

SEGURANÇA DA INFORMAÇÃO – ASPECTOS TÉCNICOS PARA DESENVOLVIMENTO DE SISTEMAS DE IDENTIFICAÇÃO AUTOMATIZADA ATRAVÉS DA EIGENFACE E REDES NEURAIIS

ALEXANDRE PEREIRA LESSA¹

O presente artigo apresenta aspectos conceituais e técnicos para desenvolvimento de sistemas de controle biométrico através da validação da face. O referente estudo aborda o processo de validação através da utilização de visão computacional e Inteligência Artificial como meio de modelar o problema apresentado, através da utilização de eigeenvetores aplicado a uma rede neural.

Palavras-chaves: Biometria, Eigenfaces, Eigenvalues, Eigenvectores e Redes Neurais.

1 INTRODUÇÃO

Em segurança, especialmente segurança física, o termo controle de acesso é uma referência à prática de permitir o acesso a uma propriedade, prédio, ou sala, apenas a pessoas autorizadas. O controle físico de acesso pode ser obtido através de pessoas (um guarda, segurança ou recepcionista); através de meios mecânicos como fechaduras e chaves; ou através de outros meios tecnológicos, como sistemas baseados em cartões de acesso entre outros.

Segundo Oliveira (2009) após os atentados terroristas de 11 de Setembro a questão de segurança pública foi elevada a tópico principal nos Estados Unidos e no Mundo, assim, novas tecnologias estão sendo pesquisadas para identificação inequívoca de indivíduos. Dentre essas tecnologias as técnicas de biometria possuem papel fundamental nesse cenário.

¹ Mestre em pesquisa Operacional e Otimização de Sistemas, Especialista em Análise de Sistema. Atualmente é analista de infraestrutura da Prefeitura Municipal de Campos dos Goytacazes, docente da Fundação de Apoio à Escola Técnica do Estado do Rio de Janeiro e professor horista da Sociedade Brasileira de Instrução. E-mail aplessa@yahoo.com.br

A biometria (do grego Bios = vida, metron = medida) é definida como o uso de características biológicas em mecanismos de identificação. Podemos citar como tais características a íris, a retina, a impressão digital, a voz, o formato do rosto e a geometria da mão, entre outros.

O processo biométrico possui vantagens significativas em relação a métodos tradicionais, como a utilização de senhas e cartões de identificações. Enquanto em tais métodos os mesmos podem ser furtados e obtidos, na tentativa de burlar o sistema de segurança, perdendo assim a característica da exclusividade das informações (ponto essencial no processo de segurança), na biometria, tais problemas são extintos ou, pelo menos, amenizados, uma vez que embora nada impeça os dispositivos de identificação biométrica de serem enganados, é extremamente difícil copiar uma característica física.

Moraes (2006) define que o processo de controle baseado em biometria é um dos métodos mais eficazes de autenticar um indivíduo e possui um vasto campo de aplicação, não apenas em sistemas computacionais, como também em qualquer tipo de controle de acesso, permitindo a criação de sistemas de controle de acesso a instituições de ensino públicas baseado em biometria, como forma de aperfeiçoar a segurança física nos prédios.

Abordam-se nesse estudo, o estado da arte da biometria no mundo, a análise de sistema geral de controle de acesso e as conclusões alcançadas durante o processo de desenvolvimento.

2 DESENVOLVIMENTO

No nosso cotidiano o reconhecimento de faces é uma das ações mais corriqueiras e recorrentes que efetuamos. Tal ação apesar de ser efetuada com naturalidade por organismos vivos evoluídos apresenta uma dificuldade inerente quando tenta ser imitada em sistemas computacionais. A utilização de Visão computacional e Inteligência Artificial para modelar a solução desse problema têm despertado um grande interesse em pesquisadores espalhados pelo mundo.

Segundo Moraes (2006), os chineses foram os precursores do uso da biometria, uma vez que, no século XIV, eles utilizavam a biometria carimbando as mãos e os dedos das crianças em papel, de forma a distinguir uma criança de outra, após o nascimento. Já na Europa a biometria é muito mais recente sendo datada do final do século XIX, mais precisamente em 1890, quando um antropologista e delegado policial de Paris, Alphonse Bertillon, encontrou na biometria uma forma de identificação de criminosos, transformando-a em uma nova área de estudo.

O sistema de Bertillon de medidas do corpo incluía medidas como o diâmetro do crânio e o comprimento do braço e de pé, sendo usado nos EUA para identificar prisioneiros até os anos 20. Henry Faulds, William Herschel e o senhor Francis Galton propuseram a identificação quantitativa através da impressão digital e de medidas faciais nos anos de 1880.

O desenvolvimento de técnicas digitais do tratamento dos sinais, nos anos 60, conduziu imediatamente a um trabalho de automatização de identificação humana. Sistemas do reconhecimento de fala e da impressão digital estavam entre os primeiros a serem explorados.

O potencial para a aplicação desta tecnologia no controle de acesso seguro a elevadores, travas pessoais e transações financeiras foi desenvolvido no princípio dos anos 60. Nos anos 70, já se consideravam o desenvolvimento e a distribuição dos sistemas baseados na geometria da mão.

Os sistemas de identificação pela retina e de verificação de assinatura vieram nos anos 80, seguidos por sistemas de identificação de faces, e mais recentemente sistemas de reconhecimento da íris, nos anos 90 (WAYMAN, 2005).

Associar uma identidade a um indivíduo é chamado identificação pessoal. O problema de resolver a identificação de uma pessoa pode ser categorizado em dois tipos de problemas fundamentais distintos com complexidades inerentes diferentes, a saber: Verificação e reconhecimento (mais popularmente conhecido como a identificação).

A verificação (autenticação) refere-se ao problema de confirmação ou de negação da identidade com base em comparações das características biométricas previamente armazenadas (é um problema de verificar se o usuário é quem ele indica ser). Já no reconhecimento, há uma busca em bases de dados para identificação do indivíduo com bases nas características biométricas apresentadas. (JAIN, 2002)

Tais características biológicas são qualificadas para serem medidas biométricas quando apresentam o seguinte conjunto de propriedades fisiológicas ou comportamentais: a universalidade, significando que todas as pessoas devem possuir tal característica; a unicidade, que indica que duas pessoas não devem possuir o mesmo conjunto de atributos relacionado a essa característica; a permanência, indicando que a característica deve ser invariante no tempo, e colecionável, que indica que a característica pode ser medida quantitativamente.

Na prática, há alguns outros requisitos importantes, como: performance, que refere a realização exata da identificação com base na arquitetura disponível; aceitabilidade, que indica a necessidade de aceitação de utilização do sistema biométrico; e a segurança, que refere o índice de dificuldade em enganar o sistema por técnicas fraudulentas.

As características biológicas são armazenadas em um *template*, que consiste de um registro com as características biométricas gravadas anteriormente do usuário. Para a formação dessa base de dados, geralmente são coletadas várias vezes as características biométricas desejadas, comparando-as, então, é gerado um *template* definitivo com base na melhor amostra recolhida. (MORAES, 2006)

Com base nele, calcula-se a resposta por parte do sistema.

Assim, podemos classificar os sistemas biométricos em dois tipos específicos. Aqueles que foram projetados para testar se as amostras submetidas são de um indivíduo conhecido pelo sistema, ou aqueles que testam se as amostras submetidas são de um indivíduo não conhecido pelo sistema.

Tendo em mente o tipo específico de sistema podemos enfocar as necessidades funcionais do mesmo. Por exemplo, em um sistema de controle de acesso espera-se que o mesmo tenha a capacidade de reconhecer um indivíduo específico para validação do acesso. Nesse sistema a principal hipótese é que aquele indivíduo está identificado internamente na base de dados. Três situações podem ocorrer: A primeira seria a situação esperada, ou seja, o sistema reconhece ou não o indivíduo submetido à análise de forma correta. A segunda e a terceira situações denotariam erros. Ou o sistema não é capaz de identificar o indivíduo submetido, apesar do mesmo existir na base de dados, o que geraria um “falso negativo” ou o sistema retornaria um positivo, apesar do indivíduo não constar na base de dados, o que geraria um “falso positivo”.

Como a situação esperada é o retorno positivo da identificação do indivíduo tal aplicação é conhecida como “sistemas de identificação positiva”, aonde o esperado é que o sistema “varre” a base de dados até o ponto aonde o indivíduo é reconhecido.

De outra forma sistemas aonde a maior probabilidade de identificação do indivíduo é um negativo, como por exemplo, sistemas de identificação de terroristas em aeroportos internacionais, são conhecidos como “sistemas de identificação negativa”. Nesses sistemas o resultado esperado é um negativo, por isso sabe-se que tais sistemas deveram “varrer” a base de dados com maior freqüência, uma vez que o esperado é que não se ache nenhum registro correspondente, impondo uma maior otimização do algoritmo para melhoria de desempenho.

2.1 SISTEMAS GERAIS DE BIOMETRIA

Embora existam dispositivos que utilizem tecnologias extensamente diferentes, há um conjunto geral de métodos que podem ser definidos. Isto permite dividir um conjunto biométrico genérico em cinco subsistemas: aquisição de dados, transmissão do exemplar, extração dos sinais, definição dos atributos e decisão através da comparação, conforme a Figura 1.

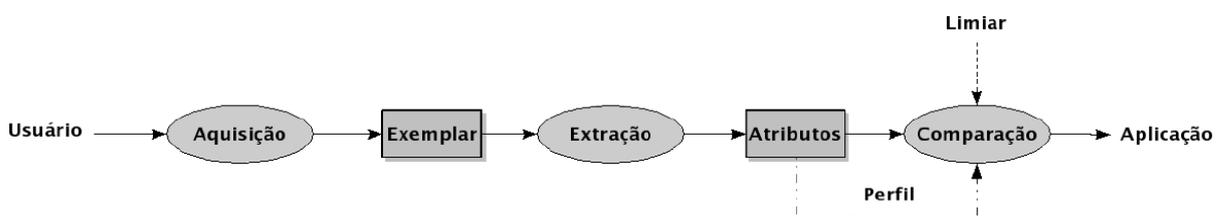


Figura 1 - Modelo Simples de Sistemas Biométricos (COSTA, 2009)

2.1.1 LEVANTAMENTO DE DADOS

O início da definição de um sistema biométrico se dá com a medida de uma característica comportamental/fisiológica. A chave de todos os sistemas é a suposição subjacente que a característica biométrica medida é distinta entre indivíduos e repetível para o mesmo indivíduo. Os problemas na medição e controle destas variações começam no subsistema de levantamento de dados. A característica do usuário deve ser apresentada a um sensor. A apresentação de toda a característica biométrica ao sensor introduz um comportamento (e, conseqüentemente) um componente psicológico a cada método biométrico.

O comportamento desses componentes pode variar extensamente entre usuários, entre aplicações, e entre testes de laboratório e o ambiente operacional. A saída do sensor, representados pelos dados de entrada modelados com base na lógica do sistema, denota uma convolução de: a medida biométrica; a maneira como a medida é apresentada; e as características técnicas do sensor.

A capacidade de repetir e a clareza da medida sofrem um impacto negativo por quaisquer mudanças em um dos fatores. Se um sistema é aberto, as características da apresentação e do sensor devem ser padronizadas para assegurar que as características biométricas coletadas com um sistema combinarão com aquelas coletadas pelo mesmo indivíduo em outro sistema.

Se o sistema for usado em uma aplicação não cooperativa, o usuário não deve poder alterar de forma intencional a biometria ou suficientemente a apresentação para impedir a combinação com os registros precedentes colhidos.

2.1.2 TRANSMISSÃO

Alguns sistemas biométricos coletam dados em uma posição, mas processam em outra, exigindo, portanto, a transmissão de dados. Se uma grande quantidade de dados é envolvida, a compressão pode ser exigida antes da transmissão

ou armazenamento para conservar espaço da largura de banda e de armazenamento.

Quando ocorre a transmissão ou armazenamento comprimido dos dados, é necessária uma etapa adicional de expansão antes de serem usados. Os processos de compressão e de expansão causam geralmente perda de qualidade no sinal restaurado. Tal perda aumenta diretamente em relação à taxa compressão empregada. A técnica de compressão usada dependerá, sobretudo, do sinal biométrico. Uma área de pesquisa interessante consiste em encontrar, para uma técnica biométrica dada, métodos da compressão com impacto mínimo no subsistema de processamento do sinal.

Segundo Wayman (2005), se um sistema é aberto, os protocolos da compressão e da transmissão devem ser padronizados de modo que cada usuário possa reconstruir o sinal original. Há padrões existentes atualmente para a compressão das impressões digitais (Quantização Wavelet escalar), das imagens faciais (JPEG), e dos dados de voz (ACELP).

2.1.3 PROCESSAMENTO DO SINAL

Após adquirir e possivelmente transmitir uma característica biométrica, devemos prepará-la para ser utilizada. O subsistema de processamento é dividido em quatro tarefas: segmentação, extração de característica, controle da qualidade, e harmonização de teste de padrão.

A segmentação é o processo de encontrar o teste biométrico de interesse dentro do sinal transmitido. Por exemplo, um sistema de reconhecimento facial deve primeiramente encontrar os limites da face ou das faces na imagem transmitida. Um sistema da verificação de fala deve encontrar a atividade de som dentro de um sinal que possa conter períodos de vozes, separando-as dos períodos sem voz. Uma vez encontrado o teste biométrico de interesse, o mesmo é extraído para um sinal maior, sendo enviado ao processo de extração de característica.

Os testes do padrão biométrico cru, mesmo depois da segmentação do sinal maior, contem distorções não repetíveis causadas pelos processos da apresentação, do sensor e da transmissão do sistema. Tais distorções, ou elementos redun-

dantes, devem ser removidos do teste do padrão biométrico ao mesmo tempo em que se preservam as qualidades que são distintivas e repetíveis.

Estas qualidades expressadas de forma matemática são conhecidas como “características”. Em um sistema de reconhecimento de fala, independente do texto, por exemplo, pode-se querer encontrar as características, tais como os relacionamentos matemáticos da freqüência nas vogais, o estado de saúde da fala, a velocidade, ou o volume e o ritmo do discurso.

A extração de característica, de uma forma geral, é uma compressão não reversível, significando que a imagem biométrica original não pode ser reconstruída a partir das características extraídas. Em alguns sistemas, a transmissão ocorre após a extração de característica para reduzir a exigência da largura de banda.

É importante que depois do processo de extração de característica (em alguns casos até mesmo antes) seja verificado se o sinal recebido do subsistema de levantamento de dados é de boa qualidade. Se as características “não fazem sentido” ou são insuficientes, conclui-se que o sinal recebido é defeituoso e uma nova amostra deve ser solicitada ao usuário enquanto o mesmo ainda se encontra no sensor. O desenvolvimento desse “processo de controle de qualidade” melhorou extremamente o desempenho de sistemas biométricos nos últimos anos. Se o controle da qualidade não puder controlar as características, cancelando-as na captura, é gerada uma “falha de registro” o que levará a erros nos resultados esperados.

A amostra, agora do tamanho muito pequeno comparado ao sinal original, será enviada ao processo da harmonização de teste de padrões para a comparação com uns ou vários “moldes” ou “modelos” previamente identificados e armazenados. O termo “molde” é utilizado para indicar características já armazenadas no sistema. As características no “molde” são do mesmo tipo que aquelas da amostra. Por exemplo, se as características da amostra são um “vetor” no sentido matemático, o “molde” armazenado também é um “vetor”.

Por sua vez o termo “modelo” é usado para indicar a construção de uma representação matemática mais complexa capaz de gerar as características de um usuário particular. Os “modelos” e as características são de diferentes tipos e estruturas matemáticas. Os “modelos” são utilizados em alguns sistemas de reconhecimento de fala e facial. Os “moldes” são usados na impressão digital, na íris, e nos sistemas de reconhecimento da geometria da mão.

O termo “registro” se refere à colocação de um “molde” ou de um “modelo” em uma base de dados (*template*). Uma vez inserido o “registro” o mesmo é associado a uma identidade, onde os dados biométricos do registro definem um “molde” ou “modelo” para o indivíduo o qual se refere.

A finalidade do processo da harmonização de teste de padrão é comparar uma amostra da característica apresentada aos dados armazenados, emitindo ao subsistema de decisão uma medida quantitativa de comparação.

Uma exceção ocorre quando há ocorrência de sistema que permite registros múltiplos. Nesse tipo de aplicação, o processo da harmonização de teste de padrão pode ser “pulado”.

Nos sistemas cooperativos de identificação de usuário ou onde há um único registro na base de dados atuais, como no exemplo do cartão magnético, o processo da harmonização de teste de padrão pode efetuar apenas uma comparação para o encontro de um único “molde” armazenado. Em todos os casos restantes, tais como a identificação em grande escala, o processo da harmonização de teste de padrão compara a amostra atual aos “moldes” ou aos “modelos” múltiplos da base de dados um de cada vez, como instruída pelo subsistema da decisão, emitindo em uma medida quantitativa a “distância” para cada comparação. No lugar de uma medida de “distância”, alguns sistemas usam medidas da “similaridade”, tais como sistemas da probabilidade máxima.

O subsistema de processamento do sinal é projetado com o objetivo de minimizar as distâncias entre modelos/moldes registrados e amostras mais novas dessas distâncias individuais, ao mesmo tempo em que deve maximizar os modelos/moldes de amostras registradas de indivíduos diferentes. No entanto, mesmo para modelos e amostras do mesmo indivíduo, raramente as distâncias serão zero, uma vez que sempre haverá alguma discrepância biométrica não repetível, devido à alteração na forma de apresentação, variação do sensor ou mesmo ruído na transmissão que permanece após o processamento.

2.1.4 ARMAZENAMENTO DE DADOS

Os moldes ou os modelos dos usuários registrados serão armazenados em uma base de dados (*template*) para a comparação pelo *matcher*² do teste de padrão às características das amostras capturadas. Para os sistemas que executam somente a harmonização “linear”, a base de dados pode ser distribuída em *smart card*, cartões de leitura ótica ou cartões magnéticos carregados por cada usuário registrado. Dependendo da política do sistema, nenhuma base de dados central precisa existir, embora nesta aplicação uma base de dados centralizada possa ser usada para detectar cartões forjados ou para reeditar cartões perdidos sem necessidade de refazer o teste de padrão biométrico.

Em uma base de dados centralizada, o sistema procura em um subespaço de dados original de dimensão N , aonde cada item representa uma face, possuindo cada uma um espaço de característica de dimensão M , que são os pixels alinhados de cada face. Para permitir a representação das características da face é necessário que N seja maior de um, levando a N combinações possíveis para pesquisa da correspondente face. Como N começa muito grande, exigências de velocidade de resposta do sistema impõem que a base de dados seja dividida em subconjuntos menores, permitindo que as comparações de amostras precisem somente ser comparadas a moldes ou modelos armazenados dentro de um subconjunto, ou mesmo permitindo a utilização de uma estrutura de dados otimizada o que leva a uma ordem vantajosa de visita para a recuperação da informação.

Estas estratégias têm o efeito de aumentar a velocidade do sistema e diminuir falsas combinações, porém tais divisões podem levar ao aumento da taxa de falsas não-combinações, ou seja, a conclusão falsa que uma face não existe no sistema, o que geraria erros no sistema. Isso significa que taxa de erro do sistema não permanece constante com o crescimento do tamanho da base de dados e em sistemas de identificação que não utilizam escala linear. Conseqüentemente, base de dados particionada/indexada representa uma decisão política complexa.

² - Matcher é um programa que determina se duas coisas (como termos, frases ou subgrafos) são equivalentes, podendo gerar uma tabela de correspondências que indica quais os elementos são equivalentes entre si.

2.1.5 DECISÃO

O subsistema de decisão é responsável por executar a busca na base de dados de acordo com uma política pré-estabelecida. Cabe a ele determinar os “positivos” ou “negativos”, ou seja, a confirmação ou não da identidade do indivíduo, de acordo com as medidas de distância ou de similaridade recebidas do subsistema de teste padrão, aceitando ou rejeitando a validação de acordo com tal política. Assim, tal política de decisão pode rejeitar a reivindicação da identidade (positivo ou negativo) de todos os usuários cujo teste do padrão não possa ser efetuado.

Para um padrão de teste existente, a política pode declarar uma combinação para qualquer distância inferior a um ponto fixo definido inicialmente, “aceitando” a reivindicação de identidade do usuário com base nessa única combinação, ou ainda, pode ser declarada a partir de um conjunto de múltiplas combinações para uma decisão de “aceite”, exigindo a existência de um grupo de distâncias inferiores.

A política de decisão empregada é uma decisão gerencial especificada nas exigências operacionais e de segurança do sistema. Geralmente, diminuir o número de falsas “não-combinações” pode levar ao aumento do número de falsas “combinações”.

Dessa forma a caracterização de uma imagem no sistema é definida através do controle de suas características médias de acordo com um calculo estatístico obtido da análise de um conjunto de faces existentes na base de dados, cabendo ao *matcher* comparar uma distância de similaridade entre as características médias do indivíduo e as características médias esperadas pelo sistema. Podemos definir que a grande problemática do reconhecimento é definir um limite superior e inferior para a oscilação das discrepâncias de tais características (do indivíduo e armazenada no sistema). Se tal margem for pequena demais o sistema gerará “falsos negativos”, se a margem for alta demais o sistema gerará “falsos positivos”. Dessa forma, um bom sistema de política deve equilibrar essa margem minimizando a ocorrência de tais erros.

Em todo caso, no teste de dispositivos biométricos, é necessário decuplar o desempenho do subsistema de tratamento dos sinais para implementação das políticas do subsistema da decisão.

2.2 IDENTIFICAÇÃO DE CARACTERÍSTICAS FACIAIS

A face é uma das biometrias mais aceitáveis, pois é um dos métodos mais comuns de identificação utilizados pelos seres humanos em suas interações visuais, além de possuir um método de aquisição de imagens não intrusivo. Dois métodos primários de identificação baseados no reconhecimento da face são: Método de aproximação de transformação, onde o universo do domínio da face é representado usando um conjunto de vetores básicos ortonormais. E o outro método é aproximação baseada em atributos. Tais atributos (nariz, os olhos, etc) são extraídos da imagem da face e a não-variação de propriedades geométricas entre as características do marco na face é usada para o reconhecimento dessas características. (JAIN, 2002)

Atualmente, a base de vetores mais populares são os *eigenfaces*: Cada *eigenface* é derivado da análise de co-variância do conjunto de imagens da face: duas faces são consideradas idênticas se são suficientemente próximas no espaço de características do *eigenface*. (ROSSA, 2005)

A análise das alterações faciais é essencial em aplicações desprovidas de sistemas de autenticações concomitantes. Torna-se desafiador desenvolver técnicas de reconhecimento faciais capazes de tolerar os efeitos do envelhecimento, de expressões faciais, de variações ligeiras no ambiente da imagem latente e de variações na pose da face no que diz respeito à câmera.

2.2.1 TÉCNICAS DE REPRESENTAÇÃO E APLICAÇÃO NO RECONHECIMENTO DE FACES

Os principais esquemas de reconhecimento facial requerem, além de representações de características de baixa-dimensão, a descrição realçada de tais características para futura classificação. Os métodos de representação começam geralmente com um procedimento para redução da dimensionalidade. Essa redução é necessária por que a elevada dimensão do espaço original faz a avaliação estatística se tornar muito difícil, e às vezes impossível, devido ao fato de que o espaço dimensional é, em sua maioria, vazio. Os métodos de discriminação tentam frequentemente conseguir a separação elevada entre testes de padrões diferentes.

A Análise de Componentes Principais³ (PCA – do inglês Principal Component Analysis) é geralmente usada para derivar representações baixo-dimensionais de imagens entradas. Especificamente, a PCA a partir de sua base da projeção ortogonal conduz diretamente à redução da extensibilidade e possivelmente à melhora na seleção de características.

De uma forma resumida o PCA é um método que tem por finalidade básica a redução de dados a partir de combinações lineares das variáveis originais levando a uma decomposição de uma matriz de dados de entradas extensa em um conjunto matrizes de dimensões inferiores, por isso o PCA é considerado um algoritmo de descorrelação de sinal, ou seja, a eliminação de parte da informação redundante em cada dimensão da imagem.

Sendo as imagens de faces altamente redundantes, uma vez que cada pixel é altamente correlacionado com outros pixels a técnica do PCA é empregada para a redução da matriz de análise para uma dimensão inferior.

Uma imagem é traduzida em um vetor unidimensional, aonde cada posição representa o valor de um pixel em uma imagem de dimensões $n \times m$. Cada vetor

^{3 3} Segundo a Opper (2022) A PCA é uma técnica de análise multivariada que pode ser usada para analisar inter-relações entre um grande número de variáveis e explicar essas variáveis em termos de suas dimensões inerentes (Componentes), seu objetivo é encontrar um meio de condensar a informação contida em várias variáveis originais em um conjunto menor de variáveis estatísticas (componentes) com uma perda mínima de informação.

é inserido em uma matriz que representa o conjunto de imagens do banco de dados (*template*), representando assim um conjunto de imagens de comparações, conforme Figura 2.

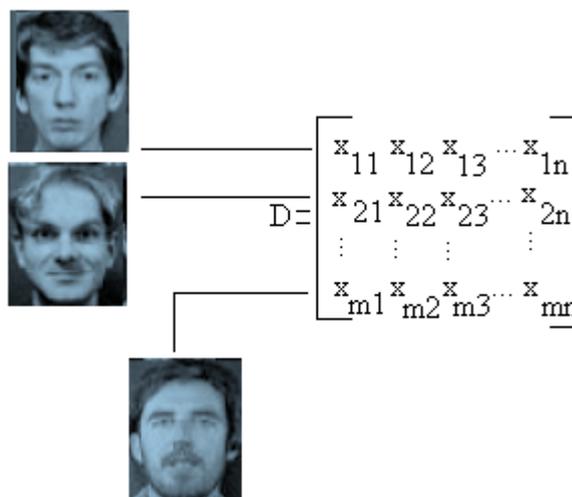


Figura 2 - Representação das figuras na Matriz

Dessa forma definimos que o objetivo da técnica do PCA é encontrar uma transformação mais representativa e geralmente mais compacta das imagens. A mesma transforma um vetor aleatório x em outro vetor y através da projeção de suas características de maior variância – o que denota os componentes principais. Estes componentes são individualmente responsáveis pela variância das observações, e neste sentido, representam a imagem com maior clareza.

Geralmente grande parte da variância dos dados é explicada por um número reduzido de componentes, sendo possível descartarem os restantes sem grande perda de informação. Assim a técnica gera uma matriz de covariância baseada na matriz de imagens geradas para detecção e posterior redução de tais redundâncias.

Aplicando-se a técnica do PCA ao reconhecimento de face (Principalmente com o método Eigenface), onde os eigenfaces correspondem aos conhecidos “vetores do eigen” associados com os valores da própria matriz de co-variância da face, é possível definir um espaço de características, ou “o espaço da face”, que reduz drasticamente a dimensionalidade do espaço original, permitindo que o reconhecimento de face seja realizado em um espaço reduzido (WAYMAN, 2005).

Dessa forma podemos definir que o método do eingenface através da técnica de PCA trata da seleção de características, atuando como um processo no qual um espaço de dados é transformado em um espaço de características, de menor dimensão, mas que ainda retenha a maior parte da informação intrínseca dos dados (redução de dimensionalidade).

Por outro lado os métodos de seleção de características tratam exatamente da escolha, dentre todos os atributos da base de dados, daqueles mais relevantes do ponto de vista da informação.

Ao utilizar-se uma rede neural a regra de Bayes permite inferir um conjunto ótimo de pesos a partir das densidades de probabilidade dos dados e dos pesos sinápticos da rede.

Dessa forma entendemos que o conjunto de características obtidas de uma matriz de imagem reduzida possui uma densidade de probabilidade condicionada à classe de padrão. Assim um vetor padrão D pertencente a uma classe w_i sendo entendida como uma extração ao acaso de uma função de probabilidade classe-condicional $p(D|w_i)$.

Conceituamos que a inferência estatística bayesiana refere-se à obtenção de conclusões sobre quantidades não observadas θ a partir de dados observados y . A mesma aborda o problema definindo a probabilidade de uma forma subjetiva, como uma medida da plausibilidade de uma proposição, condicional no conhecimento do observador.

A incerteza em relação à θ pode assumir diferentes graus, os quais se representam através de modelos probabilísticos para θ . Assim, tanto as quantidades observáveis, quanto os parâmetros do modelo estatístico são considerados quantidades aleatórias, o que constitui uma diferencia fundamental da abordagem bayesiana com a clássica, que considera o parâmetro como uma quantidade fixa e desconhecida, à qual nos aproximamos no processo de inferência. (COSTA, 2004)

De forma prática ela inicia com a especificação de um modelo probabilístico completo, através da distribuição conjunta das quantidades observáveis e não observáveis do problema.

A informação disponível sobre θ , resumida na densidade de probabilidade $p(\theta)$, é aumentada com a observação de uma quantidade aleatória y que se relaciona com θ . O teorema de Bayes fornece a regra de atualização desta informação,

sendo a relação entre as problemáticas condicionais definidas pelo teorema (Equação 2.1):

$$P(B/A, C) = \frac{P(A/B, C) \cdot P(B/C)}{P(A/C)} \quad (2.1)$$

Esse teorema pode ser interpretado como um modelo do processo de aprendizado sobre θ :

$$P(\phi/y, M) = \frac{P(y/\phi, M) \cdot P(\phi/M)}{P(y/M)} \quad (2.2)$$

Onde:

$p(y | \theta, M)$: Verossimilhança ou distribuição de probabilidade dos dados y condicionais nos parâmetros θ e nas hipóteses do modelo M , que fornece a plausibilidade de cada um dos possíveis valores do parâmetro θ ;

$p(\theta | M)$: Distribuição a priori, que expressa o conhecimento sobre os parâmetros antes de examinar os dados;

$p(y | M)$: Evidencia, ou constante de normalização;

$p(\theta | y, M)$: Distribuição a posteriori que expressa o conhecimento sobre o parâmetro após examinar os dados.

A distribuição de probabilidade a priori $p(\theta)$ reflete o conhecimento prévio sobre os parâmetros, anterior à observação dos dados (COSTA, 2004).

Segundo Marin (2006) a regra de decisão de Bayes "ótima" para a minimização do risco pode ser definido como:

$$R(w_i / D) = \sum_{j=1}^c (w_i, w_j) \cdot P(w_j / D) \quad (2.3)$$

Aonde:

c indica o conjunto de categorias e D o conjunto de características a serem analisados.

A equação acima (equação 2.3) determina o padrão de entrada D para a classe w_i para o qual o risco condicional é mínimo, onde $L(w_i, w_j)$ define a perda da função implicada sobre w_i determinada quando a classe verdadeira é w_j e $P(w_j | D)$ é a probabilidade posterior.

No caso de função perda 0/1 é definida pela distancia de Hamming. Aonde a distância entre as observações é definida pela soma ponderada dos módulos das diferenças dos valores das várias componentes do vetor. Tal cálculo possui a vanta-

gem de requisitar uma quantidade de cálculos numéricos menores que outras abordagens convencionais, além de poder também tratar vetores cujas componentes tem valores inteiros. De fato se os coeficientes de ponderação forem inteiros o valor da distância também o será. Assim ta equação é apresentada na Equação 2.4:

$$d_h = \sum_1^n \alpha_i (s_i - u_i) \quad (2.4)$$

Aonde s_i e u_i são vetores compostos com componentes da face.

Observe que o risco condicional torna-se a probabilidade condicional de classificação errônea definindo assim pela Equação 2.5:

$$L(w_i, w_j) = \begin{cases} 0, & i = j \\ 1, & i \neq j \end{cases} \quad (2.5)$$

Dessa forma para a escolha da função perda, a regra de decisão Bayes pode ser simplificada como na equação 2.6:

$$P(w_i | D) > P(w_j | D), \text{ para todo } j \neq i \quad (2.6)$$

A equação determina que o padrão de entrada D para a classe w_i .

Podem-se utilizar várias estratégias para projetar um classificador de reconhecimento de padrão estatístico, dependendo da espécie de informação disponível sobre densidades de classe-condicional. Entretanto, se todas as densidades de classe-condicional são completamente especificadas, então a regra de decisão Bayes ótima pode ser usada para projetar o classificador.

Para se detectar a localização de uma face são utilizados dois estágios, o primeiro para codificar a forma e o segundo a textura. Assim, inicialmente é codificada a face utilizando pontos importantes de limite, e, uma vez que estes pontos de controle são encontrados, estes são alinhados usando um processo de transformação da translação, escala e rotação. A média entre estes pontos de controle alinhados define a forma média da face. A próxima etapa é a triangulação das características detectadas na face relativas à forma média representada. A primeira fase define a forma, enquanto a segunda define a textura.

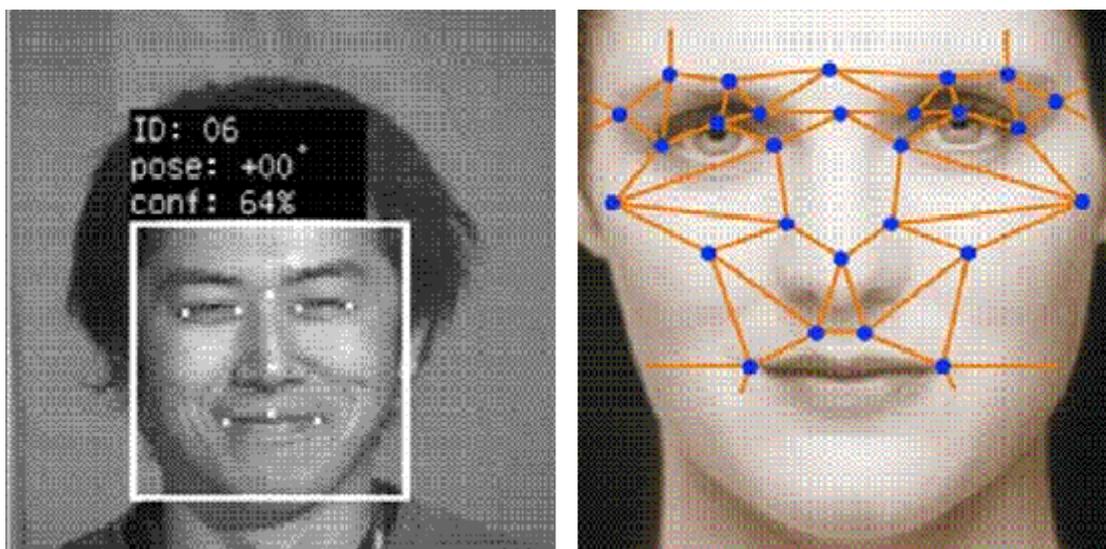


Figura 3 - A identificação e medição dos pontos baseados nas características das faces. (MORAES, 2006)

Segundo Machado (2009) o PCA é a técnica provavelmente mais utilizada para se projetar um subespaço para reconhecimento facial. A partir de um grupo de imagens treinadas, os vetores básicos do PCA são computados, partindo sempre da média computada da imagem na matriz D (Equação 2.6) e subtraída das imagens de treinamento, gerando um grupo de amostras.

Penharbel (2005) define o algoritmo de PCA da seguinte maneira:

Primeiramente as imagens são organizadas em uma matriz D onde a primeira linha será formada pelas componentes da primeira amostra, a segunda a linha formada pelas componentes da segunda amostra e a N-ésima linha será formada pelas componentes da N-ésima amostra, como na matriz abaixo:

$$D = \begin{bmatrix} c1a1 & c2a1 & cMa1 \\ c1a2 & c2a2 & cMa2 \\ \dots & \dots & \dots \\ c1aN & c2aN & cMaN \end{bmatrix} \quad (2.6)$$

Ocorre assim a transposição das imagens para o espaço de imagens, construindo uma matriz D, de tamanho (MxN), sendo M o produto da resolução horizontal (h) vezes a resolução vertical (w) da imagem original e N o número de imagens a serem utilizadas durante a fase de treinamento. Conforme definido anteriormente, o uso de eigenfaces considera esse espaço de imagens como um espaço vetorial (h x w - dimensional), e entende cada uma das linhas como um ponto neste espaço.

Após é calculado um vetor média E (Equação 2.7), formado pelas médias de cada coluna:

$$E = \mu_1, \mu_2, \dots, \mu_M \quad (2.7)$$

Achasse a dispersão dos valores de média através da subtração de cada item de cada coluna M da matriz D da média μ_M correspondente à coluna da qual o item pertença.

A partir dessa matriz é calculada uma matriz de co-variância (equação 2.8), da onde os eigenfaces devem ser extraídos.

$$COVD = \frac{1}{N-1} DD^T \quad (2.8)$$

A matriz de co-variância é obtida pela multiplicação da matriz D por sua transposta, vezes a multiplicação pelo valor real obtido pela divisão de 1 sobre N (número de elementos de D) menos 1, gerando assim uma matriz quadrada simétrica, o que é essencial para o cálculo dos auto vetores e autovalores.

O uso dessa técnica reduz os cálculos na hora de estabelecermos o reconhecimento das imagens envolvidas, através da projeção dos vetores (faces) em subespaços vetoriais, daí o uso de autovalores e auto vetores (eigenfaces). Esses auto vetores refletem em um conjunto de vetores com direções que representam a maioria das direções dos demais vetores já presentes neste espaço. Para o reconhecimento de faces isto se faz importante, na medida em que os auto vetores refletirão os componentes das imagens que são mais importantes ao reconhecimento, ou seja, aqueles que mais se fazem presentes em todas as imagens.

Dessa forma a partir da matriz de co-variância são calculados os auto vetores e autovalores gerando um vetor e uma matriz correspondente de saída.

Os autovalores correspondem efetivamente à variância das amostras no espaço de projeção e podem ser interpretados como uma estimativa da quantidade de informação descrita por cada dimensão. Assim, o que o PCA faz é tentar selecionar as dimensões com maior informação, assumindo que estas são as características mais relevantes para a representação do espaço.

Ao final ordene os auto vetores pela ordem crescente dos autovalores correspondentes.

Essa ordenação deve-se ao fato que os autovalores refletem a importância dos auto vetores, ou seja, os auto vetores com maiores autovalores refletirão em

componentes das imagens que são mais importantes do que outros com autovalores menores, permitindo inclusive que auto vetores com autovalores nulos possam ser descartados. O que tornasse importante também para a redução de dimensionalidade na hora da realização das computações em busca do reconhecimento das faces.

2.2.2 EIGENFACES

Segundo Turk (2007) muitos dos trabalhos anteriores de reconhecimento automatizado de faces ignoraram que a introdução de aspectos de estímulo do rosto⁴ são importante para sua identificação, supondo apenas que as medidas predefinidas eram relevantes e suficientes para uma identificação positiva do indivíduo, o que sugere que uma aproximação da teoria de informação da codificação e da decodificação.

Assim para que imagens possam ser entendidas, devem satisfazer as informações associadas comumente às características da face; o que enfatiza a idéia de "local e global" das características. Ou seja, um nariz possui uma dimensão relativa em relação à face e uma posição associada, que apesar de poder sofrer pequenas variações, ainda servem como características relacionadas a nossa noção intuitiva de representação da face, assim como os olhos, contornos, e o cabelo. Mesmo a inserção de características externas (óculos) ou não previstas (por exemplo, um sorriso ou uma careta) não deve influenciar na posição relativa de tais características no modelo de representação escolhida.

A intenção deste processo na teoria da informação é extrair a informação relevante em uma imagem da face, para codificá-la de forma tão eficiente quanto possível e comparar a face que foi codificada com uma base de dados de modelos similares codificados. Uma forma simples de extrair a informação contida em uma imagem de uma face é capturar de algum modo a variação em uma coleção de imagens, independente de qualquer julgamento de características, e usar esta informação para codificar e comparar imagens individuais da face.

⁴ Utilização de óculos e aspectos de fisionomia geral, como sorriso, tristeza, etc

Matematicamente, o que se deseja encontrar é as componentes principais de distribuição da face, ou os *eigenvectors*. Eles definem uma matriz de vetores calculados pela co-variância do conjunto de imagens da face, ou seja o conjunto de características de variação entre as imagens da face. Cada posição da imagem contribui para o cálculo de cada *eigenvector*, desse modo se define um tipo de face espectral que nós chamemos um *eigenface*. A *eigenface* é um “rosto” que surge da interação das variações obtidas das médias das “faces” selecionadas na base de dados, o que permite impor uma posição relativa média de cada característica no rosto. Essa seria a fase de estimação dos espaços da face.

Cada imagem da face no treinamento pode ser representada exatamente como uma combinação linear dos *eigenfaces*, sendo o número possível igual ao número de imagens das faces de treinamento. Entretanto as faces somente se aproximam usando as “melhores *eigenfaces*”, ou seja, aquelas que possuem os mais amplos *eigenvalues* e que esclarecem conseqüentemente a maioria das variações dentro da definição das imagens da face.

A idéia de usar *eigenfaces* é motivada por uma técnica desenvolvida por Sirovich e por Kirby para representação eficiente de imagens das faces usando a análise de componente principal (PCA). Defini-se que uma coleção de imagens da face pode ser reconstruída aproximadamente armazenando uma coleção pequena dos pesos para cada face e pequenos padrões de imagens, ou seja, pela implementação dos auto vetores e autovalores da matriz de imagem para treinamento do sistema. (TURK, 2007)

Dessa forma um grande número de imagens da face pode ser reconstruído por somas tornadas de uma pequena coleção de características das imagens, permitindo uma maneira eficaz de aprender e reconhecê-las através da definição das características conhecidas e comparadas aos pesos da característica necessárias a reconstrução aproximada dos pesos associados aos indivíduos conhecidos.

De todos os modelos de reconhecimento de face o modelo *eigenfaces* é o que apresenta a melhor precisão e confiabilidade. O modelo de reconhecimento de Eigen é baseado em dois estágios:

- Etapa de treinamento, na qual um número N de imagens são coletadas para se criar um modelo de imagens preto e branco M dimensional;

- Etapa operacional, em que cada imagem é transformada em um modelo M dimensional de Eigen (Vetores); este vetor é usado para a criação do *template* e serve para classificar e estabelecer a identidade do indivíduo (Figura 4). (MORAES, 2006)



Figura 4 - À esquerda temos três imagens originais e a esquerda a projeção delas no espaço de eigenfaces. (LOPES, 2005)

O treinamento das imagens da face é definido pelo conjunto:

$$\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_n \quad (2.9)$$

Com base nele criamos uma matriz Ψ contendo n cópias da média das imagens contidas em Γ definido pela equação 2.10.

$$\Psi_{[M \times N]} = \frac{1}{n} \sum_{i=1}^n \Gamma_{i, [M \times 1]} \quad (2.10)$$

Dessa forma é possível calcular os autovalores e auto vetores. Cada face Γ_i difere da média pelo vetor gerado pela Equação 2.11:

$$\Phi_{i, [M \times n]} = \Gamma_{i, [M \times n]} - \Psi_{[M \times n]} \quad (2.11)$$

Aonde Φ_i representa a diferenças entre as faces, sendo demonstrada pela Equação 2.12 (KÖRTING, 2004).

$$A_{[M \times N]} = [\Phi_1, \Phi_2, \Phi_3, \dots, \Phi_n] \quad (2.12)$$

A esse grande conjunto de vetores é então efetuado à análise componente principal, que procura um conjunto de vetores ortonormais u_n , e seus associados eigenvalues λ_k , ou seja, aqueles que descrevem melhor a distribuição dos dados.

Os vetores u_n e os valores escalares λ_k são eigenvectores e eigenvalues, respectivamente da matriz de co-variância.

A matriz de co-variância das imagens de treinamento pode então ser montada, extraídos desta matriz os eigenfaces através da relação demonstrada na Equação 2.13 (TURK, 2007).

$$C_{[M \times n]} = A_{[M \times n]} A_{[M \times n]}^T \quad (2.13)$$

Devido à complexidade computacional para encontrar os auto vetores de uma matriz $[M \times n]$ em tamanhos de imagens típicos, há uma necessidade de simplificar os cálculos adotados. Em uma matriz $L(M \times N)$, demonstrada na Equação 2.14.

$$L_{[M \times n]} = A_{[M \times n]}^T A_{[M \times n]} \quad (2.14)$$

Os primeiros n auto vetores de C podem ser expressos como uma combinação linear entre os auto vetores de L (denotados por V) e as imagens contidas em A , permitindo definir a matriz $U(M \times n)$ mostrada na Equação 2.15.

$$U_{[M \times n]} = A_{[M \times n]} V_{[M \times n]} \quad (2.15)$$

Gerasse assim os eigenfaces.

Considerando cada linha da matriz U como uma imagem, perceberemos que estas apresentam similaridades com faces humanas, ressaltando as características principais do banco de imagens de faces utilizado.

Definisse dois estágios para o reconhecimento de faces utilizando eigenfaces. O primeiro estágio seria a fase de treinamento, aonde é gerada uma matriz apresentada na Equação 2.16 com base na imagem da face de cada indivíduo.

$$\Omega_{[M \times n]} = U'_{[M \times n]} (\Gamma_{[M \times n]} \Psi_{[M \times n]}) \quad (2.16)$$

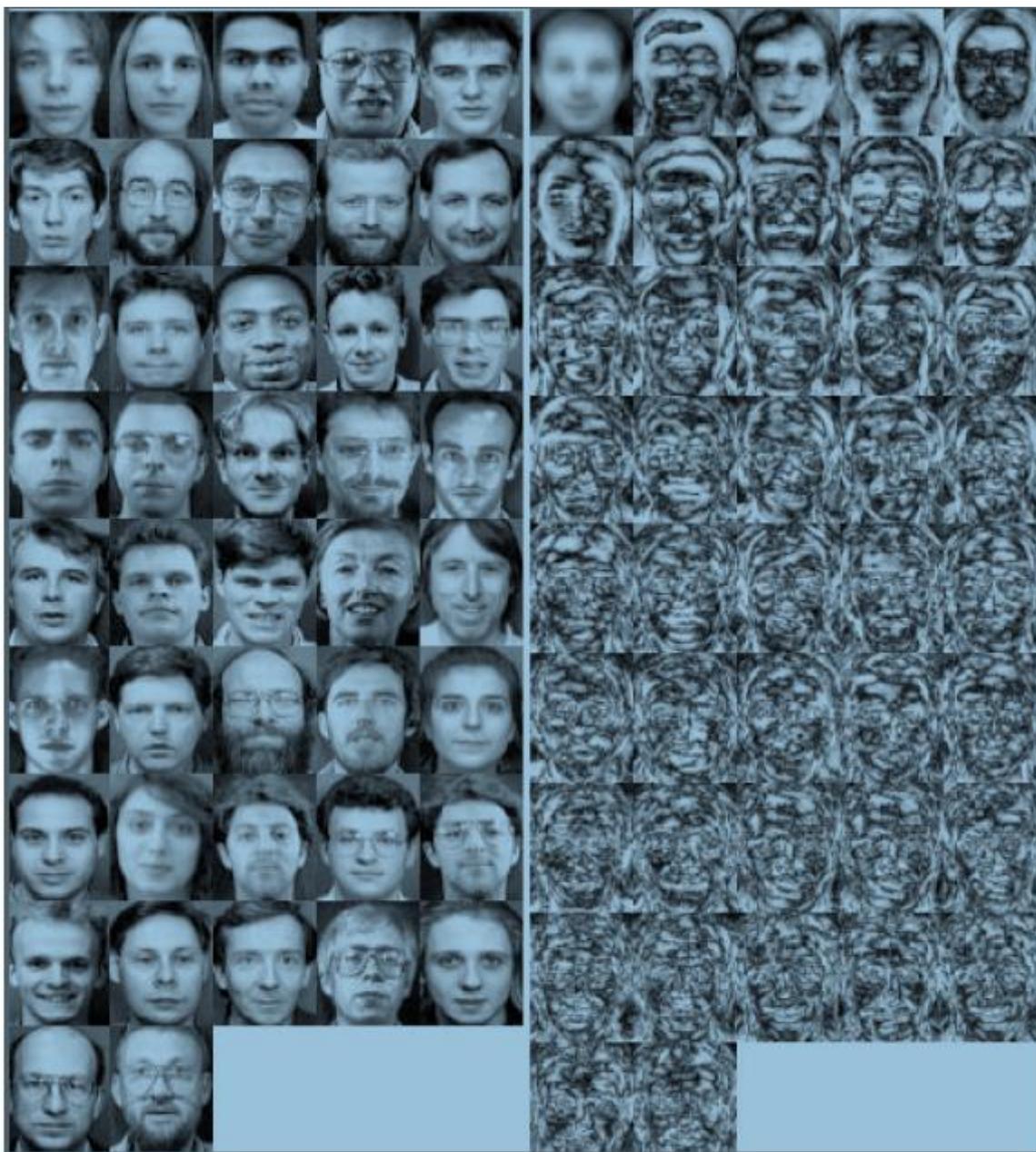


Figura 5 - À esquerda: banco de faces para criação das eigenfaces. À direita: primeira imagem é a imagem média, as restantes são as eigenfaces. (PENHARBEL,2005)

Observe que as imagens apresentadas na Figura 5 demonstram figuras

que lembram faces humanas, ou seja, imagens contendo os componentes principais das faces humanas. Dessa forma podemos definir que os eigenfaces com menores autovalores apresentam imagens sem muitas características interessantes, por isso muitas vezes são descartadas.

O segundo estágio é a fase de reconhecimento das faces.

Com base na matriz Ω é realizado as comparações para o reconhecimento das imagens. Sua dimensão muito menor do que a das imagens de entrada, denotando o principal objetivo de utilização dos eigenfaces.

Um valor de limiar Θ_c é calculado a partir da fórmula abaixo (Equação 2.17):

$$\Theta_c = \frac{1}{2} \max_{n, j}^n (\Omega_j - \Omega_k) \quad (2.17)$$

Utilizando-a para a posterior classificação das imagens, representando a maior distância aceitável entre imagens de teste e as imagens de treinamento. (PENHARBEL, 2005)

Com base em uma imagem de prova Γ , o sistema deve ser capaz de discernir se a imagem representa ou não uma face humana e em caso positivo se a mesma representa alguma pessoa cuja face está no banco de treinamento ou não. Para iniciarmos o processo, projetamos Γ no espaço de imagens através da Equação 2.18.

$$\Omega_{[nx1]} = U_{[nxM]}^T (\Gamma_{[Mx1]} - \Psi_{1,[Mx1]}) \quad (2.18)$$

É também construída através da Equação 2.19.

$$\Gamma_{f,[Mxn]} = U_{[Mxn]} \Omega_{[nx1]} + \Psi_{1,[Mx1]} \quad (2.19)$$

Para uso no cálculo posterior, a imagem de prova, guardando-a em Γ_f . Essa será utilizada para que o sistema possa realizar a distinção entre imagens de prova contendo faces humanas e imagens quaisquer. Para isso calculamos a distância (vetorial) entre a imagem de prova e sua reconstrução ε , através da Equação 2.20.

$$\varepsilon = \|\Gamma - \Gamma_f\| \quad (2.20)$$

A matriz ε define um “mapa da face”.

Além disso, precisamos saber a distância entre a imagem de prova e todas as imagens do banco de faces. Calcula-se tal distância através da equação 2.21.

$$\varepsilon_k = \|\Omega - \Omega_k\|; k = 1, \dots, n \quad (2.21)$$

Armazenamos os resultados obtidos na matriz ε_f .

Uma vez calculado todos estes valores, podemos realizar a classificação da imagem de prova, utilizando a lógica a seguir:

Se $\varepsilon \geq \theta_c$, ENTÃO a imagem não representa uma face humana;

Se $\varepsilon < \theta_c$ E $\varepsilon_k \geq \theta_c$, ENTÃO a imagem representa uma face humana, mas desconhecida pelo banco de faces;

Se $\varepsilon < \theta_c$ E $\min \varepsilon_k < \theta_c$, ENTÃO a imagem representa a face do indivíduo contido no índice k do banco de faces.

Dessa forma se o valor ε estiver abaixo de certo limite θ_c , considere a imagem Γ como sendo uma imagem de face, caso contrário considere a imagem Γ como sendo uma imagem de não face.

2.3 REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais RNA, ou redes conexionistas, são técnicas computacionais baseadas em um modelo matemático fundamentado na estrutura neural de organismos inteligentes. Essa técnica tenta emular a forma do cérebro humano processar informações através de unidades de processamentos que trabalham de maneira distribuída e paralela. (OSEAS, 2004)

Essas unidades (comumente conhecida como nós, neurônios ou células) são responsáveis pelos processamentos local dos dados obtidos através de entradas recebidas por suas conexões, que nada mais são do que canais de comunicação (arcos unidirecionais), os quais possuem atrelado um determinado peso. Cada célula possui uma única saída (axônio), a qual pode se ramificar em muitas ligações colaterais tendo o mesmo sinal de saída. Um conjunto de várias unidades de pro-

cessamento forma uma rede neural artificial, onde o seu comportamento inteligente é obtido através das interações entre as unidades de processamento da rede (MARCELO, 1996).

Podemos então caracterizar as RNA como modelos adaptativos treináveis que representam domínios complexos (não lineares) apresentando um sistema robusto para tratamento de informação. Essa robustez vem justamente da capacidade de generalização diante de informação incompleta, uma vez são capazes de efetuar um armazenamento associativo de informações. Apesar dessa característica a forma de trabalho, utilização de paralelismo no processamento leva as RNA a terem grande rapidez de processamento.

O neurônio artificial é o componente básico das RNAs. Os sinais de entradas, representada por dendritos, são multiplicados por um número, ou peso, indicativo da sua influência na saída da unidade. Executa-se então uma soma ponderada para se calcular um nível de atividade, quando esse nível atinge um limite ou função de ativação $f(.)$ é produzido um sinal de saída.

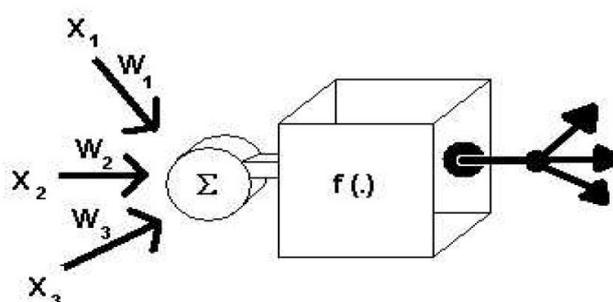


Figura 6 - Representação do Neurônio Artificial.

A Figura 6 representa um neurônio artificial aonde os sinais de entradas X_1 , X_2 e X_3 possuem pesos W_1 , W_2 e W_3 respectivamente. Seu nível de atividade é definido pelo conjunto de conexões existentes, onde aquelas que apresentam pesos positivos (excitatórias) reforçam a ativação do neurônio enquanto as de pesos negativos (inibitórias) indicam retardamento na ativação do neurônio (MARCELO, 1996).

Podemos observar na figura acima que é gerado um nível de atividade α através da Equação 2.22:

$$\alpha = \sum_{n=1}^3 X_n W_n \quad (2.22)$$

Após a obtenção do nível de atividade cabe a função de ativação $f(\cdot)$ definir o sinal de saída a ser produzido. Essa função, também chamada de função limiar ou função degrau, é a regra que define o valor da ativação de um neurônio no instante t , dependendo dos valores de ativação que podem ser discretos ($\dots, -1, 0, 1, \dots$) ou contínuos (podendo assumir valores entre 1 e 0).

Geralmente existem quatro funções de ativações que podem ser utilizadas: linear (Figura 7 A), rampa (Figura 7 B), degrau (Figura 7 C) e sigmóide (Figura 7 D) (MARCELO, 1996).

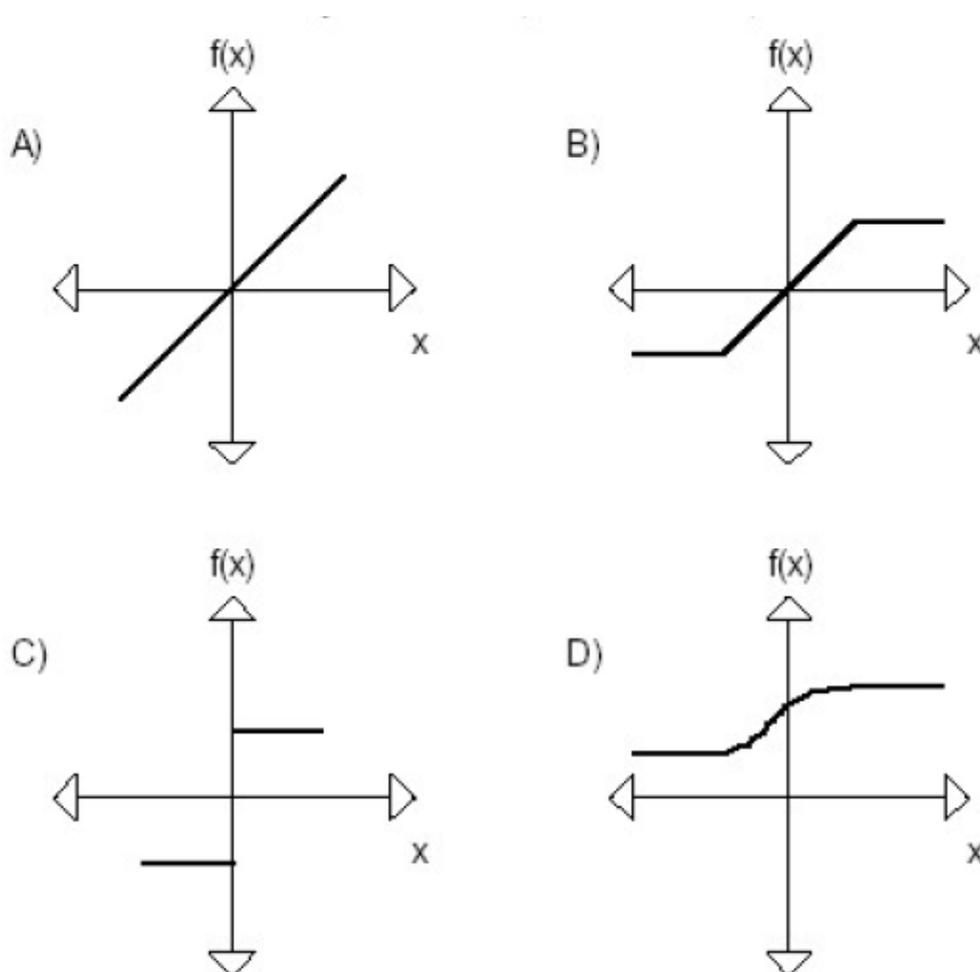


Figura 7 - Funções de ativação utilizadas

Oséas (2004) cita que a função de ativação mais comumente utilizada (99% dos casos) é a função linear definida pela equação abaixo definida (2.23).

$$net_k = A(b, w, x) = b_k + \sum (w_{kj} x_j) \quad (2.23)$$

Aonde net o estado de ativação do k-ésimo neurônio; b_k o bias do k-ésimo neurônio; w_{kj} define a j-ésima sinapse conectada ao k-ésimo neurônio.

Outra função importante de se trabalhar é a função Sigmóide [Figura 7 D], essa função permite que a entrada assuma valores contínuos no intervalo $-\infty$ até $+\infty$ os comprimindo-os para valores discretos no intervalo $[0, +1]$ (OSEAS, 2004).

A equação define matematicamente a função Sigmóide de forma logística.

$$\varphi(net) = \frac{1}{1 + e^{-x}} \quad (2.24)$$

Aonde $-x$ determina a inclinação da reta da função Sigmóide.

Podemos definir topologicamente as RNAs de acordo com as funções de suas células. Cria-se assim uma arquitetura em camadas, geralmente separadas em três grupos: Camada de Entrada, utilizadas para apresentar os dados à rede; camadas Intermediárias ou Escondidas, onde é executada a maior parte do processamento, através das conexões ponderadas; camada de Saída apresenta os resultados finais da rede (MARCELO, 1996)

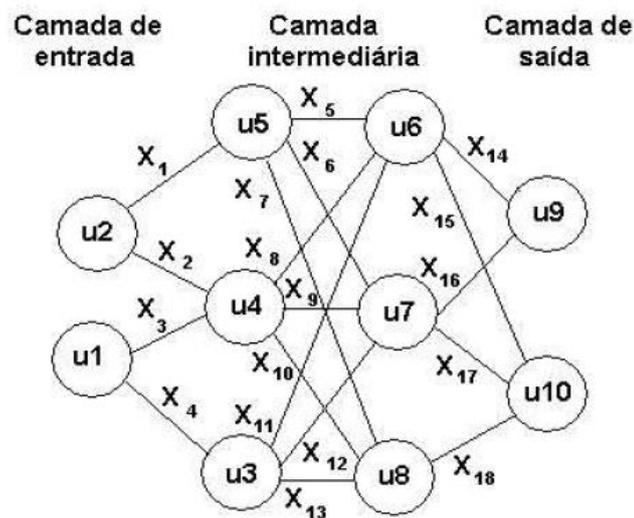


Figura 8 - Representação das camadas na RNA

A Figura 8 representa as camadas da RNA, aonde os neurônios u1 e u2 são considerados entradas da rede e os neurônios u9 e u10 saídas da rede.

Contudo o aspecto mais importante da RNA envolve a habilidade de aprender, o que em essência envolve uma mudança associada a aperfeiçoamento. Esse processo é obtido através de treinamento, que em suma altera os pesos das conexões de forma a atingir a solução geral para uma classe de problemas. Para se obter o aprendizado é utilizado um algoritmo de aprendizado, esse algoritmo define a forma de controle utilizado durante o aprendizado. No aprendizado supervisionado são utilizados rótulos externos que definem a classe correta para o padrão de entrada; já no não-supervisionado não é utilizado nenhum rótulo, uma vez que estes não são conhecidos.

Segundo Marcelo (1996) existe várias técnicas de aprendizado possíveis de serem utilizadas, como por exemplo: Correção de Erros (ajusta os pesos das conexões entre nós, na proporção da diferença entre os valores desejados e computados de cada neurônio da camada de saída); Reforço (os pesos são recompensados quando o sistema executa ações apropriadas se não são punidos).

2.3.1 RECONHECIMENTO DE FACES COM RNA

Muitos pesquisadores de detecção de faces defendem a idéia que as imagens faciais podem ser caracterizadas diretamente nos termos de intensidades do pixel. Estas imagens podem ser caracterizadas por modelos probabilísticos do conjunto de imagens da face, ou implicitamente pelas redes neurais ou mesmo por outros mecanismos. Os parâmetros para estes modelos são ajustados automaticamente a partir de imagens do exemplo ou à mão. Alguns autores utilizam-se do método de extrair características e de aplicar medidas manualmente ou automaticamente geradas para avaliar estas características.

Treinar uma rede neural para a tarefa de detecção de face é um processo desafiador por causa da dificuldade de caracterizar imagens prototípicas “não faces”, ou seja, imagens que não representam faces. Ao contrário do reconhecimento de faces, em que as classes a serem discriminadas somente pertencem a faces diferentes, no processo de detecção da face existem duas classes a serem consideradas: “imagens contendo faces” e as “imagens que não contêm faces”.

É relativamente fácil começar uma amostra, ou seja, capturar imagens gerais para criação de uma matriz de imagens representativa de imagens que somente contêm as faces. Mas para começar uma amostra representativa daqueles que não o fazem, ou seja, matrizes de imagens misturadas de faces e não faces é necessário um processo muito mais complexo. Uma das soluções para evitar o problema de usar um conjunto enorme no treinamento para “não faces” é adicionar seletivamente imagens ao treinamento ajustando à medida que se progride. Este método reduz o tamanho do treinamento necessário para o ajuste. A utilização da arbitragem entre múltiplas redes e heurísticas, limpa os resultados melhorando significativamente a exatidão do detector.

Rowley (1998) define que um primeiro componente para implementação de um sistema é um filtro que receba como a entrada uma região do pixel (por exemplo, 20x20 da imagem), e gere uma saída que varia de 1 a -1, significando a presença ou a ausência de uma face, respectivamente.

Para detectar as faces em qualquer lugar na entrada, o filtro é aplicado em cada posição na imagem. Para detecção de faces maiores do que o tamanho da janela, a imagem de entrada é reduzida repetidamente de tamanho (processo de *subsampling*), e o filtro é aplicado em cada nova imagem formada. Este filtro deve ter alguma invariância na posição e escala. A quantidade de invariância determina o número de escalas e de posições em que deve ser aplicado (VALENÇA, 2005).

3 – MODELO GERAL BÁSICO DE SISTEMA

Em geral um sistema básico para gerência de acesso baseado em faces apresenta uma interface para captura de imagens e armazenamento em banco de dados através de uma câmera integrada a um dispositivo. Para facilitar os cálculos dos valores é aplicada à imagem uma conversão para tons de cinza. Essa conversão permite que os valores de cor no padrão (RGB) sejam normalizados para valores padrões, conforme Equação 4.1.

$$\text{Cor.normalizada} = 0.3 * \text{cor.vermelho} + 0.59 * \text{c.verde} + 0.11 * \text{cor.azul} \quad (4.1)$$

Cada pixel é convertido para seu valor normalizado gerando a imagem em tons de cinza. Dessa forma não há necessidade de se preocupar com as cores em separadas, mas sim com uma cor somente o que diminui o esforço computacional.



Figura 9 - Exemplo de Tela

De forma geral, com o propósito de diminuir o esforço computacional a imagem utiliza-se uma resolução de 20x20 pixel.

O próximo passo é calcular os valores para definição dos *eigenvetores* utilizados para calcular as *eigenfaces*. Para isso vinte imagens do banco de dados são selecionadas aleatoriamente, a quantidade de escolha está diretamente ligada à resolução imposta acima e deriva da necessidade de criação de matrizes quadradas para cálculos dos auto vetores e autovalores.

Gera-se um vetor bidimensional com tamanho de 400x20, aonde 400 deriva dos pixels das imagens e os 20 da quantidade de imagens sorteadas na base de dados de faces.

Calcula-se uma matriz de valores médios de cada posição $(x_{(1,1)}+x_{(2,1)}+\dots+x_{(20,1)}/20)$, para todas as 400 posições da matriz criada e utiliza-se essa matriz para normalizar o vetor através da subtração dos valores originais da matriz pela matriz de média obtida acima.

Com a matriz normalizada é obtida a matriz de covariância. Para tanto se multiplica a matriz normalizada pela sua inversa. Note que tal multiplicação gera uma matriz que além de quadrada (20x20) é simétrica. É justamente essa matriz que é utilizada para os cálculos dos autovalores e auto vetores através do algoritmo de QL. (CAMBRIDGE, 1992)

Esse algoritmo efetua deslocamentos para determinar os autovalores e auto vetores de uma matriz real, simétrica e triangular, que é previamente reduzida. Como entrada o algoritmo recebe os elementos diagonais da matriz triangular e gera como saída os autovalores e auto vetores relacionados (algoritmo de QL – anexo A).

A saída gera dos vetores um com os autovalores e outro com os auto vetores associados. Utiliza-se então uma função de ordenação rápida (Algoritmo de QuickSort) para organizar a ordem dos auto vetores em relação aos autovalores encontrados, de forma crescente. Serão esses auto vetores que gerarão os *eigenfaces* com características gerais para a representação dentro da rede neural.

4 CONCLUSÕES

É possível observar que em sistemas de detecção e análise de faces a utilização de *eigenvalues* e *eigenvectores* na representação de imagens de faces em concomitância com as redes neurais apresenta uma boa relação de usabilidade por sistemas que necessitem implementar controles de acessos. Os algoritmos de cálculos são rápidos e apresentam um atraso tolerável para processamento da resposta esperada. De forma geral, a confiabilidade desses sistemas após um rápido treinamento apresentam resultados bastante satisfatórios.

Conclui-se assim que a implementação de um sistema utilizando redes neurais com base em características faciais simplificadas é viável e de excelente performance, não necessitando de complexos equipamentos envolvidos nem de um poder de processamento extremamente alto, uma vez que a utilização de *eigenfaces* impõe a redução da complexidade das redes neurais.

REFERÊNCIAS BIBLIOGRÁFICAS

- [CAMBRIDGE, 1992] Numeric Recipes In C: The art of Scientific Computing. 1988-1992. Cambridge University Press.
- [COSTA, 2004] Costa, Wadaed Uturbey da. Técnicas Bayesianas para Engenharia Elétrica. Universidade Federal de Minas Gerais – Depto. de Engenharia Elétrica. Janeiro, 2004.
- [COSTA, 2009] Costa, Luciano R.; Obelheiro, Rafael R.; Fraga, Joni S. Introdução a Biometria. Cap. 3. Universidade Federal de Santa Catarina, 2009.
- [JAIN, 2002] Jain, Anil; Bolle, Ruud; Pankanti, Sharath. Biometrics - Personal Identification in Networked Society. Ed. Kluwer Academic Publishers. 2002.
- [KÖRTING, 2004] Körting, Thales Sehn. Filho, Nelson Lopes Duarte. Utilizando Eigenfaces para Reconhecimento de Imagens. Fundação Universidade Federal do Rio Grande do Sul, 2004.
- [LOPES, 2005] Lopes, Eduardo Costa. Detecção de Faces e características Faciais. Telatório Técnico, PUCRS, 2005.
- [MACHADO, 2009] Machado, Bruno Bechler, et. al. Implementação de um algoritmo de reconhecimento facial usando Eigenface. Uni-BH, 2009.
- [MARCELO, 1996] Marcelo S. Portugal e Luiz Gustavo L. Fernandes. Redes neurais artificiais e previsão de séries econômicas: Uma introdução. Nova Economia, UFMG, Belo Horizonte, 1996.
- [MARIN, 2006] Marin, Luciene de Oliveira. Métodos Estatísticos no Reconhecimento de Faces. RESI-Revista Eletrônica de Sistemas de Informação, Edição 8, No2, 2006.
- [MORAES, 2006] Moraes, Alexandre Fernandes de. Método para avaliação da tecnologia biométrica na segurança de aeroportos. Dissertação de Mestrado. USP. São Paulo, 2006.
- [OLIVEIRA, 2009] Oliveira, Andréia Brandão de; Et al. Tecnologia Biométrica – Aspectos de Segurança e Privacidade. São Paulo, 2009.
- [OPPER, 2022] - <https://operdata.com.br/blog/analise-de-componentes-principais/> (último acesso em 07 de março de 2022)
- [OSEAS, 2004] Oséas Pereira Rocha. Utilização da transformada wavelet em multi-resolução para o reconhecimento de comandos de voz. Master's thesis, Universidade Cândido Mendes, 2004.

[PENHARBEL,2005] Penharbel, Eder Augusto; Wutzke, Erdiane L. G.; Silva Murilo dos S.;Bianchi, Reinaldo A.C.. E-Faces - Um classificador capaz de analisar imagens e classificá-las como faces ou não faces utilizando o método Eigenfaces. I Workshop de Visão Computacional, Centro Universitário da FEI - UNIFEI,Departamento de Ciência da Computação, 2005.

[ROSSA, 2005] Rossa, Arun; Govindarajanb, Rohin. Feature Level Fusion Using Hand and Face Biometrics. SPIE - Conference on Biometric Technology for Human Identification II, Vol. 5779, pp. 196-2004, (Orlando, USA), Março/2005.

[ROWLEY, 1998] Rowley, Henry A.;Baluja, Shumeet; Kanade, Takeo. Neural Network-Based Face Detection. IEEE, PAMI, January 1998.

[TURK, 2007] Turk, Matthew A.;Pentland, Alex P. Pentland. Face Recognition Using Eigenfaces. Vision and Modeling Group, The Media Laboratory. Massachusetts Institute of Technology,, Vol. 3, No. 1, Pages 71-86 Posted Online December 13, 2007.

[VALENÇA, 2005] Valença, Mêuser Jorge Silva. Aplicando Redes Neurais – Um guia completo.Olinda, PE. Ed. Do Autor, 2005.

[WAYMAN, 2005] Wayman, James;Jain, Anil; Maltoni, Davide; Maio, Dario. Biometric Systems. Technology, Design and Performance Evaluation. Editora Springer. 2005.

ANEXOS

ANEXO A – Algoritmo QL

Procedimento QL(Real d(), e(), z(,); Inteiro n)

Inteiro m, l, iter, i, k, contar

Real s, r, p, g, f, dd, c, b

Para contar = 0 até 2 faça

Para i = 2 até n faça $e(i - 1) = e(i)$;

$e(n) = 0.0$;

Para l = 1 até n faça

iter = 0;

Enquanto

para m = l até n - 1 faça

$dd = \text{Abs}(d(m)) + \text{Abs}(d(m + 1))$;

Se $\text{Abs}(e(m) + dd) = dd$ então saia para

próximo

Se não (m = l) então

se iter + 1 = 30 então MsgBox("Muitas interações no procedimento")

$g = (d(l + 1) - d(l)) / (2.0 * e(l))$; r = pythag(g, 1.0);

$g = d(m) - d(l) + e(l) / (g + \text{sign2}(r, g))$;

S = 1.0; c = 1.0; p = 0.0;

para i = m - 1 até l passo -1 faça

$f = s * e(i)$; $b = c * e(i)$; $e(i + 1) = \text{pythag}(f, g)$; $r = e(i + 1)$;

se r = 0 então

$d(i + 1) = d(i + 1) - p$;

$e(m) = 0.0$;

Saia do para;

fim-se;

$s = f / r$; $c = g / r$; $g = d(i + 1) - p$; $r = (d(i) - g) * s + 2.0 * c * b$;

$p = s * r$; $d(i + 1) = g + p$; $g = c * r - b$;

Para $k = 1$ até n faça

$f = z(k, i + 1)$; $z(k, i + 1) = s * z(k, i) + c * f$;

$z(k, i) = c * z(k, i) - s * f$;

próximo

próximo

se ($r = 0.0$) e ($i \geq l$) então continue enquanto

$d(l) = d(l) - p$; $e(l) = g$; $e(m) = 0.0$;

fim-se

Faça enquanto não ($m = l$)

proximo

próximo

Fim do procedimento

ANEXO B – Algoritmo da Rede Neural

Função CalcularSaida(Inteiro: itemDados, posx, posy)

Para x=1 to posy faça

CalcularSaida = CalcularSaida + (Neuronio(posx,y) * peso(posx,y));

Fim da Função

Procedimento TreinarRedes(Inteiro: y)

Inteiro i, j, Epocas;

Real resultado, esperado;

Epocas = QuantidadeTreinamento;

Para j = 1 até Epocas faça

Para i = 0 até neurônio.tamanho – 1 faça

resultado = CalcularSaida(i);

esperado = Neurônio(i, y);

Se resultado <> esperado então

CalcularAjuste(i, esperado, resultado,y);

fim-se;

próximo;

próximo;

Fim do procedimento;

Função CalcularAjuste(Inteiro itemDados, y; Real ResultadoEsperado, UltimoResultado): Real

Inteiro: i;

Para i = 0 até TerminaisEntrada – 1 faça

CalcularAjusteTerminal(itemDados, i, ResultadoEsperado, UltimoResultado, y);

Próximo

preenchePesos();

Fimfunção;

Procedimento CalcularAjusteTerminal(Inteiro itemDados, subItem, y; Real: ResultadoEsperado, UltimoResultado)

Real: $w_i = \text{peso}(\text{subItem}, y)$

Real $\text{taxa_aprendizado} = \text{taxa.texto}$

Real $\text{valor_entrada} = \text{peso}(\text{subItem}+1, y)$

Real $\text{erro} = (\text{ResultadoEsperado} - \text{UltimoResultado})$

$\text{peso}(\text{subItem}, y) = w_i + (\text{taxa_aprendizado} * \text{valor_entrada} * \text{erro})$

Fim do procedimento