede neural artificial (RNA) e análise de componentes principais (PCA).

CAPÍTULO II

FUNDAMENTAÇÃO TEORICA DOS CLASIFICADORES: ALGORITMO GENÉTICO (AG), REDE NEURAL ARTIFICIAL (RNA) E ANÁLISE DE COMPONENTES PRINCIPAIS (PCA)

2.1- Introdução

A mente humana é o mais complexo e maravilhoso sistema de organização de informações que se conhece até hoje. Como as informações são armazenadas pelo conteúdo, pode-se ver tudo que está na mente em termos de passado, presente e futuro. As pesquisas computacionais buscam insistentemente algoritmos sistematizados parecidos com a mente humana. É o conceito *lifestream*. O caminho é longo e difícil, mas os passos estão sendo realizados pela ciência.

O algoritmo genético (AG) é um método computacional de busca baseada no mecanismo de evolução natural e na genética. Pelo AG, uma população de possíveis soluções (indivíduos) para o problema em questão evolui de acordo com os conceitos da seleção natural, em que os indivíduos melhor adaptados ao ambiente têm mais chance de se reproduzirem. O algoritmo genético opera com uma população finita de indivíduos codificados, dos quais são extraídas informações utilizadas no cálculo do seu desempenho na solução do problema em questão, não necessitando empregar operações matemáticas, como equações de funções ou derivadas. Também, os AG não possuem restrições quanto ao tamanho do espaço de busca [30].

A RNA assemelha-se ao cérebro humano, pois o conhecimento é adquirido pelo aprendizado e as forças de conexão entre os neurônios são utilizadas para armazenar o conhecimento.

A PCA é uma técnica de padrão normalmente utilizada para redução de dados em reconhecimento estatístico de padrões e processamento de sinais. É um algoritmo auto-organizado que descobre padrões significativos ou características nos dados de entrada, dispondo de um conjunto de regras de natureza local que mapeia a entrada e a saída com propriedades desejáveis específicas.

Neste capítulo, é apresentado os aspectos teóricos dos classificadores: algoritmo genético (AG), rede neural artificial (RNA) e análise de componentes principais (PCA). São algoritmos que demonstraram habilidades no reconhecimento de faces, embora ainda sem guardar semelhança com mente humana.

2.2- Algoritmos Genéticos (AG)

Algoritmos genéticos são métodos de otimização e busca inspirados na teoria da evolução de Darwin que trata as possíveis soluções do problema como “indivíduos” de uma “população”, utilizando um processo evolucionário em que os indivíduos sejam solução de um problema [31, 32].

A ideia era trazer para os sistemas de computação os mecanismos de adaptação natural, utilizando os princípios da natureza como meio para a solução de problemas [33]. Na verdade, os algoritmos genéticos são um dos paradigmas de computação evolutiva. Podem ser citados ainda a programação genética, a programação evolutiva, as estratégias evolutivas e os sistemas classificadores.

A utilização de algoritmos genéticos para o tratamento de otimizações vem apresentando resultados satisfatórios para as aplicações em que os métodos tradicionais de programação não são capazes de solucionar em um tempo computacionalmente viável.

O princípio dos algoritmos genéticos é criar um conjunto inicial com várias possíveis soluções para um problema. Esse conjunto é chamado de população inicial e cada solução pertencente a essa população é chamada de indivíduo. Estes indivíduos devem ser codificados de forma a constituir seu material genético. Neste caso, considera-se que o material genético é um cromossomo, e este, é um indivíduo propriamente dito.

Essa codificação dos indivíduos visa possibilitar a aplicação de operadores genéticos e do conceito de seleção natural sobre as soluções existentes, podendo assim combinar seus materiais genéticos buscando ao longo de várias gerações, indivíduos

cada vez mais adaptados, ou seja, um mais adaptado passa a ser uma possível solução para o problema proposto.

A evolução de um algoritmo genético ocorre de acordo com uma série de parâmetros que devem ser informados inicialmente para o algoritmo. Embora existam na literatura sugestões para tal, alguns desses parâmetros não possuem regras para sua escolha, sendo assim definidos empiricamente e sofrendo ajustes de acordo com a qualidade dos resultados obtidos. Essas escolhas implicarão diretamente na forma com que o algoritmo genético irá se portar diante do problema escolhido. Por exemplo: um parâmetro relacionado às taxas de cruzamentos entre os indivíduos de uma população pode possibilitar uma convergência prematura desse conjunto de soluções ou então, em um caso totalmente oposto, fazer com que essa população nunca convirja [31, 34].

2.2.1- Desenvolvimento do algoritmo genético

Antes de começar a desenvolver um sistema dessa natureza, deve-se, primeiramente, conhecer claramente o tipo de problema de que se trata e analisar o formato de uma possível solução para ele. É importante ter certeza quanto às características de uma solução e também quanto às possíveis restrições que devem ser levadas em consideração. O passo seguinte é: como se pode codificar essa solução (indivíduo)? A codificação de uma solução é a representação da mesma por meio de um material genético, ou seja, de uma sequência de genes. Uma alternativa muito utilizada é a representação de uma solução por meio de um sistema binário. Por exemplo: uma sequência como “100101” é um indivíduo com 6 (seis) genes, possuindo 2 (dois) alelos, “0” (zero) e “1” (um).

Durante o processo de codificação, deve ser possível avaliar um indivíduo a partir de seu código genético (cromossomo). Esta é a medida de avaliação do genoma que resulta em uma nota. Esta medida chama-se adaptabilidade ou *fitness*, que é a medida de quão otimizado ou adequado é um genoma para resolver um determinado problema ou meta. Quanto maior essa nota, mais adaptado ou adequado está o genoma ao problema específico. O objetivo do algoritmo genético é otimizar as soluções para a função *fitness* específica, sendo que a população sofre a seleção baseada na nota de adaptabilidade.

Deve-se ainda visar a uma representação na qual a aplicação dos operadores genéticos (ainda não descritos) seja facilitada.

A representação deve garantir que todo o domínio de soluções possa ser alcançado. Uma vez estabelecida a forma de codificação das soluções, pode-se agora introduzir de modo mais claro e específico os conceitos de operadores genéticos e de seleção natural.

2.2.2- Operadores genéticos

A partir da observação da forma com que os seres vivos evoluem, foram desenvolvidos alguns operadores genéticos com a finalidade de manipular e transmitir as cargas genéticas dos indivíduos ao longo de várias gerações. São eles: operador de cruzamento e mutação. A seguir, são apresentados mais detalhes destes operadores.

2.2.2.1- Operador de cruzamento

O operador de cruzamento, conhecido como crossover é um operador responsável por combinar o material genético de dois indivíduos distintos, gerando dois filhos com características combinadas de seus pais. É capaz de analisar diversos locais no espaço de busca, sendo aplicado de acordo com uma taxa de crossover;

Existem algumas variações quanto à implementação do operador de crossover.

Para aplicar esse operador, realizando uma combinação de dois materiais genéticos, é necessário estabelecer um ponto chamado de ponto de corte. Assim, como o nome já indica, esse ponto corta o cromossomo em duas sequências determinando que a primeira parte de um cromossomo será combinada com a segunda parte de outro, e vice-versa.

Após essa combinação, os indivíduos gerados apresentam o mesmo tamanho de seus progenitores, sendo constituídos por uma parte de cada um deles. É comum realizar também dois cortes nos cromossomos para aplicar o operador de crossover e, sendo assim, a troca do material genético entre os dois indivíduos é entre esses dois pontos.

Essas duas implementações citadas caracterizam respectivamente o crossover com um ponto de corte e o crossover com dois pontos de corte. Uma terceira alternativa para esse operador é conhecida como crossover com máscara. Neste caso, é gerada uma máscara do mesmo tamanho de um cromossomo. Analisa-se a máscara, e os genes atribuídos ao indivíduo gerado são obtidos a partir de uma regra. Quando o valor da

máscara é verdadeiro (um), utiliza-se o gene de um pai A, quando o valor da máscara é falso (zero), utiliza-se o gene de um pai B.

2.2.2.2- Operador de mutação

O operador de mutação é responsável por garantir a diversidade da população, realizando pequenos ajustes nas soluções encontradas. Ele é capaz de evitar que o algoritmo genético fique estagnado em um mínimo local, modificando alguns genes da população de acordo com uma taxa de mutação.

A implementação do operador de mutação é bem simples. Durante uma geração, alguns genes são selecionados aleatoriamente para sofrer mutação. Quando se utiliza uma representação binária, basta inverter o gene selecionado. Por exemplo: seleciona-se um gene “n” e, se este possui valor 1 (um), então é trocado para 0 (zero). Em casos de representação diferente da binária e existindo mais de dois valores de alelos, deve-se estabelecer uma regra para a substituição do gene encontrado.

2.2.3- Métodos de seleção

Para que seja possível aplicar o operador de crossover é necessário selecionar um conjunto de indivíduos para combinar seus materiais genéticos. Esse processo de seleção é conhecido como seleção natural, já citado no início desta seção como a explicação de Charles Darwin [33] para a evolução e adaptação dos seres vivos ao longo do tempo. Para que exista uma convergência adequada da população, procura-se reproduzir da melhor maneira esse processo de seleção executado pela natureza. Para que isso seja possível, é preciso levar em consideração o grau de adaptação, fitness, de cada indivíduo como referência para determinar quais deles serão submetidos ao cruzamento. Existem várias técnicas utilizadas para simular a seleção natural. Dentre elas podem ser citadas: seleção proporcional ao desempenho, seleção por torneio, seleção por posição e seleção por truncatura. O número de indivíduos que são selecionados depende de uma taxa de cruzamento especificada no início do sistema.

2.2.4- Métodos de atualização da população

A população manipulada pelo GA é de tamanho fixo, ou seja, possui um número pré-estabelecido de indivíduos e esse número não pode ser aumentado ou diminuído.

Tendo em vista que essa característica deve ser mantida, depara-se com um novo problema: a substituição de indivíduos na população.

Quando são gerados indivíduos por meio dos operadores genéticos, esses indivíduos devem ser colocados na população e, para que o tamanho da população seja mantido é necessário que outros indivíduos sejam retirados dela.

Depois de um cruzamento existem vários métodos para realizar a substituição de indivíduos na população. Os métodos mais utilizados são [33]:

- substituição imediata: os novos indivíduos gerados substituem os seus geradores;
- substituição por fator cheio: cada novo indivíduo substitui o indivíduo mais parecido com ele na população.

2.2.5- Passos de um Algoritmo Genético

A Figura 2.1 mostra os principais de um algoritmo genético.

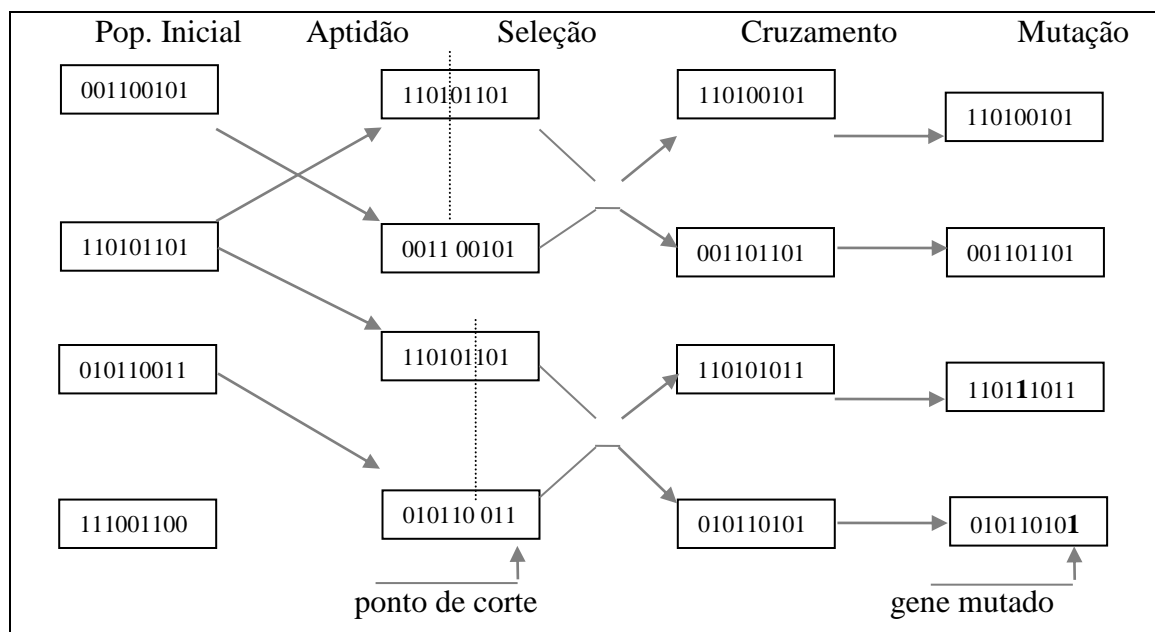


Figura 2.1: Principais passos de um algoritmo genético.

Assim, os principais passos de um algoritmo genético são:

1. Criar a população inicial gerando aleatoriamente os genes dos indivíduos;
2. Realizar a avaliação do fitness (aptidão) de cada indivíduo;
3. Submeter a população ao processo de seleção natural. Neste passo está incluso tanto a própria seleção dos indivíduos, quanto a aplicação dos operadores de cruzamento e de mutação;
4. Realizar uma avaliação do fitness (aptidão) dos indivíduos gerados pelo passo anterior;
5. Atualizar a presente população com os indivíduos gerados, mantendo o tamanho da população inalterado; e
6. Caso seja satisfeito um critério de parada, retorne o melhor indivíduo, caso contrário volte ao passo "4". O critério de parada pode ser tanto um número máximo de gerações quanto um erro mínimo desejado para uma solução.

2.3- Rede Neural Artificial (RNA)

Redes neurais artificiais são técnicas computacionais que apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes. Elas adquirem conhecimento por meio da experiência. Uma grande rede neural artificial pode ter centenas ou milhares de unidades de processamento; já o cérebro de um mamífero pode ter muitos bilhões de neurônios.

2.3.1- Funcionamento do neurônio biológico

O sistema nervoso humano é responsável pela tomada de decisões e pela adaptação do organismo ao meio ambiente. Esta função é realizada por meio de um aprendizado contínuo. Esse sistema é constituído de células que são responsáveis pelo seu funcionamento denominadas de neurônios ou unidades de entrada e de saída. Pode-se observar na Figura 2.2 [35] as partes constituintes do neurônio. O cérebro humano apresenta aproximadamente 10 bilhões de neurônios e cerca de 60 trilhões de conexões entre eles [36]. Essas entradas e saídas estão maciçamente conectadas umas com as outras na composição de uma RNA.

“Uma rede neural é um processador maciçamente distribuído paralelamente constituído de unidades de processamento simples que têm a propensão natural para armazenar conhecimento experimental e para torná-lo disponível para uso. Ela se assemelha ao cérebro humano em dois aspectos:

1. o conhecimento é adquirido pela rede por meio de um processo de aprendizagem e a partir de seu ambiente; e
2. forças de conexão entre os neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido [37].

Essas células recebem, geram e transmitem os estímulos que chegam ou partem do cérebro.

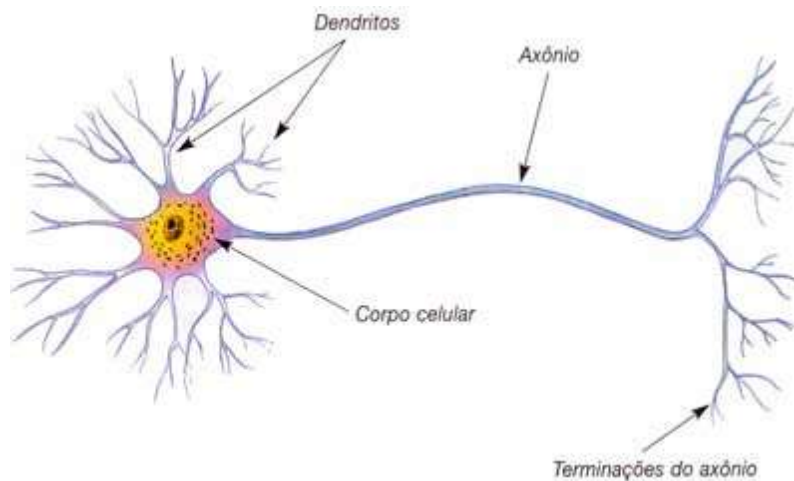


Figura 2.2: Neurônio biológico.

O neurônio é delimitado por uma fina membrana celular que possui determinadas propriedades, essenciais ao funcionamento da célula. A partir do corpo celular, projetam-se extensões filamentosas, os dendritos, e o axônio [36]. Os neurônios são definidos como células polarizadas capazes de receber sinais em seus dendritos e transmitir informações por seus axônios. Ao ser excitado, um neurônio transmite informações, por meio de impulsos chamados potenciais de ação, para outros neurônios.

Estes sinais são propagados como ondas pelo axônio da célula e convertidos para sinais químicos nas sinapses.

O neurônio biológico pode ser visto como o dispositivo computacional elementar do sistema nervoso, composto de muitas entradas e saídas. As entradas são formadas por meio das conexões sinápticas que conectam os dendritos aos axônios de

outras células nervosas. Os sinais que chegam por esses axônios são pulsos elétricos conhecidos como impulsos nervosos ou potenciais de ação e constituem a informação que o neurônio processa para produzir como saída um impulso nervoso no seu axônio.

Dependendo dos sinais enviados pelos axônios, as sinápses podem ser excitatórias ou inibitórias. Uma conexão excitatória contribui para a formação de um impulso nervoso no axônio de saída, enquanto uma sinápsse inibitória age no sentido contrário [36].

A partir do conhecimento da estrutura e do comportamento dos neurônios naturais, foram extraídas suas características fundamentais, utilizadas na criação de modelos de neurônios artificiais que simulam os reais. Estes neurônios artificiais são utilizados na formação das RNAs, se compõem-se como os principais elementos de processamento deles.

2.3.2- Modelo do neurônio artificial

Uma rede neural artificial é composta por várias unidades de processamento, cujo funcionamento é bastante simples. Geralmente, essas unidades são conectadas por canais de comunicação que estão associados a determinado peso. As unidades fazem operações apenas sobre seus dados locais que são entradas recebidas pelas suas conexões. O comportamento inteligente de uma rede neural artificial vem das interações entre as unidades de processamento da rede.

A operação de uma unidade de processamento proposta por McCulloch e Pitts em 1943, pode ser resumida da seguinte maneira:

- sinais são apresentados à entrada;
- cada sinal é multiplicado por um número ou peso que indica a sua influência na saída da unidade;
- é realizada a soma ponderada dos sinais que produz um nível de atividade; e
- se esse nível de atividade excede um certo limite (threshold), a unidade produz uma determinada resposta de saída.

Esse neurônio foi chamado de *Neurônio de McCulloch & Pitts*. A Figura 2.3 mostra esse neurônio. A operação desse neurônio pode ser resumida da seguinte forma:

- 1) os sinais são apresentados às entradas X_p ;

- 2) cada sinal é multiplicado por um peso W_p que indica a sua influência na saída do neurônio (efeito de excitação ou inibição da sinápsis);
 - 3) é realizada a soma ponderada dos sinais v_k , produzindo um nível de atividade;
- e
- 4) se esse nível excede um certo limiar, a unidade "ativa" sua saída y_k .

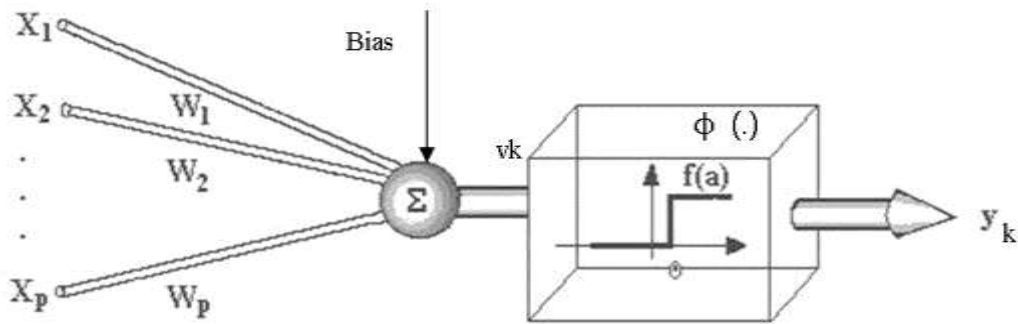


Figura 2.3: Modelo geral do neurônio de McCulloch & Pitts. Modelo geral do neurônio.

Esse modelo não apresenta somente a função de ativação limiar utilizada no neurônio de McCulloch & Pitts, mas sim uma função de ativação $\phi(.)$ que pode ser definida de várias formas. Também foi acrescentado no modelo uma polarização ou *bias*, representada por b_k , que tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação. A Equação (2.1) obtém o campo local induzido.

$$v_k = \sum_{i=1}^p W_{ki} \cdot x_i + b_k \quad (2.1)$$

A saída da rede é obtida pela Equação (2.2). Essa saída depende da função de ativação utilizada.

$$y_k = \phi(v_k) \quad (2.2)$$

onde:

y_k – saída da rede; e

$\phi(.)$: é a função de ativação do k-ésimo neurônio.

Assim, nessas equações:

i : índice do k-ésimo neurônio;

b_k : bias aplicado ao k-ésimo neurônio;

v_k : chamado de campo local induzido é a saída do combinador linear de soma do k-ésimo neurônio;

x_i : i-ésimo sinal de entrada do k-ésimo neurônio;

W_{ki} : i-ésimo peso sináptico do neurônio k; e

O parâmetro externo *bias* pode ser incluído na formulação matemática como sendo um peso sináptico associado a uma entrada de valor constante unitário.

2.3.3- Função de ativação

A função de ativação representada por $\phi(v)$ define a saída de um neurônio em termos do campo local induzido [37]. As funções utilizadas para o cálculo de ativação geralmente não são lineares para garantir a plena funcionalidade das RNAs com múltiplas camadas de neurônios. A seguir são mostradas algumas funções das RNAs.

2.3.3.1- Função limiar (degrau)

Utilizada no modelo de McCulloch e Pits, a função limiar modela a característica "tudo-ou-nada" desse neurônio e é expressa como mostrado na Equação (2.3).

$$\phi(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad (2.3)$$

onde:

$\phi(v)$ – função de ativação; e

v – valor de ativação.

Nos neurônios construídos com essa função, a saída y é igual a 0, caso o valor de ativação v seja negativo e 1 nos casos em que o valor de ativação é positivo.

A Figura 2.4 mostra uma função limiar.

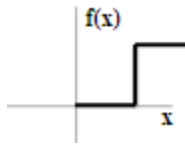


Figura 2.4: Gráfico da função de limiar

Deve-se lembrar que o valor de ativação (v) é composto pelo combinador linear e pelo bias como mostrado na Equação (2.1).

2.3.3.2- Função sigmóide

As funções mais utilizadas são as que possuem um formato sigmoidal, tais como a sigmóide, a tangente hiperbólica, seno, gaussiana.

Essa função, ao contrário da função limiar, pode assumir todos os valores entre 0 e 1. A representação mais utilizada para a função sigmóide é a função logística, definida pela Equação (2.4).

$$\phi(v) = \frac{1}{1 + e^{(-av)}} \quad (2.4)$$

onde:

a - parâmetro de inclinação da função sigmoide;

v - valor de ativação do neurônio.

A Figura 2.5 mostra uma função sigmóide.

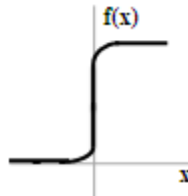


Figura 2.5: Gráfico da função sigmoide.

Porém, pode ser necessário que a função de ativação assuma valores entre 1 e -1. Esta característica traz benefícios analíticos. Para se obter tal intervalo de valores, utiliza-se a função *signum* quando se tratar de função de limiar e a função tangente hiperbólica, quando se tratar da função sigmóide, como pode-se observar a seguir.

2.3.3.3- Função signum

Esta função apresenta as mesmas características da função de limiar, porém está limitada entre 1 e -1. A função signum é representada pela Equação (2.5).

$$\phi(v) = b \frac{v}{|v|} \quad \text{para } v \neq 0 \quad (2.5)$$

onde:

b - limites inferiores e superiores ($b=|1|$ no gráfico); e

v - valor de ativação da unidade.

A Figura 2.6 mostra uma função signum.

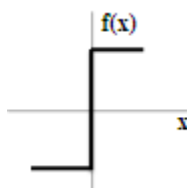


Figura 2.6: Gráfico da função signum.

2.3.3.4- Tangente hiperbólica

A função tangente hiperbólica também possui forma de "s", assumindo valores entre 1 a -1 . Ela é representada pela Equação (2.6).

$$\phi(v) = a \frac{e^{(bv)} - e^{(-bv)}}{e^{(bv)} + e^{(-bv)}} \quad (2.6)$$

onde:

a - parâmetro de inclinação da curva;

b - limites inferiores e superiores($b = |1|$ no gráfico);e

v - valor de ativação da unidade.

A Figura 2.7 mostra uma função tangente hiperbólica

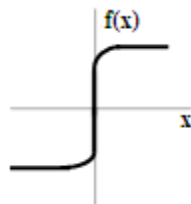


Figura 2.7: Gráfico da função tangente hiperbólica.

2.3.4- Quanto ao treinamento da RNA

Na RNA, o conhecimento é adquirido pelo processo de aprendizagem e pelas forças de conexão entre os neurônios, conhecidas como pesos sinápticos, que armazenam o conhecimento [37].

Neste trabalho, é utilizado para o treinamento o algoritmo *backpropagation* que utiliza os pares formados pela entrada e pela saída desejada para, por meio de um mecanismo de correção de erros ajustar os pesos de rede.

Trata-se da generalização do algoritmo do mínimo quadrado médio que utiliza a técnica do gradiente descendente, de forma iterativa, para minimizar uma função de custo igual à diferença média quadrática entre a saída desejada e a saída real da RNA.

A fase de treinamento, possui dois passos distintos: um passo de propagação e outro passo de retropropagação que busca minimizar o erro entre uma saída obtida pela rede e uma saída desejada. Para tal, utiliza-se o método do gradiente descendente que

requer funções de ativações que variam suavemente. Neste trabalho, utiliza-se a função de ativação tangente hiperbólica.

Os elos de conexão ou sinapse de um neurônio são caracterizados por um peso w_{kj} . Os sinais de entrada x_j são multiplicados por w_{kj} , onde k é o índice do neurônio e j é o terminal de entrada da sinapse. A junção aditiva é a responsável pela soma dos sinais de entrada ponderados pelas respectivas sinápses. A função de ativação restringe o valor de saída do neurônio, ou seja, processa o conjunto de entradas recebidas e o transforma em estado de ativação.

2.3.5- Quanto ao cálculo da RNA

Pelo algoritmo da retropropagação é possível saber como ajustar os pesos sinápticos de modo a reduzir o valor da função custo [37]. Na propagação para frente, como o sinal de entrada calcula-se a saída da rede para a iteração n . A saída do neurônio j é obtida pela Equação (2.7).

$$y_j(n) = \phi(v_j(n)) \quad (2.7)$$

Assim, $v_j(n)$ é o campo local induzido (soma ponderada de todas as entradas sinápticas acrescida do bias) do neurônio j na iteração n , obtida pela Equação (2.8).

$$v_j(n) = \sum_{i=0}^m w_{ji}(n) \cdot x_i(n) \quad (2.8)$$

onde:

$v_j(n)$ – campo local induzido do neurônio j na iteração n ;

m - número total de entradas da rede;

$w_{ji}(n)$ - peso sináptico que conecta o neurônio i ao neurônio j ; e

$x_i(n)$ - sinal de entrada do neurônio j ou, equivalentemente, o sinal de saída do neurônio i .

O sinal de saída é então comparado com a resposta desejada presente no conjunto de treinamento, de modo a se obter o sinal erro para o j -ésimo neurônio de saída, como mostrado na Equação (2.9).

$$e_j(n) = d_j(n) - y_j(n) \quad (2.9)$$

onde:

$e_j(n)$ –sinal erro para o j -ésimo neurônio de saída;

$d_j(n)$ – resposta desejada presente no conjunto de treinamento; e

$y_j(n)$ - resposta do j -ésimo neurônio de saída.

Na etapa de propagação para trás (retropropagação), o erro é propagado para trás em direção às entradas da RNA, atualizando os seus parâmetros livres conforme com uma regra de correção do erro, de modo que a saída atual da rede se aproxime da saída desejada. A forma como os parâmetros livres da rede são atualizados é obtida pela *regra delta*, como mostrado na Equação (2.10)[37].

$$\Delta w_{ji}(n) = \eta \cdot \delta_j(n) \cdot y_j(n) \quad (2.10)$$

onde:

$\Delta w_{ji}(n)$ -correção de peso aplicada na conexão entre o neurônio i e o neurônio j ;

η - parâmetro taxa de aprendizagem;

$\delta_j(n)$ - gradiente local do neurônio j ; e

$y_j(n)$ - sinal de entrada do j -ésimo neurônio ou o sinal de saída do i -ésimo neurônio.

O gradiente local do neurônio é definido como a derivada da função custo $\xi(n)$ que é a medida do desempenho de aprendizado, com relação ao campo local induzido $v_i(n)$ e aponta para as modificações necessárias aos pesos sinápticos de modo que a função custo tenha seu valor reduzido. O cálculo do gradiente local é possível tanto para neurônios da camada de saída, quanto para neurônios da camada intermediária. A taxa de aprendizagem η indica o tamanho do *passo* do algoritmo e, portanto, sua velocidade de convergência.

As equações utilizadas no cálculo do gradiente local podem ser encontradas com detalhes em [36]. Resumidamente, tem-se: o gradiente local de um neurônio localizado na camada de saída é obtido utilizando-se a Equação (2.11).

$$\delta_j(n) = e_j(n) \cdot \phi_j'(v_j(n)) \quad (2.11)$$

onde:

$\delta_j(n)$ - gradiente local de um neurônio localizado na camada de saída;

$e_j(n)$ - sinal erro para o j -ésimo neurônio de saída; e

$\phi_j'(v_j(n))$ - derivada da função de ativação.

O gradiente local de um neurônio localizado na camada intermediária j é obtido pela Equação (2.12).

$$\delta_j(n) = \phi_j'(v_j(n)) \cdot \sum_k \delta_k(n) \cdot w_{kj}(n) \quad (2.12)$$

onde:

$\delta_k(n)$ - gradiente local de um neurônio k localizado na camada de saída; e

$w_{kj}(n)$ - peso sináptico que conecta o neurônio j ao neurônio k .

A *função tangente hiperbólica* assume valores entre 1 e -1 e é representada pela Equação (2.13).

$$\phi(v) = a \tanh(b \times v) \quad (2.13)$$

onde:

$\phi(v)$ - função tangente hiperbólica assume valores entre 1 e -1;

a e b - constantes [38], elas são o parâmetro de inclinação da curva e os limites superior e inferior respectivamente; e

v - valor da ativação.

Nessa função de ativação, a transição entre os valores 1 e -1 é mais suave e contínua e sua derivada tende ao infinito. Essa função é muito popular entre os diversos desenvolvedores de redes neurais, porém alguns utilizam uma versão modificada dela que consiste em variar os limites da função de 0 até 1 (em vez de -1 a 1) com a finalidade de se obter menos custo computacionalmente, o que é desejado principalmente quando aplicado em redes neurais de grande dimensão.

2.4- Análise de Componentes Principais (PCA)

É um método que tem por finalidade básica, a análise dos dados utilizados visando a sua redução, eliminação de sobreposições e a escolha das formas mais representativas dos dados a partir das combinações lineares das variáveis originais.

O método da análise de componentes principais (PCA) é também chamado de Transformada Direta de Karhunen-Loève ou ainda Transformada Hotelling em homenagem a Kari Karhunen, Michel Loève [1907-1979] e Harold Hotelling. A PCA que foi derivada por Hotelling e por ele denominada como Método dos Componentes Principais, transforma as variáveis discretas em coeficientes descorrelacionados. Trata-se de um dos métodos estatísticos de múltiplas variáveis mais simples.

A PCA é considerada uma transformação linear ótima, dentre as transformadas de imagens, sendo muito utilizada pela comunidade de reconhecimento de padrões. Embora seja também usada em vetores de pequenas dimensões, a PCA é muito útil quando os vetores de características têm muitas dimensões e quando uma representação gráfica não é possível.

Conforme simplificado na Figura 2.8, a linha vermelha representa o componente principal como um arranjo de melhor distribuição dos dados, enquanto que a linha azul mostra o componente secundário que é perpendicular ao componente principal.

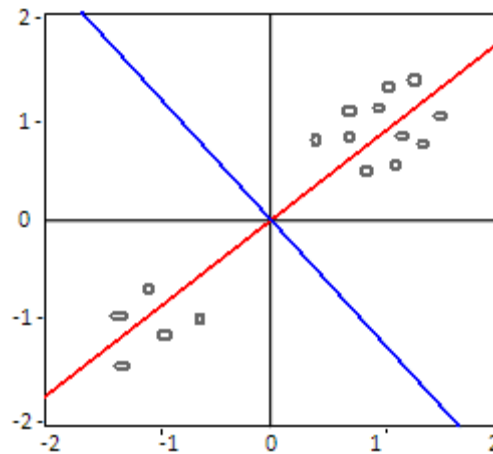


Figura 2.8: Gráfico da distribuição de dados pela PCA.

Os passos para calcular os componentes principais são:

- obter os dados ou as amostras de vetores de dimensão n ;
- calcular a média ou o vetor médio desses dados;
- subtrair a média de todos os itens de dados;
- calcular a matriz de covariância, utilizando todas as subtrações, o que resulta na média do produto de cada subtração por ela mesma cuja dimensão é $n \times n$;
- calcular os autovalores e autovetores da matriz de covariância; e
- fazer o arranjo da matriz da transformada de Hotelling.

Esses passos do algoritmo da análise de componentes principais (PCA), que consiste em um procedimento matemático, transformam uma série de variáveis correlacionadas em um número menor de variáveis não correlacionadas, chamadas componentes principais.

Um problema comum em reconhecimento estatístico de padrões é a seleção das características ou a extração de características. A seleção de características se refere a um processo no qual um espaço de dados é transformado em um espaço de características que, em teoria, tem a mesma dimensão que o espaço original de dados. Entretanto, a transformação é projetada de tal forma que o conjunto de dados pode ser representado por um número reduzido de características “efetivas” e ainda retém a maioria do conteúdo de informação intrínseco dos dados; em outras palavras, o conjunto de dados sofre uma redução de dimensionalidade [39].

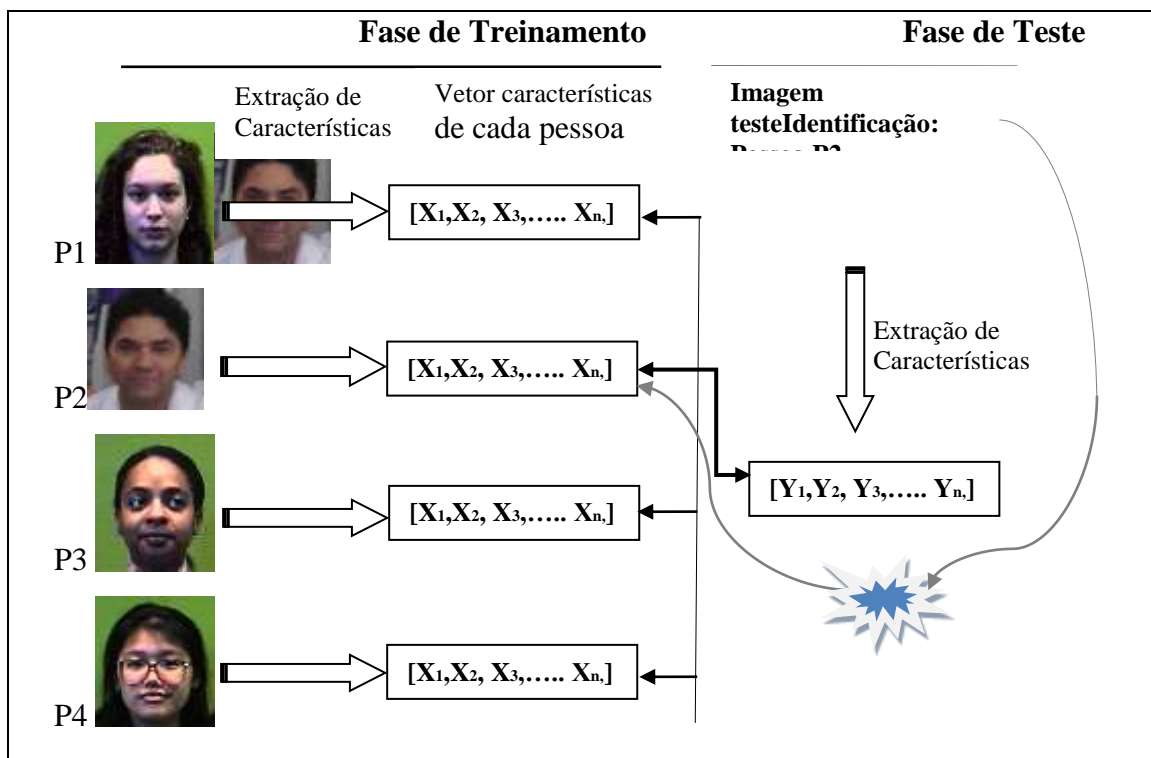
Dentro dessa perspectiva, a PCA é um forte instrumento de análise exploratória de dados e de preparo de modelos preditivos. Se um conjunto de dados multivariados é

visualizado como um conjunto de coordenadas de dados espaciais, a PCA trata esse conjunto como uma imagem de menor dimensão, uma "sombra" do objeto, mais detalhada.

Assim, a sua utilização consiste do cálculo da decomposição de autovalores das matrizes de covariâncias ou decomposição em valores singulares de uma matriz de dados de modo que a primeira linha, o elemento (0,0), seja o autovetor correspondente ao maior autovalor e assim sucessivamente até que a última linha corresponda ao menor autovalor.

O autovetor com o maior autovalor associado, corresponde ao componente principal do conjunto de dados utilizado. Isso significa que esse é o relacionamento mais significativo entre as dimensões dos dados.

A Figura 2.9 mostra um diagrama demonstrativo da PCA com imagens aleatórias.



2.9: Diagrama demonstrativo da PCA com imagens aleatórias.

2.4.1- Quanto ao treinamento da PCA

Na fase de treinamento, para se obter o conjunto de características destinado ao reconhecimento das imagens das faces, a PCA toma o arquivo base formado por j pessoas P com suas p quantidade de poses e prepara um arquivo para o treinamento, chamado de arquivo-treino. Este é formado por vetores gerados a partir do conjunto das p poses de cada pessoa P , onde cada conjunto é chamado de classe. De cada pose se extrai o vetor característica representativo $P_i, i=1,2,3,...,p$. Admitindo que uma pessoa A tenha suas poses no arquivo-treino, toma-se uma delas e chamando-a de P_A cuja dimensão da resolução do pixel é de a linhas por b colunas. Em seguida o algoritmo converte as características de P_A , por concatenação, em vetor pixel simples V_A . O comprimento do vetor V_A também é $a \times b$. Para esse projeto, o algoritmo da PCA utilizou a técnica de redução de dimensão que transforma o vetor V_A em W_A que tem a dimensão d onde $d \ll a \times b$. Desta forma, para cada pose P_i a ser treinada, a PCA gera o vetor característico reduzido W_i correspondente.

2.4.2- Quanto ao reconhecimento pela PCA

Na fase de reconhecimento ou teste, a PCA destaca um número j de pessoas para serem encontradas comparando suas características com aquelas do arquivo-treino. Assim, pega-se para teste uma pose da pessoa P_j com identidade I_j , ou seja, seu endereço dentro do arquivo-treino e calcula-se o vetor característico dessa pose, encontrando o vetor W_j . Para identificar P_j , calcula-se a similaridade entre W_j e as características de todos os vetores W_i 's do arquivo-treino. Neste trabalho, o referido cálculo é realizado em função da distância euclidiana. O vetor W_i mais semelhante a W_j é colocado na saída do reconhecedor de face. Em seguida, comparam-se as identidades I_i e I_j . Se $I_i = I_j$, ocorre a identificação da pessoa j , caso contrário, ocorre erro de classificação, ou seja, pessoa não identificada.

2.4.3- Cálculo da PCA

Suponha um arquivo-treino composto por p poses de pessoas. Cada pose tem seu vetor $P_i, i=1, 2,...,p$ convertido por concatenação em vetor pixel simples V_i na forma de

vetor linha sendo $V_i \in R^k$ em que $k = a \times b$. Aplicando a PCA para o conjunto de imagem a ser treinada, obtém-se a matriz A que contém p linhas, sendo cada linha formada por um vetor V_i . A dimensão de A é $p \times k$. Seguindo, calcula-se a matriz C_A por meio da covariância de A . A covariância sempre é medida entre duas dimensões.

Cada pose de uma face de M amostras é definida por uma matriz de *pixels* com intensidades na escala de cinza, tendo definidos os vetores $V=(V_1, V_2, V_3, \dots, V_M)$ associados aos *pixels* das poses concatenados por coluna, o vetor médio Ψ é calculado pela Equação (2.14) [40].

$$\Psi = \frac{1}{M} \sum_{i=1}^M V_i \quad (2.14)$$

onde:

Ψ – média dos vetores associados aos *pixels* das poses concatenados por coluna;

M – número de poses de faces; e

V_i -vetores associados aos *pixels* das poses concatenados por coluna.

Por meio da Equação (2.14), pode-se facilmente estender o cálculo da matriz de covariância para qualquer número de medidas e para casos de vetores com n dimensões, onde a matriz de dimensão é $n \times n$. Também podem-se calcular os desvios Φ de cada pose em relação à média, conforme Equação (2.15).

$$\Phi_i = V_i - \Psi \quad (2.15)$$

onde:

Φ_i - desvio de cada pose em relação à média;

V_i -vetores associados aos *pixels* das poses concatenados por coluna; e

Ψ - média dos vetores associados aos *pixels* das poses concatenados por coluna.

A matriz de covariância C é calculada para o conjunto de vetores definidos por Φ normalizados em relação à média, e prontos para serem utilizados no cálculo dos autovetores β_i e autovalores λ_i , conforme a Equação (2.16).

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = AA^T \quad (2.16)$$

em que a matriz $A = \{ \Phi_1, \Phi_2, \Phi_3, \dots, \Phi_i \}$

onde:

C - matriz de covariância;

M - número de poses de faces; e

Φ_i - desvio de cada pose em relação à média

A matriz da covariância é real e simétrica, sendo sempre possível encontrar um conjunto de i autovalores λ_i e os correspondentes autovetores ortonormais [41].

X_i são chamados os autovetores de AA^T , considerando-os semelhantes, conforme Equação (2.17).

$$A^T \cdot A \cdot X_i = \lambda_i \cdot X_i \quad (2.17)$$

Multiplicando-se ambos os lados da Equação (2,17) por A , tem-se a Equação (2.18).

$$A \cdot A^T \cdot A \cdot X_i = \lambda_i \cdot A \cdot X_i \quad (2.18)$$

Conclui-se, portanto, que $A \cdot X_i$ são os autovetores de $C = A \cdot A^T$.

Assim, constrói-se a matriz $D = A^T \cdot A$, resultando em uma matriz de dimensão $M \times M$ em $D_{m,n} = \Phi_i \Phi_i^T$. Portanto, são encontrados os M autovetores X_i de D . Estes autovetores determinam uma combinação linear das M poses do conjunto de treinamento para formar os autovetores β_i que são calculados conforme a Equação (2.19).

$$\beta_i = \sum_{k=1}^M X_{ik} \Phi_k = A \cdot X \quad (2.19)$$

onde:

β_i - autovetores determinados a partir da combinação linear dos autovetores

$X_i; X_{ik}$ – autovetores resultados de uma matriz de dimensão $M \times M$; e

Φ_k - conjunto de vetores normalizados em relação à média.

Assim, os cálculos da ordem dos números de *pixels* nas poses N^2 são reduzidos para o número de poses no conjunto de treinamento M .

Desta forma, obtém-se k autovalores e k autovetores associados, que podem ser colocados em ordem de acordo com sua utilidade na caracterização da variação entre as poses, uma vez que os primeiros autovetores possuem valores altos, caracterizando a sua representatividade com relação aos dados [40]. O referido ordenamento dispõe todos os autovetores com seus autovalores em ordem decrescente de autovalor e emparelhados de tal maneira que o i autovalor corresponde ao i autovetor. Na sequência, faz-se a seleção dos maiores d autovalores e autovetores emparelhados. Os autovetores selecionados são dispostos em colunas, criando-se a matriz transformação S . O objetivo é calcular os vetores características W_i onde $W_i \in R^d$ com $d \ll k$. Por meio da matriz S encontra-se W_i 's de V_i . Este cálculo é realizado utilizando-se a Equação (2.20).

$$W_i = S^T \cdot V_i^T \quad (2.20)$$

onde:

W_i - vetores características;

S – matriz criada a partir dos autovetores selecionados; e

V_i - vetores associados aos *pixels* das poses concatenados por coluna.

Na Equação (2.20) S^T e V_i^T são matrizes transpostas de S e V_i respectivamente. Cada coluna de S é um autovetor de comprimento k . Esta matriz é de comprimento igual a $a \times b$, ou seja, é a resolução da imagem de entrada. Assim, pode-se converter cada autovetor em uma imagem revertendo-se a operação de concatenação. Esses autovetores imagens convertidos, são chamados de autofaces desde que evidentemente sejam resultados de processamento de faces humanas.

Logo que são obtidos W_i 's, pode-se reconstruir a pose da pessoa i . Ao utilizar todos os autovetores k em vez de d para formar S , a imagem reconstruída é P_i . Entretanto, neste caso pretende-se utilizar a resolução simplificada com $d \ll k$. Quanto

mais autovetores forem usados na formação de S , mais próximo da imagem original é a imagem reconstruída.

2.5- Considerações Finais Deste Capítulo

Neste capítulo, foram apresentados os aspectos teóricos dos classificadores ou processadores que executam por meio de seus algoritmos a tarefa do reconhecimento de face. Foram apresentados os conceitos básicos de um algoritmo genético (AG), bem como uma visão geral da sua execução.

O algoritmo da RNA, utiliza um método local para atualizar os pesos sinápticos e os níveis de bias, ou seja, a modificação aplicada ao peso sináptico de um neurônio é confinada à vizinhança imediata daquele neurônio. Utiliza-se também um método eficiente para calcular todas as derivadas parciais da função de custo (medida de desempenho) em relação a esses parâmetros livres (pesos sinápticos e níveis de bias).

O objetivo de utilizar a PCA é realizar a compressão de dados preservando o máximo possível a informação das entradas de dados. Além disso, uma representação baseada nos primeiros componentes principais tem preferência sobre uma representação arbitrária em subespaço, pois são ordenadas em autovalores decrescentes ou em variância decrescente.

Futuramente, a biotecnologia pode mudar muitos aspectos da vida humana, inclusive alterando as características germinativas. Por isso, para o reconhecimento de face aqui tratado, os algoritmos deverão ter habilidades das emoções, ou seja, terão que aprender as artes básicas da vida, reconhecendo os impulsos da raiva, da ansiedade, da motivação, da ternura, da compaixão, da empatia e da colaboração.

O próximo capítulo apresenta os materiais, a metodologia dos testes realizados, a sequência de utilização dos classificadores e finalmente como é a apresentação dos resultados neste trabalho.

CAPÍTULO III

MATERIAIS E METODOLOGIA UTILIZADA PARA O RECONHECIMENTO DE FACES

3.1- Introdução

Nos itens 1.5 a 1.7 do capítulo 1 deste trabalho foram mostradas a importância da face como instrumento biométrico, aumentando a exposição para a sua ampla aplicabilidade na vida rotineira das pessoas e das instituições.

A partir deste capítulo é apresentada a técnica do reconhecimento de faces, por ser a face o objeto principal desta pesquisa.

Por que a face? O cérebro leva três segundos para gravar uma imagem na memória. A partir desta gravação é que são processadas as respostas do corpo. Ou seja, o ser humano enxerga a vida por meio da memória. É pela face que entram essas informações que denotam as reações. Uma imagem da face gravada a três segundos na memória tem sido a maneira mais eficiente do ser humano reconhecer as pessoas, desde o seu nascimento e desde sempre.

As variações faciais sejam por concepções lúdicas, por emoções diferenciadas ou por questões temporais, entre outras tornam o seu estudo bastante desafiador. Além, disto, existe o aspecto motivacional do autor deste trabalho que, na verdade, não se explica, tendo sido quase sempre inacessível e inconsciente, mas movido pelo sentimento da beleza que o estudo despertou.

Pode-se acrescentar também o desejo de contribuir com os estudos dos pesquisadores que buscam pelo bom desempenho dos métodos de reconhecimento de face. Tais estudos tentam implementar um sistema que seja rápido e bastante exato,

porque na maioria, as aplicações requerem reconhecimentos imediatos para as tomadas de decisões.

Por que a transformada discreta do cosseno –(DCT)? A DCT foi utilizada como base para o processamento dos algoritmos deste trabalho por apresentar a propriedade de compactação da energia de regiões próximas da origem após a execução da transformada. Com a execução desta transformada muitos coeficientes apresentam valores próximos de zero e, portanto, podem ser eliminados sem que exista perdas significativas de informações contidas nas imagens. Mais detalhes sobre a DCT estão na seção A.2 do apêndice A desta tese.

Neste trabalho, são realizadas as medições das experiências comparando o desempenho do reconhecimento de face dos algoritmos Genético (AG), Rede Neural Artificial (RNA) e Análise de Componentes Principais (PCA).

Este capítulo apresenta os materiais, a metodologia dos testes realizados, a sequência de utilização dos classificadores e finalmente como é a apresentação dos resultados neste trabalho.

A Figura 3.1 [42] metaforiza em certa medida o propósito deste trabalho ao provocar o leitor a identificar rostos humanos entre os galhos.

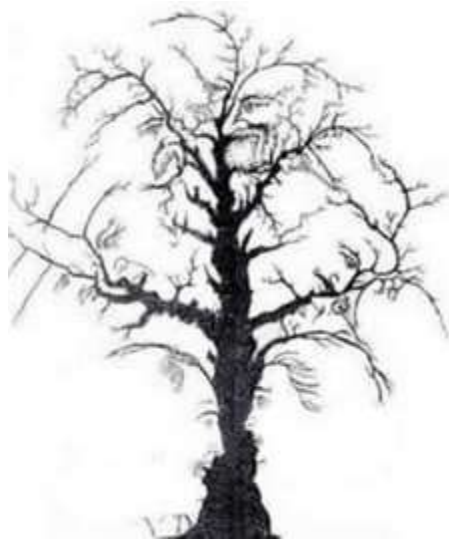


Figura 3.1: Figura da árvore com vários rostos.

3.2- Materiais

São mostradas a especificação do hardware, do software e dos bancos de imagens que são utilizados nos testes com todos os classificadores.

3.2.1- Quanto ao hardware e o software

Neste trabalho foi utilizado um notebook Dell CPU com processador Intel i7 de 2,20 GHz, memória 6 GB; sistema operacional Microsoft Windows XP de 64 bits e linguagem de programa Matlab 7.0.1 versão 2009b [43].

3.2.2- Especificação da base de dados ORL

Foi usado o banco de imagens da Olivetti Research Laboratory da Universidade Britânica de Cambridge (ORL) que contém fotos monocromáticas de 40 pessoas, cada uma com 10 diferentes poses, com um total de 400 fotos de dimensões 112 x 92 pixels no formato Portable Graymap (PGM). As imagens desse arquivo foram feitas sob condições bastante diferentes seja em iluminação, poses e expressões faciais [44].

A base de dados ORL tem as seguintes características:

- fotografias em preto e branco;
- número total de indivíduos: 40;
- número de poses por indivíduo: 10;
- número total de poses: 400;
- formato das poses: tamanho 112 x 92 pixels em pgm;
- gênero: contém poses de homens e mulheres;
- raça: contém poses de pessoas de varias raças e origens; e
- complemento: todas as pessoas são adultas, algumas usando óculos e outras usando barba.

Na Figura 3.2 está uma visão sobre a montagem e a disposição teórica das imagens das faces humanas do arquivo ORL que contém imagens de 40 pessoas (C1 a C40) onde cada pessoa tem dez diferentes poses (1 a 10).

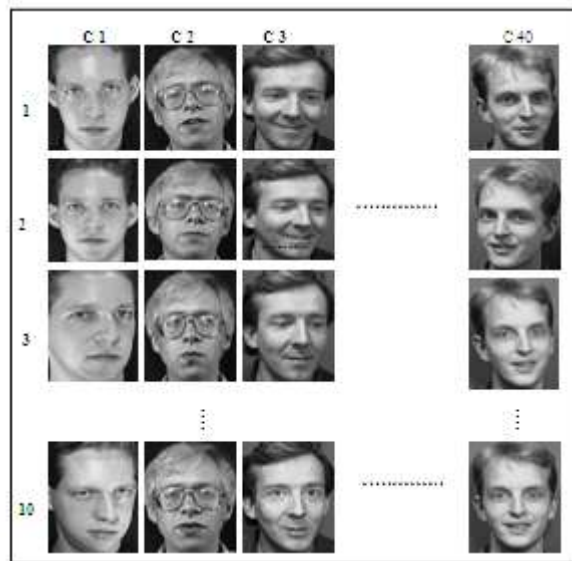


Figura 3.2: Conjunto de pessoas e suas respectivas poses do arquivo ORL.

A Figura 3.3 mostra a representação do arquivo ORL onde estão as 40 classes, onde cada uma dessas classes contém dez poses da mesma pessoa.

	C1	C2	C3		C40
1	P 1,1	P 1,2	P 1,3		P 1,40
2	P 2,1	P 2,2	P 2,3	P 2,40
3	P 3,1	P 3,2	P 3,3		P 3,40
		⋮			⋮
10	P 10,1	P 10,2	P 10,3	P 10,40

Figura 3.3: Representação do arquivo ORL com dez poses por classe.

3.2.3- Especificação da base de dados UK

Outro banco de imagens utilizado foi parte do Face Recognition Data da Universidade de Essex, Reino Unido (UK). Este banco foi preparado com imagens de 40 pessoas cada uma com 20 poses diferentes, totalizando 800 fotos de dimensões 180 x 200 pixels no formato jpeg. As fotos foram tiradas sob condições bastante diferentes seja em iluminação, poses e expressões faciais [41].

A base de dados UK tem as seguintes características:

- fotografias coloridas;
- número total de indivíduos: 120;
- número de poses por indivíduo: 20;
- número total de poses: 2400;
- formato das poses: tamanho 200 x 180 pixels em jpeg;
- gênero: contém poses de homens e mulheres;
- raça: contém poses de pessoas de várias raças e origens;
- luminosidade: iluminação, brilho, contraste e cor diferenciada; e
- complemento: todas as pessoas são adultas, algumas usando óculos e outras usando barba.

Na Figura 3.4 está uma visão geral da montagem e da disposição teórica das imagens de faces humanas do arquivo UK que contém poses de 120 pessoas (F1 a F120) onde cada pessoa tem 20 diferentes poses (1 a 20).

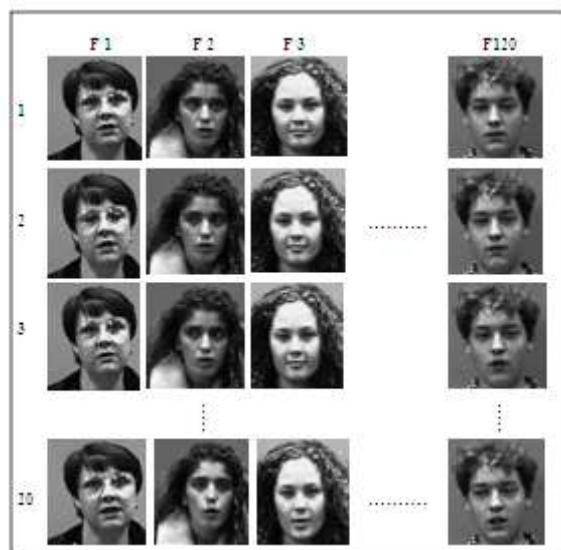


Figura 3.4: Conjunto de imagens de pessoas e suas respectivas poses extraídas do arquivo UK.

A Figura 3.5 mostra a representação do arquivo UK onde estão as 120 classes e cada uma destas classes contém 20 poses da mesma pessoa.

	F1	F2	F3		F120
1	G 1,1	G 1,2	G 1,3		G 1, 120
2	G 2,1	G 2,2	G 2,3	G 2,120
3	G 3,1	G 3,2	G 3,3		G 3,120
	⋮				⋮
20	G 20,1	G 20,2	G 20,3	G 20, 120

Figura 3.5: Representação do arquivo UK com 20 poses por classe.

3.3- Metodologia Dos Testes Realizados

Cada classificador tem a sua particularidade operacional o que distingue um do outro. Além disso, dependendo do problema, existem alterações nas respectivas implementações.

Em geral, os procedimentos consistem em fornecer várias poses de uma pessoa (classe) ao reconhecedor e este verifica se elas existem em um arquivo previamente endereçado. Isto é realizado pelo algoritmo de varredura que busca no arquivo endereçado as imagens convertidas em matrizes vetoriais, com características muito próximas daquelas matrizes das imagens fornecidas (imagens query) para o reconhecimento.

A Figura 3.6 mostra uma visão esquematizada do referido procedimento.

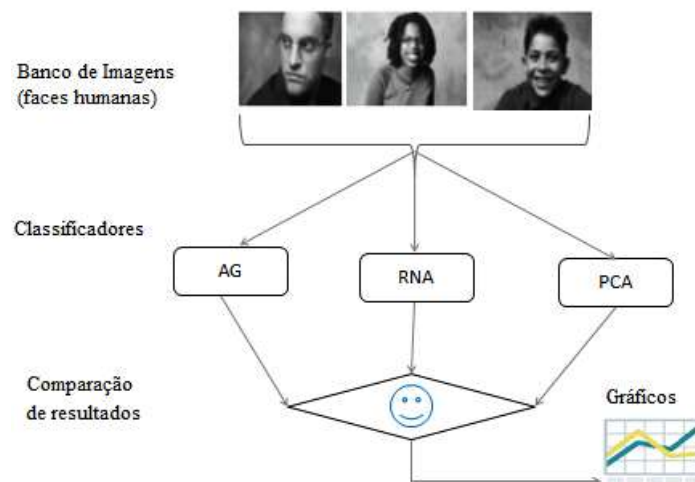


Figura 3.6: Representação do procedimento de reconhecimento de faces.

3.3.1- Metodologia do Algoritmo Genético para os arquivos ORL e UK

A metodologia proposta cria utilizando o algoritmo genético (AG) um modelo para cada classe, a partir dos coeficientes obtidos pela aplicação da DCT nas as matrizes de pixels das imagens dos arquivos ORL e UK. Cada pose de cada classe é representada por um vetor linha com os coeficientes DCT, chamado de vetor imagem P .

Ao utilizar o AG é necessário criar uma população inicial arbitrária como possível resposta, onde cada indivíduo I é tratado como cromossomo com n genes dentro do processo de evolução. Neste trabalho cada vetor I é codificado por uma sequência de dígitos binários. Em seguida cria-se para cada vetor P um vetor máscara M , resultado da combinação de cada indivíduo I com cada vetor imagem P . Nesta combinação verifica-se o valor de cada gene do cromossomo. Se o valor binário é 0, no vetor máscara M permanece 0, se é 1 atribui-se a essa posição do vetor máscara o valor do coeficiente DCT contido na mesma posição do vetor imagem P . A nota de adaptação, ou fitness, é a média da distância euclidiana entre cada vetor M com todos os vetores P . A Figura 3.7 mostra essa composição para n coeficientes (genes).

P	c1	c2	c3	c4	c5	c6	c n
I	1	1	0	0	1	0	1
M	c1	c2	0	0	c5	0	c n

Figura 3.7: Representação de um vetor imagem P , um vetor indivíduo I e um vetor máscara M para n coeficientes.

Nos bancos de faces dos arquivos ORL e UK cada indivíduo é representado por dez poses e 20 poses, respectivamente.

Inicialmente o AG gera uma população inicial aleatória de 50 indivíduos, que é avaliada e colocada em ordem crescente da nota de adaptação. O processo de avaliação é realizado calculando a média entre as distâncias euclidianas dos vetores DCT, considerando apenas as variáveis selecionadas pelo indivíduo. Quanto menor é esse valor, mais adaptado é o indivíduo.

Dessa forma, quanto maior é o número de variáveis "zeradas" pelo indivíduo, menor é a distância euclidiana e, conseqüentemente, melhor é a avaliação do indivíduo.

Neste ponto surge um problema: neste trabalho, o AG foi programado para buscar soluções com o menor erro possível e como o erro vai diminuindo proporcionalmente ao número de variáveis "zeradas", o AG pode zerar todas as variáveis dos seus indivíduos no decorrer das gerações.

Para contornar este problema, foi considerado o parâmetro *número mínimo de variáveis*, que penaliza os indivíduos que tiver em um número de variáveis "não zeradas" inferior a esse limite. Desta forma, as soluções encontradas pelo AG selecionam um número de variáveis "não zeradas" muito próximo ao limite estipulado.

Considerando, portanto, a necessidade do *número mínimo de variáveis*, a partir desse ponto inicia-se o laço de repetição dos ciclos evolutivos (gerações), que compreendem os seguintes passos:

Passo 1: A seleção natural que alguns indivíduos *I* da população criada são escolhidos para o cruzamento, onde o processo de escolha é realizado pela roleta, possibilita mais chances aos indivíduos mais adaptados. O número de indivíduos selecionados é estipulado de acordo com a probabilidade indicada pela taxa de cruzamento;

Passo 2: Os indivíduos selecionados são cruzados (em pares) gerando dois filhos de cada cruzamento;

Passo 3: Os novos indivíduos (filhos gerados pelo processo de cruzamento) são acrescentados à população;

Passo 4: Aleatoriamente são escolhidos alguns indivíduos da população para sofrerem mutação, de acordo com a probabilidade definida pela taxa de mutação externa;

Passo 5: Os indivíduos escolhidos para a mutação são copiados e passam pelo processo que altera aleatoriamente alguns de seus genes, de acordo com as probabilidades definidas na taxa de mutação interna, gerando uma sub-população mutante que também é inserida na população;

Passo 6: O conjunto da população é avaliado em um processo em que se atribui um valor de fitness para cada indivíduo;

Passo 7: A população é ordenada, do indivíduo com menor valor de fitness (mais adaptado) para o indivíduo com maior valor de fitness (menos adaptado);

Passo 8: O ciclo termina com o corte dos indivíduos menos adaptados. Com a população disposta de forma decrescente de adaptabilidade, descarta-se os indivíduos que ocuparem as posições excedentes ao tamanho da população inicial; e

Passo 9: Neste ponto o algoritmo verifica a condição de parada. Se não existir mais gerações a serem rodadas o AG salva a evolução no pacote. Caso existam mais gerações a serem executadas, o processo reinicia a partir do passo 1.

Ao terminar cada variação quanto ao uso do AG para os arquivos ORL e UK, seja com pose inteira seja com subposes, são mostrados os gráficos das taxas de acertos, dos tempos de processamentos e dos erros. As curvas de erros são valores flutuantes, ou seja, se repetir o processamento os valores dos erros em relação ao número de gerações não serão necessariamente os mesmos. Por isso padronizou-se coletar os resultados emitidos com o critério de parada na centésima geração.

3.3.1.1- Utilização do arquivo ORL pelo AG

As imagens do arquivo ORL foram utilizadas primeiro considerando a imagem inteira de cada pose e depois considerando as subposes resultantes da divisão de cada pose, como explicado a seguir:

3.3.1.1.1- Utilização de toda imagem (pose) do arquivo ORL pelo AG

Aplicou-se a DCT e foram considerados os 64 coeficientes mais significativos do canto superior esquerdo da matriz 8x8 de todas as poses de cada classe. A Figura 3.8 mostra a localização dos coeficientes da DCT utilizados para cada pose ORL e a Figura 3.9 ilustra a representação dos vetores, onde x_1 o coeficiente 1, x_2 o coeficiente 2 e assim sucessivamente até o coeficiente x_{64} . P_x é a pose de cada classe, I_x é o indivíduo criado pelo AG e M_x é o vetor máscara criado pelo AG. O índice x indica que está sendo utilizado o arquivo ORL.

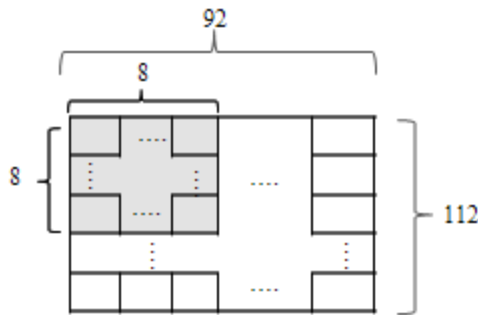


Figura 3.8: Localização dos coeficientes DCT utilizados para cada pose ORL.

Px	x1	x2	x3	x4	x5	x6	x64
Ix	1	1	0	0	1	0	1
Mx	x1	x2	0	0	x5	0	x64

Figura 3.9: Representação de um vetor imagem Px, um vetor indivíduo Ix e um vetor máscara Mx para 64 coeficientes.

3.3.1.1.2- Utilização das subposes do arquivo ORL pelo AG

Cada pose do arquivo ORL de 92 por 112 pixels foi fracionada em tamanho 8 por 8, resultando em 154 subposes X, como mostrado na Figura 3.10. Nas divisões, restaram 4 pixels ao dividir 92 por 8, referente ao número de colunas. Por isso, o AG foi programado para excluir 2 colunas de pixels de cada lado das imagens. Esses pixels estão nas bordas das poses e a ausência deles nas matrizes com valores do cinza de cada pixel não é significativa nas extrações de características das faces.

	8				
8	X1	X2	X3		X11
	X12	X13	X14	X22
	X23	X24	X25		X33
		⋮			⋮
				
	X144	X145	X146		X154

Figura 3.10: Representação de uma pose ORL de 92 colunas por 112 linhas dividida em 154 subposes X

Em cada subpose X foram retiradas as seguintes quantidades de coeficientes DCT mais significativos:

- Um coeficiente. Neste caso, cada pose de cada classe foi representada por 154 coeficientes.
- Dois coeficientes. Neste caso, cada pose de cada classe foi representada por 308 coeficientes.

- c) Quatro coeficientes. Neste caso, cada pose de cada classe foi representada por 616 coeficientes.

3.3.1.2- Utilização do arquivo UK pelo AG

As imagens do arquivo UK foram utilizadas primeiro considerando a imagem inteira de cada pose depois as subposes resultantes da divisão de cada pose, como explicado a seguir.

3.3.1.2.1- Utilização de toda imagem (pose) do arquivo UK pelo AG

Aplicou-se a DCT e foram considerados os 64 coeficientes mais significativos do canto superior esquerdo da matriz 8 x 8 de todas as poses de cada classe. A Figura 3.11 mostra a localização dos coeficientes DCT utilizados para cada pose UK e a Figura 3.12 ilustra a representação dos vetores, sendo y_1 o coeficiente 1, y_2 o coeficiente 2 e assim sucessivamente até o coeficiente y_{64} . P_y é a pose de cada classe, I_y é o indivíduo criado pelo AG e M_y é o vetor máscara criado pelo AG. O índice y indica que está sendo utilizado o arquivo UK.

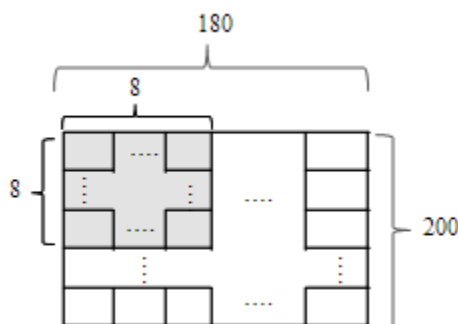


Figura 3.11: Localização dos coeficientes DCT utilizados para cada pose UK

P_y	y_1	y_2	y_3	y_4	y_5	y_6	y_{64}
I_y	1	1	0	0	1	0	1
M_y	y_1	y_2	0	0	y_5	0	y_{64}

Figura 3.12: Representação de um vetor imagem P_y , um vetor indivíduo I_y e um vetor máscara M_y para 64 coeficientes.

3.3.1.2.2- Utilização das subposes do arquivo UK pelo AG

Cada pose do arquivo UK de 180 por 200 pixels foi fracionada em tamanho 8 por 8, resultando em 550 subposes Y , como mostrado na Figura 3.13. Nas divisões, restaram 4 pixels ao dividir 180 por 8 referente ao número de colunas. Por isso, o AG foi programado para excluir 2 pixels de cada lado das imagens. Esses 4 pixels estão nas bordas das poses e a ausência deles nas matrizes com valores do cinza de cada pixel não é significativo nas extrações de características das faces.

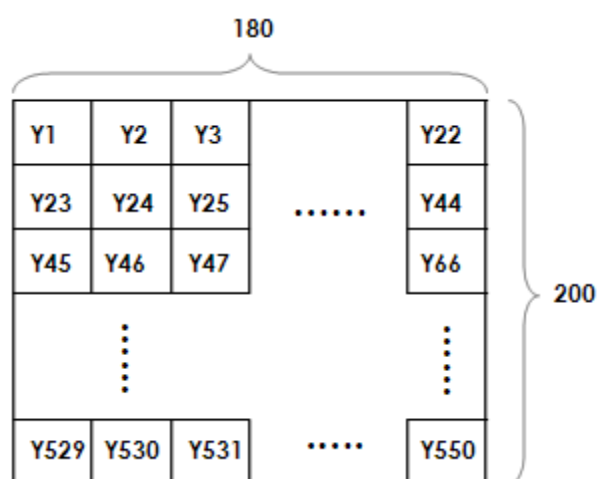


Figura 3.13: Representação de uma pose UK de 180 colunas por 200 linhas dividida em 550 subposes Y .

Em cada subpose Y referida foram retiradas as seguintes quantidades de coeficientes mais significativos:

- 1- Um coeficiente DCT. Neste caso cada pose de cada classe foi representada por 550 coeficientes.
- 2- Dois coeficientes DCT. Neste caso cada pose de cada classe foi representada por 1100 coeficientes.
- 3- Quatro coeficientes DCT. Neste caso cada pose de cada classe foi representada por 2200 coeficientes.

3.3.2- Metodologia em relação à Rede Neural Artificial

A primeira operação é o aprendizado da Rede Neural Artificial. O aprendizado ou o treinamento da RNA foi realizado pela aprendizagem por retropropagação onde os pesos são atualizados a cada *loop* de processamento comparando a saída da rede com

aquela saída desejada. O ajuste é realizado a cada *loop* ou época até apresentar o conjunto de valores de pesos que satisfaça a saída pretendida.

A rotina de cálculo da RNA utilizou matrizes de coeficientes gerados pela DCT em três etapas para cada um dos arquivos: na primeira etapa foram utilizados 64 coeficientes, na segunda etapa foram utilizados 100 coeficientes e na terceira etapa foram utilizados 256 coeficientes, para representar cada pose no processo.

3.3.2.1- A rede neural artificial para o arquivo ORL

Foram colocadas em arquivo as 400 imagens distribuídas em 40 pastas, cada pasta chamada neste trabalho de classe contendo dez poses da mesma pessoa.

Para o treinamento ou o aprendizado da rede adotou-se o critério de escolher as cinco primeiras poses de cada classe. Enquanto que o reconhecimento foi realizado em cinco experiências distintas utilizando as cinco últimas poses de cada classe. Os números de poses utilizados para o reconhecimento foram 6, 7, 8, 9 e 10. A tarefa foi verificar que classe pertencia as imagens destacadas para o reconhecimento, com a finalidade de observar a variação da taxa de exatidão, do tempo de treinamento e do tempo de reconhecimento com o aumento do número de imagens processadas utilizando a RNA.

3.3.2.2- A rede neural artificial para o arquivo UK

Considerando o arquivo UK e utilizando a RNA, foi colocado em arquivo o conjunto de imagens das 120 pessoas, contendo 20 poses cada.

Para esse aprendizado da rede adotou-se o critério de escolher as dez primeiras poses de cada pessoa F (classe), na Figura 3.5, e para o reconhecimento as dez últimas poses dessa figura.

3.3.3- Metodologia relacionada à Análise de Componentes Principais

A PCA para iniciar os cálculos faz a codificação das bases de dados, ou seja, o algoritmo toma todos os vetores imagens do arquivo treinamento que são vetores de tamanho M linhas por N colunas e retorna uma matriz com todos esses vetores em função do número de pessoas ou classes. Assim, essa nova matriz de tamanho $MN \times C_{c,i}$

onde $c=1,2,...,40$ e $i=1,2,...,10$, para o arquivo ORL ou $MN \times F_{fij}$, onde $f=1,2,...,120$ e $j=1,2,...,20$, para o arquivo UK, passa a ser a matriz representativa de cada classe.

3.3.3.1- A análise de componentes principais (PCA) para o arquivo ORL

Para a experiência com a PCA colocou-se em uma só pasta todas as 400 imagens do arquivo ORL numeradas sequencialmente de 1 a 400 de tal forma que as poses da primeira pessoa foram numeradas de 1 a 10, as da segunda pessoa de 11 a 20, as da terceira classe de 21 a 30 e assim sucessivamente até a quadragésima pessoa com imagens numeradas de 391 a 400. Fez-se uma cópia dessa pasta e foram denominados dois arquivos como: arquivo-treinamento e arquivo-reconhecimento. São nesses dois arquivos que o algoritmo da PCA atuou.

3.3.3.2- A análise de componentes principais (PCA) para o arquivo UK

Ao considerar o arquivo UK com 120 classes, o programa foi executado durante 10 dias sem interrupção e foi concluído apenas o reconhecimento de 40 classes. Mesmo utilizando uma máquina com as características mostradas na seção 3.2.1, o tempo de reconhecimento para os grandes arquivos pela PCA foi proibitivo. Pois, é uma característica desse algoritmo que não aceita fazer o treinamento com menor número de classes em relação ao total dessas classes existentes no arquivo.

Por essa razão o teste com a PCA realizado com apenas 40 classes (F1 a F40) do arquivo UK, pois à medida que aumentava o número de classes o tempo de reconhecimento aumentava vertiginosamente, chegando a semanas de processamento. Assim, foram colocadas em uma só pasta todas as 800 poses numeradas sequencialmente de 1 a 800 de tal forma que as imagens da primeira pessoa foram numeradas de 1 a 20, as da segunda pessoa de 21 a 40, as da terceira pessoa de 41 a 60, e assim sucessivamente até a centésima vigésima pessoa com imagens numeradas de 781 a 800. Fez-se uma cópia dessa pasta e denominou-se os dois arquivos como: arquivo-treinamento e arquivo-reconhecimento que o algoritmo PCA atuou.

3.4- Sequência de utilização dos classificadores

No teste os procedimentos foram divididos em três grandes blocos: no primeiro bloco, conforme Capítulo IV deste trabalho, procedeu-se ao reconhecimento das faces utilizando o algoritmo genético, no segundo bloco, o reconhecimento de faces foi realizado pela rede neural artificial, conforme Capítulo V desta tese, finalmente, no terceiro bloco, Capítulo VI deste trabalho, o reconhecimento foi realizado utilizando-se a análise de componentes Principais. Em cada um dos três blocos foram realizados os testes utilizando primeiro o arquivo ORL e depois o arquivo UK.

Dessa forma a utilização distinta de ambos os arquivos, ou seja, o uso dos dois arquivos em cada um dos algoritmos classificadores, garantiu as mesmas condições para verificar o desempenho de cada classificador ao processar cada base de dados.

3.5- Apresentação dos Resultados

Os resultados obtidos são apresentados em forma de tabelas e gráficos nos respectivos capítulos que descrevem o teste com cada algoritmo classificador. No Capítulo VII é apresentado o agrupamento dos “*dados – respostas*” onde realizadas a análise e a comparação entre os desempenhos dos métodos.

3.6- Considerações Finais Deste Capítulo

Este capítulo apresentou o hardware e o software utilizados no teste. Embora o hardware tenha sido um dos melhores em termos de computação pessoal, obviamente ele tem limitação quanto à memória e a velocidade de processamento em relação a outras máquinas do gênero. Assim, ao utilizar equipamentos mais sofisticados ocorrerá variações dos resultados, principalmente na variável tempo. Além disso, foram mostrados os arquivos das imagens da base de dados, a forma como os classificadores preparam os dados para a intervenção dos algoritmos e ainda como foram apresentados os resultados nos respectivos capítulos que descrevem a operação de cada classificador.

Dessa forma, é importante ressaltar que esses algoritmos já estão auxiliando as instituições a promover o desenvolvimento e a segurança da sociedade. A alta tecnologia cada vez mais acessível pode estimular a auto-transformação fenotípica ou antecipar em milhões de anos a teoria do equilíbrio pontuado pela qual “as espécies

vivem muito tempo sem alteração, e de repente se transformam em outros seres”. Essas mudanças repentinas perturbam a identificação corriqueira entre as pessoas.

O próximo capítulo apresenta a utilização do AG usando o arquivo ORL e a utilização do AG usando o arquivo UK.

CAPÍTULO IV

UTILIZAÇÃO DO ALGORÍTMO GENÉTICO NO RECONHECIMENTO DE FACES

4.1- Introdução

Neste trabalho o Algoritmo Genético (AG) executou o processamento computacional utilizando os arquivos ORL e UK, especificados nas seções 3.2.2 e 3.2.3 do Capítulo III desta tese. Na primeira etapa foram utilizados 64 coeficientes DCT, para representar cada pose a partir da imagem toda para ambos os arquivos. Na segunda etapa, esses coeficientes foram utilizados em diferentes quantidades para representar as poses a partir do fracionamento das imagens. Para o arquivo ORL, foram usados 154, 308 e 616 coeficientes e, para o arquivo UK, foram usados 550, 1100 e 2200 coeficientes respectivamente, conforme descrito nas seções 3.3.1.1 e 3.3.1.2 do Capítulo III desta tese.

A finalidade do algoritmo AG é identificar a classe que pertence a pose ou o grupo dessas poses, quando são fornecidas uma ou mais poses (imagens query), de qualquer uma das classes pertencentes aos arquivos ORL ou UK. O AG realiza essa tarefa inspirado na teoria de Darwin – do mecanismo das novas formas de vidas geradas, ou seja, seleção por meio das gerações.

A implementação do sistema conforme a estrutura do AG foi realizada adotando-se os seguintes parâmetros:

- População inicial: criaram-se 50 vetores indivíduos arbitrados como possíveis respostas com 64 elementos (binários);
- Máscara: para cada vetor imagem criou-se um vetor máscara;
- Adaptação: calculou-se a nota de adaptação ou fitness;

- Novo indivíduo: seleção natural probabilística, por meio da roleta;
- Taxa de cruzamento foi de 10 %;
- Taxa de mutação externa foi de 10 %;
- Taxa de mutação interna foi de 2 %;
- Número de geração foi 100 (critério de parada); e
- Número mínimo de variáveis no processo de adaptação: para o arquivo ORL utilizaram-se 30 e 50 variáveis onde foram empregados vetores característicos com 64, 154 coeficientes DCT respectivamente e 100 variáveis para os vetores com 308 e 616 coeficientes DCT. Para o arquivo UK, utilizaram-se 30 variáveis onde foram empregados vetores característicos com 64 coeficientes DCT, 50 variáveis para os vetores com 550 coeficientes DCT e 100 para vetores com 1100 coeficientes DCT e 200 variáveis para os vetores com 2200 coeficientes DCT.

Este capítulo apresenta a utilização do AG usando o arquivo ORL e a utilização do AG usando o arquivo UK.

4.2- A utilização do Algoritmo Genético utilizando o arquivo ORL

Para o arquivo ORL foram utilizados 64, 154, 308 e 616 coeficientes DCT para compor as poses de cada classe.

Variações da experiência para o treinamento e o reconhecimento pelo AG, utilizando o arquivo ORL

Os treinamentos e os reconhecimentos de faces foram realizados para o grupo de 2, 5, 10, 15, 20, 25, 30 e 40 classes tomadas de maneira sequencial para as seguintes variações, considerando a pose por classe:

1. Treinou-se com nove poses (1^a à 9^a pose) e reconheceu-se com uma pose (10^a pose);
2. Treinou-se com oito poses (1^a à 8^a pose) e reconheceu-se com duas poses (9^a e 10^a poses);
3. Treinou-se com sete poses (1^a à 7^a pose) e reconheceu-se com três poses (8^a, 9^a e 10^a poses);
4. Treinou-se com seis poses (1^a à 6^a pose) e reconheceu-se com quatro poses (7^a, 8^a, 9^a e 10^a poses); e

5. Treinou-se com cinco poses (1ª à 5ª pose) e reconheceu-se com cinco poses (6ª, 7ª, 8ª, 9ª e 10ª poses).

4.2.1- Quanto a utilização de toda a imagem (pose) ORL

A seção 3.3.1.1.1 mostra como se processa a retirada dos coeficientes DCT considerando toda a pose.

Foi realizada a operação para todas as poses de uma classe, resultando em conjunto de 10 vetores de 64 elementos. Este conjunto, mostrado na Figura 4.1, simboliza a matriz utilizada pelo AG para definir as características de uma classe, necessárias ao reconhecimento de uma imagem, onde $X_{1,1,1}$, $X_{1,1,2}$, $X_{1,1,3}$, $X_{1,1,4}, \dots, X_{1,1,64}$ significam os 64 coeficientes da primeira pose P da classe C_1 . Com este raciocínio, chega-se aos 10 vetores da primeira classe. Este procedimento foi repetido para todas as demais classes em um total de 40.

$P_{1,1}$	$X_{1,1,1}$	$X_{1,1,2}$	$X_{1,1,3}$	$X_{1,1,4}$	$X_{1,1,64}$
$P_{2,1}$	$X_{2,1,1}$	$X_{2,1,2}$	$X_{2,1,3}$	$X_{2,1,4}$	$X_{2,1,64}$
$P_{3,1}$	$X_{3,1,1}$	$X_{3,1,2}$	$X_{3,1,3}$	$X_{3,1,4}$	$X_{3,1,64}$
	\vdots	\vdots	\vdots	\vdots	...	\vdots
$P_{10,1}$	$X_{10,1,1}$	$X_{10,1,2}$	$X_{10,1,3}$	$X_{10,1,4}$	$X_{10,1,64}$

Figura 4.1: Representação de uma classe por coeficientes DCT com 10 poses. Cada pose é um vetor com 64 elementos, arquivo ORL.

Assim, ao utilizar toda a pose, a base de dados para o treinamento das características foi composta de 40 matrizes de coeficientes DCT de tamanho de 10 linhas por 64 colunas.

A Tabela 4.1 mostra o resultado do reconhecimento de face utilizando 64 coeficientes DCT para representar cada pose. Os valores contidos nessa tabela são aqueles obtidos nos testes para 2, 5, 10, 15, 20, 25, 30 e 40 classes. Até a vigésima classe obtêm-se uma taxa de acerto muito alta e a partir dessa classe a taxa de acerto diminui. Os testes mostram que o tempo de treinamento e o tempo de reconhecimento aumentam de valor com o número de classe e também aumentam de valor em função do número de poses destinadas ao treinamento e ao reconhecimento.

As variáveis da Tabela 4.1 são: *VAR* são variações da experiência para o treinamento e o reconhecimento, *NPT* é o número de poses para o treinamento, *NPR* é o número de poses para o reconhecimento de faces. Ainda, os parâmetros *AC* é a taxa de acerto em percentual, *TT* é o tempo de treinamento em segundos e *TR* é o tempo de reconhecimento em segundos.

Tabela 4.1: *Resultados obtidos do reconhecimento de face utilizando 64 coeficientes DCT de toda pose, arquivo ORL.*

V A R	N P T	N P R	Parâmetros	Classes							
				2	5	10	15	20	25	30	40
1	9	1	AC (%)	100	100	90	90	95	90	93,33	90
			TT (s)	6,711	15,034	34,029	50,549	67,75	83,821	99,747	133,09
			TR (s)	0,00519	0,0097	0,0194	0,0387	0,058	0,0859	0,1204	0,204
2	8	2	AC (%)	100	100	100	100	90	90	91,66	86,25
			TT (s)	3,58	13,251	27,482	41,271	55,76	68,571	82,35	109,57
			TR (s)	0,00716	0,0409	0,0818	0,0939	0,106	0,164	0,225	0,382
3	7	3	AC (%)	100	100	96,66	90	83,33	83,33	90	81,66
			TT (s)	4,481	11,321	22,275	33,549	45,008	56,27	67,61	89,312
			TR (s)	0,00433	0,024	0,0481	0,0998	0,1515	0,225	0,306	0,544
4	6	4	AC (%)	100	100	90	88	86,25	83,33	83,33	77,5
			TT (s)	3,547	8,821	17,042	26,135	35,482	43,512	52,52	76,72
			TR (s)	0,0073	0,0294	0,0589	0,1258	0,1927	0,291	0,391	0,693
5	5	5	AC (%)	100	100	100	90	87	84	84	80,50
			TT (s)	2,841	6,534	13,531	20,191	27,534	33,594	40,28	53,891
			TR (s)	0,0131	0,0317	0,0634	0,143	0,223	0,359	0,467	0,816

Nas Figuras 4.2 a 4.5, são mostrados os gráficos resultantes dessa experiência.

A Figura 4.2 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Nota-se pelas curvas do gráfico a tendência decrescente na taxa de acerto em relação ao aumento do número de classes. Percebe-se que para o reconhecimento com uma pose a taxa de acerto é de 100% para 2 classe flutuou, chegando a 90% para 40 classes. Isto é normal, pois, quanto maior o número de pose maior é o número de matrizes características para serem comparadas, isto aumenta a probabilidade de erro. Nota-se também a flutuação de valores entre a décima quinta a trigésima classe. A oscilação desses valores se explica pelo fato de as vezes ocorrer saturação no processamento do algoritmo e ainda pela diferença de centralização e de iluminação em certas poses dentro das respectivas classes.

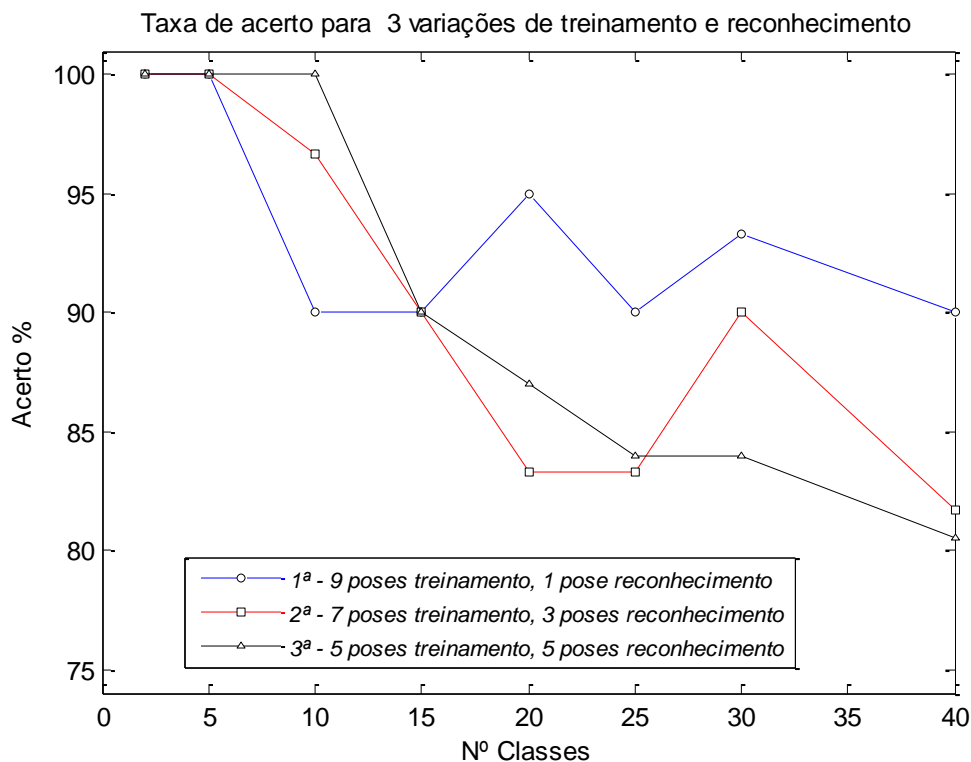


Figura 4.2: Taxa de acerto do AG considerando 64 coeficientes, arquivo ORL.

A Figura 4.3 é mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente linearmente progressivo com o aumento do número de classes. Nota-se que quanto maior o número de poses utilizados no treinamento mais acentuado é o crescimento dos valores do tempo de treinamento, 133,09 s para treinamento com 9 poses.

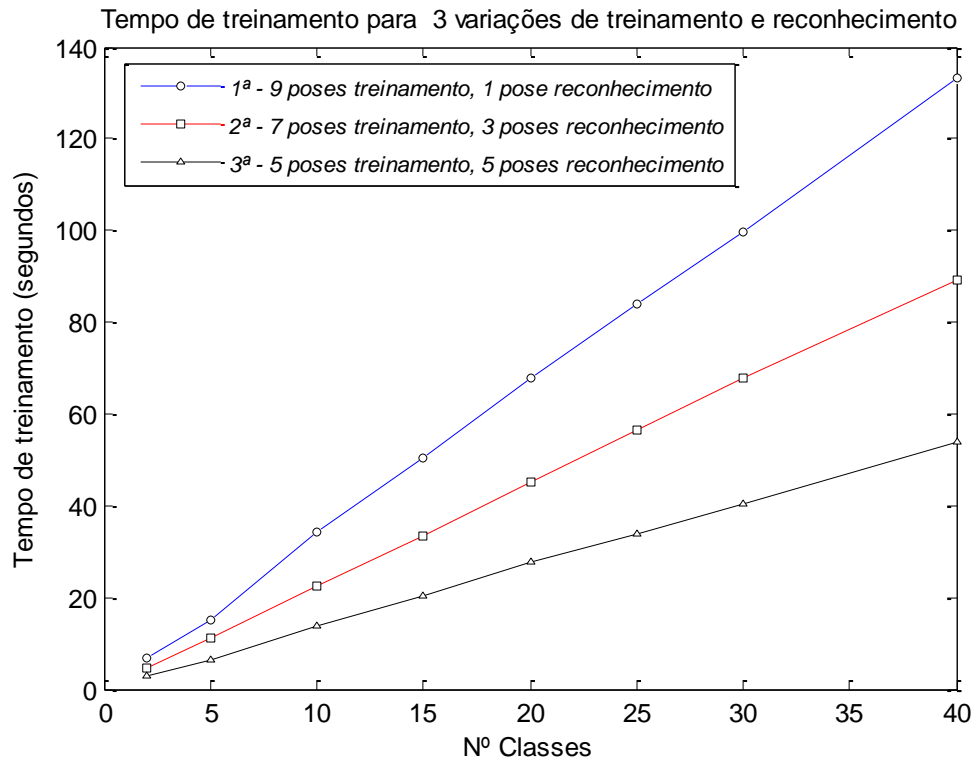


Figura 4.3: Tempo de treinamento do AG considerando 64 coeficientes, arquivo ORL.

A Figura 4.4 é mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Nota-se que mesmo com aumento do número de poses utilizados no reconhecimento o AG efetua o reconhecimento em um tempo bastante curto, 0,816 s para reconhecimento com 5 poses.

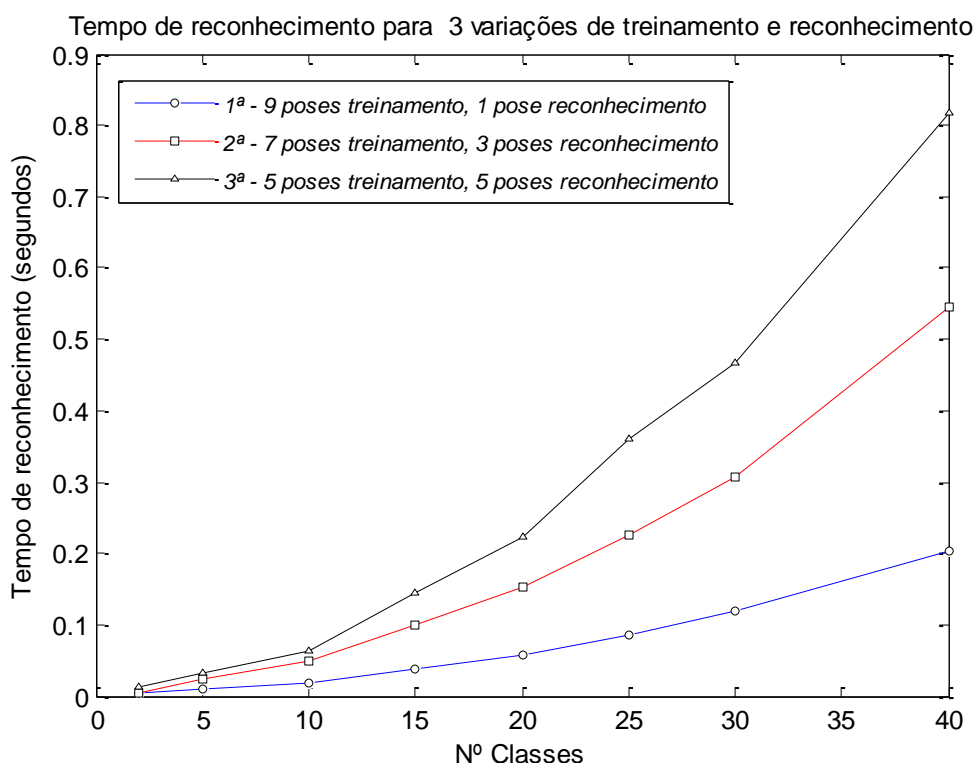


Figura 4.4: Tempo de reconhecimento do AG considerando 64 coeficientes por pose, arquivo ORL.

A Figura 4.5 é mostra o gráfico resultante dessa experiência referente ao número de erros ocorridos, considerando 5 poses para o treinamento e 5 poses para o reconhecimento e total de 40 classes. Nesse teste utilizou-se 200 poses (5 de cada classe) para o reconhecimento. O gráfico mostra a variação de erros na execução de 100 gerações para 40 classes. Assim, ao término da centésima geração registra-se a ocorrência de 39 poses identificadas erradamente. Ou seja, das 200 poses, em 39 existe erro no reconhecimento. Por isso a taxa de acerto de 80,50 % mostrada na Tabela 4.1. Nota-se que o número de erros oscila em função do número de geração processada e somente neste caso, a taxa de erro foi menor na décima terceira geração. Se alterar o número de classe, o número de poses para o treinamento e o número de poses para o reconhecimento, os valores da taxa de erro alteram também. Por isso, o sistema tem que fixar um número de geração que seja capaz de fornecer uma taxa alta de acerto mesmo com a manipulação das referidas variáveis.

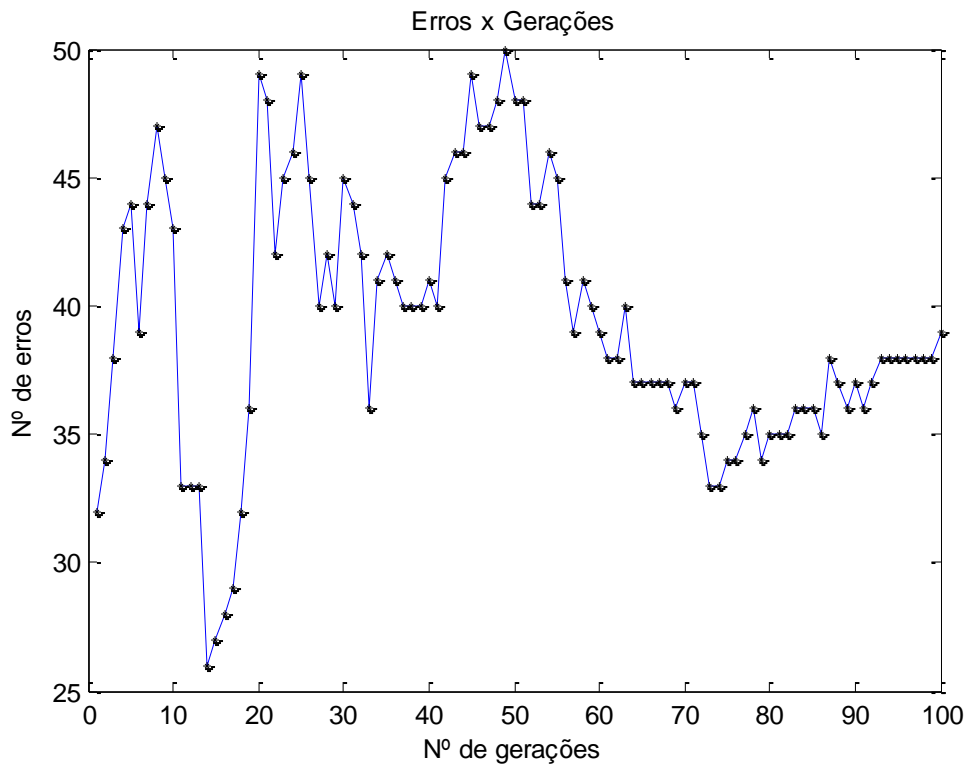


Figura 4.5: Gráficos de erros do AG em função do número de gerações sendo 5 poses para treinamento e 5 poses para o reconhecimento – usando 64 coeficientes DCT, para 40 classes, arquivo ORL.

Sobre essa variação com 64 coeficientes DCT, para 40 classes e dados de 100 gerações, ocorreu:

- cinco imagens para o treinamento por classe;
- cinco imagens para o reconhecimento por classe;
- 200 poses analisadas;
- 161 poses reconhecidas corretamente; e
- 39 poses reconhecidas incorretamente.

4.2.2- Quanto a utilização de subimagens (subposes) ORL

O preparo das subimagens foi mostrado na seção 3.3.1.1.2.

4.2.2.1- Utilizando de 154 coeficientes DCT, com um de cada subpose do arquivo ORL

Passo 1: Considerando a Figura 3.3, tomou-se primeiramente a pose $P_{1,1}$. Desta pose, que é representada Na Figura 3.10, retirou-se a subpose X_1 , que, por sua vez, está mostrada na Figura 4.6 com destaque para o coeficiente mais significativo. Pode-se observar que a subpose é composta de 64 coeficientes.

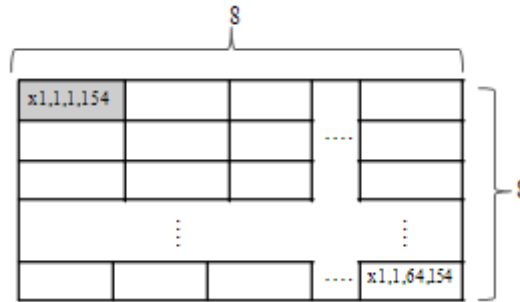


Figura 4.6: Representação da subpose X_1 , base para o vetor característica de 154 coeficientes DCT do arquivo ORL.

Passo 2: Extraíu-se da subpose X_1 o coeficiente mais significativo $x_{1,1,1,154}$, colocando-o em um vetor. Com o mesmo raciocínio, voltou-se à Figura 3.8 e utilizou-se a pose X_2 , extraíu-se o coeficiente mais significativo $x_{1,2,1,154}$ da subpose X_2 e colocou-se no mesmo vetor, assim, sucessivamente, até extrair o coeficiente $x_{1,154,1,154}$ da subpose X_{154} e colocando-o no citado vetor que é mostrado completo na Figura 4.7. Desta forma, a pose $P_{1,1}$ (primeira pose da classe C_1) foi representada por um vetor de 154 elementos.

Para melhor entendimento, segue a descrição dos coeficientes:

$x_{1,1,1,154} - x_1$ refere-se à pose 1. O número 1 seguinte representa a subpose 1. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$x_{1,2,1,154} - x_1$ refere-se à pose 1. O número 2 seguinte representa a subpose 2. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$x_{1,154,1,154} - x_1$ refere-se à pose 1. O número 154 seguinte representa a subpose 154 e o último número 1 corresponde ao primeiro coeficiente mais significativo.

O índice 154 no final de cada símbolo do coeficiente refere-se ao caso em estudo desta seção e também é o total de elementos do vetor que representa uma pose.

O símbolo $P_{1,1}$ quer dizer pose 1 da classe C_1 .

$P_{1,1}$	x1,1,1,154	x1,2,1,154	x1,3,1,154	x1,4,1,154	x1,154,1,154
-----------	------------	------------	------------	------------	-------	--------------

Figura 4.7: Vetor representativo da primeira pose P utilizando 154 coeficientes DCT mais significativo de cada subpose- arquivo ORL.

Passo 3: Repetiram-se os passos 1 e 2 para as poses $P_{2,1}$ até $P_{10,1}$ (Figura 3.3), obtendo-se o conjunto de 10 vetores que formam a matriz de coeficientes DCT representativa das 10 poses da classe C_1 . Este conjunto é mostrado na Figura 4.8.

$P_{1,1}$	x1,1,1,154	x1,2,1,154	x1,3,1,154	x1,154,1,154
$P_{2,1}$	x2,1,1,154	x2,2,1,154	x2,3,1,154	x2,154,1,154
$P_{3,1}$	x3,1,1,154	x3,2,1,154	x3,3,1,154	x3,154,1,154
	⋮	⋮	⋮	...	⋮
$P_{10,1}$	x10,1,1,154	x10,2,1,154	x10,3,1,154	x10,154,1,154

Figura 4.8: Conjunto de vetores com 154 coeficientes DCT representativo da classe C_1 - arquivo ORL.

Passo 4: Repetiram-se os passos desta seção para achar o conjunto de vetores representativos das classes C_2 , C_3 , até a classe C_{40} .

A Tabela 4.2 mostra o resultado do reconhecimento de face utilizando 154 coeficientes DCT para representar cada pose. Os valores contidos nessa tabela são aqueles obtidos nos testes para 2, 5, 10, 15, 20, 25, 30 e 40 classes. Até a vigésima classe obtêm-se uma taxa de acerto muito alta e a partir dessa classe a taxa de acerto diminui. Os testes mostram que o tempo de treinamento e o tempo de reconhecimento aumentam de valor com o número de classe e também aumentam de valor em função do número de poses destinadas ao treinamento e ao reconhecimento.

As variáveis da Tabela 4.2 são: VAR são variações da experiência para o treinamento e o reconhecimento, NPT é o número de poses para o treinamento, NPR é o número de poses para o reconhecimento de faces. Ainda, os parâmetros AC é a taxa de acerto em percentual, TT é o tempo de treinamento em segundos e TR é o tempo de reconhecimento em segundos.

Tabela 4.2: Resultados obtidos do reconhecimento de face considerando 154 coeficientes DCT arquivo ORL.

V A R	N P T	N P R	Parâmetros	Classes							
				2	5	10	15	20	25	30	40
1	9	1	AC (%)	100	100	100	100	95	100	96,66	92,50
			TT (s)	7,46	19,26	37,04	54,78	72,78	97,97	109,43	156,76
			TR (s)	0,02	0,02	0,04	0,06	0,10	0,15	0,18	0,319
2	8	2	AC (%)	100	100	100	100	92,51	96	91,66	93,75
			TT (s)	6,04	16,37	32,74	48,97	65,84	82,14	98,52	130,932
			TR (s)	0,02	0,02	0,05	0,12	0,17	0,26	0,35	0,593
3	7	3	AC (%)	100	100	100	100	93,33	92	90	87,50
			TT (s)	4,98	13,64	26,21	39,51	52,52	65,93	83,65	105,56
			TR (s)	0,02	0,04	0,07	0,13	0,21	0,37	0,49	0,8444
4	6	4	AC (%)	100	85	100	98,33	93,75	94	86,66	88,75
			TT (s)	4,09	11,15	21,53	32,18	42,94	54,02	64,91	85,869
			TR (s)	0,02	0,04	0,11	0,18	0,29	0,45	0,63	1,085
5	5	5	AC (%)	100	96	96	94,66	92	92	88	83
			TT (s)	2,99	8,25	16,33	24,57	32,71	40,97	49,45	65,862
			TR (s)	0,03	0,05	0,09	0,20	0,35	0,52	0,76	1,293

Nas Figuras 4.9 a 4.12, são mostrados os gráficos resultantes dessa experiência.

A Figura 4.2 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Nota-se pelas curvas do gráfico a tendência decrescente na taxa de acerto em relação ao aumento do número de classes. Percebe-se que para o reconhecimento com uma pose a taxa de acerto é de 100% para 2 classe flutuou, chegando a 92,50% para 40 classes. Isto é normal, pois, quanto maior o número de pose maior é o número de matrizes características para serem comparadas, isto aumenta a probabilidade de erro. Nota-se também a flutuação de valores entre a décima quinta a trigésima classe. A oscilação desses valores se explica pelo fato de as vezes ocorrer saturação no processamento do algoritmo e ainda pela diferença de centralização e de iluminação em certas poses dentro das respectivas classes.

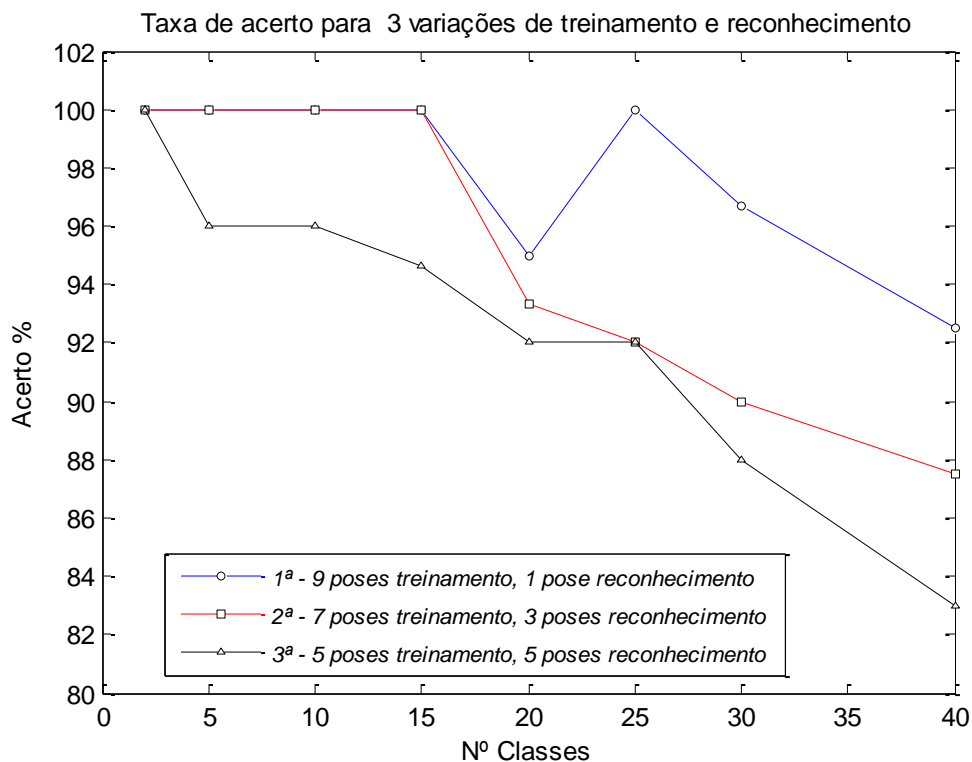


Figura 4.9: Taxa de acerto do AG considerando 154 coeficientes, arquivo ORL.

A Figura 4.10 é mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente linearmente progressivo com o aumento do número de classes. Nota-se que quanto maior o número de poses utilizados no treinamento mais acentuado é o crescimento dos valores do tempo de treinamento, 156,76 s para treinamento com 9 poses.

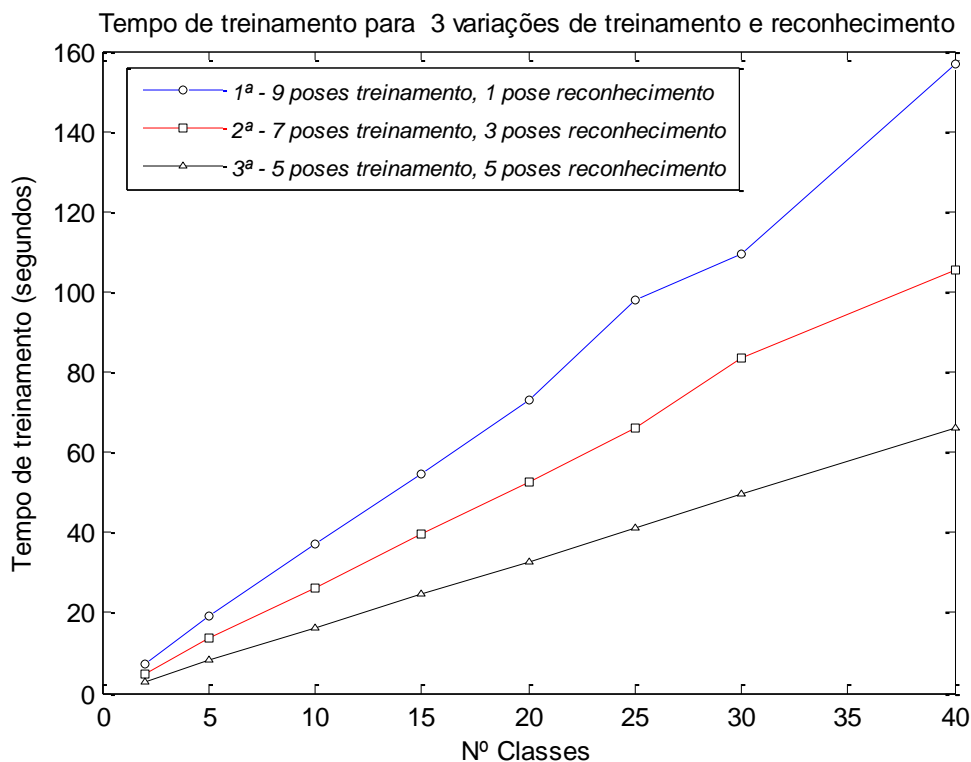


Figura 4.10: Tempo de treinamento do AG considerando 154 coeficientes, arquivo ORL.

A Figura 4.11 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Nota-se que mesmo com aumento do número de poses utilizados no reconhecimento o AG efetua o reconhecimento em um tempo bastante curto, 1,293 s, para reconhecimento com 5 poses.

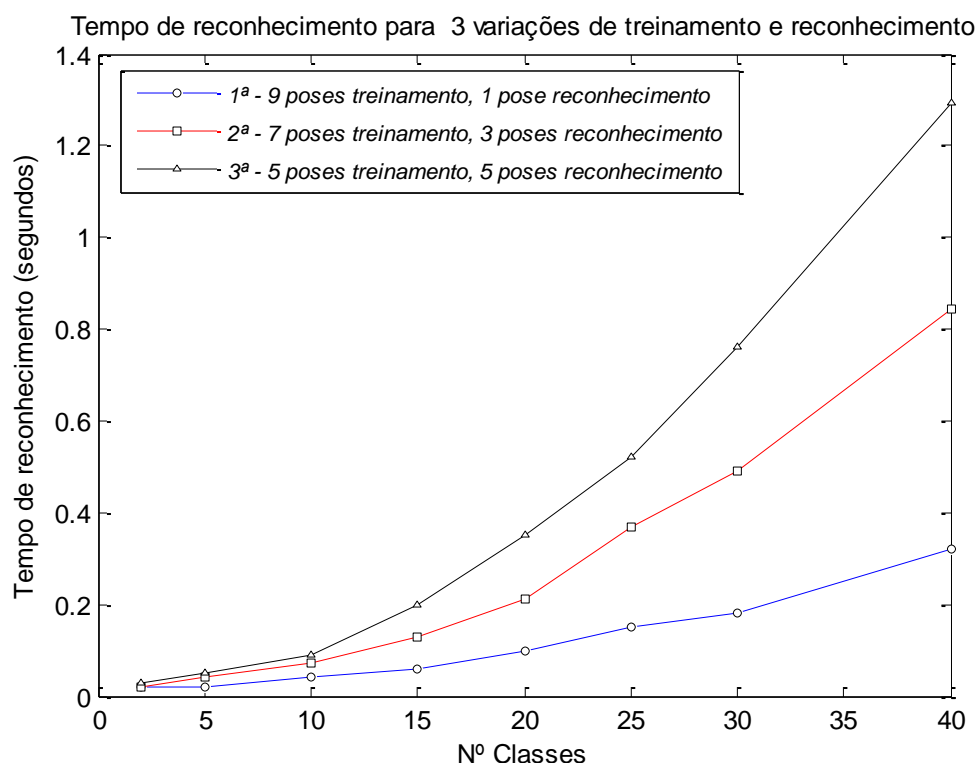


Figura 4.11: Tempo de reconhecimento do AG considerando 154 coeficientes por pose, arquivo ORL.

A Figura 4.12 mostra o gráfico resultante dessa experiência referente ao número de erros ocorridos, considerando 5 poses para o treinamento e 5 poses para o reconhecimento e total de 40 classes. Nesse teste utilizou-se 200 poses (5 de cada classe) para o reconhecimento. O gráfico mostra a variação de erros na execução de 100 gerações para 40 classes. Assim, ao término da centésima geração registra-se a ocorrência de 34 poses identificadas erradamente. Ou seja, das 200 poses, em 34 delas existe erro no reconhecimento. Por isso a taxa de acerto de 83% mostrada na Tabela 4.2. Nota-se que o número de erros oscila em função do número de geração processada e somente neste caso, a taxa de erro foi menor na décima terceira geração. Se alterar o número de classe, o número de poses para o treinamento e o número de poses para o reconhecimento, os valores da taxa de erro alteram também. Por isso, o sistema tem que fixar um número de geração que seja capaz de fornecer uma taxa alta de acerto mesmo com a manipulação das referidas variáveis.

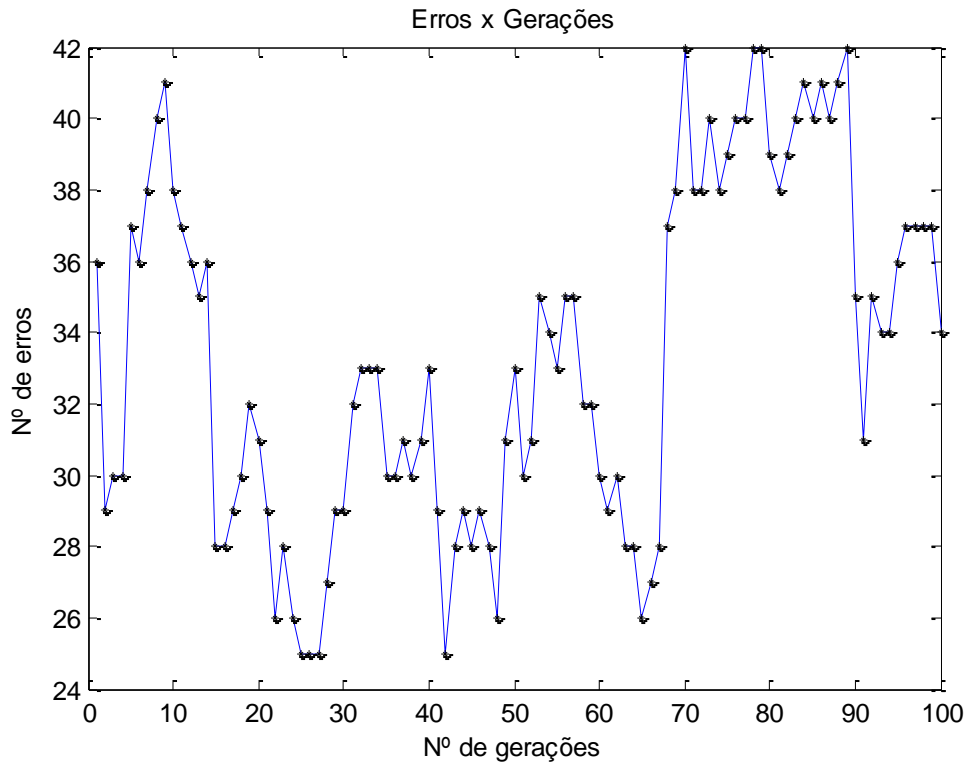


Figura 4.12: Gráficos de erros do AG em função do número de gerações sendo cinco poses para treinamento e cinco para o reconhecimento – usando 154 coeficientes DCT, para 40 classes, arquivo ORL.

Sobre essa variação com 154 coeficientes DCT, para 40 classes e dados de 100 gerações, ocorreu:

- cinco imagens para o treinamento por classe;
- cinco imagens para o reconhecimento por classe;
- 200 poses analisadas;
- 166 poses reconhecidas corretamente; e
- 34 poses reconhecidas incorretamente.

4.2.2.2- Utilizando 308 coeficientes DCT, com dois de cada subpose ORL

Passo 1: Considerando a Figura 3.3, tomou-se primeiramente a pose $P_{1,1}$. Desta pose, que é mostrada na pela Figura 3.10, retirou-se a subpose X_I , que, por sua vez, é ilustrada na Figura 4.13 com destaque para os dois coeficientes mais significativos.

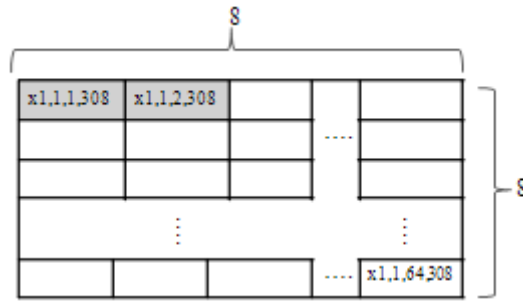


Figura 4.13: Representação da subpose X_1 , base para o vetor característico de 308 coeficientes DCT – ORL.

Passo 2: Extraíram-se da subpose X_1 os coeficientes mais significativos $x_{1,1,1,308}$ e $x_{1,1,2,308}$, colocando-os em um vetor. Com o mesmo raciocínio, voltou-se à Figura 3.10 e pegou-se a pose X_2 , extraíram-se os coeficientes mais significativos $x_{1,2,1,308}$ e $x_{1,2,2,308}$, colocando-os no mesmo vetor e, assim, sucessivamente, até extrair os coeficientes $x_{1,154,1,308}$ e $x_{1,154,2,308}$ da subpose X_{154} e colocando-os no citado vetor que é mostrado completo na Figura 4.14. Desta forma, a pose $P_{1,1}$ (primeira pose da classe C_1) foi representada por um vetor de 308 elementos.

Para melhor entendimento, segue a descrição dos coeficientes:

$x_{1,1,1,308}$ – x_1 refere-se à pose 1. O número 1 seguinte quer dizer subpose 1. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$x_{1,1,2,308}$ – x_1 refere-se a pose 1. O número 1 seguinte quer dizer subpose 1. O número 2 seguinte quer dizer o segundo coeficiente mais significativo.

$x_{1,2,1,308}$ – x_1 refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$x_{1,2,2,308}$ – x_1 refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 2 seguinte quer dizer o segundo coeficiente mais significativo.

$x_{1,154,1,308}$ – x_1 refere-se a pose 1. O número 154 seguinte quer dizer subpose 154. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$x_{1,154,2,308}$ – x_1 refere-se à pose 1. O número 154 seguinte quer dizer subpose 154. O número 2 seguinte quer dizer o segundo coeficiente mais significativo.

Ou seja, de cada subpose extraíram-se dois coeficientes.

O índice 308 no final de cada símbolo do coeficiente refere-se ao caso em estudo desta seção e também é o total de elementos do vetor que representa uma pose.

$P_{1,1}$	x1,1,1,308	x1,1,2,308	x1,2,1,308	x1,2,2,308	x1,154,1,308	x1,154,2,308
-----------	------------	------------	------------	------------	-------	--------------	--------------

Figura 4.14: Vetor representativo da primeira pose P por meio de 308 coeficientes DCT mais significativos de cada subpose- arquivo ORL.

Passo 3: Repetiram-se os passos 1 e 2 para as poses $P_{2,1}$ até $P_{10,1}$ (Figura 3.3), obtendo-se o conjunto de 10 vetores que formam a matriz de coeficientes DCT representativa das 10 poses da classe C_1 . Esse conjunto é mostrado na Figura 4.15.

$P_{1,1}$	x1,1,1,308	x1,1,2,308	x1,2,1,308	x1,2,2,308	x1,154,1,308	x1,154,2,308
$P_{2,1}$	x2,1,1,308	x2,1,2,308	x2,2,1,308	x2,2,2,308	x2,154,1,308	x2,154,2,308
$P_{3,1}$	x3,1,1,308	x3,1,2,308	x3,2,1,308	x3,2,2,308	x3,154,1,308	x3,154,2,308
	\vdots	\vdots	\vdots	\vdots	...	\vdots	\vdots
$P_{10,1}$	x10,1,1,308	x10,1,2,308	x10,2,1,308	x10,2,2,308	x10,154,1,308	x10,154,2,308

Figura 4.15: Conjunto de vetores com 308 coeficientes DCT representativo da classe C_1 - arquivo ORL.

Passo 4: Repetiram-se os passos desta seção para encontrar o conjunto de vetores representativos das classes C_2 , C_3 , até a classe C_{40} . A Tabela 4.3 mostra o resultado do reconhecimento de face utilizando-se dois coeficientes DCT da cada subpose, totalizando 308 coeficientes para representar cada pose.

A Tabela 4.3 mostra o resultado do reconhecimento de face utilizando 308 coeficientes DCT para representar cada pose. Os valores contidos nessa tabela são aqueles obtidos nos testes para 2, 5, 10, 15, 20, 25, 30 e 40 classes. Até a vigésima classe obtêm-se uma taxa de acerto muito alta e a partir dessa classe a taxa de acerto diminui. Os testes mostram que o tempo de treinamento e o tempo de reconhecimento aumentam de valor com o número de classe e também aumentam de valor em função do número de poses destinadas ao treinamento e ao reconhecimento.

As variáveis da Tabela 4.3 são: VAR são variações da experiência para o treinamento e o reconhecimento, NPT é o número de poses para o treinamento, NPR é o número de poses para o reconhecimento de faces. Ainda, os parâmetros AC é a taxa de acerto em percentual, TT é o tempo de treinamento em segundos e TR é o tempo de reconhecimento em segundos.

Tabela 4.3: Resultados obtidos do reconhecimento de face considerando 308 coeficientes DCT arquivo ORL.

V A R	N P T	N P R	Parâmetros	Classes							
				2	5	10	15	20	25	30	40
1	9	1	AC (%)	100	100	100	100	95	96	90	90
			TT (s)	10,28	24,71	52,41	77,02	97,86	128,97	145,33	195
			TR (s)	0,02	0,03	0,07	0,09	0,12	0,19	0,27	0,448
2	8	2	AC (%)	100	100	100	100	95	98	90	88,75
			TT (s)	8,61	20,65	41,60	61,04	81,29	103,46	122,08	162,90
			TR (s)	0,02	0,04	0,09	0,15	0,22	0,37	0,51	0,829
3	7	3	AC (%)	100	100	96,66	91,11	90	96	90	91,66
			TT (s)	6,95	16,87	33,94	49,88	66,38	84,42	98,85	134,43
			TR (s)	0,02	0,04	0,10	0,21	0,32	0,49	0,58	1,215
4	6	4	AC (%)	100	100	100	100	93,75	92	91,66	88,12
			TT (s)	5,51	13,53	27,22	40,05	53,48	67,55	79,39	106,79
			TR (s)	0,01	0,06	0,13	0,23	0,42	0,63	0,87	1,563
5	5	5	AC (%)	100	100	98	94,66	86	91,20	89,33	86
			TT (s)	4,39	10,76	21,50	31,63	42,23	53,60	64,24	84,83
			TR (s)	0,03	0,05	0,16	0,29	0,50	0,75	1,08	1,889

Nas Figuras 4.16 a 4.19, são mostrados os gráficos resultantes dessa experiência.

A Figura 4.16 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Nota-se pelas curvas do gráfico a tendência decrescente na taxa de acerto em relação ao aumento do número de classes. Percebe-se que para o reconhecimento com uma pose a taxa de acerto é de 100% para 2 classe flutuou, chegando a 90% para 40 classes. Isto é normal, pois, quanto maior o número de pose maior é o número de matrizes características para serem comparadas, isto aumenta a probabilidade de erro. Nota-se também a flutuação de valores entre a décima quinta a trigésima classe. A oscilação desses valores se explica pelo fato de as vezes ocorrer saturação no processamento do algoritmo e ainda pela diferença de centralização e de iluminação em certas poses dentro das respectivas classes.

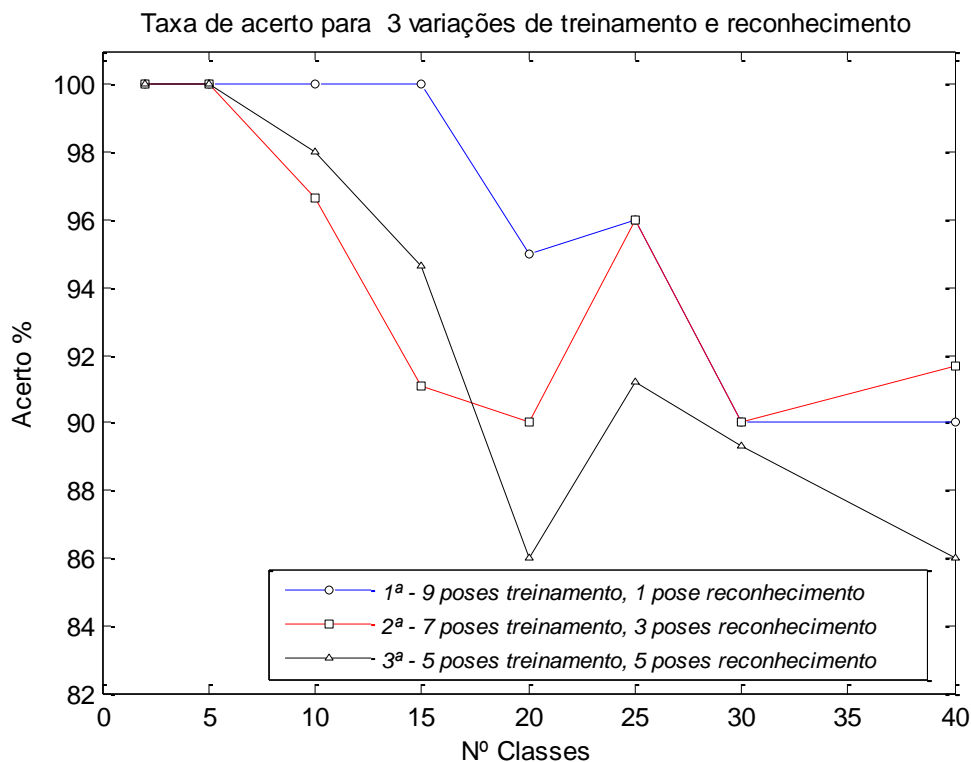


Figura 4.16: Taxa de acerto do AG considerando 308 coeficientes, arquivo ORL.

A Figura 4.17 é mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente linearmente progressivo com o aumento do número de classes. Nota-se que quanto maior o número de poses utilizados no treinamento mais acentuado é o crescimento dos valores do tempo de treinamento, 195 s para treinamento com 9 poses.

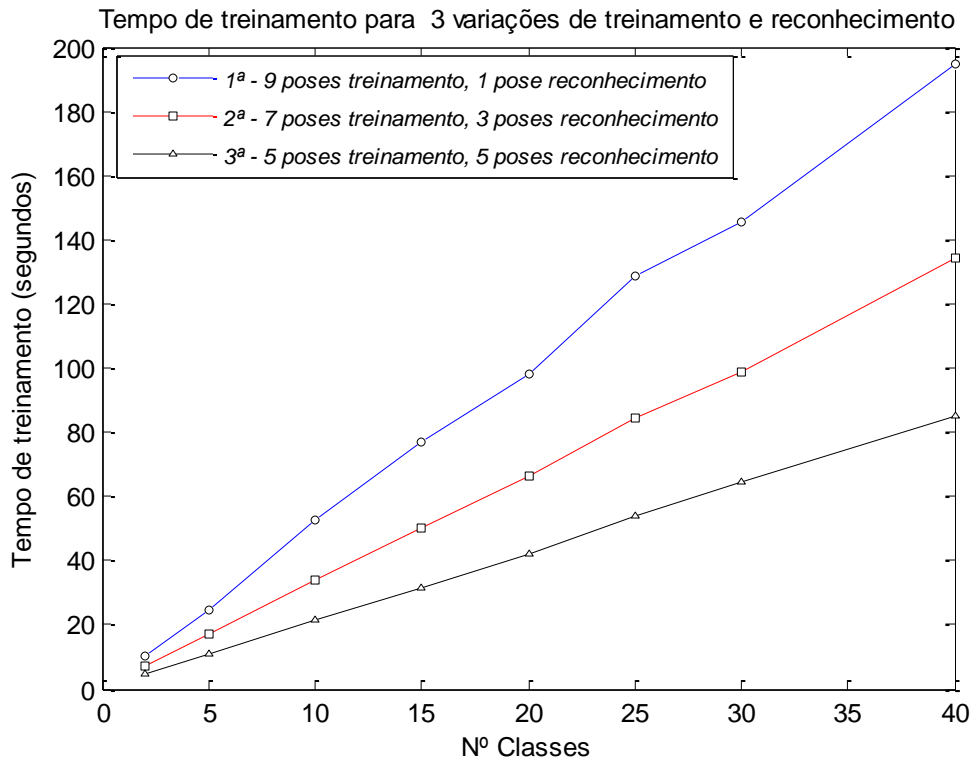


Figura 4.17: Tempo de treinamento do AG considerando 308 coeficientes, arquivo ORL.

A Figura 4.18 é mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Nota-se que mesmo com aumento do número de poses utilizados no reconhecimento o AG efetua o reconhecimento em um tempo bastante curto, 1,889 s para reconhecimento com 5 poses.

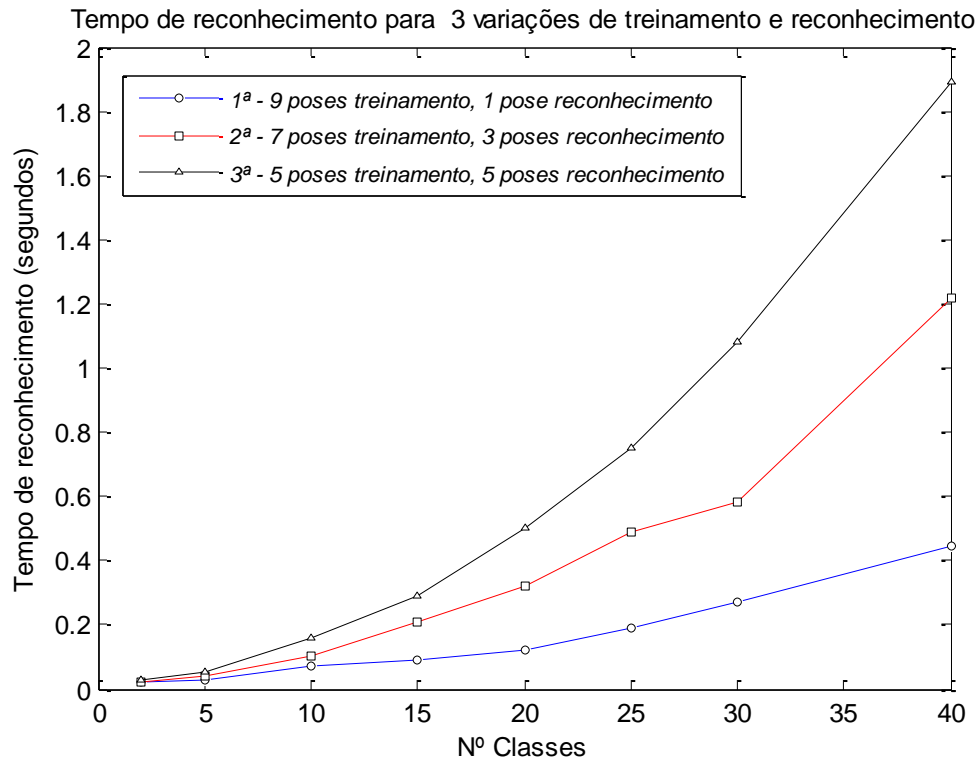


Figura 4.18: Tempo de reconhecimento do AG considerando 308 coeficientes por pose, arquivo ORL.

A Figura 4.19 mostra o gráfico resultante dessa experiência referente ao número de erros ocorridos, considerando 5 poses para o treinamento e 5 poses para o reconhecimento e total de 40 classes. Nesse teste utilizou-se 200 poses (5 de cada classe) para o reconhecimento. O gráfico mostra a variação de erros na execução de 100 gerações para 40 classes. Assim, ao término da centésima geração registra-se a ocorrência de 28 poses identificadas erradamente. Ou seja, das 200 poses, em 28 existe erro no reconhecimento. Por isso a taxa de acerto de 86 % mostrada na Tabela 4.3. Nota-se que o número de erros oscila em função do número de geração processada e somente neste caso, a taxa de erro foi menor na décima terceira geração. Se alterar o número de classe, o número de poses para o treinamento e o número de poses para o reconhecimento, os valores da taxa de erro alteram também. Por isso, o sistema tem que fixar um número de geração que seja capaz de fornecer uma taxa alta de acerto mesmo com a manipulação das referidas variáveis.

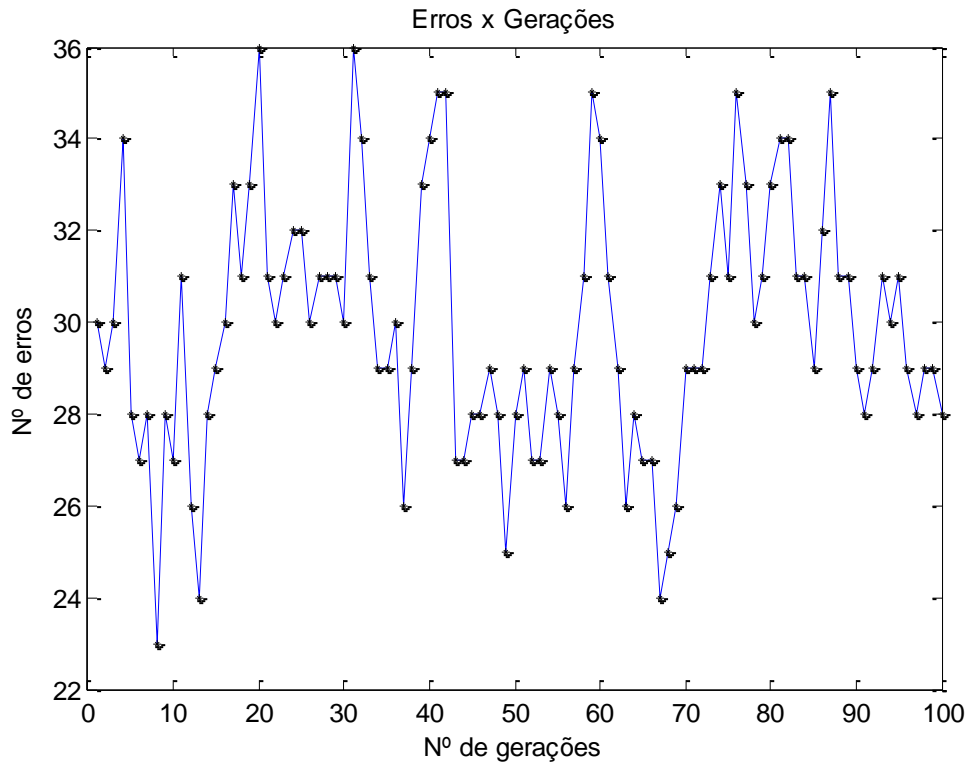


Figura 4.19: Gráficos de erros do AG em função do número de gerações sendo cinco poses para treinamento e cinco poses para o reconhecimento – usando 308 coeficientes DCT, para 40 classes, arquivo ORL.

Sobre esta variação com 308 coeficientes DCT, para 40 classes e dados de 100 gerações, obtiveram-se:

- cinco imagens para o treinamento por classe;
- cinco imagens para o reconhecimento por classe;
- 200 poses analisadas;
- 172 poses reconhecidas corretamente; e
- 28 poses reconhecidas incorretamente.

4.2.3.3- Utilizando 616 coeficientes DCT, sendo quatro são de cada subpose ORL

Passo 1: Considerando a Figura 3.3, tomou-se primeiramente a pose $P_{1,1}$. Desta pose, que é representada pela Figura 3.10, retirou-se a subpose X_1 , que, por sua vez, é mostrada na Figura 4.20 com destaque para os quatro coeficientes mais significativos.

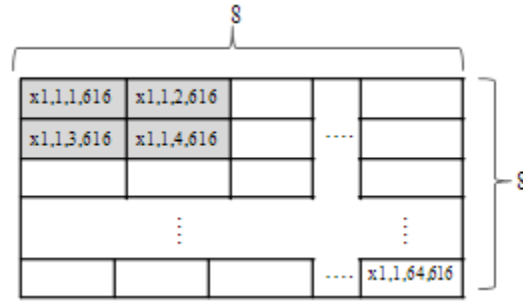


Figura 4.20: Representação da subpose X_1 , base para o vetor característica de 616 coeficientes DCT – ORL.

Passo 2: Extraíram-se da subpose X_1 os coeficientes mais significativos $x_{1,1,1,616}$, $x_{1,1,2,616}$, $x_{1,1,3,616}$ e $x_{1,1,4,616}$, colocando-os em um vetor. Depois na Figura 3.10 e pegou-se a pose X_2 , extraíndo dessa os quatro coeficientes mais significativos $x_{1,2,1,616}$, $x_{1,2,2,616}$, $x_{1,2,3,616}$ e $x_{1,2,4,616}$ colocando-os no mesmo vetor e assim sucessivamente até extrair os coeficientes $x_{1,154,1,616}$, $x_{1,154,2,616}$, $x_{1,154,3,616}$ e $x_{1,154,4,616}$ da subpose X_{154} e colocando-os nesse vetor mostrado completo na Figura 4.21. Desta forma, a pose $P_{1,1}$ (primeira pose da classe C_1) foi representada por um vetor de 616 elementos.

Para melhor entendimento, segue a descrição dos coeficientes:

$x_{1,1,1,616} - x_1$ refere-se à pose 1. O número 1 seguinte quer dizer subpose 1. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$x_{1,1,2,616} - x_1$ refere-se à pose 1. O número 1 seguinte quer dizer subpose 1. O número 2 seguinte quer dizer o segundo coeficiente mais significativo.

$x_{1,1,3,616} - x_1$ refere-se à pose 1. O número 1 seguinte quer dizer subpose 1. O número 3 seguinte quer dizer o terceiro coeficiente mais significativo.

$x_{1,1,4,616} - x_1$ refere-se à pose 1. O número 1 seguinte quer dizer subpose 1. O número 4 seguinte quer dizer o quarto coeficiente mais significativo.

$x_{1,2,1,616} - x_1$ refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$x_{1,2,2,616} - x_1$ refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 2 seguinte quer dizer o segundo coeficiente mais significativo.

$x_{1,2,3,616} - x_1$ refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 3 seguinte quer dizer o terceiro coeficiente mais significativo.

$x_{1,2,4,616} - x_1$ refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 4 seguinte quer dizer o quarto coeficiente mais significativo.

$x_{1,154,1,616} - x_1$ refere-se à pose 1. O número 154 seguinte quer dizer subpose 154. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$x_{1,154,2,616} - x_1$ refere-se à pose 1. O número 154 seguinte quer dizer subpose 154. O número 2 seguinte quer dizer o segundo coeficiente mais significativo.

$x_{1,154,3,616} - x_1$ refere-se à pose 1. O número 154 seguinte quer dizer subpose 154. O número 3 seguinte quer dizer o terceiro coeficiente mais significativo.

$x_{1,154,4,616} - x_1$ refere-se à pose 1. O número 154 seguinte quer dizer subpose 154. O número 4 seguinte quer dizer o quarto coeficiente mais significativo.

Ou seja, de cada subpose foram extraídos quatro coeficientes.

O índice 616 no final de cada símbolo do coeficiente refere-se ao caso em estudo desta seção e também é o total de elementos do vetor que representa uma pose.

$P_{1,1}$	$x_{1,1,1,616}$	$x_{1,1,2,616}$	$x_{1,1,3,616}$	$x_{1,1,4,616}$	$x_{1,154,1,616}$	$x_{1,154,2,616}$	$x_{1,154,3,616}$	$x_{1,154,4,616}$
-----------	-----------------	-----------------	-----------------	-----------------	-------	-------------------	-------------------	-------------------	-------------------

Figura 4.21: Vetor representativo da primeira pose P por meio de 616 coeficientes DCT mais significativos de cada subpose- arquivo ORL.

Passo 3: Repetiram-se os passos 1 e 2 para as poses $P_{2,1}$ até $P_{10,1}$ (Figura 3.3), obtendo-se o conjunto de 10 vetores que formam a matriz de coeficientes DCT representativa das dez poses da classe C_1 . Este conjunto é mostrado na Figura 4.22.

$P_{1,1}$	$x_{1,1,1,616}$	$x_{1,1,2,616}$	$x_{1,1,3,616}$	$x_{1,1,4,616}$	$x_{1,154,1,616}$	$x_{1,154,2,616}$	$x_{1,154,3,616}$	$x_{1,154,4,616}$
$P_{2,1}$	$x_{2,1,1,616}$	$x_{2,1,2,616}$	$x_{2,1,3,616}$	$x_{2,1,4,616}$	$x_{2,154,1,616}$	$x_{2,154,2,616}$	$x_{2,154,3,616}$	$x_{2,154,4,616}$
$P_{3,1}$	$x_{3,1,1,616}$	$x_{3,1,2,616}$	$x_{3,1,3,616}$	$x_{3,1,4,616}$	$x_{3,154,1,616}$	$x_{3,154,2,616}$	$x_{3,154,3,616}$	$x_{3,154,4,616}$
...	\vdots	\vdots	\vdots	\vdots	...	\vdots	\vdots	\vdots	\vdots
$P_{10,1}$	$x_{10,1,1,616}$	$x_{10,1,2,616}$	$x_{10,1,3,616}$	$x_{10,1,4,616}$	$x_{10,154,1,616}$	$x_{10,154,2,616}$	$x_{10,154,3,616}$	$x_{10,154,4,616}$

Figura 4.22: Conjunto de vetores com 616 coeficientes DCT representativo da classe C_1 - arquivo ORL.

Passo 4: Repetiram-se os passos desta seção para achar o conjunto de vetores representativos das classes C_2 , C_3 , até a classe C_{40} . A Tabela 4.4 mostra o resultado do reconhecimento de face utilizando-se quatro coeficientes DCT da cada subpose, totalizando 616 coeficientes para representar cada pose.

A Tabela 4.4 mostra o resultado do reconhecimento de face utilizando 616 coeficientes DCT para representar cada pose. Os valores contidos nessa tabela são

aqueles obtidos nos testes para 2, 5, 10, 15, 20, 25, 30 e 40 classes. Até a vigésima classe obtêm-se uma taxa de acerto muito alta e a partir dessa classe a taxa de acerto diminui. Os testes mostram que o tempo de treinamento e o tempo de reconhecimento aumentam de valor com o número de classe e também aumentam de valor em função do número de poses destinadas ao treinamento e ao reconhecimento.

As variáveis da Tabela 4.4 são: *VAR* são variações da experiência para o treinamento e o reconhecimento, *NPT* é o número de poses para o treinamento, *NPR* é o número de poses para o reconhecimento de faces. Ainda, os parâmetros *AC* é a taxa de acerto em percentual, *TT* é o tempo de treinamento em segundos e *TR* é o tempo de reconhecimento em segundos.

Tabela 4.4: Resultados obtidos do reconhecimento de face considerando 616 coeficientes DCT da subimagem, arquivo ORL

V A R	N P T	N P R	Parâmetros	Classes							
				2	5	10	15	20	25	30	40
1	9	1	AC (%)	100	100	100	100	96	96,00	93,33	90
			TT (s)	14,00	35,00	75,00	106,00	146	175	210,00	282,00
			TR (s)	0,03	0,06	0,16	0,27	0,33	0,44	0,49	0,713
2	8	2	AC (%)	100	100	100	100	96,00	96,00	93,33	92,50
			TT (s)	11,40	30,05	62,17	91,18	124,34	151,28	181,33	243,50
			TR (s)	0,06	0,17	0,38	0,53	0,70	0,81	0,97	1,341
3	7	3	AC (%)	100	100	100	100	95,00	94,66	88,88	87,50
			TT (s)	10,26	24,64	50,30	73,91	99,57	123,18	147,82	199,14
			TR (s)	0,10	0,25	0,45	0,73	1,23	1,23	1,57	1,961
4	6	4	AC (%)	100	100	100	100	92,50	91,00	91,00	85,00
			TT (s)	6,67	15,83	32,50	48,34	65,00	81,67	96,67	160,84
			TR (s)	0,13	0,35	0,61	1,28	1,50	1,94	2,27	2,501
5	5	5	AC (%)	100	96,00	96,00	96,00	94,00	92,00	91,33	93,00
			TT (s)	3,98	9,96	20,59	30,55	41,17	51,13	61,76	128,16
			TR (s)	0,47	0,84	1,16	1,49	1,91	2,47	2,79	3,032

Nas Figuras 4.23 a 4.26 são mostrados os gráficos resultantes desta experiência.

A Figura 4.23 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Nota-se pelas curvas do gráfico a tendência decrescente na taxa de acerto em relação ao aumento do número de classes. Percebe-se que para o reconhecimento com uma pose a taxa de acerto é de 100% para 2 classe flutuou, chegando a 90% para 40 classes. Isto é normal, pois, quanto maior o número de pose maior é o número de matrizes características para serem comparadas, isto aumenta a probabilidade de erro. Nota-se também a flutuação de valores entre a décima quinta a trigésima classe. A

oscilação desses valores se explica pelo fato de as vezes ocorrer saturação no processamento do algoritmo e ainda pela diferença de centralização e de iluminação em certas poses dentro das respectivas classes.

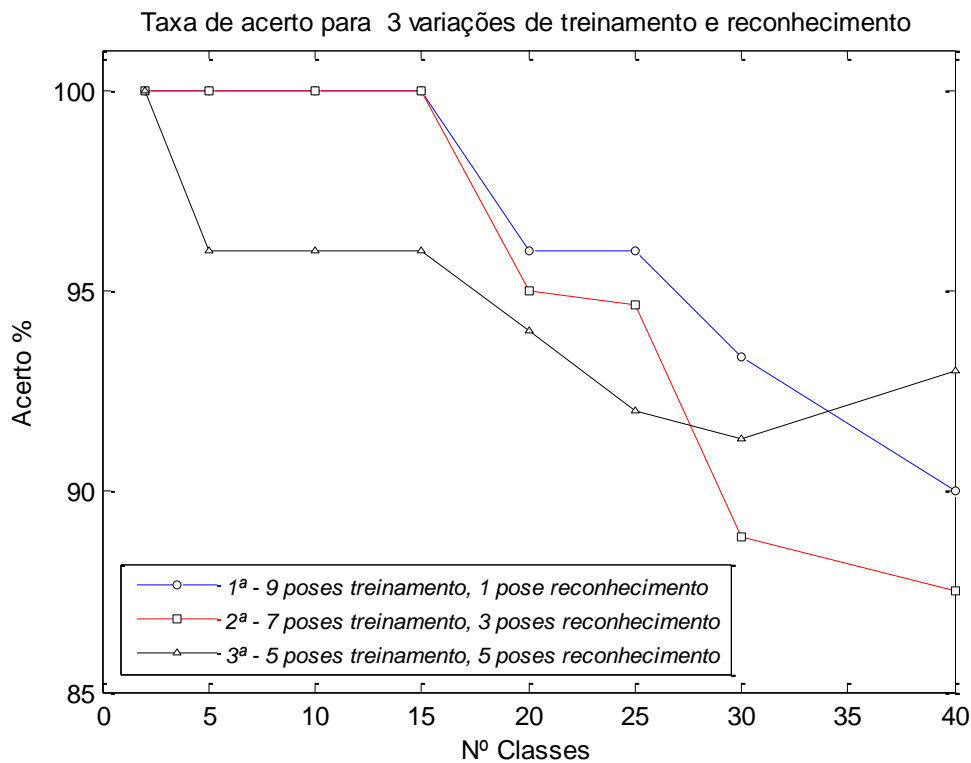


Figura 4.23: Taxa de acerto do AG considerando 616 coeficientes, arquivo ORL.

A Figura 4.24 é mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente linearmente progressivo com o aumento do número de classes. Nota-se que quanto maior o número de poses utilizados no treinamento mais acentuado é o crescimento dos valores do tempo de treinamento, 282,0 s para treinamento com 9 poses.

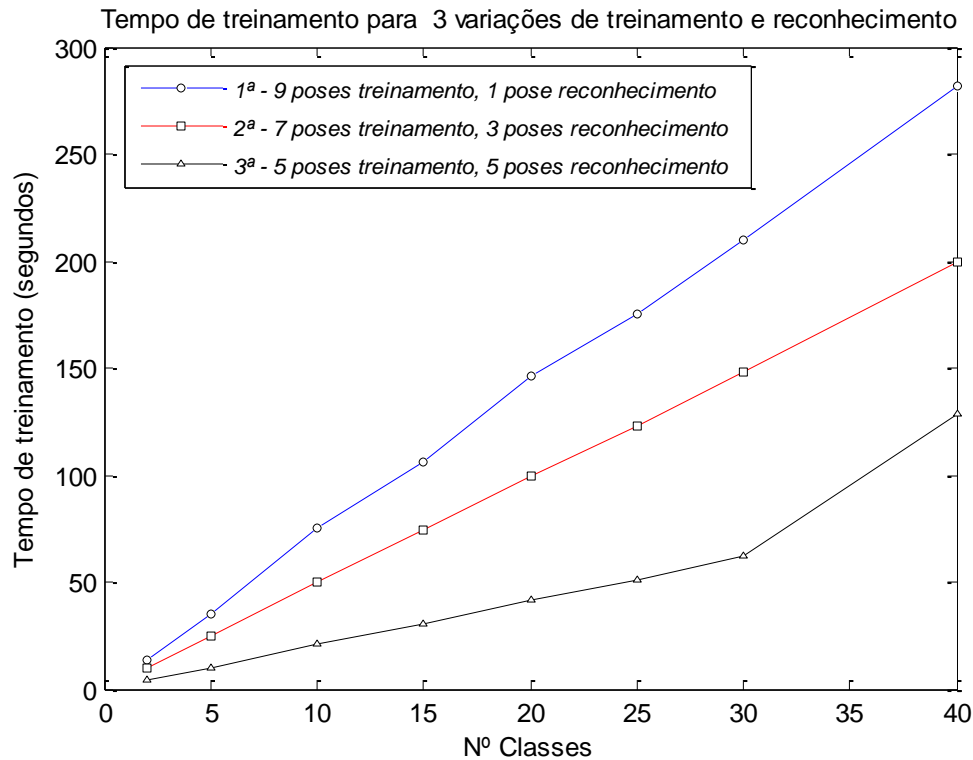


Figura 4.24: Tempo de treinamento do AG considerando 616 coeficientes, arquivo ORL.

A Figura 4.25 é mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Nota-se que mesmo com aumento do número de poses utilizados no reconhecimento o AG efetua o reconhecimento em um tempo bastante curto, 3,032 s para reconhecimento com 5 poses.

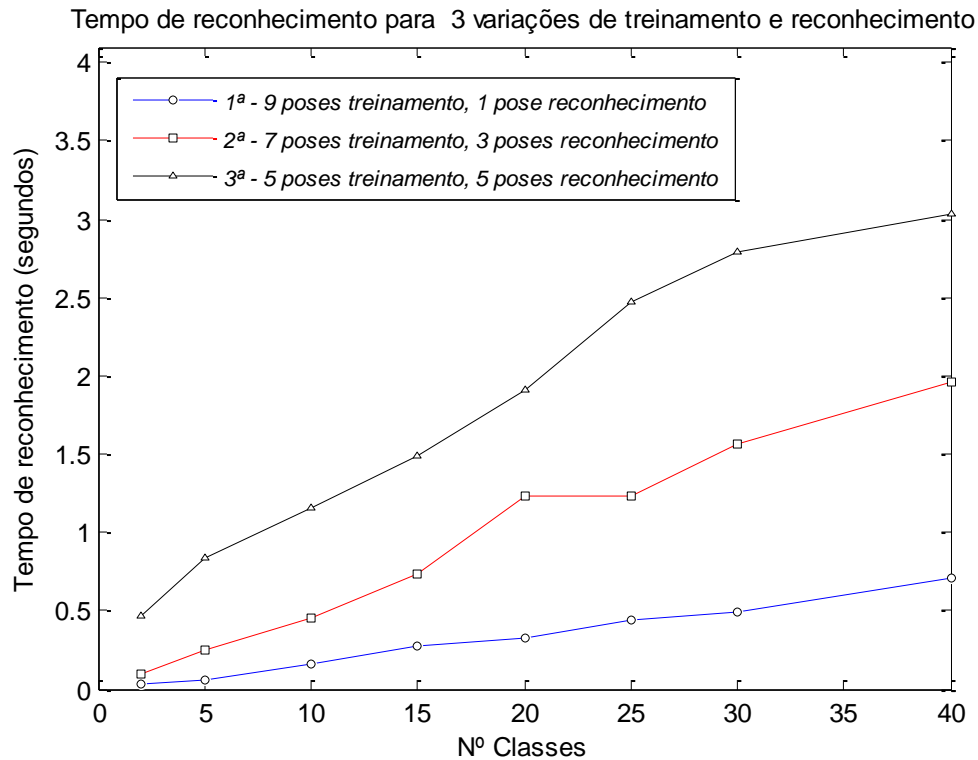


Figura 4.25: Tempo de reconhecimento do AG considerando 616 coeficientes por pose, arquivo ORL.

A Figura 4.26 mostra o gráfico resultante dessa experiência referente ao número de erros ocorridos, considerando 5 poses para o treinamento e 5 poses para o reconhecimento e total de 40 classes. Nesse teste utilizou-se 200 poses (5 de cada classe) para o reconhecimento. O gráfico mostra a variação de erros na execução de 100 gerações para 40 classes. Assim, ao término da centésima geração registra-se a ocorrência de 34 poses identificadas erradamente. Ou seja, das 200 poses, em 34 existe erro no reconhecimento. Por isso a taxa de acerto de 93 % mostrada na Tabela 4.4. Nota-se que o número de erros oscila em função do número de geração processada e somente neste caso, a taxa de erro foi menor na décima terceira geração. Se alterar o número de classe, o número de poses para o treinamento e o número de poses para o reconhecimento, os valores da taxa de erro alteram também. Por isso, o sistema tem que fixar um número de geração que seja capaz de fornecer uma taxa alta de acerto mesmo com a manipulação das referidas variáveis.

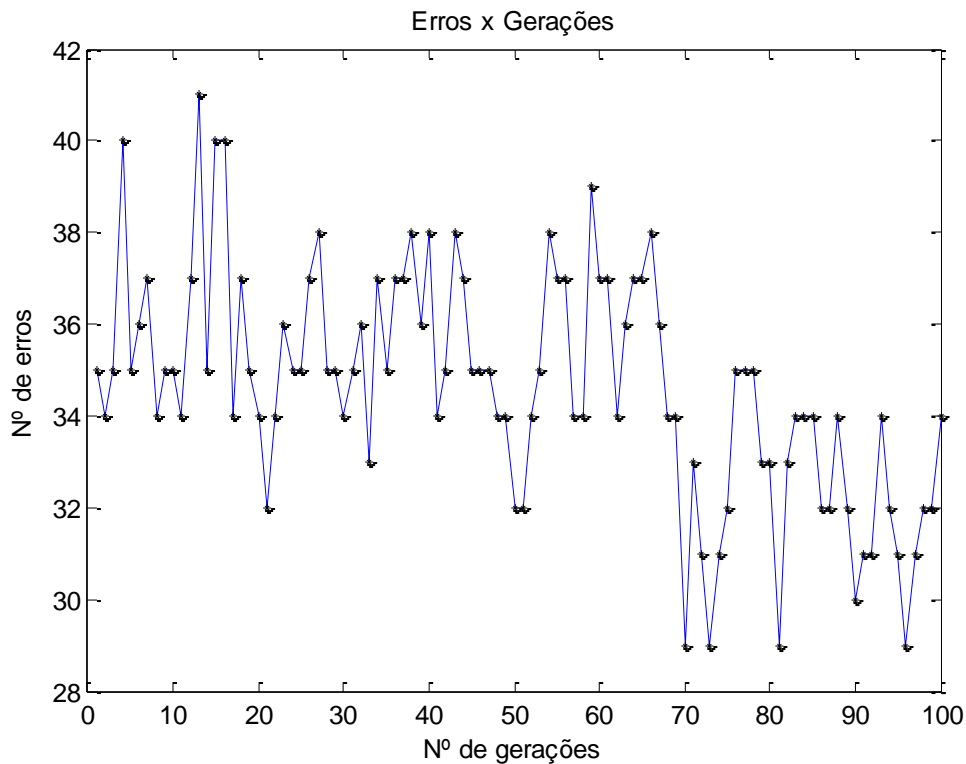


Figura 4.26: Gráficos de erros do AG em função do número de gerações sendo 5 cinco poses para treinamento e 5 cinco poses para o reconhecimento – usando 616 coeficientes DCT, para 40 classes, arquivo ORL.

Sobre esta variação com 616 coeficientes DCT, para 40 classes e dados de 100 gerações, obtiveram-se:

- cinco imagens para o treinamento por classe;
- cinco imagens para o reconhecimento por classe;
- 200 poses analisadas;
- 166 poses reconhecidas corretamente; e
- 34 poses reconhecidas incorretamente.

4.3 - A utilização do Algoritmo Genético usando o arquivo UK

Para o arquivo UK foram utilizados 64, 550, 1100 e 2200 coeficientes DCT para compor as poses de cada classe.

Variações da experiência para o treinamento e reconhecimento pelo AG, arquivo UK

Os treinamentos e os reconhecimentos de faces foram realizados para grupo de 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 e 120 classes tomadas de maneira sequencial para as seguintes variações, considerando pose por classe:

1. Treinou-se com 19 poses (1ª à 19ª pose) e reconheceu-se com uma pose (20ª pose);
2. Treinou-se com 18 poses (1ª à 18ª pose) e reconheceu-se com duas poses (19ª e 20ª poses);
3. Treinou-se com 17 poses (1ª à 17ª pose) e reconheceu-se com três poses (18ª, 19ª e 20ª pose);
4. Treinou-se com 16 poses (1ª à 16ª pose) e reconheceu-se com quatro poses (17ª, 18ª, 19ª e 20ª poses).
5. Treinou-se com 15 poses (1ª à 15ª pose) e reconheceu-se com cinco poses (16ª, 17ª, 18ª, 19ª e 20ª pose);
6. Treinou-se com 14 poses (1ª à 14ª pose) e reconheceu-se com seis poses (15ª, 16ª, 17ª, 18ª, 19ª e 20ª pose);
7. Treinou-se com 13 poses (1ª à 13ª pose) e reconheceu-se com sete poses (14ª, 15ª, 16ª, 17ª, 18ª, 19ª e 20ª poses);
8. Treinou-se com 12 poses (1ª à 12ª pose) e reconheceu-se com oito poses (13ª, 14ª, 15ª, 16ª, 17ª, 18ª, 19ª e 20ª poses);
9. Treinou-se com 11 poses (1ª à 11ª pose) e reconheceu-se com nove poses (11ª, 12ª, 13ª, 14ª, 15ª, 16ª, 17ª, 18ª, 19ª e 20ª poses); e
10. Treinou-se com 10 poses (1ª à 10ª pose) e reconheceu-se com 10 pose (10ª, 11ª, 12ª, 13ª, 14ª, 15ª, 16ª, 17ª, 18ª, 19ª e 20ª poses).

4.3.1- Quanto a utilização de toda a imagem (pose) UK

A seção 3.3.1.2.1, mostra como se processa a retirada dos coeficientes DCT considerando a pose toda.

Foi realizada a operação para todas as poses de uma classe, resultando em um conjunto de 20 vetores de 64 elementos. Esse conjunto, mostrado na Figura 4.27, simboliza a matriz utilizada pelo AG para definir as características de uma classe, necessárias ao reconhecimento de uma imagem, onde $Y_{1,1,1}$, $Y_{1,1,2}$, $Y_{1,1,3}$, $Y_{1,1,4}$, ..., $Y_{1,1,64}$

significam os 64 coeficientes da primeira pose G da classe F_1 . Com este raciocínio chega-se aos vinte vetores da primeira classe. Este procedimento foi repetido para todas as demais classes em um total de 120.

$G_{1,1}$	$Y_{1,1,1}$	$Y_{1,1,2}$	$Y_{1,1,3}$	$Y_{1,1,4}$	$Y_{1,1,64}$
$G_{2,1}$	$Y_{2,1,1}$	$Y_{2,1,2}$	$Y_{2,1,3}$	$Y_{2,1,4}$	$Y_{2,1,64}$
$G_{3,1}$	$Y_{3,1,1}$	$Y_{3,1,2}$	$Y_{3,1,3}$	$Y_{3,1,4}$	$Y_{3,1,64}$
	\vdots	\vdots	\vdots	\vdots	...	\vdots
$G_{20,1}$	$Y_{20,1,1}$	$Y_{20,1,2}$	$Y_{20,1,3}$	$Y_{20,1,4}$	$Y_{20,1,64}$

Figura 4.27: Representação de uma classe por coeficientes DCT com 20 poses. Cada pose é um vetor com 64 elementos, arquivo UK.

Assim, ao utilizar toda a pose, a base de dados para o treinamento das características foi composta de 120 matrizes de coeficientes DCT com tamanho de 20 linhas por 64 colunas.

A Tabela 4.5 mostra o resultado do reconhecimento de face utilizando 64 coeficientes DCT para representar cada pose. Os valores contidos nessa tabela são aqueles obtidos nos testes para 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 e 120 classes. As variáveis são: VAR são as variações da experiência para o treinamento e o reconhecimento, NPT é o número poses para o treinamento, NPR é o número de poses para o reconhecimento de faces. Ainda os parâmetros AC que são a taxa de acerto em percentual, TT, é o tempo de treinamento em segundos, e TR é o tempo de reconhecimento em segundos.

Nas Figuras 4.28 até a 4.31, são mostrados os gráficos resultantes desta experiência.

Tabela 4.5: Resultados obtidos do reconhecimento de face utilizando 64 coeficientes DCT da pose toda, arquivo UK.

V A R	N P T	N P R	Parâmetros	Classes												
				5	10	20	30	40	50	60	70	80	90	100	110	120
1	19	1	AC (%)	100	100	100	100	100	100	100	100	100	100	100	100	100
			TT (s)	68,639	90,6864	175,9685	273,9836	324,4064	433,2982	762,12	813,32	947,48	1045,52	1154,78	1287,84	1506,95
			TR (s)	0,0139	0,0148	0,0176	0,1349	0,2099	0,2404	0,698	0,792	1,131	1,372	1,608	1,8202	2,627
2	18	2	AC (%)	100	100	100	100	100	100	100	99,83	100	100	99,7195	99,2089	100
			TT (s)	58,737	89,8892	207,9598	287,9125	419,1756	493,8258	688,732	805,48	906,49	1010,44	1097,87	1276,69	1368,05
			TR (s)	0,0212	0,0369	0,1347	0,1798	0,5415	0,6585	1,339	1,706	2,350	2,955	3,649	4,448	5,066
3	17	3	AC (%)	100	100	100	100	99,6895	99,4492	100	99,6394	99,6394	99,6695	99,7796	99,9198	99,72
			TT (s)	53,631	103,7560	200,6470	313,1257	403,0635	498,5101	625,00	704,13	815,24	897,90	982,66	1122,95	1228,95
			TR (s)	0,0271	0,0610	0,2153	0,5694	0,8575	1,0335	1,951	2,7315	3,128	3,765	4,776	6,762	7,363
4	16	4	AC (%)	100	100	100	99,4492	98,8685	99,4492	100	99,8196	99,7996	99,8397	99,8196	99,8897	99,79
			TT (s)	46,602	89,6221	185,4146	291,8324	371,8047	444,8939	556,77	655,49	761,43	822,12	904,76	1039,33	1102,09
			TR (s)	0,0295	0,0295	0,0295	0,0293	0,0293	0,0293	2,421	3,107	4,302	5,334	6,103	8,045	9,4559
5	15	5	AC (%)	100	100	100	99,4492	99,4492	99,3190	99,33	99,5393	99,7996	99,4492	99,7496	99,7496	99,66
			TT (s)	41,388	85,4286	161,1275	258,5318	326,4120	408,1350	494,175	556,30	653,77	708,58	786,50	865,45	978,511
			TR (s)	0,0354	0,1166	0,4666	1,1249	1,7930	2,0012	2,909	4,021	5,651	6,505	8,0725	9,702	11,341
6	14	6	AC (%)	100	100	100	99,3190	98,8685	99,4492	99,16	99,4892	99,4892	9,4513	99,4492	99,4492	99,44
			TT (s)	36,58	68,9178	146,2329	199,9680	291,8277	326,3042	437,001	493,37	588,89	633,23	736,03	797,67	966,41
			TR (s)	0,0387	0,1427	0,5604	1,4929	2,2028	2,2701	3,347	4,614	5,804	7,4681	9,230	11,395	13,130
7	13	7	AC (%)	100	100	100	99,6895	99,2590	99,2590	99,28	99,4993	99,4993	99,1588	99,2089	99,2089	99,285
			TT (s)	32,071	66,1569	104,1658	199,0195	228,0601	310,1308	382,142	500,21	581,81	636,00	721,38	778,72	759,06
			TR (s)	0,048	0,1762	0,6914	1,6791	2,8074	2,9987	3,744	5,039	6,530	8,243	10,08	12,280	14,573
8	12	8	AC (%)	100	100	100	99,4492	99,7996	99,2590	99,66	98,8685	99,7996	99,7996	99,4892	99,4892	99,58
			TT (s)	28,049	39,6410	112,4325	159,5448	225,6902	278,7243	332,189	388,10	453,74	498,22	566,98	582,22	658,63
			TR (s)	0,055	0,2035	0,8187	1,9205	3,4376	3,3210	4,125	5,467	7,309	9,291	11,40	14,563	15,97
9	11	9	AC (%)	100	100	100	97,1464	99,7195	98,9987	98,88	9,6394	99,2490	99,2490	99,2089	99,2089	99,35
			TT (s)	24,019	50,2333	74,0272	150,8936	195,5767	239,7173	284,681	322,97	372,12	412,49	482,00	503,99	567,099
			TR (s)	0,056	0,2114	0,8346	1,8573	3,3260	3,5459	4,435	6,042	7,915	9,839	11,781	14,729	17,289
10	10	10	AC (%)	100	100	98,6182	99,4492	99,6194	99,6194	98,83	99,3691	99,3691	99,4492	99,4192	99,4192	99,08
			TT (s)	20,369	38,1197	81,2294	96,5465	163,1381	202,6322	259,067	287,95	348,71	392,85	440,88	482,61	481,137
			TR (s)	0,0584	0,4806	0,8592	2,5302	3,4334	4,4718	4,699	7,204	8,369	11,612	13,060	16,483	18,299

A Figura 4.28 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Percebe-se que para o reconhecimento com uma pose a taxa de acerto é de 100% para 5 até 120 classes. Nota-se também a queda na taxa de acerto na classe 20 e classe 60 para reconhecimento com 10 poses. A oscilação das curvas valores das taxas de acertos se explica pelo fato de as vezes ocorrer saturação no processamento do algoritmo e ainda pela diferença de centralização e de iluminação em certas poses dentro das respectivas classes.

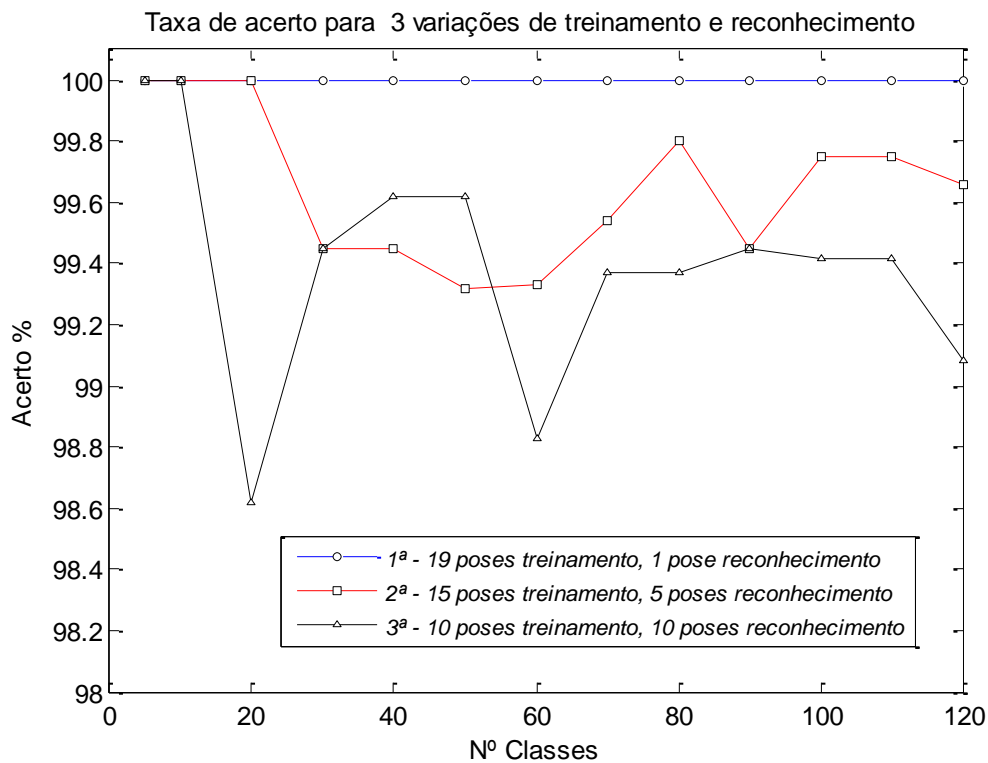


Figura 4.28: Taxa de acerto considerando 64 coeficientes por pose, arquivo UK.

A Figura 4.29 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente linearmente progressivo com o aumento do número de classes. Nota-se que quanto maior o número de poses utilizados no treinamento mais acentuado é o crescimento dos valores do tempo de treinamento, 1506,95 s para treinamento com 19 poses.

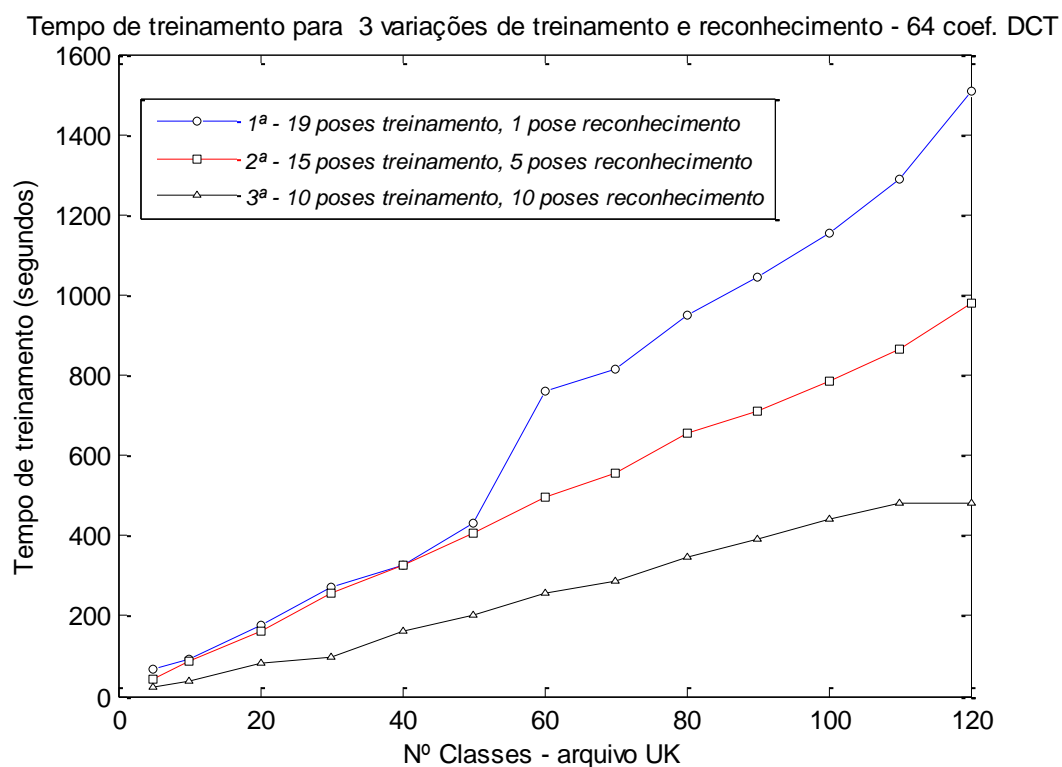


Figura 4.29: Tempo de treinamento considerando 64 coeficientes por pose, arquivo UK.

A Figura 4,30 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes do arquivo UK. Nota-se que mesmo com aumento do número de poses utilizados no reconhecimento o AG efetua o reconhecimento em um tempo bastante curto, 18,299 s para reconhecimento com 10 poses.

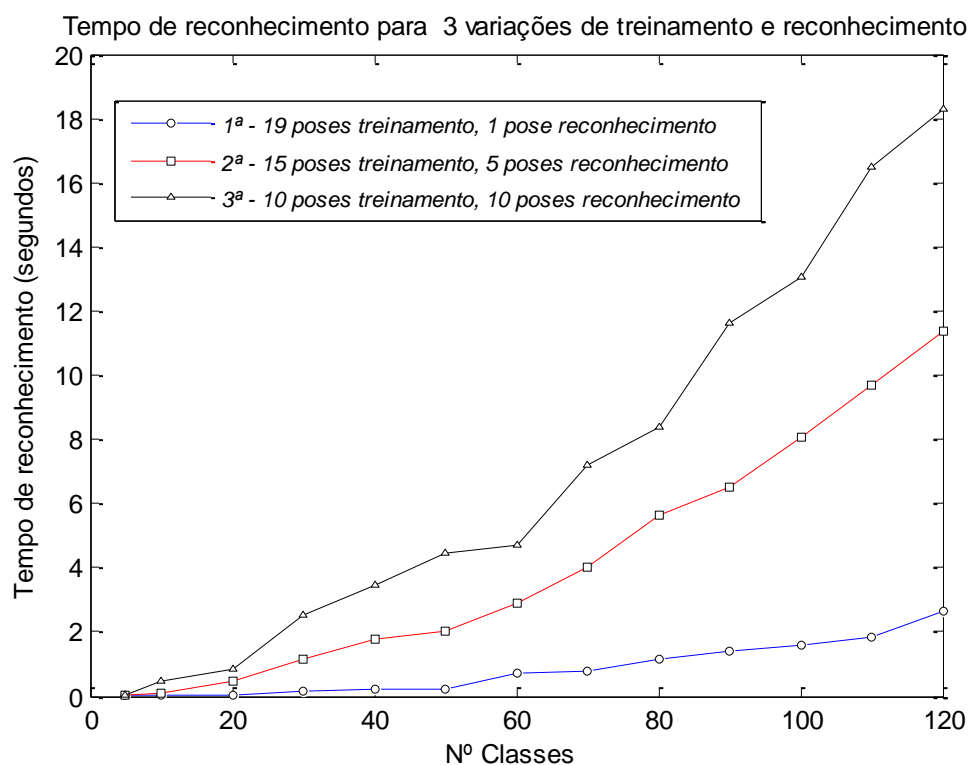


Figura 4.30: Tempo de reconhecimento do AG considerando 64 coeficientes por pose, arquivo UK.

A Figura 4.31 mostra o gráfico resultante dessa experiência referente ao número de erros ocorridos no arquivo UK, considerando 10 poses para o treinamento e 10 poses para o reconhecimento e total de 120 classes. Nesse teste utilizou-se 1200 poses (10 de cada classe) para o reconhecimento. O gráfico mostra a variação de erros na execução de 100 gerações para 120 classes. Neste reconhecimento, das 1200 poses, para 11 existe erro no reconhecimento. Por isso a taxa de acerto de 99,08 % mostrada na Tabela 4.5. Nota-se que o número de erros oscila em função do número de geração processada e somente neste caso, a taxa de erro foi menor na trigésima geração. Se alterar o número de classe, o número de poses para o treinamento e o número de poses para o reconhecimento, os valores da taxa de erro alteram também. Por isso, o sistema tem que fixar um número de geração que seja capaz de fornecer uma taxa alta de acerto mesmo com a manipulação das referidas variáveis.

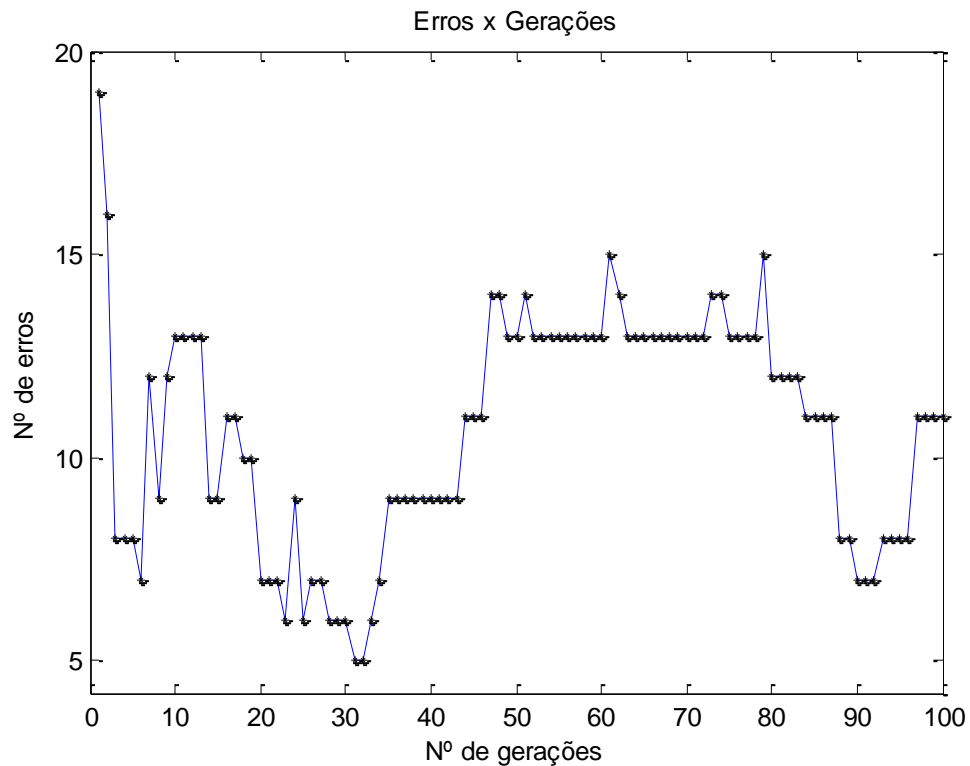


Figura 4.31: Gráficos de erros do AG em função do número de gerações sendo 10 poses para treinamento e 10 poses para o reconhecimento – usando 64 coeficientes DCT, para 120 classes, arquivo UK.

Sobre esta variação com 64 coeficientes DCT, para 120 classes e dados de 100 gerações, obtiveram-se:

- 10 imagens para o treinamento por classe;
- 10 imagens para o reconhecimento por classe;
- 1200 poses analisadas;
- 1189 poses reconhecidas corretamente; e
- 11 poses reconhecidas incorretamente.

4.3.2- Quanto utilização de subimagens (subposes) UK

O preparo das subimagens foi mostrado na seção 3.3.1.2.2.

4.3.2.1- Utilizando 550 coeficientes DCT, com um de cada subpose UK

Passo 1: Considerando a Figura 3.5, tomou-se primeiramente a pose $G_{1,1}$. Desta pose, representada na Figura 3.13, retirou-se a subpose Y_1 , que, por sua vez, está mostrada na Figura 4.32 com destaque para o coeficiente mais significativo. Pode-se observar que essa subpose é composta de 64 coeficientes.

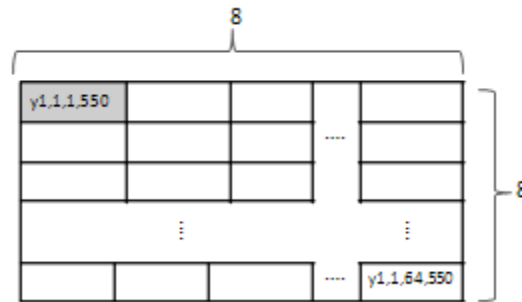


Figura 4.32: Representação da subpose Y_1 , base para o vetor característico de 550 coeficientes DCT – arquivo UK.

Passo 2: Extraíu-se da subpose Y_1 o coeficiente mais significativo $y_{1,1,1,550}$, colocando-o em um vetor. Depois na Figura 3.8 extraíu-se da pose Y_2 o coeficiente mais significativo $y_{1,2,1,550}$ e colocou-o no mesmo vetor e assim, sucessivamente, até extrair o coeficiente $y_{1,550,1,550}$ da subpose Y_{550} colocando-o nesse vetor é mostrado completo na Figura 4.33. Desta forma, a pose $G_{1,1}$ (primeira pose da classe F_1) foi representada por um vetor de 550 elementos.

Para melhor entendimento, segue a descrição dos coeficientes:

$y_{1,1,1,550} - y_1$ refere-se à pose 1. O número 1 seguinte quer dizer subpose 1. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$y_{1,2,1,550} - y_1$ refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo

$y_{1,550,1,550} - y_1$ refere-se à pose 1. O número 550 seguinte quer dizer subpose 550. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

O índice 550 no final de cada símbolo de coeficiente refere-se ao caso em estudo desta seção e também é o total de elementos do vetor que representa uma pose.

O símbolo $G_{1,1}$ representa a pose 1 da classe F_1 .

$G_{1,1}$	y1,1,1,550	y1,2,1,550	y1,3,1,550	y1,4,1,550	y1,550,1,550
-----------	------------	------------	------------	------------	-------	--------------

Figura 4.33: Vetor representativo da primeira pose G com de 550 coeficientes DCT mais significativo de cada subpose - arquivo UK.

Passo 3: Repetiram-se os passos 1 e 2 para as poses $G_{2,1}$ a $G_{20,1}$ (Figura 3.5), obtendo-se o conjunto de 20 vetores que formam a matriz de coeficientes DCT representativa das dez poses da classe F_1 . Esse conjunto é mostrado na Figura 4.34.

$G_{1,1}$	y1,1,1,550	y1,2,1,550	y1,3,1,550	y1,550,1,550
$G_{2,1}$	y2,1,1,550	y2,2,1,550	y2,3,1,550	y2,550,1,550
$G_{3,1}$	y3,1,1,550	y3,2,1,550	y3,3,1,550	y3,550,1,550
	⋮	⋮	⋮	...	⋮
$G_{20,1}$	y20,1,1,550	y20,2,1,550	y20,3,1,550	y20,550,1,550

Figura 4.34: Conjunto de vetores com 550 coeficientes DCT cada um, representativo da classe F_1 - arquivo UK.

Passo 4: Repetiram-se os passos desta seção para achar o conjunto de vetores representativos das classes F_2 , F_3 até a classe F_{120} .

A Tabela 4.6 mostra o resultado do reconhecimento de face utilizando 550 coeficientes DCT para representar cada pose. Os valores contidos nessa tabela são aqueles obtidos nos testes para 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 e 120 classes. As variáveis são: VAR são as variações da experiência para o treinamento e o reconhecimento, NPT é o número poses para o treinamento, NPR é o número de poses para o reconhecimento de faces. Ainda os parâmetros AC que são a taxa de acerto em percentual, TT , é o tempo de treinamento em segundos, e TR é o tempo de reconhecimento em segundos.

Nas Figuras 4.35 até a 4.38, são mostrados os gráficos resultantes desta experiência.

Tabela 4.6: Resultados obtidos do reconhecimento de face considerando 550 coeficientes DCT por pose, arquivo UK.

V A R	N P T	N P R	Parâmetros	Classes												
				5	10	20	30	40	50	60	70	80	90	100	110	120
1	19	1	AC (%)	100	100	100	100	100	100	100	100	100	100	100	100	100
			TT (s)	150,36	238,303	471,02	725,04	1001,59	1235,60	1325,87	1558,97	1800,58	2021,35	2367,88	2756,41	3143,022
			TR (s)	0,0351	0,0795	0,291	0,646	0,894	1,319	1,327	1,398	1,4014	1,482	5,0483	6,475	7,306
2	18	2	AC (%)	100	100	100	100	100	100	100	100	100	100	100	100	100
			TT (s)	108,52	201,58	420,15	654,25	911,61	1145,32	1297,73	1378,26	1625,32	1829,64	2013,87	2252,21	2920,37
			TR (s)	0,0429	0,08421	0,548	0,782	1,130	3,241	3,509	4,125	4,881	7,256	8,561	11,424	14,196
3	17	3	AC (%)	100	100	98,33	100	100	100	100	100	100	99,66	99,66	99,70	99,72
			TT (s)	92,35	194,80	391,34	581,30	756,25	1114,67	1194,21	1235,97	1521,32	1745,62	1997,53	2154,23	2722,72
			TR (s)	0,0537	0,104	0,827	1,907	2,451	5,007	6,847	8,547	10,549	12,547	14,457	15,256	20,795
4	16	4	AC (%)	100	100	100	99,75	99,37	99,16	99,16	99,52	99,62	99,60	99,60	99,77	99,79
			TT (s)	85,52	179,54	347,28	554,19	684,82	945,28	1150,36	1192,28	1365,91	1632,85	1789,46	1924,96	2103,48
			TR (s)	0,0734	0,245	1,152	2,203	3,954	6,294	8,912	10,673	13,045	15,254	18,265	22,695	26,98
5	15	5	AC (%)	100	100	100	99,33	99,75	99,20	99,42	99,42	99,52	99,58	99,60	99,33	99,33
			TT (s)	73,58	157,16	316,82	535,05	591,57	768,94	868,57	987,543	1287,58	1543,21	1656,56	1789,45	1865,22
			TR (s)	0,0872	0,3368	1,297	2,703	4,257	7,417	10,924	13,872	16,872	19,972	23,181	26,451	33,059
6	14	6	AC (%)	100	100	100	100	99,75	99,14	99,16	99,16	99,25	99,25	99,50	99,39	99,33
			TT (s)	69,80	135,94	285,23	451,29	537,02	631,54	798,25	901,85	1058,74	1349,83	1458,94	1598,03	1783,28
			TR (s)	0,1056	0,3951	1,523	3,015	5,593	10,837	13,456	16,873	20,458	25,687	31,457	32,627	39,476
7	13	7	AC (%)	100	100	100	100	98,0	99,14	99,38	99,38	99,04	99,40	99,42	99,31	99,28
			TT (s)	60,57	123,66	249,33	377,21	486,27	592,419	692,73	819,34	965,27	1121,5	1231,56	1359,57	1453,21
			TR (s)	0,1241	0,4438	1,7383	3,586	6,259	13,757	16,783	19,457	23,864	27,213	30,641	38,459	43,992
8	12	8	AC (%)	100	100	100	99,75	98,43	99,37	99,37	99,46	99,68	99,16	99,33	99,31	99,35
			TT (s)	54,39	118,25	210,58	310,452	416,08	456,91	611,97	702,83	839,17	921,87	1100,265	1206,14	1300,497
			TR (s)	0,1329	0,4957	1,9752	3,978	7,117	16,425	19,713	23,525	26,954	30,496	34,297	41,163	45,297
9	11	9	AC (%)	100	100	99,44	99,25	99,0	99,11	99,36	99,36	99,40	99,44	99,44	99,35	99,35
			TT (s)	45,28	99,93	187,55	267,88	328,46	387,84	500,94	650,37	749,28	834,92	920,13	980,79	1105,61
			TR (s)	0,1452	0,5553	2,107	4,361	7,843	19,699	23,694	26,763	29,764	33,697	37,927	45,297	52,638
10	10	10	AC (%)	100	100	99,33	99,0	98,75	99,11	99,166	98,71	99,25	98,88	99,33	99,36	99
			TT (s)	40,43	87,84	170,28	240,56	286,24	302,64	389,76	471,25	573,16	621,45	701,45	874,98	937,20
			TR (s)	0,1584	1,021	3,567	5,568	8,410	23,104	26,805	30,269	34,681	37,803	41,267	48,81	57,027

A Figura 4.35 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Percebe-se que para o reconhecimento com uma pose a taxa de acerto é de 100% para 5 até 120 classes. Nota-se também a queda na taxa de acerto na classe 40 e classe 80 para reconhecimento com 10 poses. A oscilação das curvas valores das taxas de acertos se explica pelo fato de as vezes ocorrer saturação no processamento do algoritmo e ainda pela diferença de centralização e de iluminação em certas poses dentro das respectivas classes.

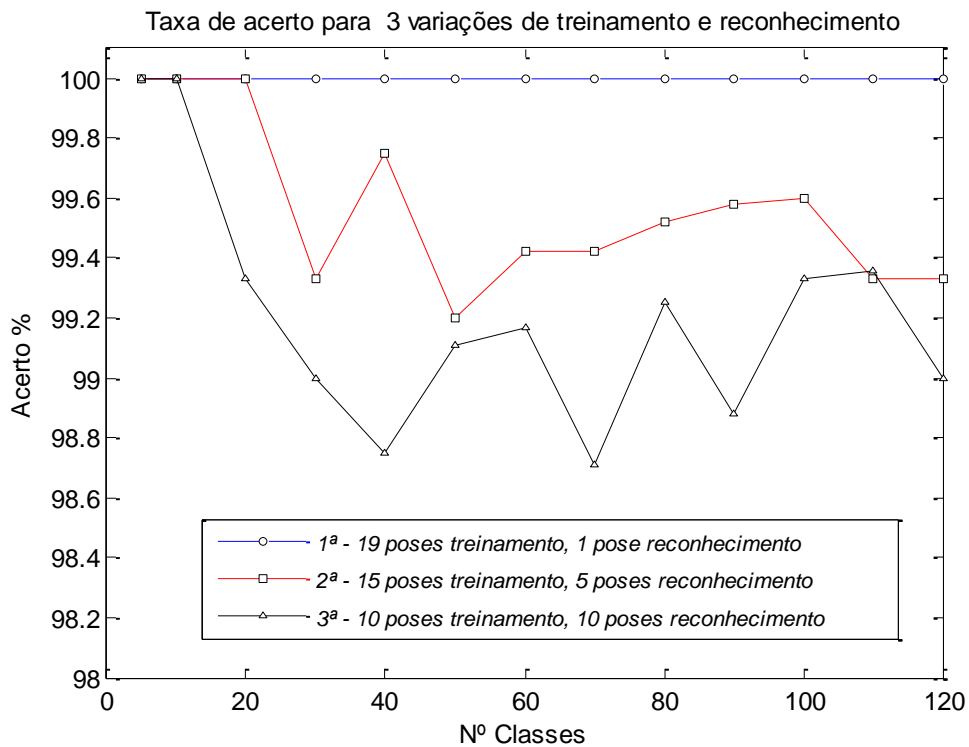


Figura 4.35: Taxa de acerto considerando 550 coeficientes por pose, arquivo UK.

A Figura 4.36 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente progressivo com o aumento do número de classes. Nota-se que quanto maior o número de poses utilizados no treinamento mais acentuado é o crescimento dos valores do tempo de treinamento, 3143,022 s para treinamento com 19 poses.

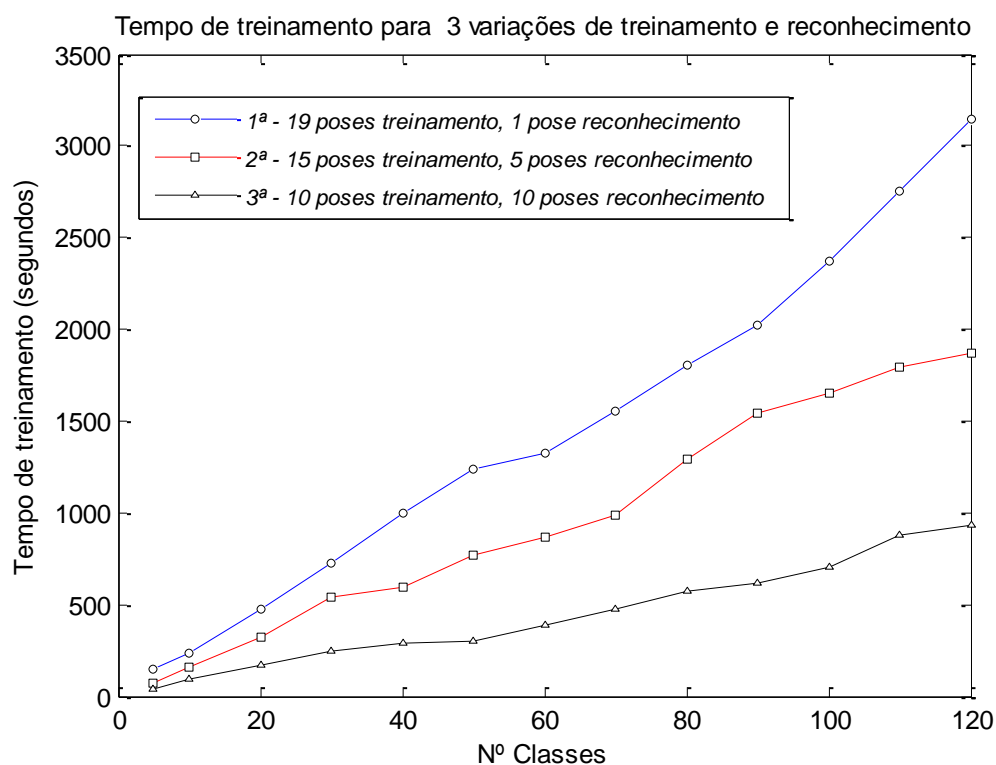


Figura 4.36: Tempo de treinamento considerando 550 coeficientes por pose, arquivo UK.

A Figura 4.37 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Nota-se que mesmo com aumento do número de poses utilizados no reconhecimento o AG efetua o reconhecimento em um tempo bastante curto, 57,027 s para reconhecimento com 10 poses.

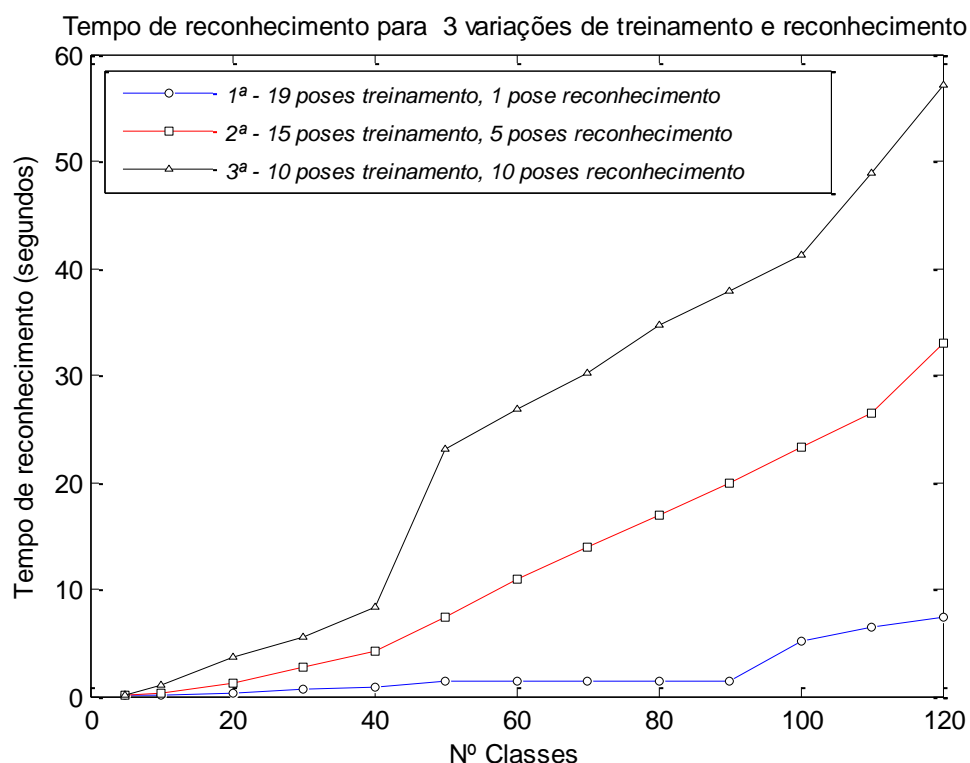


Figura 4.37: Tempo de reconhecimento do AG considerando 550 coeficientes por pose, arquivo UK.

A Figura 4.38 mostra o gráfico resultante dessa experiência referente ao número de erros ocorridos no arquivo UK, considerando 10 poses para o treinamento e 10 poses para o reconhecimento e total de 120 classes. Nesse teste utilizou-se 1200 poses (10 de cada classe) para o reconhecimento. O gráfico mostra a variação de erros na execução de 100 gerações para 120 classes. Neste reconhecimento, das 1200 poses, para 12 existe erro no reconhecimento. Por isso a taxa de acerto de 99 % mostrada na Tabela 4.6. Nota-se que o número de erros oscila em função do número de geração processada e somente neste caso, a taxa de erro foi menor na trigésima geração. Se alterar o número de classe, o número de poses para o treinamento e o número de poses para o reconhecimento, os valores da taxa de erro alteram também. Por isso, o sistema tem que fixar um número de geração que seja capaz de fornecer uma taxa alta de acerto mesmo com a manipulação das referidas variáveis.

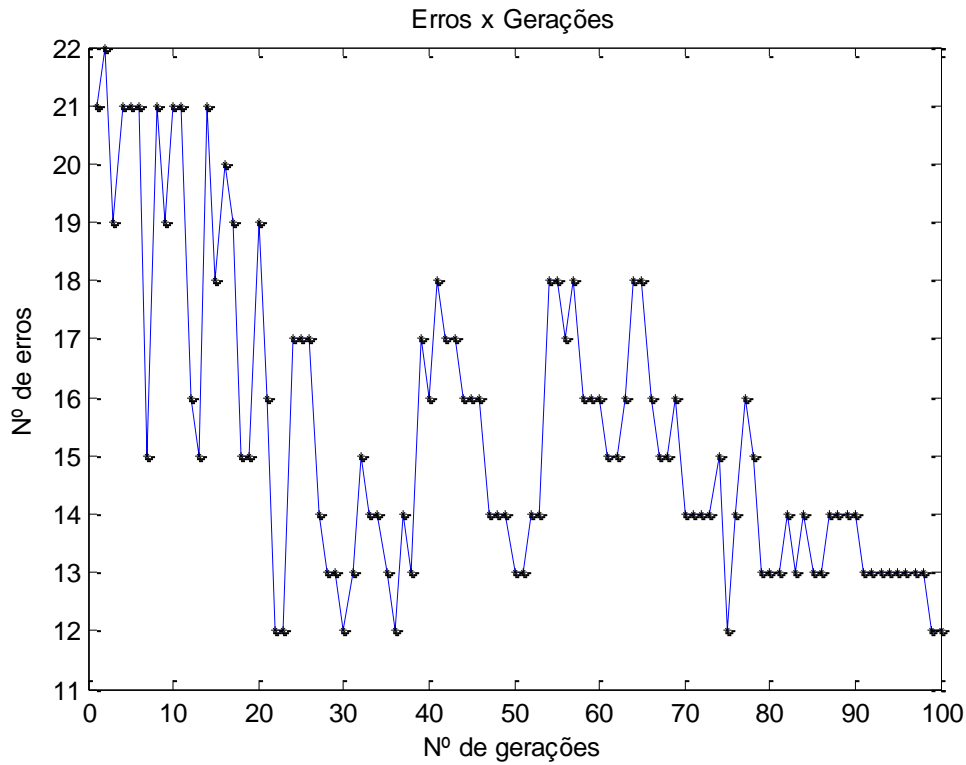


Figura 4.38: Gráfico de erros do AG em função do número de gerações onde foram utilizados 10 poses para o treinamento e 10 poses para o reconhecimento – usando 550 coeficientes DCT, para 120 classes, arquivo UK.

Sobre esta variação com 550 coeficientes DCT, para 120 classes e dados de 100 gerações, obtiveram-se:

- 10 imagens para o treinamento por classe;
- 10 imagens para o reconhecimento por classe;
- 1200 poses analisadas;
- 1188 poses reconhecidas corretamente; e
- 12 poses reconhecidas incorretamente.

4.3.2.2- Utilizando 1100 coeficientes DCT, com dois de cada subpose UK

Passo 1: Considerando a Figura 3.5, tomou-se primeiramente a pose $G_{1,1}$. Desta pose, representada na Figura 3.13, retirou-se a subpose Y_1 , ilustrada na Figura 4.39 com destaque para os dois coeficientes mais significativos.

$G_{1,1}$	y1,1,1,1100	y1,1,2,1100	y1,2,1,1100	y1,2,2,1100	y1,550,1,1100	y1,550,2,1100
-----------	-------------	-------------	-------------	-------------	-------	---------------	---------------

Figura 4.40: Vetor representativo da primeira pose G das 1100 coeficientes DCT mais significativos de cada subpose- arquivo UK.

Passo 3: Repetiram-se os passos 1 e 2 para as poses $G_{2,1}$ a $G_{20,1}$ (Figura 3.5), obtendo-se o conjunto de 20 vetores que formam a matriz de coeficientes DCT representativa das 10 poses da classe F_1 . Esse conjunto é mostrado na Figura 4.41.

$G_{1,1}$	y1,1,1,1100	y1,1,2,1100	y1,2,1,1100	y1,2,2,1100	y1,550,1,1100	y1,550,2,1100
$G_{2,1}$	y2,1,1,1100	y2,1,2,1100	y2,2,1,1100	y2,2,2,1100	y2,550,1,1100	y2,550,2,1100
$G_{3,1}$	y3,1,1,1100	y3,1,2,1100	y3,2,1,1100	y3,2,2,1100	y3,550,1,1100	y3,550,2,1100
	⋮	⋮	⋮	⋮	...	⋮	⋮
$G_{20,1}$	y10,1,1,1100	y10,1,2,1100	y10,2,1,1100	y10,2,2,1100	y10,550,1,1100	x10,550,2,1100

Figura 4.41: Conjunto de vetores com 1100 coeficientes DCT cada um, representativo da classe F_1 - arquivo UK.

Passo 4: Repetiram-se os passos desta seção para achar o conjunto de vetores representativos das classes F_2 , F_3 até a classe F_{120} .

A Tabela 4.7 mostra o resultado do reconhecimento de face utilizando 1100 coeficientes DCT para representar cada pose, arquivo UK. Os valores contidos nessa tabela são aqueles obtidos nos testes para 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 e 120 classes. As variáveis são: VAR são as variações da experiência para o treinamento e o reconhecimento, NPT é o número poses para o treinamento, NPR é o número de poses para o reconhecimento de faces. Ainda os parâmetros AC que são a taxa de acerto em percentual, TT , é o tempo de treinamento em segundos, e TR é o tempo de reconhecimento em segundos.

Nas Figuras 4.42 a 4.45, são mostrados os gráficos resultantes dessa experiência.

Tabela 4.7: Resultados obtidos do reconhecimento de face considerando 1100 coeficientes DCT por pose, arquivo UK.

V A R	N P T	N P R	Parâ- metros	Classes												
				5	10	20	30	40	50	60	70	80	90	100	110	120
1	19	1	AC (%)	100	100	100	100	100	100	100	100	100	100	100	100	100
			TT (s)	287,29	379,57	736,52	1146,764	1357,81	1813,579	2367,95	2527,04	2943,87	3248,49	3587,97	4001,39	4421,24
			TR (s)	0,123	0,1310	0,156	1,194	1,857	2,127	3,587	4,071	5,812	7,054	8,266	9,354	11,959
2	18	2	AC (%)	100	100	100	100	100	100	99,16	99,72	100	100	99,60	99,09	99,16
			TT (s)	194,45	297,58	688,45	953,14	1387,689	1634,82	2017,176	2359,11	2654,96	2959,41	3215,46	3739,21	4251,37
			TR (s)	0,149	0,259	0,947	1,264	3,806	4,628	5,948	7,582	10,443	13,127	16,211	19,76	21,547
3	17	3	AC (%)	100	100	100	100	99,57	99,33	99,40	99,52	99,52	99,55	99,66	99,80	100
			TT (s)	159,73	309,02	597,59	932,587	1200,45	1484,72	1864,23	2100,27	2431,67	2678,22	2931,04	3349,51	4002,73
			TR (s)	0,159	0,3581	1,263	3,341	5,031	6,064	8,444	11,822	13,541	16,297	20,673	29,266	35,353
4	16	4	AC (%)	100	100	100	99,33	98,75	99,33	99,58	99,70	99,68	99,72	99,70	99,77	99,58
			TT (s)	137,78	264,97	548,18	862,81	1099,25	1315,34	1611,847	1897,65	2204,34	2380,047	2619,27	3008,87	3225,33
			TR (s)	0,1697	0,457	1,8111	4,159	7,212	7,536	11,337	14,551	20,146	24,98	28,581	37,677	44,312
5	15	5	AC (%)	100	100	100	99,33	99,33	99,20	99,58	99,42	99,68	99,33	99,63	99,63	99,63
			TT (s)	120,67	249,07	469,78	753,77	951,68	1189,95	1487,99	1675,056	1968,55	2133,58	2367,44	2605,92	2841,38
			TR (s)	0,171	0,5632	2,254	5,434	8,661	9,667	13,677	18,909	26,570	30,585	37,954	45,617	54,235
6	14	6	AC (%)	100	100	100	99,20	98,75	99,33	99,44	99,37	99,37	99,44	99,33	99,33	99,16
			TT (s)	112,13	211,26	448,26	612,98	894,566	1000,25	1297,26	1464,58	1748,15	1879,77	2184,95	2367,92	2560,59
			TR (s)	0,1812	0,668	2,624	6,99	10,314	10,629	16,244	22,395	28,17	36,245	44,797	55,325	63,868
7	13	7	AC (%)	100	100	100	99,57	99,14	99,14	99,44	99,38	99,38	99,04	99,09	99,09	99,16
			TT (s)	96,22	198,48	312,52	597,102	684,23	930,46	1000,11	1309,11	1522,66	1664,49	1887,94	2037,988	2365,41
			TR (s)	0,205	0,7525	2,953	7,171	11,99	12,807	18,663	25,119	32,551	41,091	50,254	61,223	71,996
8	12	8	AC (%)	100	100	100	99,33	99,68	99,14	98,75	98,75	99,68	99,68	99,37	99,37	99,47
			TT (s)	87,527	123,70	350,85	497,86	704,267	869,76	984,49	1150,20	1344,73	1476,54	1680,33	1725,48	1958,24
			TR (s)	0,2251	0,833	3,3509	7,86	14,069	13,592	20,554	27,244	36,421	46,298	56,844	72,561	80,508
9	11	9	AC (%)	100	100	100	97,03	99,60	98,88	99,52	99,52	99,13	99,13	99,09	99,09	99,16
			TT (s)	72,66	151,96	223,94	456,469	591,64	725,17	873,51	990,989	1141,82	1265,689	1478,96	1546,43	1709,73
			TR (s)	0,247	0,9326	3,681	8,192	14,67	15,640	22,599	30,789	40,333	50,14	60,033	75,022	87,256
10	10	10	AC (%)	100	100	98,5	99,33	99,50	99,50	99,0	99,25	99,25	99,33	99,30	99,30	99,33
			TT (s)	65,66	122,88	261,84	311,22	525,88	653,19	733,52	815,31	987,337	1112,31	1248,31	1366,45	1461,908
			TR (s)	0,2689	2,213	3,956	11,65	15,809	20,59	24,364	37,357	43,397	60,222	67,715	85,470	96,185

A Figura 4.42 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Percebe-se que para o reconhecimento com uma pose a taxa de acerto é de 100% para 5 até 120 classes. Nota-se também a queda na taxa de acerto na classe 20 e classe 60 para reconhecimento com 10 poses. A oscilação das curvas valores das taxas de acertos se explica pelo fato de as vezes ocorrer saturação no processamento do algoritmo e ainda pela diferença de centralização e de iluminação em certas poses dentro das respectivas classes.

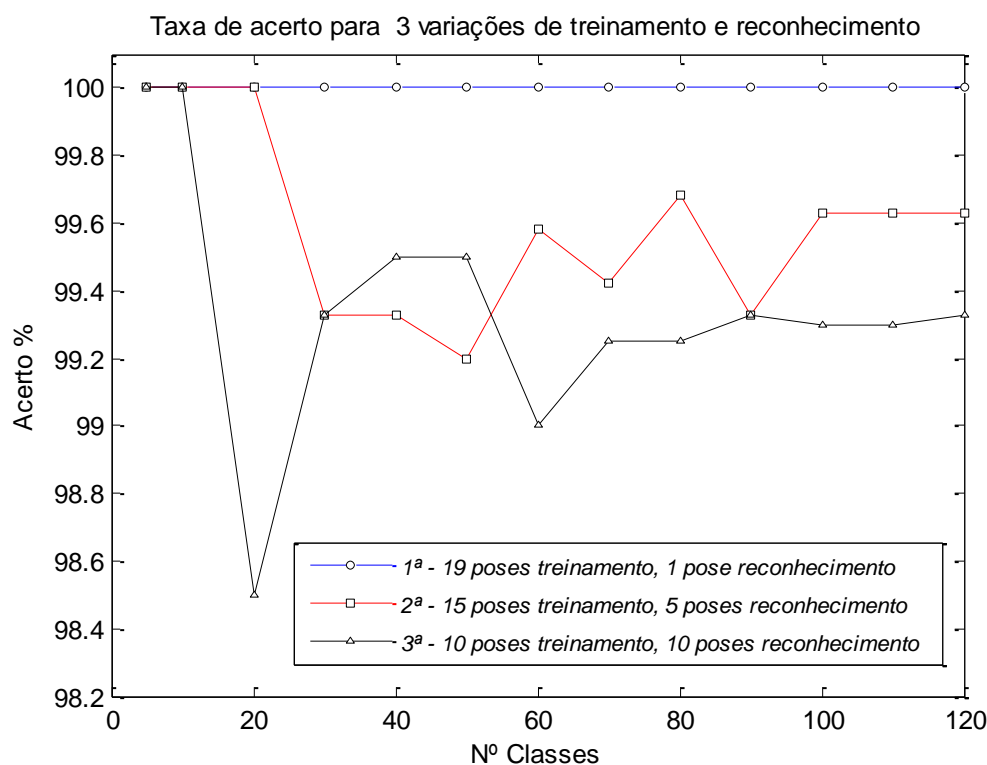


Figura 4.42: Taxa de acerto considerando 1100 coeficientes por pose, arquivo UK.

A Figura 4.43 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente linearmente progressivo com o aumento do número de classes. Nota-se que quanto maior o número de poses utilizados no treinamento mais acentuado é o crescimento dos valores do tempo de treinamento, 4421,24 s para treinamento com 19 poses.

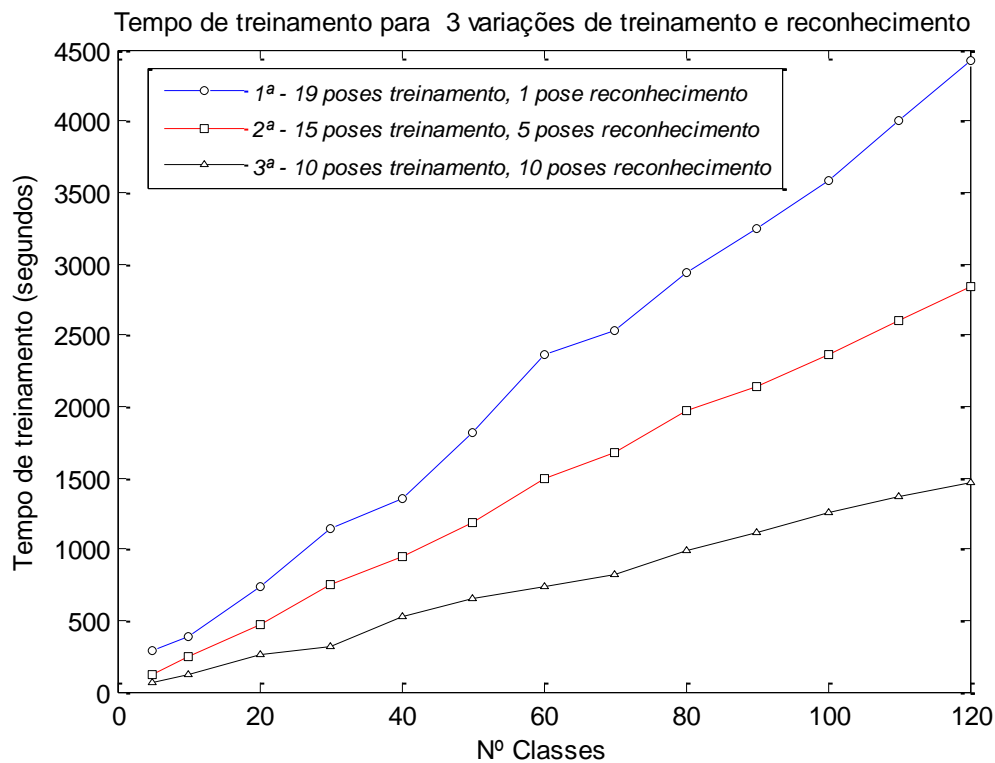


Figura 4.43: Tempo de treinamento considerando 1100 coeficientes por pose, arquivo UK.

A Figura 4.44 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Nota-se que mesmo com aumento do número de poses utilizados no reconhecimento o AG efetua o reconhecimento em um tempo bastante curto, 96,185 s para reconhecimento com 10 poses.

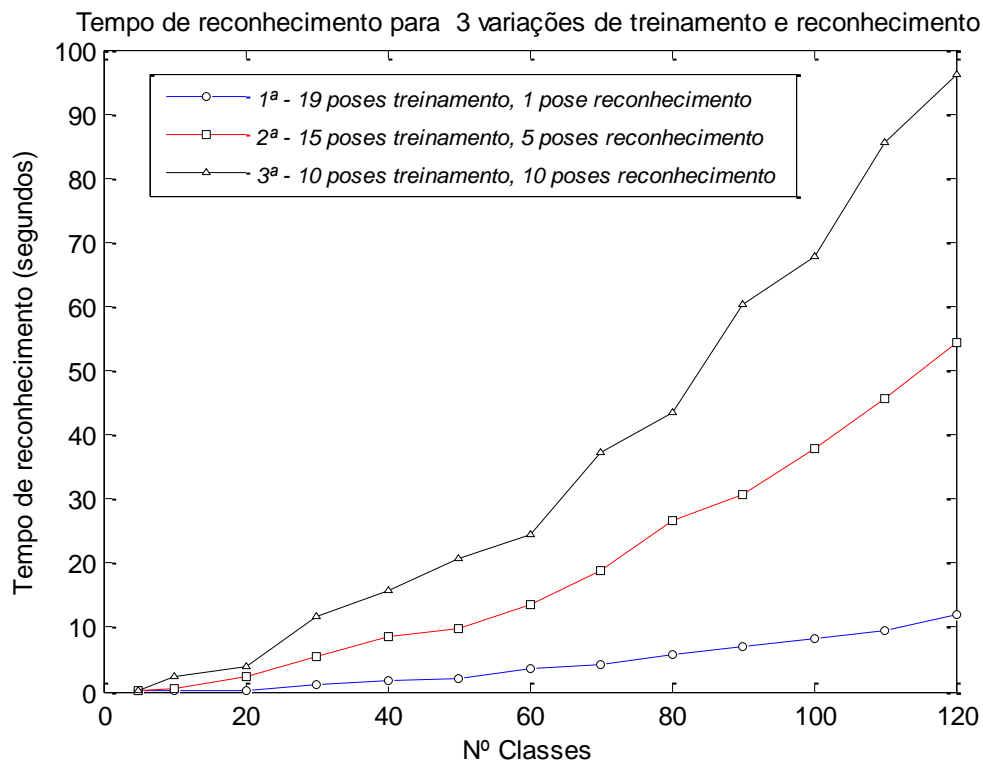


Figura 4.44: Tempo de reconhecimento do AG considerando 1100 coeficientes por pose, arquivo UK.

A Figura 4.45 mostra o gráfico resultante dessa experiência referente ao número de erro ocorrido no arquivo UK, considerando 10 poses para o treinamento e 10 poses para o reconhecimento e total de 120 classes. Nesse teste utilizou-se 1200 poses (10 de cada classe) para o reconhecimento. O gráfico mostra a variação de erros na execução de 100 gerações para 120 classes. Neste reconhecimento, das 1200 poses, para 8 existe erro no reconhecimento. Por isso a taxa de acerto de 99,33 % mostrada na Tabela 4.7. Nota-se que o número de erros oscila em função do número de geração processada. A medida que o número de geração a taxa de erro vai diminuindo. Neste caso devido o grande número de coeficientes representativos de cada pose possibilitou a alta taxa de acerto.

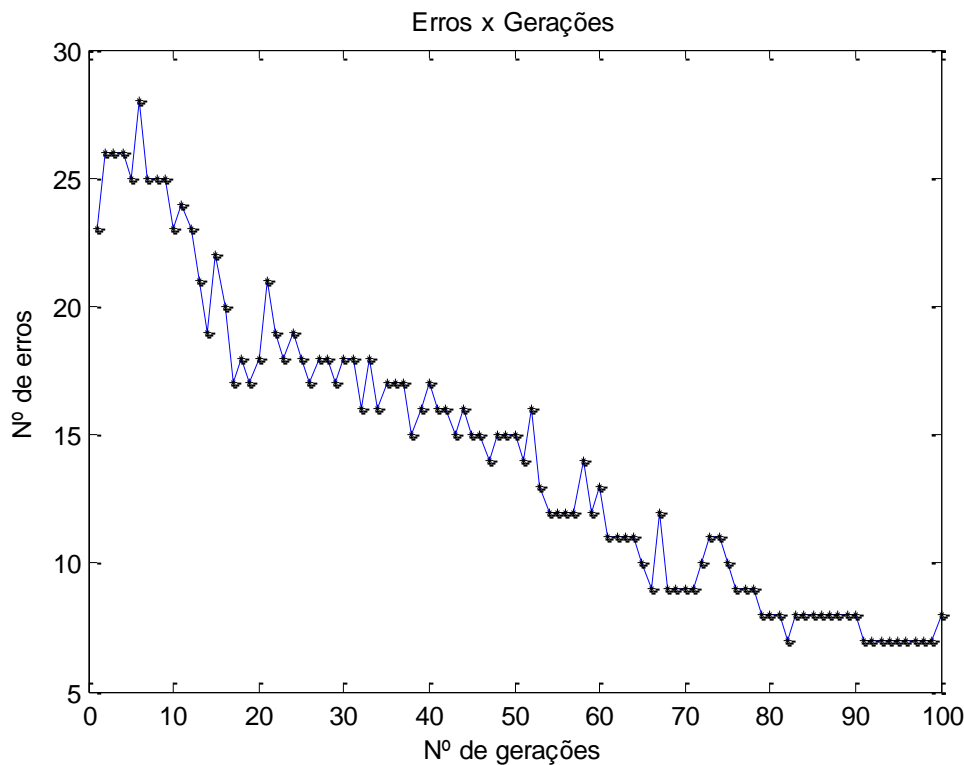


Figura 4.45: Gráfico de erros do AG em função do número de gerações onde foram utilizados 10 poses para treinamento e 10 poses para o reconhecimento – usando 1100 coeficientes DCT, para 120 classes, arquivo UK.

Sobre esta variação com 1100 coeficientes DCT, para 120 classes e dados de 100 gerações, obtiveram-se:

- 10 imagens para o treinamento por classe;
- 10 imagens para o reconhecimento por classe;
- 1200 poses analisadas;
- 1192 poses reconhecidas corretamente; e
- oito poses reconhecidas incorretamente

4.3.2.3- Utilizando 2200 coeficientes DCT, onde quatro são de cada subpose UK

Passo 1: Considerando a Figura 3.5, tomou-se primeiramente a pose $G_{I,1}$. Desta pose, representada pela Figura 3.13, retirou-se a subpose Y_I , ilustrada na Figura 4.46 com destaque para os quatro coeficientes mais significativos.

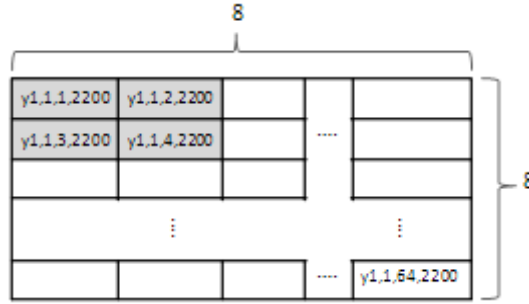


Figura 4.46: Representação da subpose Y_1 , base para o vetor característico de 2200 coeficientes DCT – UK.

Passo 2: Extraíram-se da subpose Y_1 os coeficientes mais significativos $y_{1,1,1,2200}$, $y_{1,1,2,2200}$, $y_{1,1,3,2200}$ e $y_{1,1,4,2200}$, colocando-os em um vetor. Depois da Figura 3.13 e extraiu-se a pose Y_2 , extraíram-se os quatro coeficientes mais significativos $y_{1,2,1,2200}$, $y_{1,2,2,2200}$, $y_{1,2,3,2200}$ e $y_{1,2,4,2200}$ colocando-os no mesmo vetor e assim sucessivamente até extrair os coeficientes $y_{1,550,1,2200}$, $y_{1,550,2,2200}$, $y_{1,550,3,2200}$ e $y_{1,550,4,2200}$ da subpose Y_{550} colocando-os nesse vetor mostrado completo na Figura 4.47. Desta forma, a pose $G_{1,1}$ (primeira pose da classe F_1) foi representada por um vetor de 2200 elementos.

Para melhor entendimento, segue a descrição dos coeficientes:

$y_{1,1,1,2200} - y_1$ refere-se a pose 1. O número 1 seguinte quer dizer sub-pose 1. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$y_{1,1,2,2200} - y_1$ refere-se à pose 1. O número 1 seguinte quer dizer subpose 1. O número 2 seguinte quer dizer o segundo coeficiente mais significativo.

$y_{1,1,3,2200} - y_1$ refere-se à pose 1. O número 1 seguinte quer dizer subpose e 1. O número 3 seguinte quer dizer o terceiro coeficiente mais significativo.

$y_{1,1,4,2200} - y_1$ refere-se à pose 1. O número 1 seguinte quer dizer subpose e 1. O número 2 seguinte quer dizer o quarto coeficiente mais significativo.

$y_{1,2,1,2200} - y_1$ refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$y_{1,2,2,2200} - y_1$ refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 2 seguinte quer dizer o segundo coeficiente mais significativo.

$y_{1,2,3,2200} - y_1$ refere-se à pose 1. O número 2 seguinte quer dizer subpose 2. O número 2 seguinte quer dizer o terceiro coeficiente mais significativo.

$y_{1,2,4,2200} - y_1$ refere-se a pose 1. O número 2 seguinte quer dizer sub-pose 2. O número 2 seguinte quer dizer o quarto coeficiente mais significativo.

$y_{1,550,1,2200} - y_1$ refere-se à pose 1. O número 550 seguinte quer dizer subpose 550. O número 1 seguinte quer dizer o primeiro coeficiente mais significativo.

$y_{1,550,2,2200} - y_1$ refere-se à pose 1. O número 550 seguinte quer dizer subpose 550. O número 2 seguinte quer dizer o segundo coeficiente mais significativo.

$y_{1,550,3,2200} - y_1$ refere-se à pose 1. O número 550 seguinte quer dizer subpose 550. O número 3 seguinte quer dizer o terceiro coeficiente mais significativo.

$y_{1,550,4,2200} - y_1$ refere-se à pose 1. O número 550 seguinte quer dizer subpose 550. O número 4 seguinte quer dizer o quarto coeficiente mais significativo.

Ou seja, de cada subpose extraíram-se quatro coeficientes.

O índice 2200 no final de cada símbolo de coeficiente refere-se ao caso em estudo desta seção e também é o total de elementos do vetor que representa uma pose.

$G_{1,1}$	$y_{1,1,1,2200}$	$y_{1,1,2,2200}$	$y_{1,1,3,2200}$	$y_{1,1,4,2200}$	$y_{1,550,1,2200}$	$y_{1,550,2,2200}$	$y_{1,550,3,2200}$	$y_{1,550,4,2200}$
-----------	------------------	------------------	------------------	------------------	-------	--------------------	--------------------	--------------------	--------------------

Figura 4.47: Vetor representativo da primeira pose G através de 2200 coeficientes DCT mais significativos de cada subpose- arquivo UK.

Passo 3: Repetiram-se os passos 1 e 2 para as poses $G_{2,1}$ a $G_{20,1}$ (Figura 3.5), obtendo-se o conjunto de 20 vetores que formam a matriz de coeficientes DCT representativa das vinte poses da classe F_1 . Esse conjunto é mostrado na Figura 4.48.

$G_{1,1}$	$y_{1,1,1,2200}$	$y_{1,1,2,2200}$	$y_{1,1,3,2200}$	$y_{1,1,4,2200}$	$y_{1,550,1,2200}$	$y_{1,550,2,2200}$	$y_{1,550,3,2200}$	$y_{1,550,4,2200}$
$G_{2,1}$	$y_{2,1,1,2200}$	$y_{2,1,2,2200}$	$y_{2,1,3,2200}$	$y_{2,1,4,2200}$	$y_{2,550,1,2200}$	$y_{2,550,2,2200}$	$y_{2,550,3,2200}$	$y_{2,550,4,2200}$
$G_{3,1}$	$y_{3,1,1,2200}$	$y_{3,1,2,2200}$	$y_{3,1,3,2200}$	$y_{3,1,4,2200}$	$y_{3,550,1,2200}$	$y_{3,550,2,2200}$	$y_{3,550,3,2200}$	$y_{3,550,4,2200}$
...									
$G_{20,1}$	$y_{20,1,1,2200}$	$y_{20,1,2,2200}$	$y_{20,1,3,2200}$	$y_{20,1,4,2200}$	$y_{20,550,1,2200}$	$y_{20,550,2,2200}$	$y_{20,550,3,2200}$	$y_{20,550,4,2200}$

Figura 4.48: Conjunto de vetores com 2200 coeficientes DCT cada um, representativo da classe F_1 - arquivo UK.

Passo 4: Repetiram-se os passos desta seção para achar o conjunto de vetores representativos das classes F_2 , F_3 , até a classe F_{120} .

A Tabela 4.8 mostra o resultado do reconhecimento de face utilizando 2200 coeficientes DCT para representar cada pose, arquivo UK. Os valores contidos nessa tabela são aqueles obtidos nos testes para 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 e 120 classes. As variáveis são: VAR são as variações da experiência para o treinamento e o

reconhecimento, NPT é o número poses para o treinamento, NPR é o número de poses para o reconhecimento de faces. Ainda os parâmetros AC que são a taxa de acerto em percentual, TT , é o tempo de treinamento em segundos, e TR é o tempo de reconhecimento em segundos.

Nas Figuras 4.49 até a 4.52, são mostrados os gráficos resultantes desta experiência.

Tabela 4.8: Resultados obtidos do reconhecimento de face considerando 2200 coeficientes DCT da subimagem, arquivo UK.

V A R	N P T	N P R	Parâ- metros	Classes												
				5	10	20	30	40	50	60	70	80	90	100	110	120
1	19	1	AC (%)	100	100	100	100	100	100	100	100	100	100	100	100	100
			TT (s)	344,21	688,92	1336,8	2081,38	2464,43	3291,65	4145,83	4586,58	5343,12	5896,01	6512,17	7262,52	8214,03
			TR (s)	0,0427	0,24	0,28	2,17	3,38	3,87	5,670	7,50	10,71	13,00	15,23	17,24	21,868
2	18	2	AC (%)	100	100	100	100	100	100	100	99,6005	99,8801	99,8801	99,4806	98,9712	99,58
			TT (s)	310,48	540,11	1249,5	1729,95	2518,66	2967,20	3763,53	4281,78	4818,75	5371,33	5836,06	6786,67	7471,31
			TR (s)	0,0858	0,47	1,72	2,30	6,93	8,42	10,995	13,97	19,25	24,19	29,88	36,42	42,813
3	17	3	AC (%)	100	100	100	100	99,4507	99,2109	94,44	99,4007	99,4007	99,4307	99,5406	99,6804	99,44
			TT (s)	283,07	560,87	1084,6	1692,65	2178,82	2694,77	3376,34	3811,99	4413,48	4860,97	5319,84	6079,36	6795,408
			TR (s)	0,1206	0,65	2,30	6,08	9,16	11,04	16,018	21,79	24,96	30,04	38,10	53,94	62,790
4	16	4	AC (%)	100	100	100	99,2109	98,6316	99,2109	99,16	99,5805	99,5605	99,6005	99,5805	99,6504	99,79
			TT (s)	267,56	480,92	994,95	1566,00	1995,14	2387,34	3135,41	3444,23	4000,88	4319,79	4753,98	5461,10	6192,184
			TR (s)	0,1565	0,83	3,30	7,57	13,13	13,72	21,032	26,82	37,13	46,04	52,67	69,44	82,182
5	15	5	AC (%)	100	100	100	99,2109	99,2109	99,0811	99,33	99,3008	99,5605	99,2109	99,5106	99,5106	99,66
			TT (s)	230,80	452,07	852,65	1368,09	1727,30	2159,76	2668,70	3040,23	3572,92	3872,45	4296,366	4729,74	5321,69
			TR (s)	0,1923	1,03	4,10	9,89	15,76	17,59	25,612	34,85	48,97	56,37	69,95	84,07	100,022
6	14	6	AC (%)	100	100	100	99,0811	98,6316	99,2109	99,16	99,2509	99,2509	99,4287	99,2109	99,2109	99,44
			TT (s)	202,50	383,44	813,59	1112,56	1623,64	1815,45	2320,765	2658,21	3172,89	3411,78	3965,68	4297,77	4674,31
			TR (s)	0,227	1,22	4,78	12,72	18,77	19,34	29,688	41,27	51,92	66,80	82,56	101,96	118,010
7	13	7	AC (%)	100	100	100	99,4507	99,0212	99,0212	98,57	99,2609	99,2609	98,9213	98,9712	98,9712	99,166
			TT (s)	178,38	360,25	567,22	1083,74	1241,88	1688,78	2047,88	2376,03	2763,63	3021,05	3426,61	3698,95	4118,89
			TR (s)	0,259	1,37	5,37	13,05	21,82	23,31	33,946	46,29	59,99	75,73	92,62	112,83	134,015
8	12	8	AC (%)	100	100	100	99,2109	99,5605	99,0212	99,37	98,6316	99,5605	99,5605	99,2509	99,2509	99,58
			TT (s)	154,77	224,52	636,79	903,62	1278,24	1578,61	1845,77	2087,61	2440,68	2679,92	3049,80	3131,75	3580,16
			TR (s)	0,2873	1,52	6,10	14,31	25,61	24,74	37,66	50,21	67,12	85,33	104,76	133,73	150,556
9	11	9	AC (%)	100	100	100	96,9137	99,4806	98,7615	99,25	99,4007	99,0112	99,0112	98,9712	98,9712	98,88
			TT (s)	133,21	275,81	406,45	828,49	1073,83	1316,18	1546,71	1798,65	2072,40	2297,23	2684,31	2806,77	3110,193
			TR (s)	0,3166	1,70	6,70	14,91	26,70	28,46	41,26	56,74	74,33	92,41	110,64	138,27	163,55
10	10	10	AC (%)	100	100	98,382	99,2109	99,3807	99,3807	99,33	99,1310	99,1310	99,2109	99,1810	99,1810	98,916
			TT (s)	116,30	223,03	475,25	564,86	954,47	1185,54	1323,78	1479,79	1792,02	2018,84	2265,68	2480,11	2655,69
			TR (s)	0,3751	4,03	7,20	21,20	28,77	37,47	44,445	68,85	79,98	110,99	124,80	157,52	175,791

A Figura 4.49 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Percebe-se que para o reconhecimento com uma pose a taxa de acerto é de 100% para 5 até 120 classes. Nota-se também a queda na taxa de acerto na classe 20 para reconhecimento com 10 poses. A oscilação das curvas valores das taxas de acertos se explica pelo fato de as vezes ocorrer saturação no processamento do algoritmo e ainda pela diferença de centralização e de iluminação em certas poses dentro das respectivas classes.

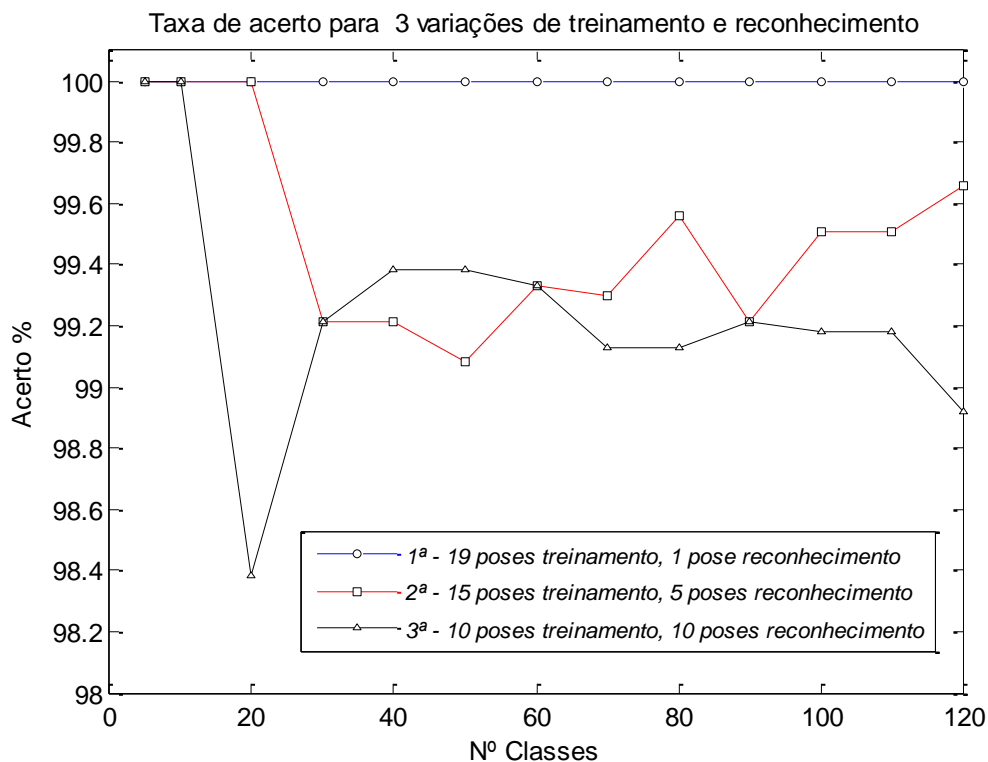


Figura 4.49: Taxa de acerto considerando 2200 coeficientes por pose, arquivo UK.

A Figura 4.50 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente linearmente progressivo com o aumento do número de classes. Nota-se que quanto maior o número de poses utilizados no treinamento mais acentuado é o crescimento dos valores do tempo de treinamento, 8214,03 s para treinamento com 19 poses.

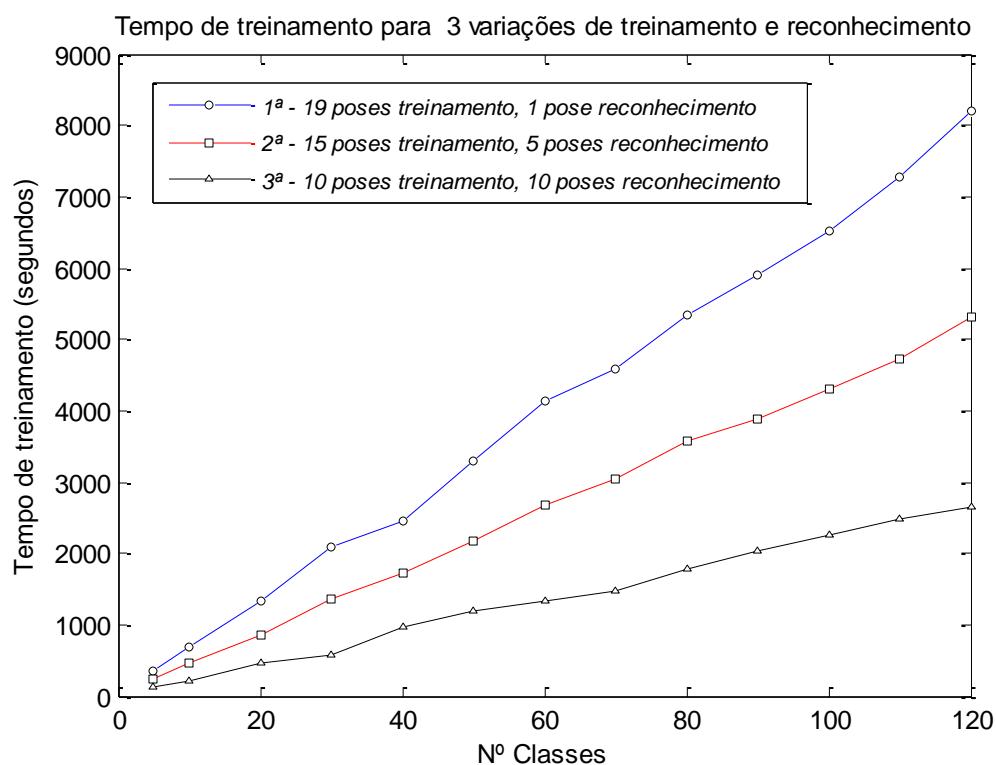


Figura 4.50: Tempo de treinamento considerando 2200 coeficientes por pose, arquivo UK.

A Figura 4,51 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Nota-se que mesmo com aumento do número de poses utilizados no reconhecimento o AG efetua o reconhecimento em um tempo bastante curto, 175,791 s para reconhecimento com 10 poses.

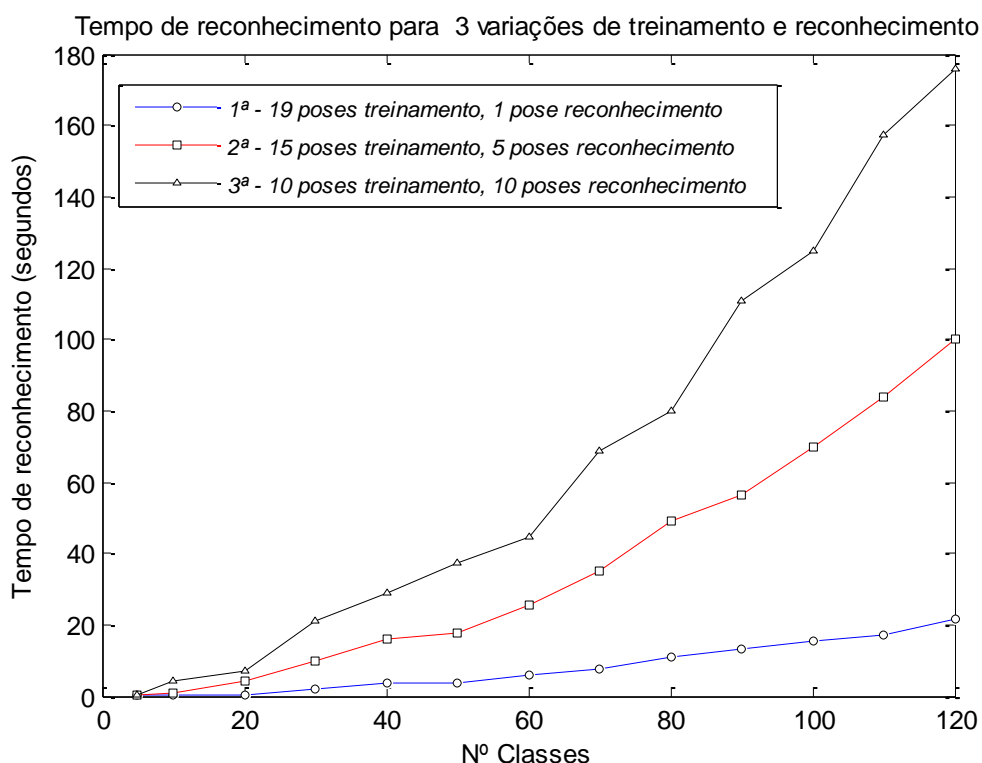


Figura 4.51: Tempo de reconhecimento do AG considerando 2200 coeficientes por pose, arquivo UK.

A Figura 4.52 mostra o gráfico resultante dessa experiência referente ao número de erros ocorridos no arquivo UK, considerando 10 poses para o treinamento e 10 poses para o reconhecimento e total de 120 classes. Nesse teste utilizou-se 1200 poses (10 de cada classe) para o reconhecimento. O gráfico mostra a variação de erros na execução de 100 gerações para 120 classes. Neste reconhecimento, das 1200 poses, para 13 existe erro no reconhecimento. Por isso a taxa de acerto de 98,916 % mostrada na Tabela 4.8. Nota-se que o número de erros oscila em função do número de geração processada e somente neste caso, a taxa de erro foi menor na trigésima geração. Se alterar o número de classe, o número de poses para o treinamento e o número de poses para o reconhecimento, os valores da taxa de erro alteram também. Por isso, o sistema tem que fixar um número de geração que seja capaz de fornecer uma taxa alta de acerto mesmo com a manipulação das referidas variáveis.

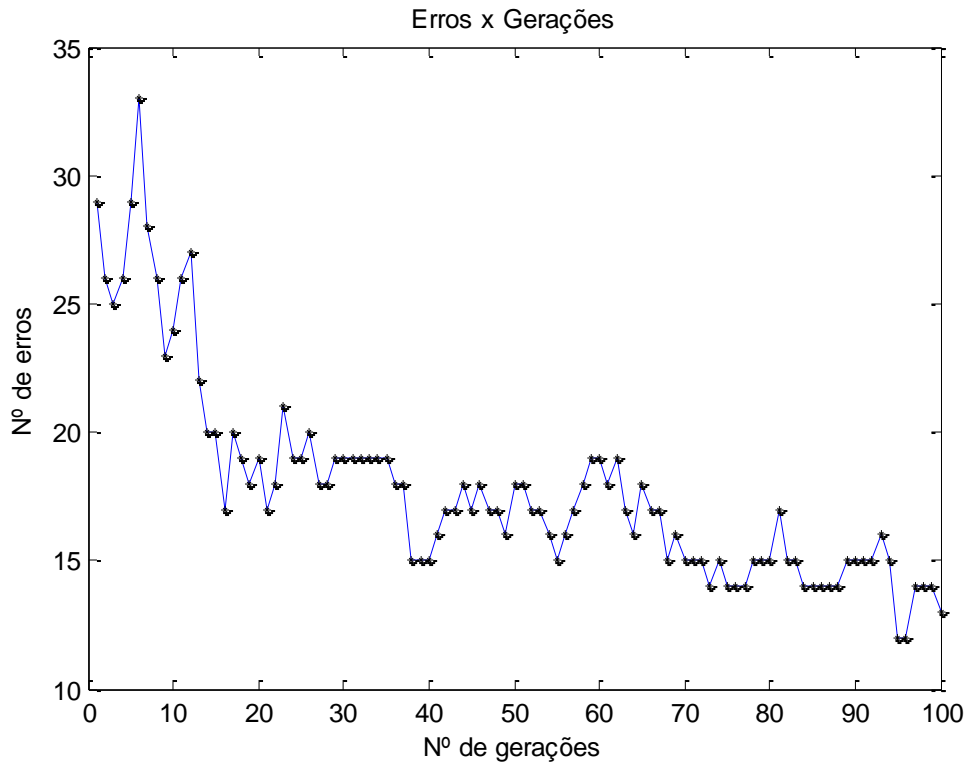


Figura 4.52: Gráficos de erros do AG em função do número de gerações onde foram utilizados 10 poses para treinamento e 10 poses para o reconhecimento – usando 2200 coeficientes DCT, para 120 classes, arquivo UK.

Sobre esta variação com 2200 coeficientes DCT, para 120 classes e dados de 100 gerações, obtiveram-se:

- 10 imagens para o treinamento por classe;
- 10 imagens para o reconhecimento por classe;
- 1200 poses analisadas;
- 1187 poses reconhecidas corretamente; e
- 13 poses reconhecidas incorretamente.

4.4- Considerações Finais Deste Capítulo

A taxa de acerto do AG aumenta com o aumento do número de imagens no arquivo. O arquivo UK tem muito mais imagens que o ORL e, mesmo assim, ambos apresentaram taxas de acerto semelhantes. O AG mostrou ser um instrumento capaz de

manter aceitabilidade mesmo com o aumento de dados a serem processados. Ele é um algoritmo que trabalha bem com grande massa de dados. Embora o tempo de treinamento ou a modelagem sejam razoavelmente longos, a facilidade de preservar a condição modelada é um fator positivo. E o mais importante, a etapa de reconhecimento, foi realizada em um tempo irrisório independentemente da quantidade de classe a ser analisada. Esta constatação o remete de imediato para o campo da aplicação prática e da pesquisa.

O ser humano vai continuar se transformando. Pela teoria Gaia (Deusa grega de Terra), “a Terra pode ser considerada como um sistema vivo único e autorregulado” [51]. Como todos outros primatas, os humanos não são obra de Deus, mas de centenas de milhões de anos de interação entre os micróbios altamente responsivos [51]. Conclui-se que, como o tempo transformou os humanos, a alta tecnologia, cada vez mais disponível, também o fará. Desta forma, necessário se faz o constate refino dos algoritmos de reconhecimento de faces, para incluir também o histórico da transformação das imagens das faces.

O próximo capítulo mostra os testes realizados com a rede neural artificial.

CAPÍTULO V

UTILIZAÇÃO DA REDE NEURAL ARTIFICIAL NO RECONHECIMENTO DE FACES

5.1- Introdução

Nos testes realizados com a Rede Neural Artificial (RNA) os dados foram obtidos dos arquivos ORL e UK, especificados nas seções 3.3.2 do Capítulo 3 deste trabalho.

As quantidades dos coeficientes DCT da transformada das matrizes com valores de pixels de todas as poses de cada classe foram 64, 100 e 256 coeficientes para ambos os arquivos para representar cada pose, conforme descrito nas seções 3.3.1.1 e 3.3.1.2 do Capítulo 3 desta tese.

O que se pretendeu foi que o algoritmo identificasse a que classe pertencia a pose ou o grupo destas quando fosse considerada uma ou mais poses (imagens query), de qualquer uma das classes pertencentes aos arquivos.

A implementação do sistema conforme a estrutura da RNA foi feita adotando os seguintes parâmetros:

1. variáveis de entradas: 200 vetores característicos;
2. camada intermediária: 30 neurônios;
3. épocas: até 5000 loops; e
4. aproximação por erro: taxa de 0,01.

A rotina de cálculo da RNA utilizou matrizes de coeficientes gerados pela DCT em três etapas para cada um dos arquivos: na primeira etapa foram utilizados 64 coeficientes, na segunda etapa utilizaram-se 100 coeficiente de cada e na terceira etapa foram utilizados 256 coeficientes de cada pose.

Este capítulo mostra os testes realizados com a rede neural artificial.

5.2- Experiência com a Rede Neural Artificial utilizando o arquivo ORL

Nesta seção, os testes com a RNA foram realizados com o arquivo ORL utilizando primeiro 64 coeficientes, em seguida 100 coeficientes e por último 256 coeficientes DCT por pose.

5.2.1- Experiência utilizando 64 coeficientes DCT, arquivo ORL

Após o cálculo da DCT, utilizaram-se 64 coeficientes mais significativos para compor cada pose.

A Tabela 5.1 mostra os resultados das experiências com a RNA usando cinco poses para o treinamento e de uma a cinco poses para o reconhecimento de face utilizando 64 coeficientes DCT para representar cada pose. Tem-se também *NPT* como número de poses para o treinamento, *NPR* número de poses para o reconhecimento de faces e ainda os parâmetros *AC* que são a taxa de acerto em percentual, *TT*, tempo de treinamento em segundos, e *TR* que é o tempo de reconhecimento em segundos. Pode-se observar na Tabela 5.1 que a taxa de acerto tem os valores muito próximos para todas as classes. Para 40 classes o menor tempo de reconhecimento foi de 0,60 s e o tempo de treinamento foi de 270,29 s.

Tabela 5.1: Resultados da experiência com a RNA usando cinco poses para o treinamento e de uma a cinco poses para o reconhecimento para o nº de 64 coeficientes – ORL.

NPT	NPR	Parâ- metros	Classes							
			2	5	10	15	20	25	30	40
5	1	AC	97,00	99,60	99,50	99,70	99,55	99,75	99,70	99,81
		TR	0,03	0,11	0,13	0,20	0,30	0,35	0,53	0,60
	2	AC	96,50	99,60	99,60	99,73	99,65	99,76	99,73	99,83
		TR	0,03	0,10	0,16	0,25	0,35	0,40	0,53	0,68
	3	AC	97,33	98,80	99,86	99,64	99,66	99,74	99,73	99,86
		TR	0,04	0,12	0,26	0,36	0,49	0,60	0,73	1,01
	4	AC	97,50	99,80	99,85	99,75	99,71	99,78	99,72	99,81
		TR	0,08	0,15	0,31	0,46	0,74	0,74	0,97	1,41
	5	AC	100	99,84	99,76	99,69	99,67	99,80	99,76	99,85
		TR	0,09	0,18	0,43	0,56	0,83	0,93	1,24	1,75
TT			1,33	2,93	8,16	13,87	21,35	33,13	70,17	270,29

Da Figura 5.1 a 5.3, são mostrados os gráficos resultantes desta experiência, primeiro das taxas de acertos e, em seguida dos tempos de reconhecimento e de treinamento.

A Figura 5.1 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Pode-se observar nessa figura que para o reconhecimento realizado com uma pose a 5 poses a taxa de acerto está em torno de 99% para 2, 5, 10, 15, 20, 25, 30 e 40 classes. Observa-se também a menor taxa de acerto é 97 % quando foi realizado os testes para 2 classes com um poses para o reconhecimento.

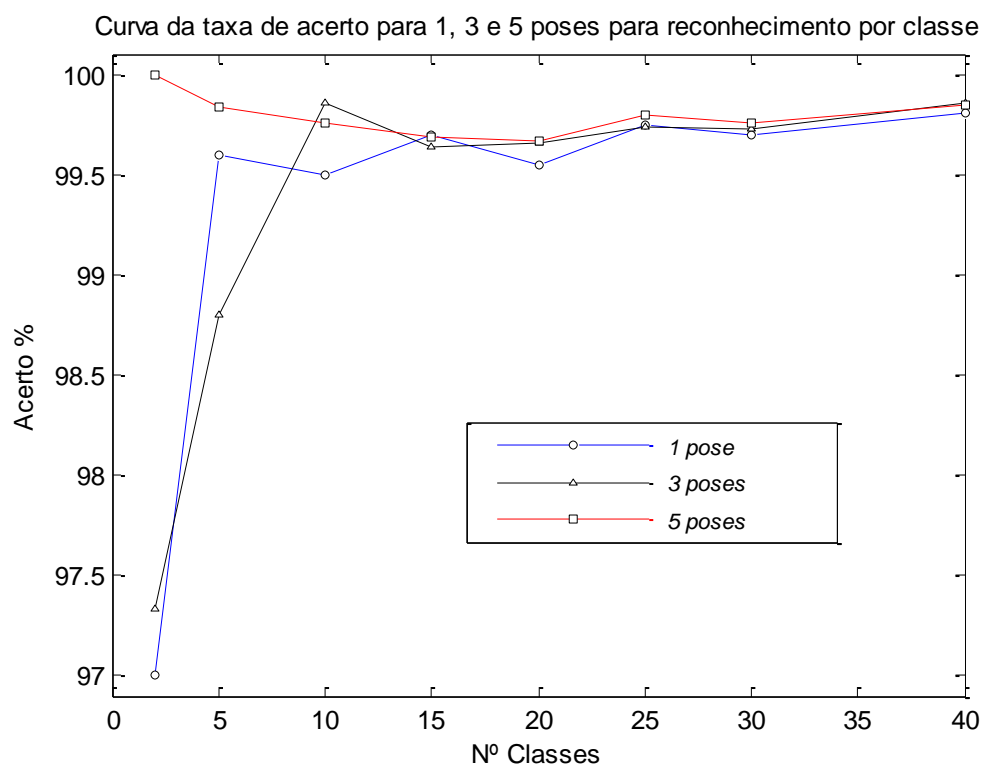


Figura 5.1: Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, três e cinco poses para o reconhecimento, 64 coeficientes- ORL.

A Figura 5.2 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Observa-se que mesmo com o aumento do número de poses utilizados no reconhecimento, a RNA efetua o reconhecimento em um tempo bastante curto, 1,75 s para reconhecimento com 5 poses sendo o maior tempo de reconhecimento.

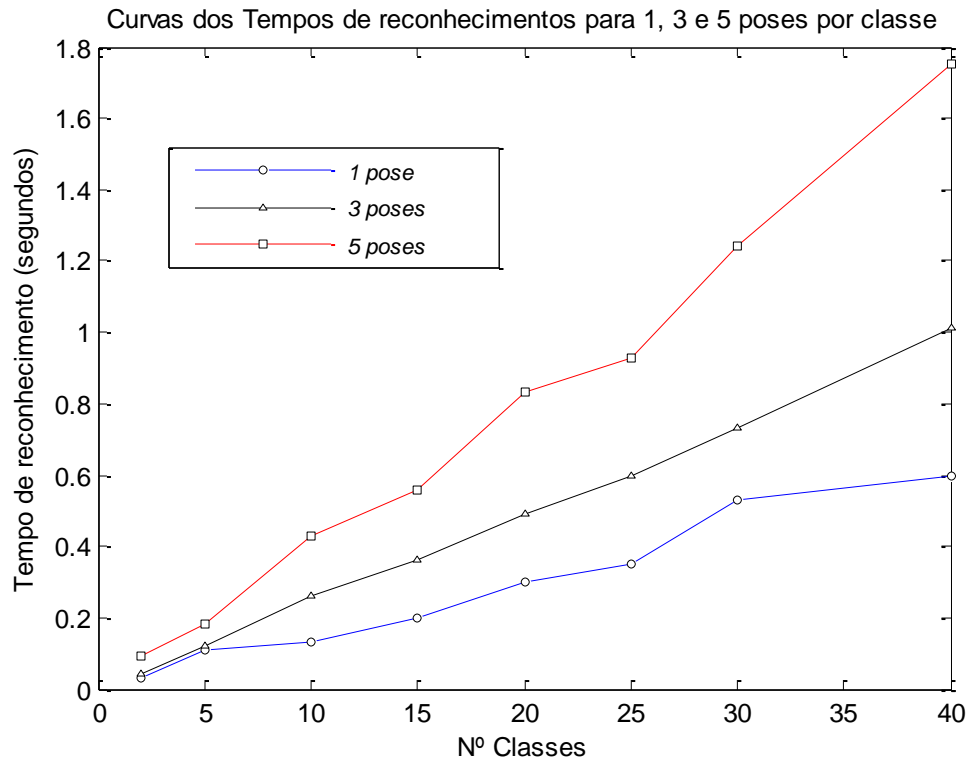


Figura 5.2: Curvas dos tempos de reconhecimentos da RNA em função do n° de classes utilizando uma, três e cinco poses, 64 coeficientes – ORL.

A Figura 5.3 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente em relação ao número de classes. Observa-se que o treinamento é realizado de uma só vez para cada número de classe. O tempo de treinamento varia de 1,33 s a 270,29 s para a quantidade 2 e 40 classes respectivamente.

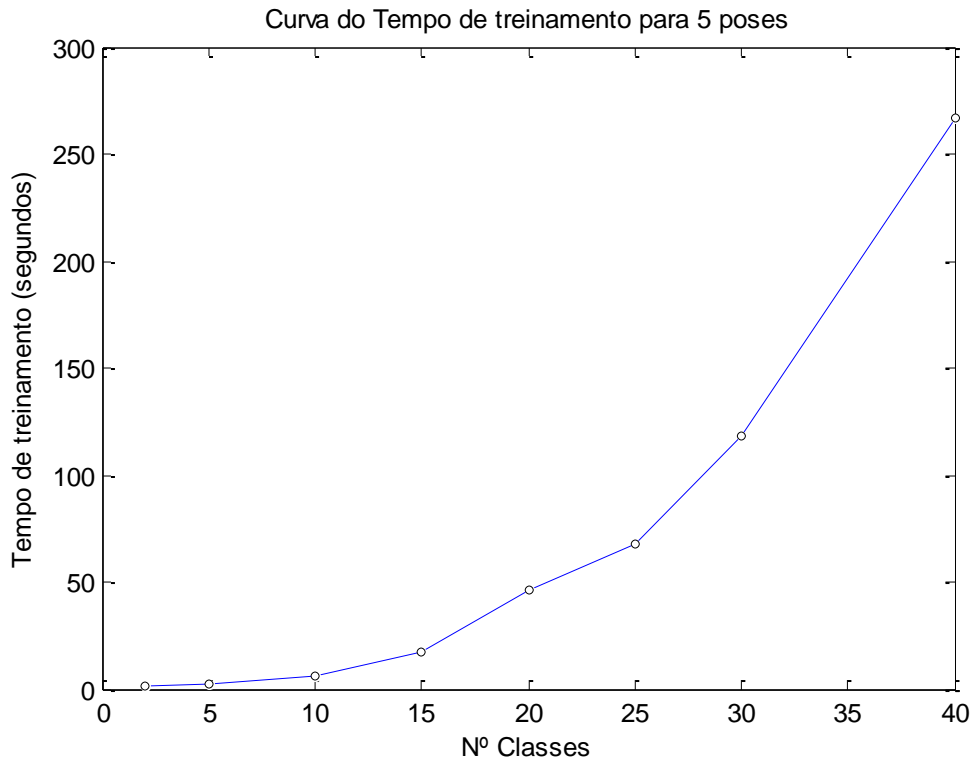


Figura 5.3: Curva do tempo de treinamento da RNA em função do nº de classes para 64 coeficientes por pose – ORL.

5.2.2- Experiência utilizando 100 coeficientes DCT utilizando o arquivo ORL

Após o cálculo da DCT utilizaram-se 100 coeficientes. Aqueles localizados no canto superior esquerdo com dimensão 10 por 10.

A Tabela 5.2 mostra os resultados das experiências com a RNA usando cinco poses para o treinamento e de uma a cinco poses para o reconhecimento de face utilizando 100 coeficientes DCT para representar cada pose. Tem-se também NPT é o número poses para o treinamento, NPR é o número de poses para o reconhecimento de faces e ainda os parâmetros AC é a taxa de acerto em percentual, TT é o tempo de treinamento em segundos, e TR é o tempo de reconhecimento em segundos. Pode-se observar na Tabela 5.2 que a taxa de acerto tem os valores muito próximos para todas as classes. Para 40 classes o menor tempo de reconhecimento foi de 0,64 s e o tempo de treinamento foi de 224,60 s.

Tabela 5.2: Resultados da experiência com a RNA utilizando cinco poses para o treinamento e de um a cinco poses para o reconhecimento para o n° de 100 coeficientes- ORL.

NPT	NPR	Parâ- metros	Classes							
			2	5	10	15	20	25	30	40
5	1	AC	97,25	98,30	99,00	99,70	99,80	98,90	99,80	99,90
		TR	0,05	0,10	0,17	0,27	0,25	0,31	0,45	0,64
	2	AC	97,50	99,20	98,90	99,60	99,67	99,80	99,80	99,87
		TR	0,05	0,10	0,18	0,27	0,31	0,40	0,50	0,68
	3	AC	96,66	98,40	99,53	99,61	99,63	99,81	99,79	99,85
		TR	0,07	0,14	0,24	0,38	0,46	0,58	0,73	1,00
	4	AC	96,25	99,20	99,30	99,75	99,66	99,76	99,82	99,84
		TR	0,09	0,17	0,30	0,48	0,60	0,77	0,95	1,32
	5	AC	95,00	98,08	99,40	99,57	99,71	99,77	99,85	99,87
		TR	0,10	0,20	0,40	0,60	0,74	0,95	1,20	1,70
TT			0,93	1,77	5,66	16,83	30,06	79,68	153,81	224,60

As Figuras 5.4 a 5.6 mostram os gráficos resultantes desta experiência, primeiramente referentes às taxas de acertos e, em seguida, aos tempos de reconhecimento e de treinamento.

A Figura 5.4 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Pode-se observar nessa figura que para o reconhecimento realizado com uma pose a 5 poses a taxa de acerto é menor para menor número de classes, até 15 classes e estabiliza com valor acima de 99,5% após os testes com 15 classes. Com 40 classes a taxa de acerto é a maior para as três curvas traçadas. Observa-se também a menor taxa de acerto é 95 % quando foi realizado os testes para 2 classes com uma poses para o reconhecimento.

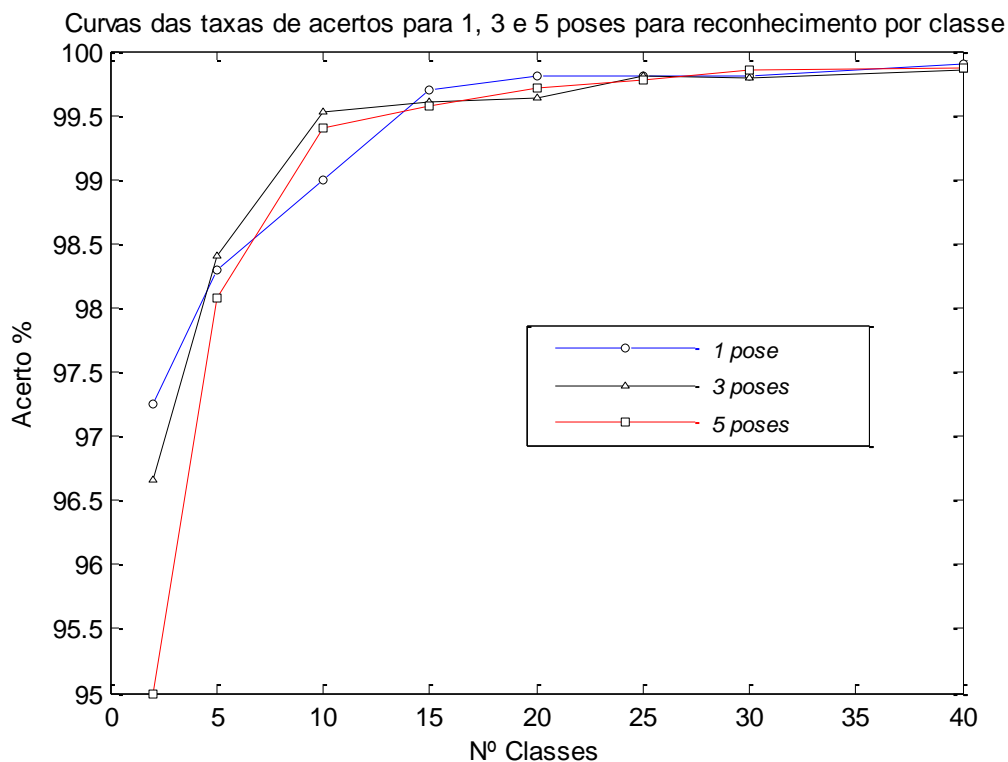


Figura 5.4: Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, três e cinco poses para o reconhecimento, 100 coeficientes- ORL.

A Figura 5.5 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Observa-se que mesmo com o aumento do número de poses utilizados no reconhecimento, a RNA efetua o reconhecimento em um tempo bastante curto, 1,70 s para reconhecimento com 5 poses que é o maior tempo de reconhecimento da Tabela 5.2.

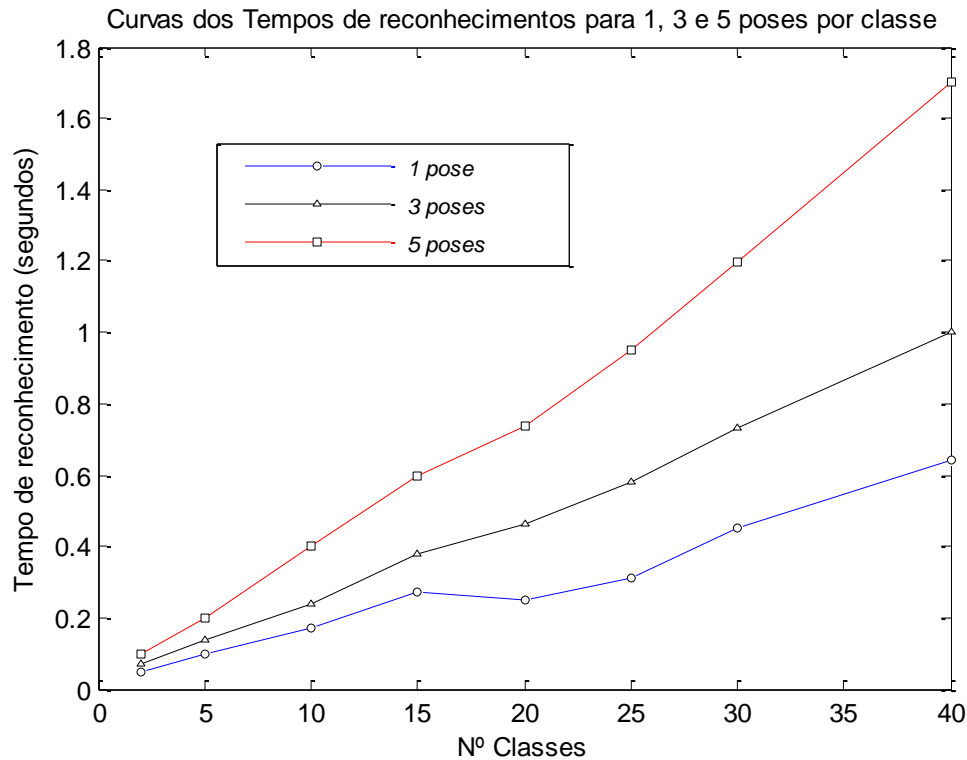


Figura 5.5: Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, três e cinco poses, 100 coeficientes – ORL.

A Figura 5.6 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente em relação ao número de classes. Observa-se que o treinamento de uma só vez para cada número de classe. O tempo de treinamento varia de 0,93 s a 224,60 s para a quantidade de 2 e 40 classes, respectivamente.

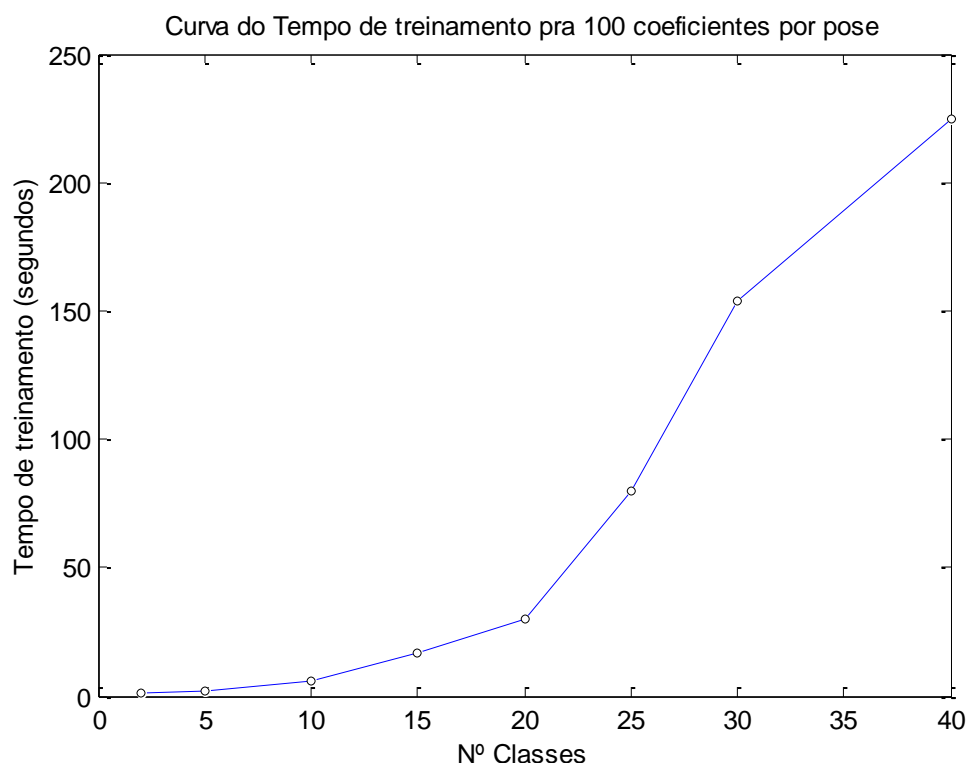


Figura 5.6: Curva do tempo de treinamento da RNA em função do nº de classes para 100 coeficientes por pose – ORL.

5.2.3- Experiência utilizando 256 coeficientes DCT, arquivo ORL

Após o cálculo da DCT, foram utilizados 256 coeficientes. Aqueles localizados no canto superior esquerdo com a dimensão 16 por 16.

A Tabela 5.3 mostra os resultados das experiências com a RNA usando cinco poses para o treinamento e de uma, três e cinco poses para o reconhecimento de face utilizando 256 coeficientes DCT para representar cada pose. Tem-se também NPT como número poses para o treinamento, NPR número de poses para o reconhecimento de faces e ainda os parâmetros AC que são a taxa de acerto em percentual, TT , tempo de treinamento em segundos, e TR que é o tempo de reconhecimento em segundos. Pode-se observar na Tabela 5.3 que a taxa de acerto tem os valores muito próximos após a quantidade de 5 classes. Para as 40 classes o menor tempo de reconhecimento foi de 0,37 s e o tempo de treinamento foi de 267,05 s.

Tabela 5.3: Resultados da experiência com a RNA utilizando cinco poses para o treinamento e de uma, três e cinco poses para o reconhecimento para o n° de 256 coeficientes- ORL.

NPT	NPR	Parâ- metros	Classes							
			2	5	10	15	20	25	30	40
5	1	AC	95,00	99,60	99,80	99,30	99,80	99,80	99,70	99,77
		TR	0,06	0,10	0,17	0,15	0,29	0,25	0,48	0,37
	2	AC	95,00	99,60	99,70	99,42	99,60	99,77	99,73	99,76
		TR	0,06	0,09	0,18	0,29	0,37	0,43	0,57	0,49
	3	AC	96,66	99,73	99,46	99,61	99,68	99,78	99,76	99,80
		TR	0,07	0,14	0,29	0,42	0,52	0,62	0,88	1,29
	4	AC	95,00	99,40	99,65	99,53	99,65	99,71	99,82	99,82
		TR	0,08	0,16	0,33	0,56	0,78	0,80	1,16	1,68
	5	AC	99,00	99,52	99,72	99,52	99,71	99,78	99,82	99,82
		TR	0,09	0,19	0,41	0,66	0,89	0,99	1,39	2,02
TT			1,14	2,68	6,29	17,54	46,40	67,56	118,14	267,05

Nas Figuras 5.7 a 5.9 são mostrados os gráficos resultantes desta experiência, primeiramente referentes às taxas de acertos e, em seguida, aos tempos de reconhecimento e de treinamento.

A Figura 5.7 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Pode-se observar nessa figura que para o reconhecimento realizado com 1, 3 e 5 poses a taxa de acerto é menor para 2 a 5 classes a partir da qual flutua em torno 99,5%. Observa-se também a menor taxa de acerto é 95 % quando foi realizado os testes para 2 classes com um poses para o reconhecimento.

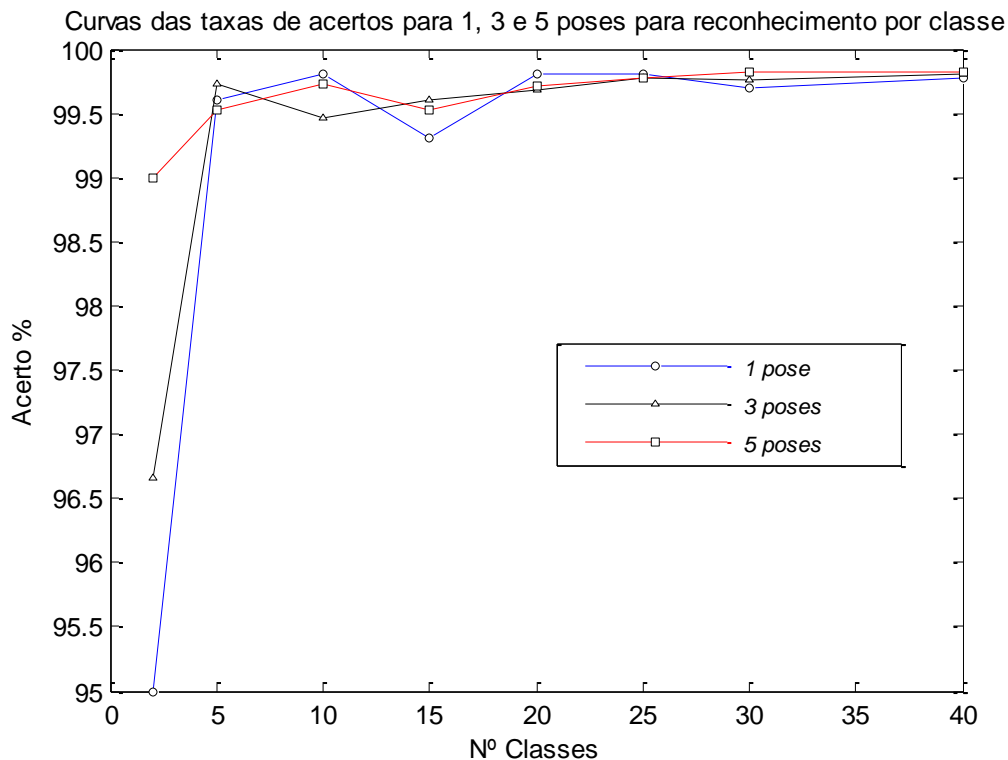


Figura 5.7: Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, três e cinco poses para o reconhecimento, 256 coeficientes – ORL.

A Figura 5.8 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Observa-se que mesmo com o aumento do número de poses utilizados no reconhecimento, a RNA efetua o reconhecimento em um tempo bastante curto, 2,02 s para reconhecimento com 5 poses que é o maior tempo de reconhecimento da Tabela 5.3.

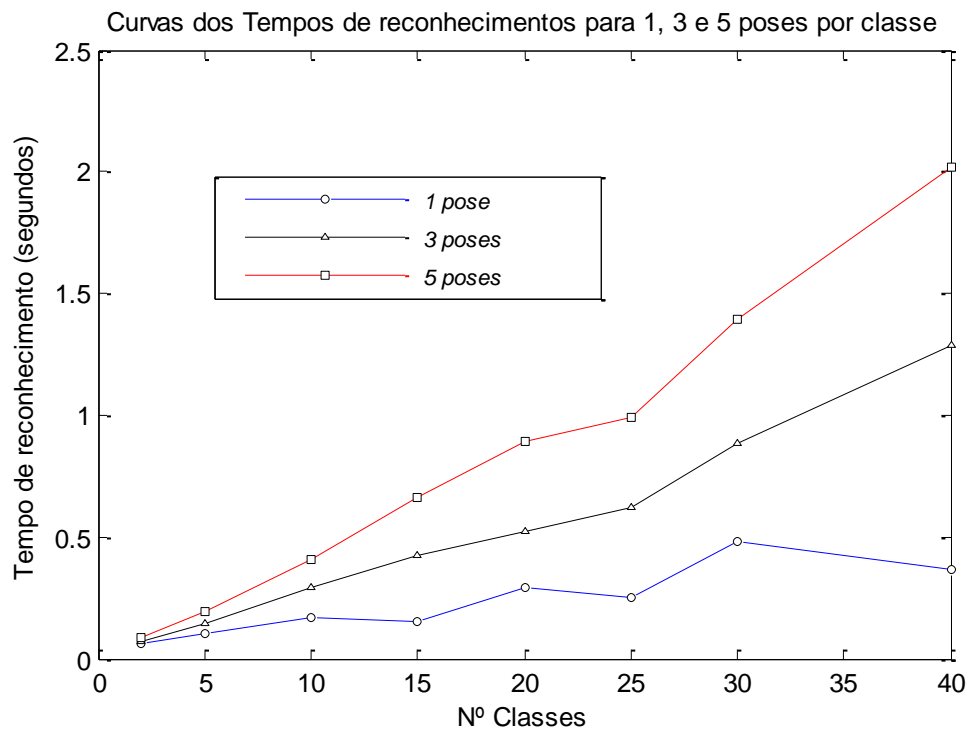


Figura 5.8: Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, três e cinco poses, 256 coeficientes – ORL.

A Figura 5.9 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente em relação ao número de classes. Observa-se que o treinamento de uma só vez para cada número de classe. O tempo de treinamento varia de 1,44 s a 267,05 s para a quantidade de classes 2 e 40, respectivamente.

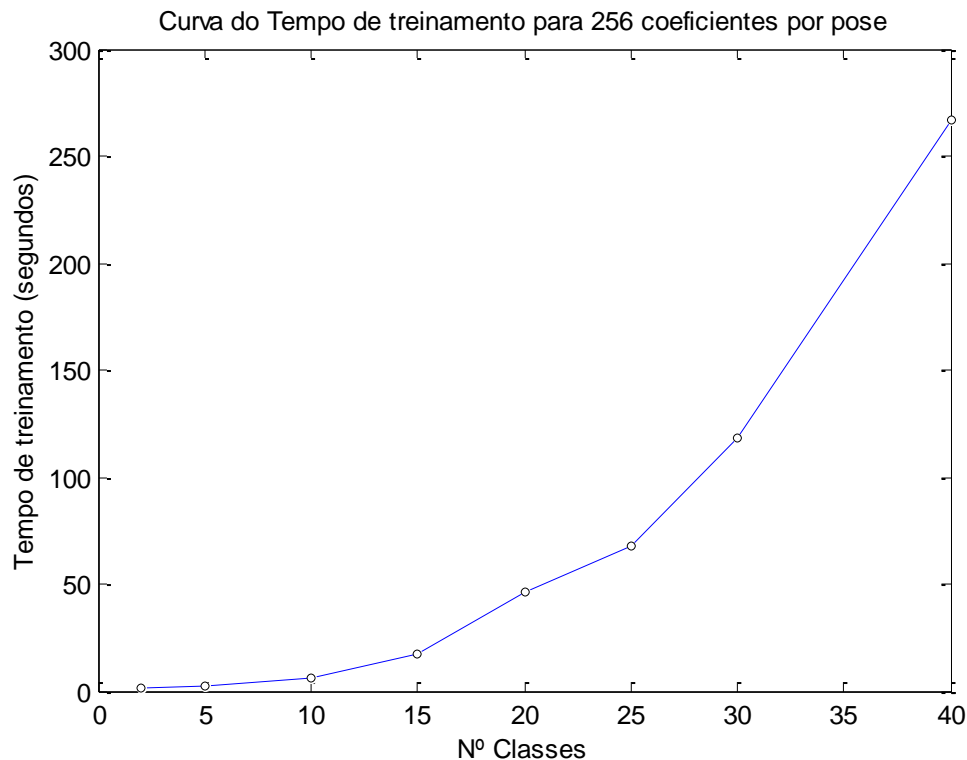


Figura 5.9: Curva do tempo de treinamento da RNA em função do nº de classes para 256 coeficientes por pose – ORL.

5.3- Experiência com a Rede Neural Artificial utilizando o arquivo UK

Nesta seção, se tem a utilização, pela RNA, do arquivo UK utilizando primeiramente 64 coeficientes, em seguida 100 coeficientes e, por último, 256 coeficientes DCT por pose.

5.3.1- A experiência com 64 coeficientes DCT utilizando arquivo UK

Após o cálculo da DCT, utilizaram-se 64 coeficientes mais significativos para compor cada pose.

As Tabelas 5.4 e 5.5 mostram os resultados das experiências com a RNA usando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de face utilizando 64 coeficientes DCT para representar cada pose. Tem-se também *NPT* como número poses para o treinamento, *NPR* número de poses para o reconhecimento de faces e ainda os parâmetros *AC* que são a taxa de acerto em percentual, *TT*, tempo de treinamento em segundos, e *TR* que é o tempo de reconhecimento em segundos. Pode-se observar na Tabela 5.4 e Tabela 5.5 que a taxa de acerto tem os valores são próximos após a quantidade de 25 classes. Para as 120 classes o menor tempo de reconhecimento para reconhecimento com uma imagem query foi de 0,177 s e o tempo de treinamento foi de 4800 s.

Tabela 5.4: Resultados da experiência com a RNA usando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de duas a 40 classes, 64 coeficientes DCT – UK.

NPT	NPR	Parâ- Metros	Classes							
			2	5	10	15	20	25	30	40
10	1	AC	100	80	99	95,33	92,50	99,20	96	95,50
		TR	0,0007	0,0047	0,0042	0,0109	0,0103	0,0157	0,0211	0,0308
	2	AC	100	82	99	95,33	93,00	99,60	96,33	96
		TR	0,0007	0,0054	0,0042	0,0067	0,0101	0,0158	0,0214	0,0309
	3	AC	100	84	99	95,33	94	100	97,66	96,25
		TR	0,0007	0,0055	0,0042	0,0070	0,0103	0,0159	0,0178	0,0337
	4	AC	100	86	99	95,33	95	100	97,66	96,75
		TR	0,0007	0,0045	0,0043	0,0070	0,0108	0,0156	0,0180	0,0321
	5	AC	100	88	99	95,33	95	100	97,66	97
		TR	0,0007	0,0078	0,0044	0,0071	0,0101	0,0157	0,0181	0,0314
	6	AC	100	90	99	96,66	96	100	97,66	97,50
		TR	0,0007	0,0019	0,0042	0,0102	0,0131	0,0135	0,0206	0,0272
	7	AC	100	92	100	98,66	98	100	98	98,25
		TR	0,0007	0,0023	0,0070	0,0108	0,0127	0,0137	0,0213	0,0273
	8	AC	100	94	100	99,33	100	100	98,66	98,25
		TR	0,0007	0,0049	0,0072	0,0105	0,0137	0,0161	0,0239	0,0298
	9	AC	100	96	100	99,33	100	100	99,33	98,25
		TR	0,0013	0,0060	0,0077	0,0110	0,0139	0,0105	0,0279	0,0308
	10	AC	100	98	100	100	100	100	99,66	99,75
		TR	0,0015	0,0021	0,0044	0,0070	0,0108	0,0158	0,0196	0,0275
TT			1,33	2,93	8,16	13,87	21,35	33,13	70,17	121,83

Tabela 5.5: Resultados da experiência com a RNA usando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de 50 a 120 classes, 64 coeficientes DCT – UK.

NPT	NPR	Parâ- metros	Classes							
			50	60	70	80	90	100	110	120
10	1	AC	97,80	97	98,42	96,87	97,66	97,20	97,81	96,08
		TR	0,0458	0,0696	0,0703	0,0843	0,1110	0,1299	0,1557	0,177
	2	AC	98,46	97,28	98,71	97,12	97,77	97,80	97,90	96,55
		TR	0,0423	0,0697	0,069	0,0848	0,1106	0,1290	0,1455	0,180
	3	AC	99	97,71	99,28	97,50	98,11	98,20	98,18	97,16
		TR	0,0387	0,0697	0,0701	0,0934	0,1104	0,1290	0,1474	0,176
	4	AC	99	98,42	99,42	98,12	98,77	98,60	98,27	97,66
		TR	0,0420	0,0701	0,0706	0,0897	0,1124	0,1296	0,1621	0,183
	5	AC	99	98,85	99,71	98,25	99,11	98,70	98,63	97,16
		TR	0,0417	0,0698	0,0703	0,0949	0,1108	0,1288	0,1548	0,185
	6	AC	99,40	99	99,85	98,37	99,22	98,90	98,81	98,50
		TR	0,0440	0,0703	0,0704	0,0913	0,1104	0,1289	0,1643	0,183
	7	AC	99,60	99,57	100	98,75	99,33	99,30	98,81	98,83
		TR	0,0443	0,0695	0,0702	0,090	0,1119	0,1292	0,1619	0,172
	8	AC	99,80	99,71	100	99,12	99,44	99,40	99	99,16
		TR	0,0377	0,0695	0,0703	0,0866	0,1109	0,1295	0,1499	0,181
	9	AC	99,80	99,85	100	99,50	99,55	99,70	99,27	99,25
		TR	0,0379	0,0694	0,0704	0,0880	0,1110	0,1305	0,1610	0,181
	10	AC	99,80	100	100	99,62	99,77	99,70	99,81	99,75
		TR	0,0386	0,0703	0,0710	0,0920	0,1114	0,1351	0,1495	0,171
TT			368,35	959,58	1047,0	1174,5	1737,3	3003,1	3476,90	4800

Da Figura 5.10 a 5.12, estão mostrados os gráficos resultantes desta experiência, primeiramente referentes às taxas de acertos e em seguida aos tempos de reconhecimento e de treinamento.

A Figura 5.10 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Pode-se observar nessa figura que para o reconhecimento realizado com 1, 5 e 10 poses a taxa de acerto é menor para 5 classes. Tem uma queda nessa variável em torno da classe 20, cujo valor é 92,50 % para reconhecimento com uma pose, a partir da qual flutua em torno 98%. Observa-se também a menor taxa de acerto é 100 % quando foi realizado os testes para 2 classes com um poses para o reconhecimento.

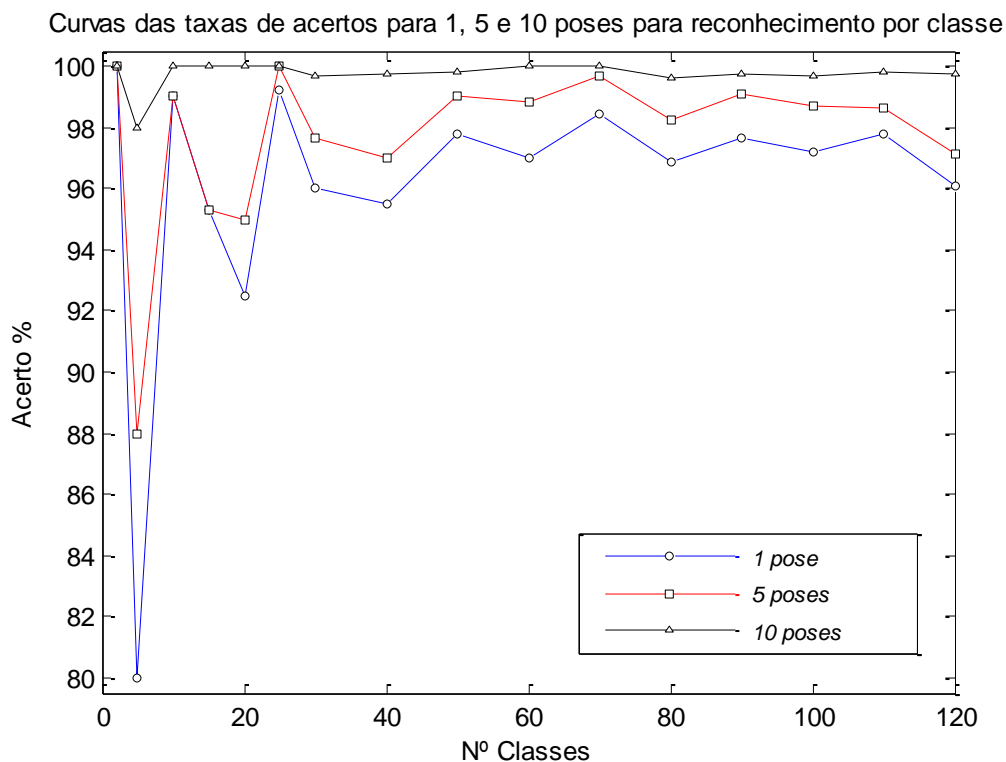


Figura 5.10: Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, 64 coeficientes- UK.

A Figura 5.11 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Observa-se que mesmo com o aumento do número de poses utilizados no reconhecimento, a RNA efetua o reconhecimento em um tempo bastante curto, em torno de 0,180 s para reconhecimento com as 10 poses realizando os testes com 120 classes, Tabela 5.5.

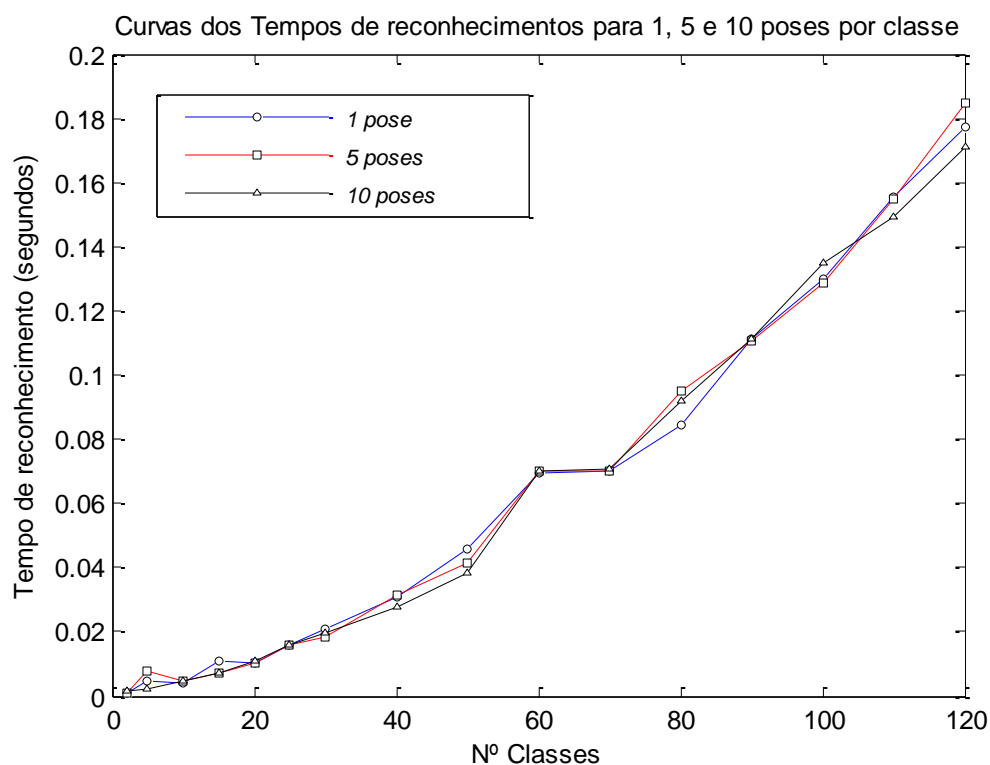


Figura 5.11: Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, cinco e 10 poses, 64 coeficientes – UK.

A Figura 5.12 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente em relação ao número de classes. Observa-se que o treinamento de uma só vez para cada número de classe. O tempo de treinamento varia de 1,33 s a 4800 s para a quantidade de classes 2 e 120, respectivamente.

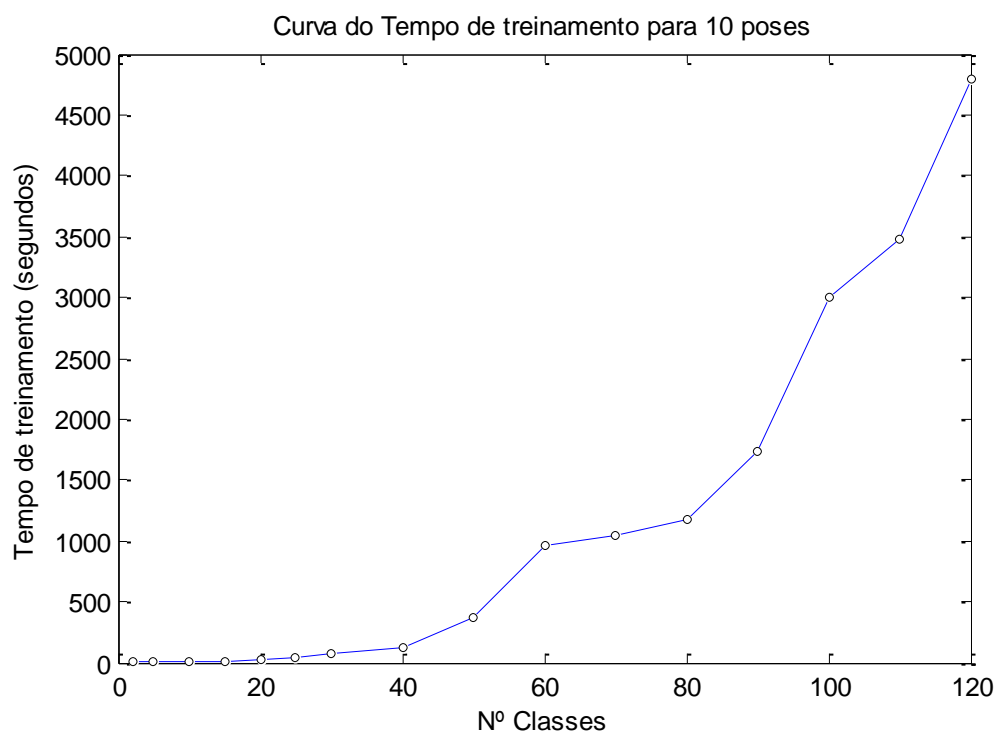


Figura 5.12: Curva do tempo de treinamento da RNA em função do nº de classes para 64 coeficientes por pose – UK.

5.3.2- Experiência com 100 coeficientes DCT, utilizando arquivo UK

Após o cálculo da DCT, utilizaram-se 100 coeficientes mais significativos para compor cada pose.

As Tabelas 5.6 e 5.7 mostram os resultados das experiências com a RNA usando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de face utilizando 100 coeficientes DCT para representar cada pose. Tem-se também *NPT* como número de poses para o treinamento, *NPR* número de poses para o reconhecimento de faces e ainda os parâmetros *AC* que são a taxa de acerto em percentual, *TT*, tempo de treinamento em segundos, e *TR* que é o tempo de reconhecimento em segundos. Pode-se observar na Tabela 5.6 e Tabela 5.7 que a taxa de acerto tem os valores com variação acentuada para as 120 classes. Para as 120 classes o menor tempo de reconhecimento foi em torno de 0,192 s e o tempo de treinamento foi de 5408 s, Tabela 5.6.

Tabela 5.6: Resultados da experiência com a RNA usando 10 poses para o treinamento e de 1 uma a 10 poses para o reconhecimento de 2 duas a 40 classes, 100 coeficientes DCT – UK.

NPT	NPR	Parâ- metros	Classes							
			2	5	10	15	20	25	30	40
10	1	AC	100	100	96	97,33	96	98	98,33	98
		TR	0,0061	0,0052	0,0051	0,0077	0,0117	0,0149	0,0224	0,0322
	2	AC	100	100	96,20	97,33	96,50	98	98,33	98,25
		TR	0,0034	0,0048	0,0047	0,0078	0,0112	0,0149	0,0221	0,0321
	3	AC	100	100	96,50	97,33	96,50	98	98,66	98,50
		TR	0,0041	0,0056	0,0050	0,0077	0,0143	0,0149	0,0222	0,0321
	4	AC	100	100	97	97,33	97	98	99	99,50
		TR	0,0035	0,0055	0,0048	0,0077	0,0142	0,0192	0,0231	0,0322
	5	AC	100	100	97,85	97,33	97	98	99,33	99
		TR	0,0050	0,0053	0,0050	0,0198	0,0142	0,0177	0,0223	0,0325
	6	AC	100	100	98	98	97,50	98,40	99,33	99,75
		TR	0,0033	0,0075	0,0050	0,0104	0,0139	0,0179	0,0222	0,0333
	7	AC	100	100	98,50	98,66	98	99,20	99,33	100
		TR	0,0062	0,0023	0,0047	0,0109	0,0139	0,0177	0,0223	0,0292
	8	AC	100	100	99	99,33	99,50	99,60	99,66	100
		TR	0,0046	0,0022	0,0050	0,0120	0,0144	0,0181	0,0225	0,0298
	9	AC	100	100	99	99,33	99,50	99,60	99,66	100
		TR	0,0043	0,0023	0,0048	0,0078	0,0152	0,0188	0,0230	0,0293
	10	AC	100	100	100	100	100	100	99,66	100
		TR	0,0011	0,0024	0,0054	0,0080	0,0114	0,0152	0,0196	0,0293
TT			1,50	3,20	10	24	35	61,10	98,73	200

Tabela 5.7: Resultados da experiência com a RNA usando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de duas a 120 classes, 100 coeficientes DCT – UK.

NPT	NPR	Parâ- metros	Classes							
			50	60	70	80	90	100	110	120
10	1	AC	96	99	97,57	98,25	96,33	98	98,18	97,16
		TR	0,0480	0,0571	0,0734	0,0932	0,1153	0,1430	0,1618	0,1929
	2	AC	96,50	99,16	97,85	98,25	96,55	98,20	98,63	97,50
		TR	0,0522	0,0583	0,0737	0,0932	0,1153	0,1405	0,1641	0,1911
	3	AC	96,80	99,66	98,28	98,37	97	99,60	99,09	98
		TR	0,0418	0,0565	0,0785	0,0985	0,1164	0,1403	0,1646	0,1935
	4	AC	97	9 9,83	98,57	98,50	97,33	99	99,36	98,60
		TR	0,0418	0,0572	0,0728	0,0941	0,1169	0,1431	0,1624	0,1945
	5	AC	97,40	100	98,71	98,62	97,16	99,10	99,45	98,91
		TR	0,0421	0,0565	0,0724	0,0942	0,1150	0,1417	0,1654	0,1985
	6	AC	98	100	98,85	98,75	98	99,60	99,45	99,08
		TR	0,0465	0,0538	0,0744	0,0960	0,1163	0,1420	0,1627	0,1933
	7	AC	99	100	99,14	99	98,44	99,60	99,54	99,16
		TR	0,0422	0,0563	0,0744	0,0963	0,1150	0,1423	0,1640	0,1939
	8	AC	99,40	100	99,42	99,37	99	99,60	99,54	99,25
		TR	0,0479	0,0569	0,0732	0,0932	0,1151	0,1428	0,1600	0,1946
	9	AC	99,40	100	99,71	99,62	99,11	99,80	99,72	99,58
		TR	0,0486	0,0575	0,0738	0,0931	0,1154	0,1428	0,1622	0,1928
	10	AC	99,40	100	99,85	99,87	99,66	99,90	99,90	99,75
		TR	0,0454	0,0540	0,0705	0,0922	0,1161	0,1416	0,1637	0,1941
TT			397	930	1265	2023,3	2664	4036	4359	5408

Da Figura 5.13 até a 5.15, são mostrados os gráficos resultantes desta experiência, primeiramente referentes às taxas de acertos e, em seguida, aos tempos de reconhecimento e de treinamento.

A Figura 5.13 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Pode-se observar nessa figura que para o reconhecimento realizado com uma poses a taxa de acerto é menor para 10, 20 e 50 classes. Observa-se também que a taxa de acerto é acima de 99 % nos testes realizados para as 120 classes com 10 poses para o reconhecimento.

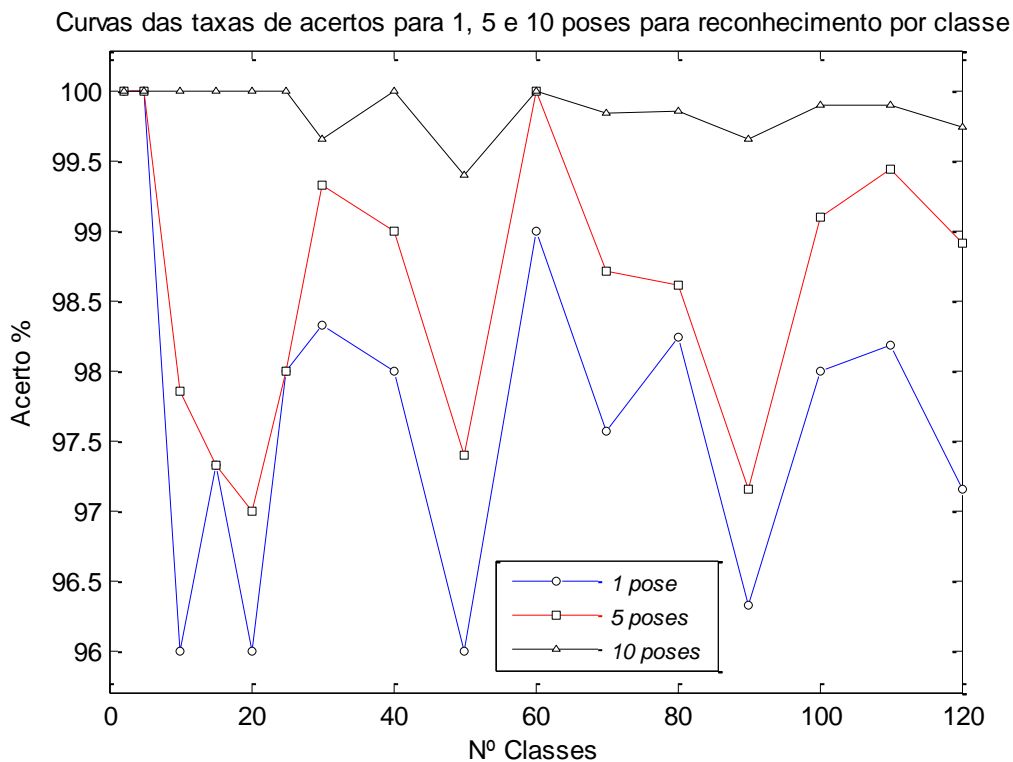


Figura 5.13: Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, 100 coeficientes- UK.

A Figura 5.14 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Observa-se que mesmo com o aumento do número de poses utilizados no reconhecimento, a RNA efetua o reconhecimento em um tempo bastante curto, em torno de 0,1930 s para reconhecimento com as 10 poses realizando os testes com 120 classes, Tabela 5.7.

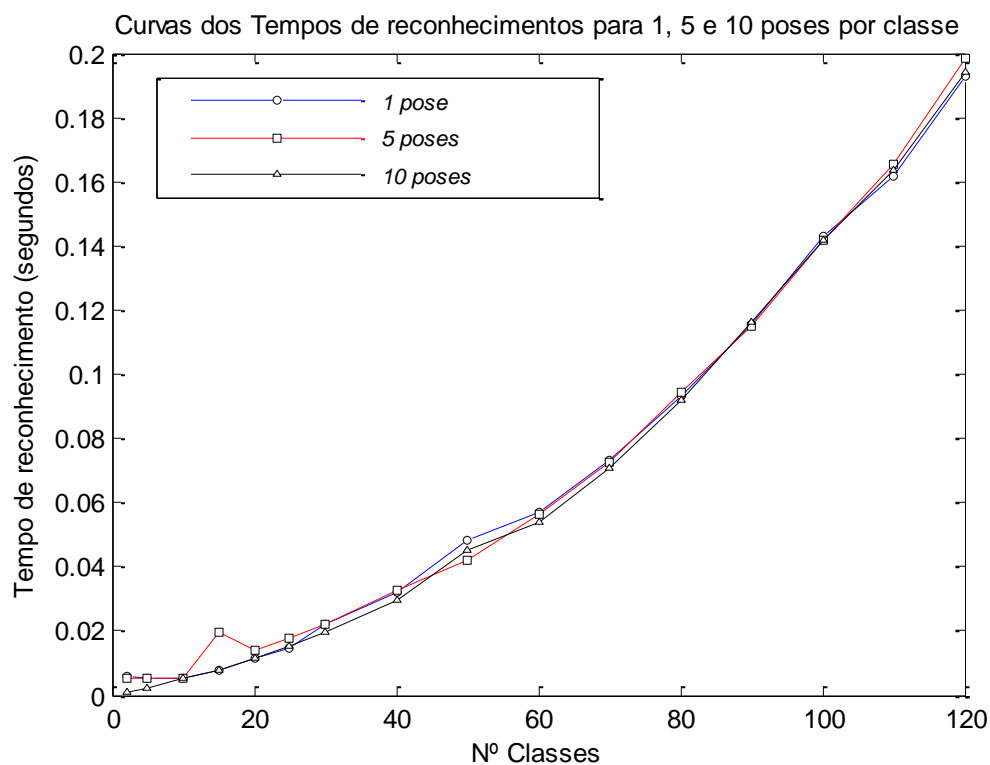


Figura 5.14: Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, cinco e 10 poses, 100 coeficientes – UK.

A Figura 5.15 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente em relação ao número de classes. Observa-se que o treinamento de uma só vez para cada número de classe. O tempo de treinamento varia de 1,50 s a 5408 s para a quantidade de classes 2 e 120, respectivamente, Tabela 5.6 e Tabela 5.7.

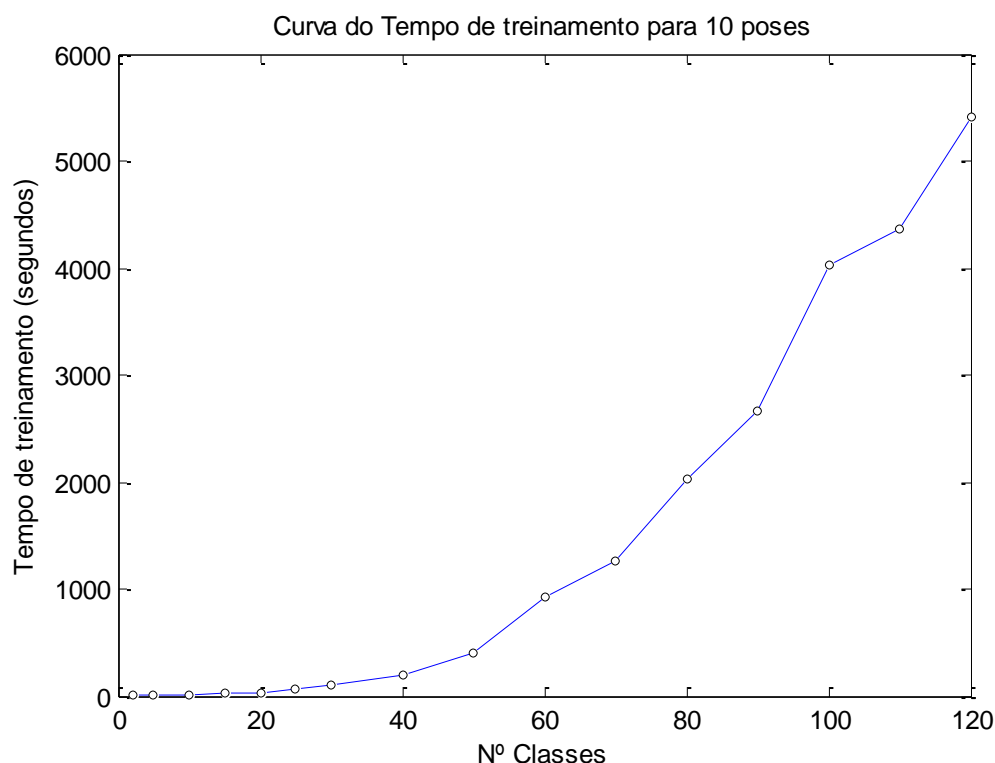


Figura 5.15: Curva do tempo de treinamento da RNA em função do nº de classes para 100 coeficientes por pose – UK.

5.3.3- Experiência com 256 coeficientes DCT, utilizando arquivo UK

Após o cálculo da DCT para compor cada pose forma utilizaram-se 256 coeficientes mais significativos.

As Tabelas 5.8 e 5.9 mostram os resultados com a RNA utilizando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de face usando 256 coeficientes DCT para representar cada pose. Tem-se também NPT que é o número de poses para o treinamento, NPR é o número de poses para o reconhecimento de faces e ainda os parâmetros AC que é a taxa de acerto em percentual, TT , é o tempo de treinamento em segundos, e TR é o tempo de reconhecimento em segundos. Pode-se observar na Tabela 5.8 e Tabela 5.9 que a taxa de acerto tem os valores com variação acentuada para as 120 classes. Para as 120 classes o menor tempo de reconhecimento foi em torno de 0,220 s e o tempo de treinamento foi de 15132 s, Tabela 5.9.

Tabela 5.8: Resultados da experiência com a RNA usando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de duas a 40 classes, 256 coeficientes DCT – UK.

NPT	NPR	Parâ- metros	Classes							
			2	5	10	15	20	25	30	40
10	1	AC	100	98	100	96,66	96,50	96,20	96	96
		TR	0,0041	0,0040	0,0080	0,0166	0,0205	0,0304	0,0317	0,0448
	2	AC	100	98	100	96,66	97,50	96,60	96	96,50
		TR	0,0049	0,0030	0,0082	0,0174	0,0212	0,0269	0,0322	0,0461
	3	AC	100	100	100	96,66	97,80	96,40	96,66	96,25
		TR	0,0040	0,0040	0,0120	0,0165	0,0174	0,0260	0,0291	0,0443
	4	AC	100	100	100	96,66	97,50	97,60	97,33	96,75
		TR	0,0044	0,0037	0,0134	0,0164	0,0177	0,0265	0,0363	0,0439
	5	AC	100	100	100	98	97,50	98,40	98	96,75
		TR	0,0151	0,0041	0,0134	0,0166	0,0293	0,0300	0,0346	0,0443
	6	AC	100	100	100	99,33	97,50	98,80	98,33	97
		TR	0,0040	0,0048	0,0119	0,0166	0,0180	0,0260	0,0293	0,0445
	7	AC	100	100	100	99,33	98	99,20	99	98
		TR	0,0049	0,0064	0,0175	0,0167	0,0222	0,0258	0,0345	0,0447
	8	AC	100	100	100	100	98,50	99,60	99,66	99
		TR	0,0045	0,0078	0,0171	0,0171	0,0179	0,0288	0,0352	0,0448
	9	AC	100	100	100	100	99,50	99,60	99,66	99,25
		TR	0,0059	0,0082	0,0123	0,0173	0,0182	0,0272	0,0289	0,0463
	10	AC	100	100	100	100	100	100	100	99,75
		TR	0,0017	0,0048	0,0093	0,0142	0,0184	0,0226	0,0365	0,0412
TT			2,04	5,06	18,71	48,32	91,95	173,09	278,67	791

Tabela 5.9: Resultados da experiência com a RNA usando 10 poses para o treinamento e de uma a 10 poses para o reconhecimento de duas a 120 classes, 256 coeficientes DCT – UK

NPT	NPR	Parâ- metros	Classes							
			50	60	70	80	90	100	110	120
10	1	AC	96,20	97,16	98,57	97,50	96,88	97,70	96	96,08
		TR	0,0561	0,0799	0,1037	0,1150	0,1307	0,1605	0,1969	0,2157
	2	AC	96,80	97,33	98,57	97,62	97,22	97,60	96,18	96,50
		TR	0,060	0,0826	0,1065	0,1153	0,1434	0,1595	0,1966	0,2254
	3	AC	96,80	97,50	98,85	97,75	97,66	97,80	96,90	97,08
		TR	0,0615	0,0852	0,1030	0,1164	0,1458	0,1597	0,1962	0,2251
	4	AC	97,60	97,83	99	98	98,11	98	97,36	97,75
		TR	0,0591	0,0818	0,1006	0,1157	0,1452	0,1502	0,1963	0,2255
	5	AC	98	98,33	99,28	98,25	98,44	98,50	97,80	98,25
		TR	0,0597	0,0818	0,1035	0,1159	0,1354	0,1602	0,1968	0,2262
	6	AC	98,40	98,33	99,42	98,62	98,66	99,10	98,09	98,75
		TR	0,0558	0,0719	0,0947	0,1157	0,1505	0,1599	0,1977	0,2264
	7	AC	98,80	99,16	99,71	99,12	98,88	99,50	98,63	99,08
		TR	0,0591	0,0843	0,0814	0,1158	0,1402	0,1600	0,1971	0,2260
	8	AC	99,20	99,66	99,85	99,50	99,33	99,70	99,09	99,33
		TR	0,0589	0,0766	0,0943	0,1157	0,1578	0,1694	0,1965	0,2252
	9	AC	99,60	98,83	99,85	99,62	99,77	99,80	99,54	99,41
		TR	0,0589	0,0811	0,1018	0,1159	0,1500	0,1596	0,1963	0,2246
	10	AC	99,60	100	99,85	99,75	99,88	100	99,81	99,75
		TR	0,0590	0,0793	0,1006	0,1161	0,1483	0,1612	0,1971	0,2258
TT			963,59	1777	3437	5310	7114	10840	14465	15132

Das Figuras 5.16 a 5.18, são mostrados os gráficos resultantes desta experiência, primeiramente referentes às taxas de acertos e, em seguida, aos tempos de reconhecimento e de treinamento.

A Figura 5.16 mostra o gráfico resultante dessa experiência referente às taxas de acertos. Pode-se observar nessa figura que para o reconhecimento realizado com uma poses a taxa de acerto decresce a partir da classe 5, com menores valores entre as classes 25 e 40, voltando a esses valores a partir do emprego de 110 a 120 classes. é menor para 10, 20 e 50 classes. Neste caso recomenda-se o reconhecimento com 10 poses, pois os resultados foram acima 99 % nos testes realizados para as 120 classes com 10 poses.

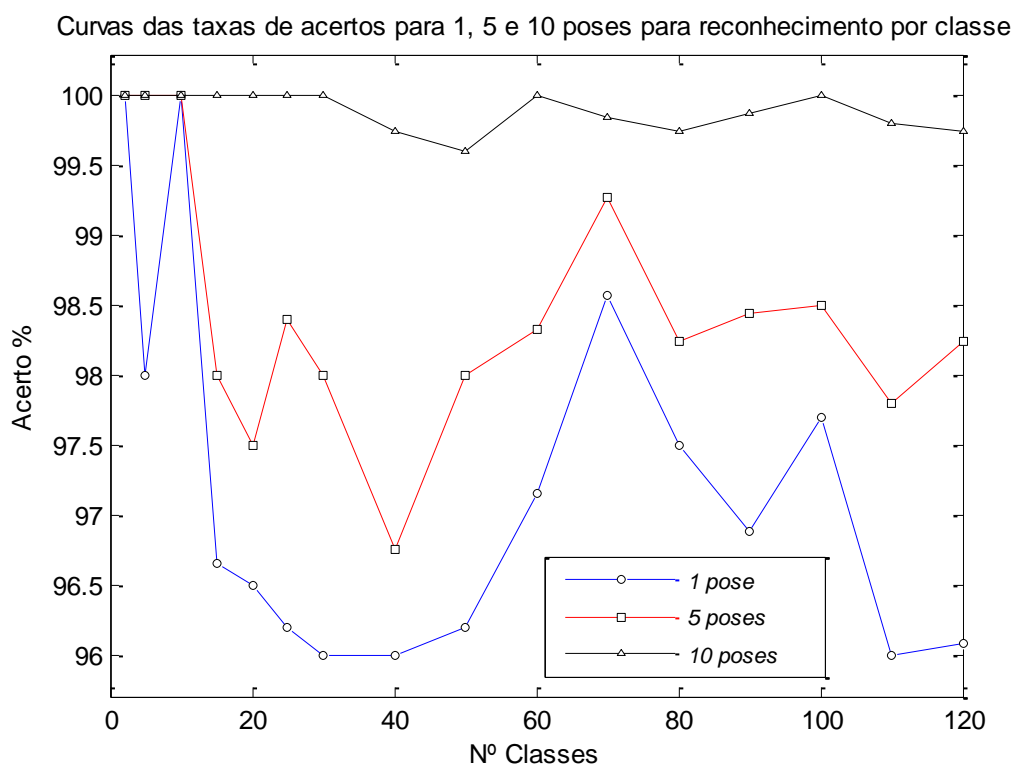


Figura 5.16: Curvas das taxas de acertos da RNA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, 256 coeficientes- UK.

A Figura 5.17 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Observa-se que mesmo com o aumento do número de poses utilizados no reconhecimento, a RNA efetua o reconhecimento em um

tempo bastante curto, em torno de 0,220 s para reconhecimento com as 10 poses realizando os testes com 120 classes, Tabela 5.9.

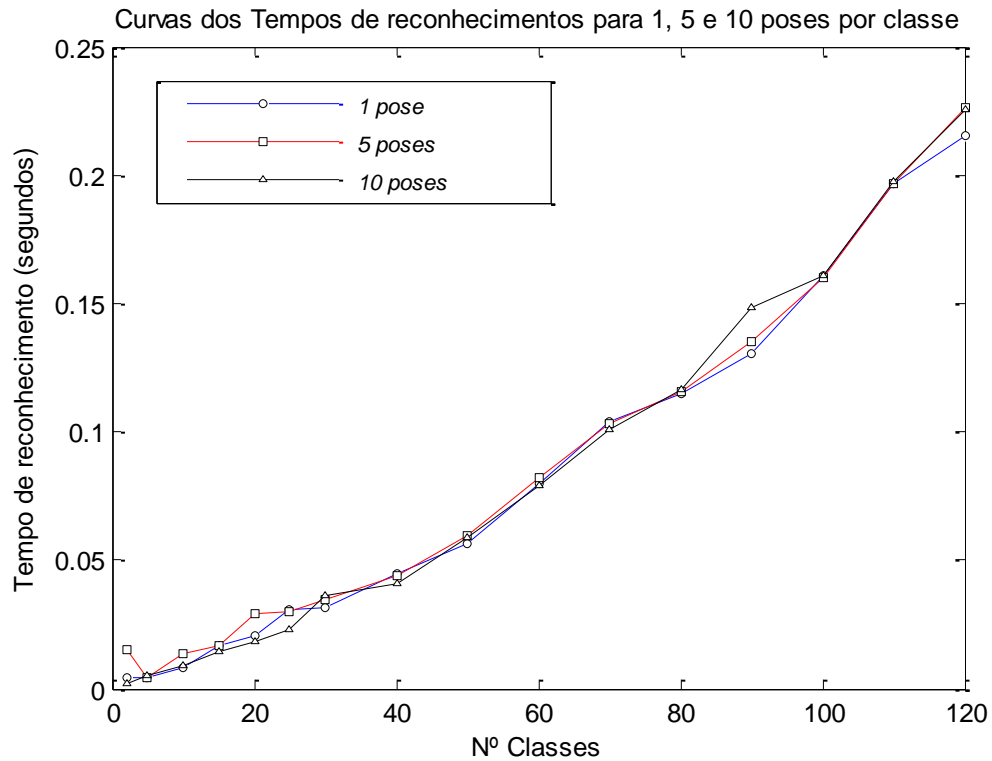


Figura 5.17: Curvas dos tempos de reconhecimentos da RNA em função do nº de classes utilizando uma, cinco poses, 256 coeficientes – UK.

A Figura 5.18 mostra o gráfico resultante dessa experiência referente aos tempos de treinamentos. Essa variável tem um comportamento ascendente em relação ao número de classes. Observa-se que o treinamento de uma só vez para cada número de classe. O tempo de treinamento varia de 2,04 s a 15132 s para a quantidade de classes 2 e 120, respectivamente, Tabela 5.8 e Tabela 5.9.

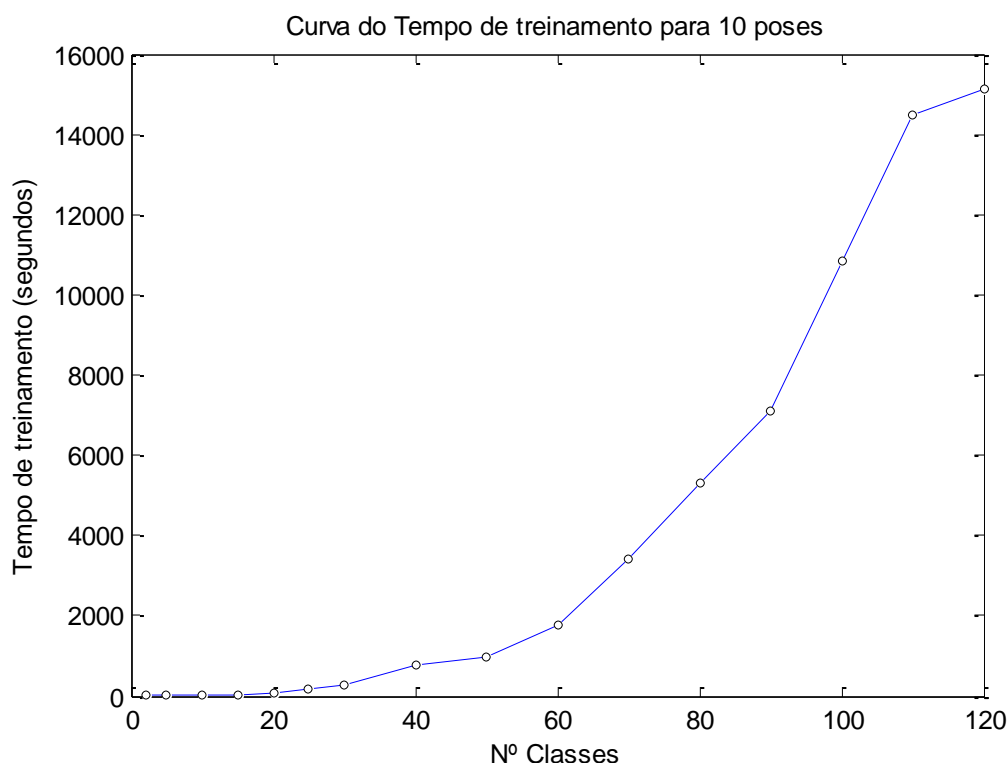


Figura 5.18: Curva do tempo de treinamento da RNA em função do nº de classes para 256 coeficientes por pose – UK.

5.4- Considerações finais deste Capítulo

Nos testes realizados a RNA apresentou uma taxa de acerto em torno de 98 % . O arquivo UK tem o dobro de imagens do ORL e, mesmo assim, a taxa de acerto para ambos os arquivos foram muito próximas. Mesmo com o aumento vertiginoso do tempo de treinamento em relação ao aumento do número de classes, o mais importante, é que a etapa de reconhecimento, foi realizada em um tempo bastante pequeno, mesmo com o acréscimo da quantidade de classes analisadas.

Novas disciplinas nas fronteiras da biologia, em especial, a genética do comportamento, a psicologia da evolução e a neurociência cognitiva, estão estabelecendo uma ponte a entre natureza e a sociedade na forma de uma compreensão científica da natureza humana. A face humana reflete esta evolução, seja no componente temporal, seja no componente emocional. A identificação das pessoas pela sociedade é fundamental para o equilíbrio existencial, já que existe uma dependência mútua,

necessária a sobrevivência da espécie. A RNA é um dos algoritmos poderosos que vem contribuindo neste processo.

O próximo capítulo mostra os testes realizados com a análise dos componentes principais.

CAPÍTULO VI

UTILIZAÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS NO RECONHECIMENTO DE FACES

6.1- Introdução

O algoritmo da análise de componentes principais (PCA), neste trabalho, consistiu em executar o processamento computacional alimentados por dados dos arquivos ORL e UK, especificados nas seções 3.3.3 do Capítulo III. De acordo com a seção 2.4 do Capítulo II, pode-se entender o princípio de funcionamento da PCA.

A codificação do arquivo no processo de treinamento foi feita pela PCA utilizando todas as imagens os arquivos ORL e UK, ou seja, o treinamento é foi realizado para todas as classes. O algoritmo não treina a parte de um todo.

O que se pretendeu foi que o algoritmo identificasse a que classe pertencia a pose ou o grupo destas quando fosse considerada uma ou mais poses (imagens query), de qualquer uma das classes pertencentes a esses arquivos. A Figura 2.9 do Capítulo II mostra um diagrama adaptado de como se realiza o processamento da PCA.

A PCA executou esta tarefa promovendo a concentração dos dados resultantes em componentes principais após as transformações lineares dos mesmos.

Este capítulo apresenta os testes realizados com a análise de componentes principais.

6.2- Experiência com a Análise de Componentes Principais (PCA) utilizando o arquivo ORL

Para a experiência com a PCA executaram-se os procedimentos apresentados na seção 3.3.3.1 do Capítulo III.

No primeiro momento, para o treinamento, foram processados os dados para a construção da matriz $MN \times C$, onde M é o número de linhas, N é o número de colunas e C é o número de classes cujo tempo foi de 39 segundos. Na sequência iniciou-se a fase de reconhecimento que se baseia no acesso às bases codificadas para buscar no arquivo-treinamento características de uma pose retiradas do arquivo-reconhecimento. O tempo para fazer essa operação depende do número de classes estipuladas e também dos diferentes números de poses a serem procuradas. Assim, para observar a variação do tempo de reconhecimento com o aumento do número de poses processadas, fez-se a experiência também em etapas utilizando-se de uma a 10 poses para o reconhecimento. É importante se lembrar que, considerando o número de poses destinadas ao reconhecimento, esta quantidade foi utilizada em todas as classe existentes no arquivo.

Na Tabela 6.1, estão os resultados desta experiência onde NPR é o número de poses utilizadas por vez para o reconhecimento de faces em segundos, AC é a taxa de acerto em percentual e TR é o tempo de reconhecimento em segundos. Pode-se observar na Tabela 6.1 que a taxa de acerto tem os valores de 100% para quase todas as quantidades de classes testadas. Para 40 classes a taxa de acerto tem uma ligeira queda no entanto fica acima de 99%. Já o tempo de reconhecimento é proibitivo, chegando a 15047,42 s para 10 poses para o reconhecimento.

Tabela 6.1: Resultado da experiência com a PCA utilizando de uma a 10 poses para o reconhecimento de faces – arquivo ORL.

NP	Parâmetros	Classe							
		2	5	10	15	20	25	30	40
1	AC (%)	100	100	100	100	100	100	100	99
	TR (s)	23,15	73,25	100,05	168,25	267,12	341	412,1	575,01
2	AC (%)	100	100	100	100	100	100	100	98,75
	TR (s)	33,95	84,95	170,05	258,75	347,52	434	521,99	695,32
3	AC (%)	100	100	100	100	100	100	100	99,16
	TR (s)	49,01	128,05	256,17	389,16	522,15	652,50	782,86	1042,70
4	AC (%)	100	100	100	100	100	100	100	99,37
	TR (s)	64,01	170,05	342,24	519,62	697,00	873	1050,46	1390,55
5	AC (%)	100	100	100	100	100	100	100	99,50
	TR (s)	89,01	215,74	428,72	649,48	870,25	1086,0	1301,78	1733,57
6	AC (%)	100	100	100	100	100	100	100	99,58
	TR (s)	108,01	259,64	516,79	780,71	1044,64	1310,7	1576,68	2088,62
7	AC (%)	100	100	100	100	100	100	100	99,64
	TR (s)	129,01	309,97	600,56	903,05	1205,55	1513,5	1821,51	2428,84
8	AC (%)	100	100	100	100	100	100	100	99,68
	TR (s)	145,42	354,97	685,71	1029,80	1373,79	1736,5	2099,29	9239,57
9	AC (%)	100	100	100	100	100	100	100	99,72
	TR (s)	153,02	362,45	694,28	1102,16	1459,37	2523,08	4367,29	10648,69
10	AC (%)	100	100	100	100	100	100	100	99,72
	TR (s)	166,02	407,98	777,74	1236,4	1695,05	3685,9	7676,67	15047,42

As Figuras 6.1 e 6.2 mostram os gráficos resultantes desta experiência, primeiro referentes às taxas de acertos e, em seguida, os tempos de reconhecimento. O treinamento é realizado de uma vez só para todo o arquivo. Neste caso com a duração de 39 s.

Pode-se observar na Figura 6.1 que para o reconhecimento realizado com uma pose a 10 poses a taxa de acerto é muito alta acima 99% para 2, 5, 10, 15, 20, 25, 30 e 40 classes. Observa-se também a menor taxa de acerto é 99,15 % quando foi realizado os testes para 40 classes com 10 poses para o reconhecimento.

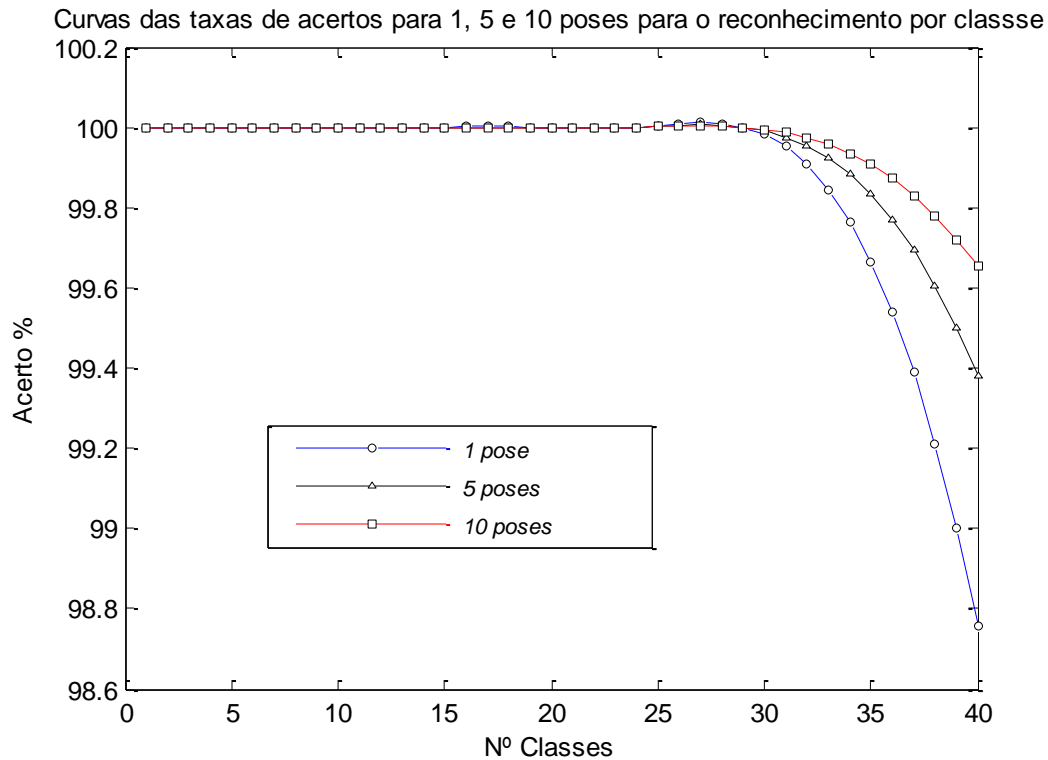


Figura 6.1: Curvas das taxas de acertos da PCA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, arquivo ORL.

A Figura 6.2 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Observa-se que mesmo com o aumento do número de poses utilizados no reconhecimento a PCA fica extremamente lenta no reconhecimento resultando em um tempo muito alto. Observa-se que para o teste com 2 classes o tempo reconhecimento com um pose de 23,15 s e para 10 poses é 166,02 s.

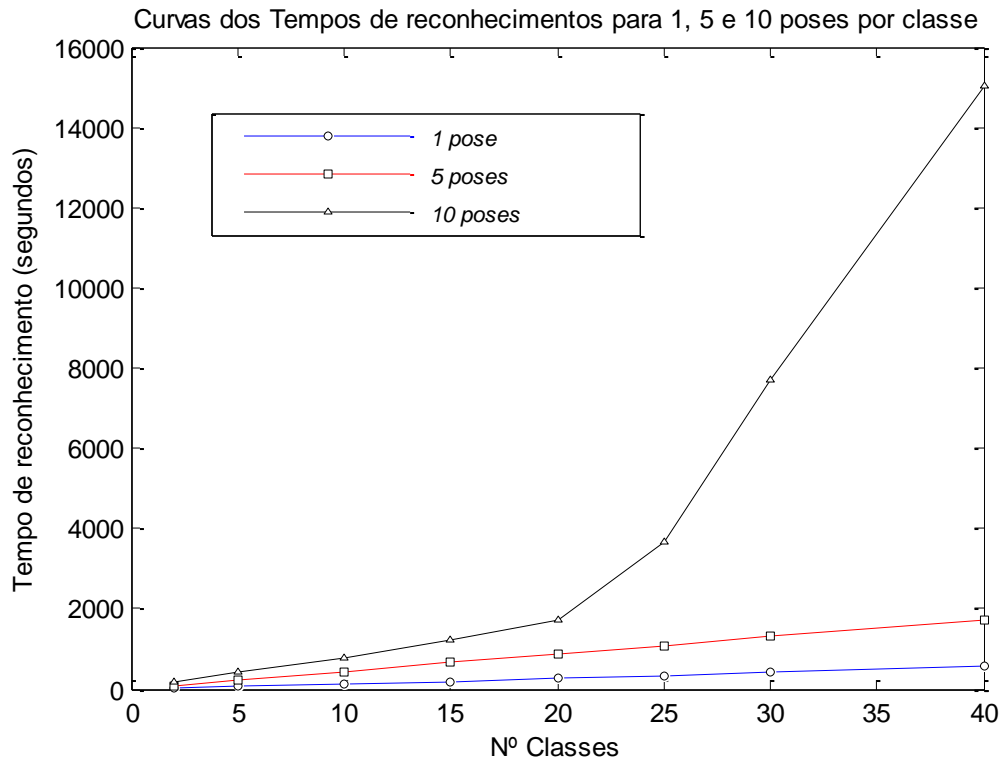


Figura 6.2: Curvas dos tempos de reconhecimentos da PCA em função do n° de classes utilizando uma, cinco e 10 poses, arquivo ORL.

6.3- Experiência com a Análise de Componentes Principais (PCA) utilizando arquivo UK

Para a experiência com a PCA, foram executados os procedimentos apresentados na seção 3.3.3.2 do capítulo III.

No primeiro momento, para o treinamento, processaram-se os dados UK para a construção da matriz $MN \times F$, onde M é o número de linhas, N é o número de colunas e F é o número de classes cujo tempo foi de 216 segundos. Na sequência, iniciou-se a fase de reconhecimento que se baseia no acesso às bases codificadas para buscar no arquivo-treinamento características de uma pose retiradas do arquivo-reconhecimento. O tempo para fazer essa operação depende do número de classes estipuladas e também dos diferentes números de poses a serem procuradas. Assim, para observar a variação do tempo de reconhecimento com o aumento do número de poses processadas, fez-se a experiência também em etapas utilizando de uma a 10 poses para o reconhecimento.

É importante lembrar que, considerando um número de poses destinadas ao reconhecimento, esta quantidade foi utilizada em todas as classe existentes no arquivo.

A Tabela 6.2, mostra os resultados desta experiência onde *NPR* é o número de poses utilizadas por vez para o reconhecimento de faces em segundos, *AC* é a taxa de acerto em percentual e *TR* é o tempo de reconhecimento em segundos. Pode-se observar na Tabela 6.2 que a taxa de acerto tem os valores acima de 99% para todas as quantidades de classes testadas. Para 40 classes a taxa de acerto tem uma ligeira queda no entanto fica acima de 99%. Já o tempo de reconhecimento é proibitivo, chegando a 16676,97 s para 10 poses para o reconhecimento.

Tabela 6.2: Resultado da experiência com a PCA utilizando de uma a 10 poses para o reconhecimento de faces - arquivo UK.

NP	Parâmetros	Classes							
		2	5	10	15	20	25	30	40
1	AC (%)	100	100	100	100	100	100	100	100
	TR (s)	42,26	92,25	120,54	186,25	291,62	372,27	469,21	608,21
2	AC (%)	100	100	100	100	100	100	100	99,50
	TR (s)	55,53	106,23	192,57	280,29	369,81	459,87	542,96	716,28
3	AC (%)	100	100	100	100	100	100	100	98,75
	TR (s)	84,64	135,12	225,24	302,10	391,20	480,32	574,56	741,28
4	AC (%)	100	100	100	100	100	100	100	99,50
	TR (s)	123,96	194,58	345,89	478,11	611,85	741	871	1131,54
5	AC (%)	100	100	100	100	100	100	100	99,50
	TR (s)	363,33	469,23	641,56	818,78	996,45	1172,0	1349,85	1689,26
6	AC (%)	100	100	100	100	100	100	100	99,50
	TR (s)	442,62	567,54	780,89	980,25	1200,87	1417,29	1627,79	2061,25
7	AC (%)	100	100	100	100	100	100	100	99,75
	TR (s)	518,61	668,13	925,81	1189,53	1453,31	1719,26	1986,64	2498,45
8	AC (%)	100	100	100	100	100	100	100	99,64
	TR (s)	558,15	738,61	1047,25	1350,57	1658,23	1976,86	2306,73	2946,59
9	AC (%)	100	100	100	100	100	100	100	99,50
	TR (s)	590,10	805,19	1135,17	1465,89	1795,58	2158,47	2512,34	11061,39
10	AC (%)	100	100	100	100	100	100	99,75	99,15
	TR (s)	665,68	906,52	1276,39	1816,41	2276,59	3276,15	8276,96	16676,97

Nas Figuras 6.3 e 6.4, são mostrados os gráficos resultantes desta experiência, primeiramente referentes às taxas de acertos e, em seguida, aos tempos de reconhecimento. O treinamento é realizado de uma vez só para todo o arquivo. Neste caso com a duração de 216 s.

Pode-se observar na Figura 6.3 que para o reconhecimento realizado com uma a 10 poses a taxa de acerto é muito alta acima 99%. Observa-se também a taxa de acerto

tem uma ligeira redução nos testes feitos com 30 a 40 com 10 poses para o reconhecimento.

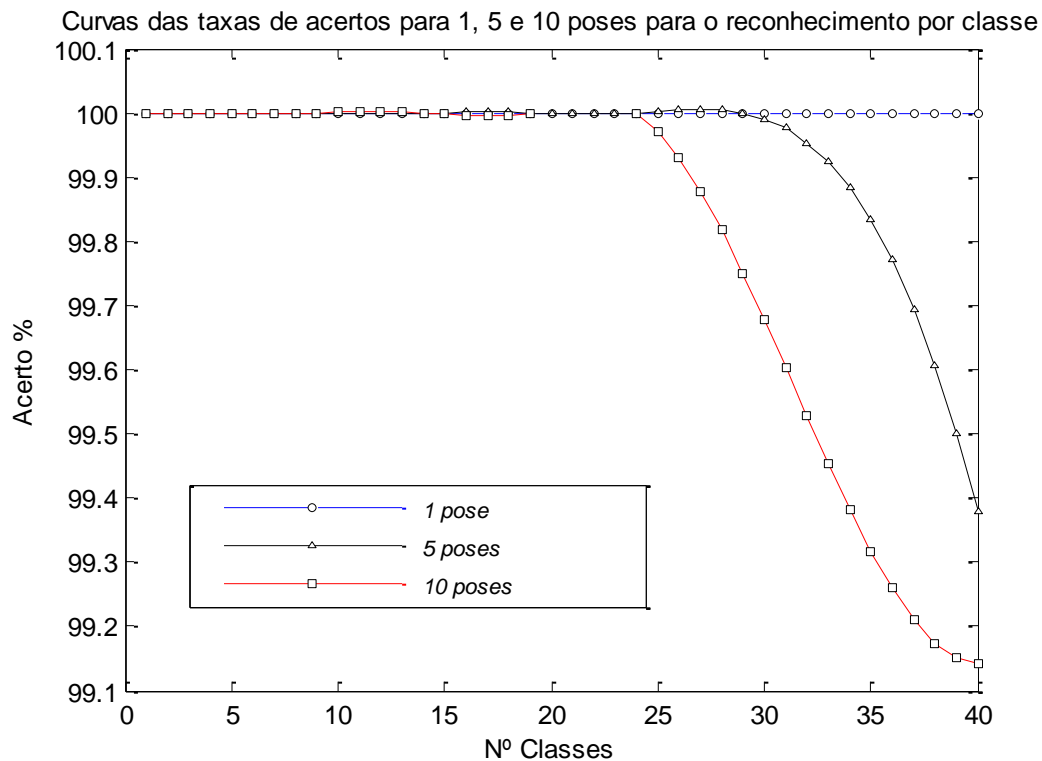


Figura 6.3: Curvas das taxas de acertos da PCA em função do nº de classes utilizando uma, cinco e 10 poses para o reconhecimento, arquivo UK.

A Figura 6.4 mostra o gráfico resultante dessa experiência referente aos tempos de reconhecimentos. Essa variável tem um comportamento ascendente progressivo não linear com o aumento do número de classes. Observa-se que mesmo com o aumento do número de poses utilizados no reconhecimento a PCA fica extremamente lenta no reconhecimento resultando em um tempo muito alto. Observa-se que para o teste somente com 2 classes o tempo reconhecimento com um pose de 42,26 s e para 10 poses é 665,68 s.

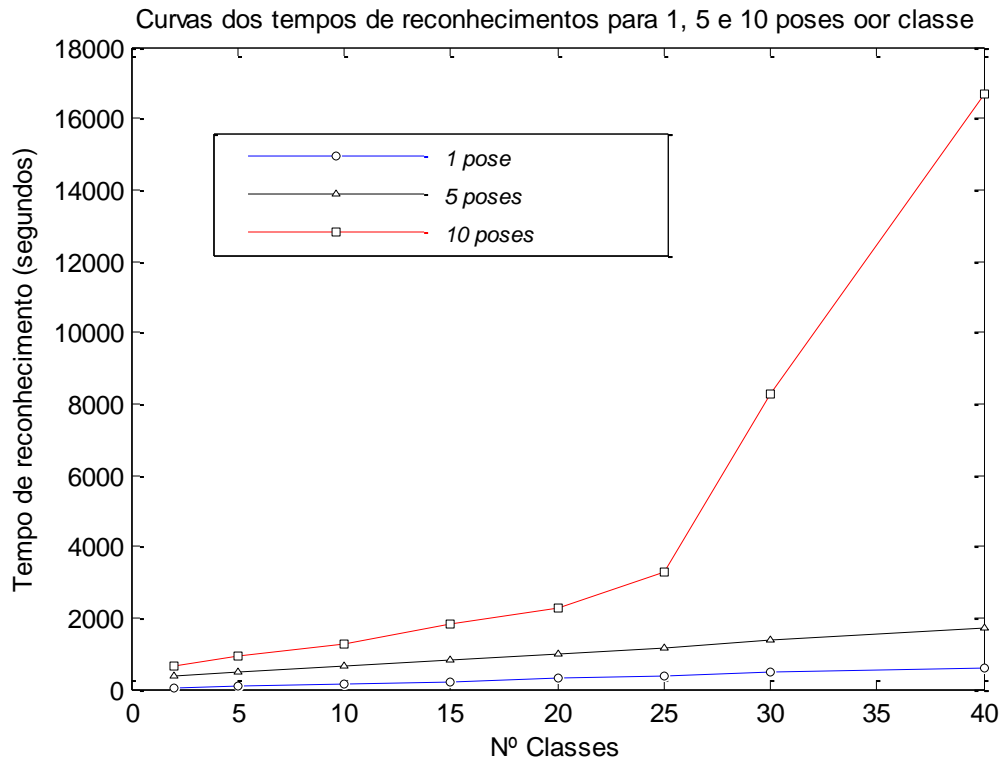


Figura 6.4: Curvas dos tempos de reconhecimentos da PCA em função do nº de classes utilizando uma, cinco e 10 poses, arquivo UK.

6.4- Considerações Finais Deste Capítulo

A PCA tem uma taxa de acerto muito alta e estável, média em torno de 99 %. O arquivo UK tem o dobro das imagens do ORL e, mesmo assim, as taxas de acertos para ambos os arquivos foram muito próximas. Mesmo o tempo de treinamento sendo pequeno em relação ao aumento do número de classes, a etapa mais importante, é o reconhecimento, realizada em um tempo muito grande mesmo com poucas classes e poucas poses utilizadas para o reconhecimento. Pelo exposto na seção 3.3.3.2 do Capítulo III, a experiência com 120 classes não foi concluída, pois o tempo necessário para reconhecer uma pose dentre as 120 classes muito maior do o limite da prática operacional.

O reconhecimento utilizando a PCA é para casos que se dispõe de tempo para o reconhecimento de faces e de máquinas poderosas para o processamento em termos de velocidade e paginação, ou seja, grande memória virtual.

Mesmo assim, a PCA tem sido muito utilizadas nas pesquisas sobre reconhecimentos de faces. Embora naturalmente ocorram erros nas sucessivas tentativas que o processo requer, não se deve desconsiderá-la, pois é muito grande o desafio a ser enfrentado em função do rumo quase sempre gritantemente injusto que acompanha a transformação da evolução humana.

O próximo capítulo mostra a comparação dos resultados obtidos entre os algoritmos Genético, Rede Neural Artificial e Análise de Componentes Principais.

CAPÍTULO VII

COMPARAÇÃO DOS RESULTADOS OBTIDOS ENTRE OS ALGORITMOS GENÉTICO, REDE NEURAL E ANÁLISE DE COMPONENTES PRINCIPAIS

7.1- Introdução

As implementações das etapas de reconhecimento de faces deste trabalho foram realizadas utilizando os Algoritmos Genético (AG), Rede Neural Artificial (RNA) e Análise de Componentes Principais (PCA).

Este capítulo mostra o resumo dos testes realizados com cada um dos três algoritmos e depois a comparação dos resultados obtidos apontando as vantagens de cada um deles.

7.2 - Resumo da experiência com o Algoritmo Genético

Nesta seção são mostrados os resumos do processamento com o AG para os arquivos ORL e UK.

7.2.1- Resumo da Experiência com o Algoritmo Genético para o arquivo ORL

Esta seção apresenta o resumo dos resultados obtidos utilizando o AG considerando apenas os valores relacionados ao processamento para o reconhecimento de faces com cinco poses de cada classe mostrados nas Tabelas 4.1 a 4.4 do Capítulo IV, arquivo ORL.

A Tabela 7.1 mostra os dados relacionados a 64, 154, 305 e 616 coeficientes DCT representados pelo número de coeficientes (variável *NC*). Estes dados são relacionados ao reconhecimento com cinco poses. Observa-se que quanto maior o número de coeficientes a tendência foi de maior a taxa de acerto. Os tempos de processamento são obviamente crescentes com o número de classes.

Tabela 7.1: Resultados obtidos do reconhecimento da face, AG, cinco poses para reconhecimento, arquivo ORL para 2 a 40 classes.

NC *	Parâmetros	Classes							
		2	5	10	15	20	25	30	40
64	Acerto (%)	100	100	100	90	87	84	84	80,50
	T,Trein, (s)	2,841	6,534	13,531	20,191	27,534	33,594	40,28	53,891
	T, Rec, (s),	0,0131	0,0317	0,0634	0,143	0,223	0,359	0,467	0,816
154	Acerto (%)	100	96	96	94,66	92	92	88	83
	T,Trein, (s)	2,99	8,25	16,33	24,57	32,71	40,97	49,45	65,862
	T, Rec, (s),	0,03	0,05	0,09	0,20	0,35	0,52	0,76	1,293
308	Acerto (%)	100	100	98	94,66	86	91,20	89,33	86
	T,Trein, (s)	4,39	10,76	21,50	31,63	42,23	53,60	64,24	84,83
	T, Rec, (s),	0,03	0,05	0,16	0,29	0,50	0,75	1,08	1,889
616	Acerto (%)	100	96,00	96,00	96,00	94,00	92,00	91,33	93,00
	T,Trein, (s)	3,98	9,96	20,59	30,55	41,17	51,13	61,76	128,16
	T, Rec, (s),	0,47	0,84	1,16	1,49	1,91	2,47	2,79	3,032

* número de coeficientes

As Figuras 7.1 e 7.2 mostram os gráficos comparativos entre os diferentes números de coeficientes empregados nesta experiência. Primeiro é ilustrado a comparação entre as taxas de acerto e depois entre os tempos de reconhecimentos, conforme a Tabela 7.1.

A Figura 7.1 mostra as curvas das taxas de acerto. Observa-se que entre 2 a 10 classes a taxa foi de 100 % a partir da qual caiu acentuadamente chegando a 84,50 % com 40 classes. O melhor comportamento desta variável foi para 616 coeficientes cujos valores das taxas de acerto ficaram acima de 91 % .

Comparação das taxas de acertos usando 5 poses no reconhecimento por cada classe

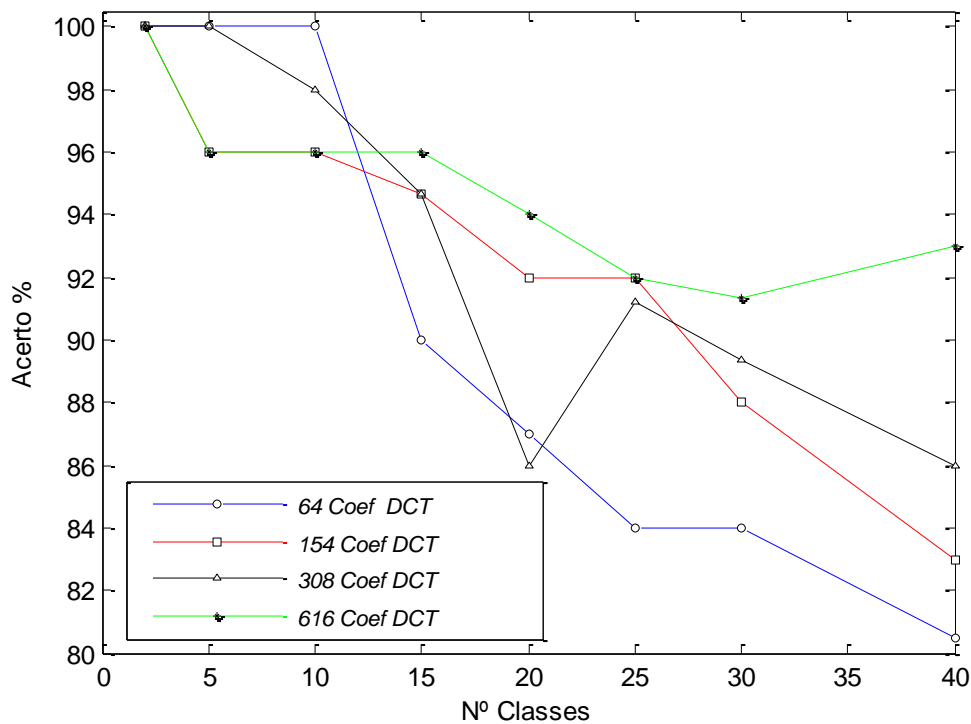


Figura 7.1: Comparação entre as taxas de acerto do AG considerando 64, 154, 308 e 616 coeficientes, arquivo ORL.

A Figura 7.2 mostra as curvas dos tempos de reconhecimento. Observa-se que a partir do processamento com 15 classes o tempo aumenta mais rapidamente. O maior tempo foi para 616 coeficientes cujos valores para 2 e 40 classes são 0,47 s e 3,032 s, respectivamente.

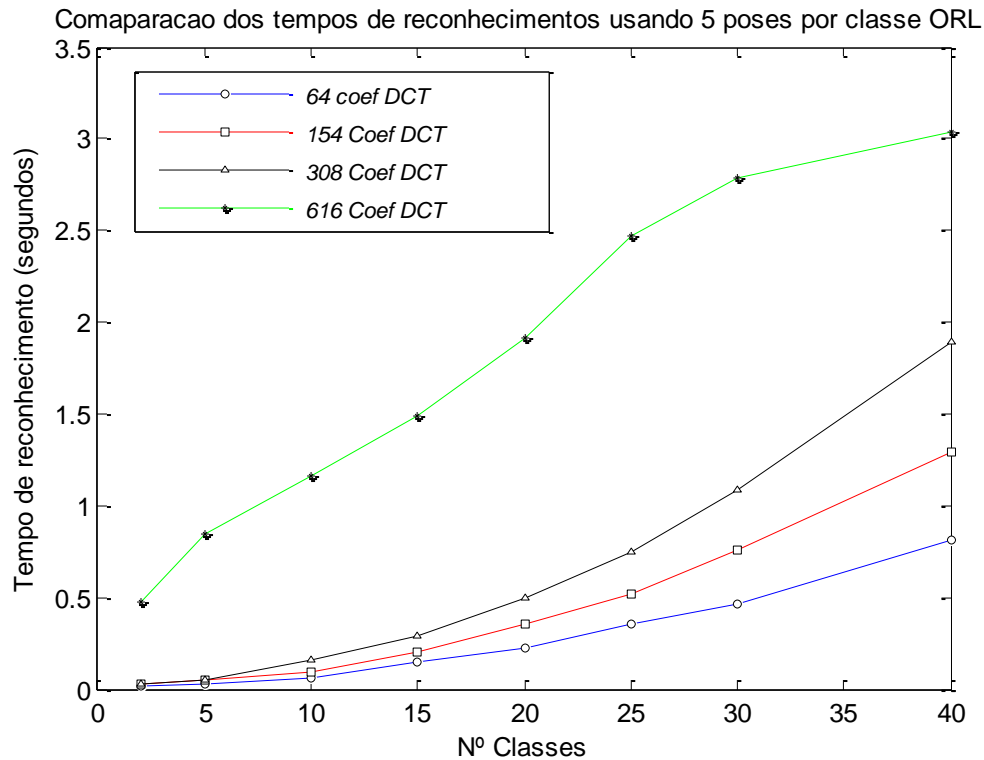


Figura 7.2: Comparação entre os tempos de reconhecimentos do AG considerando 64, 154, 308 e 616 coeficientes, arquivo ORL.

7.2.2- Resumo da experiência com Algoritmo Genético para o arquivo UK

Esta seção apresenta o resumo do processamento com o AG considerando apenas os valores relacionados ao processamento para o reconhecimento de faces com 10 poses de cada classe, obtidos das Tabelas 4.5 a 4.8 do Capítulo IV, arquivo UK.

As Tabelas 7.2 e 7.3 mostram os dados relacionados para os casos 64, 550, 1100 e 2200 coeficientes DCT representados pela variável NC . Estes dados são relacionados ao reconhecimento com 10 poses. Observa-se que há uma flutuação da taxa de acerto acima de 98 % . Os tempos de processamento são obviamente crescentes com o número de classes.

Tabela 7.2: Resultados obtidos do reconhecimento da face, AG, 10 poses para reconhecimento, arquivo UK para 5 a 60 classes.

NC *	Parâmetros	Classes						
		5	10	20	30	40	50	60
64	Acerto (%)	100	100	98,6182	99,4492	99,6194	99,6194	98,83
	T.Trein. (s)	20,369	38,119	81,229	96,546	163,1381	202,6322	259,067
	T. Rec. (s).	0,0584	0,4806	0,8592	2,5302	3,4334	4,4718	4,699
550	Acerto (%)	100	100	99,33	99,0	98,75	99,11	99,166
	T.Trein. (s)	40,43	87,84	170,28	240,56	286,24	302,64	389,76
	T. Rec. (s).	0,1584	1,021	3,567	5,568	8,410	23,104	26,805
1100	Acerto (%)	100	100	98,5	99,33	99,50	99,50	99,0
	T.Trein. (s)	65,66	122,88	261,84	311,22	525,88	653,19	733,52
	T. Rec. (s).	0,2689	2,213	3,956	11,65	15,809	20,59	24,364
2200	Acerto (%)	100	100	98,382	99,2109	99,3807	99,3807	99,33
	T.Trein. (s)	116,30	223,03	475,25	564,86	954,47	1185,54	1323,78
	T. Rec. (s).	0,3751	4,03	7,20	21,20	28,77	37,47	44,445

**número de coeficientes*

Tabela 7.3: Resultados obtidos do reconhecimento da face, AG, 10 poses para o reconhecimento, o arquivo UK para 70 a 120 classes.

NC *	Parâmetros	Classes					
		70	80	90	100	110	120
64	Acerto (%)	99,3691	99,3691	99,4492	99,4192	99,4192	99,08
	T.Trein. (s)	287,95	348,71	392,85	440,88	482,61	490,137
	T. Rec. (s).	7,204	8,369	11,612	13,060	16,483	18,299
550	Acerto (%)	99,25	98,88	99,33	99,36	99	99,25
	T.Trein. (s)	573,16	621,45	701,45	874,98	937,20	573,16
	T. Rec. (s).	34,681	37,803	41,267	48,81	57,027	34,681
1100	Acerto (%)	99,25	99,33	99,30	99,30	99,33	99,25
	T.Trein. (s)	987,337	1112,31	1248,31	1366,45	1461,908	987,337
	T. Rec. (s).	43,397	60,222	67,715	85,470	96,185	43,397
2200	Acerto (%)	99,1310	99,1310	99,2109	99,1810	99,1810	98,916
	T.Trein. (s)	1479,79	1792,02	2018,84	2265,68	2480,11	2655,69
	T. Rec. (s).	68,85	79,98	110,99	124,80	157,52	175,791

**número de coeficientes*

As Figuras 7.3 e 7.4 mostram os gráficos comparativos entre os diferentes números de coeficientes empregados nesta experiência. Primeiramente a comparação entre taxas de acertos e depois entre os tempos de reconhecimentos.

A Figura 7.3 mostra as curvas das taxas de acertos. Observa-se que entre 5 a 20 classes a taxa cai bruscamente elevando-se a partir de então para um patamar flutuante acima de 98,50. O melhor comportamento desta variável, no geral, é para 2200 coeficientes, embora com 20 classe é registrado o menor valor das taxas de acertos que é 98,38 % .

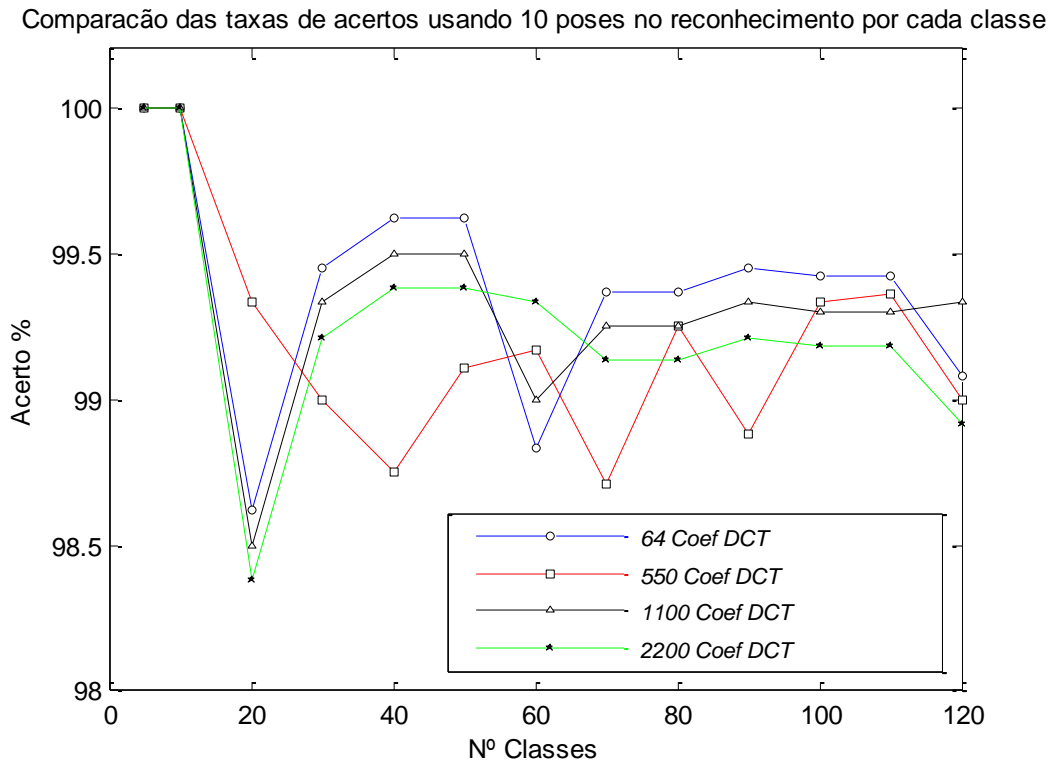


Figura 7.3: Comparação entre as taxas de acertos do AG considerando 64, 550, 1100 e 2200 coeficientes, arquivo UK.

A Figura 7.4 mostra as curvas dos tempos de reconhecimento, conforme as Tabelas 7.2 e 7.3. Observa-se que a partir do processamento com 80 classes o tempo aumenta mais rapidamente. O maior tempo foi para 2200 coeficientes cujos valores para 5 e 120 classes são 0,375 s e 175,791 s, respectivamente.

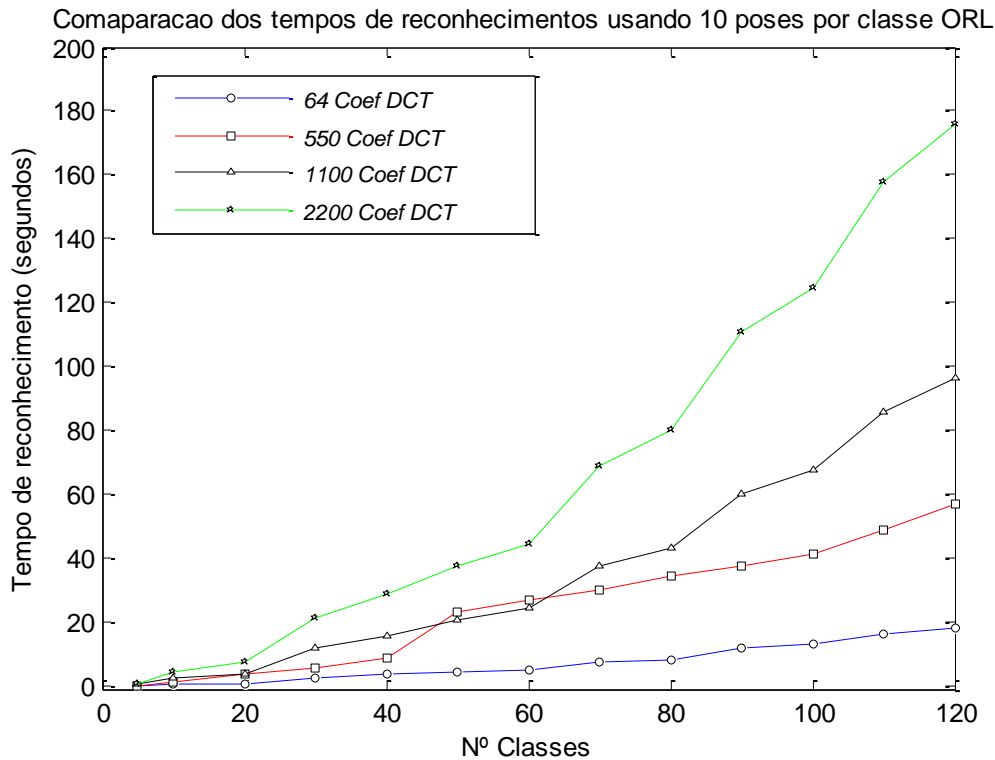


Figura 7.4: Comparação entre os tempos de reconhecimentos do AG considerando 64, 550, 1100 e 2200 coeficientes, arquivo UK.

7.3 - Resumo da experiência com a Rede Neural Artificial

Nesta seção, são mostrados os resumos utilizando o processamento com o RNA para os arquivos ORL e UK.

7.3.1- Resumo da Experiência com a Rede Neural Artificial para o arquivo ORL

Esta seção apresenta o resumo do processamento com a RNA considerando apenas os valores relacionados ao processamento para o reconhecimento de faces com 5 poses de cada classe, obtidos das Tabelas 5.1 a 5.3 do Capítulo V, arquivo ORL.

A Tabela 7.4 mostra os dados relacionados aos casos para 64, 100 e 256 coeficientes DCT representados pela variável NC . Estes dados são relacionados ao reconhecimento com cinco poses. Observa-se independente do número de coeficientes a taxa de acerto mantém-se em valores acima 99 % exceto no processamento de 2 classes

com 100 coeficientes que a taxa é de 95 %. Os tempos de processamento são obviamente crescentes com o número de classes embora muito pequeno.

Tabela 7.4: Resultados obtidos do reconhecimento de face com a RNA, 5 poses para reconhecimento, arquivo ORL para 2 a 40 classes.

NC *	Parâmetros	Classes							
		2	5	10	15	20	25	30	40
64	Acerto (%)	100	99,84	99,76	99,69	99,67	99,80	99,76	99,85
	T. Rec. (s).	0,09	0,18	0,43	0,56	0,83	0,93	1,24	1,75
	T.Trein. (s)	1,33	2,93	8,16	13,87	21,35	33,13	70,17	270,29
100	Acerto (%)	95,00	98,08	99,40	99,57	99,71	99,77	99,85	99,87
	T. Rec. (s).	0,10	0,20	0,40	0,60	0,74	0,95	1,20	1,70
	T.Trein. (s)	0,93	1,77	5,66	16,83	30,06	79,68	153,81	224,60
256	Acerto (%)	99,00	99,52	99,72	99,52	99,71	99,78	99,82	99,82
	T. Rec. (s).	0,09	0,19	0,41	0,66	0,89	0,99	1,39	2,02
	T.Trein. (s)	1,14	2,68	6,29	17,54	46,40	67,56	118,14	267,05

*número de coeficientes

As Figuras 7.5 e 7.6 mostram os gráficos comparativos entre os diferentes números de coeficientes empregados nesta experiência. Primeiro é mostrado a comparação entre as taxas de acertos e depois entre os tempos de reconhecimentos.

A Figura 7.5 mostra as curvas das taxas de acertos. Observa-se que entre 2 a 5 classes a taxa é crescente apenas para o teste com 100 coeficientes. Para as demais situações as taxas de acerto são mantidas acima de 99 % .

Comparacao das Taxas de Acerto da RNA, reconhecimento com 5 poses por classe ORL

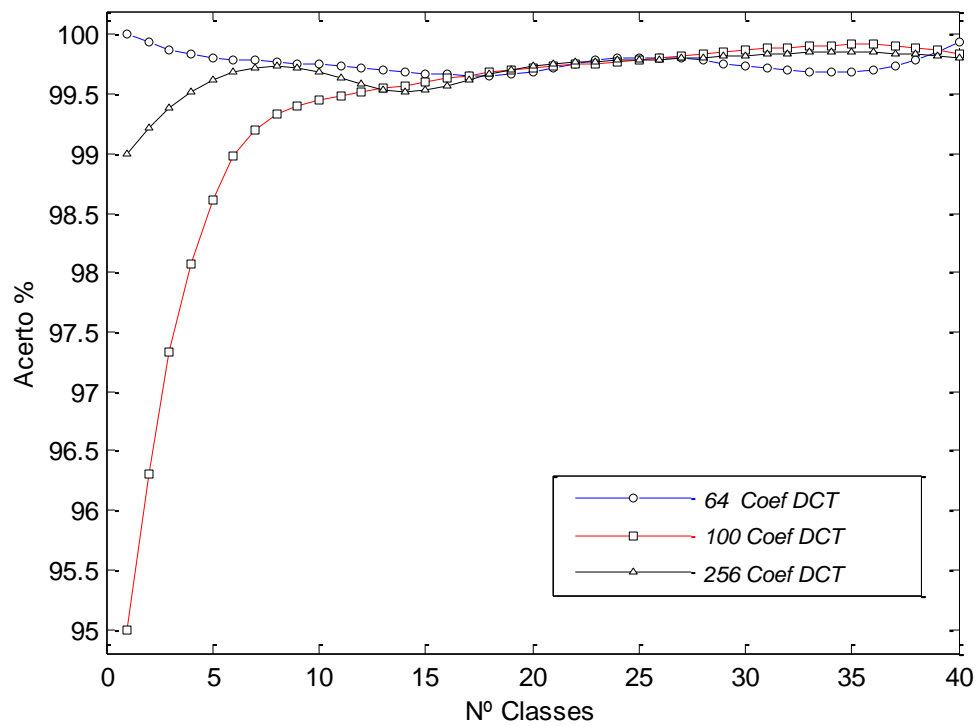


Figura 7.5: Comparação entre as taxas de acerto do RNA considerando 64, 100 e 256 coeficientes, arquivo ORL.

A Figura 7.6 mostra as curvas dos tempos de reconhecimento, conforme a Tabelas 7.4. O maior tempo foi para 256 coeficientes cujos valores para 2 e 40 classes são 0,09 s e 2,02 s, respectivamente.

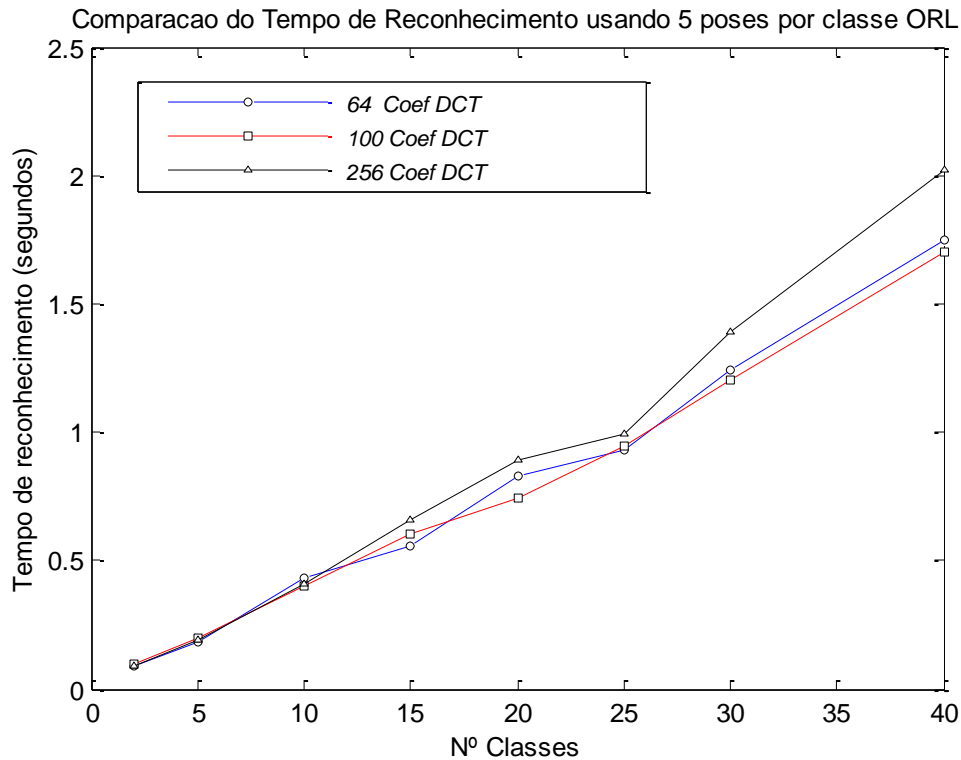


Figura 7.6: Comparação entre os tempos de reconhecimentos da RNA considerando 64, 100 e 256 coeficientes, arquivo ORL.

7.3.2- Resumo da experiência com a Rede Neural Artificial para o arquivo UK

Esta seção apresenta o resumo do processamento com a RNA considerando apenas os valores relacionados ao processamento para o reconhecimento de faces com 10 poses de cada classe obtidos das Tabelas 5.4 a 5.9 do Capítulo V, arquivo UK.

As Tabelas 7.5 e 7.6 mostram os dados relacionados para 64, 100 e 256 coeficientes DCT representados pela variável *NC*. Estes dados são relacionados ao reconhecimento com 10 poses. Observa-se independente do número de coeficientes que a taxa de acerto mantém-se em valores acima 99 % exceto no processamento de 5 classes com 64 coeficientes que a taxa é de 98 %. Os tempos de reconhecimento são obviamente crescentes com o número de classes embora muito pequeno.

Tabela 7.5: Resultados obtidos do reconhecimento da face, RNA, 10 poses para reconhecimento, arquivo UK para 5 a 60 classes.

NC *	Parâmetros	Classes						
		5	10	20	30	40	50	60
64	Acerto (%)	98	100	100	99,66	99,75	99,80	100
	T.Trein. (s)	2,93	8,16	21,35	70,17	121,83	368,35	959,58
	T. Rec. (s)	0,0021	0,0044	0,0108	0,0196	0,0275	0,0386	0,0703
100	Acerto (%)	100	100	100	99,66	100	99,40	100
	T.Trein. (s)	3,20	10	35	98,73	200	397	930
	T. Rec. (s)	0,0024	0,0054	0,0114	0,0196	0,0293	0,0454	0,0540
256	Acerto (%)	100	100	100	100	99,75	99,60	100
	T.Trein. (s)	5,06	18,71	91,95	278,67	791	963,59	1777
	T. Rec. (s)	0,0048	0,0093	0,0184	0,0365	0,0412	0,0590	0,0793

*número de coeficientes

Tabela 7.6: Resultados obtidos do reconhecimento de face, RNA, 10 poses para reconhecimento, arquivo UK para 70 a 120 classes.

NC *	Parâmetros	Classes					
		70	80	90	100	110	120
64	Acerto (%)	100	99,62	99,77	99,70	99,81	99,75
	T.Trein. (s)	1047,0	1174,5	1737,3	3003,1	3476,90	4800
	T. Rec. (s)	0,0710	0,0920	0,1114	0,1351	0,1495	0,171
100	Acerto (%)	99,85	99,87	99,66	99,90	99,90	99,75
	T.Trein. (s)	1265	2023,3	2664	4036	4359	5408
	T. Rec. (s)	0,0705	0,0922	0,1161	0,1416	0,1637	0,1941
256	Acerto (%)	99,85	99,75	99,88	100	99,81	99,75
	T.Trein. (s)	3437	5310	7114	10840	14465	15132
	T. Rec. (s)	0,1006	0,1161	0,1483	0,1612	0,1971	0,2258

*número de coeficientes

Nas Figuras 7.7 e 7.8 são mostrados os gráficos comparativos entre os diferentes números de coeficientes empregados nesta experiência. Primeiro é mostrado a comparação entre as taxas de acertos e depois entre os tempos de reconhecimentos.

A Figura 7.7 mostra as curvas das taxas de acertos. Observa-se que há uma pequena queda, 98 %, nessa taxa para 5 classes e 64 coeficientes. Nas demais situações as taxas de acerto são superiores a 99 %.

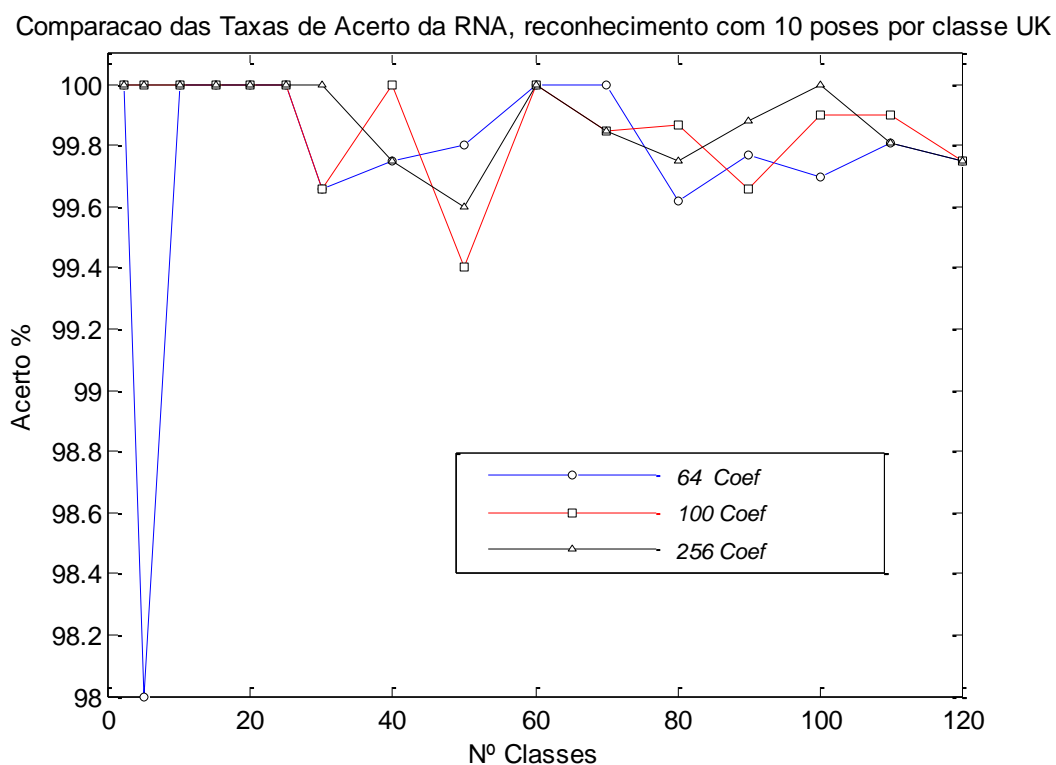


Figura 7.7: Comparação entre as taxas de acerto da RNA considerando 64, 10 e 256 coeficientes, arquivo UK.

A Figura 7.8 mostra as curvas dos tempos de reconhecimento, conforme as Tabelas 7.5 e 7.6. O maior tempo foi para 256 coeficientes cujos valores para 5 e 120 classes são 0,0048 s e 0,2258 s, respectivamente.

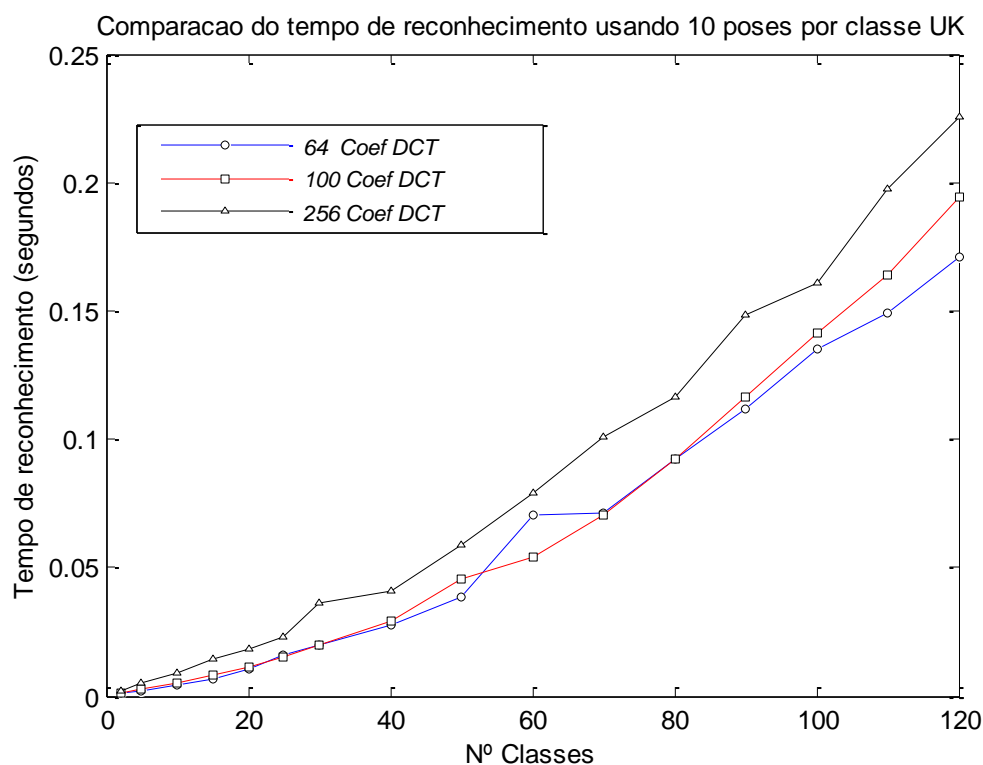


Figura 7.8: Comparação entre os tempos de reconhecimentos da RNA considerando 64, 10 e 256 coeficientes, arquivo UK.

7.4 - Resumo da experiência com a Análise de Componentes Principais

Esta seção mostra os resumos utilizando o processamento com a PCA para os arquivos ORL e UK.

7.4.1- Resumo da experiência com a Análise de Componentes Principais para o arquivo ORL

Esta seção apresenta o resumo do processamento com a PCA considerando os valores relacionados ao processamento para o reconhecimento de faces obtidos da Tabela 6.1, Capítulo VI, arquivo ORL.

A Tabela 7.7 mostra os dados relacionados aos casos utilizando apenas 10 poses para o reconhecimento de faces. O tempo de treinamento considerado uma só vez para as 40 classes foi de 39 segundos. Pode-se observar na Tabela 7.7 que as taxas de acerto

apresentam valores praticamente sem erro, pois atingem 100%. Para 2 classes o menor tempo de reconhecimento foi de 166,02 s.

Tabela 7.7: Resultados obtidos de reconhecimento de face, PCA, 10 poses para reconhecimento, arquivo ORL para 2 a 40 classes.

Parâmetros	Classes							
	2	5	10	15	20	25	30	40
Acerto %	100	100	100	100	100	100	100	99,72
T. Rec. (s)	166,02	407,98	777,74	1236,4	1695,05	3685,9	7676,6	15047,4

7.4.2- Resumo da experiência com a Análise de Componentes Principais para o arquivo UK

Esta seção apresenta o resumo do processamento com a PCA considerando apenas os valores relacionados ao processamento para o reconhecimento de faces obtidos da Tabela 6.2 do Capítulo VI, arquivo UK.

A Tabela 7.8 mostra os dados relacionados aos citados casos, usando 10 poses para o reconhecimento de faces. O tempo de treinamento considerado de uma só vez para as 40 classes foi de 216 segundos. Pode-se observar na Tabela 7.8 que as taxas de acerto apresentam valores de 100% até no processamento com 25 classes, com ligeira queda de 30 a 40 classes. Para 2 classes o menor tempo de reconhecimento foi de 665,68 s.

Tabela 7.8: Resultados obtidos de reconhecimento de face, PCA, 10 poses para reconhecimento, arquivo UK para 2 a 40 classes.

Parâmetros	Classes							
	2	5	10	15	20	25	30	40
Acerto %	100	100	100	100	100	100	99,75	99,15
T.Rec. (s)	665,68	906,52	1276,4	1816,4	2276,59	3276,15	8276,9	16676,97

Observa-se que o tempo de reconhecimento torna-se muito grande à medida que aumenta o número de classe.

As Figuras 7.9 e 7.10 mostram os gráficos comparativos entre os diferentes números de coeficientes empregados nesta experiência. Primeiro é mostrada a comparação entre taxas de acertos e depois entre os tempos de reconhecimentos. Observando que os tempos de treinamento são 39 segundos e 216 segundos para os arquivos ORL e UK, respectivamente.

A Figura 7.9 mostra o gráfico resultante da comparação entre taxas de acertos. Pode-se observar nessa figura que embora a taxa de acerto esteja acima de 99% ela é melhor para o arquivo ORL que o arquivo UK. Naquele caso a taxa começa a decrescer a partir do processamento com 30 classes, enquanto que para este o decréscimo começa com 25 classes.

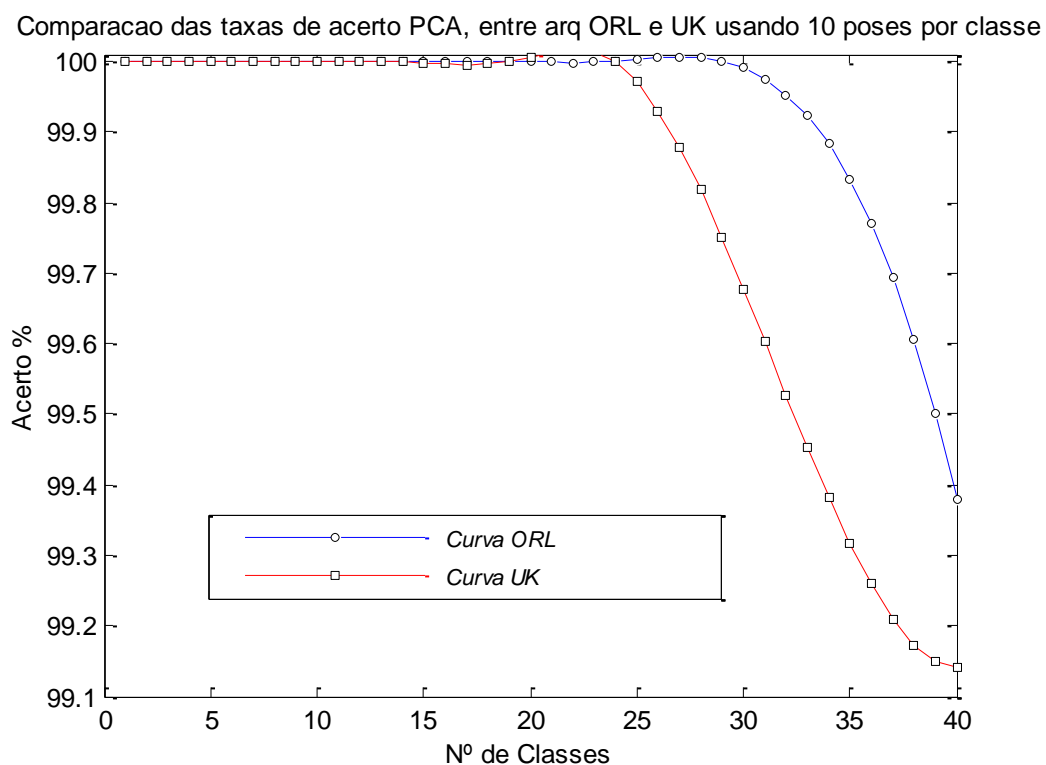


Figura 7.9: Comparação entre as taxas de acertos da PCA para os arquivos ORL e UK.

A Figura 7.10 mostra o gráfico comparativo entre os tempos de reconhecimentos. Observa-se que a partir do processamento com 25 classes, para o arquivo UK, o tempo de reconhecimento inicia uma brusca ascensão, chegando a 16676,97 s para 40 classes.

Comparação do tempo de reconhecimento PCA, entre arq ORL e UK usando 10 poses por classe

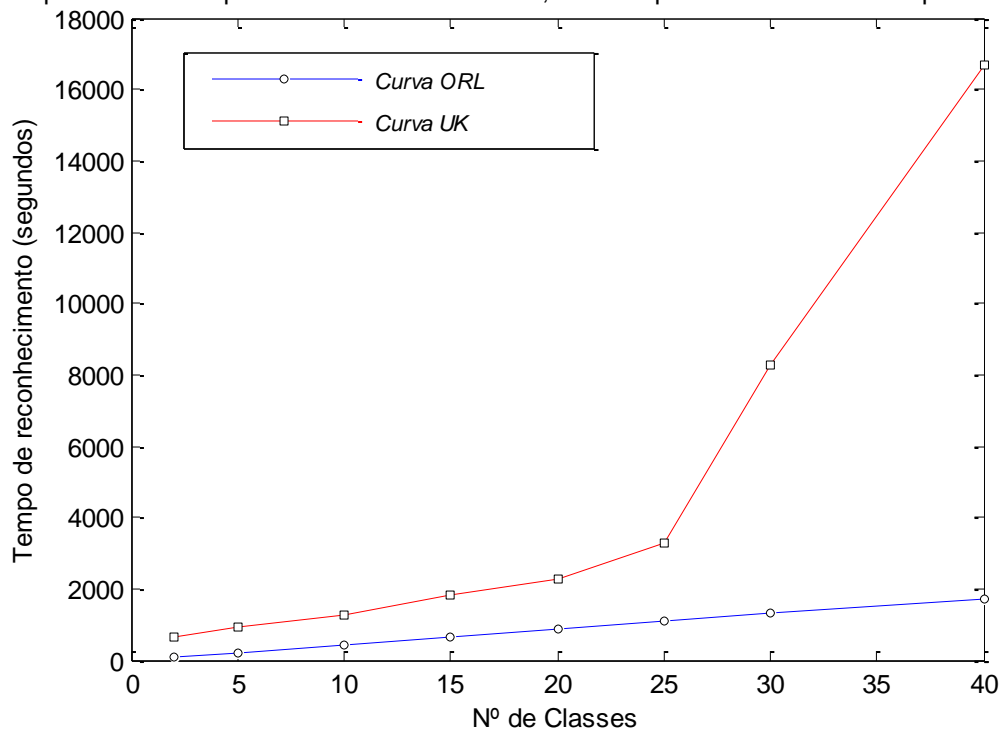


Figura 7.10: Comparação entre os tempos de reconhecimentos PCA para os arquivos ORL e UK.

7.5- Comparação Dos Resultados Entre os Métodos Do AG, RNA e PCA

Esta seção mostra as comparações do desempenho dos três métodos, primeiro para o arquivo ORL e em seguida para o arquivo UK.

7.5.1- Comparação dos resultados entre os métodos do AG, RNA e PCA para o arquivo ORL

São mostrados apenas os resultados obtidos nos testes com 64 coeficientes DCT por pose de cada classe.

A Tabela 7.9 mostra o desempenho entre os classificadores AG (resultados da Tabela 7.1), RNA (resultados da Tabela 7.4) e PCA (resultados da Tabela 6.1) quando foram utilizados 5 poses para o reconhecimento por classe. Ressaltando mais uma vez que a PCA treina todas as 40 classes por vez. Pode-se observar na Tabela 7.9 que os menores valores das taxas de acerto foram através do AG a partir do processamento com 10 classes enquanto que a RNA e a PCA mantiveram os valores acima de 99 %.

Tabela 7.9: Resultados obtidos entre os classificadores AG (100 gerações), RNA e PCA considerando 5 poses para o reconhecimento, tempo em segundos - arquivo ORL.

Método	Parâmetros	Classes							
		2	5	10	15	20	25	30	40
AG 64 coef	Acerto (%)	100	100	100	90	87	84	84	80,50
	T.Trein. (s)	2,841	6,534	13,531	20,191	27,534	33,594	40,28	53,891
	T. Rec. (s)	0,0131	0,0317	0,0634	0,143	0,223	0,359	0,467	0,816
RNA 64 coef	Acerto (%)	100	99,84	99,76	99,69	99,67	99,80	99,76	99,85
	T.Trein. (s)	1,33	2,93	8,16	13,87	21,35	33,13	70,17	270,29
	T. Rec. (s)	0,09	0,18	0,43	0,56	0,83	0,93	1,24	1,75
PCA	Acerto (%)	100	100	100	100	100	100	100	99,50
	T.Trein. (s)	39	39	39	39	39	39	39	39
	T. Rec. (s)	89,01	215,74	428,72	649,48	870,25	1086,0	1301,78	1733,57

Nas Figuras 7.11 a 7.13 são mostrados os gráficos comparativos relacionados a Tabela 7.7. Primeiramente a comparação entre taxas de acertos, depois entre os tempos de reconhecimentos e na sequência os tempos de treinamentos.

A Figura 7.11 mostra o gráfico resultante da comparação entre taxas de acertos com os resultados mostrados na Tabela 7.9. Pode-se observar nessa figura que o AG apresenta taxa de acerto decrescente a partir do processamento com 10 classes com 100 % até 40 classes com 80,50 %. Já a RNA e A PCA mostraram estáveis com mais de 99 % de acertos, quase que sem erros.

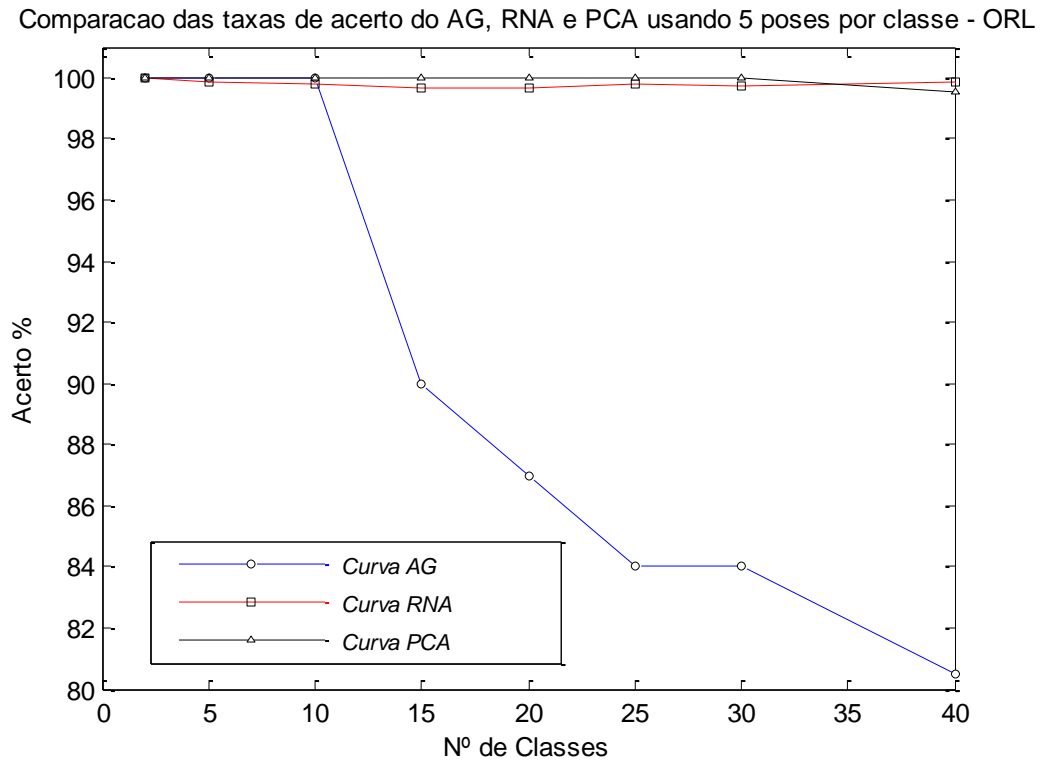


Figura 7.11: Comparação das taxas de acertos entre AG, RNA e PCA - arquivo ORL

A Figura 7.12 mostra os gráficos comparativos dos tempos de reconhecimentos entre o AG, RNA e PCA. Observa-se que para 40 classes tanto o AG quanto a RNA tem uma bom desempenho ao realizarem o reconhecimento em tempo pequeno, 0,816 s e 1,75 s respectivamente. A PCA por sua vez, para o mesmo número de classes realiza o reconhecimento em 1733,57 s.

Comparação do tempo reconhecimento do AG, RNA e PCA usando 5 poses por classe - ORL

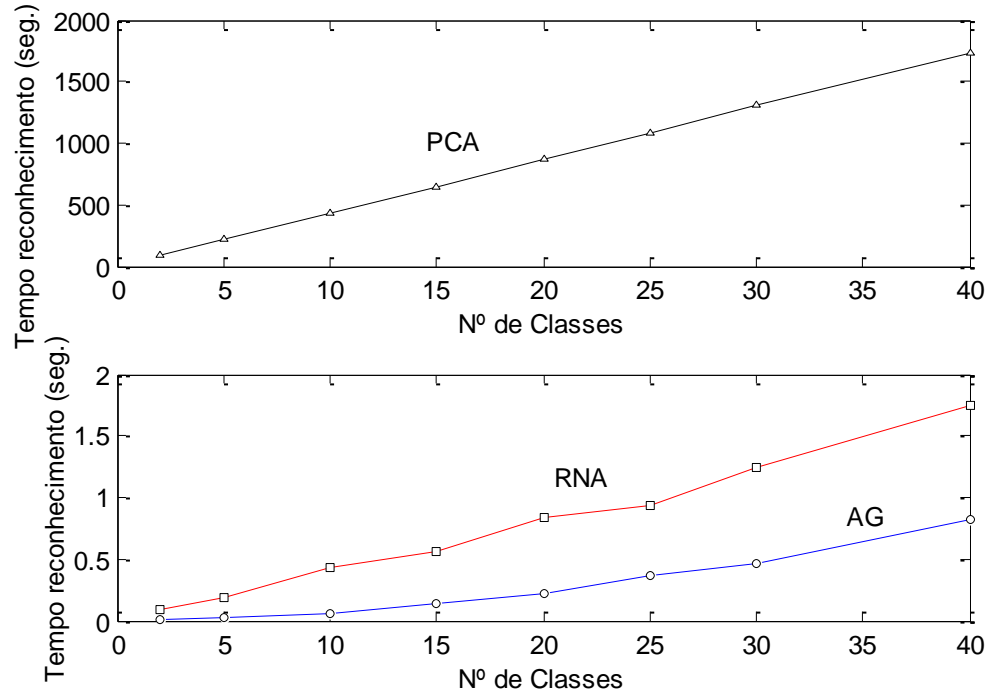


Figura 7.12: Comparação dos tempos de reconhecimentos entre o AG, a RNA e a PCA - arquivo ORL.

A Figura 7.13 mostra os gráficos comparativos referentes aos tempos de treinamentos. Observa-se o bom desempenho da PCA com um tempo único de 39 s, pois ela treina o arquivo uma só vez treinamento de uma só vez. Para 40 classes o AG e a RNA realizam o treinamento em 53,891 s e 270,29 s, respectivamente.

Comparação do tempo de treinamento do AG, RNA e PCA usando 5 poses por classe - ORL

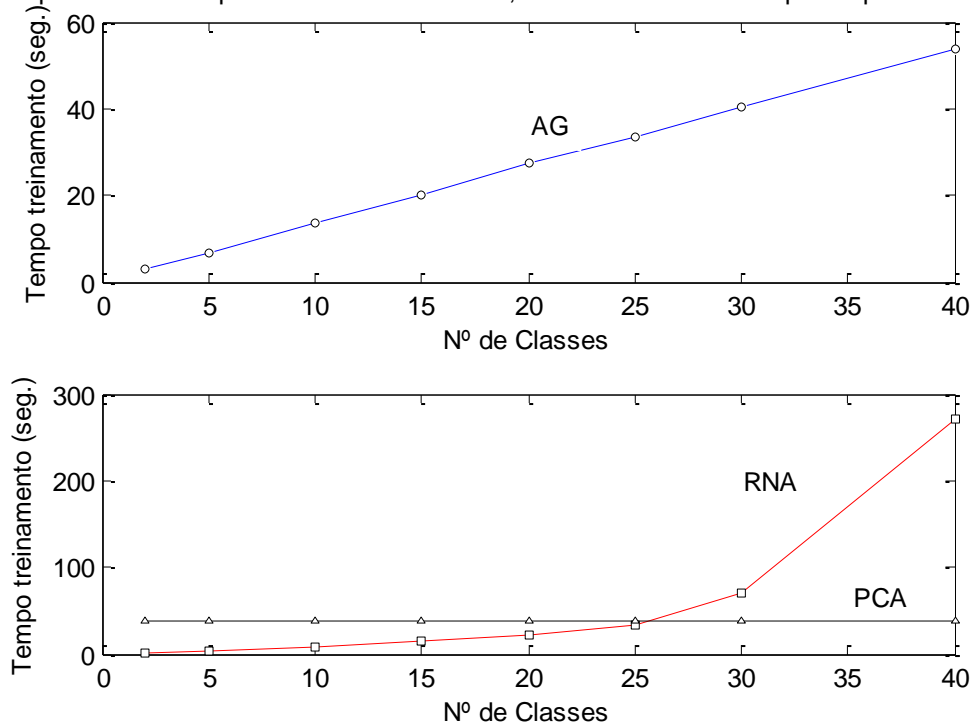


Figura 7.13: Comparação dos tempos de treinamentos entre o AG, a RNA e a PCA - arquivo ORL.

7.5.2- Comparação dos resultados entre os métodos do AG, da RNA e da PCA para o arquivo UK

São considerados apenas os resultados obtidos nos testes com 64 coeficientes DCT por pose de cada classe.

A Tabela 7.10 mostra os resultados obtidos entre os classificadores AG (resultados da Tabela 7.2), RNA (resultados da Tabela 7.5) e PCA (resultados da Tabela 7.8) considerando de 10 poses para o reconhecimento por classe. Ressaltando que mesmo tendo resultados do AG e da RNA para 120 classes, são mostrados apenas os de 40 classes porque o reconhecimento utilizando a PCA para um número maior de classes é impraticável. Pode-se observar na Tabela 7.10 que as taxas de acerto apresentadas pela PCA é um pouco melhor embora os três algoritmos obtêm alto valores de acertos. O tempo de reconhecimento da PCA é de 16676,97 s para 40 classes.

Tabela 7.10: Resultados obtidos utilizando a da RNA, PCA e AG (100 gerações) considerando 10 poses para o reconhecimento, tempo em segundos - arquivo UK.

Método	Parâmetros	Classes				
		5	10	20	30	40
AG 64 coef	Acerto (%)	100	100	98,6182	99,4492	99,6194
	T.Trein. (s)	20,69	38,119	81,229	96,546	163,138
	T. Rec. (s)	0,0584	0,4806	0,8592	2,5302	3,4334
RNA 64 coef	Acerto (%)	98	100	100	99,66	99,75
	T.Trein. (s)	2,93	8,16	21,35	70,17	121,83
	T. Rec. (s)	0,0021	0,0044	0,0108	0,0196	0,0275
PCA	Acerto (%)	100	100	100	99,75	99,15
	T.Trein. (s)	216	216	216	216	216
	T. Rec. (s)	906,52	1276,39	2276,59	8276,96	16676,97

As Figuras 7.14 a 7.16 mostram os gráficos comparativos dos resultados da Tabela 7.10. Primeiro é a comparação entre as taxas de acertos depois entre os tempos de reconhecimentos e na sequência os tempos de treinamentos.

Na Figura 7.14 pode-se observar que o AG e a PCA apresentam altas taxas de acerto acima de 99 %. Já a RNA apresenta a menor taxa que é de 98 %.

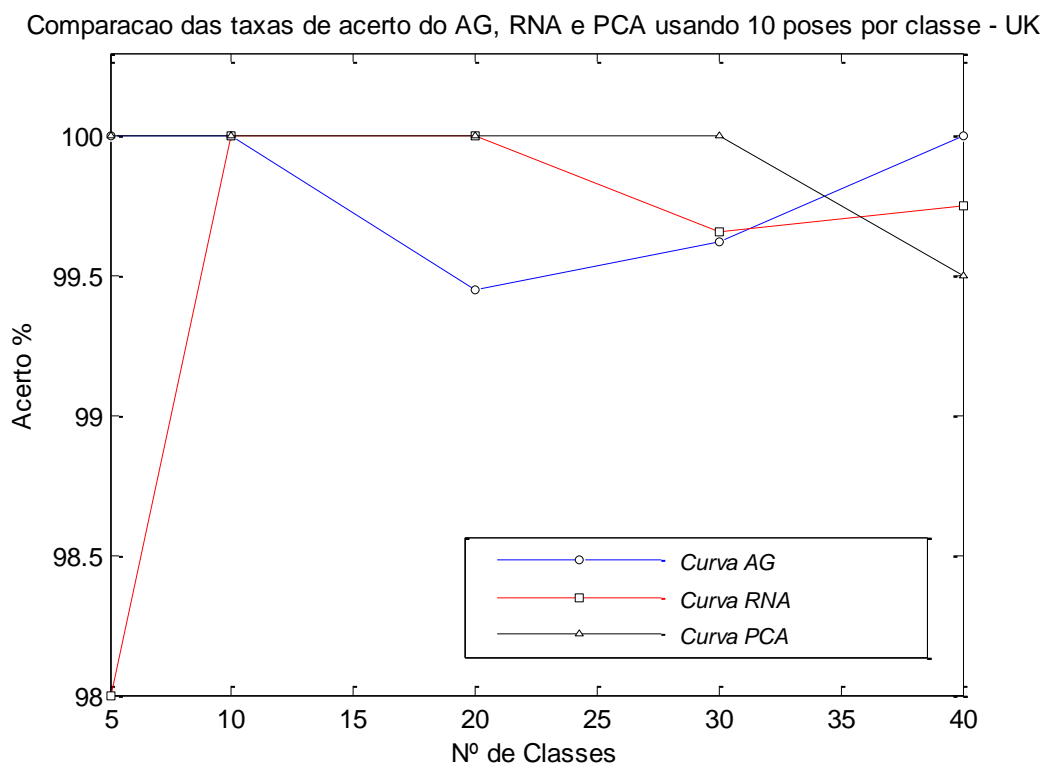


Figura 7.14: Comparação das taxas de acertos entre o AG, a RNA e a PCA - arquivo UK.

A Figura 7.15 mostra os gráficos comparativos dos tempos de reconhecimentos entre o AG, RNA e PCA. Observa-se que para 40 classes tanto o AG quanto a RNA tem um bom desempenho ao realizarem o reconhecimento em tempo pequeno, 3,4334 s e 0,0275 s respectivamente. A PCA por sua vez, para o mesmo número de classes realiza o reconhecimento em 16676,97 s.

Comparação do tempo de reconhecimento do AG, RNA e PCA usando 10 poses por classe - UK

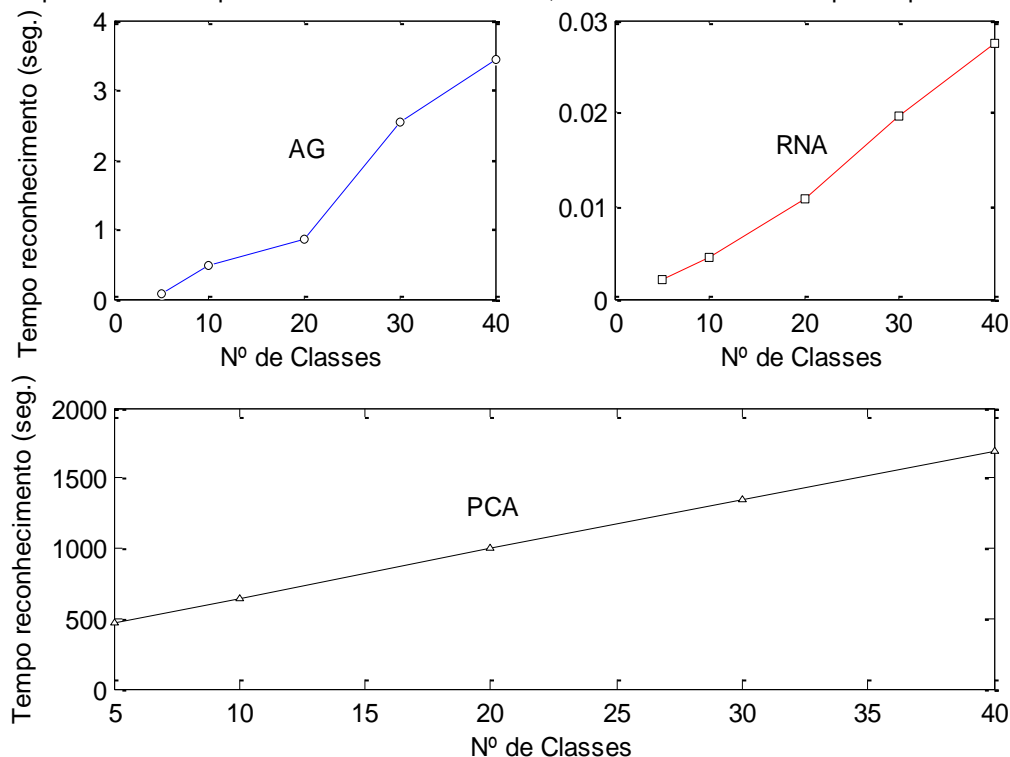


Figura 7.15: Comparação dos tempos de reconhecimentos entre o AG, a RNA e a PCA - arquivo UK.

A Figura 7.16 mostra os gráficos comparativos referentes aos tempos de treinamentos. Observa-se o bom desempenho da PCA com um tempo único de 216 s, pois ela treina o arquivo uma só vez treinamento de uma só vez. Para 40 classes o AG e a RNA realizam o treinamento em 163,138 s e 121,83 s, respectivamente.

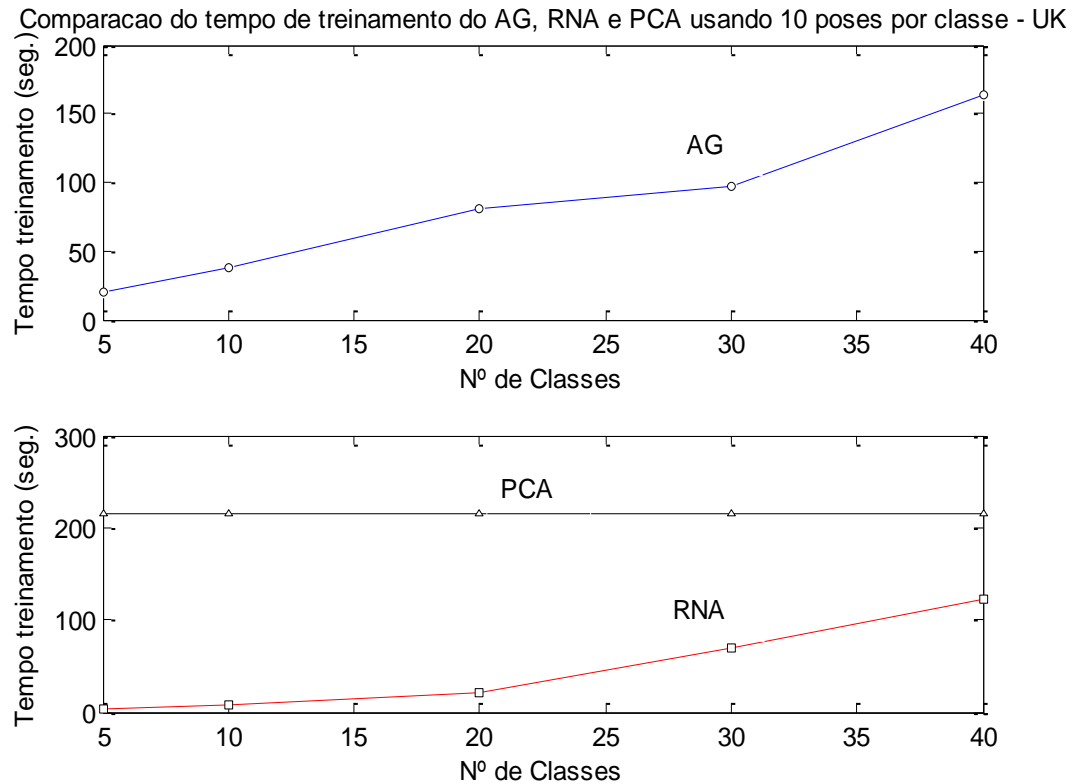


Figura 7.16: Comparação dos tempos de treinamentos entre o AG, a RNA e a PCA – arquivo UK.

7.6- Amostras de comparações de imagens reconhecidas erradas pelos algoritmos

A Figura 7.17 mostra as imagens de poses que o AG reconheceu errado. Primeiro da esquerda para a direita, a comparação entre duas poses do arquivo ORL e depois duas poses do arquivo UK.

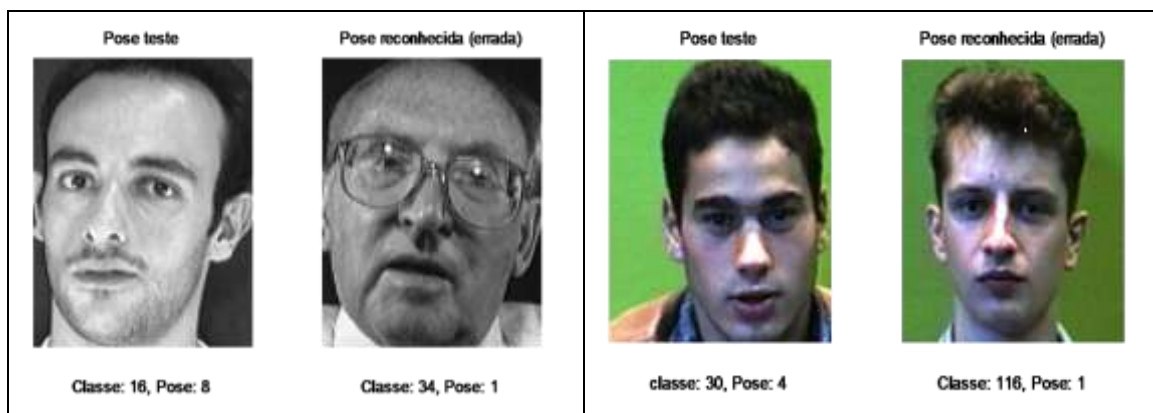


Figura 7.17: Comparação das imagens de faces reconhecidas erradas pelo AG. As duas primeiras faces no arquivo ORL e as outras duas no arquivo UK.

A Figura 7.18 mostra as imagens de poses que a RNA reconheceu errado. Primeiro da esquerda para a direita, a comparação entre duas poses do arquivo ORL e depois duas poses do arquivo UK.

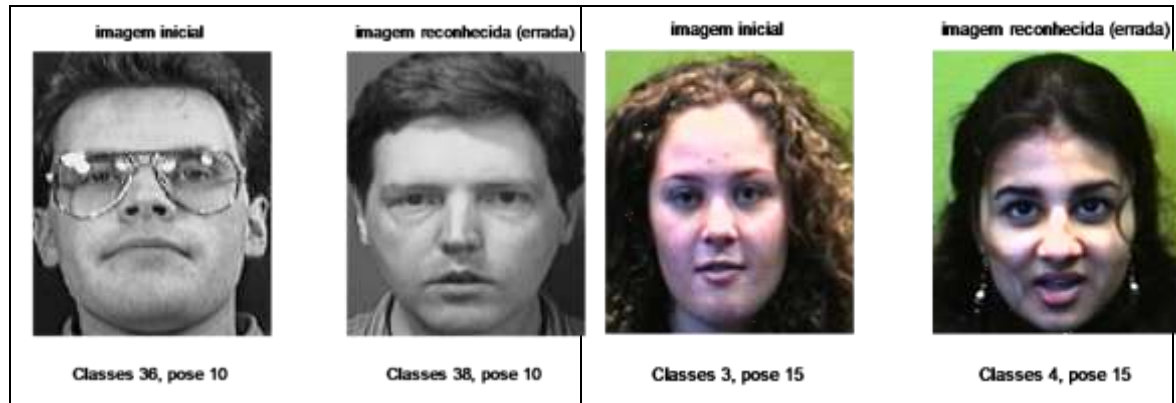


Figura 7.18: Comparação das imagens de faces reconhecidas erradas pela RNA. As duas primeiras faces são do arquivo ORL e as outras duas do arquivo UK

A Figura 7.19 mostra as imagens de poses que a PCA reconheceu errado. Primeiro da esquerda para a direita, a comparação entre duas poses do arquivo ORL e depois duas poses do arquivo UK.



Figura 7.19: Comparação das imagens de faces reconhecidas erradas pela PCA. As duas primeiras faces do arquivo ORL e as outras duas do arquivo UK

7.7- Análises Dos Resultados

A taxa de acerto do AG executado com 100 gerações, para o arquivo ORL, apresentou diminuição de valores em torno da classe dez. Para o arquivo UK a

diminuição de valores ocorreu na faixa entre a décima quinta classe e a vigésima quinta classe. Para esses arquivos, as taxas de acertos do AG aumentaram bem menos do que as taxas de acertos da RNA e da PCA que estão muito próximas d 100 %.

O tempo de reconhecimento o AG para o arquivo ORL mostrou vantagens ao ser muito pequeno mesmo com número de classes elevado. A RNA apresentou o tempo ascendente constante com relação ao aumento do número de classe, enquanto que a PCA apresentou pior desempenho nesta variável, ou seja, o tempo de reconhecimento é muito longo.

Com relação ao tempo de treinamento tanto para arquivo ORL quanto para o arquivo UK, o AG apresentou o tempo de treinamento em ascendência forte crescente com o número de classes. A RNA apresentou tempo de treinamento significativo após a vigésima a classe. A PCA mostrou-se interessante neste particular. O tempo de treinamento é estável. A PCA codifica todo o arquivo de uma só vez, deixando os dados do arquivo codificado disponíveis para os testes de reconhecimentos com qualquer número de classe constante no arquivo.

Os testes realizados foi verificado que o nível de acerto é muito próximo nos métodos que utilizam diferentes números de coeficientes DCT. Quanto maior é o número de coeficientes DCT utilizados mais alto o patamar de flutuação da taxa de acerto.

7.8- Sobre O Desempenho E Aplicações Dos Classificadores

Foi analisado e comparado o desempenho de cada classificador. A Análise de Componentes Principais pelo seu alto percentual de acerto deve ser empregada em projetos especiais em que se podem aguardar longo tempo por uma resposta. Serve como instrumento auxiliar de outros algoritmos, por exemplo, para os momentos nos que não se pode ter dúvida quanto ao reconhecimento de uma pessoa.

Para uma aplicação que é necessária resposta imediata, os classificadores Algoritmo Genético e Rede Neural Artificial são os mais adequados. Considerando o tempo de resposta, a relação custo x benefício é muito favorável ao AG e a RNA que apresenta alta taxa de acerto e baixo tempo de reconhecimento mesmo ao aumentar o número de poses no banco de imagens, como foi observado ao utilizar o arquivo UK.

Nas aplicações práticas, o AG possui a vantagem de funcionar mesmo que as imagens são aleatórias, obtidas da rotina diária das pessoas. Como mostrado na seção B.2 do apêndice B deste trabalho, utilizou-se fotos caseiras com tamanho muito maior dos que aqueles utilizados na experiência, sem nenhum trabalho de melhoria e mesmo assim o AG obteve a taxa de acerto um percentual de acerto aceitável, conforme Figura B.2, para as circunstâncias (características do arquivo). Esta mesma experiência ao ser executada tanto pela PCA quanto pela RNA não foi concluída, pois, para aquela faltou memória virtual (mesmo com sistema operacional de 64 Bits) e para este, deu erro em função do tamanho de cada pose (4320 x 3240 pixels).

Em outra aplicação prática simples, mostrada na seção B.3 do apêndice B deste trabalho, foi considerado o princípio da interatividade. Neste caso o AG ao atingir uma taxa de acerto estipulada, aciona uma placa lógica programável que por sua vez aciona um dispositivo externo.

Portanto, os testes mostraram maior versatilidade do AG. Conclui-se que ele pode realizar reconhecimento de faces utilizando diversos bancos de imagens.

7.9- Considerações Finais Deste Capítulo

A taxa de acerto do AG aumenta e são semelhantes mesmo tratando de arquivos diferentes o ORL e o UK. Embora o tempo de treinamento seja razoavelmente longo a etapa de reconhecimento é realizada em um tempo o que o remete de imediato para o campo da aplicação prática e da pesquisa.

Nos testes realizados a RNA apresentou uma taxa de acerto em torno de 98 %. O arquivo UK tem o dobro de imagens do ORL e, mesmo assim, a taxa de acerto para ambos os arquivos foram muito próximas. Mesmo com o aumento vertiginoso do tempo de treinamento em relação ao aumento do número de classes, o mais importante, é que a etapa de reconhecimento, foi realizada em um tempo bastante pequeno.

A PCA tem uma taxa de acerto muito alta e estável, média em torno de 99 %. O arquivo UK tem o dobro das imagens do ORL e, mesmo assim, as taxas de acertos para ambos os arquivos foram muito próximas. Mesmo o tempo de treinamento sendo pequeno em relação ao aumento do número de classes, a etapa mais importante, é o reconhecimento, que é realizada em um tempo muito grande mesmo com poucas classes e poucas poses utilizadas para o reconhecimento. A PCA é recomendada para casos que se dispõe de tempo para o reconhecimento de faces e de máquinas poderosas

para o processamento em termos de velocidade e paginação, ou seja, grande memória virtual.

O próximo capítulo mostra as conclusões, as contribuições, as sugestões de trabalhos futuros e publicações do autor desta tese.

CAPÍTULO VIII

CONCLUSÕES, CONTRIBUIÇÕES, SUGESTÃO DE TABALHOS FUTUROS E PUBLICAÇÕES

8.1- As implementações deste trabalho

O Capítulo I apresentou o estado da arte, onde foi mostrado as aplicações da biometria nas atividades da sociedade, incluindo os desafios e dificuldades a serem superados. Também, referencia vários trabalhos realizados para o reconhecimento de faces, conforme exemplificados na seção 1.8 do capítulo I. Os métodos muito utilizados foram a RNA e a PCA, no entanto pouquíssimos trabalhos foram realizados utilizando AG.

O Capítulo III apresentou os fundamentos teóricos dos classificadores: Algoritmo Genético, Rede Neural Artificial e Análise de Componentes Principais.

O Capítulo III apresentou os materiais comuns às experiências realizadas, a metodologia utilizada com os três classificadores e as especificações das bases de dados, ou seja, os dois arquivos de imagens de faces humanas que são completamente diferentes um do outro.

O Capítulo IV apresentou a experiência com o Algoritmo Genético utilizando os dados dos arquivos ORL e UK, aplicando-se métodos de otimização e busca inspirados na teoria da evolução de Darwin. Esta teoria trata as possíveis soluções do problema como “indivíduos” de uma “população” ao executar um processo evolucionário de geração em geração até encontrar a solução do problema. Para tanto foram utilizados os coeficientes DCT das poses (inteiras) de cada classe e depois os coeficientes das

subposes. Estes procedimentos foram descritos nas seções 3.3.1.1 e 3.3.1.2 do Capítulo III. Os motivos da escolha da DCT estão descritos na seção 3.1 do Capítulo III.

No Capítulo V foi mostrado o desempenho do algoritmo da Rede Neural Artificial no processo de reconhecimento de faces com o objetivo principal de obter os modelos com boa capacidade de generalização tendo como entrada um conjunto de dados DCT de cada face dos arquivos. Trata-se de uma ferramenta que faz a classificação e/ou a predição por meio de um conjunto de treinamento composto por pares de entrada e de saída da rede (x, y) , saída esta ajustada por meio de pesos que pela iteração vai modificando-a até atingir ou aproximar da saída desejada. Nas seções 3.3.2.1 e 3.3.2.2 do Capítulo III foram mostrados como as diferentes poses de faces dos arquivos em estudo foram utilizadas.

O Capítulo VI mostrou o reconhecimento de faces utilizando o algoritmo da Análise de Componentes Principais que transforma uma série de variáveis correlacionadas em um menor número de outras não correlacionadas chamadas componentes principais. Desta forma o conjunto de dados pode ser representado por um número reduzido de características e ainda retém a maioria do conteúdo da informação intrínseco dos dados; ou seja, o conjunto de dados sofre uma redução de dimensão. Neste caso o emprego da Análise de Componentes Principais foi utilizado como um instrumento de análise exploratória das faces humanas contidas nos arquivos ORL e UK e para obter modelos preditivos. A utilização constituiu do cálculo da decomposição de autovalores de dados, matrizes de covariâncias ou decomposição em valores singulares de uma matriz de dados. Nas seções 3.3.3.1 e 3.3.3.2 do Capítulo III foram mostrados como as diferentes poses de faces contidas nos arquivos em estudo foram utilizadas.

O Capítulo VII mostrou o resumo dos testes realizados com cada um dos três algoritmos, o Algoritmo Genético, a Rede Neural Artificial, e a Análise de Componentes Principais e depois a comparação dos resultados obtidos apontando as vantagens de cada um deles.

Este capítulo mostra as conclusões, contribuições, sugestões de trabalhos futuros e publicações do autor desta tese.

8.2- Conclusões e Contribuições

O objetivo deste trabalho é contribuir para a área de pesquisa de sistemas de reconhecimento biométrico de indivíduos utilizando as características da face humana.

Inicialmente, uma contribuição que se pôde realizar foi à elaboração do estado da arte que cita as inúmeras aplicações de um sistema de reconhecimento de faces e alguns trabalhos publicados e os resultados obtidos. O estado da arte serve para se obter uma visão geral da aplicação prática do reconhecimento biométrico e também adquirir muitas informações para atualizar com o assunto.

Outra contribuição desta tese foi analisar o comportamento do Algoritmo Genético utilizando o fracionamento das poses das faces humanas. Esse fracionamento busca as respostas por meio de processos evolutivos ao transformar aos poucos uma determinada resposta arbitrária até esta aproximar-se de uma verdadeira. Apesar do princípio de operação do AG ser de natureza biológica o Algoritmo Genético ainda é pouco utilizado para o reconhecimento biométrico. Neste sentido, este trabalho forneceu mais uma fonte de pesquisa trazendo as experiências com testes sistemáticos comparando os desempenhos tanto do Algoritmo Genético como da Rede Neural e da Análise de Componentes Principais.

É bastante complexo o estudo de reconhecimento de faces, devido as suas inúmeras variáveis e inúmeras condições que as faces estão expostas, que mesmo com o avanço tecnológico e com os aplicativos destinados à biometria, ainda é possível o desenvolvimento de novas técnicas. Desenvolver estas novas técnicas é uma tarefa de Universidades, e de equipes de pesquisadores com a finalidade de disponibilizar para a sociedade mais segurança e bem estar.

8.3- Sugestões de trabalhos futuros

Este trabalho tem por objetivo principal apresentar e comparar as três técnicas de reconhecimento de faces, onde duas delas: a RNA e a PCA já foram utilizadas em vários trabalhos de pesquisas desta natureza, enquanto o AG tem poucas referências.

Longe de esgotar em si mesmo, este trabalho pela multiplicidade de condições, possibilita pesquisas em reconhecimento de faces, como por exemplo:

- A utilização de diferentes técnicas e/ou métodos nas etapas do reconhecimento de faces, como por exemplo, a Quantização Vetorial (QV) e a Transformada de *Wavelet*. A QV é uma técnica de compressão em que as amostras do conjunto de dados são agrupadas em vetores k -dimensionais e a Transformada de *Wavelet* é utilizada nos algoritmos de normalização da iluminação de uma figura;

- A utilização da Quantização Vetorial e a PCA. O resultado da QV é um centróide representativo de cada classe. Com essa representação consegue-se diminuir o tempo e o tamanho da matriz característica;
- A utilização da Regressão por Mínimos Quadrados Parciais (PLS) trabalha simultaneamente com as informações espectrais e as concentrações no processo de calibração. Por sua vez a calibração obtém padrões por meio de medidas realizadas em uma série de amostras de concentrações. e com estas amostras são previstas as concentrações de novas amostras [52];
- A utilização de outras linguagens de programação diferente do Matlab empregada neste trabalho. Por exemplo, a linguagem C, C++ ou Java podem tornar o processamento de alguns algoritmos mais rápido;
- Desenvolvimento software e hardware para a localização de pessoas em uma multidão;
- A utilização neste trabalho de outros bancos de imagens;
- A utilização no AG de interações ecológicas. Existem comportamentos na natureza que contribuem com o ambiente para que duas determinadas espécies maximizem seus resultados [49].
- A utilização no AG de interações sociais. A grande maioria dos animais possuem regras de comportamento onde existem castas e/ou lideranças que buscam maximizar a sobrevivência do grupo, melhorar a qualidade da prole gerada ou simplesmente otimizar a relação da espécie com o meio ambiente [49].
- A utilização de novas idéias genéticas. Os algoritmos genéticos usam ainda, somente os conceitos de cruzamento e mutação. Podem ser considerados nos AG os introns, os transposons, os genes diplóides com dominância e codominância, diferenciação sexual e outros [50];
- Esta tese pode inspirar outros trabalhos no sentido de particularizar o processo de reconhecimento de faces sejam nos aspectos geométricos entre seus órgãos

como as colorações da pele e ainda em um viés mais desafiador que é buscar o reconhecimento utilizando os sentimentos e emoções que as faces revelam.

Finalmente, fazendo uma breve reflexão sobre o tema desta tese, existe um critério “estético” que parece ter sido à base da cultura grega (aproximadamente no século V a.C.) e que, pelas “mãos” de pensadores do século XV d.C. (onde o mais conhecido é Leonardo Da Vinci), em pleno Renascimento, construíram os fundamentos das artes e das ciências ocidentais. Como a imagem é parte preponderante dessa visão e está ligada ao objetivo do “ato de reconhecer pessoas”, foi oportuno este trabalho ao processar imagens de rostos humanos, por três algoritmos distintos, na busca de sua identificação. Mesmo contrapondo as tintas e os pinceis de outrora, os mecanismos de hoje se mostram com a mesma ansiedade desmedida da humanidade ao desejar reconhecer-se por meio das suas imagens. A relevância desta pesquisa para a engenharia é evidente na medida em que abre uma perspectiva inovadora do trato das imagens como o objeto de reconhecimento automático de pessoas com eficiência. Várias áreas estão em aberto, tanto teórica quanto prática, cujo desdobramento é o desafio aos pesquisadores que devem responder a essa demanda criando *hardwares* e *softwares* cada vez mais avançados.

8.4- Publicações

- Revista Hispeci & Lema, ISSN 1980-2536, ano III, n. 3, 2012.

Título: “RECONHECIMENTO DE FACE UTILIZANDO BANCO DE IMAGENS MONOCROMÁTICAS E COLORIDAS ATRAVÉS DOS MÉTODOS DA ANÁLISE DO COMPONENTE PRINCIPAL (PCA) E DA REDE NEURAL ARTIFICIAL (RNA)”.

- Contemporary Engineering Sciences, ISSN 1314-7641, Vol. 5, n. 10, 463-472, 2012.

Título: “RECOGNITION OF HUMAN FACES THROUGH GENETIC ALGORITHM”.

Aceito para publicação:

Revista do IEEE America Latina

Título: ” RECOGNITION OF HUMAN FACES THROUGH GENETIC ALGORITHM USING THE DISCRETE COSINE TRANSFORM”.

REFERÊNCIAS BIBLIOGRAFICAS

- [1] Supercomputador. <http://www.top500.org/list/2010/11/100>. Disponível em 19 de novembro de 2010.
- [2] Dantas, G. F. L. (2003). Sistemas Biométricos de Identificação pela Imagem Facial.
- [3] Impressão digital. <http://pt.dreamstime.com/imagem-de-stock-impress-atildeo-digital-image3203111>. Disponível em 31 de março de 2011.
- [4] A. Roddy and J.D. Stosz 1997. *Fingerprint Features: Statistical analysis and system Performance Estimates*. IEEE, vol. 85, Sept 1997, pp. 1390-1397.
- [5] Chips de silício. <http://w4.siemens.de/newsline.d/pressfor/end99101.htm>. Disponível em 19 de novembro de 2010.
- [6] Daugman, J. *High confidence visual recognition of person by a test of statistical independence*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 11, 1993.
- [7] Wildes, R. P.; Asmuth, J.; Green, G.; Hsu, S.; Kolczynski, R.; Mtey, J. & McBride, S. *A system for automated iris recognition*. Proceedings of the IEEE Workshop on applications of Computer Vision, 1994.
- [8] Íris. <http://clinicaolhosdiadema.com.br/olhos.html>. Disponível em 31 de março de 2011.

- [9] Globo ocular. <http://www.nucleomg.com.br/index.php?id=03>.Disponível em 31 de março de 2011.
- [10] Mckenna,Stephen J. and gong, Shaogang 1996. *Tracking Faces*. Second Internacional Conference on Automated Face and Gesture Recognition, October 1996, Vermont, USA.
- [11] A. W. Senior, 1996. *Face and Feature Fiding for a Face Recognition System In Proceedings of Audio and Video based Biometric Person Authentication '99*,pp. 154-159. March 1999, Washington, D.C., USA.
- [12] Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, et al 1999. *Face Recognition by Elastic Bunch Graph Matching*. Conference on Computer Analysis of Images and Patterns, CAIP'97, Kiel.
- [13] Cognitec. <http://www.cognitec-systems.de/index.html>. Disponível em 19 de abril de 2011.
- [14] FaceIt-Hist. http://www.identix.com/company/comp_history.html. Disponível em 19 de abril de 2011.
- [15] Y. Rui, T.S. Huang, and S.-F. Chang. *Image retrieval: current techniques, promising directions and open issues*.J. Visual Communication and Image Representation, 10-4:39–62, 1999.
- [16] A. Kudhinsky, C. Pering, M.L. Creech, D. Freeze, B. Serra, and J. Gvvizdka. *FotoFile: a consumer multimedia organization and retrieval system*. In Proceedings of CHI'99, pages 496–503, 1999.
- [17] P. Navarrete and J. del Solar. *Interactive face retrieval using self-organizing maps*. In Proceedings,2002 Int. Joint Conf. on Neural Networks: IJCNN2002, 2002.

- [18] R. Houghton. *Named faces: putting names to faces*. IEEE Intelligence Systems, 14-5:45–50, 1999.
- [19] FaceIt. <http://www.identix.com>. Disponível em 19 de abril de 2011.
- [20] H. Wactlar, T.K.M. Smith, and S. Stevens. *Intelligence access to digital video: informedia project*. IEEE Computer, 29-5:46–52, 1996.
- [21] H.Wang and S.-F. Chang. *A highly efficient system for automatic face region detection in mpeg video sequences*. IEEE Trans. on Circuits and Systems for Video Technology, Special Issue on Multimedia Systems and Technologies, 7-4:615–628, 1997.
- [22] Y. Wu and T. Huang. *Nonstationary color tracking for vision-based human computer interaction*. IEEE Transactions on Neural Networks, 13-4, 2002.
- [23] I. Cohen, N. Sebe, F.G. Cozman, M.C. Cirelo, and T. Huang. *Learning Bayesian network classifiers for facial expression recognition using both labeled and unlabeled data*. In Proceedings of the IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR 2002), 2003.
- [24] Mendonça, M. M. *Comparação entre métodos de normalização da iluminação utilizados para melhorar a taxa do reconhecimento facial*. Dissertação apresentada à Escola de Engenharia de São Carlos, da Universidade de São Paulo.
- [25] Kim. K. *Face recognition using Principal Component Analysis (PCA)*, University of Maryland, USA.
- [26] Matos, F. M. S. *Reconhecimento de Faces utilizando a Transformada Discreta do Cosseno*. Dissertação apresentada à Universidade Federal da Paraíba.

- [27] Amarapur Basavaraj. *The Facial Features Extraction for Face Recognition Based on Geometrical Approach*, P. D. A. College of Engineering Gulbarga, India.
- [28] L. Ming,. *Feature Selection of Face Recognition based on Improved Chaos Genetic Algorithm*, University of Technology LanZhou, China.
- [29] Alves, W. J. B., *Reconhecimento de Pessoas por meio de Algoritmo Genético aplicado em Medidas das Proporções Áureas da Face Humana*. Faculdades Integradas - Fafibe, Bebedouro, São Paulo.
- [30] Tanomaru, J. *Motivação, fundamentos e aplicações de algoritmos genéticos*. Anais do II Congresso Brasileiro de Redes Neurais, 1995.
- [31] Liden, Ricardo. *Algoritmos Genéticos*. Ed. Ciência Moderna, 3ª edição, 2012.
- [32] Pozo, A.; Cavaleiro, A. F.; Isihida, C.; Spinosa, E.; Rodrigues, E.. *Computação Evolutiva*(2005). Departamento de Informática, Universidade Federal do Paraná.
- [33] Yepes, I. *Sistemas Inteligentes: Uma incursão aos algoritmos genéticos*, 2005. Projeto ISIS UFRS.
- [34] Explicação do Algoritmo Genético. www.professor.webizu.org/ga.Disponível em 09 de novembro de 2011.
- [35] Neurônio. <http://www.google.com.br/search?q=neuronio>.Disponível em 12/01/2011.
- [36] Braga, Antônio de Pádua; Carvalho, André Carlos Ponce de Leon Ferreira de; Ludermir, Teresa Bernarda. *Redes Neurais Artificiais: teoria e aplicações*. Rio de Janeiro: LTC, 2000.
- [37] Haykin, Simon. *Redes Neurais - Princípios e Prática*, Bookman, 2001.

- [38] Kovács, Zsolt László. *Redes Neurais Artificiais: Fundamentos e Aplicações*. 3ª edição, São Paulo: Livraria da física, 2002.
- [39] Kim, Kyungnam. *Face Recognition using Principle Component Analysis*.
- [40] Santos, A. R. *Identificação de faces humanas através de PCA-LDA e Rede Neurais SOM*, 2005. USP – Departamento de Engenharia Elétrica.
- [41] Anton, H., Rorres C. *Álgebra Linear com Aplicações*, Bookman, 2004.
- [42] Árvore Faces. <http://www.eternoretorno.com/2008/09/11/figuras-e-imagens-de-ilusao-de-otica/>. Disponível em 18 de janeiro de 2012.
- [43] MATLAB *User's Guide*. The MathWorks, 2009a.
- [44] Cambridge Laboratórios AT&T.
<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>. Disponível em 30 de julho de 2010.
- [45] Computer Vision Science Research Projects.
<http://cswww.essex.ac.uk/mv/allfaces/index.html>. Disponível em 30 de julho de 2010.
- [46] Rao, K Y P. *Discrete Cosine Transform – Algorithms, Advantages, Applications*, San Diego, Academic Press, 1990.
- [47] Batista, L. V. *Compressão de Sinais Eletrocardiográficos Baseada na Transformada Cosseno Discreta*. Tese de Doutorado, Pós-Graduação em Engenharia Elétrica, UFPB, Campina Grande, Brasil, 2002.
- [48] Pedrini, Hélio. *Análise de Imagens Digitais – Princípios, Algoritmos e Aplicações*.

- [49] Mitchell, M. *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, EUA, 1996.
- [50] Alberts, B. Johnson, A. et al. *Molecular Biology of the Cell*, 5ª Edição, Garland Publishing, Boston, EUA, 2007.
- [51] Lovelock, James; Margulis, Lynn. *Gaia – Uma teoria do conhecimento*; Ed. Gaia, 2ª Edição, 2000.
- [52] Konzen, P.H.A. et al. *Otimização de métodos de controle de qualidade de fármacos usando algoritmo genético e busca tabu*, Pesquisa Operacional, v.23, n.1, rio de Janeiro, 2003.

Apêndice A

Introdução a Transformada Discreta do Cosseno

A.1 Introdução

As transformadas integrais em especial, são utilizadas para reduzir a complexidade de problemas matemáticos, como por exemplo, a resolução de equações diferenciais e integrais. Um caso bem conhecido é a transformada de Fourier, que decompõe um sinal em seus componentes de frequência, ou seja, uma função complexa da frequência angular.

O desenvolvimento de algoritmos rápidos para o cálculo da transformada de Fourier em 1965 representou um grande impulso na aplicação de transformadas em processamento digital de sinais [46]. Similarmente, o surgimento em 1974 da Transformada Discreta do Cosseno (*Discrete Cosine Transform* - DCT) e seus algoritmos eficientes impulsionaram em especial aplicações em processamento de sinais, processamento de imagens e compressão de dados [46]. Desde então foram verificados experimentos com resultados promissores utilizando DCT em problemas como reconhecimento de padrões, compressão, filtragem e reconstrução.

Uma característica importante de algumas transformadas refere-se a sua capacidade de gerar coeficientes descorrelacionados, concentrando a maior parte da energia do sinal em um número reduzido de coeficientes, ou seja, corresponde a uma abordagem capaz de reduzir a correlação presente nos sinais. Quando o sinal se comporta como um processo estocástico, ou seja, com origem em processos não determinados ou eventos aleatórios, com determinadas propriedades, pode-se mostrar que a Transformada *Karhunen-Loève Transform* (KLT) é ótima, pois obtém descorrelação máxima e concentra a energia em um número mínimo de coeficientes

[24]. A KLT não é objeto de estudo neste trabalho, mas pode-se dizer que a sua implementação é muito elaborada e exige a estimação da matriz de covariância além do cálculo da transformada propriamente dita [47]. Neste trabalho foi utilizada a transformada sub-ótima DCT.

A DCT é uma função linear e inversível, $\mathbb{R} \rightarrow \mathbb{R}$ que expressa sinais como uma soma de funções discretas do cosseno. O sinal original é convertido para o domínio da frequência pela DCT e é possível converter o sinal de volta para o domínio do tempo aplicando-se a DCT inversa.

A.2 Transformada Discreta do Cosseno

Desde que foi apresentada por Ahmed et al. (1974) [48], a transformada discreta do cosseno (DCT) tem sido explorada amplamente pela comunidade de processamento de imagens, principalmente nas áreas de compressão, filtragem e extração de características. A DCT apresenta a propriedade da compactação da energia de regiões próximas da origem após a execução da transformada. Muitos coeficientes apresentam valores próximos a zero e, portanto, podem ser eliminados sem que ocorra perda significativa de informações contidas na imagem.

O conceito de energia está relacionado com o tamanho dos vetores de entrada e saída da transformada, sua definição é mostrada na Equação (A.1). Considera-se que uma transformada conserva a energia quando $\|x\|^2 = \|y\|^2$, para uma transformação da forma $y^T = Ax^T$. Com a DCT, os coeficientes que apresentam valores mais significativos se localizam nas áreas mais próximas da origem, facilitando dessa maneira a determinação das regiões que podem ser excluídas quando o objetivo é a compressão de dados.

$$\|v\| = \sum_{k=0}^{n-1} |v_k|^2 \quad (\text{A.1})$$

Considerando que os coeficientes localizados em regiões distantes da origem apresentam valores próximos à zero, mesmo com a eliminação destes, a quantidade de energia não sofre redução significativa. A Equação (A.2) mostra o núcleo da transformada do cosseno para uma função de entrada composta por N elementos, onde $n, u=0,1,\dots,N-1$ e $k(\alpha)=1/\sqrt{2}$ se $\alpha=0$.

$$b_{u,n} = k(n) \sqrt{\frac{2}{N}} \cos\left(\frac{\pi \cdot n(u + \frac{1}{2})}{N}\right) \quad (\text{A.2})$$

O núcleo da DCT pertence a um conjunto de representações obtidas a partir dos autovetores extraídos das variações da matriz mostrada na Equação (A.3).

$$A = \begin{bmatrix} \Theta & \Theta & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \cdot & \cdot & \cdot & \cdot \\ & & & -1 & 2 & -1 \\ & & & & \Theta & \Theta \end{bmatrix} \quad (\text{A.3})$$

A partir da matriz mostrada na Equação (A.4), obtém-se a DCT proposta por Ahmed et al, denominada de DCT. Esta transformada é amplamente utilizada em processamento e análise de imagens. Outros três núcleos são apresentados nas Equações (A.5) a (A.7) para o mesmo $k(\alpha)$ definido anteriormente.

$$A = \begin{bmatrix} -1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \cdot & \cdot & \cdot & \cdot \\ & & & -1 & 2 & -1 \\ & & & & -1 & -1 \end{bmatrix} \quad (\text{A.4})$$

DCT1

$$b_{u,n} = k(u)k(n)\sqrt{\frac{2}{N}} \cdot \cos\left(\frac{\pi \cdot nu}{N}\right) \quad (\text{A.5})$$

DCT3

$$b_{u,n} = k(u)\sqrt{\frac{2}{N}} \cdot \cos\left(\frac{\pi \cdot n(u + \frac{1}{2})}{N}\right) \quad (\text{A.6})$$

DCT4

$$b_{u,n} = \sqrt{\frac{2}{N}} \cdot \cos\left(\frac{\pi(n + \frac{1}{2})(u + \frac{1}{2})}{N}\right) \quad (\text{A.7})$$

Em *Discrete Cosine Transform – Applications* [46] existem quatro definições para DCT: DCT1, DCT2, DCT3 e DCT4. Neste trabalho a DCT2 é chamada apenas como DCT. Ela tem sido mais utilizada em processamento de sinais e de imagens devido a sua forte capacidade de compactação de energia: muitas informações do sinal tendem a se concentrar em poucos componentes de baixa frequência. Por esse motivo, o padrão JPEG usa a DCT na sua implementação. Nesse contexto, a imagem original corresponde à matriz de escala de cinza $f(x, y)$, de dimensões $a \times b$, com $a*b = N$. A; a aplicação da DCT resulta na matriz $F(u, v)$, também de dimensão $a \times b$. As variáveis x e y são as coordenadas no domínio espacial e u e v são as coordenadas no domínio da frequência.

A transformada unidimensional do cosseno e sua inversa para um vetor de entrada composto por N elementos são mostradas nas Equações (A.8) e (A.9), respectivamente, onde $n, u=0,1,\dots,N-1$ e b representa o núcleo da transformada.

$$y_u = \sum_{n=0}^{N-1} x_n b_{u,n} \quad (\text{A.8})$$

$$x_n = \sum_{u=0}^{N-1} y_u b_{n,u} \quad (\text{A.9})$$

A Figura A.1 ilustra o resultado da aplicação da DCT utilizando a linguagem Matlab 7.2, em uma imagem de uma face humana de dimensão 112 x 92 pixels do arquivo ORL [44]. As Figuras A.1 (a) e (b) mostram as imagens da face original e a imagem resultante da aplicação da DCT, respectivamente. Na Figura A1 (b) é ilustrado em destaque a área onde concentra os coeficientes de maior significado que é a área com maior concentração de energia

A Figura A.2 ilustra o resultado da aplicação da DCT utilizando a linguagem Matlab 7.2, em uma a imagem de uma face humana de dimensão 180 x 200 pixels do arquivo UK [45]. As Figuras A.2 (a) e (b) mostram a imagem da face em preto e branco, porém a imagem original é colorida é a imagem resultante da aplicação da DCT, respectivamente. Na Figura A.2. é ilustrado em destaque a área onde concentra os coeficientes de maior significado que é a área com maior concentração de energia

Os dados da Tabela A.1 mostram o nível de cinza da imagem da Figura A.1 nas coordenadas de (1,1) a (8,8) e a Tabela A.2 ilustra os coeficientes da DCT. A Tabela A.3 mostra os níveis de cinza da imagem da Figura A.2 nas coordenadas de (1,1) a (8,8) e a Tabela A.4 ilustra os coeficientes da DCT.

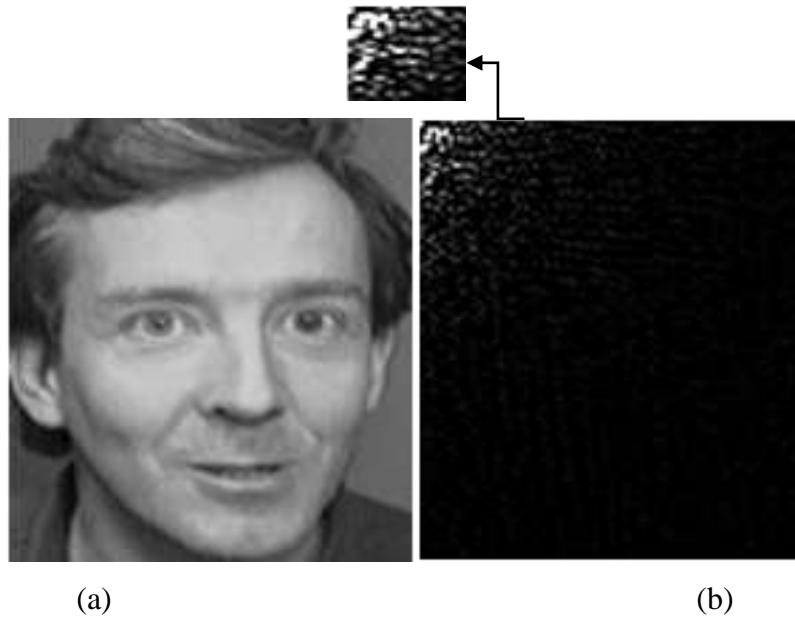


Figura A.1 - (a) Imagem original do arquivo ORL. (b) resultado da aplicação da DCT.

Tabela A.1-Escala de cinza da imagem da Figura A.1, coordenadas (1,1) a (8,8).

100	88	86	95	83	88	91	92
-107	110	04	110	96	75	83	91
-102	-100	-116	117	86	93	78	72
122	-124	-119	-92	127	82	94	81
110	116	101	-128	-91	-105	108	94
-126	126	122	101	97	-114	-105	-123
104	107	118	117	107	99	120	-117
98	94	85	102	108	117	116	114

Tabela A.2- Coeficientes da DCT da imagem da Figura A.1.

1,1939	-0,0484	-0,2842	-0,0027	-0,0481	0,0059	0,0086	0,0015
0,0226	-0,0061	-0,0042	0,0038	-0,0136	0,0035	-0,0077	-0,0025
-0,2019	-0,0038	0,0850	0,0209	0,0806	-0,0042	0,0382	-0,0265
-0,0037	-0,0151	0,1137	-0,0495	0,0431	0,0549	0,0329	-0,0001
-0,0575	0,0422	0,0797	-0,0027	0,0357	-0,0367	-0,0171	0,0311
-0,0817	0,0042	0,0085	-0,0391	0,0216	0,0031	-0,0074	0,0105
0,0163	0,0514	0,0160	-0,0223	-0,0333	0,0110	-0,0196	0,0144
0,1147	0,0307	-0,0313	-0,0153	0,0032	0,0364	0,0016	0,0029

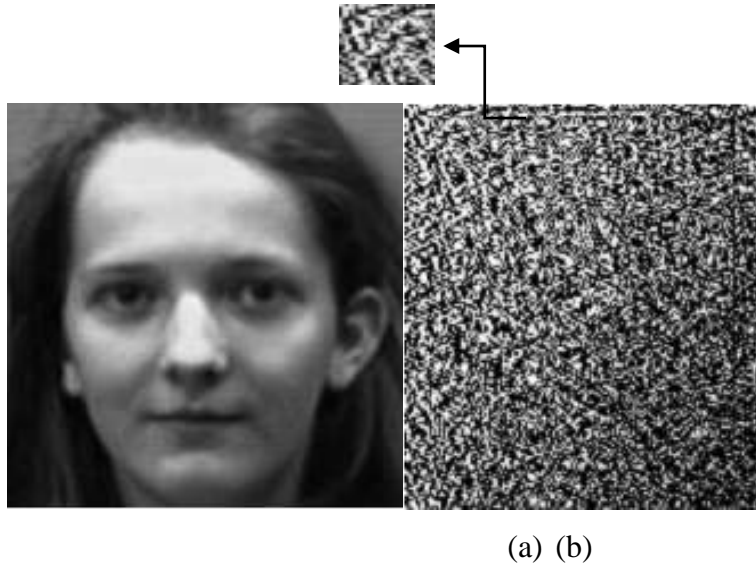


Figura A.2- (a) Imagem original do arquivo UK. (b) resultado da aplicação da DCT.

Tabela A.3-Escala de cinza da imagem da Figura A.2, nas coordenadas de (1,1) a (8,8).

91	93	95	96	96	95	93	91
96	97	98	99	99	98	97	96
98	98	98	98	98	98	98	98
6	96	95	95	95	95	96	96
95	95	94	93	93	94	95	95
6	96	96	96	96	96	96	96
5	96	98	98	98	98	96	95
93	95	97	98	98	97	95	93

Tabela A.4- Coeficientes da DCT da imagem da Figura A.2.

1,6770	0,1485	-0,4435	-0,3565	0,0273	-0,0781	0,0788	0,028
0,4299	0,0300	0,1246	-0,0371	0,0385	0,0369	-0,0241	0,0493
-0,1671	-0,0744	0,2916	0,0965	0,0536	0,1069	-0,0256	0,0531
-0,0003	-0,1183	0,1877	0,1752	-0,1387	0,0109	-0,0391	-0,1004
-0,1888	-0,0708	0,1082	0,1532	-0,0277	0,0429	-0,0204	-0,1409
-0,0750	-0,0254	0,0290	0,0295	-0,0060	-0,0944	0,0433	0,0287
0,1900	0,0529	-0,1584	-0,0285	-0,0275	-0,0495	-0,0103	0,0139
0,0903	0,0301	-0,0870	-0,0594	0,0326	0,0411	-0,0206	-0,0367

Apêndice B

Experiências Práticas: Banco de Imagens Familiar e Interação com o Meio Externo

B.1 Introdução

Com intuito de experimentar o algoritmo genético em uma situação prática, foram realizadas duas experiências simples. A primeira com um banco de fotografias familiar cujas poses foram retiradas da vida rotineira das pessoas. A segunda situação prática foi a instalação de uma controladora lógica programável (CLP) para transformar em uma ação externa o comando emitido pelo AG. Em ambas as situações os resultados foram positivos, primeiro se considerar que as fotos foram tiradas por amadores, sem nenhuma preocupação com a iluminação e sem nenhuma melhoria na qualidade das mesmas. Em segundo pelo fato do AG, após fazer a crítica lógica da taxa de acerto, aciona a CLP e esta por sua vez responde executando uma ação externa, realizando, portanto, o princípio da interatividade.

B.2 Experiência com fotografia familiar

Este arquivo possui fotografias de 13 pessoas e de um animal. Cada ente foi representado por 10 poses. As fotografias foram tiradas com uma câmera fotográfica da marca Olympus FE-4030 14 megapixels. Como convencionou-se ao longo deste trabalho, cada ente representado foi chamado de classe. Portanto, neste caso o banco de fotografias familiar possui 14 classes com 10 poses cada. A Figura B.1 mostra uma amostra desse arquivo.

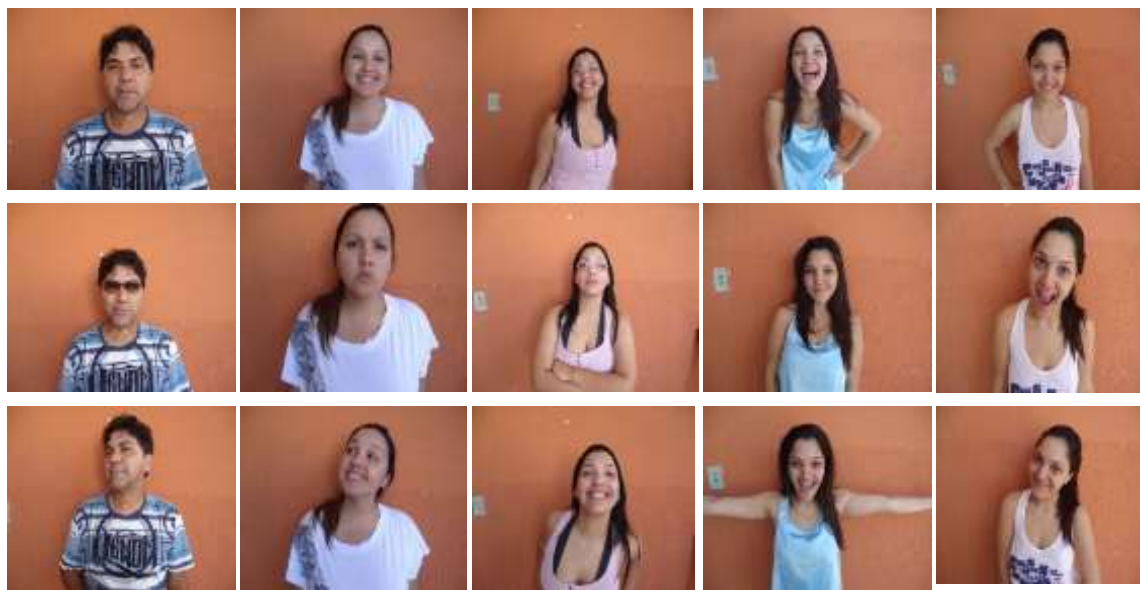


Figura B.1: Amostra das poses do arquivo familiar.

B.2.1 Resultados

O Algoritmo Genético (AG), executou o processamento do arquivo familiar especificado na seção B.2. O tamanho cada pose é de 3240 linhas por 4320 colunas e cada uma delas foi subdividida em 192 partes de 256 pixels. Foi aplicada a transformada direta do cosseno (DCT) nas subposes das quais foram retirados 4 coeficientes dessa transformada formando o vetor representativo de cada pose com 768 coeficientes.

A finalidade do AG foi identificar a classe de cada pose ou o grupo delas, quando fosse considerada uma ou mais poses (imagens query) de qualquer uma das classes pertencentes ao arquivo familiar.

A implementação do sistema conforme a estrutura do AG foi realizada considerando os seguintes parâmetros:

- População inicial: Criou-se 50 vetores indivíduos arbitrados;
- Máscara: Para cada vetor imagem criou-se um vetor máscara;
- Adaptação: Calculou-se a nota de adaptação ou fitness;
- Novo indivíduo: Seleção natural probabilística, através de roleta;
- Taxa de Cruzamento foi de 10 %;
- Taxa de Mutação Externa foi de 10 %;
- Taxa de Mutação Interna foi de 2 %;

- Número de Geração foi 100 (critério de parada); e
- Número mínimo de variáveis no processo de adaptação: 256.

Variações da experiência para o treinamento e o reconhecimento pelo AG utilizando o arquivo familiar

Os treinamentos e os reconhecimentos das faces foram realizados para os grupos de 2, 5, 10, 14 classes tomadas de maneira sequencial para as seguintes variações considerando a pose por classe:

- Treinou-se com nove poses (1ª à 9ª pose) e reconheceu-se com uma pose (10ª pose).
- Treinou-se com cinco poses (1ª à 5ª pose) e reconheceu-se com cinco poses (6ª a 10ª poses).

A Tabela B.1 mostra os resultados obtidos nos testes realizados para o reconhecimento de face utilizando o arquivo familiar onde *VAR* são as variações da experiência para o treinamento e o reconhecimento, *NPT* é o número poses para o treinamento, *NPR* é o número de poses para o reconhecimento de faces. Ainda têm-se que os parâmetros *AC* é a taxa de acerto em percentual, *TT* é o tempo de treinamento em segundos e *TR* é o tempo de reconhecimento em segundos.

Pode-se observar na Tabela B.1 que a taxa de acerto tem os valores de 100% para a quantidade de duas classes testadas. A taxa de acerto diminui bruscamente a partir de duas classes até 14 classes. Já o tempo de reconhecimento é muito pequeno sendo 0,445 s para 5 poses para o reconhecimento.

Tabela B.1: Resultados obtidos do reconhecimento de face utilizando 64 coeficientes da DCT de toda pose do arquivo família.

V A R	N P T	N P R	Parâmetros	Classes			
				2	5	10	14
1	9	1	AC (%)	100	80	70	70
			TT (s)	16,598	41,056	81,09	113,36
			TR (s)	0,0043	0,0186	0,059	0,107
5	5	5	AC (%)	100	80	70	65
			TT (s)	7,56	19,00	37,26	51,046
			TR (s)	0,0158	0,0709	0,237	0,445

As Figuras B.2 a B.4 mostram os gráficos dos resultados obtidos utilizando o arquivo familiar. Primeiro das taxas de acertos, depois os tempos de treinamento e de reconhecimentos e em seguida as poses reconhecidas erradamente.

Na Figura B.2 pode-se observar que o AG apresenta uma curva descendente onde mostra alta taxa de acerto para poucas classes e a diminuição brusca desta com o aumento de classes. Considerando uma pose para o reconhecimento a taxa de acerto varia de 100 % para duas classes a 70 % para 14 classes.

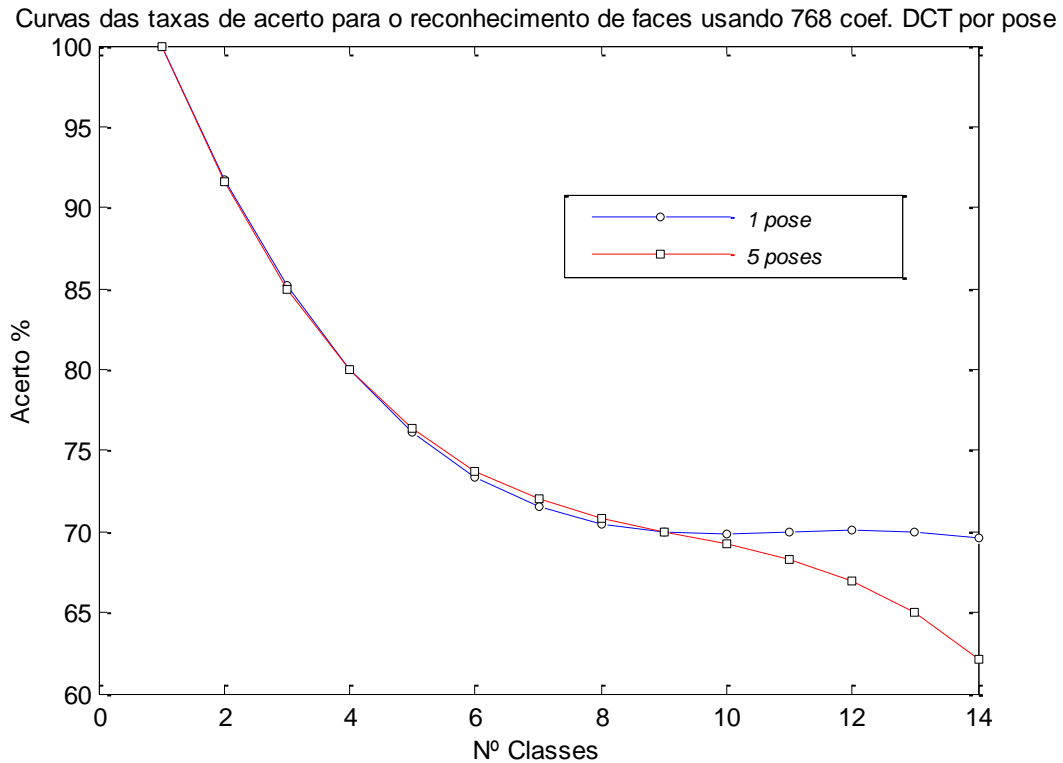


Figura B.2: Taxa de acerto considerando 768 coeficientes DCT por pose, arquivo familiar.

A Figura B.3 mostra o gráfico resultante dessa experiência referente aos tempos de treinamento e de reconhecimento. Quanto ao tempo de treinamento, observa-se que ele é maior para 9 poses para treinamento (1 pose para o reconhecimento) 113,36 s e menor para uma pose para 5 poses para o treinamento (5 poses para o reconhecimento) 51,046 s. Quanto ao tempo de reconhecimento mesmo sendo pequeno no geral, ele é maior para o reconhecimento com 5 poses, sendo 0,0158 s para 2 classes e 0,445 s para 14 classes..

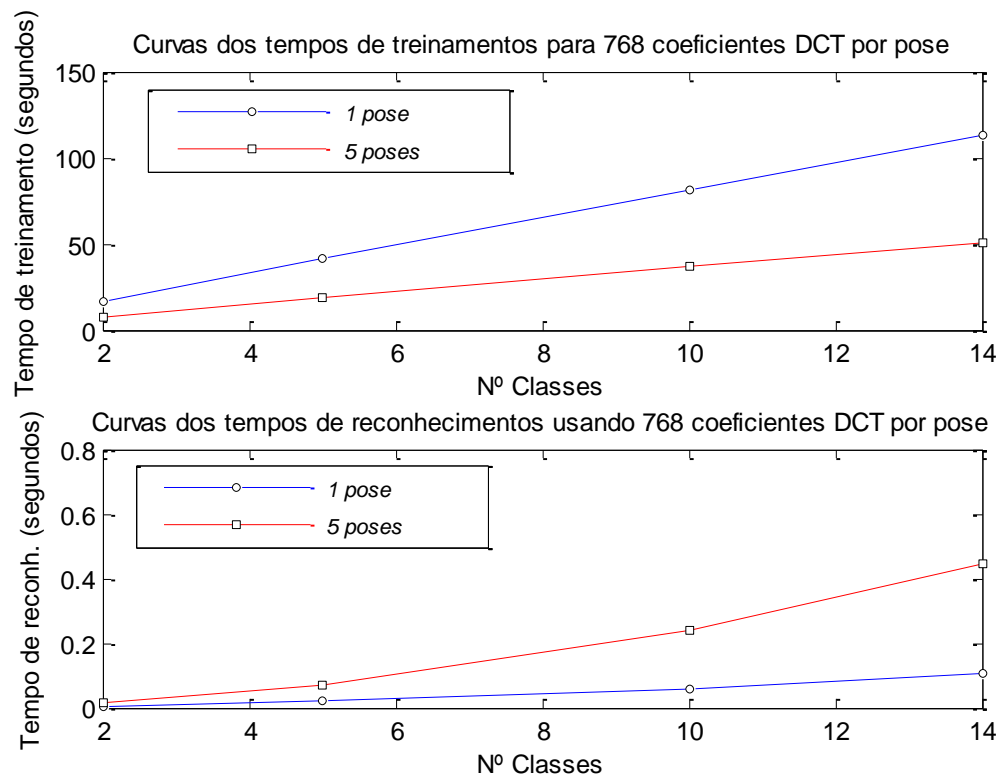


Figura B.3: Tempo de treinamento e de reconhecimento considerando 768 coeficientes DCT por pose, arquivo familiar.

A Figura B.4 mostra as imagens resultantes de um dos reconhecimentos errados. A pose 7 da classe 7 foi dado como semelhante a pose 1 da classe 10.



Figura B.4: A mostra de poses cujo reconhecimento foi errado

B.3 Experiência Com a Controladora Lógica Programável

Nesta experiência foi montado o circuito, mostrado na Figura B.5. Os componentes utilizados nele foram: um notebook com as especificações descritas na seção 3.2.1 do Capítulo III, uma placa CLP Andri MUSB 8E4RL e simbolizando os alarmes de uma aplicação real, uma pequena caixa de som e uma lâmpada.



Figura B.5: Circuito do disparo dos alarmes utilizando o AG.

Desta forma os “alarmes” do circuito da Figura B.5 foram acionados, após o processamento do programa AG reconhecer uma pose de acordo com a taxa de acerto desejada.