



UNIVERSIDADE FEDERAL DE PERNAMBUCO
GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
CENTRO DE INFORMÁTICA

**ESTRATÉGIAS PARA RECONHECER FACES
BASEADAS NA ANÁLISE DE COMPONENTES
PRINCIPAIS**

TRABALHO DE GRADUAÇÃO

Aluno: José Francisco Pereira (jfp@cin.ufpe.br)

Orientador: George Darmiton da Cunha Cavalcanti (gdcc@cin.ufpe.br)

Recife, janeiro de 2008

Resumo

A utilização de medidas biométricas para controle de acesso ou como forma de interação entre homens e máquinas tem crescido bastante, especialmente com o surgimento de novos dispositivos e mídias que exigem formas de interação diferente. Neste contexto o reconhecimento de faces destaca-se pela sua alta precisão e por ser uma das técnicas de medidas biométricas menos intrusivas.

Este trabalho de graduação consiste de uma avaliação de algoritmos baseados em PCA para o reconhecimento de faces. Fatores como taxa de acerto, redução da dimensionalidade e eficiência no armazenamento dos dados serão considerados na análise. Serão realizados experimentos sobre bancos de imagens muito utilizados por pesquisadores da área e que exploram problemas muito comumente encontrados em sistemas reais.

Dedico este trabalho a uma guerreira fundamental em minha trajetória de vida, sem a qual eu jamais teria alcançado êxito: a minha mãe Josefa.

Agradecimentos

Primeiramente agradeço à família pelo apoio incondicional durante toda a vida, pelo esforço em me dar a infra-estrutura necessária para conquistar meus objetivos.

Agradeço à minha noiva Cintya pelo companheirismo, pelos puxões de orelha nas horas certas, por todas as palavras de ânimo, em fim, muito obrigado por tudo.

Aos “irmãos da faculdade” pelas conversas, distrações, discussões e brigas que só contribuíram para o meu amadurecimento como profissional e como pessoa.

Aos professores e funcionários do Centro de Informática por me proporcionarem uma graduação de altíssimo nível. Agradeço em especial ao professor George Darmiton pela oportunidade e orientação neste primeiro trabalho juntos.

*“De tudo ficam três coisas:
a certeza de que estamos sempre começando...
a certeza de que é preciso continuar...
a certeza de que seremos
interrompidos antes de terminar...
portanto, devemos
fazer da interrupção um caminho novo...
da queda um passo da dança...
do medo, uma escada...
do sonho, uma ponte...
da procura... um encontro”*

Fernando Sabino

Sumário

Índice de Imagens	7
Índice de Tabelas	9
1. Introdução	10
2. Sistemas de Reconhecimento de Faces	12
2.1. Arquitetura de um SRF	12
2.2. Estado da Arte	13
3. Métodos e Técnicas	18
3.1. PCA (eigenfaces).....	18
3.2. PCA Bidimensional	20
3.3. PCA Modular (MPCA).....	23
4. Resultados.....	27
4.1. Bancos de dados.....	27
4.1.1. ORL	27
4.1.2. UMIST.....	28
4.1.3. Yale.....	28
4.2. Experimentos	29
4.2.1. Resultados para a base UMIST.....	30
4.2.2. Resultados para a base YALE.....	34
4.2.3. Resultados para a base ORL	38
4.2.4. Classificador	43
5. Considerações Finais	44
6. Referências	46
Anexo I – Tabelas de Resultados.....	48
Anexo II – Ferramenta Desenvolvida (hE).....	57

Índice de Imagens

1.	Fig. 2.1 – visão geral da arquitetura de um sistema de reconhecimento de faces.....	13
2.	Fig. 2.2 - projeção do ponto x sobre a linha x_1x_2	15
3.	Fig. 2.3 - HMM para o reconhecimento de faces.	17
4.	Fig. 3.1 - Divisão de uma imagem da base de faces Yale para valores de $N = 4$ e 16	24
5.	Fig. 4.1 - Exemplos de imagens da base de testes ORL.....	27
6.	Fig. 4.2 - Exemplos de imagens da base de testes UMIST	28
7.	Fig. 4.3 - Exemplos de imagens da base de testes UMIST	29
8.	Fig. 4.4 – Relacionamento entre os autovetores e a taxa de acerto correspondente à aplicação do PCA tradicional sobre a base de faces UMIST com faces selecionadas aleatoriamente.	30
9.	Fig. 4.5– Exemplo de imagens da base UMIST utilizada para treinamento e testes durante os experimentos	30
10.	Fig. 4.6 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o PCA tradicional sobre a base de faces UMIST utilizando cinco imagens para treinamento e cinco para teste	32
11.	Fig. 4.7– Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o IMPCA sobre a base de faces UMIST utilizando cinco imagens para treinamento e cinco para teste	33
12.	Fig. 4.8– Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o MPCA sobre a base de faces UMIST utilizando cinco imagens para treinamento e cinco para teste.	35
13.	Fig. 4.9– Exemplo de imagens da base Yale utilizada para treinamento e testes durante os experimentos	36
14.	Fig. 4.10 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o PCA tradicional sobre a base de faces Yale utilizando cinco imagens para treinamento e cinco para teste.	36
15.	Fig. 4.11– Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o IMPCA sobre a base de faces Yale utilizando cinco imagens para treinamento e cinco para teste.	37
16.	Fig. 4.12– Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o MPCA sobre a base de faces Yale subdividindo-se a imagem em 4 partes.....	39
17.	Fig. 4.13– Exemplo de imagens da base ORL utilizada para treinamento e testes durante os experimentos	39

18. Fig. 4.14– Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o PCA tradicional sobre a base de faces ORL utilizando cinco imagens para treinamento e cinco para teste	40
19. Fig. 4.15– Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o IMPCA sobre a base de faces ORL utilizando cinco imagens para treinamento e cinco para teste.	41
20. Fig. 4.16– Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o MPCA sobre a base de faces ORL subdividindo-se a imagem em 4 partes.....	42

Índice de Tabelas

1. Tab. 4.1– Resumo dos dados obtidos executando-se as técnicas sobre a base de dados UMIST. Os dados em destaque representam os melhores resultados da técnica.	33
2. Tab. 4.2– Resumo dos dados obtidos executando-se as técnicas sobre a base de faces Yale. Os dados em destaque representam os melhores resultados da técnica.	37
3. Tab. 4.3– Resumo dos dados obtidos executando-se as técnicas sobre a base de faces ORL. Os dados em destaque representam os melhores resultados da técnica.	41
4. Tab. 4.4– Número de coeficientes utilizados para representar uma característica da face.....	42
21. Tab. 4.5– Tempo médio utilizado pelas técnicas para processamento das imagens e o desvio padrão entre parênteses.	42

1. Introdução

Cada vez mais cresce a necessidade de controle de acesso a lugares, serviços e informações. É crescente também a busca por soluções mais eficientes na identificação dos indivíduos. Neste contexto, a biometria, que consiste no uso de características biológicas como mecanismo de identificação, tem sido utilizada com resultados bastante promissores. A biometria parte do princípio de que o corpo humano é sua senha e que todo indivíduo é único. Dentre as características utilizadas para identificação podem ser destacadas a íris, a retina, a impressão digital, o formato dos rostos ou até mesmo a geometria da mão.

Outro uso bastante promissor diz respeito à interação entre homens e máquinas onde ações e gestos podem ser utilizados como entrada do sistema. Além disso, novos dispositivos e mídias exigem novas formas de interação além dos tradicionais teclados, mouses e telas. O entendimento, por parte dos sistemas, do estado emocional e ações dos usuários também será muito beneficiada com o desenvolvimento das tecnologias de reconhecimento baseadas em biometria.

O reconhecimento de faces é um método de reconhecimento de indivíduos baseados em suas características faciais. Ela destaca-se por ser uma das técnicas biométricas que exigem menos interação do usuário. Pesquisas na área vêm sendo conduzidas a mais de trinta anos e como consequência o estado da tecnologia de reconhecimento de faces está bastante avançado.

Diversas aplicações que fazem uso de técnicas de extração e reconhecimento de faces estão no mercado. Estes sistemas são aplicados em diversos contextos como controle de acesso a universidades ou a parques de diversões, identificação de usuários de cartão de crédito ou ainda reconhecimento de criminosos. Neste caso uma imagem captada é processada e comparada com uma base de imagens pré-existente.

A tecnologia de reconhecimento de faces envolve duas grandes áreas de pesquisa. A primeira dela é bastante utilizada em vários outros campos de processamento de imagens, ela diz respeito à segmentação da imagem e detecção de padrões, no caso, detecção de faces. A outra técnica envolvida diz respeito à representação e reconhecimento de faces encontradas na etapa anterior, de modo que possa ser utilizada para comparação com faces da base de conhecimento. O reconhecimento juntamente com a detecção e extração de faces de imagens constituem uma área de pesquisa fundamental em visão computacional e análise de padrões.

A segmentação ou detecção de faces, segundo [9], é o ato de localizar as faces presentes em uma imagem sendo considerada como uma etapa de pré-processamento e independente do sistema de reconhecimento. Ainda segundo o autor esta etapa não se trata de um procedimento simples, pois em imagens, muitos objetos têm forma semelhante a faces. Há diversas técnicas de reconhecimento de padrões de objetos em imagens e especificamente para reconhecimento de faces, entretanto, elas não serão analisadas nesse trabalho.

A etapa de reconhecimento de faces é responsável por representar a imagem segmentada na etapa anterior e compará-la às faces da base de conhecimento. Na maior parte das técnicas a imagem a ser processada tem sua informação transformada em uma nova representação que facilite o processo de comparação entre as imagens e diminua os custos de processamento e armazenamento, esta nova representação geralmente é menor e com dados mais representativos. Esta etapa do reconhecimento pode ser feita sobre imagens ou vídeo.

Há diversas abordagens que podem ser utilizadas para o reconhecimento de faces, elas, em geral, se caracterizam em função das regiões da face que serão utilizadas para análise e representação. Pode-se utilizar toda a informação presente na imagem ou apenas os dados de regiões mais representativas. Existem diversas técnicas de extração de características e representação de faces humanas em cada uma das abordagens. Dentre elas destacam-se as técnicas baseadas em PCA (do inglês, *Principal Component Analysis*) como as mais utilizadas e com resultados mais promissores.

Duas técnicas recentes baseadas em PCA para extração de características e reconhecimento de faces serão estudadas neste trabalho. Elas sugerem melhorias sobre diversos aspectos como custo de processamento e melhoria no desempenho de reconhecimento de faces afetadas por variações de iluminação, expressões faciais e posicionamento da face.

No próximo capítulo será detalhada a arquitetura de um sistema de reconhecimento e as principais técnicas de representação utilizadas. No capítulo 3 serão apresentadas as técnicas em estudo neste trabalho. No capítulo 4 serão detalhados os experimentos e os resultados obtidos além das bases de faces utilizadas. E, por fim, no capítulo 5 será feita uma análise crítica sobre cada uma das técnicas em estudo.

2. Sistemas de Reconhecimento de Faces

O reconhecimento de face é uma das poucas tecnologias biométrica de reconhecimento que reúne duas excelentes características, possui elevada taxa de reconhecimento e é pouco intrusiva no que concerne à interação entre os usuários e o sistema. Desta forma, a tecnologia pode ser utilizada de forma transparente e com resultado muito satisfatório. Neste contexto, sistemas de reconhecimento de faces têm ganhado cada vez mais atenção pelos setores industriais e acadêmicos. Estes sistemas podem ser complexos e com grande volume de dados como os utilizados em investigações criminais ou simples e com base de conhecimento menor como os softwares embarcados em celulares que reconhecem a face do proprietário.

Independentemente da aplicação e do tamanho da base de conhecimento utilizada pode-se destacar um núcleo comum a estes sistemas. Especialmente no que diz respeito as duas macro-etapas de todo o processo: a detecção das faces em imagens e a representação e reconhecimento delas.

2.1. Arquitetura de um SRF

O conjunto de passos padrão para reconhecimento de faces é constituído pelas seguintes etapas:

- *Aquisição da Imagem*: etapa na qual a imagem contendo a face é obtida do meio externo. Geralmente esta etapa exige o uso de equipamentos para obtenção da imagem do ambiente.
- *Detecção da Face*: nesta etapa a imagem obtida é segmentada e todas as faces presentes na imagem são extraídas. Geralmente esta etapa pode se dividida em outra responsável por melhorar a qualidade da imagem obtida.
- *Extração de características*: para cada imagem extraída na etapa anterior é realizada uma mudança na representação das faces. Ela deixa de ser representada por uma matriz de pixels e passa a ser representada por um conjunto de características, geralmente numéricas, de menor tamanho. Esta etapa é primordial para o desempenho do sistema e muito relacionada às etapas seguintes do processamento das imagens.
- *Reconhecimento da face*: é responsável pela identificação ou verificação das faces obtidas nas etapas anteriores e por este motivo é bastante dependente das demais etapas. Geralmente é utilizado um classificador para comparação entre as imagens da base de conhecimento e as utilizadas no teste o que influencia diretamente no desempenho dos sistemas.

Uma visão geral da arquitetura de um sistema de reconhecimento de faces pode ser visto no diagrama a seguir:

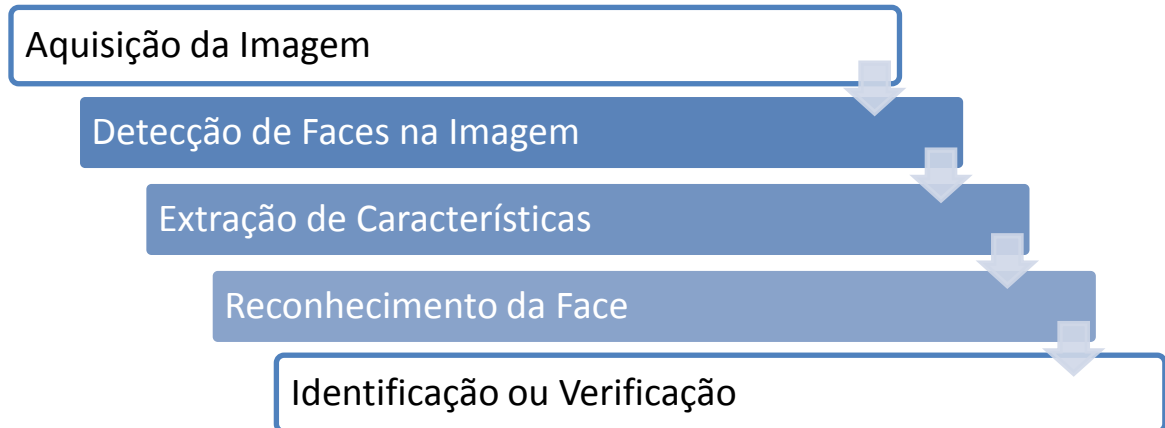


Fig. 2.1 – visão geral da arquitetura de um sistema de reconhecimento de faces.

Neste trabalho serão abordadas apenas as etapas de extração de características e de reconhecimento de faces. Entretanto, há problemas a serem superados nas etapas de aquisição e detecção de faces nas imagens. A solução desses problemas é de fundamental importância para o bom desempenho do sistema.

Na etapa de aquisição da imagem destacam-se, entre outros, as condições de iluminação do ambiente, a qualidade dos dispositivos de captura e a detecção do melhor angulo para obtenção das imagens.

A extração das faces presentes na imagem é bastante complexa. A detecção está associada a um grande conjunto de variáveis que interfere diretamente na qualidade da extração. Nesta etapa é necessário reconhecer uma determinada região da imagem (um conjunto de pixels) como sendo uma face. A detecção de faces está condicionada a fatores como o posicionamento da face em relação ao plano da imagem, o uso de objetos ou a presença de elementos da própria face como barba ou bigode, expressões faciais, posicionamento do dispositivo de captura de imagens além de variações da própria imagem como qualidade e escala dos objetos.

Existem diversas técnicas para tratar os problemas relacionados à detecção, entretanto, elas não serão analisadas, pois neste trabalho presume-se que todas as faces já tenham sido extraídas e normalizadas para uma resolução comum.

2.2. Estado da Arte

Uma vez segmentadas todas as faces presentes na imagem inicia-se o processo de extração de características e reconhecimento. Normalmente a extração de características está intrinsecamente relacionada à etapa de reconhecimento.

Em relação à abordagem utilizada no reconhecimento de faces podemos classificar as técnicas em três grandes grupos, a abordagem holística, a abordagem baseada em característica e a abordagem híbrida.

Na abordagem holística toda a informação da face, ou da imagem que contém a face, é utilizada como entrada do sistema de reconhecimento. Alguns exemplos de técnicas que utilizam a abordagem são *eigenfaces*, que se destaca como base das técnicas mais promissoras no reconhecimento de faces, *eigenfaces* probabilísticas, *fisherfaces*, Proximidade das Linhas de Características (NFL, do inglês *Nearest Feature Lines*) e Análise de Componentes Independentes (ICA, do inglês *Independent Component Analysis*). Todas as técnicas são baseadas no método de Análise de Componentes Principais (PCA) utilizado para representar o conjunto de dados em uma dimensão menor.

Nas abordagens baseadas em características apenas algumas regiões da face como nariz, boca e olhos são segmentados e utilizados como entrada para classificadores estruturais. *Hidden Markov Model* e *dynamic link architecture* são exemplos de técnicas que pertencentes a esta categoria.

A abordagem híbrida é inspirada na forma como o sistema de visão humano percebe as características locais e a informação global presente na face, mesclando as duas abordagens citadas anteriormente. Entre as técnicas que utilizam a abordagem podemos destacar a *eigenfaces* modular, características locais híbridas e métodos baseados em componentes.

As diversas técnicas existentes para o processamento e reconhecimento de faces foram agrupadas sob as três abordagens considerando apenas o tratamento que dão às informações presentes na imagem. Dentro de uma mesma abordagem há técnicas que fazem uso de diversas tecnologias e teorias diferentes. Entre estas tecnologias pode-se citar o uso de redes neurais, cadeias de Markov, estatística e combinações destas. A seguir serão apresentadas quatro técnicas como exemplo de cada uma das abordagens citadas.

Eigenfaces

É um método de análise estatística e extração de características de faces humanas. A técnica baseia-se no fato de que qualquer face humana pode ser representada por uma combinação linear de *eigenfaces*. Entenda *eigenfaces* como sendo um conjunto de autovetores extraídos da matriz de covariância de todas as faces que constituem a base de dados. O trabalho [7] publicado por Pentland a mais de uma década tornou-se padrão na área de pesquisa servindo como referência de desempenho para novas técnicas até hoje.

No método as imagens são tratadas como vetores normalizados, isto é, que foram subtraídos da média. Destes vetores é gerada a matriz de covariância da qual são extraídos os autovetores (*eigenfaces*). Desta forma, qualquer imagem pode ser projetada no espaço-face definido simplesmente multiplicando os vetores normalizados das imagens pelos *eigenfaces* extraídos da matriz de covariância.

Como a técnica aplica PCA diretamente ela não destrói qualquer informação da imagem sendo possível uma posterior reconstrução das imagens. Embora a técnica garanta uma elevada taxa de reconhecimento ela é bastante sensível a modificações de iluminação, pose e tamanho das faces.

Proximidade de Linha de Características

Em seu trabalho [10] publicado em 1999 Li apresenta uma técnica holística baseada em Proximidade de Linhas de Características (NFL, do inglês *Nearest Feature Lines*) para tratar alguns dos problemas apresentados pela técnica eigenfaces.

A técnica representa cada imagem como um vetor, ou um ponto no espaço, e define o conceito de *feature* (característica) como sendo a representação de uma classe. Para se obter uma *feature* são necessários pelo menos dois protótipos de uma mesma classe. Definindo x_1 e x_2 como dois protótipos quaisquer de uma classe da base, pode-se definir um vetor (x_1x_2) passando pelos dois pontos do espaço. Este vetor é conhecido por linha de característica (FL, do inglês *Feature Line*) da classe à qual as imagens pertencem. O vetor definido pelos pontos representa uma aproximação entre os dois protótipos (imagens) que podem ter sido obtidos em diferentes condições de iluminação, posicionamento da face ou com diferença de escala.

Para uma classe de teste qualquer x sua classificação é baseada na distância entre seu ponto no espaço (característica) e as linhas de características dos vetores da base (Fig. 2.2). Esta distância é definida pela equação $d(x, x_1x_2) = |x - p|$. A classe representada pela linha mais próxima do ponto x será atribuída ao protótipo de teste.

Experimentos feitos por [10] com a técnica sobre bases de faces de Cambridge, Bern, Yale e Harvard comprovaram o melhor desempenho da técnica quando aplicada sobre imagens com variação de iluminação, expressões faciais e posicionamento da face quando comparados com a técnica de eigenfaces.

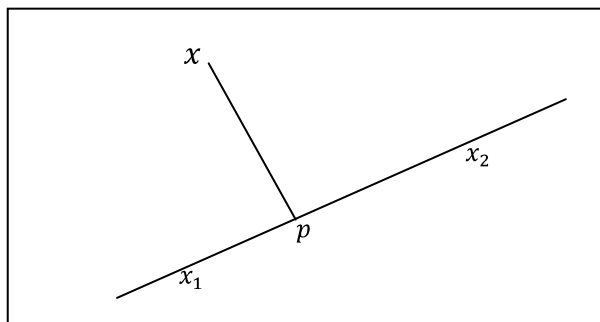


Fig. 2.2 – projeção do ponto x sobre a linha x_1x_2 .

Fisherfaces

Belhumeur [12] propôs o método *fisherfaces* utilizando PCA e Análise de Discriminantes Lineares de Fisher (LDA, do inglês *Fisher's Linear Discriminant Analysis*) para produzir um subespaço muito semelhante ao definido na técnica de *eigenfaces*. Este método visa solucionar um dos principais problemas presentes na técnica *eigenfaces*. Nela a dispersão total das imagens é maximizada de forma que pode influenciar a dispersão entre os protótipos de uma mesma classe. Uma maior separação entre os elementos de uma mesma classe pode acarretar em aumento da taxa de erro de classificação quando há variação significativa de iluminação e posicionamento da face entre imagens de uma mesma classe.

Considerando que grandes variações de iluminação e pose de uma mesma face são geralmente mais complexas de serem tratadas do que as diferenças entre faces de indivíduos diferentes nas mesmas condições, um sistema de reconhecimento de faces mais robusto deveria ser capaz de tratar este problema. O método *fisherfaces* toma proveito de informações intra-classe e procura minimizar a variação entre protótipos da mesma classe e conseqüentemente, maximiza a separação entre elementos de classes distintas. Nestas condições, os problemas de classificação erradas decorrentes de variações ocorridas em imagens da mesma face serão menores, melhorando o desempenho de sistema que utilizam a técnica.

Modelos Escondidos de Markov

Os resultados obtidos com a aplicação dos Modelos Escondidos de Markov (HMM, do inglês *Hidden Markov Models*) como técnica de reconhecimento de faces tem obtido bastante animadores, especialmente em faces com variações de iluminação, expressões faciais e orientação. HMM é um conjunto de modelos estatísticos usado para caracterizar propriedade de sinais. Já vem sendo utilizado há bastante tempo, com muito êxito, para reconhecimento de fala e de caracteres onde os dados possuem apenas uma dimensão.

O sistema de reconhecimento a ser modelado é entendido como sendo um Processo de Markov [13] com parâmetros desconhecidos, sendo o objetivo encontrar a melhor configuração para os parâmetros escondidos a partir dos modelos observáveis. Cada estado da HMM tem uma distribuição de probabilidade que define as possíveis saídas, visto que cada estado é observável no modelo de Markov regular.

Em [11] o autor utilizou a técnica HMM para reconhecimento de faces baseado nos vetores de características obtidos pela transformada discreta do cosseno (DCT, do inglês *two dimensional discrete cosine transformation*). O autor tomou proveito das propriedades de compressão da DCT para extrair as características da imagem e reduzir a dimensionalidade dos dados. A Fig. 2.3 apresenta as transições e a estrutura da cadeia aplicada ao reconhecimento de faces.

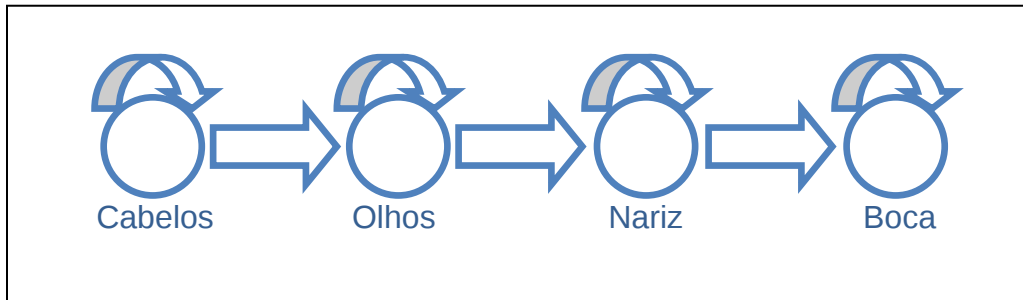


Fig. 2.3 – HMM para o reconhecimento de faces. Estados e possíveis transições.

Para o tratamento de faces humanas a técnica extrai os principais componentes como cabelo, olhos, nariz, boca e queixo e os representa como estados da cadeia unidimensional. Cada um desses componentes tem sua dimensionalidade reduzida com o uso do DCT. Desta forma, os vetores de observação podem ser reduzidos significativamente tornando o sistema muito eficiente do ponto de vista de tempo de processamento e mantendo uma elevada taxa de reconhecimento. A convergência do modelo é baseada na mudança nos valores da matriz de probabilidades observáveis. Estes valores de estimativa são obtidos através dos vetores de observação associados a cada estado, que nada mais são do que a resposta de um determinado estado do modelo para um padrão apresentado.

Os métodos apresentados neste capítulo apresentam melhorias sobre a técnica eigenfaces no que diz respeito ao tratamento de imagens com variações muito presentes em sistemas de reconhecimento reais. É notória a melhoria na qualidade e desempenho dos sistemas de reconhecimento de faces atuais, entretanto, ainda há um longo caminho a se percorrer para tornar o desempenho do sistema independente das variações sofridas pelas imagens.

3. Métodos e Técnicas

Desde a formalização da técnica, baseada em PCA, para extração de características e representação de faces feita por Pentland [7], em 1991, ela se tornou referência para outros trabalhos de reconhecimento de face. Logo foram propostas diversas modificações da técnica visando adaptá-la a variações de iluminação e posicionamento, escala, tratamento matemático para modificações de iluminação nas imagens sem adicionar a informação à base dentre outras. Recentemente surgiram abordagens mais complexas que embora sejam baseadas em PCA apresentam melhores resultados em função de modificações na extração de características e na representação das faces, dentre elas podemos destacar o ICA, o Kernel PCA e o PCA Bidimensional que melhoraram a o desempenho da técnica formalizada por Pentland [7].

Neste trabalho serão analisadas duas novas abordagens de extração de características e representação de faces desenvolvidas recentemente além da abordagem tradicional. Uma delas, conhecida como PCA Modular (MPCA), baseia-se na subdivisão das imagens das faces para tirar proveito de variações parciais sofridas pelas imagens. Já a técnica PCA Bidimensional (IMPCA ou 2DPCA) rompe com a representação tradicional feita pelas técnicas baseadas em PCA e manipulam imagens como matrizes conseguindo trabalhar em dimensão menor e conseqüentemente tornando os dados mais representativos das amostras. As técnicas analisadas serão detalhadas nas seções seguintes.

3.1. PCA (eigenfaces)

Eigenfaces (auto-faces) é uma técnica de reconhecimento de padrões em faces baseado em análise estatística de imagens de várias faces. Mais precisamente, é um conjunto de autovetores obtidos a partir de uma matriz de covariância de vetores que representam. Desta forma, a face de qualquer indivíduo pode ser representada por uma combinação linear destes autovetores.

Esta abordagem de reconhecimento de faces foi inicialmente proposta por Pentland [7] motivado pela aplicação da técnica de análise de componentes principais (PCA) no reconhecimento de faces. A técnica havia sido modificada por Karhunen e Loeve, em 1990, que melhoraram a forma de redução de dimensionalidade das faces e tornou-se conhecida também como expansão de Karhunen-Loeve.

O processamento das imagens das faces pela técnica inicia-se pela representação adotada. Nela todas as imagens da base de dados são representadas como um longo vetor de dados, em lugar da tradicional forma de matriz. O conjunto das imagens forma um espaço (espaço-imagem) onde cada face é representada por um ponto.

Uma vez que as faces possuem estrutura bastante similar os vetores que as representam serão co-relacionados e tendem a se localizar em regiões próximas uns dos outros no espaço. Esta proximidade entre os vetores em alta dimensão, em um primeiro momento, dificulta a separação das imagens em

classes distintas. A idéia por traz da representação das imagens através dos autovetores da matriz de covariância das faces é encontrar um subespaço onde vetores de menor dimensionalidade possam representar as faces.

Calculando as representações das faces

Considerando que as imagens utilizadas nos experimentos foram normalizadas para a resolução de $m \times n$ pixels. Elas devem ser representadas por vetores de dimensão $m \cdot n$, ou um ponto no espaço de mesma dimensão. Nos experimentos todas as imagens foram normalizadas para a resolução de 92×112 sendo necessários 10304 coeficientes para representá-las

Desta forma um conjunto de faces será representado por um conjunto de pontos no espaço (imagem-espaço). Pela similaridade estrutural entre as faces (presença de olhos, nariz, boca etc.) estes pontos não estarão distribuídos aleatoriamente no espaço de forma que será possível representá-los em uma menor dimensão. Os vetores base deste novo subespaço podem ser os autovetores associados à matriz de covariância das imagens da base.

Considere que as imagens $I_1, I_2, \dots, I_M$ representam as M imagens presentes na base de treinamento do experimento. A média das imagens é obtida da seguinte forma:

$$\bar{A} = \frac{1}{M} \sum_{i=1}^M (I_i) \quad (1)$$

Todas as imagens da base são normalizadas de acordo com a média calculada.

$$Y_i = I_i - \bar{A} \quad (2)$$

Desta forma, é definido um conjunto de vetores $Y_1, Y_2, \dots, Y_M$, centralizados na média, que serão utilizados para cálculo da matriz de covariância dos exemplos de treinamento.

$$C = \frac{1}{M} \sum_{i=1}^M (Y_i \cdot Y_i^T) \quad (3)$$

Os autovetores da matriz de covariância C são extraídos e os M' mais significantes, que possuem os maiores autovalores associados, são escolhidos. Com estes autovetores são calculados os pesos de cada imagem que constitui a base de treinamento através da fórmula

$$W_{iK} = E_K^T \cdot (Y_i) \quad \forall i, K \quad (4)$$

Sabendo que E_K 's são os autovetores associados aos maiores autovalores e K varia de 1 a M' .

Da equação (4) percebe-se que a representação das imagens W_i é feita através de pesos indexados pela quantidade (K) de autovetores extraídos da matriz de covariância. Desta forma, torna-se perceptível a relação entre o número de componentes principais adotados na representação da imagem e a quantidade de autovetores extraídos da matriz de covariância. Para cada imagem da base será obtido um vetor W_i de dimensionalidade K , sendo K o número de componentes principais utilizados na representação.

Método de classificação

Uma vez extraídos os vetores de componentes principais das imagens é utilizado um classificador baseado na técnica k -nn para classificação. O cálculo entre os vetores de características é definido pela expressão a seguir:

$$d(W_i, W_j) = \|W_i - W_j\|$$

Sabendo que $\|W_i - W_j\|$ representa a distância euclidiana entre os dois vetores e componente principais W_i e W_j .

A classe do vetor da base W_i que possuir menor distância do vetor de teste W_j será atribuída como classe da imagem W_j , ou seja, adota-se $k = 1$ resultando em um classificador do tipo 1-nn.

3.2. PCA Bidimensional

Nas técnicas baseadas em PCA as matrizes de imagens precisam ser transformadas em vetores unidimensionais antes de ser submetido à técnica. Como resultado desta necessidade, os vetores-imagem resultantes da transformação trabalham em espaços de dimensão muito alta. Nestas condições, torna-se difícil analisar a matriz de covariância precisamente, em função do seu tamanho e a quantidade relativamente pequena de exemplos de treinamento. Fazendo uso das técnicas de Decomposição em Valores Singulares [4] (do inglês, *Singular Value Decomposition* - SVD) os autovetores podem ser calculados sem a necessidade de geração e processamento da matriz de covariância. Contudo, isso não implica que os autovetores podem ser avaliados de forma precisa, visto que são estatisticamente determinados pela matriz de covariância, independentemente do método adotado para obtê-los.

Com base nestas obrigações impostas pelas técnicas de representação de imagens baseadas em PCA foi desenvolvida, em 2004, a técnica 2DPCA (do inglês, *two-dimensional principal component analysis*) por Zhang et al [3]. Na técnica proposta, diferentemente da representação utilizada na PCA, as imagens são representadas através de matrizes bidimensionais, desta forma, a etapa de

transformação da matriz em um vetor é dispensada. Além disso, diferentemente da matriz de covariância produzida pela técnica PCA a matriz gerada pela técnica é bem pequena. Como resultado, torna-se mais fácil avaliar a matriz mais precisamente, por representar melhor a base de imagens e requerer menos tempo de processamento na avaliação.

Calculando as representações das faces

O processo de extração de características das faces se diferencia levemente das técnicas baseadas no PCA tradicional apresentadas nas seções anteriores. Especialmente no que diz respeito à representação internada das imagens adotada pela técnica.

Considere uma imagem I_i de uma face de tamanho $m \times n$ representada por uma matriz A qualquer e um vetor X n -dimensional sobre a qual a imagem será projetada.

$$Y = AX \quad (10)$$

Com isso, obteremos o vetor Y m -dimensional que é chamado vetor de características projetada da imagem A . A avaliação da qualidade da projeção feita pelo vetor X pode ser feita através da medida de dispersão total das imagens projetadas. A dispersão total dos vetores produzidos na projeção das imagens pode ser medida através do traço da matriz de covariância gerada a partir destes vetores. Desta forma, o seguinte critério pode ser adotado:

$$J(X) = \text{tr}(S_x) \quad (11)$$

Sendo que S_x corresponde à matriz covariância dos vetores de características obtidos das imagens da base de treinamento e $\text{tr}(S_x)$ corresponde ao traço da matriz de covariância. A maximização do critério definido em (11) resulta em encontrar um vetor X , sobre o qual as imagens serão projetadas, que maximize a dispersão total dos vetores de características obtidos. A matriz de covariância S_x pode ser obtida por:

$$S_x = E(Y - EY)(Y - EY)^T = E[AX - E(AX)][AX - E(AX)]^T$$

$$S_x = E[(A - EA)X][(A - EA)X]^T$$

Logo,

$$\text{tr}(S_x) = X^T G_t X$$

Onde G_t pode ser expresso como:

$$G_t = E[(A - EA)^T(A - EA)]$$

A matriz G_t definida acima representa a matriz de covariância das imagens da base. De sua definição percebe-se que G_t é uma matriz quadrada não negativa de tamanho $n \times n$. A matriz G_t pode ser obtida diretamente através

das imagens da base de treinamento. Supondo que há M exemplos de imagens na base de treinamento, a j -ésima imagem de treinamento é representada por uma matriz de tamanho $m \times n$ denominada A_j ($j = 1, 2, \dots, M$) e a média de todas as imagens de treinamento é representada por $\bar{A}$. Com isso, podemos definir G_t como sendo:

$$G_t = \frac{1}{M} \sum_{j=1}^M (A_j - \bar{A})^T (A_j - \bar{A})$$

Desta forma, o critério definido em (11) resume-se a:

$$J(X) = X^T G_t X$$

O critério acima é conhecido como critério generalizado de dispersão total (do inglês, *generalized total scatter criterion*). O vetor unitário X que maximiza a expressão é o que maximiza a dispersão total das imagens projetadas sobre ele.

O vetor de projeção ótima X_{opt} é o que maximiza $J(X)$, isto é, o autovetor de G_t correspondente aos maiores autovalores. Assim como utilizado nas técnicas baseadas em PCA, em geral se escolhe um conjunto de vetores ótimos para extração dos pesos. Estes vetores devem ser ortonormais e maximizar o critério originalmente definido em (11). Desta forma, os vetores de projeção ótima, $X_1, X_2, \dots, X_d$ são os autovetores de G_t correspondentes aos d maiores autovalores.

Os vetores de projeção da técnica 2DPCA, $X_1, X_2, \dots, X_d$, serão utilizados para extração de características das imagens das faces. Para uma imagem A pode-se extrair os vetores de características a partir da seguinte expressão:

$$Y_k = A X_k, k = 1, 2, \dots, d$$

Com isso, é obtido um conjunto de vetores de características $Y_1, Y_2, \dots, Y_k$, chamados de vetores de componentes principais da imagem A . É importante notar que cada componente principal extraído com a técnica é um vetor, enquanto que na técnica PCA cada componente é um valor escalar (coeficiente).

Os vetores de componentes principais extraídos são utilizados para construir a matriz $B = [Y_1, \dots, Y_k]$ de tamanho $m \times d$ conhecida como a matriz de características da imagem.

Método de classificação

Após a transformação das imagens em matrizes de características é utilizado um classificador baseado na técnica k -nn. A distância calculada entre as matrizes de características, $B_i = [Y_1^{(i)}, \dots, Y_d^{(i)}]$ e $B_j = [Y_1^{(j)}, \dots, Y_d^{(j)}]$ é definida por

$$d(B_i, B_j) = \sum_{k=1}^d \|Y_k^{(i)} - Y_k^{(j)}\|$$

Sabendo que $\|Y_k^{(i)} - Y_k^{(j)}\|$ representa a distância euclidiana entre os dois vetores e componente principais $Y_k^{(i)}$ e $Y_k^{(j)}$.

A classe de uma matriz qualquer da base B_k que possuir menor distância da matriz de teste qualquer B_i será atribuída como classe da imagem B_i . Caracterizando assim um classificador do tipo 1-*nn*.

3.3. PCA Modular (MPCA)

O método de representação de faces baseados em PCA descrito na seção anterior tem seu desempenho bastante abalado sob condições de variação de pose e iluminação. Isso ocorre porque a técnica considera toda a informação presente na imagem, desta forma, pequenas variações que ocorrem apenas em partes da imagem acarretam mudanças em toda sua representação. Estas modificações na representação acabam diferenciando bastante a faces com variações das demais faces do mesmo indivíduo em condições normais de pose e iluminação dificultando sua classificação.

Neste contexto a técnica de PCA Modular (MPCA) proposta por Gottumukkal e Asari [1], em 2002, surge para melhorar esta deficiência do PCA tradicional pela subdivisão da imagem original em partes menores. A técnica considera que se a imagem for dividida em regiões menores e os vetores de pesos forem extraídos para cada uma das partes estes pesos conseguirão representar mais fielmente a informação local de cada região da face.

Variações como iluminação, expressões faciais e até mesmo rotação da face, em geral, comprometem apenas parte das informações presentes na imagem deixando outras regiões intactas, isto é, com a mesma informação das faces em condições normais. Com isso, as regiões não afetadas pelas alterações irão influenciar a classificação da imagem como pertencente à classe da imagem em condições normais. Por este motivo, é esperada uma melhoria na taxa de reconhecimento da técnica quando comparada ao PCA tradicional.

Calculando as representações das faces

Por utilizarem uma base teórica comum, os procedimentos necessários para extração das características da imagem utilizada pela técnica PCA Modular são muito parecidos com os utilizados pela técnica *eigenfaces*. Entretanto, há algumas peculiaridades a serem abordadas, especialmente a forma como as imagens são subdivididas.

Na técnica MPCA, a imagem é dividida em N partes menores. Desta forma, considerando uma imagem quadrada de tamanho L por L o tamanho de

cada sub-imagem será L^2/N pixels. Matematicamente, estas imagens podem ser representadas como:

$$I_{xij}(m,n) = I_x\left(\frac{L}{\sqrt{N}}(j-1) + m, \frac{L}{\sqrt{N}}(i-1) + n\right) \quad \forall i,j \quad (5)$$

Sabendo-se que x varia de 1 a M , sendo M o número de imagens no conjunto de treinamento, j e i variam de 1 a $\sqrt{N}$, sendo N o número de sub-imagens e m e n variando de 1 a $\frac{L}{\sqrt{N}}$. A função $I_{xij}(m,n)$ pode ser interpretada como sendo o valor da intensidade do pixel de coordenada (m,n) da sub-imagem de índice i,j da imagem original I_x .



Fig. 3.1 – Divisão de uma imagem da base de faces Yale para valores de $N = 4$ e 16 .

Outra característica importante da técnica é o cálculo da média das imagens e da matriz de covariância. Na técnica, há apenas uma média, independentemente da quantidade de partes na qual uma imagem tenha sido dividida e posteriormente apenas uma matriz de covariância é gerada. Embora o motivo de utilizar apenas uma média e uma matriz de covariância não tenha sido explicado pelos autores da técnica entende-se que esta abordagem visa relacionar as partes da imagem com o todo. Desta forma, o PCA Modular não se trata apenas da aplicação do PCA em cada parte que constitui a imagem e sim de relacionar cada parte com a imagem original e posteriormente realizar os cálculos necessários à representação de cada uma delas.

Desta forma, a média de todas as imagens é obtida como sendo:

$$\bar{A} = \frac{1}{M \cdot N} \sum_{x=1}^M \sum_{i=1}^{\sqrt{N}} \sum_{j=1}^{\sqrt{N}} (I_{xij}) \quad (6)$$

Examinando esta equação se percebe uma grande semelhança com a equação (1) apresentada na técnica PCA apenas adicionado da soma das partes que constituem a imagem original. Assim como na PCA tradicional apenas uma média é gerada para toda a base de treinamento.

Obtida a imagem média (ponto médio no espaço-face) as imagens da base de treinamento podem ser normalizadas. Matematicamente esta normalização equivale a uma transformação linear de troca de bases, ou seja, o

ponto de origem do espaço é transladado para o ponto médio e as normas de todos os vetores são modificadas. Esta normalização é feita pela expressão a seguir:

$$Y_{xij} = I_{xij} - \bar{A} \quad \forall x, i, j \quad (7)$$

Após a obtenção das médias podemos calcular a matriz de covariância das imagens da base.

$$C = \frac{1}{M \cdot N} \sum_{x=1}^M \sum_{i=1}^{\sqrt{N}} \sum_{j=1}^{\sqrt{N}} (Y_{xij} \cdot Y_{xij}^T) \quad (8)$$

Novamente, a similaridade das fórmulas é perceptível. A fórmula anterior se assemelha bastante com a forma de cálculo da matriz de covariância da técnica PCA (3) adicionado do somatório das partes que constituem cada imagem. Esta característica impõe à técnica a necessidade de que todas as partes tenham o mesmo tamanho para que possam ser utilizadas. Ao mesmo tempo, o fato de todos os cálculos serem feitos com base em vetores de dados menores o tamanho da matriz de covariância diminui proporcionalmente. Com isso, é obtido um ganho no tempo de processamento em função da extração de autovetores e autovalores ser feita sobre uma matriz menor.

A próxima etapa do processamento é extrair os autovetores da matriz de covariância associados aos maiores autovalores. Estes autovetores serão representados por $E_1, E_2, \dots, E_K$. Os pesos são calculados através da equação a seguir:

$$W_{xijk} = E_K^T \cdot (Y_{xij}) \quad \forall x, i, j, K \quad (9)$$

Sabendo-se que E_K^T representa a transposta do k -ésimo autovetor extraído da matriz de covariância e K assume os valores $1, 2, 3, \dots, M'$, x é o índice da imagem processada e assume valores de 1 a M , sendo M o número de imagens presentes na base de treinamento, j e i assumem valores de 1 a $\sqrt{N}$, onde N é a quantidade de partes que a imagem foi dividida. Da definição, percebe-se que cada face I_x será representada por $j \times i$ vetores de características W_x cada um com dimensionalidade K .

Desta forma, para uma única imagem I_x que tenha sido dividida em dezesseis partes ($N = 16$) e se queira utilizar dez componentes principais ($K = 10$) na sua representação, serão necessários dezesseis vetores de dez dimensões cada, totalizando assim cento e sessenta características ($N \times K = 160$). Logo, para representar imagens da base de treinamento com a técnica é necessário mais capacidade de armazenamento embora possa se chegar a um resultado de forma mais rápida e com melhor taxa de reconhecimento.

Método de classificação

Após a transformação das imagens em vetores de características principais é utilizado um classificador baseado na técnica k -nn. A distância calculada entre os componentes principais das imagens, $B_i = [W_{i,1}, W_{i,2}, \dots, W_{i,N}]$ e $B_j = [W_{j,1}, W_{j,2}, \dots, W_{j,N}]$ é definida por

$$d(B_i, B_j) = \sum_{k=1}^N \|W_{i,k} - W_{j,k}\|$$

Sabendo que $\|W_{i,k} - W_{j,k}\|$ representa a distância euclidiana entre os vetores de componentes principais da k -ésima parte da matriz de componentes W_i e W_j . Pela definição da distância é perceptível que ela se baseia na soma das distâncias das sub-imagens correspondentes.

A classe de uma imagem qualquer da base de treinamento B_j que possuir menor distância da imagem de teste B_i será atribuída como classe da imagem B_i , caracterizando dessa forma um classificador do tipo 1-nn.

4. Resultados

As técnicas de extração e representação de características de faces analisadas neste trabalho foram aplicadas em três bancos de imagens bem conhecidos na área (ORL, UMIST e Yale). Os resultados e a análise das técnicas em estudo serão detalhados nas próximas seções.

4.1. Bancos de dados

Os bancos de imagens foram selecionados para uso neste projeto em função da frequência com que aparecem em trabalhos relacionados e por explorar variações muito comuns em sistemas reais tais como: iluminação, rotações, expressões faciais e uso de objetos sobre a face. Cada uma das bases será detalhada a seguir.

4.1.1. ORL

A base de dados ORL [14] contém imagens de quarenta indivíduos cada um com dez imagens. Estas imagens foram tiradas em laboratório sob condições controladas de luminosidade e plano de fundo. As fotos de alguns indivíduos foram tiradas em momentos distintos, entretanto, devido ao pouco intervalo de tempo seus feitos não são perceptíveis. Expressões faciais como olhos abertos ou fechados e sorrisos além de detalhes faciais como a presença ou não de óculos também variam.

As imagens foram obtidas com pequenas variações de inclinação da face e rotação de aproximadamente 20° . É perceptível também uma variação de escala de aproximadamente 10% para alguns indivíduos. Todas as imagens estão em escala de cinza e estão normalizadas para a resolução de 92×112 pixels.



Fig. 4.1 – Exemplos de imagens da base de faces ORL

4.1.2. UMIST

A base de dados UMIST [15] é composta por 564 imagens em escala de cinza de 20 indivíduos obtidas em ambiente controlado com fundo homogêneo. Na versão normalizada da base as imagens possuem a resolução de 92 x 112 pixels. Nela cada indivíduo possui uma quantidade de imagens diferente podendo variar de 19 (indivíduo 1g) a 48 imagens (indivíduo 1s). A base será utilizada para teste de desempenho das técnicas quanto a grandes variações de rotação das faces.



Fig. 4.2 – Exemplos de imagens da base de faces UMIST

4.1.3. Yale

A base de dados Yale [16] é composta por 165 imagens em escala de cinza de 15 indivíduos – onze imagens diferentes por indivíduo. As imagens não foram obtidas em ambientes totalmente controlados e possuem variações do plano de fundo. Esta base explora diferentes expressões faciais (cansado, surpreso, sorrindo, triste etc.) o uso de objetos sobre face além de diferentes configurações de iluminação.

As imagens oficiais, que são disponibilizadas no site, possuem a resolução de 320 x 243 com as faces não centralizadas e localizadas em diferentes posições nas imagens. As faces de todas as imagens foram extraídas, à mão, centralizadas e normalizadas para a resolução de 92 x 112 pixels.



Fig. 4.3 – Exemplos de imagens da base de faces Yale

4.2. Experimentos

As três técnicas (PCA, PCA Modular e PCA Bidimensional) foram aplicadas sobre as bases apresentadas na seção anterior sob diferentes configurações de treinamento e teste. Em todos os experimentos foram medidos os tempos necessários para processamento e classificação das imagens do experimento. Para a realização dos experimentos todas as bases foram normalizadas para uma resolução comum (92 x 112 pixels).

Foram realizados testes variando as quantidades de imagens utilizadas para treinamento e testes. Em um conjunto de testes, o k -ésimo experimento utilizou as k primeiras imagens para treinamento e as demais para teste. Além disso, foram feitos experimentos modificando a quantidade de autovetores (M') utilizados na representação das faces. Foram analisadas e escolhidas as melhores configurações de cada técnica ponderando apenas a quantidade de autovetores e a taxa de acerto, pois, no escopo deste trabalho, o custo computacional não foi considerado como critério para escolha da configuração.

Com base nos resultados é perceptível que o aumento do número de autovetores implica no crescimento da taxa de reconhecimento, entretanto, o custo computacional, além do uso de memória, também cresce com o aumento da dimensionalidade.

O gráfico abaixo apresenta a taxa de acerto e quantidade de informação (soma dos k maiores autovalores normalizada em 100) em relação à quantidade de componentes principais. Os valores foram obtidos com base na aplicação da técnica PCA tradicional sobre a base de dados UMIST utilizando dez imagens selecionadas aleatoriamente sendo cinco imagens para treinamento e cinco para testes.

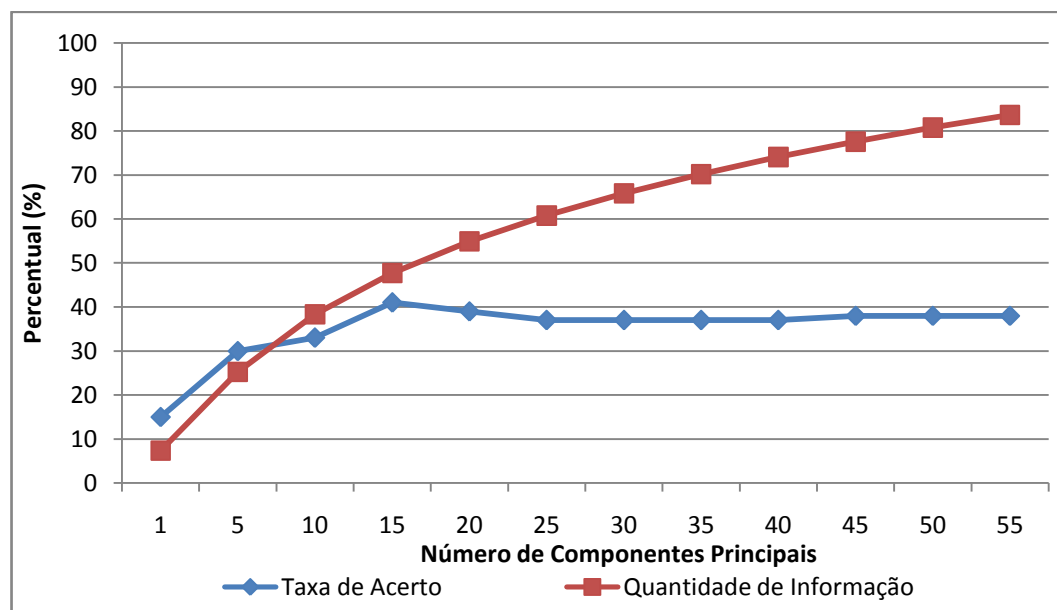


Fig. 4.4 – Relacionamento entre os autovetores e a taxa de acerto correspondente à aplicação do PCA tradicional sobre faces da base UMIST com faces selecionadas aleatoriamente.

Nas classificações não foram utilizados threshold para os testes logo, as faces são classificadas como pertencendo a uma das classes apresentadas, de forma que não há rejeição por parte do sistema.

4.2.1. Resultados para a base UMIST

A utilização desta base visa testar o desempenho das técnicas sobre grandes variações de posicionamento da face dos indivíduos. Esta base contém faces com rotações de até 180 graus.

Em função da variação na quantidade de imagens por indivíduos da base, foram escolhidas aleatoriamente dez imagens de cada indivíduo. Dentre as dez imagens selecionadas foram utilizadas, em um primeiro teste, cinco imagens para treinamento e as demais foram utilizadas para teste. Todos os testes realizados sobre a base foram feitos utilizando as mesmas imagens de treinamento e teste visando comparar o desempenho das técnicas nas mesmas condições.

A seguir são apresentadas algumas das imagens utilizadas para o treinamento e para o teste:



Fig. 4.5 – Exemplo de imagens da base UMIST utilizada para treinamento e testes durante os experimentos.

PCA

Aplicando-se o PCA tradicional sobre as imagens da base UMIST conforme descrito anteriormente os resultados obtidos foram:

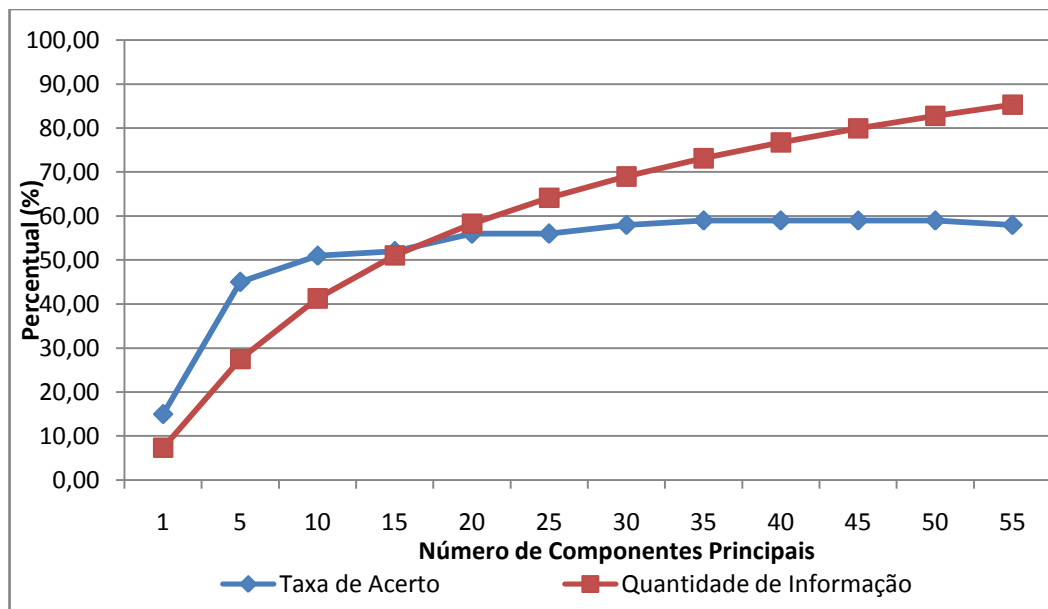


Fig. 4.6 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o PCA tradicional sobre a base de faces UMIST utilizando cinco imagens para treinamento e cinco para teste.

A partir do gráfico se pode observar que o melhor resultado para a base UMIST com o conjunto de teste mencionado foi obtido com 35 componentes principais que mantêm cerca de 73% da quantidade de informação do conjunto além do que a taxa de acerto se mantém estável em aproximadamente 59% com 30 ou mais componentes.

Outra informação importante que se percebe no gráfico é a distribuição homogênea da quantidade de informação dos primeiros autovalores obtidos com a técnica. Isso implica em um equilíbrio de relevância de cada componente e tem como principal consequência a necessidade de um maior número de componentes para representar as faces.

PCA BIDIMENSIONAL

Aplicando-se a técnica de PCA Bidimensional (2DPCA ou IMPCA) obtém-se o seguinte resultado:

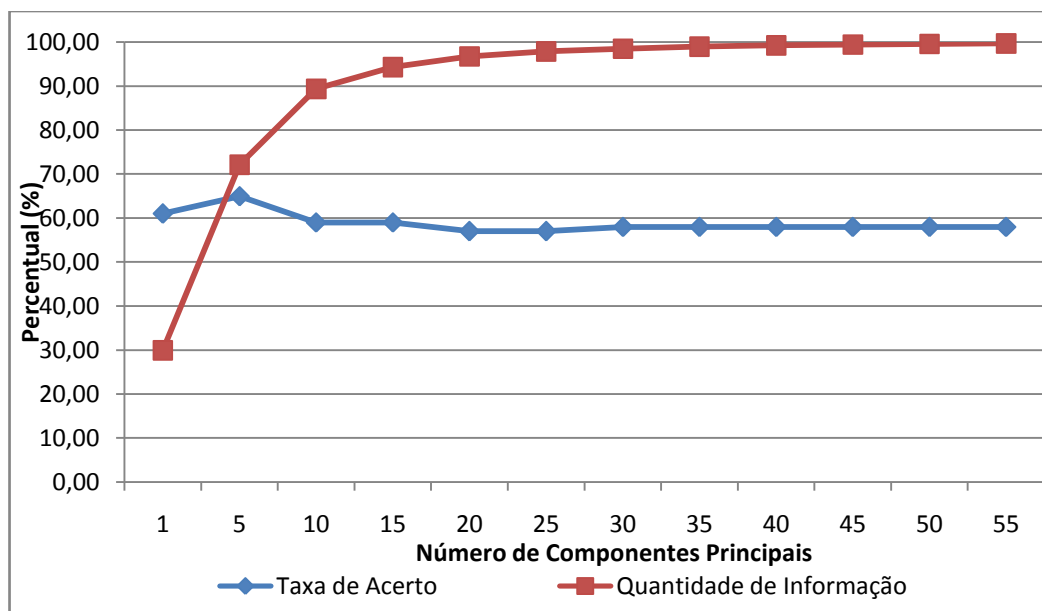


Fig. 4.7 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o IMPCA sobre a base de faces UMIST utilizando cinco imagens para treinamento e cinco para teste.

O gráfico mostra que, com baixa dimensão, a taxa de reconhecimento da técnica é bastante elevada quando comparado com o PCA. Entretanto, esta taxa diminui ao se aumentar o número de componentes e se mantém estável para valor igual ou superior a 20.

É perceptível também a maior relevância dos primeiros componentes em relação ao PCA, isso é, maior quantidade de informação contida nas primeiras componentes. Isso se deve em parte pela maior representatividade das faces pela matriz de covariância obtida durante o processamento. Além disso, a matriz é consideravelmente menor que a gerada pelo PCA resultando em menos autovetores implicando em um aumento na quantidade de informação associada a cada componente.

Vale ressaltar também que para a base de dados UMIST a técnica representa cada componente com 112 coeficientes. Desta forma, o melhor resultado obtido com a técnica equivale ao uso de 112 componentes utilizando-se a técnica PCA tradicional. Comparando-se os resultados obtidos com o mesmo número de coeficientes pela técnica de PCA (56%) e PCA Bidimensional (58%) percebe-se um desempenho um pouco superior do IMPCA.

PCA MODULAR

A técnica de PCA modular consiste em uma modificação do PCA tradicional visando tirar proveito de variações parciais das imagens das faces. Aplicando-se a técnica sobre a base de dados observou-se o seguinte resultado:

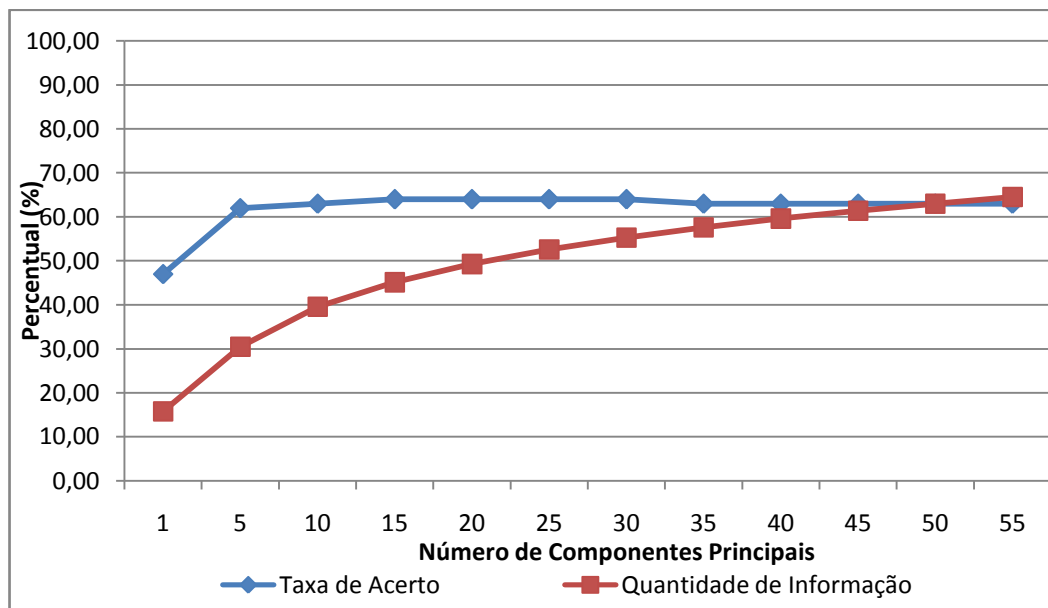


Fig. 4.8 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o MPCA sobre a base de faces UMIST utilizando cinco imagens para treinamento e cinco para teste.

A seguir é exibida uma tabela com o resumo dos resultados obtidos para a base de dados UMITS:

N°. de Componentes	PCA		IMPCA		PCA Modular	
	Acerto(%)	% de Informação	Acerto(%)	% de Informação	Acerto(%)	% de Informação
1	15,00	7,36	62,00	29,95	61,00	29,95
5	45,00	27,51	65,00	72,14	62,00	30,49
10	51,00	41,30	59,00	89,39	63,00	39,60
15	52,00	51,08	59,00	94,33	64,00	45,16
20	56,00	58,29	57,00	96,75	64,00	49,31
25	56,00	64,17	57,00	97,88	64,00	52,60
30	58,00	69,02	58,00	98,53	64,00	55,32
35	59,00	73,15	58,00	98,97	63,00	57,64
40	59,00	76,74	58,00	99,27	63,00	59,62
45	59,00	79,96	58,00	99,47	63,00	61,39
50	59,00	82,80	58,00	99,62	63,00	63,02
55	58,00	85,37	58,00	99,72	63,00	64,54

Tab. 4.1 – Resumo dos dados obtidos executando-se as técnicas sobre a base de dados UMIST. Os dados em destaque representam os melhores resultados da técnica.

Fig. 4.5 mostra algumas imagens utilizadas para treinamento e teste. Estas imagens foram escolhidas para testar as técnicas sob variações de posicionamento das faces para as quais elas não foram treinadas. Em função disso, os resultados obtidos pelas técnicas sobre o conjunto de testes foram muito baixos.

Nenhum dos resultados alcançados foi superior a 64%. Sendo o melhor desempenho obtido pelo PCA Bidimensional. O desempenho inferior do MPCA (60%) ocorreu como consequência do tipo de variação sofrida pelas imagens. Em variações de rotação da face toda a imagem sofre modificação, desta forma, algoritmos que exploram pouca ou nenhuma variação em partes da imagem decaem seu desempenho.

Em todos os testes realizados sobre esta base a técnica de PCA Bidimensional obteve um desempenho superior às demais. Além disso, o tempo necessário para processamento foi muito inferior ao necessário para processamento pelas demais técnicas.

Quando a quantidade de informação utilizada na representação das imagens com as técnicas os valores tendem a se manter constantes sem ganho de desempenho. O valor da taxa de reconhecimento se aproxima de 39% oscilando em um ponto percentual à medida que a quantidade aumenta.

4.2.2. Resultados para a base YALE

A utilização desta base visa testar as técnicas sob variações de configurações de iluminação e mudanças de expressões faciais (cansado, surpreso, sorrindo, triste, etc.) além do uso ou não de objetos, como óculos, sobre a face.

Para esta base foram utilizadas as onze imagens de cada indivíduo. Em um primeiro teste, cinco imagens para treinamento e as demais foram utilizadas para teste. Novamente os experimentos feitos sobre a base utilizaram as mesmas imagens de treinamento e teste para todas as técnicas. Os resultados, detalhados a seguir, foram obtidos utilizando apenas as quatro primeiras imagens para treino e as demais para teste. Esta configuração apresentou os melhores resultados obtidos com a base.

A seguir são apresentadas algumas das imagens utilizadas para treino e teste:



Fig. 4.9 – Exemplo de imagens da base Yale utilizada para treinamento e testes durante os experimentos.

PCA

Aplicando-se a técnica PCA sobre a base de dados obtivemos o seguinte resultado:

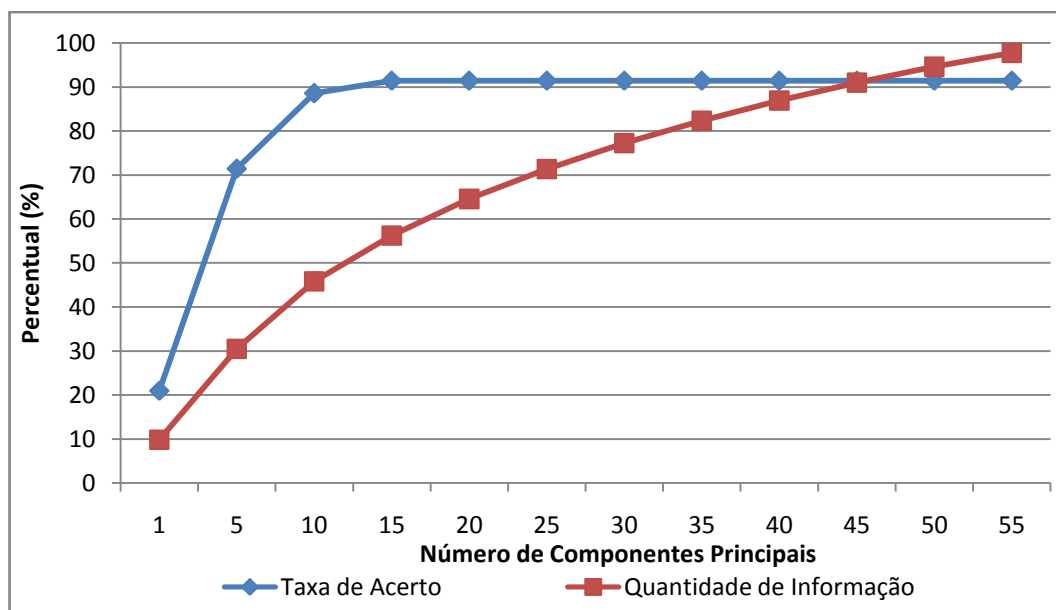


Fig. 4.10 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o PCA tradicional sobre a base de faces Yale utilizando cinco imagens para treinamento e cinco para teste.

É perceptível a estabilidade da taxa de acerto utilizando-se 20 ou mais componentes principais. A taxa de acerto se estabiliza em aproximadamente 91% utilizando pelo menos 56% da quantidade de informação da base.

PCA BIDIMENSIONAL (2DPCA ou IMPCA)

Aplicando-se a técnica de PCA Bidimensional obtivemos o seguinte resultado:

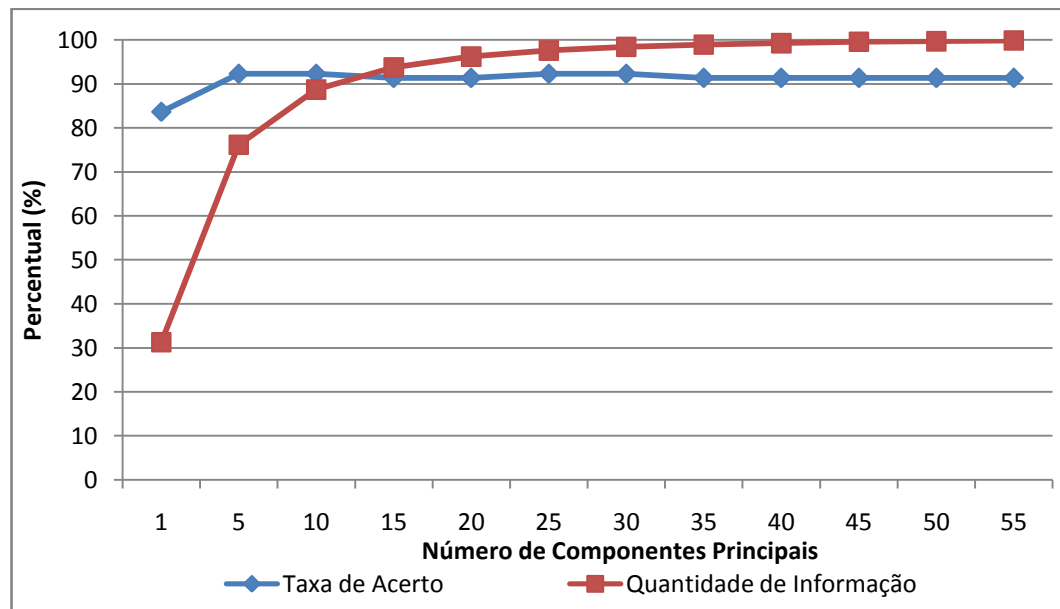


Fig. 4.11 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o IMPCA sobre a base de faces Yale utilizando cinco imagens para treinamento e cinco para teste.

Do resultado percebe-se claramente o elevado grau de significância que os primeiros componentes possuem. Cinco deles representam aproximadamente 77% da informação da imagem. Isso ocorre em função da maior representatividade dos dados pela matriz de covariância obtida durante o processamento das imagens além do seu tamanho reduzido.

Outro ponto em destaque no gráfico é a elevada taxa de acerto obtida com apenas um componente quando comparada com a técnica PCA. Novamente, isso ocorre em função de um componente do PCA Bidimensional ser composto por 112 coeficientes, conforme descrito na seção 3.2.

PCA MODULAR

A técnica de PCA modular foi desenvolvida procurando tirar proveito de regiões da face que tenham sofrido pouca ou nenhuma modificação em

decorrência de modificações na imagem. Por isso, presume-se que aplicando a técnica de PCA Modular sobre imagens com variações de iluminação e expressões faciais o ganho de desempenho seja mais elevado que o obtido quando o PCA tradicional é aplicado.

A seguir são apresentados os resultados obtidos com o uso da técnica sobre a base de dados Yale:

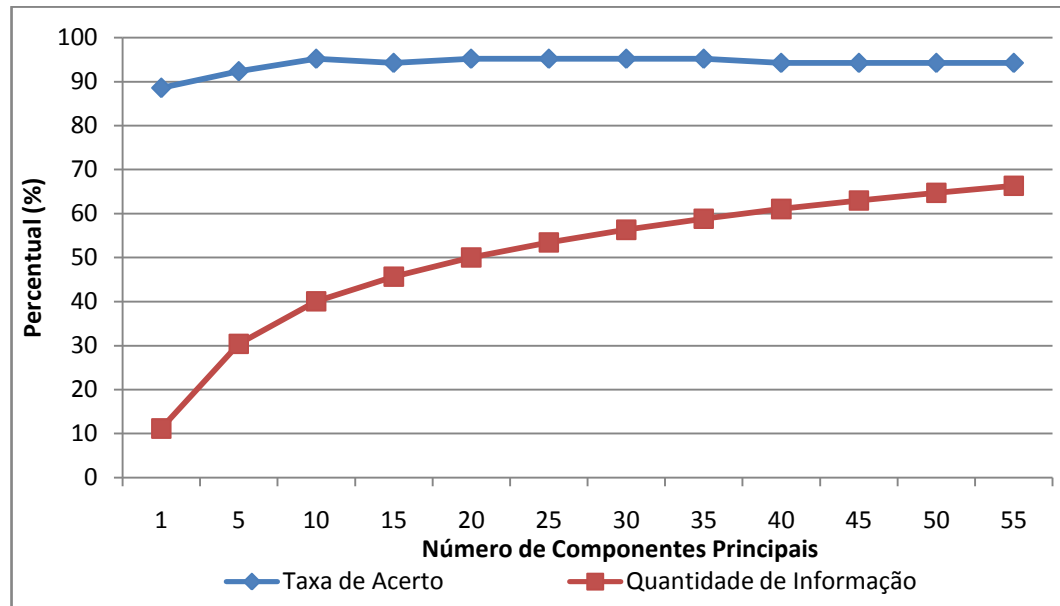


Fig. 4.12 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o MPCA sobre a base de faces Yale subdividindo-se a imagem em 4 partes.

O valor da taxa de acerto obtida com a técnica MPCA utilizando-se cerca de 100% da quantidade de informação mantém-se estável em aproximadamente 94% com variações de aproximadamente dois pontos percentuais. Desta forma o melhor resultado alcançado pela técnica foi obtido utilizando apenas dez componentes principais e aproximadamente 40% da quantidade de informação.

A seguir é exibida uma tabela com o resumo dos resultados obtidos para a base de dados YALE:

N°. de Componentes	PCA		IMPCA		PCA Modular	
	Acerto(%)	% de Informação	Acerto(%)	% de Informação	Acerto(%)	% de Informação
1	20,95	9,83	83,65	31,30	88,57	11,18
5	71,43	30,49	92,31	76,20	92,38	30,44
10	88,57	45,85	92,31	88,73	95,24	40,11
15	91,43	56,25	91,35	93,76	94,29	45,68
20	91,43	64,57	91,35	96,22	95,24	50,05
25	91,43	71,38	92,31	97,59	95,24	53,48
30	91,43	77,26	92,31	98,42	95,24	56,37

35	91,43	82,37	91,35	98,94	95,24	58,86
40	91,43	86,92	91,35	99,27	94,29	61,07
45	91,43	91,01	91,35	99,61	94,29	63,03
50	91,43	94,62	91,35	99,71	94,29	64,77
55	91,43	97,79	91,35	99,91	94,29	66,35

Tab. 4.2 – Resumo dos dados obtidos executando-se as técnicas sobre a base de faces Yale. Os dados em destaque representam os melhores resultados da técnica.

Dos resultados apresentados em **Tab. 4.2** pode-se perceber o melhor desempenho da técnica de PCA Modular. O melhor desempenho da técnica pode ser atribuído ao tipo de variação presente pelas imagens da base Yale. As variações presentes nas imagens da base, iluminação e expressões faciais, não afetam todas as partes da imagem preservando as informações contidas em algumas delas. A técnica PCA Modular explorar estas mudanças parciais nas imagens para elevar seu desempenho.

O IMPCA e o PCA consideram a informação global da imagem. Desta forma, pequenas variações, como sombras ou expressões faciais, que afetam algumas regiões da imagem acarretam em modificação em toda a representação da face e sua posterior classificação.

4.2.3. Resultados para a base ORL

A utilização da base ORL tem por finalidade analisar as técnicas com imagens que sofrem variações de expressões faciais, pequenas rotações, variações de escala e uso ou não de objetos sobre a face.

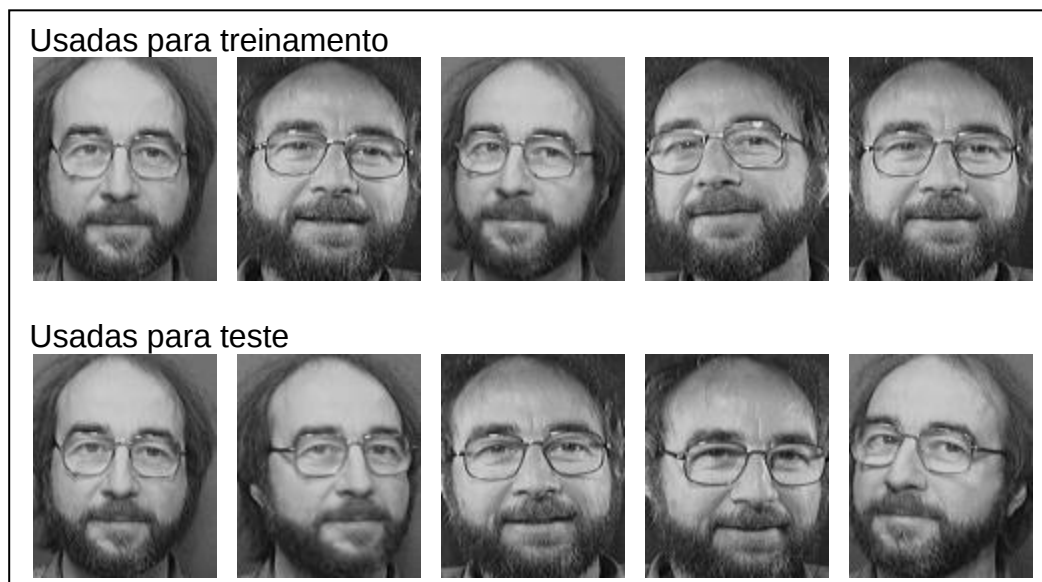


Fig. 4.13 – Exemplo de imagens da base ORL utilizada para treinamento e testes durante os experimentos.

Todas as imagens da base foram utilizadas nos testes. Foram feitos vários experimentos variando o número de imagens de treinamento e de teste, entretanto, os resultados detalhados a seguir foram obtidos utilizando-se cinco imagens para treinamento e as demais para teste. Todos os resultados apresentados nos experimentos foram obtidos utilizando o mesmo conjunto de imagens. Em **Fig. 4.13** são apresentadas algumas das imagens utilizadas para treinamento e teste.

PCA

Aplicando-se a técnica sobre a base de dados foram obtidos os seguintes resultados:

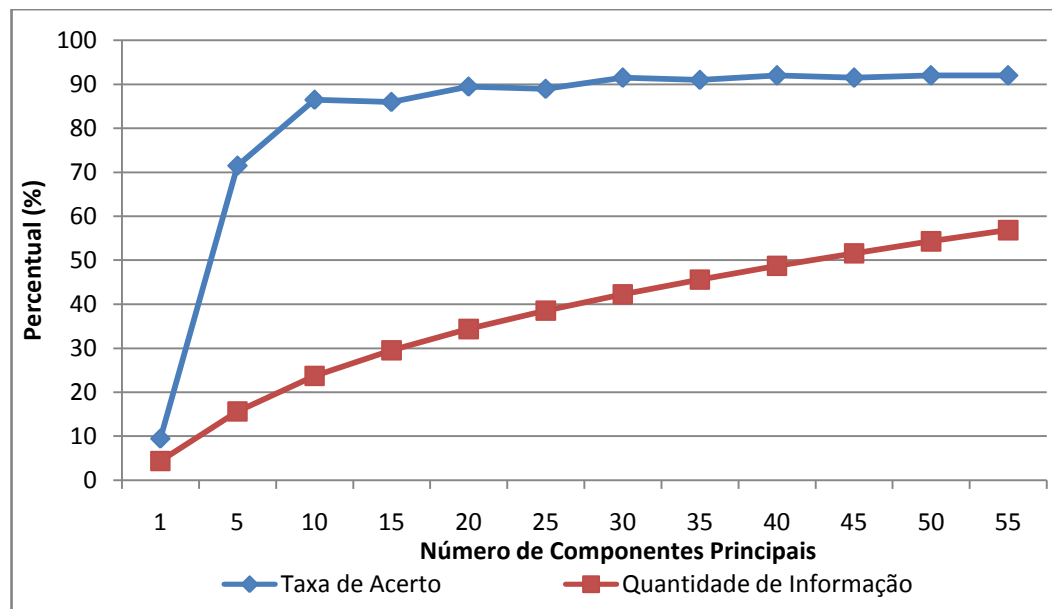


Fig. 4.14 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o PCA tradicional sobre a base de faces ORL utilizando cinco imagens para treinamento e cinco para teste.

É perceptível o elevado aumento na taxa de acerto obtido com a técnica sobre a base quando se modifica a quantidade de componentes de um para cinco imagens. É perceptível também a distribuição homogênea da quantidade de informação entre as primeiras componentes. Isso acarreta na necessidade de se utilizar muitos componentes para se representar e reconhecer a imagem adequadamente.

PCA BIDIMENSIONAL (2DPCA OU IMPCA)

Aplicando-se a técnica IMPCA sobre a base percebe-se uma maior concentração da quantidade de informação nos primeiros componentes da

imagem resultando em uma elevada taxa de reconhecimento ainda com poucos componentes. Além disso, com dez ou mais características (que no experimento resulta em 1120 ou mais coeficientes) a taxa de acerto mantém-se bastante estável.

Segue o gráfico mostrando os resultados obtidos nos experimentos.

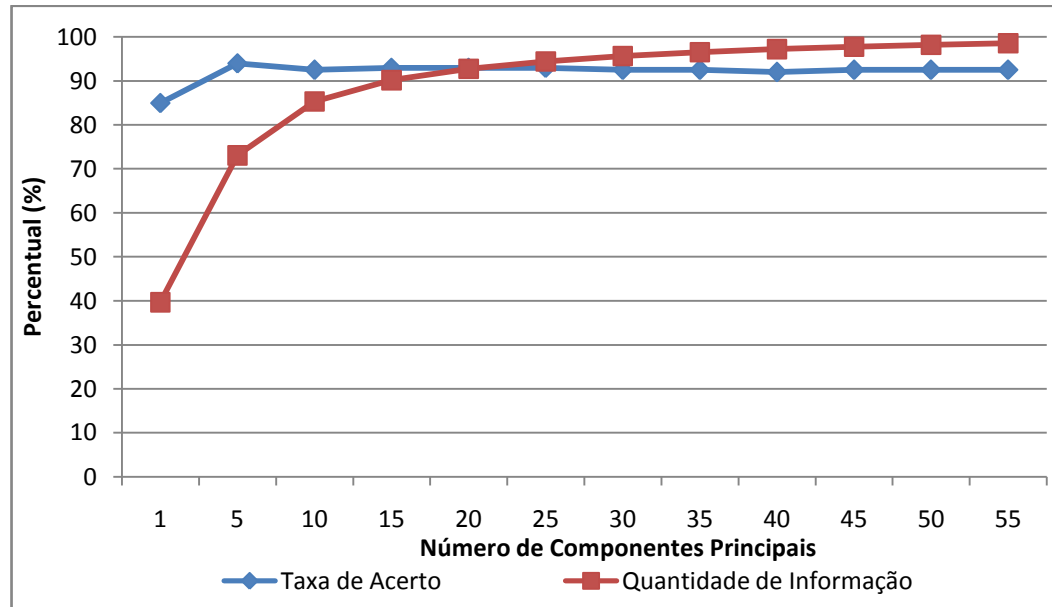


Fig. 4.15 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o IMPCA sobre a base de faces ORL utilizando cinco imagens para treinamento e cinco para teste.

PCA MODULAR

Como resultado da aplicação da técnica sobre a base de dados obteve-se:

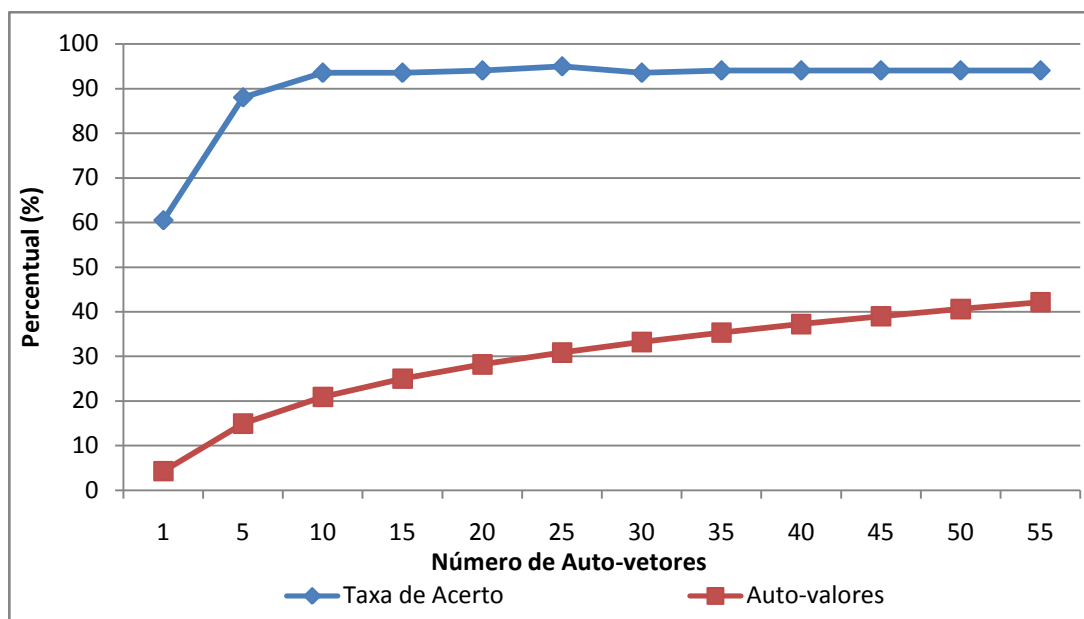


Fig. 4.16 – Apresentação da taxa de acerto e quantidade de informação em relação à quantidade de componentes principais aplicando-se o MPCA sobre a base de faces ORL subdividindo-se a imagem em 4 partes.

Os resultados com o MPCA mostram, a exemplo do que ocorreu com o PCA tradicional, uma distribuição homogênea entre os primeiros componentes utilizados na representação da imagem. Entretanto, a técnica MPCA explora melhor as informações que não sofreram modificação nas imagens. Isso resulta em um ganho considerável, mesmo com baixa quantidade de informação, quando comparado ao obtido com o PCA tradicional.

N°. de Componentes	PCA		IMPCA		PCA Modular	
	Acerto(%)	% de Informação	Acerto(%)	% de Informação	Acerto(%)	% de Informação
1	9,50	4,41	85,00	39,69	88,00	14,94
5	71,50	15,64	94,00	73,08	96,00	20,86
10	86,50	23,71	92,50	85,34	96,50	27,79
15	86,00	29,55	93,00	90,18	95,00	32,21
20	89,50	34,40	93,00	92,74	94,50	35,66
25	89,00	38,55	93,00	94,45	94,50	38,47
30	91,50	42,26	92,50	95,66	94,50	40,96
35	91,00	45,63	92,50	96,55	95,00	43,11
40	92,00	48,74	92,00	97,24	95,00	45,07
45	91,50	51,62	92,50	97,78	95,00	46,86
50	92,00	54,32	92,50	98,21	95,00	48,53
55	92,00	56,86	92,50	98,57	95,00	50,06

Tab. 4.3 – Resumo dos dados obtidos executando-se as técnicas sobre a base de faces ORL. Os dados em destaque representam os melhores resultados da técnica.

Os resultados deste experimento ficaram muito próximos para todas as técnicas, novamente, com a técnica de PCA Modular conseguiu o melhor resultado entre as técnicas testadas em função do tipo de variação presente nas imagens da base.

Representação das imagens

Embora as técnicas sejam todas baseadas em PCA, e possuam um núcleo comum a forma de extrair e representar as informações de uma imagem, elas se diferenciam bastante em outros aspectos do processo de reconhecimento. No tocante, a representação adotada por algumas técnicas estudadas explicam seu elevado desempenho. Principalmente com baixa dimensão as técnicas de MPCA e IMPCA alcançam resultados muito aquém dos obtidos fazendo uso do PCA tradicional. Isso ocorre devido à quantidade de coeficientes utilizados na representação das componentes da imagem.

A seguir é apresentada uma tabela com as quantidades de coeficientes (valores numéricos) utilizados para cada componente da imagem processada pela técnica.

PCA	IMPCA	PCA Modular
1	$1 \times d$	$1 \times n$

Tab. 4.4 – Número de coeficientes utilizados para representar uma característica da face.

Sabendo-se que n é a quantidade de imagens em que a face foi dividida e d é a quantidade de coeficientes que constituem os vetores resultantes da multiplicação das matrizes de imagens com os autovetores da matriz de covariância. Em relação à resolução da imagem d pode ser entendido como sendo a quantidade de pixels presentes em uma coluna da imagem.

Aplicando-se a regra para as imagens utilizadas no experimento, que foram normalizadas para a resolução de 92 x 112 pixels um componente da técnica IMPCA é composto por 112 coeficientes e um componente da técnica MPCA cujas imagens foram divididas em 16 partes é formada por 16 coeficientes.

Outro fator importante a ser ponderado diz respeito ao tempo de processamento utilizado por elas. A seguir é relacionado o tempo médio, em segundos (s), gasto no processamento das imagens e seu desvio padrão.

PCA	IMPCA	PCA Modular
142 (99) s	20 (14) s	120 (75) s

Tab. 4.4 – Tempo médio utilizado pelas técnicas para processamento das imagens e o desvio padrão entre parênteses.

Pelos resultados apresentados em Tab. 4.3 e Tab. 4.4 percebe-se que a técnica IMPCA é aproximadamente dez vezes mais rápida que as demais.

Entretanto, requer bem mais memória para representar as faces quando comparado ao PCA tradicional e também que a necessária pela técnica MPCA com $N < D$. Sabendo-se que D é a quantidade de pixels em uma coluna da imagem e N a quantidade de imagens na qual a face foi dividida.

4.2.4. Classificador

Durante todos os experimentos foi utilizado um único classificador. Ele é baseado na distância euclidiana entre os elementos e faz uso da técnica k -nn com k igual a um, ou seja, 1- nn . Desta forma, a classe que possuir o elemento mais próximo do que está sendo classificado será atribuída ao protótipo de teste.

Há duas peculiaridades no cálculo das distâncias entre faces que merecem destaque:

- A distância entre as representações obtidas com o uso do MPCA é calculada a partir da soma das distâncias das sub-imagens correspondentes de cada face. Ou seja, para duas faces que tenham sido divididas em dezesseis partes a distância entre elas é calculada como sendo a soma entre as distâncias das suas dezesseis sub-imagens correspondentes;
- O cálculo da distância para as representações resultantes da técnica IMPCA também é diferenciado. A distância é calculada como sendo a soma das distâncias dos itens correspondentes no vetor de características de cada imagem. Nos experimentos cada item é composto por um vetor de dimensão 112.

5. Considerações Finais

Desde a publicação dos trabalhos de Pentland [7] diversas outras abordagens foram desenvolvidas para lidar com dois grandes problemas presentes em sistemas reais de reconhecimento de faces como as variações na iluminação e de posicionamento da face nas imagens.

As duas novas técnicas analisadas neste trabalho (MPCA e IMPCA) foram comparadas com a técnica de eigenfaces (PCA) e apresentaram resultados superiores em todos os experimentos.

A técnica IMPCA, que propõe nova forma de representação para as imagens das faces, tem pelo menos três vantagens quando comparada à técnica de PCA tradicional. Primeiro, ela é baseada na representação matricial das imagens não sendo necessária a etapa de transformação das matrizes em vetores. Segundo, a técnica apresentou melhores resultados que o PCA em todos os experimentos realizados. E por último, a técnica IMPCA é computacionalmente mais eficiente que o PCA tradicional. Ela diminui significativamente o tempo necessário para extração de características das imagens. Contudo a técnica não é eficiente do ponto de vista de armazenamento necessitando de muitos coeficientes para representação de uma única característica da imagem.

A melhoria da taxa de reconhecimento do IMPCA pode ser explicada pela sua melhor aplicabilidade a problemas com poucos protótipos visto que o tamanho da matriz de covariância gerada é menor e os autovetores utilizados na representação das faces são estatisticamente dependentes dela.

Os resultados apresentados pela técnica MPCA foram muito superiores aos alcançados pela técnica PCA tradicional quando aplicados sobre bases com grande variação de pose e iluminação. Desta forma a técnica MPCA pode ser utilizada como alternativa para o PCA tradicional especialmente quando aplicada em sistemas de reconhecimento sujeitos a grandes variações de iluminação e expressões faciais.

O melhor desempenho do MPCA sobre o IMPCA nas bases de dados Yale e ORL é explicado em função do tratamento modular dados às faces pela técnica MPCA e pelos tipos de variações às quais as faces das bases foram submetidas. Embora o IMPCA seja um método robusto e represente melhor as informações das bases (Tab. 4.2 e Tab. 4.3) ele é bastante afetado pelo tipo de variação presente nas imagens o que resulta em problemas em sua representação e classificação.

As técnicas PCA e MPCA podem ser combinadas gerando assim uma técnica bidimensional modular com a finalidade de aproveitar o melhor das duas técnicas. Esta técnica dividiria a imagem em partes menores pela aplicação da técnica MPCA e representaria cada uma das partes mais fielmente fazendo uso dos procedimentos presentes na técnica (IMPCA).

Outro ponto importante a ser trabalhado, e que não foi esclarecido pelos autores da técnica MPCA, diz respeito ao uso de apenas uma média das imagens e de apenas uma matriz de covariância independentemente do número de partes que a imagem tenha sido dividida. Embora esta característica sirva para relacionar as partes da imagem com o todo ela também pode acarretar no aumento da dispersão intra-classe da imagem. A utilização de mais de uma média e mais de uma matriz de covariância será alvo de estudo na continuação deste trabalho.

Estudos de novos métodos de reconhecimento baseados em PCA juntamente com técnicas de transformação de imagens dentre elas *Fourier* e *Wavelets* vem sendo utilizadas com bastante êxito no reconhecimento de imagens de faces. Estas técnicas constituem outra vertente dos trabalhos futuros a serem desenvolvidos nesse projeto.

Em termos de desempenho, o estado atual das tecnologias de reconhecimento de faces está muito aquém da capacidade humana especialmente em imagens com variações na iluminação e expressões faciais. Além disso, há necessidade de tratamento para outras variações ainda mais complexas como as modificações na face causadas pelo tempo. Entretanto, o aumento recente do número de pesquisas na área resultarão, mesmo que em longo prazo, em melhorias nas tecnologias de reconhecimento de modo a tornar nossas vidas mais simples e convenientes.

6. Referências

- [1] GOTTMUKKAL, R. e ASARI, V. K. *An improved face recognition technique based on modular PCA approach*. Pattern Recognition Letters 15 (2004) pp. 429-436, 2004.
- [2] WANG, L., WANG, X., ZHANG, X. e FENG, J, *The equivalence of two-dimensional PCA to line-based PCA*. Pattern Recognition Letters 16 (2005) pp. 57-60, 2005.
- [3] YANG, J., ZHANG, D., FRANGI, A.F. e YANG, J. *Two-dimensional PCA: A new approach to appearance-Based Face Representation and Recognition*. IEEE transactions on pattern analysis and machine intelligence, vol. 26, No. 1, pp. 131 - 137, 2004.
- [4] ROCHA, L.M., *Singular Value Decomposition and principal component analysis in A Pratical Approach in Microarray Data Analysis*. Portland State University – Ph.D. Program, Kluwer: Norwel, pp. 91-109, MA, 2003.
- [5] BARLETT, M.S, MOVELLAN J.R. e BEVERIDGE, J.R., *Face recognition by Independent Component Analysis*. IEEE Trans. Neural Networks, vol. 13, no. 6, pp. 1450-1464, 2002.
- [6] SHLENS, J. *A Tutorial on Principal Component Analysis*. Universidade da Califórnia, La Jolla 2005.
- [7] M. Turk, A. Pentland. *Eigenfaces for Recognition*. Journal of Cognitive Neuroscience. Vol. 3, No. 1. 71-86, 1991.
- [8] MARIM, L.O. e Barreto J.M., *Reconhecimento de Faces*, Universidade Federal e Santa Catarina – Tese de Mestrado, pp. 57-66, 2003
- [9] BIANCHI, M. F., *Extração de características de imagens de faces humanas através de Wavelets, PCA e IMPCA*, Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, pp. 37-63, 2006
- [10] LI, S. Z., Lu, J., *Face recognition using the nearest feature line method*, IEEE Trans. Neural Networks, , pp. 439-443, 1999
- [11] KIM, H. H. M., *Survey paper: Face Detection and Face Recognition*, 2004
- [12] BELHUMEUR, v., *eigenfaces vs. fisherfaces: recognition using class specific linear projections*, IEEE Trans. on PAMI, V.19,pp. 711-720, 1997

[13] BASTOS, G., *HMM's para identificação de regiões promotoras de genes de reparação*, Trabalho de Graduação, Centro de Informática – Universidade Federal de Pernambuco, Pernambuco, pp. 10-16, 2003

[14] GRAHAM, D. H., ALLINSON N.M., *Face Recognition: From theory to applications (UMIST Face Database)*, NATO ASI Series F, Computer and Systems Science, Vol. 163, pp. 446-456, 1998

[15] GEORGHIADES, A.S. and BELHUMEUR, P.N. and KRIEGMAN, D.J., *From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose*, IEEE Trans. Pattern Anal. Mach. Intelligence, Vol. 23, pp. 643-660, 2001

[16] *The ORL Database of Faces*, Cambridge University - Engineering Department, 1994

Anexo I – Tabelas de Resultados

A seguir serão apresentadas as tabelas com os resultados obtidos nos experimentos. Serão apresentados os resultados obtidos nas três bases (Yale, ORL e UMIST, nesta ordem) variando-se a quantidade de imagens para treinamento entre 1 e 5. Além disso, para o caso específico da técnica MPCA serão apresentados os resultados obtidos através da divisão da imagem para diferentes valores de N .

Yale (PCA)

N°. de Componentes	PCA (4)		PCA (3)		PCA (2)		PCA (1)	
	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação
5	71,43	30,49	74,17	31,86	71,11	37,61	44,67	54,30
10	88,57	45,85	86,67	49,11	86,67	58,02	64,67	83,83
15	91,43	56,25	86,67	61,49	86,67	72,90	-	-
20	91,43	64,57	89,17	71,19	86,67	84,58	-	-
25	91,43	71,38	89,17	79,35	87,41	94,02	-	-
30	91,43	77,26	89,17	86,12	-	-	-	-
35	91,43	82,37	89,17	91,93	-	-	-	-
40	91,43	86,92	89,17	96,84	-	-	-	-
45	91,43	91,01	-	-	-	-	-	-
50	91,43	94,62	-	-	-	-	-	-
55	91,43	97,79	-	-	-	-	-	-
60	-	-	-	-	-	-	-	-
65	-	-	-	-	-	-	-	-
70	-	-	-	-	-	-	-	-
75	-	-	-	-	-	-	-	-
80	-	-	-	-	-	-	-	-
85	-	-	-	-	-	-	-	-
90	-	-	-	-	-	-	-	-
95	-	-	-	-	-	-	-	-
100	-	-	-	-	-	-	-	-

Tab. AI.1 – Resultados obtidos pela aplicação da técnica PCA sobre a base Yale.

Yale (MPCA)

Nº. de Componentes	MPCA (N=4)		MPCA (N=16)		MPCA (N=64)	
	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação
5	86,67	27,59	92,38	30,44	93,33	48,07
10	88,57	36,83	95,24	40,11	94,29	59,42
15	92,38	43,16	94,29	45,68	94,29	65,96
20	92,38	48,02	95,24	50,05	94,29	70,67
25	92,38	51,97	95,24	53,48	94,29	74,26
30	92,38	55,41	95,24	56,37	94,29	77,21
35	93,33	58,45	95,24	58,86	94,29	79,67
40	94,29	61,19	94,29	61,07	94,29	81,82
45	94,29	63,65	94,29	63,03	94,29	83,66
50	94,29	65,87	94,29	64,77	94,29	85,30
55	94,29	67,91	94,29	66,35	94,29	86,73
60	94,29	69,81	94,29	67,80	94,29	88,04
65	94,29	71,59	93,33	69,15	94,29	89,22
70	95,24	73,26	93,33	70,39	94,29	90,30
75	94,29	74,83	93,33	71,55	94,29	91,29
80	94,29	76,31	93,33	72,65	94,29	92,22
85	94,29	77,70	93,33	73,69	94,29	93,07
90	94,29	79,01	93,33	74,65	94,29	93,87
95	95,24	80,26	93,33	75,55	94,29	94,61
100	95,24	81,46	93,33	76,40	94,29	95,30

Tab. AI.3 – Resultados obtidos pela aplicação da técnica IMPCA sobre a base Yale utilizando-se quatro elementos para treinamento e variando-se os valores de N (N = 4, 16, 64).

Yale (IMPCA)

Nº. de Comp.	IMPCA (4)		IMPCA (3)		IMPCA (2)		IMPCA (1)	
	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação
5	92,38	76,10	87,50	71,82	90,37	71,92	65,33	73,68
10	92,38	88,70	88,33	86,35	88,89	86,57	70,00	88,13
15	91,43	93,73	88,33	92,08	88,89	92,25	69,33	93,44
20	91,43	96,20	89,17	95,16	89,63	95,25	70,00	96,11
25	91,43	97,59	89,17	96,88	88,89	97,01	69,33	97,62
30	92,38	98,42	89,17	97,96	88,89	98,04	69,33	98,47
35	91,43	98,94	88,33	98,64	88,15	98,68	69,33	98,99
40	91,43	99,28	88,33	99,07	87,41	99,10	69,33	99,32
45	91,43	99,50	88,33	99,35	87,41	99,37	69,33	99,53
50	91,43	99,65	88,33	99,55	87,41	99,56	69,33	99,68
55	91,43	99,75	88,33	99,68	87,41	99,69	69,33	99,78
60	91,43	99,83	88,33	99,78	87,41	99,78	69,33	99,85
65	91,43	99,88	88,33	99,85	87,41	99,85	69,33	99,90
70	91,43	99,92	88,33	99,90	87,41	99,90	69,33	99,93
75	91,43	99,95	88,33	99,94	87,41	99,94	69,33	99,96
80	91,43	99,98	88,33	99,97	87,41	99,97	69,33	99,98
85	91,43	99,99	88,33	99,99	87,41	99,99	69,33	99,99
90	91,43	100,00	88,33	100,00	87,41	100,00	69,33	100,00
95	-	-	-	-	-	-	-	-
100	-	-	-	-	-	-	-	-

Tab. AI.2 – Resultados obtidos pela aplicação da técnica IMPCA sobre a base Yale utilizando-se respectivamente 4, 3, 2 e 1 elementos na base de treinamento.

ORL (PCA)

N°. de Comp.	PCA (5)		PCA (4)		PCA (3)		PCA (2)		PCA (1)	
	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação
5	71,50	15,64	67,92	16,96	67,86	18,83	60,31	21,98	51,39	29,75
10	86,50	23,71	83,33	25,80	79,64	28,50	74,06	33,20	64,44	46,04
15	86,00	29,55	82,08	32,23	80,71	35,79	75,63	41,94	64,44	58,65
20	89,50	34,40	84,58	37,56	83,21	41,88	76,56	49,44	66,39	69,55
25	89,00	38,55	85,83	42,19	83,21	47,13	76,88	56,02	67,78	79,09
30	91,50	42,26	87,50	46,29	84,29	51,85	76,56	61,85	69,17	87,57
35	91,00	45,63	87,08	50,03	84,29	56,14	77,19	67,14	69,44	94,88
40	92,00	48,74	88,33	53,45	85,71	60,11	77,50	71,97	-	-
45	91,50	51,62	88,33	56,61	84,29	63,80	77,19	76,50	-	-
50	92,00	54,32	87,50	59,60	85,00	67,24	76,88	80,73	-	-
55	92,00	56,86	88,33	62,39	84,29	70,49	77,19	84,67	-	-
60	91,00	59,28	88,33	65,06	83,93	73,57	77,81	88,39	-	-
65	91,50	61,57	88,33	67,61	84,29	76,51	77,50	91,84	-	-
70	92,00	63,76	87,08	70,03	84,29	79,29	77,81	95,05	-	-
75	92,00	65,86	86,67	72,35	85,00	81,97	78,75	97,99	-	-
80	91,50	67,88	87,08	74,56	85,71	84,49	78,75	100,00	-	-
85	91,50	69,82	86,67	76,69	85,36	86,89	-	-	-	-
90	91,50	71,69	86,25	78,73	85,71	89,17	-	-	-	-
95	91,00	73,48	86,25	80,70	85,71	91,36	-	-	-	-
100	91,50	75,23	87,08	82,59	85,71	93,43	-	-	-	-

Tab. AI.4 – Resultados obtidos pela aplicação da técnica PCA sobre a base Yale.

ORL (MPCA)

N°. de Comp.	MPCA (N=4)		MPCA (N=8)		MPCA (N=64)	
	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação
5	88,00	14,94	96,00	20,86	94,00	36,52
10	93,50	20,92	96,50	27,79	94,00	46,11
15	93,50	25,01	95,00	32,21	94,00	52,41
20	94,00	28,20	94,50	35,66	94,00	57,20
25	95,00	30,86	94,50	38,47	94,00	61,18
30	93,50	33,22	94,50	40,96	93,50	64,64
35	94,00	35,34	95,00	43,11	93,50	67,64
40	94,00	37,24	95,00	45,07	93,50	70,38
45	94,00	38,98	95,00	46,86	93,50	72,82
50	94,00	40,60	95,00	48,53	93,50	75,04
55	94,00	42,12	95,00	50,06	93,50	77,14
60	94,00	43,56	95,00	51,50	93,50	79,07
65	94,00	44,90	95,00	52,85	93,00	80,87
70	93,50	46,16	95,00	54,13	93,00	82,59
75	93,50	47,35	95,00	55,33	93,00	84,20
80	93,50	48,50	94,00	56,48	93,00	85,71
85	93,50	49,60	94,00	57,59	93,00	87,13
90	93,50	50,64	94,00	58,64	92,50	88,47
95	93,50	51,65	94,00	59,66	92,50	89,72
100	94,00	52,63	94,00	60,63	92,50	90,91

Tab. A1.6 – Resultados obtidos pela aplicação da técnica IMPCA sobre a base ORL utilizando-se cinco elementos para treinamento e variando-se os valores de N (N = 4, 16, 64).

ORL (IMPCA)

N°. de Comp.	IMPCA (5)		IMPCA (4)		IMPCA (3)		IMPCA (2)		IMPCA (1)	
	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação
5	94,00	73,08	90,42	72,91	89,29	72,73	83,75	72,76	72,50	73,23
10	92,50	85,34	89,17	85,22	87,86	85,08	81,25	85,03	71,39	85,46
15	93,00	90,18	90,42	90,08	87,86	89,98	79,69	89,93	71,67	90,33
20	93,00	92,74	90,42	92,67	88,21	92,57	78,75	92,54	71,11	92,91
25	93,00	94,45	89,58	94,39	87,50	94,30	79,38	94,29	71,39	94,61
30	92,50	95,66	89,58	95,62	87,50	95,54	78,75	95,53	71,11	95,80
35	92,50	96,55	89,17	96,52	87,14	96,46	78,75	96,44	71,11	96,66
40	92,00	97,24	89,17	97,21	87,50	97,16	78,75	97,15	71,11	97,35
45	92,50	97,78	88,75	97,76	86,79	97,72	78,75	97,72	71,11	97,90
50	92,50	98,21	88,75	98,20	86,79	98,16	78,75	98,16	71,11	98,32
55	92,50	98,57	88,75	98,56	86,43	98,53	78,75	98,53	71,11	98,66
60	92,50	98,87	88,75	98,86	86,43	98,84	78,75	98,84	70,83	98,96
65	92,50	99,14	88,75	99,14	86,43	99,12	78,75	99,12	70,83	99,21
70	92,50	99,37	88,75	99,36	86,43	99,35	78,75	99,35	70,83	99,43
75	92,50	99,56	88,33	99,56	86,43	99,55	78,75	99,55	70,83	99,60
80	92,50	99,71	88,33	99,72	86,43	99,71	78,75	99,71	70,83	99,75
85	92,50	99,85	88,33	99,85	86,43	99,85	78,75	99,85	70,83	99,87
90	92,50	99,96	88,33	99,96	86,43	99,96	78,75	99,96	70,83	99,97
95	-	-	-	-	-	-	-	-	-	-
100	-	-	-	-	-	-	-	-	-	-

Tab. AI.5 – Resultados obtidos pela aplicação da técnica IMPCA sobre a base ORL utilizando-se respectivamente 5, 4, 3, 2 e 1 elementos na base de treinamento.

UMIST (PCA)

Nº. de Comp.	PCA (5)		PCA (4)		PCA (3)		PCA (2)		PCA (1)	
	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação
5	45,00	27,51	57,50	30,03	58,57	33,25	56,25	38,56	60,56	47,77
10	51,00	41,30	57,50	45,03	59,29	49,96	61,88	57,80	62,22	71,75
15	52,00	51,08	58,33	55,49	60,00	61,56	61,88	71,44	62,22	89,17
20	56,00	58,29	57,50	63,27	60,71	70,46	61,88	82,17	-	-
25	56,00	64,17	56,67	69,59	59,29	77,28	62,50	89,16	-	-
30	58,00	69,02	57,50	74,65	60,00	82,65	62,50	94,27	-	-
35	59,00	73,15	58,33	78,90	60,00	86,93	62,50	98,01	-	-
40	59,00	76,74	60,00	82,58	60,00	90,58	62,50	100,00	-	-
45	59,00	79,96	59,17	85,87	60,00	93,73	-	-	-	-
50	59,00	82,80	60,00	88,77	60,00	96,52	-	-	-	-
55	58,00	85,37	60,00	91,29	60,00	98,70	-	-	-	-
60	59,00	87,74	60,00	93,57	60,00	100,00	-	-	-	-
65	58,00	89,88	60,00	95,66	-	-	-	-	-	-
70	58,00	91,80	60,00	97,53	-	-	-	-	-	-
75	57,00	93,57	60,00	99,04	-	-	-	-	-	-
80	58,00	95,20	60,00	100,00	-	-	-	-	-	-
85	58,00	96,70	-	-	-	-	-	-	-	-
90	58,00	98,08	-	-	-	-	-	-	-	-
95	58,00	99,24	-	-	-	-	-	-	-	-
100	58,00	100,00	-	-	-	-	-	-	-	-

Tab. A1.7 – Resultados obtidos pela aplicação da técnica PCA sobre a base UMIST.

UMIST (MPCA)

Nº. de Comp.	MPCA (N=4)		MPCA (N=8)		MPCA (N=64)	
	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação
5	59,00	24,10	62,00	30,49	64,00	48,48
10	59,00	33,49	63,00	39,60	64,00	58,82
15	60,00	39,68	64,00	45,16	64,00	65,14
20	61,00	44,25	64,00	49,31	64,00	69,67
25	59,00	47,99	64,00	52,60	64,00	73,31
30	58,00	51,19	64,00	55,32	64,00	76,21
35	59,00	54,03	63,00	57,64	64,00	78,67
40	59,00	56,48	63,00	59,62	64,00	80,80
45	58,00	58,68	63,00	61,39	64,00	82,64
50	58,00	60,67	63,00	63,02	64,00	84,27
55	58,00	62,48	63,00	64,54	64,00	85,73
60	58,00	64,13	63,00	65,94	64,00	87,07
65	58,00	65,66	63,00	67,24	64,00	88,32
70	58,00	67,09	63,00	68,45	64,00	89,47
75	58,00	68,44	63,00	69,59	64,00	90,54
80	58,00	69,71	63,00	70,66	64,00	91,54
85	58,00	70,92	63,00	71,66	64,00	92,46
90	58,00	72,06	63,00	72,62	64,00	93,32
95	58,00	73,15	63,00	73,52	64,00	94,12
100	58,00	74,18	63,00	74,38	64,00	94,86

Tab. A1.6 – Resultados obtidos pela aplicação da técnica IMPCA sobre a base UMIST utilizando-se quatro elementos para treinamento e variando-se os valores de N (N = 4, 16, 64).

UMIST (IMPCA)

N°. de Comp.	IMPCA (5)		IMPCA (4)		IMPCA (3)		IMPCA (2)		IMPCA (1)	
	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação	Acerto (%)	% de Informação
5	65,00	72,14	65,00	72,45	66,43	72,59	67,50	73,17	66,67	73,44
10	59,00	89,39	63,33	89,24	59,29	89,16	62,50	89,25	63,89	89,31
15	59,00	94,33	60,83	94,22	60,71	94,13	63,13	94,15	64,44	94,25
20	57,00	96,75	60,00	96,67	60,71	96,60	62,50	96,59	63,89	96,68
25	57,00	97,88	60,00	97,83	60,71	97,80	62,50	97,81	64,44	97,91
30	58,00	98,53	60,00	98,50	60,00	98,48	62,50	98,49	64,44	98,58
35	58,00	98,97	60,00	98,95	60,00	98,94	62,50	98,95	64,44	99,02
40	58,00	99,27	60,00	99,26	60,00	99,24	62,50	99,25	64,44	99,30
45	58,00	99,47	60,00	99,46	60,00	99,45	62,50	99,46	64,44	99,50
50	58,00	99,62	60,00	99,61	60,00	99,60	62,50	99,61	64,44	99,64
55	58,00	99,72	60,00	99,72	60,00	99,72	62,50	99,72	64,44	99,74
60	58,00	99,80	60,00	99,79	60,00	99,79	62,50	99,80	64,44	99,81
65	58,00	99,85	60,00	99,85	60,00	99,85	62,50	99,85	64,44	99,87
70	58,00	99,90	60,00	99,90	60,00	99,90	62,50	99,90	64,44	99,91
75	58,00	99,94	60,00	99,94	60,00	99,93	62,50	99,93	64,44	99,94
80	58,00	99,96	60,00	99,96	60,00	99,96	62,50	99,96	64,44	99,97
85	58,00	99,98	60,00	99,98	60,00	99,98	62,50	99,98	64,44	99,98
90	58,00	100,00	60,00	100,00	60,00	100,00	62,50	100,00	64,44	100,00
95	-	-	-	-	-	-	-	-	-	-
100	-	-	-	-	-	-	-	-	-	-

Tab. A1.8 – Resultados obtidos pela aplicação da técnica IMPCA sobre a base UMIST utilizando-se respectivamente 5, 4, 3, 2 e 1 elementos na base de treinamento.

Anexo II – Ferramenta Desenvolvida (hE)

Neste trabalho foram analisadas três técnicas de reconhecimento de faces baseadas em PCA. Para cada uma delas foi codificado um algoritmo de extração de características e reconhecimento de faces. Visando facilitar o processo de análise destas técnicas foi desenvolvida uma ferramenta gráfica para seleção dos bancos de faces e seleção das imagens a serem utilizadas como base de conhecimento e teste de cada uma das técnicas.

O aplicativo é composto de três interfaces. Uma delas (**Fig. AII.1**) destina-se à seleção dos protótipos a serem utilizados no experimento e as demais (**Fig. AII.2** e **Fig. AII.3**) para visualização dos resultados obtidos.

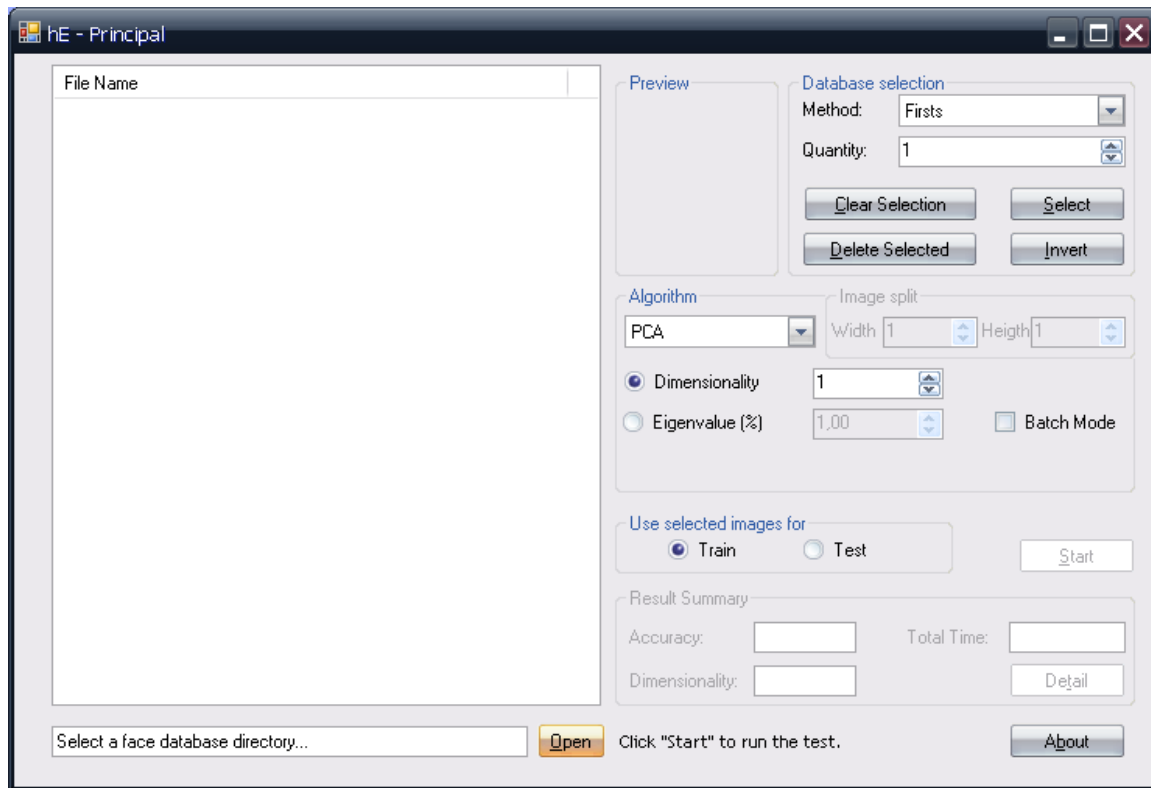
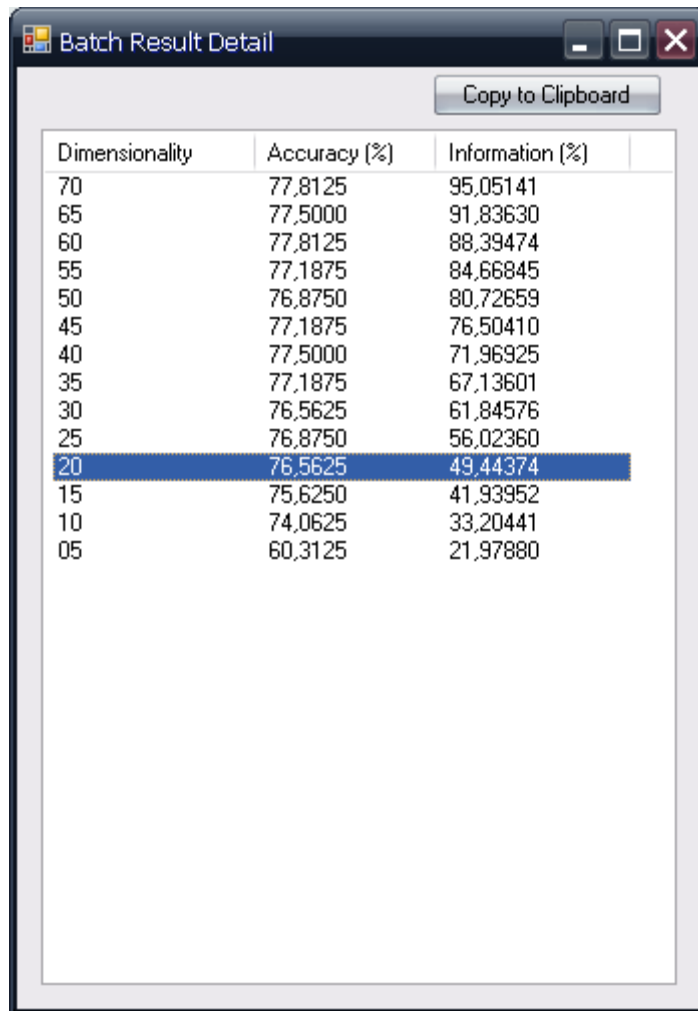


Fig. AII.1 – Tela inicial do aplicativo onde podem ser selecionadas as bases e os protótipos a serem utilizados no experimento.

Na interface inicial (**Fig. AII.1**) podem ser selecionados os protótipos que constituirão a base de conhecimento e os que serão utilizados para teste. Há opções para excluir protótipos que não serão utilizados no experimento. A técnica a ser utilizada assim como sua parametrização também pode ser escolhida nesta interface.

Há ainda a possibilidade de realização de experimentos em modo *batch* onde se escolhe o número de componentes principais inicial, o número máximo de componentes (ou o percentual de informação a ser preservado) e o

incremento a ser feito em cada iteração. Por exemplo, escolhendo-se 5 componentes principais inicial, 70 como número máximo de componentes e um incremento de 5 componentes por iteração o seguinte resultado pode ser obtido:



Dimensionality	Accuracy (%)	Information (%)
70	77,8125	95,05141
65	77,5000	91,83630
60	77,8125	88,39474
55	77,1875	84,66845
50	76,8750	80,72659
45	77,1875	76,50410
40	77,5000	71,96925
35	77,1875	67,13601
30	76,5625	61,84576
25	76,8750	56,02360
20	76,5625	49,44374
15	75,6250	41,93952
10	74,0625	33,20441
05	60,3125	21,97880

Fig. AII.2 – Interface para visualização dos resultados em modo *batch*.

A interface (**Fig. AII.2**) é utilizada para visualização dos resultados dos experimentos em modo *batch* para cada teste realizado é indicado o número de componentes principais utilizados, sua taxa de acerto e o percentual de informação utilizado no teste.

A interface (**Fig. AII.3**) exibe os detalhes do experimento. Exibe a matriz de confusão do experimento assim como as imagens de uma mesma classe que foram classificadas erroneamente ao se selecionar a linha correspondente da matriz de confusão.

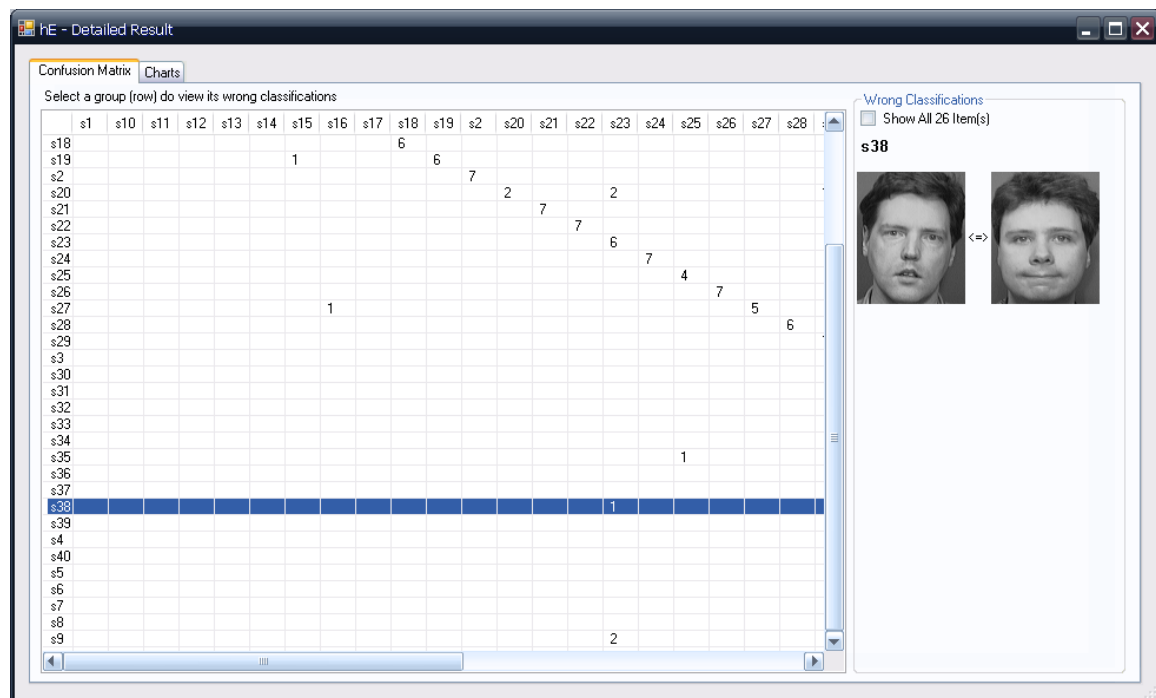


Fig. AII.3 – Interface para visualização da matriz de confusão do experimento e as imagens classificadas erroneamente.