

RECONHECIMENTO DE FACES USANDO ANÁLISE DE COMPONENTES PRINCIPAIS E MORFOLOGIA MATEMÁTICA

Trabalho de Conclusão de Curso

Engenharia da Computação

Gilliano Ginno Silva de Menezes
Orientador: Wellington Pinheiro dos Santos
Co-orientador: George Darminton Cavalcanti, UFPE



UNIVERSIDADE
DE PERNAMBUCO

**GILLIANO GINNO SILVA DE
MENEZES**

**RECONHECIMENTO DE FACES
USANDO ANÁLISE DE
COMPONENTES PRINCIPAIS E
MORFOLOGIA MATEMÁTICA**

Monografia apresentada como requisito parcial para obtenção do diploma de Bacharel em Engenharia da Computação pela Escola Politécnica de Pernambuco – Universidade de Pernambuco.

Recife, novembro 2009.

A minha família.

Agradecimentos

Agradeço a Deus pela oportunidade em sua misericórdia para que eu pudesse realizar mais esse projeto de vida. Também sou bastante grato aos meus pais, Gildenor e Verônica, pelo constante apoio, suporte e carinho necessários para a realização deste sonho. Também não poderei esquecer jamais do companheirismo prestado pelos meus três irmãos ao longo dessa jornada.

Agradeço especialmente à Rafaella, que com muito amor, me ajudou a passar por mais esta importante etapa de minha vida. Que venham outras, sempre ao seu lado. Agradeço também a minha mais nova família, que apesar de não unirem-se a mim por laços sanguíneos, uniram-se por um laço ainda mais inquebrantável, o da fé.

Agradeço também aos meus avôs, sogros (Edivaldo e Sandra) e aos meus cunhados (Anderson e Raquel). A meus tios e tias, primos e primas, meu muito obrigado.

Agradeço a todos ingressantes (e anexados) da turma 2005.1 pela amizade, incentivo e torcida.

Serei sempre muito grato pela solicitude de meus orientadores, prof. Wellington e prof. George. A forma como me orientaram ao longo deste semestre me servirá de inspiração para o resto de minha vida acadêmica e profissional.

Resumo

Devido à crescente necessidade de sistemas de seguranças e de controle de acesso a áreas restritas, sistemas de reconhecimento de pessoas vem ganhando destaque dentro e fora da academia. Técnicas de reconhecimento estão sendo criadas e aperfeiçoadas, com a crescente necessidade do aumento de segurança. Embora existam sistemas de reconhecimento por outros tipos de biometria, há um grande esforço por empresas e universidades no campo de reconhecimento facial, já que este é um método que não exige contato físico com o dispositivo e não é necessário o consentimento da pessoa que está sendo analisada. Este trabalho apresenta duas abordagens para solução do problema de reconhecimento de faces. Uma delas usa a Análise de Componentes Principais e outra emprega a Morfologia Matemática. No final deste trabalho discutem-se as vantagens de uso de cada uma delas, bem como uma comparação dos resultados alcançados pelas duas abordagens.

Abstract

Due to the increasing need for security systems and to control people access to restrict areas, recognition systems are gaining attention within and outside the academy. Recognition techniques are being created and improved, with growing need for increased security. Although there are recognition systems of other types of biometrics, there is a great effort by companies and universities in the field of face recognition as this is a method that does not require physical contact with the device and is not needed the consent of the person being examined. This paper presents two approaches to solve the problem of Face recognition. One uses the Principal Component Analysis and the other uses the Mathematical Morphology. At the end of this paper is discussed the advantages of using each one, and a comparison of the results achieved by the two approaches is made.

Sumário

Resumo	i
Abstract	ii
Sumário	iii
Índice de Figuras	v
Índice de Tabelas	vii
Tabela de Símbolos e Siglas	viii
Capítulo 1 Introdução	9
1.1 Motivação e Problema.....	9
1.2 Objetivos	10
1.3 Organização do Trabalho	10
Capítulo 2 Fundamentação Teórica	12
2.1 Análise de Componentes Principais	12
2.2 Elementos de Processamento de Imagens	13
2.2.1 Representação de imagens digitais.....	14
2.2.2 Morfologia Matemática	14
2.2.3 Elemento Estruturante	15
2.2.4 Morfologia para imagens em tons de cinza	15
2.2.5 Dilatação para imagens em tons de cinza	16
2.2.6 Erosão para imagens em tons de cinza.....	19
2.2.7 Abertura para imagens em tons de cinza	21
2.2.8 Espectro de Padrões	22
Capítulo 3 Reconhecimento de faces	25
3.1 Tarefas na Identificação de uma Face	26
3.2 Sistema proposto e análise dos resultados	27
3.3 Reconhecimento de Faces utilizando Morfologia Matemática.....	27

3.4	Reconhecimento de Faces utilizando PCA	34
Capítulo 4	Conclusão e Trabalhos Futuros	38
Bibliografia		39
Apêndice A	Fundamentação Matemática	41

Índice de Figuras

Figura 1.	Representação dos eixos em uma imagem digital. Um ponto $f(x,y)$ é identificado (Figura adaptada de [4]).....	14
Figura 2.	Exemplo de elementos estruturantes: (a) em coluna, (b) em linha, (c) em cruz, (d) quadrado.	15
Figura 3.	Aplicando-se dilatação numa imagem utilizando uma matriz 3×3 como elemento estruturante [5].....	17
Figura 4.	Exemplo de dilatação em tons de cinza: (a) imagem original, (b) elemento estruturante em cruz com origem no centro e (c) imagem resultante da dilatação de (a) por (b)	17
Figura 5.	Exemplo de dilatação em tons de cinza em imagens. (a) Indica a imagem original e (b) imagem dilatada.....	18
Figura 6.	Detalhe de realce em áreas claras na imagem dilatada (b) em relação à imagem original (a).....	18
Figura 7.	Exemplo de erosão em tons de cinza: (a) imagem original, (b) elemento estruturante em cruz com origem no centro e (c) imagem resultante da erosão de (a) por (b).	20
Figura 8.	Exemplo de erosão em tons de cinza. (a) Indica a imagem original e (b) imagem erodida.....	20
Figura 9.	Análise da figura 8 (b).	21
Figura 10.	Resultado da operação de abertura sobre uma imagem em tons de cinza por elemento estruturante 5×5 em forma de quadrado com intensidade 1 e origem no centro: (a) imagem original e (b) imagem resultante após a aplicação de Abertura.....	22
Figura 11.	O Espectro de Padrões da imagem original da figura 12 com Abertura com elemento estruturante em forma de disco com raio 5.....	23
Figura 12.	Os resíduos após uma série de aplicações consecutivas da operação morfológica de Abertura na imagem.	24

Figura 13.	Etapas do problema de reconhecimento automático de faces [8].	26
Figura 14.	Arquitetura do modelo proposto para reconhecimento por Morfologia Matemática.	28
Figura 15.	Cenário ocorrido na primeira bateria de testes.	30
Figura 16.	Imagens de treinamento e imagem de teste de uma mesma pessoa para a primeira bateria de testes.	30
Figura 17.	Exemplo de cenário de testes. Uma das imagens do subconjunto de imagens de treinamento deve ser escolhida para ser considerado um acerto.	31
Figura 18.	Gráfico de imagens da base de treinamento x Distância Euclidiana.	31
Figura 19.	Imagem selecionada (em destaque) pelo critério de menor Distância Euclidiana entre os Espectros de Padrões.	32
Figura 20.	Resultado do reconhecimento por Morfologia Matemática.	33
Figura 21.	Arquitetura do modelo Proposto para Reconhecimento por Análise de Componentes Principais.	35
Figura 22.	Reconhecimento incorreto ocorrido na segunda bateria de testes.	36

Índice de Tabelas

Tabela 1. Aplicações típicas de reconhecimento de faces	25
--	----

Tabela de Símbolos e Siglas

PCA – Principal Components Analysis (Análise de Componentes Principais)

PDI – Processamento Digital de Imagens

Capítulo 1

Introdução

1.1 Motivação e Problema

Ao longo dos últimos anos, têm-se observado um aumento substancial de sistemas de segurança, os quais fazem uso da biometria, para regulamentação de acesso, o controle de ponto e a identificação de pessoas. A biometria se enquadra no campo da identificação e autenticação de usuários, mas foca na utilização de recursos estatísticos na análise e síntese de dados biológicos. Do grego bios = vida e metron = medida, a biometria propõe o estudo estatístico de características físicas e comportamentais exclusivas de cada ser humano.

Principalmente após o fatídico dia de 11 de setembro de 2001, governos de todo o globo tem financiado pesquisas para reconhecimento, identificação e rastreamento de suspeitos em áreas de grande circulação de pessoas como portos, aeroportos e estações de metrô a fim de combater o terrorismo [1]. Além disso, o aumento de crimes como clonagem de cartões de crédito e débito tem fomentado pesquisas para combater quadrilhas especializadas em roubos de dados de clientes dos bancos. Para combater esse tipo de crime, vários bancos estão implantando em sua rede de caixas eletrônicos sistemas de biometria, que identificam o cliente pela palma da mão ou por impressões digitais.

Uma das técnicas de reconhecimento biométrico mais comum é o reconhecimento pela biometria da face. Embora essa técnica não seja tão precisa quanto outros métodos de reconhecimento, como reconhecimento por impressões digitais, ela possui algumas vantagens. Por ser um método menos intrusivo e não ser necessário contato físico com algum dispositivo biométrico, o reconhecimento pela biometria da face tem atraído pesquisadores do campo da visão computacional. Entretanto, as largas variações nas faces humanas, devido à pose, iluminação e expressões faciais resultam em uma distribuição altamente complexa e deterioram o

desempenho desse tipo de reconhecimento [2], demonstrando, assim, que este é um campo aberto para novas pesquisas.

1.2 Objetivos

O Trabalho de Conclusão de Curso em questão tem como principal objetivo modelar e implementar uma ambiente para reconhecimento de faces, baseado em imagens de um conjunto de treinamento. A meta principal é construir uma prova de conceito a partir da modelagem e implementação de um sistema de reconhecimento usando Análise de Componentes Principais e outro sistema de reconhecimento por Morfologia Matemática, comparando os resultados obtidos por cada uma dessas técnicas.

As etapas de desenvolvimento deste trabalho são:

- Revisar literatura sobre reconhecimento de faces e estudar as técnicas utilizadas para este objetivo;
- Apresentar um módulo de aquisição de imagens para criação do conjunto de treinamento e apresentação de uma nova imagem a ser reconhecida;
- Modelar e implementar um módulo de Processamento Digital de Imagem (PDI) para processamento das imagens obtidas usando Morfologia Matemática;
- Modelar e implementar um módulo de Processamento Digital de Imagem (PDI) para processamento das imagens obtidas usando Análise de Componentes Principais;
- Comparar os resultados obtidos por cada uma das técnicas.

1.3 Organização do Trabalho

O primeiro capítulo desta monografia contém o texto introdutório sobre o trabalho, abordando motivação, problema, objetivos e metas. Além disso, apresenta uma visão geral sobre o sistema proposto.

O segundo capítulo reúne os principais conceitos necessários para a fundamentação teórica e compreensão do trabalho proposto. Para tal, apresenta os conceitos acerca da Análise de Componentes Principais e Morfologia Matemática.

Encontra-se no terceiro capítulo deste trabalho os principais conceitos a respeito do problema de reconhecimento de faces. Além disso, este capítulo

apresenta o modelo proposto de reconhecimento pela biometria da face através das técnicas de Análise de Componentes Principais e Morfologia Matemática bem como os resultados alcançados.

No último capítulo são apresentados os resultados alcançados com o sistema proposto, dificuldades encontradas e conclusões obtidas. Uma discussão acerca de trabalhos futuros também é apresentada.

Capítulo 2

Fundamentação Teórica

2.1 Análise de Componentes Principais

A análise de componentes principais (também conhecida como a transformação de Karhunen-Loève na teoria da comunicação) é uma ferramenta bastante utilizada em reconhecimento estatístico de padrões para seleção 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 exatamente 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 reter a maioria do conteúdo da informação intrínseco dos dados por maximizar a taxa de redução da variância; em outras palavras, o conjunto de dados sofre uma redução de dimensionalidade.

Análise de Componentes Principais pode ser descrita como uma transformação de um dado conjunto de vetores de entrada (variáveis) com o mesmo tamanho K formados a partir do vetor de n dimensões $\vec{x} = [x_1, x_2, \dots, x_n]^T$ em um vetor y de acordo com

$$\vec{y} = A(\vec{x} - m_x) \quad (1)$$

Este ponto de vista nos habilita a expressar uma fórmula simples (1), mas é necessário ter em mente que cada linha do vetor x consiste em K valores pertencentes a uma entrada. O vetor m_x na equação (1) é o vetor de valores médios de todas as variáveis de entrada definidas pela relação

$$m_x = E\{x\} = \frac{1}{K} \sum_{k=1}^K x_k \quad (2)$$

A matriz A definida na equação (1) é determinada pela matriz de covariância C_x . Linhas na matriz A são formadas pelos autovetores e de C_x ordenados em ordem decrescente de acordo com os autovalores correspondentes (Para uma descrição mais detalhada sobre o que são matrizes de covariância, autovalores e autovetores, veja o apêndice A). A avaliação da matriz C_x é possível de acordo com a relação

$$C_x = E\{(x - m_x)(x - m_x)^T\} = \frac{1}{K} \sum_{k=1}^K x_k x_k^T - m_x m_x^T \quad (3)$$

Como o vetor \vec{x} de entradas variáveis é de n dimensões, é óbvio que a dimensão de C_x é $n \times n$. Os elementos $C_x(i, i)$ situados na diagonal principal são as variâncias

$$C_x(i, i) = E\{(x_i - m_i)^2\} \quad (4)$$

de \vec{x} e os outros valores $C_x(i, j)$ determinam a covariância entre as variáveis de entrada x_i, x_j .

$$C_x(i, j) = E\{(x_i - m_i)(x_j - m_j)\} \quad (5)$$

As linhas de A na equação (1) são ortonormais, então a inversão da PCA é possível de acordo com a relação

$$\vec{x} = A^T \vec{y} + m_x \quad (6)$$

O núcleo da PCA definida pela equação (1) possui algumas outras propriedades interessantes, que podem ser usadas em processamento de sinais e de imagens para alcançar vários objetivos, como os mencionados nesta monografia.

2.2 Elementos de Processamento de Imagens

A seguir serão apresentados os alguns dos conceitos fundamentais sobre processamento de imagens, que serão utilizados no restante deste documento.

2.2.1 Representação de imagens digitais

O termo imagem monocromática ou imagem refere-se à função bidimensional de intensidade de luz $f(x, y)$ onde x e y representam coordenadas espaciais da imagem. O valor de f para qualquer ponto (x, y) da imagem, é proporcional ao brilho ou níveis de cinza daquela imagem naquele ponto [3]. Na figura 1 abaixo, temos um exemplo de imagem representada nos eixos das coordenadas x e y .

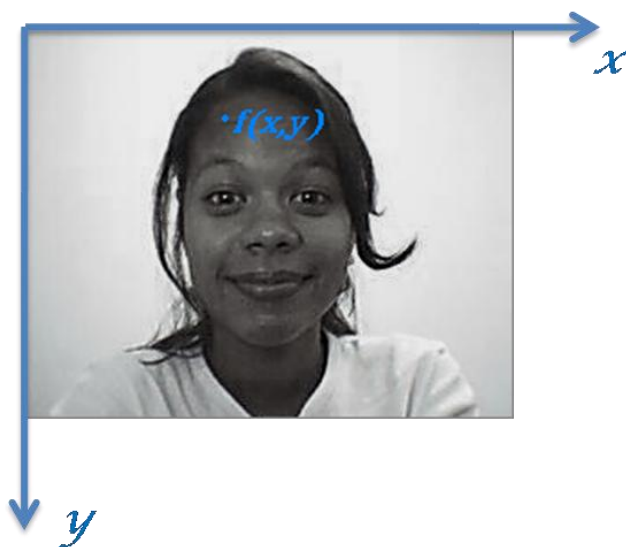


Figura 1. Representação dos eixos em uma imagem digital. Um ponto $f(x, y)$ é identificado (Figura adaptada de [4])

Uma imagem digital é uma imagem $f(x, y)$ discretizada tanto em coordenadas espaciais quanto em brilho. Uma imagem digital pode ser considerada como sendo uma matriz cujos índices de linhas e de colunas identificam um ponto na imagem e o correspondente valor do elemento da matriz identifica o nível de cinza naquele ponto. Os elementos dessa matriz digital são chamados de “pixels” (picture elements) [4].

2.2.2 Morfologia Matemática

Morfologia Matemática, originalmente desenvolvida em 1964 por Georges Matheron e Jean Serra na École des Mines de Paris, é uma ferramenta que se concentra nas estruturas geométricas da imagem para extração de componentes que são úteis para representação e descrição da forma de uma região, como

fronteiras e esqueletos. Algumas operações matemáticas podem ser aplicadas a conjuntos de pixels a fim de intensificar aspectos das formas tal que eles possam ser contados ou reconhecidos [5].

Esta técnica, baseada na teoria dos conjuntos, oferece uma abordagem unificada e poderosa para inúmeros problemas de processamento de imagens [3]. Os conjuntos em Morfologia Matemática representam as formas dos objetos em uma imagem. Por exemplo, o conjunto de todos os pixels pretos em uma imagem binária é uma descrição completa dessa imagem.

As operações na Morfologia Matemática são realizadas através de operadores elementares tais como a dilatação e a erosão. As técnicas morfológicas para pré e pós-processamento, como filtragem morfológica, afinamento e poda (“pruning”) também são muito exploradas atualmente [4].

2.2.3 Elemento Estruturante

O elemento estruturante é um conjunto definido e conhecido (forma e tamanho), que é comparado, a partir de uma transformação, ao conjunto desconhecido da imagem. Ele pode assumir várias formas e sua origem pode ser definida em qualquer ponto [4]. Para o caso de processamento de imagens em tons de cinza, os elementos do conjunto elemento estruturante podem assumir valores intermediários entre 0 e 255. A figura 2 exemplifica algumas configurações de elementos estruturantes. Vale ressaltar que não necessariamente o elemento estruturante é uma matriz retangular ou quadrada.

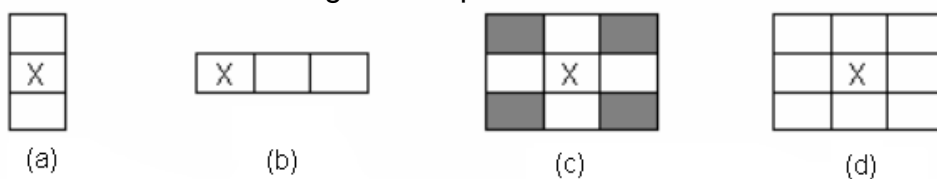


Figura 2. Exemplo de elementos estruturantes: (a) em coluna, (b) em linha, (c) em cruz, (d) quadrado.

2.2.4 Morfologia para imagens em tons de cinza

As imagens em nível de cinza são consideradas como conjuntos tridimensionais nos quais os operadores morfológicos podem ser aplicados. Nesse caso, dois componentes de cada elemento do conjunto se referem às coordenadas do pixel, enquanto o terceiro corresponde ao valor discreto de intensidade.

No contexto de imagens com essa característica (níveis de cinza) as funções são descritas na forma $f(x, y)$ para a imagem de entrada e $b(x, y)$ para o elemento estruturante. A suposição é que estas funções são discretas, isto é, se Z denota o conjunto de inteiros reais, a hipótese é que (x, y) são inteiros de $Z \times Z$ e que f e b são funções que assumem um valor representando o nível de cinza para cada par de coordenadas distintos (x, y) [3].

2.2.5 Dilatação para imagens em tons de cinza

Utilizando um elemento estruturante onde todos os valores são positivos para efetuar uma dilatação numa imagem em níveis de cinza, obtém-se como imagem resultante uma imagem com detalhes escuros atenuados e detalhes claros acentuados. A operação de dilatação em tons de cinza fornece como saída uma imagem com as mesmas dimensões da imagem de entrada e pixels com valores de 0 a 255, para o caso de uma imagem original com 256 níveis de cinza.

Matematicamente, dilatação em escala de cinza de uma imagem f por um elemento estruturante b , denotado por $f \oplus b$ é definido por

$$(f \oplus b)(s, t) = \max\{f(s - x, t - y) + b(x, y) \mid (s - x), (t - y) \in D_f; (x, y) \in D_b\}$$

onde D_f e D_b são os domínios de f e b respectivamente.

Isso significa que o algoritmo para aplicação de dilatação em imagens em tons de cinza segue os seguintes passos:

1. O elemento estrutural é posicionado no início da imagem a ser aplicada a dilatação;
2. É feita a comparação da soma das intensidades dos pixels da imagem original com as intensidades dos pixels correspondentes da matriz do elemento estruturante;
3. O pixel correspondente da imagem de saída terá a intensidade da maior soma desta comparação. Caso a maior soma ultrapassar 255 o valor assumido para a intensidade do pixel correspondente será 255.

4. A matriz do elemento estruturante é posicionada no elemento seguinte da matriz que representa a imagem original, retornando-se ao passo de número 2. Caso o último elemento da matriz da imagem original já tenha sido visitado pela matriz do elemento estruturante, o processo de dilatação está concluído.

A figura 3 ilustra a execução do algoritmo de dilatação utilizando um elemento estruturante de dimensões 3×3 . A matriz do elemento estruturante está posicionada “sobre” a imagem original sendo realizada a comparação da soma das intensidades dos pixels das duas matrizes.

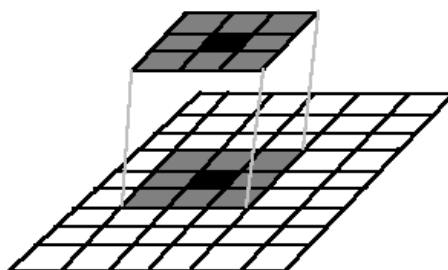


Figura 3. Aplicando-se dilatação numa imagem utilizando uma matriz 3×3 como elemento estruturante [5].

Podemos observar na figura 4 que, se o elemento estruturante tem um valor uniforme, ou seja, todos os pontos do elemento que estão ligados possuem a mesma intensidade, então esta operação é equivalente a tomar o valor máximo da região da imagem definida pelo elemento estruturante somado a este valor uniforme do elemento estruturante [4].

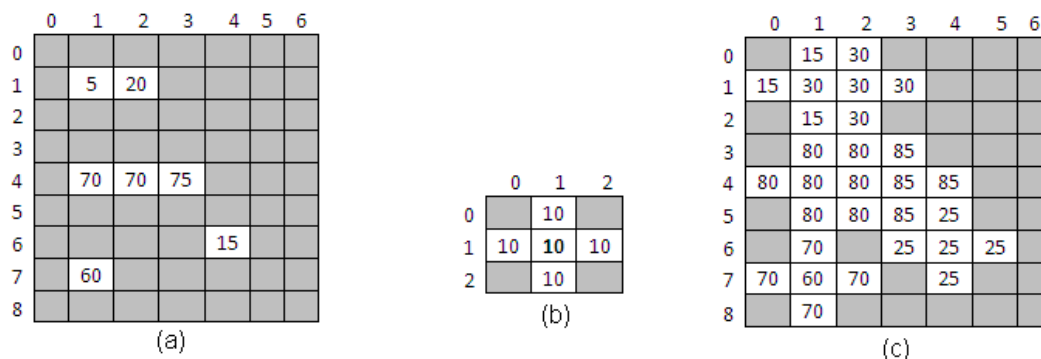


Figura 4. Exemplo de dilatação em tons de cinza: (a) imagem original, (b) elemento estruturante em cruz com origem no centro e (c) imagem resultante da dilatação de (a) por (b)

As imagens da figura 5 é um exemplo da aplicação da dilatação sobre uma imagem em tons de cinza, utilizando uma matriz quadrada 5×5 como elemento estruturante.



(a)



(b)

Figura 5. Exemplo de dilatação em tons de cinza em imagens. (a) Indica a imagem original e (b) imagem dilatada.

Observe que após a aplicação de dilatação na imagem original (a) da figura 5, foi alcançado o efeito de uma imagem mais clara (já que todos os valores do elemento estruturante eram positivos) e foram eliminados alguns detalhes escuros. Na figura 6 vemos o detalhe da área clara no telhado da casa sendo realçado.



(a)



(b)

Figura 6. Detalhe de realce em áreas claras na imagem dilatada (b) em relação à imagem original (a).

2.2.6 Erosão para imagens em tons de cinza

A imagem resultante da aplicação de uma Erosão não somente aparentará ser mais clara que a original, mas os tamanhos das áreas mais escuras serão reduzidos.

Matematicamente, erosão em escala de cinza de uma imagem f por um elemento estruturante b , denotado por $f \ominus b$ é definido por

$$(f \ominus b)(s, t) = \min\{f(s + x, t + y) - b(x, y) \mid (s + x), (t + y) \in D_f; (x, y) \in D_b\}$$

onde D_f e D_b são os domínios de f e b , respectivamente.

Isso significa que o algoritmo para aplicação de erosão em imagens em tons de cinza segue os seguintes passos:

1. O elemento estrutural é posicionado no início da imagem a ser aplicada a dilatação;
2. É feita a comparação da subtração das intensidades dos pixels da imagem original com as intensidades dos pixels correspondentes da matriz do elemento estruturante;
3. O pixel correspondente da imagem de saída terá a intensidade da menor subtração desta comparação. Caso a menor subtração seja menor que 0 (zero) o valor assumido para a intensidade do pixel correspondente será 0 (zero).
4. A matriz do elemento estruturante é posicionada no elemento seguinte da matriz que representa a imagem original, retornando-se ao passo de número 2. Caso o último elemento da matriz da imagem original já tenha sido visitado pela matriz do elemento estruturante, o processo de erosão está concluído.

Podemos observar na figura 7 que, se o elemento estruturante tem um valor uniforme, ou seja, todos os pontos do elemento que estão ligados possuem a mesma intensidade, então esta operação é equivalente a tomar o valor mínimo da região da imagem definida pelo elemento estruturante subtraído do valor uniforme do elemento estruturante [4].

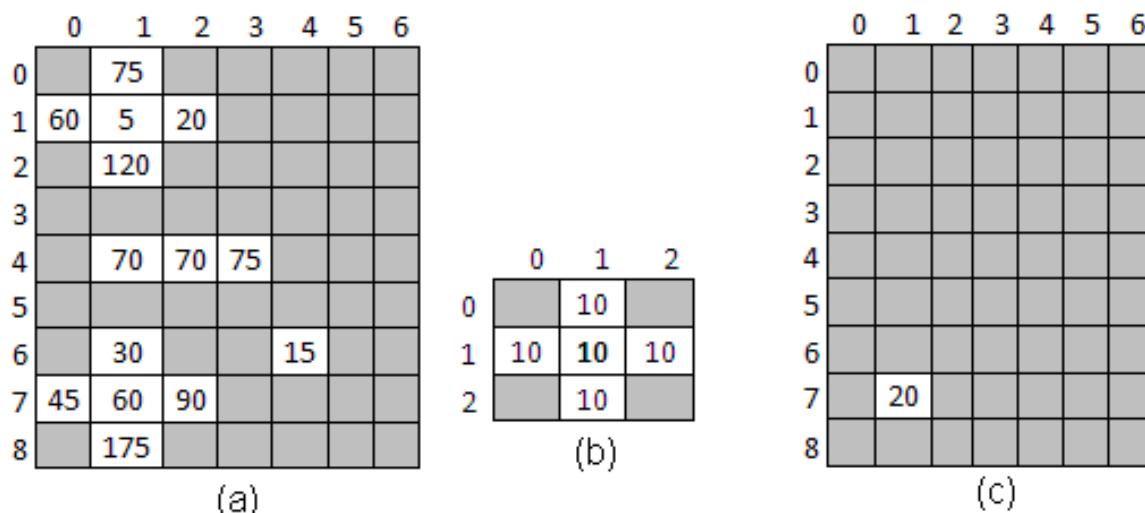


Figura 7. Exemplo de erosão em tons de cinza: (a) imagem original, (b) elemento estruturante em cruz com origem no centro e (c) imagem resultante da erosão de (a) por (b).

As imagens da figura 8 é um exemplo da aplicação da erosão sobre uma imagem em tons de cinza, utilizando uma matriz quadrada 5×5 como elemento estruturante.



Figura 8. Exemplo de erosão em tons de cinza. (a) Indica a imagem original e (b) imagem erodida.

Na figura 9 correspondente à análise da figura 8(b) podemos observar na área circundada em vermelho que uma forma da figura 8(a) desaparece após

aplicação da erosão, o que exemplifica o fato da erosão remover detalhes claros da imagem.



Figura 9. Análise da figura 8 (b).

2.2.7 Abertura para imagens em tons de cinza

A abertura de uma imagem f pela subimagem (elemento estruturante) b , denotado por $f \circ b$, é

$$f \circ b = (f \ominus b) \oplus b$$

ou seja, a operação de abertura de uma imagem é simplesmente a erosão de f por b , seguido por uma dilatação do resultado por b [3]. A figura 10 nos mostra um exemplo de aplicação de abertura em uma imagem em tons de cinza. Essa operação elimina as protuberâncias da imagem de forma mais suave que a erosão.



Figura 10. Resultado da operação de abertura sobre uma imagem em tons de cinza por elemento estruturante 5×5 em forma de quadrado com intensidade 1 e origem no centro: (a) imagem original e (b) imagem resultante após a aplicação de abertura.

2.2.8 Espectro de Padrões

O Espectro de Padrões obtém o histograma da distribuição dos tamanhos dos vários objetos que compõem a imagem. Os parâmetros obtidos a partir deste espectro permitem desenvolver uma análise quantitativa do conteúdo da imagem.

Se for aplicada uma abertura numa imagem utilizando-se um elemento estruturante B , alguns elementos devem desaparecer na nova imagem obtida. Se for aplicada a mesma operação morfológica (abertura) usando desta vez um elemento estruturante maior, então mais elementos na imagem na imagem original A vão desaparecer. Desta forma, é possível determinar como o número de pixels eliminados aumenta quando a imagem é morfológicamente aberta usando $nB = B \oplus B \oplus B \dots \oplus B$ n vezes. O gráfico resultante do número de pixels eliminados *versus* o tamanho do elemento estruturante é chamado de Espectro de Padrões ou Curva Granulométrica [10].

O Espectro de Padrões para imagens em tons de cinza é criado como a seguir:

1. Aplicar uma Abertura na imagem com um elemento estruturante adequado;
2. Somar todos os valores em escala de cinza para todos os pixels da imagem obtida e dividir pela soma de todos os valores dos pixels da imagem original, obtendo assim λ ;
3. Calcular $1 - \lambda$ e colocar o valor da diferença no n -ésimo elemento do vetor que representa a distribuição acumulada (começando pelo índice zero);
4. Calcular o valor da diferença entre o n -ésimo elemento da distribuição acumulada e o valor do n -ésimo -1 elemento da distribuição acumulada, e inserir o valor encontrado no n -ésimo elemento do vetor que representa o Espectro de Padrões;
5. Repetir os passos 1 a 4 com a imagem resultante da aplicação da abertura como sendo a nova imagem original (n incrementa de 1).

Um exemplo: Figura 11 mostra o Espectro de Padrões da primeira imagem na série de imagens da figura 12. Nesta figura é possível ver os resíduos depois de uma operação morfológica de Abertura. Inicialmente, os menores objetos são eliminados, para só então os maiores.

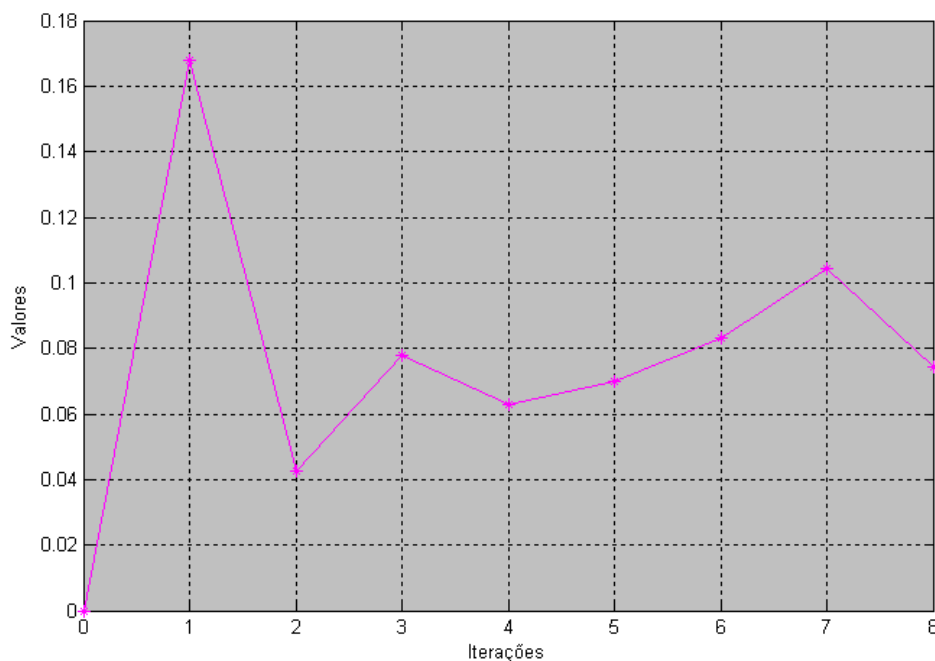


Figura 11. O Espectro de Padrões da imagem original da figura 12 com Abertura com elemento estruturante em forma de disco com raio 5.

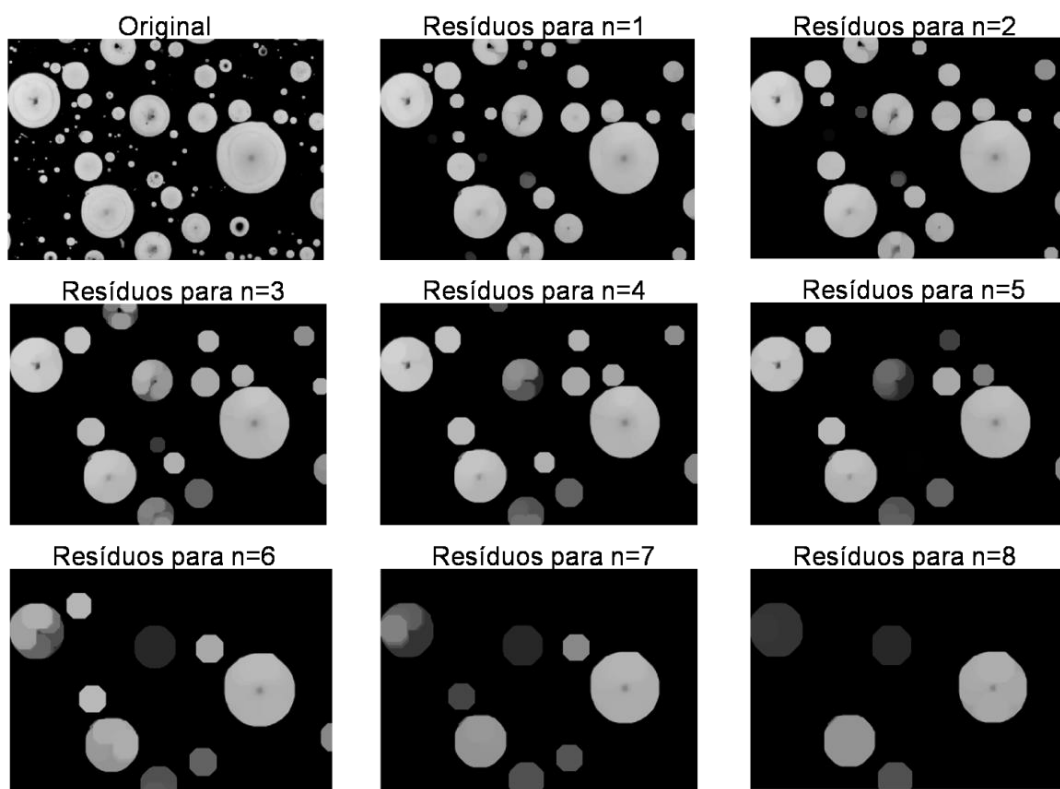


Figura 12. Os resíduos após uma série de aplicações consecutivas da operação morfológica de Abertura na imagem.

Capítulo 3

Reconhecimento de faces

Reconhecimento de faces tem se tornado uma área popular de pesquisa em visão computacional e uma das aplicações de análise de imagens de maior interesse. Devido à natureza do problema, não somente pesquisadores da área da ciência da computação estão interessados nesse campo de pesquisa, mas também neurocientistas e psicologistas. É opinião geral que os avanços em visão computacional irão proporcionar às áreas da neurociência e psicologia uma melhor compreensão de como o cérebro humano funciona [6]. Embora métodos extremamente confiáveis de identificação biométrica existam, como por exemplo, análises de impressões digitais e leitura de retina ou íris, esses métodos contam com a cooperação dos participantes, ao passo que um sistema de identificação de pessoas baseados em análises de imagens frontais ou de perfil da face é frequentemente efetiva sem a cooperação ou conhecimento do participante. A tabela 1 lista algumas das aplicações de reconhecimento de faces [7].

Tabela 1. Aplicações típicas de reconhecimento de faces [7].

Áreas	Aplicações específicas
Biometria	Licença de motorista, Imigração, Registro de Identificação Nacional, Passaporte, Título de Eleitor
Segurança da informação	Logon do Sistema Operacional, Aplicações de Segurança, Segurança de Banco de Dados, Encriptação de arquivos, Segurança da Intranet, Acesso à Internet, Registros Médicos, Caixas Eletrônicos
Aplicação da Lei e Vigilância	Vigilância Avançada por meio de Vídeo, Controle CCTV, Controle de entrada, Análise de pós-evento, Furtos, Rastreamento de Suspeitos e Investigação
Smart Cards	Segurança do Valor Armazenado, Autenticação do Usuário
Controle de Acesso	Facilidade de Acesso, Acesso de Veículos

Um enunciado geral do problema de reconhecimento facial (em visão computacional) pode ser formulado da seguinte forma: Dado uma fotografia ou

imagens de vídeo de uma cena, identificar ou verificar uma ou mais pessoas na cena usando um banco de dados de imagens de faces [6].

3.1 Tarefas na Identificação de uma Face

O reconhecimento de faces é uma particularização do problema geral de reconhecimento de padrões [8], cujo objetivo é classificar informações baseado ou em conhecimento a priori ou em informações estatísticas extraídas dos padrões [9]. Pode ser ilustrado por um diagrama como o mostrado na figura 13.

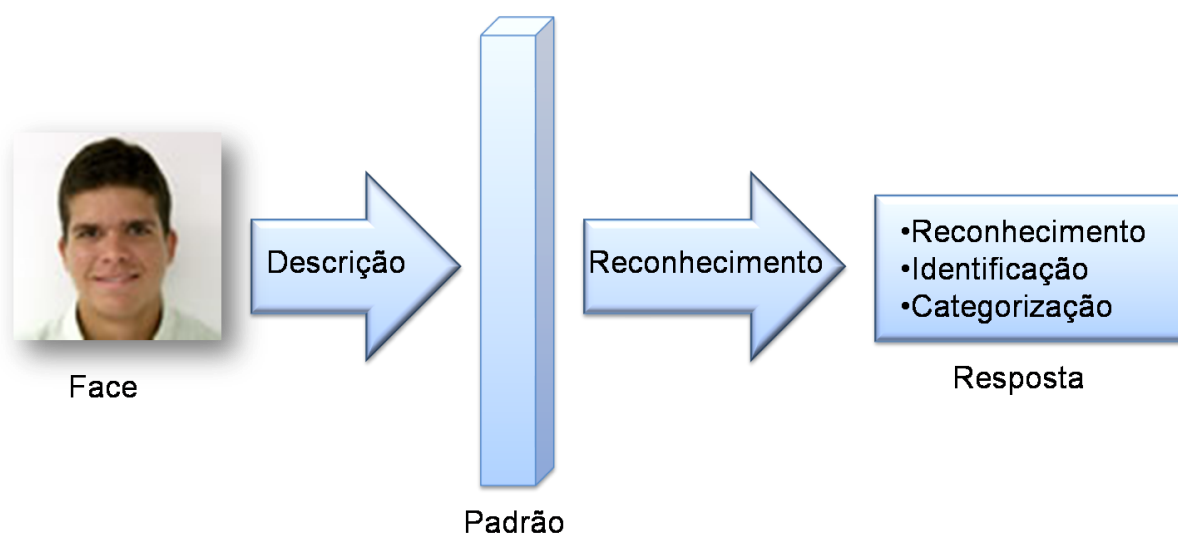


Figura 13. Etapas do problema de reconhecimento automático de faces [8].

Segundo [11] e [12], São pelo menos quatro tarefas relacionadas com o reconhecimento de faces:

1. **Classificação:** Consiste na identificação de uma face x assumindo-se que ela pertence a uma pessoa do conjunto Ω , sendo Ω o conjunto de todas as classes (ou pessoas) existentes. Em outras palavras, assumindo-se que x , um padrão originário de uma face cuja classificação é desconhecida, pode ser classificado como um padrão de alguma classe w_i , tal que $w_i \in \Omega$, a tarefa de classificação consiste em determinar o valor de i .

2. **Conhecido-desconhecido:** Objetiva decidir se a face é ou não um membro de Ω , ou seja, se x pode ser classificado como um padrão de alguma classe de Ω .

3. **Verificação:** Dado que a identidade w_i de uma face x foi determinada através de outro meio não visual, essa tarefa busca confirmar a identidade dessa pessoa usando imagens de face, ou seja, confirmar se x é da classe w_i . Isso equivale à tarefa “conhecido-desconhecido” com apenas uma pessoa conhecida dentro da base de dados do conjunto de treinamento.

4. **Reconhecimento completo:** visa determinar se uma face é de uma classe Ω e, em caso positivo, determinar sua identidade w_i .

3.2 Sistema proposto e análise dos resultados

O objetivo deste trabalho é comparar duas técnicas para reconhecimento de faces, sendo uma delas bem sedimentada para esse objetivo enquanto a outra é uma proposta nova. Atualmente, existem inúmeros artigos sobre técnicas de reconhecimento de biometria da face utilizando a Análise de Componentes Principais, discutida no capítulo 2. No entanto, foi verificada uma ausência de técnicas que usem morfologia da face para reconhecimento. Assim, neste capítulo, é descrito o modelo utilizado para extração de características através destas duas técnicas para reconhecimento de faces e posterior comparação dos resultados obtidos.

A prova de conceito dessas duas técnicas foi desenvolvida usando-se a *software* MATLAB versão R2009a. O programa possui diversas extensões (chamadas *toolboxes*), incluindo muitos recursos para processamento de imagens [13].

3.3 Reconhecimento de Faces utilizando Morfologia Matemática

A técnica de Morfologia Matemática, apresentada no capítulo 2, procura analisar as estruturas geométricas da imagem para extração de componentes que são úteis para representação e descrição da forma de uma região. Assim, componentes como: orientação dos olhos, tamanho e formato do nariz, uso ou não de óculos, posição das sobrancelhas em relação ao restante do rosto, são

elementos que são considerados para a verificação de semelhanças entre imagens de faces.

A estrutura utilizada pelo nosso modelo está ilustrada na figura 14. Inicialmente, um conjunto de imagens de faces é usado para compor o conjunto de treinamento. Logo em seguida, é calculada o Espectro de Padrões para cada uma das imagens da base de treinamento. O mesmo processo é feito para a imagem de teste selecionada, gerando o Espectro de Padrões para a imagem de teste. Então, é calculada a Distância Euclidiana entre o Espectro de Padrões da imagem de teste e o Espectro de Padrões para cada uma das imagens da base de treinamento. A menor Distância Euclidiana encontrada é o critério utilizado para o sistema associar a imagem de teste a alguma imagem pertencente à base de treinamento. Com isso, a imagem de teste é reconhecida pela imagem da base de testes que tiver um Espectro de Padrões mais semelhante.

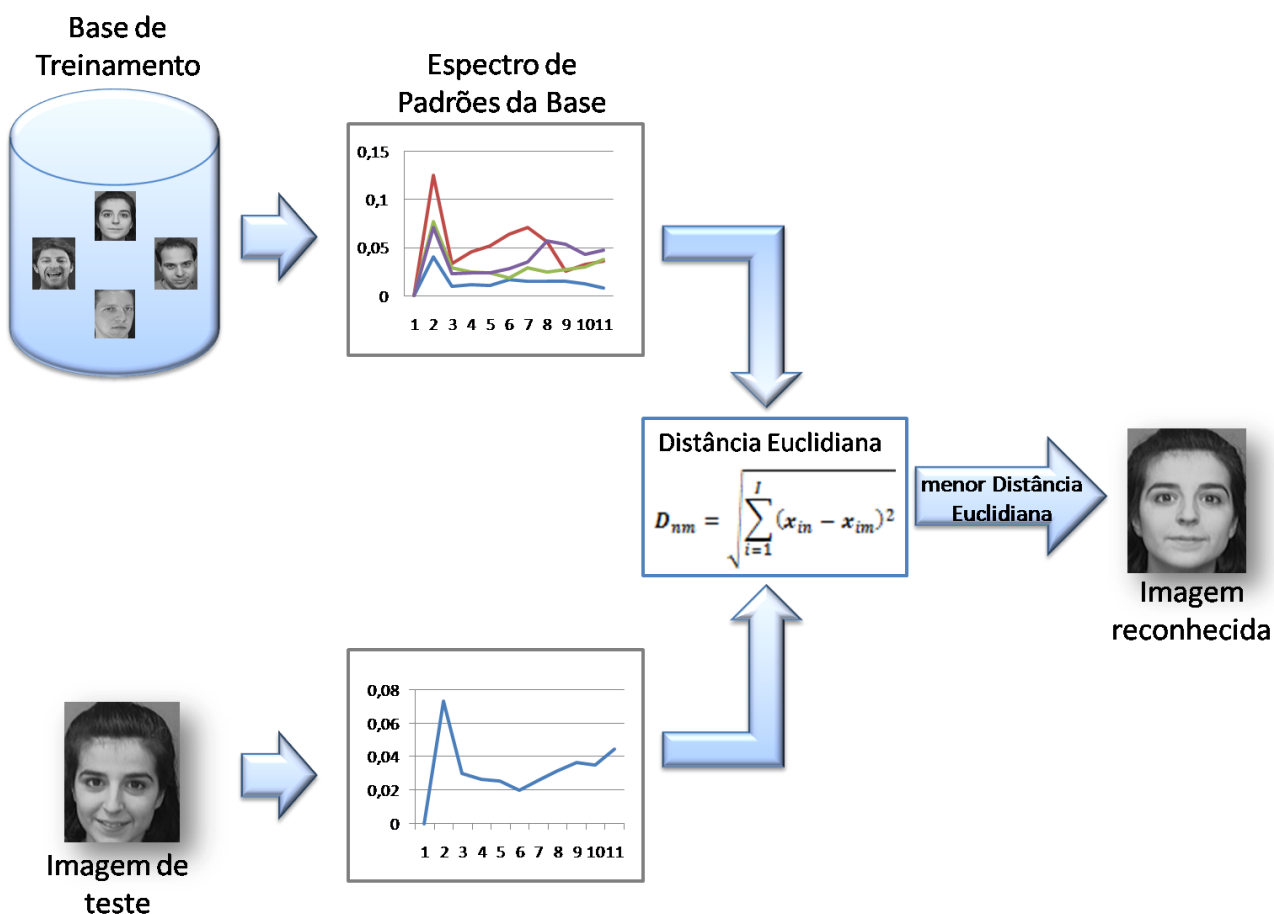


Figura 14. Arquitetura do modelo proposto para reconhecimento por Morfologia Matemática.

A base de treinamento para a prova de conceito do modelo proposto foi composta inicialmente de apenas vinte imagens, em tons de cinza, sendo duas imagens da face para cada uma das dez pessoas representadas. O conjunto de imagens utilizadas foi extraído da base criada pela *AT&T Laboratories*, conhecido como Banco de Imagens de Face ORL [14]. Nesse banco, há dez imagens diferentes de cada um dos quarenta indivíduos distintos. Para alguns dos sujeitos, as imagens foram tiradas em momentos diferentes, variando a iluminação, expressões faciais (olhos abertos/fechados, sorrindo/não sorrindo) e detalhes faciais (com/sem óculos). Todas as imagens foram tiradas contra um fundo escuro homogêneo com as pessoas em uma posição vertical, posição frontal (com tolerância de alguns movimentos laterais).

Para a primeira bateria de testes, com base de treinamento composta de vinte imagens de faces, foram selecionadas dez imagens de faces pertencentes a pessoas distintas para verificar a eficiência da técnica. Nessa situação, o sistema só obteve 50% de acerto, mas algumas características interessantes devem ser ressaltadas.

Uma das imagens de entrada do sistema foi a imagem de teste da figura 15. Após calcular o Espectro de Padrões para a imagem de entrada, foi selecionada a imagem da Base de Treinamento com a menor Distância Euclidiana entre seu respectivo Espectro de Padrões e o Espectro da imagem de entrada. A imagem que atendeu a esses requisitos foi a imagem de saída da figura 15. Embora a técnica não tenha acertado nesse cenário, verificamos que a ocorrência do erro se deu devido aos aspectos morfológicos das duas imagens: ambas as faces estão posicionadas de uma forma semelhante (de frente para a câmera), estão com os olhos abertos, tamanho do nariz bastante similar, inclinação da face semelhante e uso dos mesmos óculos nas duas imagens. Outro fator que contribuiu para o reconhecimento incorreto foi as outras imagens de treinamento da pessoa da imagem de teste. Essas imagens da face possuíam aspectos morfológicos bastantes distintos em relação à imagem de teste. A figura 16 nos mostra as duas imagens pertencentes à base de treinamento da mesma pessoa da imagem de teste.

Para se obter o Espectro de Padrões das imagens, foi calculada a Abertura para cada uma delas, usando-se como Elemento Estruturante uma matriz quadrada

3x3 positiva. Foram realizadas quarenta iterações no cálculo de Abertura para se gerar o Espectro de Padrões.

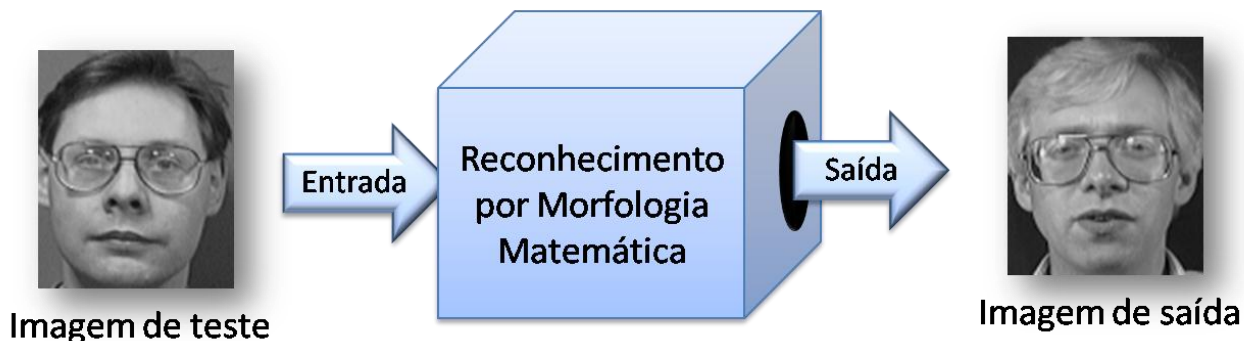


Figura 15. Cenário ocorrido na primeira bateria de testes.



Figura 16. Imagens de treinamento e imagem de teste de uma mesma pessoa para a primeira bateria de testes.

Após uma investigação a respeito do baixo número de respostas corretas da primeira bateria de testes (apenas 50% de acerto), foi ampliada a base de treinamento de vinte imagens para noventa imagens de faces, todas representando as mesmas dez pessoas do conjunto inicial. Assim, cada pessoa era representada com nove imagens na Base de Treinamento. Com isso, a taxa de acerto do sistema subiu dos 50% iniciais, para o 100% de acerto na segunda bateria de testes.

Foram realizados testes com imagens de dez pessoas. Na base de treinamento há nove imagens para cada uma delas, sendo estas distintas entre si e entre as imagens de testes. Para o primeiro teste, foi selecionada a imagem representada pela face à esquerda da figura 17. No conjunto de noventa imagens,

nove pertenciam a esta mesma pessoa, e também está representada na mesma figura.



Figura 17. Exemplo de cenário de testes. Uma das imagens do subconjunto de imagens de treinamento deve ser escolhida para ser considerado um acerto.

Na figura 18, o gráfico explicita as Distâncias Euclidianas entre todas as noventa imagens que compõem a base de treinamento e a imagem de entrada, sendo esta a imagem de teste à esquerda da figura 17. Observa-se facilmente, através do gráfico, que a imagem com a menor Distância Euclidiana é a sétima imagem da base de treinamento. Essa imagem corresponde à imagem em destaque da figura 19.

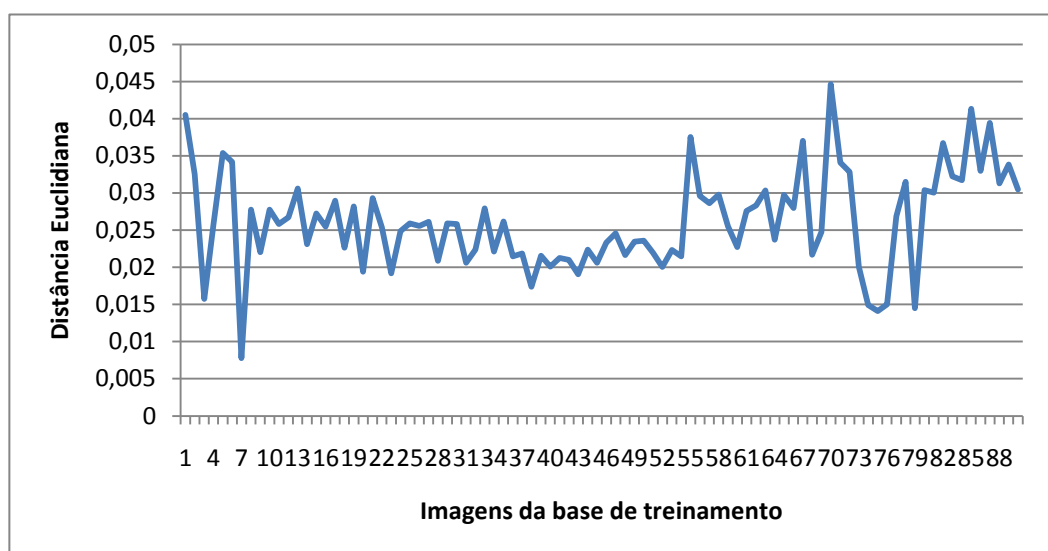


Figura 18. Gráfico de imagens da base de treinamento x Distância Euclidiana.

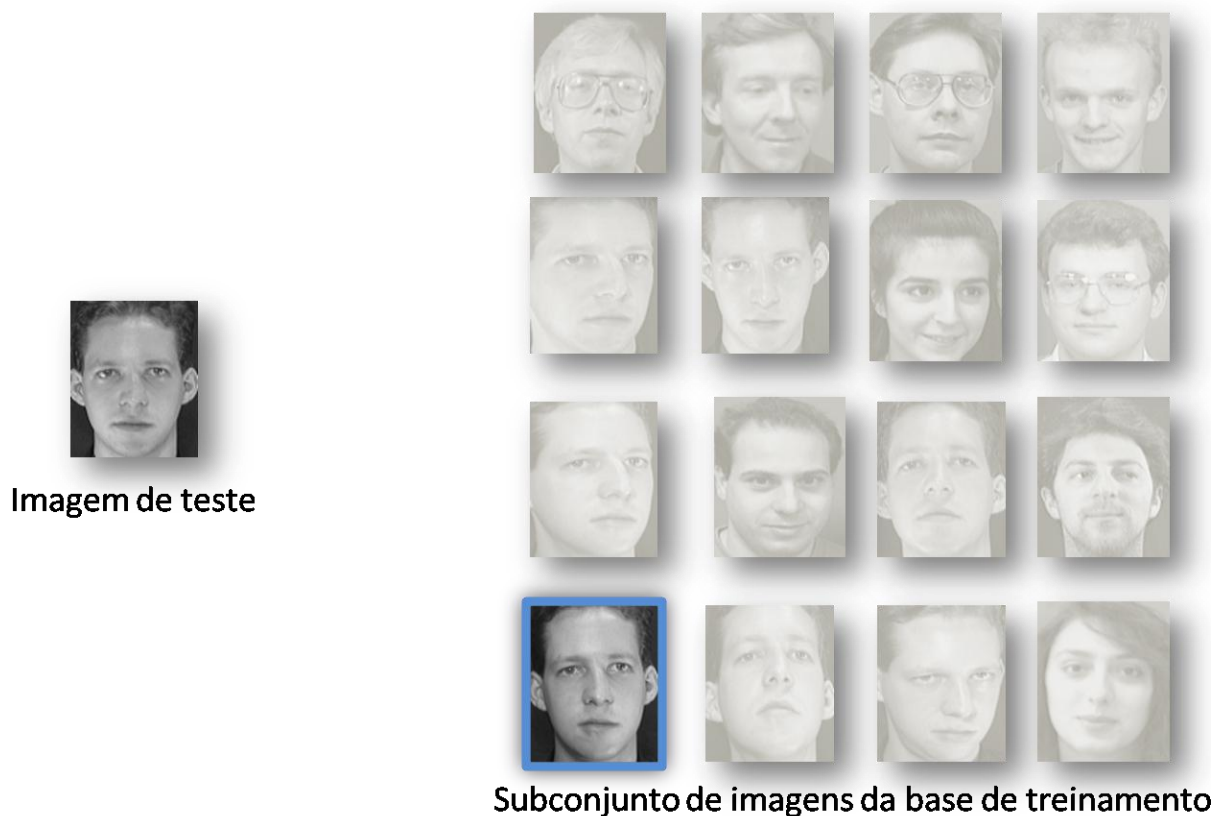


Figura 19. Imagem selecionada (em destaque) pelo critério de menor Distância Euclidiana entre os Espectros de Padrões.

Na figura 20 são apresentados os resultados obtidos para todas as imagens de testes, onde as imagens de faces à esquerda pertencem às entradas e as imagens à direita correspondem às imagens reconhecidas. Para se obter o Espectro de Padrões das imagens, foi calculada a Abertura para cada uma delas, usando-se como Elemento Estruturante uma matriz quadrada 3×3 positiva. Foram realizadas quarenta iterações no cálculo de Abertura para se gerar o Espectro de Padrões. A escolha desse número de iterações se deve à natureza do Espectro de Padrões; a partir de um número de iterações maior, a variação do Espectro de Padrões é mínima (assumindo valores da distribuição acumulada da imagem muito próximas a zero) para todas as imagens observadas, o que seria irrelevante na comparação dos diferentes espectros.

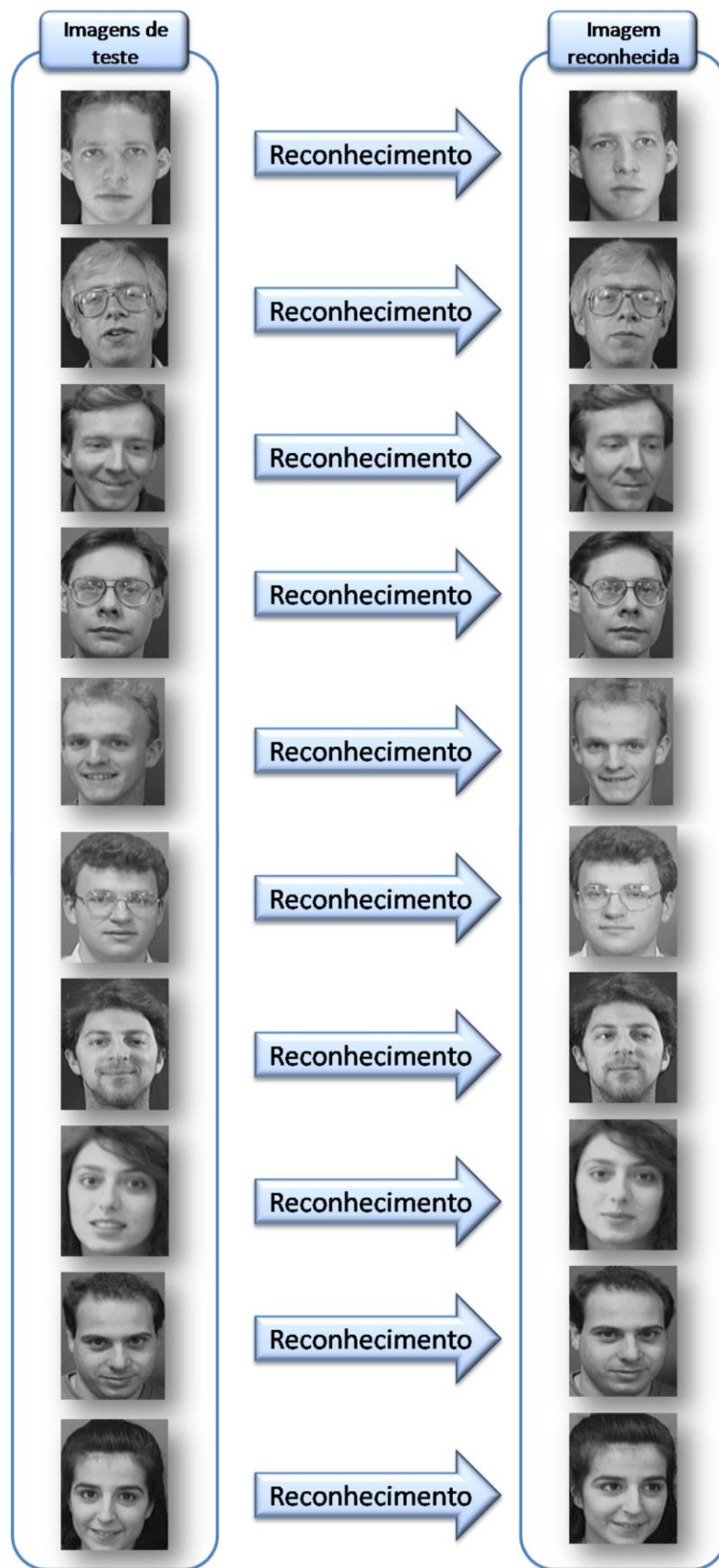


Figura 20. Resultado do reconhecimento por Morfologia Matemática.

3.4 Reconhecimento de Faces utilizando PCA

A Análise de Componentes Principais, apresentada no capítulo 2, é uma ferramenta bastante utilizada em reconhecimento estatístico de padrões para seleção de características, como os da face. A aplicação desta técnica é recomendada quando se deseja identificar padrões em dados, e expressá-los de forma que as similaridades e diferenças sejam realçadas.

Na figura 21 são ilustrados os componentes do sistema proposto, para prova de conceito da técnica de reconhecimento por PCA. Assim como no sistema de reconhecimento de faces por Morfologia Matemática, um conjunto de imagens de faces é selecionado para compor o conjunto de treinamento, representado na figura pelo módulo de Base de Treinamento. Uma dessas imagens deverá ser selecionada de acordo com a imagem de entrada do sistema. Essa imagem deverá corresponder a pessoa da imagem de entrada, para caracterizar um reconhecimento bem sucedido.

No módulo “Análise de Componentes Principais” será utilizada PCA para determinar as características mais discriminantes entre todas as imagens das faces na Base de Treinamento. Ali é calculado a imagem média da Base de Treinamento, a matriz de covariância, os Autovetores da matriz de covariância e o vetor de características. A mesma operação também deverá ocorrer para a imagem de entrada do sistema, representada na mesma figura pela Imagem de Teste.

Em seguida, será feita uma comparação entre características principais, ressaltadas pelo módulo de Análise de Componentes Principais, por medir a Distância Euclidiana entre elas. A menor Distância Euclidiana encontrada é o critério utilizado para o sistema associar a imagem de teste a alguma imagem pertencente à base de treinamento, finalizando o processo de reconhecimento por PCA.

Para a primeira bateria de testes, foi utilizado a mesma Base de Treinamento, com vinte imagens de faces, e as mesmas imagens de entrada da primeira bateria de testes realizadas para o sistema de reconhecimento por Morfologia Matemática. Nesse cenário, a técnica de reconhecimento por PCA obteve uma taxa de acerto de 90% no reconhecimento, contra os 50% dos primeiros testes com a técnica de Morfologia Matemática.

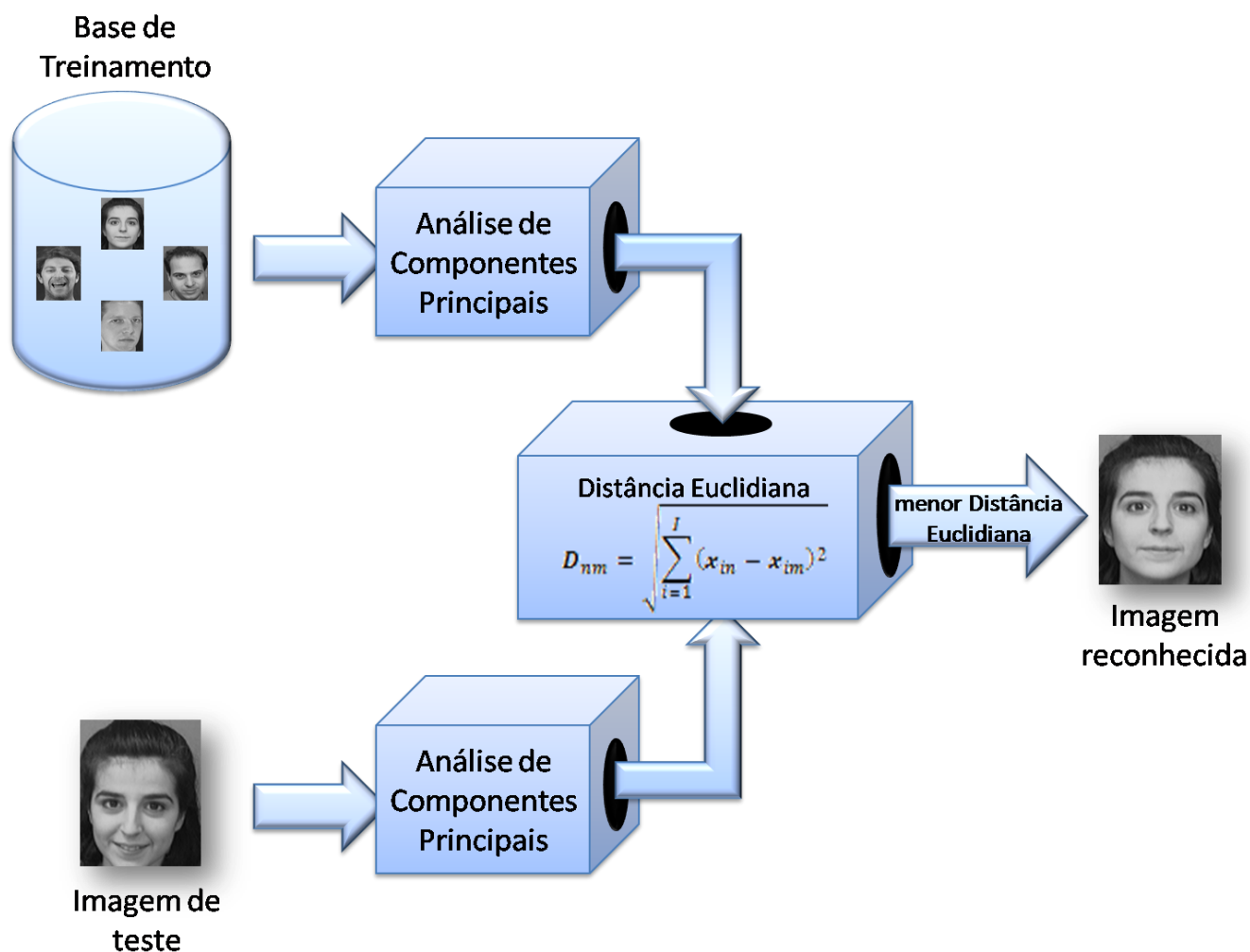


Figura 21. Arquitetura do modelo Proposto para Reconhecimento por Análise de Componentes Principais.

Na segunda bateria de testes, utilizou-se a mesma Base de Treinamento, com noventa imagens de faces, e as mesmas imagens de entrada da segunda bateria de testes realizadas para o sistema de reconhecimento por Morfologia Matemática. Nesse cenário, a técnica de reconhecimento por PCA manteve uma taxa de acerto de 90% no reconhecimento, apenas falhando pelo caso ilustrado na figura 22. Como discutido anteriormente, para este mesmo cenário, a técnica de reconhecimento por Morfologia Matemática obteve um acerto em 100% dos casos apresentados.

É importante ressaltar que apesar de uma taxa de acerto ligeiramente inferior, a técnica de reconhecimento por Análise de Componentes Principais demandou um tempo de processamento muito pequeno, da ordem de alguns segundos, quando comparado à técnica de reconhecimento por Morfologia Matemática. Neste trabalho

não será realizada a comparação das técnicas com respeito ao seu desempenho. Portanto, uma comparação dos tempos de execução de cada técnica não serão apresentados. Além disso, é interessante observar que, na segunda bateria de testes, em apenas um caso as duas técnicas concordaram em qual imagem da Base de Treinamento teria uma similaridade maior em relação à imagem de teste.



Figura 22. Reconhecimento incorreto ocorrido na segunda bateria de testes.

Utilizou-se Morfologia Matemática para identificar estruturas geométricas da imagem para extração de componentes que são úteis para representação e descrição da forma da face. Os Espectros de Padrões são únicos para cada imagem da face, porém, para uma mesma pessoa estes tendem a serem semelhantes. Com isso, foi feita uma comparação entre os Espectros da imagem de entrada e das imagens já previamente conhecidas a fim de identificar a imagem que melhor se assemelha àquela de entrada. Porém, para uma base pequena de imagens conhecidas, esta abordagem só foi bem sucedida (acertou o reconhecimento) em apenas 50% dos casos. No entanto, para uma base maior, a taxa de acerto do sistema alcançou os 100%.

Para comparação com a abordagem proposta, foi implementada uma técnica estatística de reconhecimento de padrões para seleção de características já bastante utilizada para o problema de reconhecimento de faces. Tal técnica, conhecida como Análise de Componentes Principais (*Principal Components Analysis - PCA*) é utilizada para a identificação da pessoa a partir de uma imagem de sua própria face. A mesma base de imagens conhecidas utilizadas nas duas baterias de testes da abordagem anterior foi utilizada com objetivo de comparar a eficácia do reconhecimento. Para a primeira bateria de testes, com uma base reduzida de

imagens conhecidas, a técnica de PCA mostrou-se melhor do que a técnica de Morfologia Matemática, com 90% de acerto contra os 50% da outra técnica. Entretanto, na segunda bateria de testes, onde foi utilizada uma base estendida, a técnica de reconhecimento pela Análise dos Componentes Principais não se saiu melhor do que a outra técnica, com 90% de acertos no reconhecimento, contra o 100% de acerto da técnica que utilizou Morfologia Matemática.

Capítulo 4

Conclusão e Trabalhos Futuros

Este trabalho de conclusão de curso propôs a comparação de duas técnicas para o problema de reconhecimento de faces. A primeira delas é uma abordagem nova para este tipo de problema, que visa identificar nas imagens da face aspectos morfológicos relevantes para reconhecimento de faces. A segunda delas é uma técnica já bastante utilizada no campo de reconhecimento de padrões.

Como trabalho futuro pretende-se expandir a quantidade de pessoas representadas na base, para um número bem maior do que o utilizado na segunda bateria de testes, a fim de verificar a eficácia da abordagem de Morfologia Matemática frente a outras técnicas mais utilizadas para este tipo de problema, tendo-se uma Base de Treinamento compatível com a maioria das aplicações. Também deverá ser buscadas imagens com aspectos morfológicos ainda mais diferenciados, como uso ou não de barba, faces com alterações devido a cirurgias plásticas ou uso de acessórios como óculos escuros. Além disso, outra forma de comparação dos Espectros de Padrões deverá ser utilizada, como por exemplo, as Redes Neurais Artificiais [15].

Bibliografia

- [1] ALAA ELEYAN e HASAN DEMIREL. PCA and LDA based Neural Networks for Human Face Recognition, I-Tech Education and Publishing, cap. 6, p. 93-106, 2007, Viena, AUT.
- [2] DAO-QING DAI e HONG YAN. Wavelets and Face Recognition, I-Tech Education and Publishing, cap. 4, p.59-74, 2007, Viena, AUT.
- [3] GONZALES, R.C.; WOODS R.E. Digital Image Processing, Addison-Wesley, 1992.
- [4] VIEIRA, E.A.; MELO, R.H.C. Granulometria: Uma Aplicação para Contagem e Medição de Grãos em Imagens Digitais, Monografia apresentada ao Departamento de Ciência da Computação da Universidade Federal Fluminense, 2004.
- [5] MELLO, C.A.B. Morfologia Matemática, Notas de aula na Graduação em Engenharia da Computação - UPE, Poli, DSC, 2008.
- [6] Face Recogniton Home Page, “General Info”, Disponível em: <http://www.face-rec.org/general-info/> Último acesso em 19 de outubro, 2009.
- [7] ZHAO, W.; CHELLAPPA, R.; ROSENFELD, A.; PHILLIPS, P.J.; Face Recognition: A Literature Survey, ACM Computing Surveys, 2003, pp. 399-45.
- [8] OLIVEIRA, Y. G.. Classificação de Metodologias para Reconhecimento Automático de Faces Humanas. Dissertação de Mestrado, Universidade Federal de Minas Gerais, 1997.
- [9] Wikipedia, “Reconhecimento de Padrões”, disponível em: http://pt.wikipedia.org/wiki/Reconhecimento_de_padrões. Último acesso em 20 outubro, 2009.
- [10] LEDDA, A.; QUINTELIER, J.; SAMYN, P.; DE BAETS, P.; PHILIPS, W.; Quantitative Image Analysis with Mathematical Morphology, Proceedings of ProRISC 2003, pp. 399-406, 2003, Veldhoven, Holanda.
- [11] GONG, S.; MCKENNA, S.; PSARROU, A. Dynamic Vision: From Images to Face Recognition. Imperial College Press, UK, 2005.

- [12] CAMPOS, T.E.; Técnicas de Seleção de Características com Aplicações e Reconhecimento de Faces. Dissertação de Mestrado, Universidade de São Paulo, 2001.
- [13] R.C. Gonzalez, R.E. Woods, and S.L. Eddins, Digital Image Processing Using Matlab, Prentic Hall, 2004.
- [14] AT&T Laboratories, “The Database of Faces”, disponível em: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>. Último acesso em 16 de novembro, 2009.
- [15] Haykin, S; Neural Networks A Comprehensive Foudation, Person Education, 2005.

Apêndice A

Fundamentação Matemática

A Matriz de Covariância

Se tivermos um conjunto de dados com mais de 2 dimensões, há mais de uma covariância que pode ser calculada. Por exemplo, a partir de um conjunto de dados de 3 dimensões (dimensões x, y, z) é possível calcular $cov(x, y)$, $cov(x, z)$ e $cov(y, z)$, sendo $cov(x, y)$ a covariância entre x e y . De fato, para um conjunto de dados de n dimensões, é possível calcular $\frac{n!}{(n-2)! \cdot 2}$ valores de covariância diferentes.

Uma forma usual para se ter todos os valores de covariância possíveis entre todas as dimensões é por calcularem-se todas e inseri-las numa matriz. Então, a definição para a matriz de covariância para um conjunto de dados com n dimensões é:

$$C^{n \times n} = (C_{ij}, C_{ij} = cov(Dim_i, Dim_j)),$$

onde $C^{n \times n}$ é uma matriz com n linhas e n colunas, e Dim_x é a x -ésima dimensão. Em resumo, se temos um conjunto de dados de n dimensões, então a matriz tem n linhas e colunas (e, portanto é uma matriz quadrada) e cada entrada da matriz é o resultado do cálculo da covariância entre das dimensões separadas. Por exemplo, a entrada na linha 2, coluna 3, é o valor da covariância calculado entre a 2ª dimensão e a 3ª dimensão.

A seguir temos uma matriz de covariância para um conjunto de dados imaginário de 3 dimensões, usando as dimensões usuais x, y, z . Logo, a matriz de covariância tem 3 linhas e 3 colunas, e os valores são:

$$C = \begin{pmatrix} cov(x, x) & cov(x, y) & cov(x, z) \\ cov(y, x) & cov(y, y) & cov(y, z) \\ cov(z, x) & cov(z, y) & cov(z, z) \end{pmatrix}$$

Alguns pontos a se observar: A diagonal principal é formada pelos valores das covariâncias entre uma das dimensões e ela própria. Essas são as variâncias para essa dimensão.

Autovetores e autovalores

É uma transformação especial da forma $T: V \rightarrow W$

$$T(v) = \lambda v \quad (1)$$

Onde, λ é o autovalor (escalar) e v é autovetor (se $v \geq 0$). Como toda transformação linear pode ser escrita pela multiplicação de uma matriz por um vetor então:

$$T(v) = Av \quad (2)$$

Igualando (1) e (2), encontramos o seguinte: $Av = \lambda v$ ou $Av - \lambda v = 0$, o que resulta no sistema homogêneo:

$$(A - \lambda I)v = 0 \quad (3)$$

Onde A é $n \times n$, $v = 0$ é sempre solução (trivial). Os vetores $v \geq 0$ para os quais existe um λ que resolve a equação (3) são chamados de autovetores da matriz A e os valores de λ , que conjuntamente com v resolvem a equação são chamados de autovalores da matriz A associados aos respectivos autovetores.

Para que a equação (3) tenha solução além da trivial é necessário que o determinante da matriz dos coeficientes seja zero, ou seja, $\det(A - \lambda I) = 0$. O que resulta em um polinômio de grau n em λ , conhecido como polinômio característico. As raízes do polinômio característico são os autovalores da matriz A . Para se encontrar os autovetores basta substituir o valor do autovalor na equação original e encontrar o autovetor. O autovalor será, então, associado ao autovetor encontrado.

Na verdade, o autovetor encontrado forma uma base para o espaço de solução da equação (3), dado o respectivo autovalor. Logo, qualquer múltiplo do autovetor também é um autovetor.

Em resumo, se temos uma matriz canônica A que representa um operador linear T , temos:

- Autovalores λ de T ou de A : são as raízes da equação
$$\det(A - \lambda I) = 0.$$
- Autovetores de v de T ou de A , são as soluções da equação
$$Av = \lambda v \text{ ou } Av - \lambda v = 0$$