UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

JAIRO LUCAS DE MORAES

CONTROLE DE ACESSO BASEADO EM BIOMETRIA FACIAL

VITÓRIA
2010

JAIRO LUCAS DE MORAES

CONTROLE DE ACESSO BASEADO EM BIOMETRIA FACIAL

Dissertação apresentada ao Programa de Pós-Graduação em Informática do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Informática.

VITÓRIA
2010

JAIRO LUCAS DE MORAES

CONTROLE DE ACESSO BASEADO EM BIOMETRIA FACIAL

COMISSÃO EXAMINADORA

Prof. Dr. Alberto Ferreira de Souza
Universidade Federal do Espírito Santo
Orientador

Prof. Dr. Elias de Oliveira
Universidade Federal do Espírito Santo

Dr. Fábio Daros de Freitas
Receita Federal do Brasil

Vitória, _____ de _____ de _____.

Cidadão do mundo é cidadão de lugar nenhum.

Jairo Lucas de Moraes

Dedico este trabalho a meus pais Isaltina e Resoly e a meus irmãos Jane e Jaime, pelo incentivo e apoio incondicional, não somente neste trabalho, mas em toda a vida.

Agradeço ao meu orientador Prof. Alberto Ferreira de Souza, pela paciência e ensinamentos que tornaram este trabalho possível.

Ao amigo Nuno Rasseli, meu “consultor” em C#, cuja ajuda foi de grande valia para este trabalho.

Ao amigo Osvaldo Medina pelo incentivo e apoio.

RESUMO

A tarefa de reconhecimento facial é uma das tarefas mais corriqueiras e naturais executadas pelos seres humanos. Apesar de simples para nós, esta tarefa tem se mostrado um grande desafio para pesquisadores nas áreas de inteligência artificial e visão computacional.

As pesquisas na área de detecção e reconhecimento de objetos, e mais especificamente aquelas voltadas à face humana, têm aumentado muito nos últimos anos principalmente devido à sua aplicabilidade em áreas tais como: segurança pública, controle de acesso, autenticação continua em redes de computadores, entre outras. Nesta dissertação, investigamos a viabilidade de um sistema para controle de acesso que usa unicamente a biometria da face como chave de acesso. Neste caso, o controle de acesso deixa de ser baseado em “algo que o individuo tem” ou “algo que o individuo sabe” e passa a ser baseado no próprio indivíduo.

Para investigar a viabilidade de um sistema de controle baseado unicamente na biometria da face, desenvolvemos um protótipo de tal sistema que atua de forma completamente automática, sendo capaz de detectar uma face em uma imagem estática ou em vídeo e de efetuar o reconhecimento da face sem nenhum tipo de intervenção humana. Para detectar a face utilizamos uma abordagem já referência na literatura (Viola Jones) e para o reconhecimento utilizamos redes neurais sem peso do tipo *Virtual Generalizing RAM* (*VG-RAM* WNN). Por fim, para o controle de acesso, empregamos técnicas probabilísticas Bayesianas.

Os resultados obtidos são promissores. Simulamos o controle de acesso de 50, 100 e 200 usuários a determinado recurso e, com o conjunto de 200 usuários, o sistema conseguiu autenticar corretamente 93% dos usuários com um *FAR* (*False Acceptance Rate*) de apenas 0,77%, com o conjunto de 100 usuários o sistema conseguiu autenticar corretamente 90,25% com um *FAR* de 1,79%, e com o conjunto de 50 usuários o sistema autenticou corretamente 93,11% com um *FAR* de 4,76%.

ABSTRACT

Face recognition is one of the most ordinary and natural task carried out by humans. Even though it is a simple task, it has proved to be a major challenge for artificial intelligence and computer vision researchers.

Researches on object detection and recognition, and more specifically those related to human face, has greatly increased their demand in recent years mainly due to its employ in applications such as: public safety, access control, continuous authentication on computer networks, among others. This dissertation studies the feasibility of a system for access control using only facial biometrics as access key. In this case, access control would no longer be based "Something that the person has" or "something that the individual knows" but it becomes the person itself.

To inquire the feasibility of an access control system based only in face biometric, we developed a prototype of this system that operates fully automatically, being able to detect a face in a static image or in a video and then perform the recognition of that face, with no human intervention. We use a well known approach [Viola01] for the task of face detection and *VG-RAM* WNN (Virtual Generalizing Random Access Memory Weightless Neural Networks) for the recognition assignment. Lastly, we employed Bayesian probabilistic techniques for the access control problem.

The obtained results are promising. The access control to given a resource was simulated for a number of 50, 100 and 200 users. For the set of 200 users the system was able to authenticate correctly 93.00% of the users with a *FAR* (*False Acceptance Rate*) of only 0.77%, for the set of 100 users the system was able to authenticate correctly 90.25% of the users with a *FAR* of 1.79%, and for the set of 50 users the system properly authenticated 93.11% of users with a *FAR* de 4.76%.

SUMÁRIO

LISTA DE SIGLAS E ABREVIATURAS

EER	*Equal Error Rate*
FAR	*False Acceptance Rate*
FRR	*False Rejection Rate*
VP	Verdadeiro Positivo
VN	Verdadeiro Negativo
FP	Falso Positivo
FN	Falso Negativo
RNSP	Redes Neurais Sem Peso
ROC	*Receiver Operating Characteristic*
RAM	Randon Access Memory
VG-RAM	*Virtual Generalizing RAM*
LCAD	Laboratório de Computação de Alto Desempenho
UFES	Universidade Federal do Espírito Santo

LISTA DE FIGURAS

LISTA DE TABELAS

1 INTRODUÇÃO

O processo de identificação ou verificação eletrônica da identidade de pessoas tem se tornado cada vez mais corriqueiro. Atualmente, compreende desde o acesso à garagem de edifícios por meio de cartões magnéticos até a digitação de códigos de identificação e senha eletrônicas em sites de bancos. Em vários níveis, mais ou menos transparentes, invariavelmente, o indivíduo, hoje, convive com os processos de identificação e verificação de identidade no seu dia a dia.

Segundo [Hong98], as abordagens convencionais de identificação e verificação são baseadas em “algo que você sabe” — como uma senha ou número de identificação — ou “o que você tem” — como um cartão de identificação. Infelizmente, para diversas aplicações, estes métodos podem não ser suficientemente seguros para garantir uma identificação pessoal, pois senhas e cartões de acessos podem ser roubados ou falsificados.

Os sistemas biométricos de identificação — que compreendem dispositivos de captura de informações biométricas: imagem da face, imagem da íris, impressão digital etc., bancos de dados e software para o armazenamento e manipulação destas informações [Vetter10] — fazem a identificação ou verificação de indivíduos baseados nas características físicas e comportamentais dos mesmos. Ou seja: os próprios indivíduos passam a ser a “chave de identificação”, o que torna o processo mais transparente e menos sujeito a fraudes.

Em função de sua grande aplicabilidade em diversas áreas — segurança pública, autenticação contínua em redes de computadores, controle de acesso etc. —, os sistemas de identificação biométrica estão ganhando cada vez mais atenção de pesquisadores na academia e indústria [Yang02, Zhao03].

Entre as várias alternativas de informação biométrica que hoje podem ser capturadas por dispositivos parte de sistemas de identificação, a imagem da face é umas das que mais tem se destacado. Isso tem ocorrido porque os dispositivos de captura — câmeras digitais — operam de forma não invasiva, são de baixo custo e fácil utilização. Além disso, o aumento contínuo do desempenho dos processadores nas últimas décadas permitiu o uso de algoritmos mais

sofisticados, robustos, confiáveis — e com um tempo de resposta aceitável — no processo de detecção e reconhecimento de faces.

Na Figura 1, está mostrado o esquema geral de um Sistema de Reconhecimento de Faces [Zhao03]. Ele é constituído de três módulos principais: detecção de face, extração das características da face e reconhecimento de face.

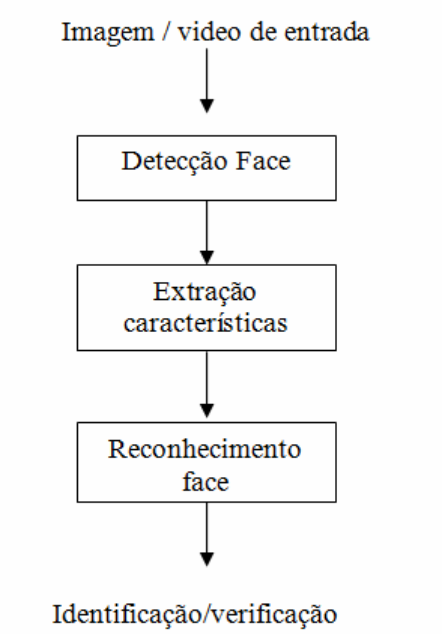

Figura 1 — Esquema genérico de um sistema de reconhecimento de face
Fonte: Adaptado de [Zhao03]

Neste estudo, investigamos o problema de controle de acesso utilizando a biometria da face como chave única de acesso.

1.1 Motivação

A principal motivação para a realização deste trabalho é a pequena quantidade de trabalhos científicos na área de controle de acesso que utilizem unicamente a biometria da face como chave de identificação. Existem muitos trabalhos na área conhecida como reconhecimento de faces, mas não no seu emprego para o controle de acesso.

Para deixar claro este ponto, é importante definir os termos detecção de faces, reconhecimento de faces, verificação de face, autenticação de face e identificação de face no contexto de sistemas de identificação biométricos, além do controle de acesso no mesmo contexto.

Detecção de face (*face detection*)

O problema de detecção de face pode ser caracterizado como “dada uma imagem arbitrária de entrada, determinar se existe ou não uma face humana na imagem e, caso exista, retornar as coordenadas onde foi encontrada a face humana” [Sung94, Yang02].

Reconhecimento de face (*face recognition*)

O problema de reconhecimento de face pode ser caracterizado como “dado uma imagem de entrada de uma face, comparar a face de entrada com uma biblioteca de modelos de faces conhecidas, e reportar se uma equivalência foi encontrada” [Yang02]. Os sistemas biométricos de reconhecimento de face podem operar em dois modos distintos [Hong98, Zhao03, Jain04].

a — Verificação (*Verification*) : O Indivíduo fornece seus dados biométricos e um código de identificação: nome, CPF, identificação funcional etc. O sistema examina se os dados biométricos de entrada são aqueles pertencentes ao indivíduo cuja identidade é reivindicada [Hong98, Zhao03, Jain04].
Ou seja: é feita uma comparação um-para-um, em que o sistema deve buscar uma resposta binária para a pergunta “Eu sou quem reivindico ser?” Computacionalmente, isto significa que não é necessário examinar toda a base de conhecimento a fim de verificar a veracidade da reivindicação. Alguns autores, como [Yang02], definem este processo como autenticação de face (*face authentication*).

b — Identificação (*Identification*) : O Indivíduo fornece seus dados biométricos ao sistema, que deve examinar toda a sua base de conhecimento a fim de encontrar um individuo com características equivalentes. O sistema deve responder à pergunta “Quem sou eu?”

O problema de identificação é mais complexo que o de verificação, pois, além de garantir a mesma acurácia exigida na verificação, deve percorrer toda a base de conhe-

cimento e fazer a identificação, o que torna o tempo de resposta um problema a ser tratado [Hong98].

Na figura 2, é mostrado um diagrama da estrutura do modulo de reconhecimento de face.

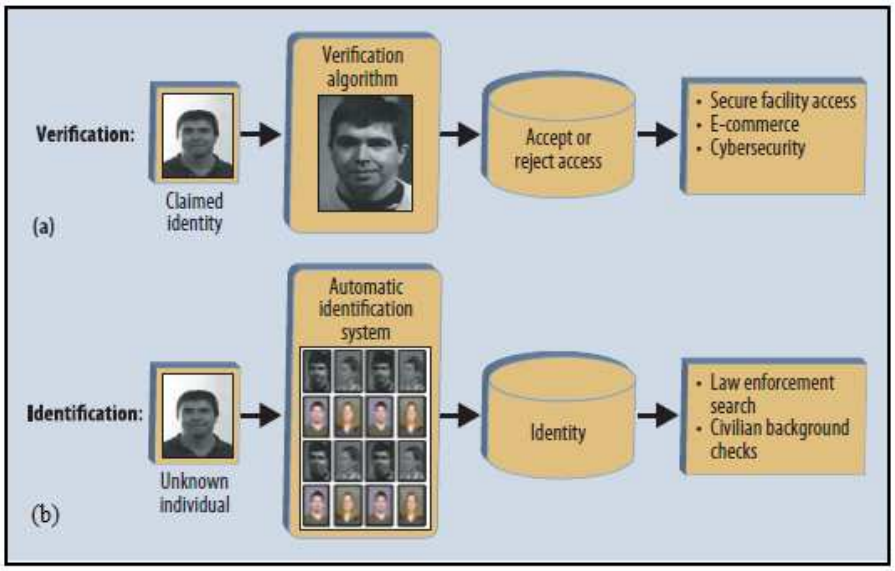

Figura 2 — Estrutura da tarefa de reconhecimento de face
Fonte: Adaptado de [Chellappa, Zhao10]

Controle de acesso

O problema de controle de acesso, no contexto deste trabalho, será caracterizado como “dada uma imagem de entrada de uma face, examinar a base de conhecimento sobre as pessoas que possuem acesso a determinados recursos ou ambiente, e reportar se a face de entrada pertence a uma destas pessoas ou não”. O sistema deve fornecer uma resposta binária para a pergunta “possuo acesso a este recurso?” Caso haja uma resposta afirmativa, o sistema poderá identificar o individuo, respondendo à pergunta “Quem sou eu?”.

É importante observar que o problema de controle de acesso difere do de identificação de face. Um sistema de identificação sempre irá retornar uma face que contenha as características mais equivalentes à face de entrada, ainda que a mesma não esteja na base de conhecimento (falsa identificação), já um sistema de controle de acesso irá retornar uma resposta negativa nos casos onde a face de entrada não se encontra na base de conhecimento.

1.2 Objetivos

O objetivo principal deste trabalho é a investigação do problema de controle de acesso utilizando unicamente a biometria da face como chave de acesso. Foi utilizada a técnica de Viola e Jones [Viola01] para a detecção de faces, redes neurais sem peso (RNSP) para o reconhecimento facial e técnicas probabilísticas para o controle de acesso.

1.3 Contribuições

A principal contribuição deste trabalho foi o desenvolvimento de um protótipo funcional de um sistema de controle de acesso que utiliza exclusivamente a biometria da face como chave de acesso.

Este trabalho se diferencia de outras abordagens relacionadas ao reconhecimento de face, tais como [Tolba05], [Miller94], por tratar o controle de acesso como problema binário: o sistema deve responder se determinado usuário tem ou não acesso a determinado recurso ou ambiente, usando, para isso, somente os dados biométricos da face do usuário.

Os trabalhos citados tratam do problema da identificação, sempre retornando a face que possui a maior similaridade com o modelo fornecido, ou do problema de verificação, que valida ou não uma identidade reivindicada, implicando que o usuário deva ter um código de identificação além dos dados biométrico da face.

1.4 Estrutura do Trabalho

Esta dissertação está divida da seguinte forma:

Capítulo 2

Explanação sobre biometria e os principais métodos de identificação biométrica utilizados atualmente, métricas utilizadas e medidas de performance.

Capítulo 3

Explanação detalhada dos métodos de detecção de objetos, mais especificamente a detecção de face e mãos humanas, técnicas usadas neste trabalho. Neste capítulo será abordada, com maior ênfase, a técnica proposta por Viola e Jones [Viola01], implementada como parte deste trabalho.

Capítulo 4

Explanação geral sobre o funcionamento de Redes Neurais Sem Peso (RNSP), Redes *VG-RAM*, reconhecimento de faces e uma visão geral do sistema proposto.

Capítulo 5

Metodologia usada, especificando bases de dados, software e hardware no desenvolvimento e testes, métricas utilizadas e conjuntos de avaliações testados.

Capítulo 6

Experimentos efetuados e os resultados obtidos.

Capítulos 7 e 8

Discussão sobre os resultados com as principais conclusões e sugestões de trabalhos futuros.

2 BIOMETRIA

O avanço da tecnologia em várias áreas possibilitou à sociedade moderna oferecer as mais diversas facilidades aos seus indivíduos. Hoje, é possível efetuar transações financeiras sem sair de casa, fazer reuniões com pessoas que estão a milhares de quilômetros de distância, assistir aulas e palestras sendo proferidas em outro país ou viajar de um continente a outro em poucas horas.

Porém, todas essas conveniências, e um número cada vez maior de pessoas usufruindo as mesmas, tornaram indispensável o uso de mecanismos de identificação pessoal, cada vez mais robustos, que possam comprovar que um indivíduo realmente é quem alega ser.

Estes mecanismos, que se apresentam na forma de cartões magnéticos, senhas pessoais, cartões de identidade, passaporte etc., trazem também uma série de problemas associados, tais como perda, falsificação, empréstimo, dificuldade de memorização ou armazenamento de vários códigos, dentre outros.

O processo de identificação pessoal baseado em biometria tenta minimizar estes problemas, pois ele deixa de ser baseado em algo que o individuo tem, ou algo que o individuo sabe, e passa a considerar o próprio indivíduo como código de identificação.

O termo biometria tem origem na união das palavras gregas *bios* — que significa vida — e *metron*, que significa mensuração, e pode ser definida em um contexto geral como a ciência que estuda a mensuração de características dos seres vivos [Holanda09].

A biometria explora o fato que certas características físicas ou comportamentais dos seres vivos poderem ser usadas, de maneira confiável, para diferenciar um ente de seus pares. Jain [Jain07] define biometria como sendo o reconhecimento pessoal baseado em características comportamentais ou fisiológicas de um indivíduo.

2.1 Sistemas de Identificação Baseados em Biometria

Sistemas de identificação baseados em biometria são basicamente sistemas de reconhecimento que, dado uma informação biométrica de entrada, conseguem distinguir padrões e separá-los em diferentes classes ou categorias.

Entre as principais características anatômicas, fisiológicas e comportamentais usadas nos sistemas biométricos pode-se citar impressão digital, geometria das mãos, aparência facial, temperatura da face, íris, retina, voz, assinatura, padrão de andar e arcada dentária, dentre outros.

A figura 3 mostra algumas destas características:

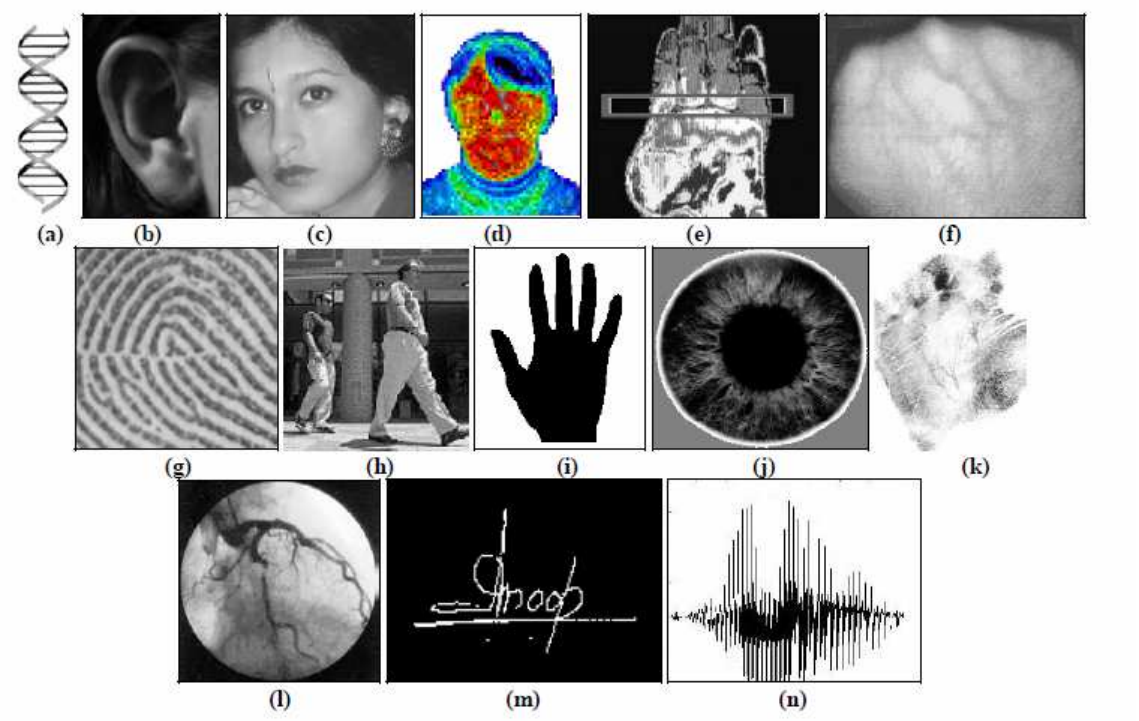

Figura 3 — Exemplos de características biométricas:
a — DNA ; b — Anatomia da orelha; c — Aparência facial; d — Termograma facial; e — Termograma das mãos; f — Veias das mãos; g — Impressão digital; h — Análise da forma de caminhar; i — Anatomia das mãos; j — Análise da íris; k — Análise da impressão da palma da mão; l — Análise da retina; m — Assinatura; n — Análise do padrão de voz
Fonte: [Jain04]

Na prática, qualquer característica anatômica, fisiológica ou mesmo comportamental de um ser humano pode ser usada como padrão em um sistema biométrico, desde que ofereça as seguintes propriedades [Jain04]:

• Universalidade — Todas as pessoas devem possuir;

• Unicidade — Deve ser única em cada pessoa, ou seja, possuir padrões diferentes em pessoas diferentes;

• Coletabilidade — Deve ser passível de ser coletada e medida

• Permanência — Deve permanecer invariável durante certo período de tempo.

Em sistemas biométricos de identificação usados em aplicações no mundo real, outras questões devem ser levadas em consideração, tais como:

• Performance — permitir processar com o tempo e a precisão necessária para a aplicação;

• Aceitabilidade — grau com que as pessoas aceitam fornecer as características ao sistema;

• Impostura de difícil imitação — quanto menor o grau de impostura, mais difícil imitar.

Atualmente, nenhuma característica biométrica apresenta todas as propriedades no seu maior grau. Ou seja: não existe a “melhor característica” para identificação de um indivíduo. A escolha depende da aplicação em que será usada.

A próxima seção apresenta um descritivo com as principais características biométricas utilizadas atualmente.

2.2 Comparação entre Características Biométricas

Nenhuma característica biométrica atende as necessidades de todas as aplicações. A correspondência entre característica e aplicação é determinada pela forma como esta aplicação vai funcionar e os pontos fortes apresentados por determinada característica. Abaixo, é formulado um breve resumo das características mais comumente usadas.

DNA — Deoxyribo do Ácido Nucleico

Atualmente, é das características biométricas que apresentam maior unicidade e permanência. Excetuando-se gêmeos idênticos, que possuem cadeias de DNA idênticas, cada indivíduo tem a sua, única e exclusiva, que não se altera ao longo da vida. Seus pontos fracos são a dificuldade de coleta e a aceitabilidade. O reconhecimento de DNA envolve processo químico extremamente complexo, caro e especializado, limitando grande parte do seu uso a práticas forenses [Jain04].

Orelhas

Utiliza a anatomia da orelha para identificar indivíduos, em abordagens pouco comuns. Os pontos fortes são aceitabilidade e permanência; os fracos, unicidade e performance. Maiores informações sobre esta técnica podem ser obtidas em [Kyong03].

Termograma facial e de mãos

O padrão de calor irradiado pelo corpo humano é característica individual de cada pessoa, e pode ser captado por câmara de infravermelho. Sistemas baseados em imagens termográficas não requerem contato nem a cooperação do indivíduo. Porém, a captura das imagens ainda é um desafio em ambientes não controlados, pois a mesma é afetada por fontes de calor que, eventualmente, possam estar próximas ao individuo. Seus pontos fortes são universalidade, impostura e unicidade. O principal ponto negativo é a permanência, visto que o estado de saúde — ou mesmo o emocional — do individuo pode afetar o calor emanado pelo corpo [Jain04].

Impressão digital

Característica mais utilizada em sistemas de identificação automatizados em larga escala. Sua popularidade se deve, em parte, ao baixo custo dos aparelhos coletores e à performance razoável no processo de identificação. Apesar da impressão digital não mudar naturalmente ao longo dos anos, é suscetível a fatores ambientais a que são submetidos os indivíduos, podendo fazer com que seja alterada ou deteriorada. Trabalhadores braçais, por exemplo, podem ter impressão digital alterada constantemente em função de cortes profundos ou outros ferimentos nos dedos [Jain04].

Forma de caminhar (marcha)

Apesar de não ser muito distinta, é suficientemente discriminatória para permitir aplicação em sistema de pouca segurança. Possui baixa taxa de permanência, podendo mudar constantemente em função do peso, fraturas nas articulações ou embriaguês do indivíduo. O ponto forte é a fácil coletabilidade, sendo feita sem a cooperação do individuo e com simples câmara digital. Informações mais detalhadas sobre esta técnica podem ser obtidas em [Gafurov06].

Anatomia da mão

Utiliza medições como forma da palma e comprimento e largura dos dedos. Bastante usada, pois, apesar de simples, possui alta unicidade e alta permanência (desde que coletada na vida adulta). Seus pontos fracos são alto grau de impostura e difícil coletabilidade, pois necessita de aparelho especifico com tamanho razoavelmente grande, o que inviabiliza a incorporação da técnica em desktops, por exemplo [Jain04].

Íris

Formada durante o desenvolvimento fetal, se estabiliza nos dois primeiros anos de vida. Sua textura é extremamente complexa e traz informações úteis para serem usadas no reconhecimento facial. Apresenta alta unicidade, sendo distinta mesmo em gêmeos idênticos. Possui baixo grau de impostura, pois é extremamente difícil, mesmo cirurgicamente, alterar a textura da íris. Seu ponto fraco fica por conta da coletabilidade, pois necessita de hardware caro e complexo e também da cooperação do indivíduo [Jain04].

Retina

O scanner de retina se baseia na análise das artérias e veias existentes no olho humano. É, supostamente, a mais segura característica biométrica do ser humano, sendo impossível reproduzir seus padrões. Possui alta unicidade, sendo diferente em cada olho e mesmo em gêmeos idênticos. É uma técnica pouco utilizada em função da difícil coletabilidade, dependendo de hardware especifico e complexo. Além disso, é bastante invasiva e necessita de cooperação total do indivíduo a ser identificado [Jain04].

Assinatura

Característica biométrica comportamental do individuo. Apesar de ser a mais utilizada em métodos de identificação não automatizados — devido à fácil coletabilidade e aceitabilidade —, é fraca em relação à impostura e à permanência. Falsários profissionais são capazes de reproduzir a maioria das assinaturas. Além disso, são substancialmente afetadas por condições físicas e emocionais do indivíduo [Jain04].

Voz

Combinação de biometria fisiológica e comportamental. Não muda em períodos curtos, mas pode ser afetada por fatores como um simples resfriado, estado emocional e ruídos de fundo. Possui baixa unicidade, não sendo recomendada para identificação em larga escala. O ponto forte é coletabilidade e aceitabilidade, além do baixo custo dos coletores. Geralmente indicada para verificação de identidade em conversas telefônicas. [Jain04].

Face

Característica biométrica mais usada pelos seres humanos para fazer uma identificação pessoal. A gama de aplicações usando esta característica vai desde aplicações para reconhecimento de faces estáticas em ambiente controlado até verificação de identidade em imagens em tempo real e com fundo complexo. As abordagens mais populares usadas no problema de reconhecimento de face se baseiam na localização e análise de atributos faciais como olhos, nariz e boca, ou na análise global da mesma, representada como combinação ponderada de uma série de faces canônicas.

Embora o desempenho dos sistemas comerciais que utilizam esta propriedade biométrica para reconhecimento seja razoável, impõem restrições sobre como a imagem da face foi obtida, necessitando fundo controlado, iluminação razoavelmente ajustada e poses no ângulo esperado [Phillips02]. Alguns autores — como [Tolba05] — acreditam que aspectos faciais somente têm bons resultados se combinados com outras características biométricas.

Os pontos fortes são fácil coletabilidade, sendo possível fazer sem a cooperação da pessoa, alta aceitabilidade e universalidade. Os sistemas de identificação baseados no reconhecimento da face utilizam quatro passos básicos:

- Detectar se existe uma face em uma imagem ou seqüencia de vídeo;
- Definir a localização da mesma;
- Extrair as características e comparar com base de conhecimento existente; e.
- Retornar o padrão armazenado que mais se aproximada da imagem de entrada.

Este trabalho utiliza as características biométricas da face para efetuar o controle de acesso. Nos próximos capítulos, detalhamos as técnicas usadas e resultados obtidos.

2.3 Medidas de Desempenho de Sistemas Biométricos

Sistemas biométricos são projetados para procurar similaridade entre um padrão biométrico de entrada e padrões armazenados em uma base de conhecimento. Este nível de similaridade raramente será de 100%. Pela própria essência, a biometria tem grande variação intraclasse, onde padrões do mesmo indivíduo podem variar devido a poses ou uso ou não de barba, óculos, maquiagem etc.

Em função desta variação, estes sistemas adotam algum tipo de pontuação (grau de crença, ranking, maior voto etc.) para definir se a amostra apresentada coincide ou não com determinado padrão. Quanto maior a pontuação, maior a certeza que a amostra de entrada pertence a determinado padrão.

No restante desta seção, sempre que nos referirmos a sistemas biométricos, estaremos falando exclusivamente de sistemas biométricos voltados para o controle de acesso ou sistemas de verificação de identidade.

Um sistema biométrico de controle de acesso — assim como sistemas de verificação de identidade — deve retornar uma resposta binária: aceitar o indivíduo como "legítimo" (validando a identidade reclamada ou dando acesso a determinado recurso) ou rejeitar o mesmo como "impostor". Esta decisão é tomada utilizando-se um limite conhecido como "limiar de casamento" (matching threshold). É ele que define o grau de certeza (pontuação, crença...) a ser utilizado para optar entre as classes legítimas e impostoras.

A partir desta premissa, podem ocorrer quatro situações [Bradley97, Fawcett06]:

a — Verdadeiro Positivo (True Positive — TP)

O padrão é verdadeiro e o classificador o classifica como tal, confirmando sua identidade ou o acesso a determinado recurso.

b — Verdadeiro Negativo (True Negative — TN)

O padrão é falso e o classificador o classifica com tal, recusando a identidade ou negando o acesso a determinado recurso.

c — Falso Positivo (False Positive — FP)

O padrão é falso, porém o classificador o classifica como verdadeiro. Neste caso um "impostor" tem a identidade confirmada ou o acesso a determinado recurso liberado.

d — Falso Negativo (False Negative — FN)

O padrão é verdadeiro, mas o classificador o classifica como falso. Neste caso um usuário "legítimo" tem a identidade recusada ou o acesso a determinado recurso negado.

A figura 4 mostra a matriz de confusão, também chamada de Tabela de Contingência, que ilustra graficamente as situações citadas acima.

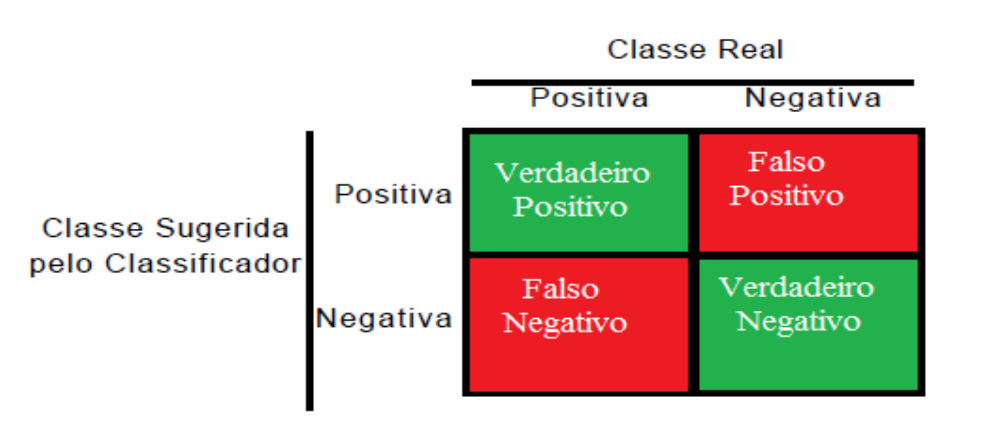

Figura 4 — Tabela de Contingência com as quatro situações possíveis para um classificador binário
Fonte: Adaptado de [Fawcett06]

A partir da Tabela de Contingência mostrada na figura 4, temos as métricas básicas mais usadas para medir o desempenho de um classificador binário (restrito a duas classes):

• *FAR* — Taxa de Falsa Aceitação (*False Acceptance Rate*, ou segundo [Phillips08], False Alarm Rate) ou, ainda, Taxa de Falsos Positivos (TFP): É a Probabilidade de um "impostor" ser considerado "legítimo". É dada pela divisão do total de falsos positivos (FP) pelo número total de negativos (FP + VN).

$$FAR = \frac{FP}{FP + VN}$$

• *FRR* — Taxa de Falsa Rejeição (*False Rejection Rate — FRR*) ou Taxa de Falsos Negativos (TFN): É Probabilidade de indivíduo "legítimo" ser considerado "impostor". É dada pela divisão do total de falsos negativos (FN) pelo total de positivos (VP + FN).

$$FRR = \frac{FN}{FN + VP}$$

• *Recall* ou Taxa de Verdadeiros Positivos : É a taxa de exemplos corretamente classificados. É dado pela divisão do total dos verdadeiros positivos pelo número total de positivos.

$$\text{Recall ou Taxa de verdadeiro Positivo} = \frac{VP}{FN + VP}$$

As taxas mencionadas acima são diretamente influenciadas pelo limiar de decisão escolhido.

A figura 5 mostra gráfico hipotético de distribuição entre classes genuína e impostora.

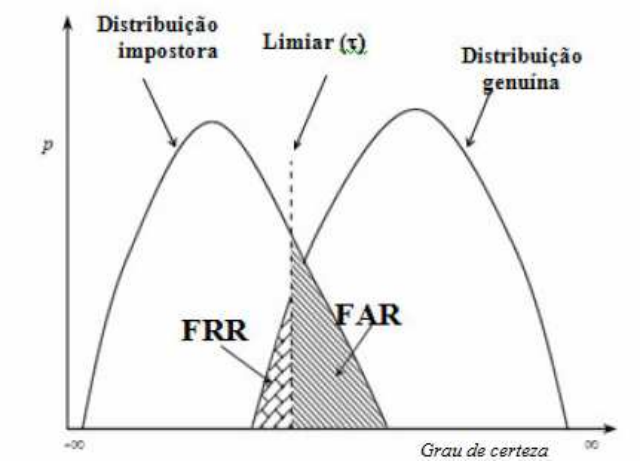

Figura 5 — Gráfico hipotético com distribuição entre classes genuína e impostora. O limiar t define a fronteira que separa as classes. As áreas pintadas mostram os intervalos em que existe ocorrência de falsos positivos e falsos negativos.
Fonte: Adaptado de [Jain04].

Pode ser observado que, ao deslocar o limiar para a direita no eixo x (grau de certeza), a taxa de falsos positivos (*FAR*) diminui e a de falsos negativos (*FRR*) aumenta. Se o limiar for deslocado na direção contrária, o efeito é inverso. Existe um limiar para o qual o valor das taxas *FAR* e *FRR* tem valores iguais. É chamado de *EER* (*Equal Error Rate*) e conseguido através da curva *ROC* (*Receiver Operating Characteristic*).

É interessante ressaltar que as importâncias das taxas de *FRR* e *FAR* não são, necessariamente, as mesmas. Variam de acordo com a aplicação. Aquelas que requerem alta segurança devem priorizar uma *FAR* mínima, mesmo que, para isso, tenha que conviver com número razoavelmente elevado de *FRR*.

As curvas *ROC* (*Receiver Operating Characteristic*) são originárias do período da Segunda Guerra Mundial, quando começaram a ser utilizadas para avaliar a precisão dos operadores de radares (chamados na época de *receiver operators*) em diferenciar sinais verdadeiros de ruídos captados pelos equipamentos.

A partir da década de 70, a aplicação das curvas *ROC* se disseminou amplamente na área médica, usada para definir riscos de determinados procedimentos e testes de diagnósticos. Posteriormente, pesquisadores da área de reconhecimento de padrões passaram a utilizar seus conceitos para avaliação de classificadores binários.

Atualmente, a curva *ROC* é dos principais métodos para definir limites de cortes levando em conta a necessidade de cada aplicação, e também é um dos modelos mais usados para comparação de performance entre classificadores binários.

Uma curva *ROC* é basicamente um gráfico de duas dimensões onde cada ponto (x, y) geralmente é representado pelas taxas *FAR* e *Recall*, respectivamente. Como o *Recall* e o *FAR* dependem diretamente de um limiar de decisão — definidor da fronteira entre os usuários "legítimos" e "impostores" —, uma curva *ROC* é formada variando-se este limiar dentro de determinado espaço, plotando-se no gráfico cada ponto (*FAR, Recall*) obtido com o limiar adotado.

Um classificador perfeito teria um ponto nas coordenadas (0,1) para qualquer variação limiar de decisão adotado. Ou seja: 0% de falsos positivos e 100% de verdadeiros positivos.

Segundo Fawcett [Fawcett06], informalmente, pode-se dizer que um ponto β em uma curva *ROC* é melhor que um ponto θ — se este encontra-se à esquerda e acima de θ, considerando-se uma aplicação onde os falsos positivos tenham a mesma importância dos falsos negativos.

A figura 6 mostra um gráfico de curva *ROC* para três classificadores diferentes

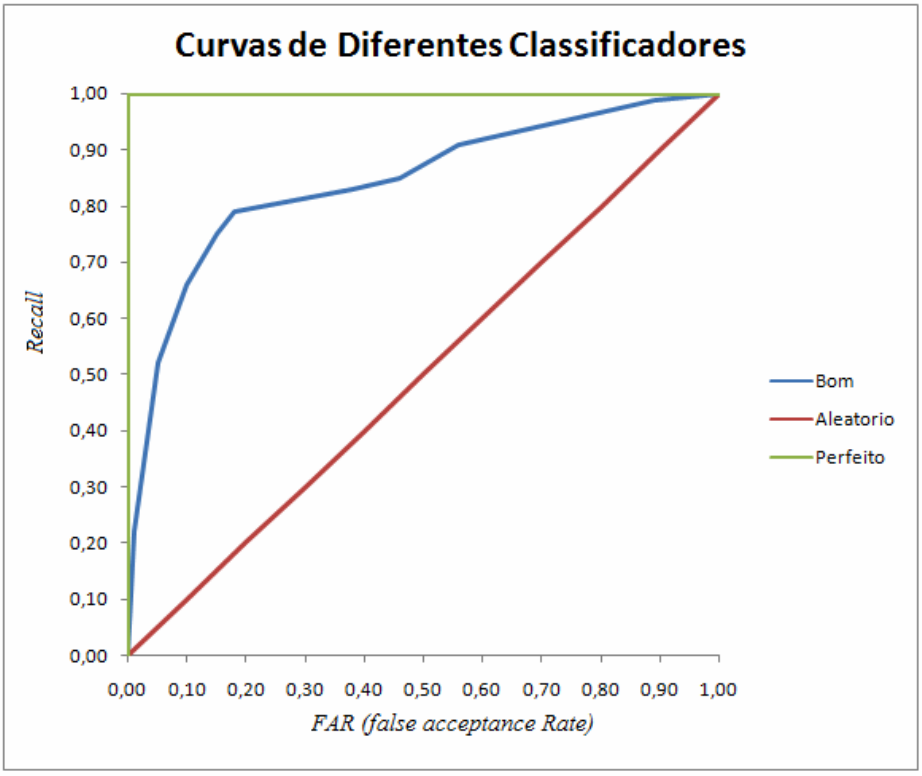

Figura 6 — Gráfico da curva *ROC* de três classificadores hipotéticos. O eixo X indica a taxa de *FAR* (*False Acceptance Rate*); o Y, a taxa de verdadeiros positivos (*Recall*).

Neste capítulo, fizemos um breve resumo sobre os principais conceitos e métricas relacionadas a sistemas biométricos. Nos próximos, estes conceitos e métricas serão largamente utilizados.

3 DETECÇÃO AUTOMÁTICA DE OBJETOS EM IMAGENS

Segundo Zhang [Zhang06], a detecção automática de objetos em uma imagem é um dos problemas mais desafiadores da área de visão computacional. O problema é assim definido: Dado uma classe de objetos de interesse T (pedestres, rostos, automóveis, edifícios...) e uma imagem P, detecção de objetos é o processo de determinar se existem instâncias de T em P, e, em caso afirmativo, retornar os locais onde T é encontrado em P.” [Zhang06]

A principal dificuldade na detecção de objetos surge da grande variabilidade da aparência dos objetos pertencentes à classe de interesse. Um carro, por exemplo, pode variar de tamanho, cor, formato e continua sendo um objeto da classe carro. Além disso, um sistema automático de detecção de objetos deve ser capaz de determinar a presença, ou não, de objetos com tamanhos diferentes, em diferentes graus de rotação e translação, com fundos complexos e com diferentes tipos de iluminação e cores [Zhang06].

Muitas abordagens têm sido propostas para o problema de detecção de objetos, sendo que a maioria delas são baseadas em modelos de treinamento estatísticos, onde primeiro as imagens de exemplo são representadas por um conjunto de características, para depois ser utilizado algum método de aprendizado para localizar os objetos de interesse da classe.

Detecção da face humana e das mãos são generalizações do problema de detecção de objetos, e serão detalhadas a seguir.

3.1 Detecção de Faces

Um dos primeiros trabalhos sobre detecção automática de face foi apresentado por Sakai, Nagao e Kaned [Sakai72] em 1972. A técnica consistia em criar uma imagem binária representando os contornos da figura, para depois extrair as características relacionadas com a face humana.

A extração das características era baseada na geometria facial, iniciando-se com a localização do topo da cabeça na matriz de contorno da imagem e, em passos seguintes, procurando bochechas, nariz, boca, queixo, contorno da boca, linhas do lado da face, linha do nariz, olhos e o eixo da face.

Naquele trabalho, está reportada um taxa de detecção de aproximadamente 91%, sem especificar a taxa de falsos positivos. A técnica proposta funciona somente em faces sem a presença de óculos e sem barba. Nos testes com modelos de óculos e usando barba cerrada foi reportado que não conseguiu detectar nenhuma das faces apresentadas [Sakai72].

Esta área de pesquisa ficou bastante estagnada até meados dos anos 90, quando voltou a receber atenção significativa dos pesquisadores [Hong98]. Segundo Zhao, [Zhao03] isso pode ser evidenciado pelo surgimento de várias conferências sobre o tema — *International Conference on Automatic Face and Gesture Recognition* (AFGR), em 1995, e *International Conference on áudio and Video-Based Authentication* (AVBPA), em 1997 — e o surgimento de diversos sistemas de avaliação empírica das técnicas de reconhecimento visual — FERET, em 1998 [Phillips98]; XM2VTS, 1999; FRVT 2000 [Phillips00], 2000; e FRVT, 2003 [Phillips03].

Zhao [Zhao03] atribui este crescente interesse a basicamente dois fatores: larga aplicabilidade dessas técnicas em sistemas comerciais e de segurança pública e grande disponibilidade de recursos e tecnologias que se tornaram viáveis nos últimos 30 anos. Atualmente, existem diversas autores que propõem várias técnicas para o problema relatado. Dentre eles:

• Sung [Sung94] — abordagem baseada em aprendizado por exemplos, formando clusters com o padrão de distribuição das faces, representando faces e não faces.

• Rowley [Rowley98] — uso de técnicas de redes neurais.

• Peng [Peng05] — combina várias técnicas para determinar o centro dos olhos, sendo voltada especificamente para a detecção sem presença de óculos.

• Waring [Waring05] — utiliza histogramas espectrais e SVM.

• Viola e Jones [Viola01] — utilizam imagem integral e combinação de classificadores.

A técnica escolhida para a implementação deste trabalho foi à proposta por Viola e Jones [Viola01]. Além de ser das mais rápidas — uma vez que o nível de performance na detecção da face é crucial para o modelo proposto —, oferece alto índice de acerto, com baixa taxa de falsos positivos. Além disso, esta técnica permite utilizar a mesma abordagem para a tarefa de detecção de mãos, também necessária neste trabalho. Nas próximas páginas, detalharemos esta técnica.

3.2 Detecção de Faces em Tempo Real

A abordagem proposta por Viola e Jones [Viola01b] é dos métodos mais usados e citados pela comunidade cientifica. Ela traz três grandes inovações em relação a outras técnicas:

1 — Representação da imagem de entrada em forma de imagem integral, permitindo que as características (features) usadas pelo detector sejam processadas em uma única passagem pela imagem.

2 — Construção de um classificador simples e eficiente, baseado em uma adaptação do Adaboost, que seleciona um pequeno número de características criticas em um conjunto muito grande de características possíveis.

3 — Método para combinar classificadores em cascata usando adaptação do AdaBoost, que permite descartar imagens de fundo de forma rápida e eficiente.

3.2.1 Imagem Integral

As características (*features*) neste modelo são a soma dos valores dos *pixels* na escala de cinza em determinada região da imagem. A maior motivação para o uso de características — em vez dos *pixels* do objeto procurado (no caso a face) — está no fato do número das primeiras serem bem menor que os dos segundos em um objeto, o que torna o processamento muito mais rápido.

A rotina de detecção está baseada nas principais características do objeto, previamente extraídas em fase de treinamento. Tais características são suficientemente distintivas para diferenciar um de outro, pois cada conjunto de características encontradas em um objeto será altamente distintivo em relação ao conjunto de características encontradas em outro objeto diferente.

Viola e Jones [Viola01] propõem a utilização de três formatos, conhecidas como características *haar*:

> **Característica de dois retângulos :** Valor definido pela diferença entre a soma dos valores dos *pixels* de duas regiões retangulares e adjacentes do mesmo tamanho.
>
> **Característica de três retângulos:** Soma dos valores dos *pixels* de um retângulo central menos a soma dos valores dos *pixels* de dois retângulos externos.
>
> **Característica de quatro retângulos**: Diferença entre os valores dos pares diagonais de retângulos.

A figura 7 exemplifica os três formatos de características haar.

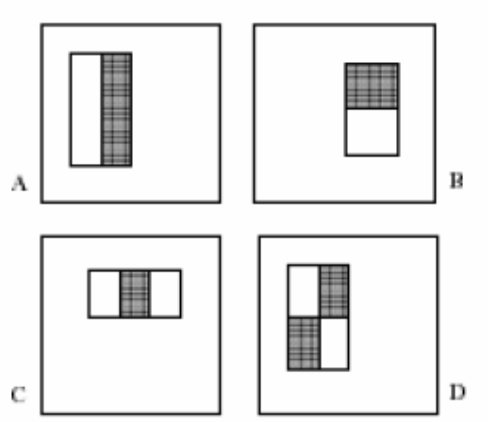

Figura 7 — Formato das características Harr. (a) — Dois retângulos na vertical ; (b) — Dois retângulos na horizontal; (c) — Três retângulos; (d) — Quatro retângulos.
Fonte : Adaptado de [Viola01]

Com o objetivo de otimizar o cálculo das características harr, Viola e Jones propõem o uso de uma representação intermediária da imagem, representação chamada de Imagem Integral.

Nesta forma de representação, cada ponto (x, y) da imagem contém o somatório da origem da imagem até sua localização. Esta representação pode ser calculada em uma única passada na imagem original usando a equação:

$$ii(x, y) = \sum_{x' \leq x,\, y \leq y} i(x', y')$$

Sendo, *ii* (x,y) é o valor da imagem integral num ponto (x,y) da imagem e *i* (x', y') o valor do pixel na escala de cinza no ponto (x', y') [Viola01].

Uma vez calculado o valor da imagem integral, é possível encontrar o valor de qualquer área retangular utilizando apenas os quatro pontos dos vértices da área desejada.

A figura 8 mostra um exemplo do cálculo de uma área usando imagem integral.

• O ponto P1 é a soma dos *pixels* da área A;
• O ponto P2 é a soma dos *pixels* da área A com os *pixels* da área B;
• O ponto P3 é a soma dos *pixels* da área C com os *pixels* da área A; e,
• O ponto P4 é a soma dos *pixels* de todas as áreas.

Ou seja, a área B é dada por P2 – P1, a área C é dada por P3 – P1 e a área D é dada por (P4 + P1) – (P2 + P3)

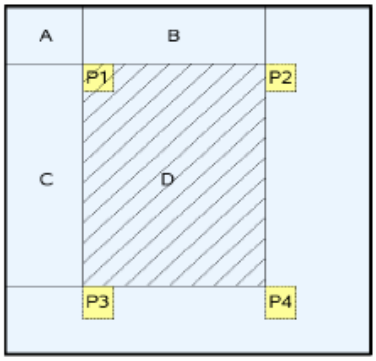

Figura 8 — Cálculo de área usando imagem integral
Fonte: Adaptado de [Viola01]

3.2.2 Seleção de Características

O algoritmo escolhido por Viola e Jones [Viola01] para a seleção das principais características do objeto procurado e treinamento foi adaptação do AdaBoost[1]. O algoritmo é treinado com imagens positivas e negativas do objeto procurado (no nosso caso, faces e não faces).

Segundo os autores, apesar de cada imagem possuir milhares de características, somente é necessário usar um número muito pequeno dessas características para diferenciar o objeto procurado.

A imagem do objeto é percorrida em várias escalas diferentes (subjanelas) e, a cada leitura, o algoritmo "fraco" é projetado para selecionar a característica de um único retângulo que melhor separa os exemplos positivos e negativos (face e não face).

Para cada característica, o algoritmo determina a função para um ponto de corte ótimo, de modo a minimizar o número de exemplos classificados incorretamente.
O classificador fraco (h_j x) consiste em uma característica (f_j), um ponto de corte (θ_j) e a paridade (p_j) indicando o sinal da desigualdade.

$$h_j(x) = \begin{cases} 1 & \text{if } p_j f_j(x) < p_j \theta_j \\ 0 & \text{otherwise} \end{cases}$$

A figura 9 é mostra exemplo de duas características extraídas pelo algoritmo proposto pelos autores para a face de uma pessoa. A figura 10, o pseudocódigo do algoritmo.

[1] O AdaBoost é um ensemble onde vários classificadores de baixo desempenho, chamados de classificadores fracos, são combinados para formar um classificador com melhor desempenho. Para maiores informações ver [Schapire00].

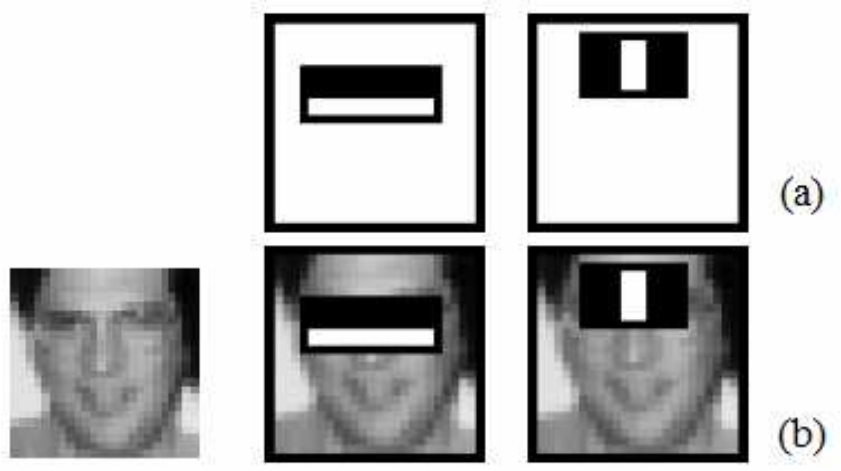

Figura 9 — Exemplo de seleção de características de uma face
(a) — Primeira e segunda características selecionada pelo Adaboost ; (b) — Características sobrepostas em uma face qualquer.
Observa-se que a primeira característica evidencia que a região dos olhos é mais escura que a região das bochechas; a segunda evidência a diferença de intensidade entre os olhos e o nariz.
Fonte: Extraído de [Viola01]

- Given example images $(x_1, y_1), \ldots, (x_n, y_n)$ where $y_i = 0, 1$ for negative and positive examples respectively.
- Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $y_i = 0, 1$ respectively, where m and l are the number of negatives and positives respectively.
- For $t = 1, \ldots, T$:
  1. Normalize the weights,
     $$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}$$
     so that w_t is a probability distribution.
  2. For each feature, j, train a classifier h_j which is restricted to using a single feature. The error is evaluated with respect to w_t, $\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$.
  3. Choose the classifier, h_t, with the lowest error ϵ_t.
  4. Update the weights:
     $$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$$
     where $e_i = 0$ if example x_i is classified correctly, $e_i = 1$ otherwise, and $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$.
- The final strong classifier is:
  $$h(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases}$$
  where $\alpha_t = \log \frac{1}{\beta_t}$

Figura 10 – Pseudocódigo do algoritmo de treinamento e seleção de características, proposto por Viola e Jones.
Fonte: Extraído de [Viola01]

3.2.3 Classificadores em Cascata

Após a fase de extração de características — e ainda na fase de treinamento do classificador — é feita uma combinação de classificadores em cascata que, assim como na fase de extração de características, utiliza uma adaptação do AdaBoost.

A estrutura em cascata reflete o fato de que, na maioria esmagadora das subjanelas, não existe o objeto procurado. O objetivo nas fases iniciais é, usando poucas características, descartar o máximo possível de subjanelas que não contém o objeto procurado.

No exemplo citado pelos autores, o classificador inicial está formatado para utilizar apenas duas características (mostradas na figura 10). O limiar de corte é definido de modo que possa minimizar a ocorrência de falsos negativos. É um classificador bastante flexível em relação ao formato do objeto procurado: face.

Segundo os autores, utilizando apenas estas duas características, é possível eliminar 60% das subjanelas que não contém o objeto procurado, mantendo 100% daquelas que podem conter o mesmo.

Todas as subjanelas não descartadas são processadas pelo próximo classificador, que utiliza cinco características e um limiar de corte mais rigoroso. Este consegue descartar 80% das janelas que não contém o objeto procurado e mantêm 100% das janelas que possam conter o objeto. Em seguida, as janelas não descartadas são processadas pelo próximo classificador, desta vez usando 20 características.

Este processo é repetido até o número de camadas (classificadores) definidas pelo usuário. A figura 11 mostra o esquema da detecção em cascata.

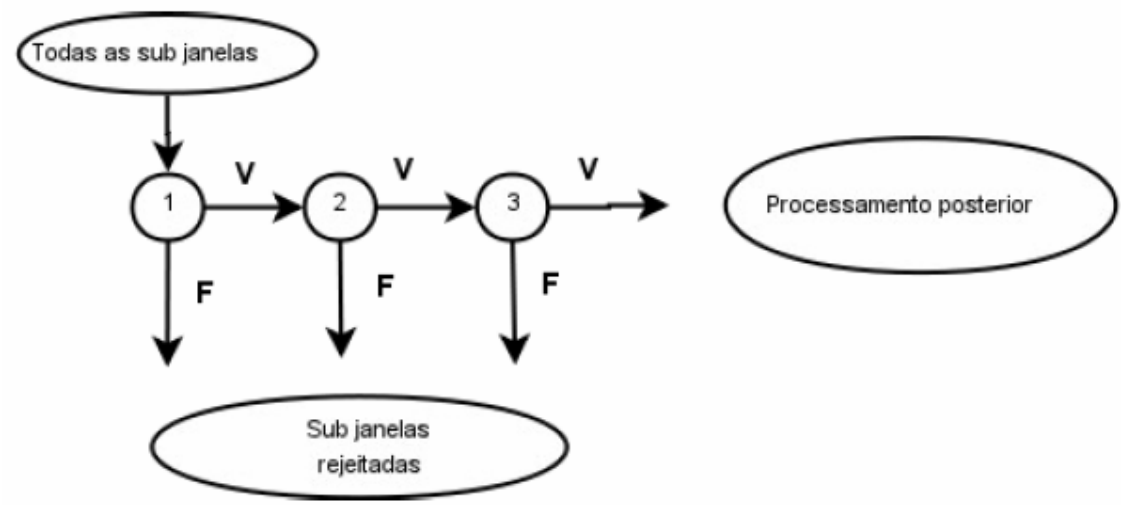

Figura 11- Esquema da detecção em cascata
Fonte: Traduzido de [Viola01]

O número de características usadas em cada camada e o número de camadas da cascata deve ser definido conforme a performance que se espera do detector, tentando sempre equacionar alta taxa de detecção, baixa taxa de falsos positivos e performance do tempo de processamento durante o uso do classificador.

Dado uma cascata de classificadores treinada, a taxa de falsos positivos é obtida através da equação:

$$F = \prod_{i=1}^{K} f_i,$$

Onde:

F — Taxa de falsos positivos;

K — Número de classificadores; e,

f_i — Taxa de falso positivo do i-ésimo classificador dos exemplos que conseguem passar por ele.

Da mesma forma, a taxa de detecção (verdadeiros positivos) é dada por:

$$D = \prod_{i=1}^{K} d_i,$$

Onde:

D — Taxa de detecção do classificador,

K — Número de classificadores; e,
d_i — Taxa de detecção do i-ésimo classificador dos exemplos que conseguem passar por ele.

Assim, pode-se definir uma arquitetura muito simples para produzir uma cascata de classificadores de acordo com as necessidades da aplicação. Informando as taxas mínimas de d_i e f_i, o número de características pode ser aumentado até que atinjam as taxas desejadas em cada estágio.

As taxas são obtidas testando-se o detector corrente sobre um conjunto de validação. Se as taxas obtidas não forem atendidas, uma nova camada pode ser adicionada à cascata. O resultado final do método proposto pelos autores é extremamente rápido e eficiente na detecção de objetos, e tem sido largamente utilizado em aplicações no mundo real. Porém, a fase de treinamento — em que é "ensinado" ao detector as formas do objeto a ser detectado — é extremamente lenta. No trabalho dos autores, publicado em 2001, o treinamento para reconhecimento de faces humanas levou várias semanas de processamento em uma máquina comum, com a velocidade de 466mhz.
Mesmo levando-se em conta o grande aumento na performance dos computadores atuais, esta tarefa ainda demandaria dezenas de horas, ou até mesmo dias.

Este fato não é impeditivo para o uso método, visto que a etapa de treinamento é realizada uma única vez para cada objeto. Além disso, é bastante comum encontrar arquivos de treinamentos já prontos para os objetos mais comumente usados, como face, mãos, olhos, cadeiras, mesas etc.

3.2.4 A Busca Pelo Objeto

A busca pelo objeto procurado é feita varrendo a imagem várias vezes, em tamanhos diferentes. Em cada passagem é analisada uma janela com tamanho diferente. Isso é chamado de "janela de busca". O usuário pode definir o tamanho mínimo desta janela de busca.

Como a imagem deve ser percorrida em várias locações, o usuário pode definir a quantidade de *pixels* β que a mesma será deslocada em cada locação. Ela é deslocada em β *pixels* na horizontal enquanto houver tamanho suficiente para o deslocamento no eixo x. E no mesmo valor

de *pixels* β no eixo y enquanto houver tamanho suficiente para o deslocamento no eixo y.

O valor de β deve ser escolhido com bastante critério: alto, fará com que menos janelas de buscas sejam processadas, aumentando a performance, porém diminuindo o número de objetos verdadeiros detectados; baixo, teria efeito contrário.

A imagem também é percorrida em várias escalas diferente. Assim como β define um valor para o deslocamento da janela, θ irá definir um fator de escalonamento da janela. Por exemplo: θ = 0.10 fará com que cada vez que a janela de busca terminar de percorrer o eixo x e o eixo y, a mesma seja aumentada em 10% antes de iniciar uma nova varredura da imagem.

Neste trabalho foi utilizada a técnica descrita nesta seção para a tarefa de detecção da face, detecção dos olhos e detecção das mãos. No capitulo referente ao experimento, detalharemos os parâmetros utilizados em cada caso.

4 CONTROLE DE ACESSO BASEADO EM BIOMETRIA FACIAL

A solução para controle de acesso proposta neste trabalho foi baseada em um algoritmo totalmente automático — detecta rosto e olhos sem o auxílio de ser humano —, com resposta em tempo real — que utiliza redes neurais sem peso para a tarefa de reconhecimento da face, — e técnicas probabilísticas para definir o controle de acesso

Neste capítulo, abordaremos os principais pontos relacionados a estes temas, bem como a arquitetura geral do sistema.

4.1 Reconhecimento de Face

Apesar de ser das tarefas mais corriqueiras executadas pelos seres humanos, o reconhecimento de face ainda é um grande desafio para pesquisadores na área de visão computacional e reconhecimento de padrões.

O processo natural de reconhecimento de faces em humanos é tão complexo que pesquisadores como Jain [Jain04] acreditam que a mesmo é processo dedicado do cérebro, existindo área exclusivamente dedicada a ele.

Além das características faciais, o cérebro utiliza ainda outras informações no processo de reconhecimento, inclusive considerando o contexto onde a mesma esta inserida, já em um processo de reconhecimento automatizado, o sistema pode contar com uma única informação, que é a face, o que torna o problema realmente desafiador.

O problema de reconhecimento de faces é um caso particular do reconhecimento de padrões e tem sido tratado com especial interesse pelos pesquisadores devido à sua grande aplicabilidade, desde aquelas relacionadas com segurança pública até como a que está sendo proposta neste trabalho.

Sendo um subproblema de reconhecimento de padrões, um sistema automático de reconhecimento de faces segue três fases fundamentais:

1. **Detecção e localização da face:** Esta fase é responsável por detectar a face em uma imagem ou seqüencia de vídeo e definir sua localização.

2. **Extração das características faciais:** A face em uma imagem ou seqüencia de vídeo, estará inserida no ambiente que a rodeia, que pode ser mais ou menos complexo. Esta fase é responsável por isolar a face o máximo possível do restante do ambiente, geralmente isso é feito “recortando” a imagem nos limites onde a face foi detectada, e extraindo as principais características que serão usadas na próxima fase.

3. **Reconhecimento da face**: As características da face que foram extraídas na fase anterior são comparadas com as características de faces previamente cadastradas no sistema, e retornado o padrão que mais de assemelha a face procurada.

Atualmente, existem vários tipos de abordagens diferentes para o problema de reconhecimento de faces. Neste trabalho, usaremos Redes Neurais sem Peso (RNSP), descritas na próxima seção.

4.2 Redes Neurais Sem Peso (RNSP)

Redes neurais artificiais é um modelo computacional com raízes em muitas disciplinas, como Matemática, Neurociência, Estatística, Física, Ciência da Computação e Engenharia (Haykin 2001). Trata-se de modelo computacional inspirado na estrutura paralela e fortemente conectada dos neurônios do cérebro humano.

A figura 12 mostra o esquema simplificado de um neurônio humano.

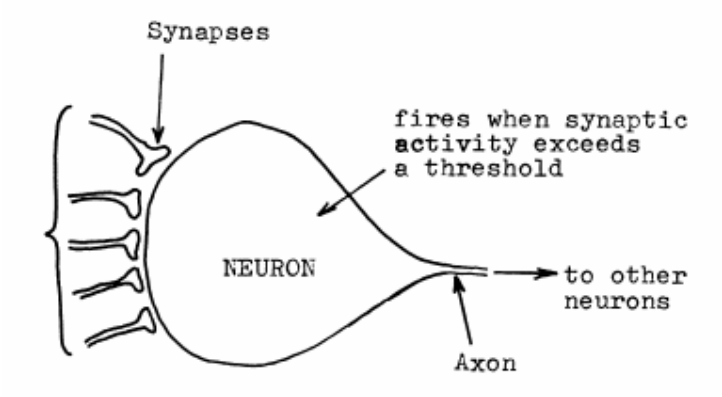

Figura 12 – Esquema simplificado de um neurônio
Fonte: Extraído de [Aleksander67]

Simulando as conexões dos neurônios do cérebro humano, as operações realizadas por uma rede neural artificial são feitas através de uma associação de elementos de processamento e conexões. O elemento básico do processamento de uma rede neural artificial é chamado de neurônio, ou nodo. A Figura 13 mostra o diagrama básico do funcionamento de um neurônio artificial:

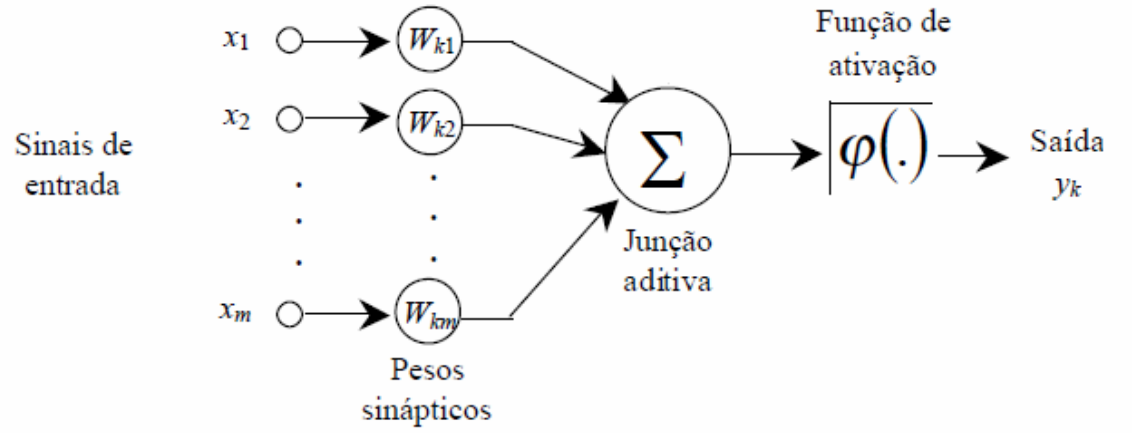

Figura 13 – Modelo de um neurônio artificial
Fonte: Adaptado de [Haykin01]

Conforme pode ser visto na figura 12, um neurônio artificial possui três elementos básicos: [Haykin01]:

- **Sinapses ou elos de conexão :** Cada Sinapse é caracterizada por um peso ou força própria. Especificamente, um sinal x*j* na entrada da sinapse *j* conectada ao neurônio *k*

é multiplicado pelo peso sináptico W*kj*. É importante notar a maneira como são escritos os índices do peso sináptico W*kj*. O primeiro índice se refere ao neurônio em questão, e o segundo, ao terminal de entrada da sinapse à qual o peso se refere.

- **Somador :** Utilizado para somar os sinais de entrada, ponderados pelas respectivas sinapses do neurônio.

- **Função de ativação :** Computa a saída y*k* em função da saída do somador. Geralmente, a função de ativação normaliza a saída entre o intervalo fechado [0, 1] ou alternativamente [-1, 1].

Rede neural sem peso (RNSP), também conhecida como rede baseada em RAM (*Random Access Memories*) é um tipo de rede neural que não armazena conhecimento em suas conexões, mas em memórias do tipo RAM dentro dos nodos da rede, ou neurônios.

Estes neurônios operam com valores de entrada binários e usam RAM como tabelas-verdade: as sinapses de cada neurônio coletam um vetor de bits da entrada da rede, que é usado como o endereço da RAM. O valor armazenado neste endereço é a saída do neurônio. O treinamento pode ser feito em um único passo e consiste basicamente em armazenar a saída desejada no endereço associado com o vetor de entrada do neurônio.

Apesar da sua notável simplicidade, as RNSP são muito efetivas como ferramentas de reconhecimento de padrões — oferecendo treinamento e testes rápidos e fácil implementação. No entanto, se a entrada da rede for muito grande, o tamanho da memória dos neurônios da RNSP torna-se proibitivo, dado que tem de ser igual a 2^n, onde n é o tamanho da entrada.

As redes *Virtual Generalizing RAM* (*VG-RAM*) são redes neurais baseadas em RAM que somente requerem capacidade de memória para armazenar os dados relacionados ao conjunto de treinamento.

Os neurônios *VG-RAM* armazenam os pares entrada-saída observados durante o treinamento, em vez de apenas a saída. Na fase de testes, as memórias dos neurônios *VG-RAM* são pesquisadas mediante a comparação entre a entrada apresentada à rede e todas as entradas nos pares entrada-saída aprendidos.

A saída de cada neurônio *VG-RAM* é determinada pela saída do par cuja entrada é a mais próxima da entrada apresentada. A métrica adotada pelos neurônios *VG-RAM* é à distância de *Hamming*, que calcula o número de bits diferentes entre dois vetores de bits de igual tamanho.

A figura 14 ilustra a tabela-verdade de um neurônio *VG-RAM* com três sinapses (1, 2 e 3). A tabela-verdade possui três pares de entradas-saídas armazenados durante a fase de treinamento. Na fase de testes, quando é apresentada uma entrada à rede, o algoritmo calcula a distância de *Hamming* entre o vetor que está sendo apresentado à rede e todos os pares entrada-saída que possui armazenado, e retorna o valor armazenado na saída do neurônio que possuir a menor distância de *Hamming*.

No exemplo da figura 14, seria retornada a classe B, pois é a que possui a menor distância para o vetor de entrada.

	Sinapse 1	Sinapse 2	Sinapse 3	Saida Neuronio	Distância Hamming
Par Entrada/Saida 1	1	1	0	Classe A	2
Par Entrada/Saida 2	0	0	1	Classe B	1
Par Entrada/Saida 3	0	1	0	Classe C	3
Entrada (fase teste)	1	0	1		

Figura 14 – Exemplo da tabela verdade de um neurônio *VG-RAM*

Na próxima seção, detalharemos o uso das redes *VG-RAM* no problema de reconhecimento da face, técnica esta adotada neste trabalho.

4.3 Reconhecimento de Faces Com RNSP - *VG-RAM*

Neste trabalho, usamos *VG-RAM* para extração das características e reconhecimento da face. Graças à arquitetura da *VG-RAM* empregada, não é necessário extrair características específicas da face (olhos, nariz, boca etc.) para efetuar o reconhecimento. Em vez disso, usaremos neurônios especializados em monitorar regiões específicas da face.

Esta arquitetura usa uma única matriz bidimensional com m x n neurônios, onde cada neurônio $n_{i,j}$ possui um conjunto de sinapses W= (w_1,w_2,...w_w), conectadas a uma entrada bidimensional Φ da rede, composta de u×v elementos de entrada.

O padrão de cada conexão sináptica de neurônio $n_{i,j}$ e $\Omega_{i,j,\sigma}$ (W) segue uma distribuição normal bidimensional com variância σ^2 centrada no pixel $\varphi_{\mu k,\mu l}$, onde μk = i.u/m e μl = j.v/n. Isto é: as coordenadas *k* e *l* dos elementos de Φ aos quais $n_{i,j}$ conecta-se via W seguem as PDF's (Probabilistic Density Function):

$$\omega_{\mu_k,\sigma^2}(k) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(k-\mu_k)^2}{2\sigma^2}} \qquad (1)$$

$$\omega_{\mu_l,\sigma^2}(l) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(l-\mu_l)^2}{2\sigma^2}} \qquad (2)$$

σ é um parâmetro de nossa rede neural. Assim, cada neurônio $n_{i,j}$ monitora uma região especifica da entrada Φ , sendo "especializado" em características da face mapeadas para aquela região.

As sinapses de uma rede *VG-RAM* conseguem ler apenas um bit na sua entrada (0 ou 1). Para possibilitar o uso com imagens — quando um pixel pode assumir uma série de valores —, usamos células *Minchinton* [Mitchell98], em que cada sinapse, w_t, forma uma célula *Minchinton* com a próxima w_{t+1} ($w_{|W|}$ forma uma célula *Minchinton* com w_1).

Cada uma destas células *Minchinton* retorna 1 se a sinapse w_t está conectada a um elemento da entrada Φ, φ_{ij}, cujo valor, x_t, seja maior que o valor do elemento φ_{kl} ao qual a sinapse w_{t+1} está conectada; caso contrário, retorna zero.

A figura 15 mostra um diagrama da arquitetura proposta.

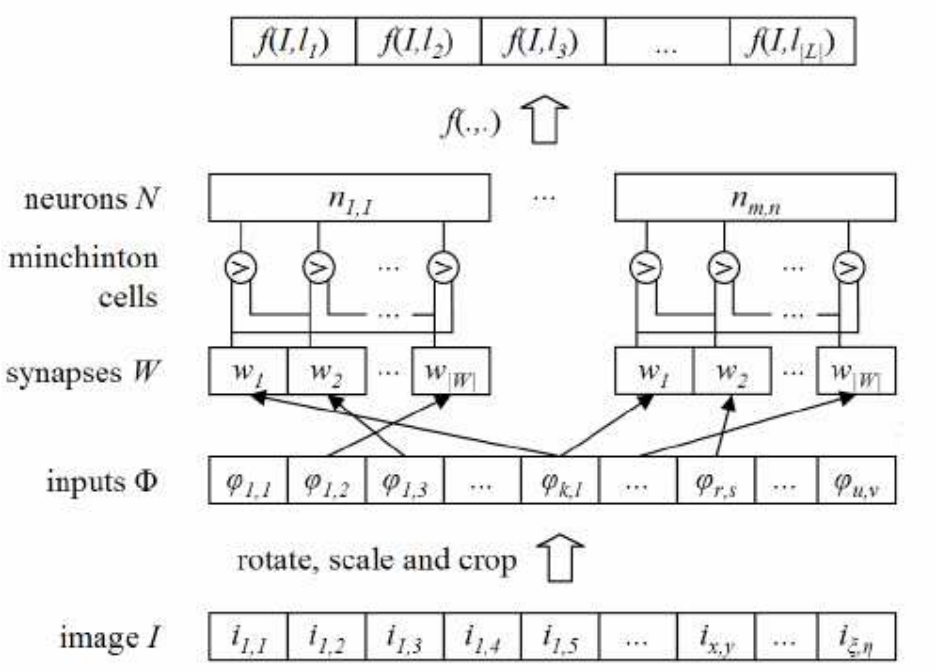

Figura 15 – Diagrama da Arquitetura da *VG-RAM*
Fonte: Extraído de [DeSouza08]

Para que uma imagem de entrada possa ser usada na *VG-RAM*, é necessário que a mesma seja adaptada para que seu tamanho fique compatível com a entrada da rede. Isso é feito de forma automática, seguindo os seguintes passos:

- O Sistema localiza e marca a posição da face na imagem
- De posse da localização da face, localiza os olhos na imagem
- Baseado na posição da face e dos olhos, rotaciona a imagem para que os olhos fiquem alinhados, escalona para o tamanho da entrada da rede e recorta a mesma;
- Antes de a imagem ser enviada à rede, é aplicado um filtro gaussiano para reduzir as imperfeições geradas pelas transformações.

A figura 16 mostra este processo:

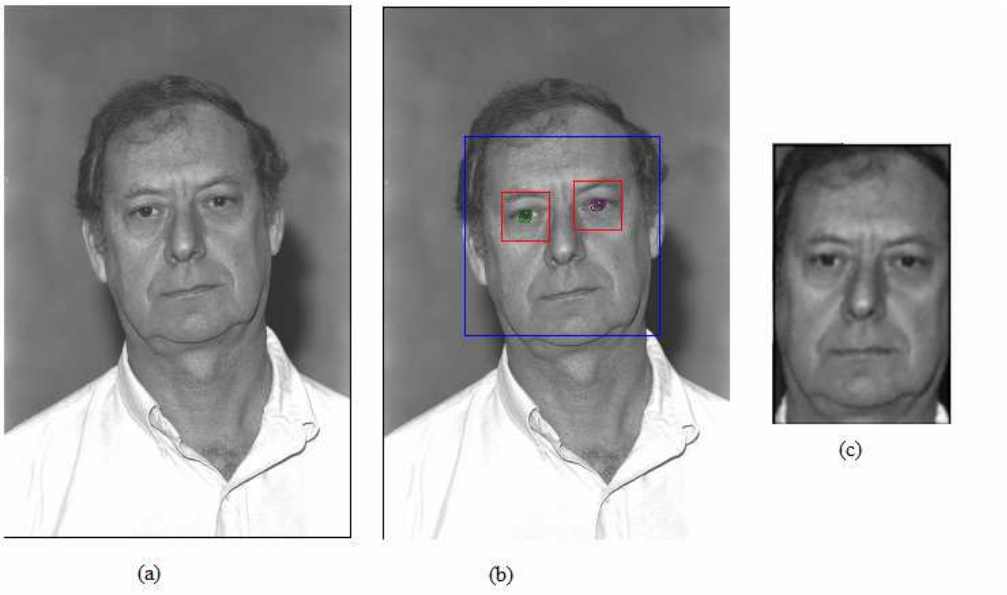

Figura 16 – Processo de aquisição da imagem: (a) Imagem original; (b) Após detecção da face e dos olhos; e, (c) imagem recortada, rotacionada e filtrada.

Para que a rede *VG-RAM* seja usada, ainda precisa ser treinada. Nesta fase, uma imagem I_x da pessoa $_P$ — depois de passar pelas transformações descritas antes — tem todos seus *pixels* copiados para a entrada da rede Φ e todos seus neurônios de saída tem seu valor L_P associados com a face da pessoa $_P$. Este processo é repetido para todas as imagens da pessoa $_P$ e, sucessivamente, para todas as pessoas no conjunto de treino.

4.4 Visão Geral da Solução Proposta

O protótipo funcional do sistema de controle de acesso desenvolvido neste trabalho foi projetado para trabalhar tanto com imagens de vídeo, como com imagens estáticas (fotos).

Quando operando com imagens de vídeo, o sistema analisa continuamente a imagem em busca de uma mão – chave escolhida para solicitação de acesso a determinado recurso – se a mesma for encontrada, o sistema desencadeia o seguinte processo:

a) Detecção e localização da face e olhos na imagem capturada;
b) Processamento da imagem;
c) Reconhecimento da face capturada;

d) Decisão baseada no limiar de decisão adotado, se a face de entrada pertence ou não a um usuário “conhecido”, ou seja, que esta em sua base de conhecimento.

Quando operando com imagens estáticas (fotos), o sistema, após receber uma imagem de entrada com a face de um suposto usuário, inicia o mesmo processo definido acima.

A figura 17 mostra um fluxograma com a visão geral da solução proposta:

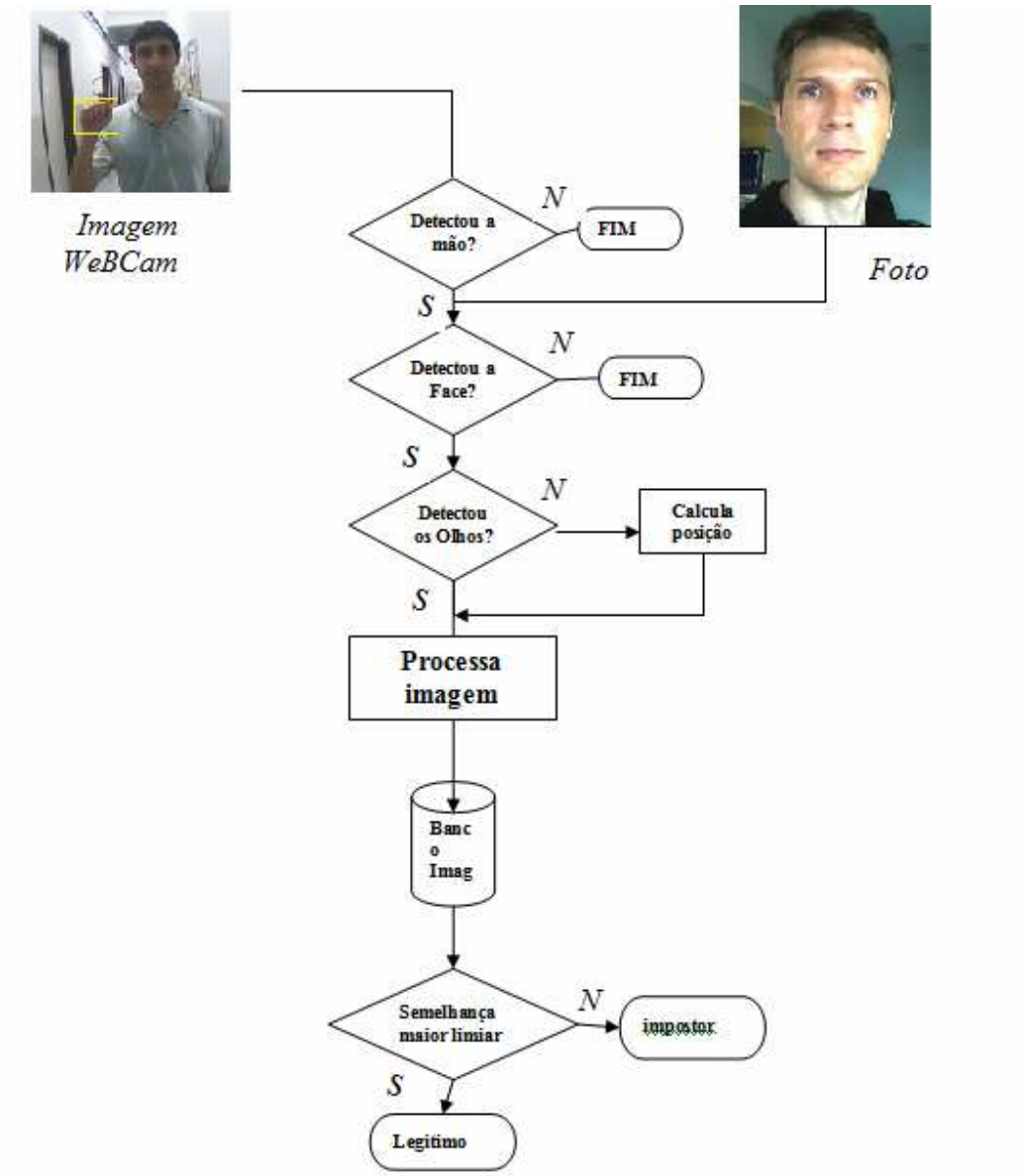

Figura 17 — Fluxograma do processo de controle de acesso

A figura 18 mostra a tela do sistema com o resultado de uma solicitação de acesso.

Figura 18 – Tela do sistema com um usuário solicitando acesso ao LCAD.

Na próxima seção iremos detalhar os principais pontos deste processo.

4.4.1 Detecção da Mão

Como na solução proposta neste trabalho não existe identificação prévia do indivíduo através de chave de identificação solicitando acesso a determinado recurso, a chave escolhida para funcionar como solicitação de acesso foi à mão.

Sempre que o mesmo quiser acesso a determinado recurso (no caso do projeto, as dependências do LCAD – Laboratório de Computação de Alto Desempenho da UFES) deve se posicionar em frente à câmara e levantar a mão fechada. Uma vez detectada a mão na imagem, o algoritmo dá inicio ao processo de identificar o usuário como “legítimo” ou “impostor”.

Para a tarefa de detecção da mão foi utilizado a abordagem para detecção de objetos proposta por Viola e Jones, e detalhada no capítulo 2.

Após testes empíricos de calibração do sistema, foi definido um valor de 5% para o escalonamento da janela de busca (seção 3.1.4) e um tamanho mínimo de 10x10 *pixels* para o objeto.

Com os parâmetros acima, conseguimos uma detecção de aproximadamente 91% — com uma taxa de *FAR* em torno de 10% —, o que atende as necessidades do projeto.

4.4.2 Detecção da Face e Olhos

Na tarefa de detecção da face e olhos também usamos a abordagem de Viola e Jones. Nos testes de calibração do sistema, definimos uma janela de busca com escalonamento de 3% e tamanho mínimo de 30x30 *pixels* para a face. Para os olhos, estes valores foram definidos em: 2% de escalonamento e tamanho mínimo de 15 x 15.

Na detecção da face, os valores mostraram-se perfeitos, sendo possível detectar 100% das faces nas imagens treinadas com uma taxa de *FAR* próxima de zero. Nos olhos, a taxa de detecção ficou em torno de 85% a 92%, com uma taxa de *FAR* em torno de 12%. Esta baixa performance se deve basicamente ao fato de vários modelos estarem usando óculos ou de olhos parcialmente fechados nas imagens.
Visto que as coordenadas dos olhos são ponto vital na tarefa de identificação da face, estes números não atendiam as necessidades do projeto, e nos levaram a adotar uma segunda forma para calcular a posição dos olhos, baseada em geometria facial.

O primeiro passo foi definir se o processo principal conseguiu detectar a posição dos olhos, e se estas coordenadas são coerentes. Para isso, adotamos as seguintes regras:

1 — O número de olhos detectados deve ser igual a dois;

2 — A distância entre as coordenadas no eixo Y (altura) do olho direito e do olho esquerdo não pode ser superior a 30 *pixels*;

3 — A distância mínima entre as coordenadas no eixo X (largura) do olho direito e do olho esquerdo deve ser no mínimo 60 *pixels*; e,

4 — Se β é a altura total da face detectada, ambos os olhos devem possuir valor menor que $\beta/2$ para suas coordenadas Y.

Caso algumas dessas regras não fossem atendidas, é porque os olhos não foram detectados, ou foram detectados incorretamente. Neste caso, obtínhamos as coordenadas dos olhos com a seguinte fórmula:

Seja F(x,y) o ponto superior direito da face detectada, β sua largura e θ sua altura. A localização dos olhos é dada por:

$OD(x) = F(x) + \beta / 3.25$
$OD(y) = F(y) + \theta / 2.6$

$OE(x) = F(x) + \beta - (\beta / 3.25)$
$OE(y) = F(y) + \theta / 2.6$

É importante ressaltar que a definição “olho esquerdo” e “olho direito” é dada segundo a perspectiva do modelo, nunca do observador.

Estes valores foram aferidos em diversas medições empíricas, e mostraram bons resultados, retornando valor bem aproximado da localização exata dos olhos.

A figura 19 mostra duas imagens onde os olhos foram corretamente detectados usando características *haar*. A figura 20, casos onde este método falhou e foi utilizado o método alternativo.

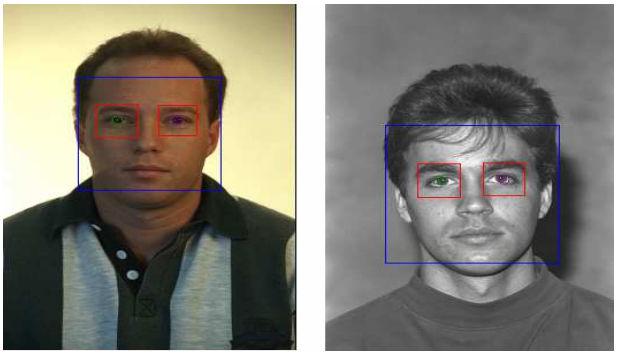

Figura 19 — Rosto e olhos detectados corretamente usando características *Haar*

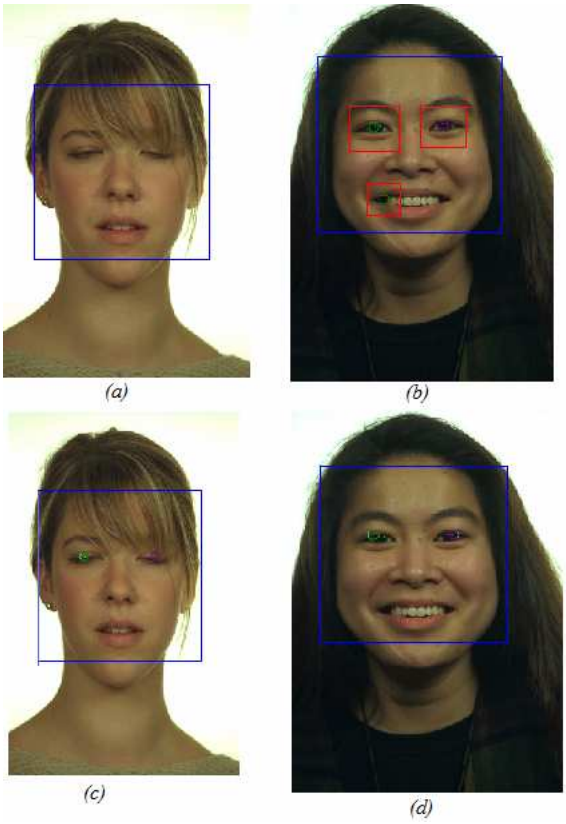

Figura 20: a — Olhos não detectados; b — Detectados incorretamente ; c e d — **D**etecção feita pelo método alternativo

No próximo capitulo, será detalhada a metodologia usada neste trabalho.

5 METODOLOGIA

Neste capítulo, descrevemos a metodologia usada no desenvolvimento desse trabalho. Serão detalhadas as bases de dados, definição dos conjuntos de avaliação, hardware e software, limiar de decisão e métrica utilizadas na avaliação de performance do sistema proposto.

5.1 Base de Dados

Os experimentos efetuados utilizaram dois tipos de imagens: estáticas e dinâmicas. Estas imagens tiveram duas fontes distintas, que estão detalhadas nas próximas seções.

5.1.1 Imagens Estáticas

Todas as imagens estáticas utilizadas neste trabalho foram extraídas da FERET. A FERET (Face Recognition Technology) é um dos mais conhecidos banco de imagens faciais, sendo que a versão à qual tivemos acesso para o desenvolvimento deste trabalho contava com 3.700 imagens de 991 indivíduos diferentes.

Cada indivíduo possui, no mínimo, duas fotos frontais e um número variado de fotos em ângulos que variam de 15 a 90 graus. Somente as fotos frontais foram consideradas. Existem ainda 228 indivíduos que possuem imagens com mais de um ano entre as seções de coleta das fotos.

Todas as fotos possuem um tamanho padrão (767x512 *pixels*), sendo que parte das mesmas é colorida e outra em tons de cinza (não fizemos distinção entre umas e outras). Todas colhidas em ambiente interno e com iluminação razoavelmente controlada. Os modelos podem apresentar mudança no visual entre as seções de fotos, podendo variar o uso de óculos, barba, maquiagem e diferentes penteados. As figuras 21 e 22 mostram exemplos dessas fotos.

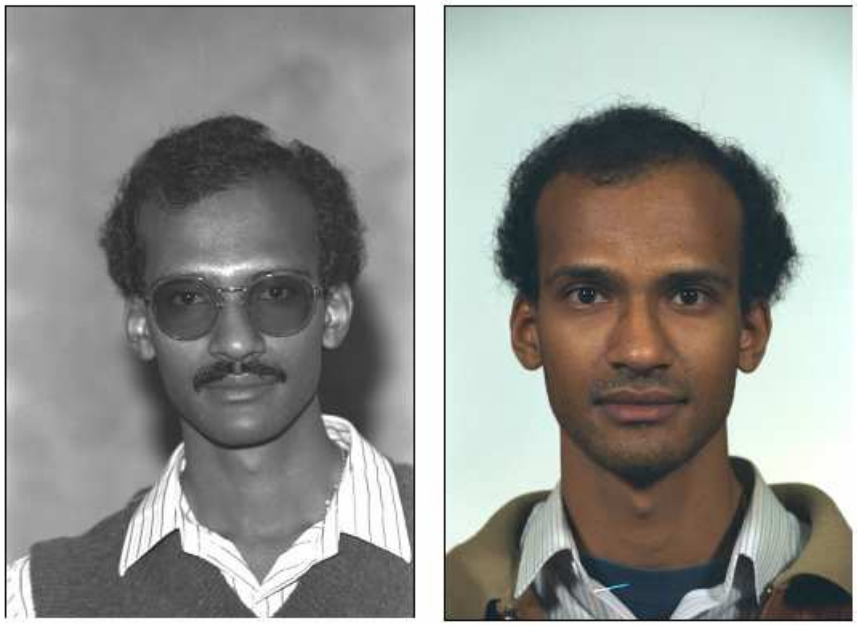

Figura 21 — Fotos da mesma pessoa adquiridas com mais de um ano de intervalo entre as seções.

Figura 22 — Fotos da mesma pessoa adquiridas com um curto intervalo entre as seções

A partir desta base, foram montados vários conjuntos de avaliação com as seguintes imagens e denominações:

1) Conjunto PessoasFB

Composto somente por imagens frontais, denominadas poses *fa* e *fb*, coletadas com intervalo máximo de dois meses entre as seções, com variação nas feições e/ou visual (óculos, barba, penteado etc.).

A partir deste conjunto foram criados dois subconjuntos: Teste e Treino.

• Subconjunto de Treino : 991 imagens da pose *fa* escolhidas sequencialmente do conjunto PessoasFB, sendo uma imagem por pessoa, não existindo duplicidade de pessoas.

• Subconjunto de Teste : 991 imagens da pose *fb* escolhidas sequencialmente do conjunto PessoasFB, sendo uma imagem por pessoas, não existindo duplicidade de pessoas. Todas as pessoas deste subconjunto estão no subconjunto de treino.

O conjunto de avaliação listado acima foi criado atendendo o protocolo Feret [Phillips00] para fins específicos na comparação com algoritmos de identificação de faces publicados na Feret.

Vale lembrar que este trabalho se refere ao problema de controle de acesso. A comparação com algoritmos que tratam do problema de identificação de face será efetuada unicamente para se ter algum parâmetro de performance do trabalho, visto que não foram encontradas publicações com performance especificas de controle de acesso via reconhecimento facial.

2) Conjuntos usados nos teste de controle de acesso

Foram criados mais três conjuntos de avaliação, usados especificamente para aferir a performance do controle de acesso neste trabalho. Estes conjuntos foram denominados CA1, CA2, CA3 e seguem o seguinte padrão:

Conjunto CA1 : A partir da base de dados FERET, foram extraídas duas imagens frontais (poses *fa* e *fb*) de cada uma das 991 pessoas constantes na base, e montados 10

subconjuntos denominados *folder*. Cada *folder* possui dois subconjuntos: Treino e Teste, assim definidos:

• Subconjunto de Treino: 50 pessoas (pose *fa*) colhidas sequencialmente do conjunto CA1 entre as pessoas ainda não usadas nos subconjuntos de testes dos *folders* anteriores. Cada pessoa possui uma única imagem e não existe a duplicação de pessoas. Assim sendo, o subconjunto de treino do primeiro *folder* foi composto pelas primeiras 50 pessoas do conjunto CA1; o segundo, pelas pessoas da posição 101ª à 150ª — e assim sucessivamente.

• Subconjunto de Teste: 100 pessoas (pose fb) colhidas sequencialmente do conjunto CA1, sendo uma imagem por pessoa, sem duplicidade de pessoas. Assim sendo, o subconjunto de teste do primeiro *folder* foi composto pelas primeiras 100 pessoas do conjunto CA1; o do segundo, pelas pessoas da 101ª à 200ª posição — e assim sucessivamente. Cada subconjunto de teste $T(_i)$ possui as mesmas 50 pessoas do subconjunto de treino $TR\ (_i)$ — com imagens diferentes — e 50 outras pessoas diferentes. O conjunto de teste do último *folder* possui somente 91 pessoas.

Conjunto CA2 : A partir da base de dados FERET, foram extraídas duas imagens frontais (poses *fa* e *fb*) de cada uma das 991 pessoas constantes da base e com as mesmas foram montados cinco subconjuntos denominados *folders*. Cada *folder* possui dois subconjuntos: Treino e Teste, assim definidos:

• Subconjunto de Treino: 100 pessoas (pose *fa*) colhidas sequencialmente do conjunto CA2 entre as ainda não usadas nos subconjuntos de testes dos *folders* anteriores. Cada pessoa possui uma única imagem e não existe duplicidade de pessoas. Assim, o subconjunto de treino do primeiro *folder* foi composto pelas primeiras 100 pessoas do conjunto CA2; o segundo, pelas pessoas da 201ª à 300ª posição — e assim sucessivamente.

• Subconjunto de Teste: 200 pessoas (pose *fb*) colhidas sequencialmente do conjunto CA2, sendo uma imagem por pessoa, sem duplicidade de pessoas. Assim, o subconjunto de teste do primeiro *folder* foi composto pelas primeiras 200 pessoas do conjunto CA2; o do segundo, pelas pessoas da 201ª à 400ª po-

sição — e assim sucessivamente. Assim, cada subconjunto de teste T(i) possui as mesmas 100 pessoas do subconjunto de treino TR (i) — com imagens diferentes — e 100 outras pessoas diferentes. O conjunto de teste do último *folder* possui somente 191 pessoas.

Conjunto CA3 : A partir da base de dados FERET, foram extraídas duas imagens frontais (poses *fa* e *fb*) de cada uma das 991 pessoas constantes da base e com as mesmas foram montados dois *folders*. Cada *folder* possui dois subconjuntos: Treino e Teste, assim definidos:

• Subconjunto de Treino: O subconjunto de treino do primeiro *folder* é formado pelas primeiras 200 pessoas (pose *fa*) colhidas sequencialmente do conjunto CA3 entre as pessoas ainda não usadas nos subconjuntos de testes dos *folders* anteriores. Cada pessoa possui uma única imagem e não existe duplicidade de pessoas. O segundo subconjunto de treino é composto pelas pessoas da 401ª à 600ª.

• Subconjunto de Teste: O subconjunto de testes do primeiro *folder* é formado pelas primeiras 400 pessoas (pose *fb*) colhidas sequencialmente do conjunto CA3, sendo uma imagem por pessoa, sem duplicidade de pessoas. O subconjunto de teste do segundo *folder* é formado pelas 591 pessoas restantes.

5.1.2 Imagens de vídeo

Através das imagens de vídeo capturadas pela webcam, foi criado o conjunto de avaliação denominado **PessoasCAM**$_2$ e este foi dividido em três subconjuntos:

• Subconjunto de Treino: As imagens deste subconjunto foram feitas com a cooperação do indivíduo, que se posicionava olhando para a câmara, a uma distância de aproximadamente dois metros. Possui 29 indivíduos, com uma imagem de cada, capturadas de forma controlada, com a pessoa imóvel. Os pertencentes a este subconjunto são considerados usuários “legítimos”.

• Subconjunto de Testes I: Criado em tempo real, a partir dos acessos ao LCAD da UFES. Para entrar, a pessoa deveria posicionar-se olhando para a webcam e levantar a mão direita para solicitar a entrada. O sistema processava a imagem e liberava ou não o acesso, registrando neste subconjunto a imagem do usuário e a decisão tomada. Este subconjunto possui tentativas de acesso tanto de usuários "legítimos" — que estão no subconjunto de treino — como de usuários "impostores" — que não estão no subconjunto de treino.

Cada acesso efetuado por um usuário é registrado de forma independente, logo, poderá haver várias imagens de um mesmo usuário neste subconjunto, tantas quanto o número de acessos efetuados pelo mesmo. Ao todo, este conjunto possui 46 acessos.

As figuras 23a e 23b mostram imagens usadas nos subconjuntos de Treino e Teste I, respectivamente.

Figura 23 - a — Imagem principal usada no treino; b — Imagem de um procedimento real, em que a pessoa solicita acesso ao LCAD

O que se conseguiu de cada uma das bases está descrito e comentado no capítulo referente aos resultados.

5.2 Hardware

Foram usados dois ambientes de trabalho, com as seguintes configurações:

Configuração 1: Usada no desenvolvimento, processamento, testes e treinos com as imagens estáticas.

- Notebook HP, com processador Pentium Core 2 Duo de 2 Gigahertz de velocidade e 4 Gigabytes de memória RAM.
- Webcam marca HP com definição de 1.2 Megapixels.
- Sistema operacional Windows Vista Professional

Configuração 2: Usada no processamento e reconhecimento das imagens de vídeo capturadas nos testes efetuados no LCAD.

- Desktop Dell, com processador Athom com 1 Gigahertz de velocidade e 1 Gigabyte de memória RAM.
- Webcam DMH – Hard Cam com definição de 300 *pixels*.
- Sistema Operacional Windows XP Professional.

5.3 Software

Neste trabalho foram utilizados somente softwares de livre distribuição (freeware) ou versões gratuitas limitadas (shareware) de softwares comerciais.

EMGU CV

Biblioteca de processamento de imagens de distribuição gratuita (freeware) e código fonte aberto. Ferramenta foi escolhida pela facilidade que oferece na implementação dos métodos de detecção de objetos (ver capítulo 2) e manipulação de imagens — e pelo fato de ter o código fontes aberto, o que permite a modificação de certas funções para os objetivos específicos deste trabalho.

Disponível em http://sourceforge.net/projects/emgucv/.

Microsoft Visual C# 2008 – Express Edition

O Visual C# (pronuncia-se C sharp) é uma linguagem visual de desenvolvimento orientada a objetos oferecida pela Microsoft. A versão Express Edition é distribuída em forma de Shareware, sob limitações, como, por exemplo, não aceitar acesso a banco de dados remoto ou suporte a dispositivos móveis. Nenhuma delas foi impeditiva neste projeto. O Visual C# foi escolhido pela produtividade oferecida e também pelo fato do mesmo estar sendo usado como ferramenta pelo LCAD em outras linhas de pesquisa na área de visão artificial, o que permite uma fácil troca de conhecimento entre as pesquisas.
Disponível em http://msdn.microsoft.com/pt-br/vcsharp/default.aspx.

Microsoft SQL Server 2008 – Express Edition

O SQL Server é um gerenciador de banco de dados relacional oferecido pela Microsoft. A versão Express Edition, distribuída em forma de shareware, apresenta limitações que tiveram que ser contornadas para o desenvolvimento deste trabalho. A principal delas é a de tamanho que esta versão impõe ao banco de dados, sendo possível gerenciar o máximo de 2 Gigabytes. Devido à grande quantidade de fotos (aproximadamente 3.700) e ao tamanho das mesmas (aproximadamente, 1 Megabyte cada um), foi necessário separar as mesmas em dois bancos diferentes. Este fato não impactou em nenhum dado ou estatística publicada. A escolha do SQL Server se deve basicamente ao fato do mesmo de ser um dos poucos bancos de dados de grande porte a oferecer versões gratuitas e com facilidade de integração do mesmo com a linguagem escolhida.
Disponível em http://www.microsoft.com/sqlserver/2008/pt/br/default.aspx.

5.4 Limiar de Decisão

Conforme citado no capítulo 2, classificadores que tratam do problema de controle de acesso devem tomar uma decisão binária de conceder ou não acesso a determinado recurso. Para is-

so, usam um limiar de decisão — ou match limiar — que separa usuários “legítimos” de “impostores”.

Neste trabalho, este limiar foi baseado em decisões probabilísticas, a partir do Teorema de Bayes, que permite calcular uma probabilidade condicional a *posteriori* baseado em probabilidades já conhecidas (a *priori*).

A probabilidade condicional relaciona a ocorrência de um evento (a) à ocorrência de outro (b)

$$P(a|b) = x$$

A equação acima pode ser lida como “a probabilidade de acontecer o evento a dado que **aconteceu** o evento b é x.”
Exemplo:

Seja “a” o evento “dentes cariados”
Seja “b” o evento “dor de dente”

$$P(a|b) = 0.8$$

A probabilidade em é lida como “dado que um paciente esta com dor de dente, o mesmo tem a probabilidade de 0,8 de estar com dentes cariados.”

Baseados nas regras básicas de probabilidade condicional, o Teorema de Bayes é formulado como [Magalhães02]:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

P(A|B) é a probabilidade procurada. Define as chances de acontecer o evento A, dado que ocorreu o evento B. Esta é a probabilidade a *posteriori*, ou seja, ainda não conhecida.

P (B|A) é a probabilidade de acontecer o evento B dado que ocorreu o evento A. Esta é uma probabilidade condicional; sua ocorrência esta relacionada à ocorrência do evento A. Deve ser previamente conhecida.

P(A) é a probabilidade de ocorrer o evento A. Ela é incondicional, pois não depende de outro evento. Deve ser previamente conhecida.

P(B) é a probabilidade de ocorrência do evento B. Ela é incondicional, pois não depende de outro evento. Deve ser previamente conhecida.

As probabilidades a *priori* dos parâmetros (PA/PB), P(A), e P(B) são obtidas através de tabela previamente criada durante a fase de calibração do sistema, onde, através de experimentos empíricos, é obtida a distribuição de usuários "legítimos" e "impostores" entre vários intervalos de crença.

A tabela 1 mostra esta separação de intervalos para o *folder* F1 do conjunto de avaliação CA1. A figura 24 mostra graficamente esta mesma distribuição.

Tabela 1 — Distribuição de usuários "legítimos" e "impostores" entre vários intervalos crença do classificador

Probabilidade de distribuição dos acessos			
Intervalo [..[	Usuários Legitimos	Usuários Impostores	Total Acessos
0 a 4 %	0,00%	0,00%	0,00%
4 a 6%	0,00%	4,00%	2,00%
6 a 8%	0,00%	32,00%	16,00%
8 a 10%	0,00%	36,00%	18,00%
10 a 12%	1,00%	10,00%	5,50%
12 a 14%	3,00%	9,00%	6,00%
14 a 16%	3,00%	5,00%	4,00%
16 a 18%	3,00%	3,00%	3,00%
18 a 20%	3,00%	1,00%	2,00%
20 a 22%	2,00%	0,00%	1,00%
22 a 24%	2,00%	0,00%	1,00%
24 a 26%	1,00%	0,00%	0,50%
26 a 28%	1,00%	0,00%	0,50%
28 a 30%	1,00%	0,00%	0,50%
30 a 32%	1,00%	0,00%	0,50%
32 a 34%	1,00%	0,00%	0,50%
32 a 36%	1,00%	0,00%	0,50%
36 a 38%	1,00%	0,00%	0,50%
38 a 40%	1,00%	0,00%	0,50%
40 a 42%	1,00%	0,00%	0,50%
42 a 45%	4,00%	0,00%	2,00%
45 a 50%	2,00%	0,00%	1,00%
50 a 60%	10,00%	0,00%	5,00%
60 a 70%	12,00%	0,00%	6,00%
>70%	46,00%	0,00%	23,00%
Total	50,00%	50,00%	100,00%

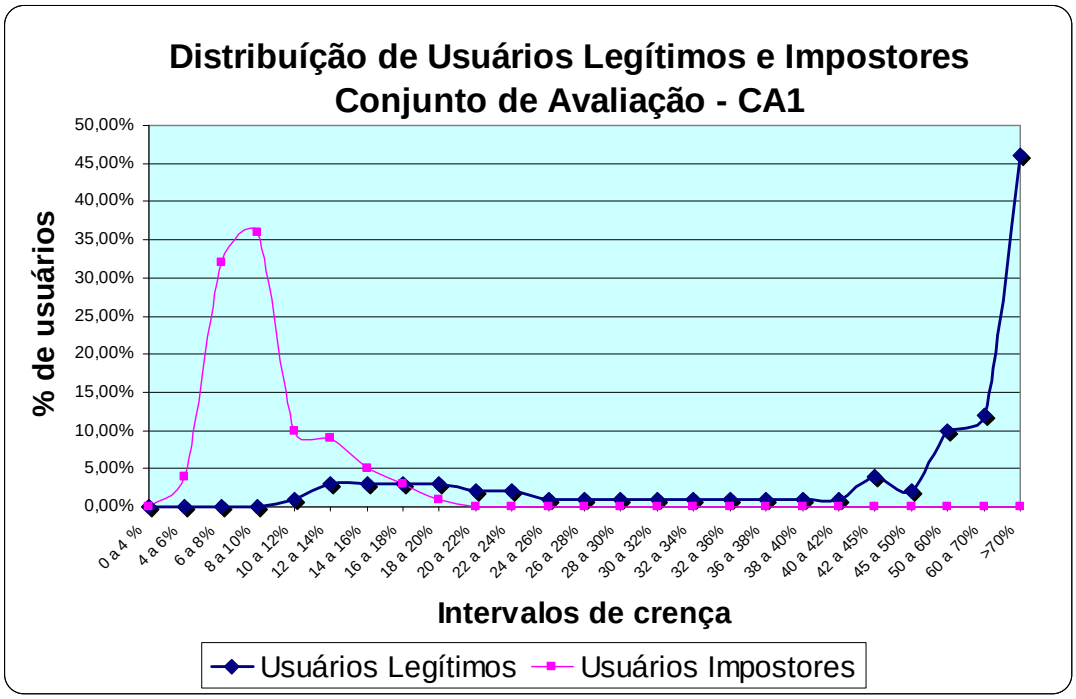

Figura 24 — Gráfico com a distribuição de "legítimos" e "impostores" para o conjunto de avaliação CA1.

No gráfico, é possível observar que, à medida que a crença do classificador aumenta, o número de usuários "impostores" diminui. Após os 8% de crença, acontece queda drástica no número de "impostores" — praticamente não existindo "impostores" nos intervalos de crença acima de 16%. Também é possível observar aumento drástico no número de usuários "legítimos" após o intervalo de crença de 45%.

O uso do Teorema de Bayes para calcular a probabilidade a *posteriori* de um usuário ser "legítimo", sendo que o classificador o classificou com tal, acontece da seguinte forma:

Em um exemplo hipotético, o classificador considerou o usuário "legítimo" com crença de 18%. O que procuramos saber é qual a probabilidade dele ser realmente "legítimo", ou seja, a probabilidade do classificador estar correto. Podemos formular a hipótese da seguinte forma:

• **A** - Indica a probabilidade incondicional do usuário ser "legítimo";

• **B** - Indica probabilidade incondicional de um usuário "legítimo" ou não ser classificado no intervalo de 18% a 20% (a tabela usa intervalo inicial fechado e final aberto); e,

• **P (B|A)** - Indica probabilidade do usuário ser classificado no intervalo de 18% a 20% dado que é “legítimo”.

Da tabela 1 temos:

• P (A) = 50,00%
• P (B) = 2,00 %
• P (B|A) = 3,00%

Aplicando o Teorema de Bayes:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)}$$

$$P\,(A|B) = \frac{0{,}03 \ * \ 0{,}50}{0{,}02} = 0{,}75$$

Assim sendo, podemos afirmar que a probabilidade do classificador estar correto e o usuário ser realmente “legítimo” é de 75,00%. Esta é a certeza final do classificador. A decisão de autenticar ou não dependerá do limiar de decisão adotado no sistema.

A tabela 2 mostra os intervalos de crenças definidos na tabela 1 com o acréscimo de uma coluna onde é mostrado o resultado do Teorema de Baeys para cada intervalo. O limiar de decisão é sempre confrontado com o valor desta coluna.

Tabela 2 —Intervalos de crença com a resposta final do classificador para o *folder* CA1-F1 após calcular a probabilidade a posteriori de um indivíduo ser “legítimo”.

Probabilidade de distribuição dos acessos				Probabildade a *Posteriori da certeza do classificador*
Intervalo [..)	Usuários Legitimos	Usuários Impostores	Total Acessos	
0 a 4 %	0,00%	0,00%	0,00%	0,00%
4 a 6%	0,00%	4,00%	2,00%	0,00%
6 a 8%	0,00%	32,00%	16,00%	0,00%
8 a 10%	0,00%	36,00%	18,00%	0,00%
10 a 12%	1,00%	10,00%	5,50%	9,09%
12 a 14%	3,00%	9,00%	6,00%	25,00%
14 a 16%	3,00%	5,00%	4,00%	37,50%
16 a 18%	3,00%	3,00%	3,00%	50,00%
18 a 20%	3,00%	1,00%	2,00%	75,00%
20 a 22%	2,00%	0,00%	1,00%	100,00%
22 a 24%	2,00%	0,00%	1,00%	100,00%
24 a 26%	1,00%	0,00%	0,50%	100,00%
26 a 28%	1,00%	0,00%	0,50%	100,00%
28 a 30%	1,00%	0,00%	0,50%	100,00%
30 a 32%	1,00%	0,00%	0,50%	100,00%
32 a 34%	1,00%	0,00%	0,50%	100,00%
32 a 36%	1,00%	0,00%	0,50%	100,00%
36 a 38%	1,00%	0,00%	0,50%	100,00%
38 a 40%	1,00%	0,00%	0,50%	100,00%
40 a 42%	1,00%	0,00%	0,50%	100,00%
42 a 45%	4,00%	0,00%	2,00%	100,00%
45 a 50%	2,00%	0,00%	1,00%	100,00%
50 a 60%	10,00%	0,00%	5,00%	100,00%
60 a 70%	12,00%	0,00%	6,00%	100,00%
>70%	46,00%	0,00%	23,00%	100,00%
Total	50,00%	50,00%	100,00%	

5.5 Métricas e Avaliações

O principal objetivo de um sistema de controle de acesso baseado em biometria facial é ter a capacidade de distinguir faces de pessoas “conhecidas”, previamente “ensinadas” ao sistema, como sendo usuários “legítimos”, de faces desconhecidas, nunca vistas, que seriam os usuários “impostores”.

Um sistema ideal daria acesso a determinado recurso a todas as faces “conhecidas”, e negaria a todas as “desconhecidas”. Para avaliar a performance no seu objetivo principal, são fundamentais quatro métricas:

1 — Taxa de Verdadeiros Positivos ou *Recall*;

2 — Taxa de Falsos Positivos ou *FAR* (*False Acceptance Rate*);

3 — Taxa de Verdadeiro Negativo; e,

4 — Taxa Falso Negativo.

1 e 3, quanto maior melhor; 2 e 4, quanto menor melhor. Para maiores detalhes sobre essas métricas e forma de apresentação das mesmas, consulte a seção 2.3.

6 EXPERIMENTOS E RESULTADOS

O objetivo destas avaliações é confirmar — ou refutar — a hipótese que é possível construir um sistema de controle de acesso eficiente baseado unicamente na biometria facial, e dentro de que limites esta hipótese é verdadeira.

Todos os experimentos foram realizados utilizando os conjuntos de avaliação descritos na seção 6.1.2 e 6.1.3, com seus resultados reportados em tabelas descritivas e em gráficos de curvas *ROC* (*Receiver Operating Characteristic*) usando as métricas descritas na seção anterior.

Em todos os experimentos nossa rede neural usou 1024 neurônios e 512 sinapses.

6.1 Controle de Acesso de 50 Usuários

Neste conjunto de testes, avaliamos a performance do sistema em controlar o acesso de 50 usuários a determinado recurso, ambiente similar ao encontrado em uma pequena empresa.

Para isso, foi usado o conjunto de avaliação CA1 conforme descrito na seção 6.1.1. Este conjunto é formado por imagens de 991 pessoas e foi dividido em nove *folders* com 100 pessoas cada e um *folder* com 91 pessoas.

Inicialmente, usamos o primeiro *folder* para treinar o sistema com as 50 primeiras pessoas contidas no mesmo. Em seguida, apresentamos ao sistema um conjunto de teste formado por todas as pessoas no *folder* — ou seja, 100 pessoas, sendo que 50 delas são as mesmas que foram apresentadas no treinamento (com imagens diferentes) e 50 são completamente desconhecidas. O desempenho ideal do sistema seria conceder acesso às 50 conhecidas e negar para as 50 desconhecidas.

Este primeiro *folder* será usado para montar a tabela com a distribuição de usuários “legítimos” e “impostores” entre vários intervalos crença do classificador (tabela 1) e para definir o limiar de decisão que será usado para este conjunto de teste. Usamos uma curva *ROC* para

ajustar o limiar de decisão, para que o mesmo ofereça o melhor desempenho possível. Uma vez ajustado, esse limiar se manterá inalterado para todos os *folders* do conjunto de avaliação. O resultado foi reportado separadamente para cada *folder*, sendo medido o número e taxa de Verdadeiro Positivo, Verdadeiro Negativo, Falso Positivo e Falso Negativo. O desempenho do sistema será a média dos resultados reportados em cada *folder*.

A figura 25 mostra a curva *ROC* com as taxas de *Recall* e *FAR* do *folder* CA1-F1 para todos os limiares de decisão.

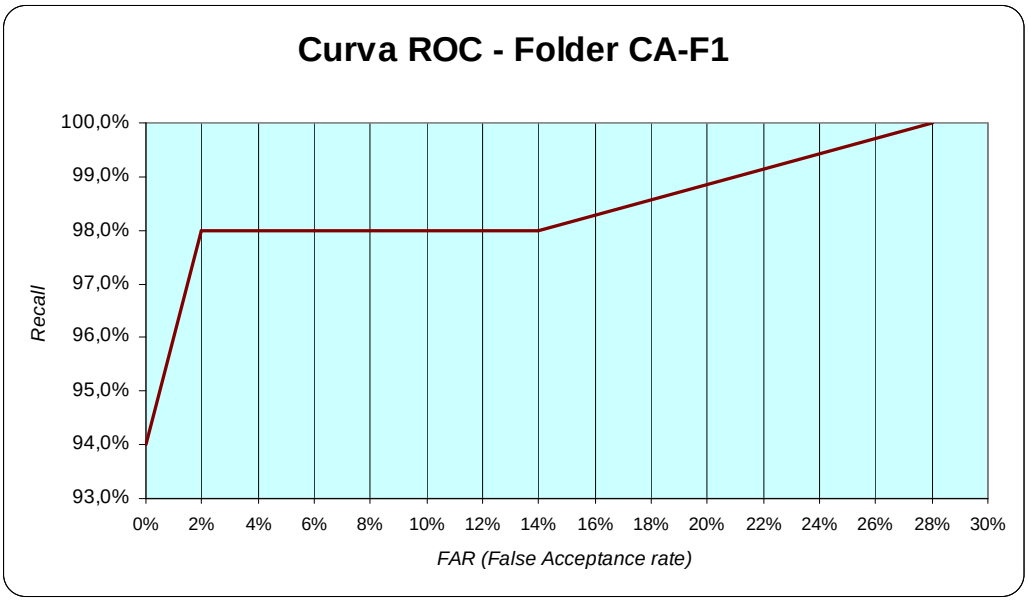

Figura 25 — Gráfico com a curva *ROC* para o conjunto de avaliação CA1-F1.

O eixo X do gráfico informa a taxa de *FAR* (*False Acceptance Rate*) ou falsos positivos. Este percentual indica a taxa de "impostores" que conseguiram acesso indevido a determinado recurso. O eixo Y informa a taxa de *Recall*, ou verdadeiros positivos, informando o percentual de usuários "legítimos" a que o sistema concedeu acesso a determinado recurso. Informalmente, pode-se considerar que, em um gráfico de curva *ROC*, quanto mais acima e mais à esquerda o ponto estiver, melhor é o seu desempenho.

No gráfico, é possível observar que, neste *folder*, o sistema conseguiu autenticar corretamente 94% dos "legítimos", com uma taxa zero de *FAR*. Com um *FAR* de 2%, é possível autenticar 98% "legítimos".

A tabela 3 mostra as taxas de *FAR* e *Recall* do *folder* CA1-F1 para todos os limiares de decisão. O Limiar de Decisão indica qual certeza mínima deve ser retornada pelo classificador para que o indivíduo seja considerado “legítimo” (seção 5.4).

Vale lembrar que o limiar de decisão é comparado sempre com a certeza retornada pelo classificador depois de calcular a probabilidade a *posteriori*, ou seja, após a aplicação do teorema de Bayes (Seção 5.4).

Tabela 3 — taxas de *FAR* e *Recall* para todos os limiares de decisão do *folder* CA1-F1

Limiar de Decisão	Taxa de Recall	Taxa de FAR
>=09 %	100,00%	28,00%
>=25%	98,00%	14,00%
>=35%	98,00%	6,00%
>=50%	98,00%	2,00%
>=70%	94,00%	0,00%

Conforme citado nas seções 6.4 e 2.3, o limiar de decisão adotado é muito importante, pois um usuário será considerado “legítimo” ou “impostor” baseado nele. Para os próximos subconjuntos de testes, fixamos nosso limiar de decisão em 50%, que, neste *folder*, retornou um *recall* de 98% com um *FAR* de apenas 2%.

Esta escolha é uma decisão de projeto e depende da aplicação. Para uma aplicação onde o foco fosse maximizar a segurança, seria mais interessante fixar o limiar de decisão em 70%. Assim, seria esperada uma taxa zero de *FAR*, mesmo que, para isso, fosse necessário conviver com uma taxa de falsos negativos de 6%.

Com o limiar de decisão estipulado, avaliamos o restante dos *folders* utilizando sempre um subconjunto de treino formado pelas 50 primeiras pessoas do *folder* e um subconjunto de testes formando por todo o *folder*.

Em todos os subconjuntos de testes (com exceção do *folder* 10) haverá sempre 50 pessoas conhecidas (“apresentadas” ao sistema na fase de treinamento) e 50 desconhecidas. A figura 26 mostra o resultado para os *folders* F2 a F10 do conjunto de avaliação CA1. Os *labels* “Total Positivos” e “Total Negativos” indicam, respectivamente, o número real de usuários “legítimos” e “impostores” no *folder*.

Folder CA1-2		
Total Positivos	50	
Total Negativos	50	
Total Folder	100	
Verdadeiro Positivo	46	92,00%
Verdadeiro Negativo	48	96,00%
Falso Positivo	2	4,00%
Falso Negativo	4	8,00%
Total	100	

Folder CA1-3		
Total Positivos	50	
Total Negativos	50	
Total Folder	100	
Verdadeiro Positivo	48	96,00%
Verdadeiro Negativo	47	94,00%
Falso Positivo	3	6,00%
Falso Negativo	2	4,00%
Total	100	

Folder CA1-4		
Total Positivos	50	
Total Negativos	50	
Total Folder	100	
Verdadeiro Positivo	46	92,00%
Verdadeiro Negativo	49	98,00%
Falso Positivo	1	2,00%
Falso Negativo	4	8,00%
Total	100	

Folder CA1-5		
Total Positivos	50	
Total Negativos	50	
Total Folder	100	
Verdadeiro Positivo	44	88,00%
Verdadeiro Negativo	49	98,00%
Falso Positivo	1	2,00%
Falso Negativo	6	12,00%
Total	100	

Folder CA1-6		
Total Positivos	50	
Total Negativos	50	
Total Folder	100	
Verdadeiro Positivo	49	98,00%
Verdadeiro Negativo	49	98,00%
Falso Positivo	1	2,00%
Falso Negativo	1	2,00%
Total	100	

Folder CA1-7		
Total Positivos	50	
Total Negativos	50	
Total Folder	100	
Verdadeiro Positivo	48	96,00%
Verdadeiro Negativo	46	92,00%
Falso Positivo	4	8,00%
Falso Negativo	2	4,00%
Total	100	

Folder CA1-8		
Total Positivos	50	
Total Negativos	50	
Total Folder	100	
Verdadeiro Positivo	48	96,00%
Verdadeiro Negativo	47	94,00%
Falso Positivo	3	6,00%
Falso Negativo	2	4,00%
Total	100	

Folder CA1-9		
Total Positivos	50	
Total Negativos	50	
Total Folder	100	
Verdadeiro Positivo	45	90,00%
Verdadeiro Negativo	47	94,00%
Falso Positivo	3	6,00%
Falso Negativo	5	10,00%
Total	100	

Folder CA1-10		
Total Positivos	50	
Total Negativos	41	
Total Folder	91	
Verdadeiro Positivo	45	90,00%
Verdadeiro Negativo	38	92,68%
Falso Positivo	3	7,32%
Falso Negativo	5	10,00%
Total	91	

Figura 26 — Resultados de todos os *folders* do conjunto de avaliação CA1.

Na figura 26, é possível ainda observar que, em apenas um dos *folders*, a taxa de *recall* (verdadeiros positivos) ficou abaixo de 90% — e com um *FAR* de somente 2%. Na média, o sistema apresentou taxa de *recall* de mais de 93% e *FAR* em torno de 4%. Isso é mostrado na tabela 4.

Tabela 4 — Desempenho geral do conjunto de avaliação CA1.

Desempenho CA1		
Total Positivos	450	
Total Negativos	441	
Total Testados	891	
Verdadeiro Positivo	419	93,11%
Verdadeiro Negativo	420	95,24%
Falso Positivo	21	4,76%
Falso Negativo	31	6,89%
Total	891	

A figura 27 mostra a curva *ROC* para os *folders* de melhor e de pior desempenho, que foram o

F6 e F10, respectivamente. Também é mostrado o ponto onde ocorre o desempenho médio do sistema.

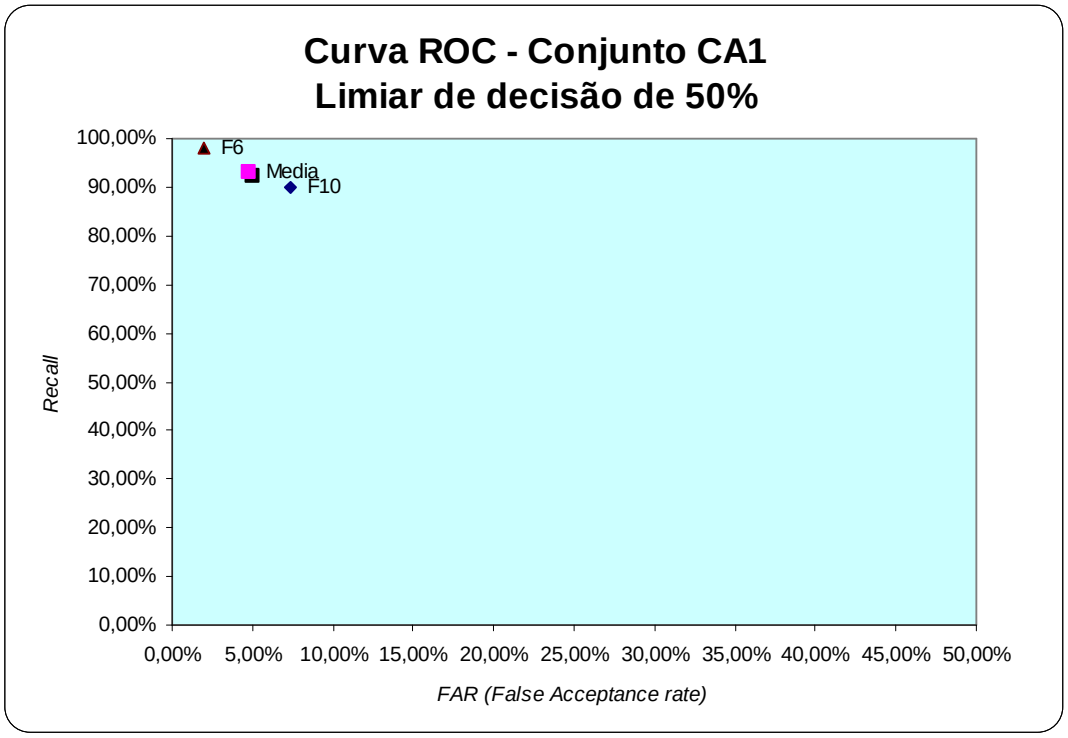

Figura 27 — Curva *ROC* para os *folders* F1, F10 e a média geral de desempenho do conjunto CA1 usando um limiar de decisão de 50%.

Os resultados apresentados mostram que a o sistema se mostrou bastante estável em todos os *folders*, tendo performance bem animadora para algo baseado unicamente em características biométricas e que não utiliza nenhum equipamento especifico para a coleta dessas características.

6.2 Controle de Acesso de 100 Usuários

Neste conjunto de testes, avaliamos a performance do sistema em controlar o acesso de 100 indivíduos a determinado recurso, ambiente similar ao encontrado em uma empresa de médio porte.

Para isso, foi usado o conjunto de avaliação CA2 conforme descrito na seção 6.1.1. Este conjunto é formado por imagens de 991 pessoas e foi dividido em 4 *folders* com 200 pessoas e um *folder* com 191.

Inicialmente, usamos o primeiro *folder* para treinar o sistema com as 100 primeiras pessoas contidas no mesmo. Em seguida, apresentaremos ao sistema um subconjunto de teste formado por todas as pessoas no *folder* — ou seja, 200 —, sendo que 100 delas são as mesmas apresentadas no treinamento (com imagens diferentes) e 100 completamente desconhecidas.

Seguindo o mesmo procedimento adotado no conjunto de avaliação CA1, usamos o primeiro *folder* para montar a tabela com a distribuição de usuários "legítimos" e "impostores" entre vários intervalos crença do classificador, e para ajustar o limiar de decisão usado neste conjunto de avaliação. Após ajustados, estes parâmetros, se manterão inalterados para todos os *folders* do conjunto.

O resultado foi reportado separadamente para cada *folder*, sendo medido o número e taxa de Verdadeiro Positivo, Verdadeiro Negativo, Falso Positivo e Falso Negativo. O desempenho do sistema será a média dos resultados reportados em cada *folder*.

A tabela 5 mostra a distribuição dos usuários nos intervalos de crença do classificador para este conjunto de avaliação.

Tabela 5 — Distribuição de usuários “legítimos” e “impostores” entre vários intervalos de crença do classificador para o *folder* CA2-F1

Probabilidade de distribuição dos acessos			
Intervalo [..)	Usuários Legitimos	Usuários Impostores	Total Acessos
0 a 4 %	0,00%	0,00%	0,00%
4 a 6%	0,00%	17,00%	8,50%
6 a 8%	0,00%	43,00%	21,50%
8 a 10%	3,00%	22,00%	12,50%
10 a 12%	2,00%	10,00%	6,00%
12 a 14%	2,00%	6,00%	4,00%
14 a 16%	1,00%	1,00%	1,00%
16 a 18%	2,00%	1,00%	1,50%
18 a 20%	1,00%	0,00%	0,50%
20 a 22%	1,00%	0,00%	0,50%
22 a 24%	3,00%	0,00%	1,50%
24 a 26%	1,00%	0,00%	0,50%
26 a 28%	1,00%	0,00%	0,50%
28 a 30%	4,00%	0,00%	2,00%
30 a 32%	1,00%	0,00%	0,50%
32 a 34%	4,00%	0,00%	2,00%
32 a 36%	1,00%	0,00%	0,50%
36 a 38%	3,00%	0,00%	1,50%
38 a 40%	1,00%	0,00%	0,50%
40 a 42%	1,00%	0,00%	0,50%
42 a 45%	2,00%	0,00%	1,00%
45 a 50%	3,00%	0,00%	1,50%
50 a 60%	17,00%	0,00%	8,50%
60 a 70%	18,00%	0,00%	9,00%
>70%	28,00%	0,00%	14,00%
Total	50,00%	50,00%	100,00%

A figura 28 mostra a curva *ROC* com as taxas de *FAR* e *Recall* do *folder* CA2-F1 para todos os limiares de decisão.

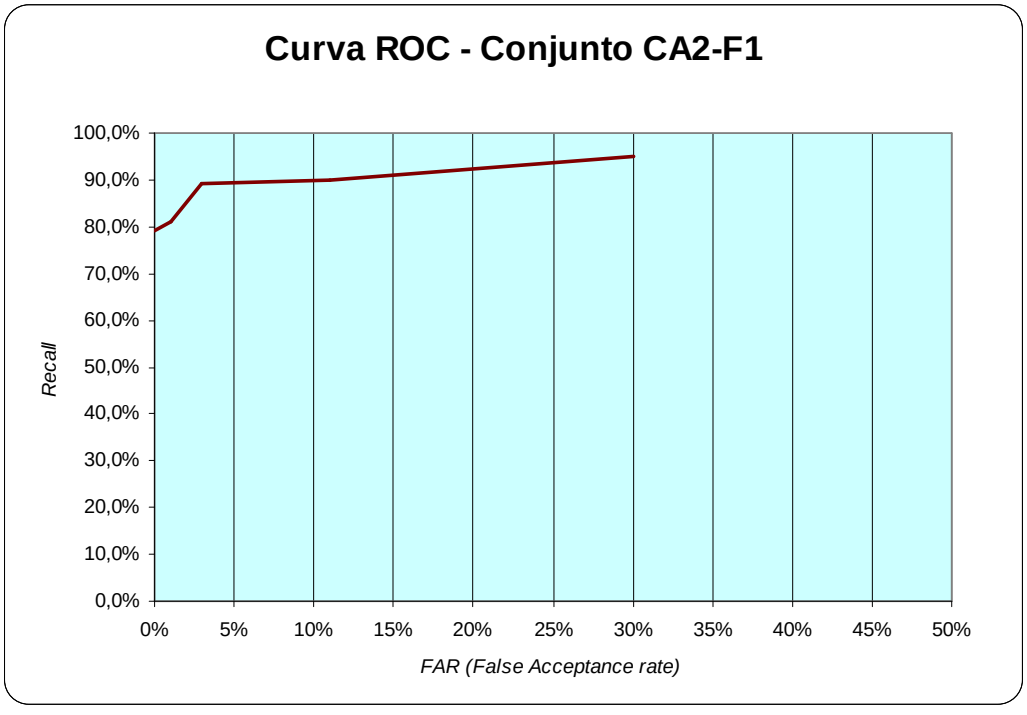

Figura 28 — Curva *ROC* com todos os limiares de decisão para o *folder* CA2-F1.

No gráfico, é possível observar que, nesta base, o sistema consegue autenticar somente 79% dos indivíduos se for desejada uma taxa zero de *FAR*. Com uma taxa de *FAR* de 3%, o consegue autenticar corretamente 90%. A tabela 6 mostra as taxas de *FAR* e *Recall* do *folder* CA2-F1 para todos os limiares de decisão.

Tabela 6 — Taxas de *FAR* e *Recall* para todos os limiares de decisão do *folder* CA2-F1.

Limiar de Decisão	Taxa de Recall	Taxa de FAR
>=12 %	95,00%	30,00%
>=16%	91,00%	11,00%
>=25%	90,00%	3,00%
>=50%	81,00%	1,00%
>=65%	79,00%	0,00%

Para este conjunto de teste foi escolhido um limiar de decisão de 25%. Com isso esperávamos uma taxa de *Recall* de 90% e um *FAR* de 3%.

Com o limiar de decisão estipulado, avaliamos o restante dos *folders* utilizando sempre um subconjunto de treino formado pelas 100 primeiras pessoas do *folder* e um subconjunto de testes formado por todo o *folder*. Em todos os subconjuntos de testes (com exceção do *folder* 5) sempre existem 100 pessoas conhecidas (apresentadas ao sistema na fase de treinamento) e 100 desconhecidas. A figura 29 mostra o resultado de todos os *folders* do conjunto de avaliação CA2.

Folder CA2-2		
Total Positivos	100	
Total Negativos	100	
Total Folder	200	
Verdadeiro Positivo	91	91,00%
Verdadeiro Negativo	97	97,00%
Falso Positivo	3	3,00%
Falso Negativo	9	9,00%
Total	200	

Folder CA2-3		
Total Positivos	100	
Total Negativos	100	
Total Folder	200	
Verdadeiro Positivo	91	91,00%
Verdadeiro Negativo	99	99,00%
Falso Positivo	1	1,00%
Falso Negativo	9	9,00%
	0	
Total	200	

Folder CA2-4		
Total Positivos	100	
Total Negativos	100	
Total Folder	200	
Verdadeiro Positivo	90	90,00%
Verdadeiro Negativo	99	99,00%
Falso Positivo	1	1,00%
Falso Negativo	10	10,00%
Total	200	

Folder CA2-5		
Total Positivos	100	
Total Negativos	91	
Total Folder	191	
Verdadeiro Positivo	89	89,00%
Verdadeiro Negativo	89	97,80%
Falso Positivo	2	2,20%
Falso Negativo	11	11,00%
Total	191	

Figura 29 — Resultado dos *folders* do conjunto de avaliação CA2.

O desempenho dos *folders* deste conjunto de avaliação se manteve com poucas oscilações nas taxas de *recall* e *FAR*. Apenas um deles ficou com a taxa de *recall* abaixo de 90%. Na tabela 7 é mostrado o desempenho médio do sistema para todo o conjunto CA2.

Tabela 7 — Desempenho médio para o conjunto de avaliação CA2.

Desempenho CA2		
Total Positivos	400	
Total Negativos	391	
Total Testados	791	
Verdadeiro Positivo	361	90,25%
Verdadeiro Negativo	384	98,21%
Falso Positivo	7	1,79%
Falso Negativo	39	9,75%
Total	791	

A figura 30 mostra o gráfico com a curva *ROC* para os *folders* de melhor e pior desempenho — F3 e F5 respectivamente —, bem como a média do desempenho para este conjunto de avaliação.

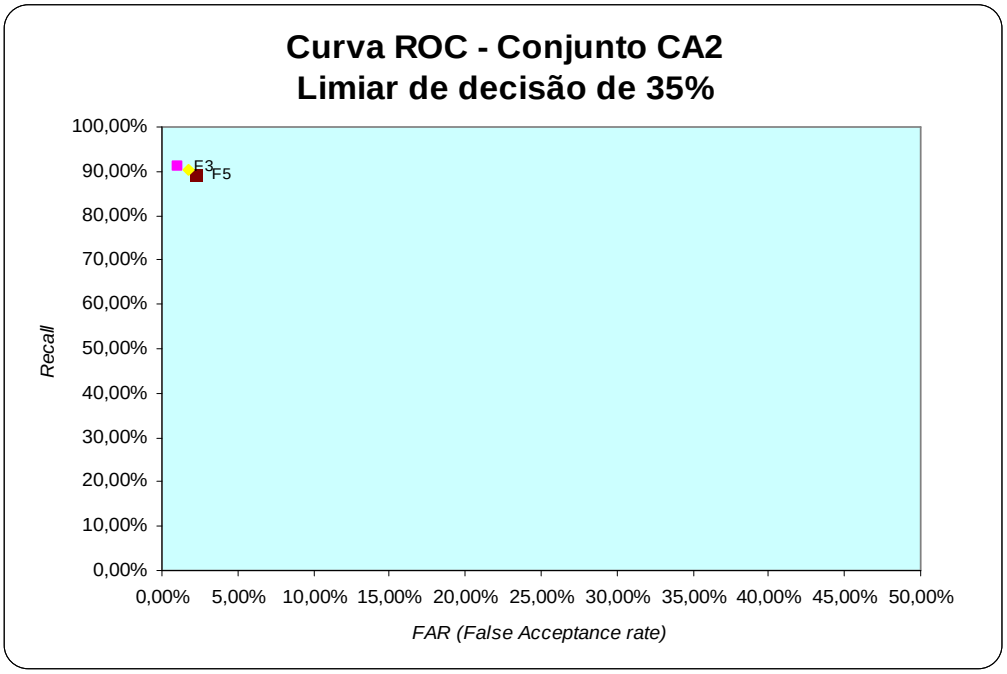

Figura 30 — Curva *ROC* para os *folders* F3, F5 e a média geral de desempenho do conjunto CA2 usando um limiar de decisão de 25%.

Os resultados apresentados por este conjunto de teste mostram que o sistema teve queda no número de corretamente autenticados em comparação com o conjunto anterior; Porém, a taxa de *FAR* diminuiu mais de 50% em relação à apresentada no conjunto anterior.

Para aplicações onde a segurança fosse priorizada, a performance deste conjunto de avaliação seria mais interessante que a do conjunto CA1, pois oferece um *FAR* abaixo de 2%.

6.3 Controle de Acesso de 200 Usuários

Neste conjunto de testes, avaliamos a performance do sistema em controlar o acesso de 200 indivíduos a determinado recurso, ambiente similar ao encontrado em uma empresa de grande porte.

Para isso, foi usado o conjunto de avaliação CA3 conforme descrito na seção 6.1.1. Este conjunto é formado por imagens de 991 pessoas e foi divido em dois *folders*, sendo o primeiro com 400 pessoas e o segundo, com 591.

Inicialmente, usamos o primeiro *folder* para treinar o sistema com as 200 primeiras pessoas contidas no mesmo. Em seguida, apresentamos ao sistema um subconjunto de teste formado por todas as pessoas no *folder*— ou seja, 400 pessoas, sendo 200 delas as mesmas que foram apresentadas no treinamento (com imagens diferentes) e 200 completamente desconhecidas.

Seguindo o mesmo procedimento adotado no conjunto de avaliação CA1 e CA2, usamos o primeiro *folder* para gerar a tabela com as crenças do classificador e ajustar o limiar de decisão. Este, uma vez ajustado, foi usado para testar o segundo *folder* do conjunto.

O resultado reportado considera somente o resultado do segundo *folder* para medir o desempenho do sistema. A tabela 8 mostra a distribuição dos usuários "legítimos" e "impostores" nos intervalos de crença do classificador para o conjunto CA3-F1.

Tabela 8 — Distribuição de usuários "legítimos" e "impostores" entre vários intervalos crença do classificador

Probabilidade de distribuição dos acessos			
Intervalo [..)	Usuários Legitimos	Usuários Impostores	Total Acessos
0 a 4 %	0,00%	1,50%	0,75%
4 a 6%	3,00%	32,00%	17,50%
6 a 8%	4,00%	34,50%	19,25%
8 a 10%	2,00%	17,50%	9,75%
10 a 12%	1,00%	8,50%	4,75%
12 a 14%	2,00%	4,00%	3,00%
14 a 16%	1,50%	1,00%	1,25%
16 a 18%	1,50%	0,50%	1,00%
18 a 20%	3,00%	0,50%	1,75%
20 a 22%	2,00%	0,00%	1,00%
22 a 24%	2,50%	0,00%	1,25%
24 a 26%	2,50%	0,00%	1,25%
26 a 28%	0,50%	0,00%	0,25%
28 a 30%	1,00%	0,00%	0,50%
30 a 32%	2,50%	0,00%	1,25%
32 a 34%	3,00%	0,00%	1,50%
32 a 36%	1,50%	0,00%	0,75%
36 a 38%	0,50%	0,00%	0,25%
38 a 40%	1,50%	0,00%	0,75%
40 a 42%	1,00%	0,00%	0,50%
42 a 45%	2,00%	0,00%	1,00%
45 a 50%	10,00%	0,00%	5,00%
50 a 60%	14,00%	0,00%	7,00%
60 a 70%	13,00%	0,00%	6,50%
>70%	24,50%	0,00%	12,25%
Total	50,00%	50,00%	100,00%

A figura 31 mostra a curva *ROC* com as taxas de *Recall* e *FAR* do *folder* CA3-F1 para todos os limiares de decisão.

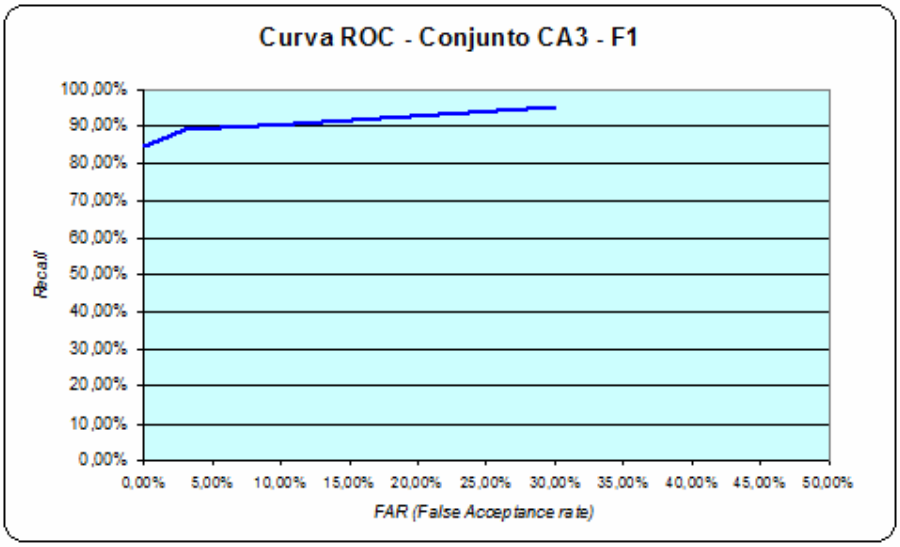

Figura 31 — Curva *ROC* para todos os limiares de decisão do *folder* CA3-F1

A tabela 9 mostra as taxas de *FAR* e *Recall* do *folder* CA3-F1 para todos os limiares de decisão.

Tabela 9 — Taxas de *FAR* e *Recall* do *folder* CA3-F1 para todos os limiares de decisão do *folder* CA3-F1.

Limiar de Decisão	Taxa de Recall	Taxa de FAR
>=08%	95,00%	30,00%
>=11%	91,00%	11,00%
>=33%	89,00%	3,00%
>=60%	86,00%	1,00%
>=75%	85,00%	0,00%

Para este conjunto de avaliação, adotamos um limiar de decisão de 33%. Com isso, esperamos minimizar o *FAR* no próximo *folder* a ser testado.

Com o limiar de decisão estipulado, avaliamos o segundo *folder*, formado por 591 pessoas, sendo 200 delas conhecidas (foram apresentadas ao sistema na fase de treinamento) e 391 desconhecidas. A tabela 10 mostra o resultado da avaliação deste segundo *folder*.

Tabela 10 — Desempenho para o conjunto de avaliação CA3-F2

Desempenho CA3-F2		
Total Positivos	200	
Total Negativos	391	
Total Testados	591	
Verdadeiro Positivo	186	93,00%
Verdadeiro Negativo	388	99,23%
Falso Positivo	3	0,77%
Falso Negativo	14	7,00%
Total	591	

Como pode ser observado, este conjunto obteve excelente performance, principalmente considerando o número de indivíduos razoavelmente alto. Este conjunto de avaliação conseguiu autenticar corretamente 93,0% — com uma taxa de *FAR* de apenas 0,77%.

6.4 Controle de Acesso Baseado em Imagens de Vídeo

Neste conjunto, avaliamos a performance do sistema utilizando imagens de vídeo em um ambiente real, sem controle de poses do usuário. Neste experimento foi utilizado uma webcam comum, com resolução de 300 *pixels* (seção 5.2)

Neste experimento foi utilizado o conjunto de avaliação pessoasCAM, descrito em 5.1.2. Durante três dias, o sistema monitorou a porta do LCAD através de uma webcam. Neste período, sempre que o usuário desejasse acesso ao LCAD deveria se posicionar em frente à câmara e levantar a mão, o sistema depois de capturar e processar a imagem “concedia” ou não o acesso a LCAD. Neste procedimento, sempre era registrando a foto do usuário a decisão tomada pelo sistema.
Vale lembrar que este procedimento era opcional, não existindo na prática, nenhuma barreira para que o usuário tivesse acesso ao LCAD sem se submeter ao procedimento.

No final do experimento, foram armazenados 46 acessos, sendo que desses, 36 eram referentes a acessos de usuários “legítimos” – que faziam parte do subconjunto de treino – e 10 eram de usuários desconhecidos, nunca vistos pelo sistema.

Em função da pouca quantidade de dados, torna-se inviável criar um subconjunto de calibração a fim de estipular um limiar de decisão, por isso, o resultado dessa avaliação será mostrado para todos os limiares de decisão.

A tabela 11 mostra a distribuição de usuários “legítimos” e “impostores” nos vários intervalos de crença do classificador para o conjunto de avaliação PessoasCAM.

Tabela 11- Distribuição usuário “legítimos” e “Impostores” nas várias faixas de crença do classificador para o conjunto PessoasCAM

Probabilidade de distribuição dos acessos			
Intervalo [..)	Usuários Legitimos	Usuários Impostores	Total Acessos
0 a 4 %	0,00%	0,00%	0,00%
4 a 6%	0,00%	0,00%	0,00%
6 a 8%	0,00%	0,00%	0,00%
8 a 10%	0,00%	0,00%	0,00%
10 a 12%	0,00%	0,00%	0,00%
12 a 14%	2,78%	30,00%	8,70%
14 a 16%	5,56%	30,00%	10,87%
16 a 18%	8,33%	30,00%	13,04%
18 a 20%	8,33%	10,00%	8,70%
20 a 22%	5,56%	0,00%	4,35%
22 a 24%	2,78%	0,00%	2,17%
24 a 26%	5,56%	0,00%	4,35%
26 a 28%	2,78%	0,00%	2,17%
28 a 30%	5,56%	0,00%	4,35%
30 a 32%	5,56%	0,00%	4,35%
32 a 34%	2,78%	0,00%	2,17%
32 a 36%	2,78%	0,00%	2,17%
36 a 38%	5,56%	0,00%	4,35%
38 a 40%	2,78%	0,00%	2,17%
40 a 42%	2,78%	0,00%	2,17%
42 a 45%	5,56%	0,00%	4,35%
45 a 50%	2,78%	0,00%	2,17%
50 a 60%	2,78%	0,00%	2,17%
60 a 70%	0,00%	0,00%	0,00%
>70%	19,44%	0,00%	15,22%
Total	78,26%	21,74%	100,00%

A figura 32 mostra o gráfico com a curva ROC para todos os limiares de decisão. A tabela 12 mostra as taxas de *FAR* e *recall* gerados para cada limiar de decisão adotado.

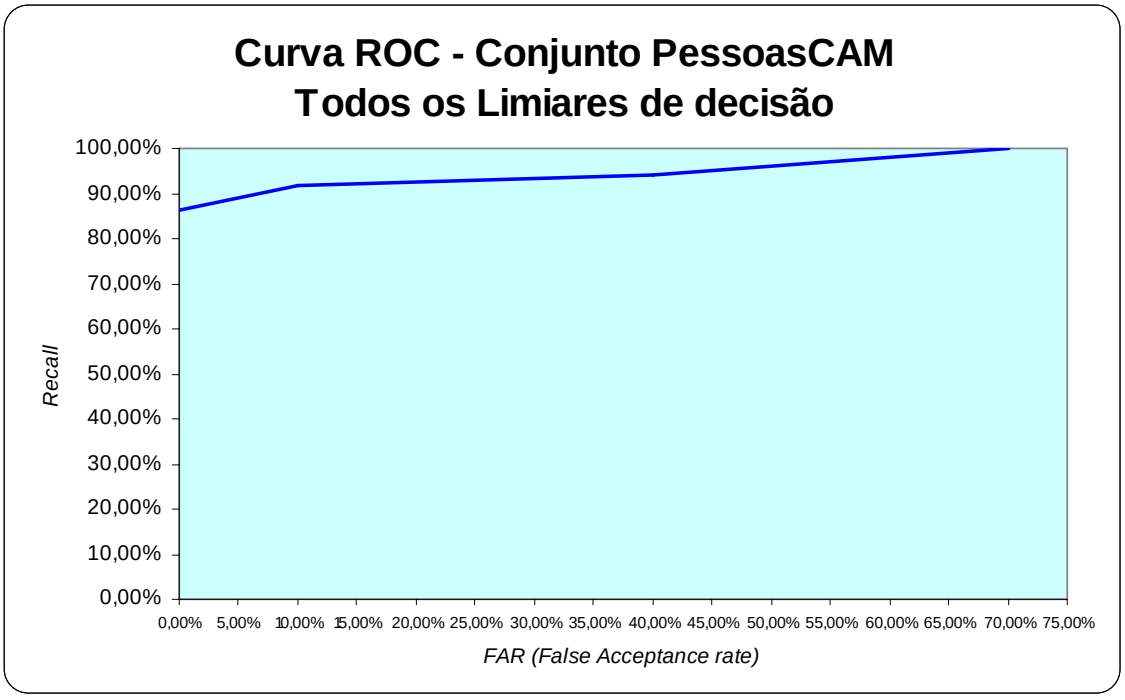

Figura 32 – Gráfico com a Curva ROC para todos os limiares de decisão do conjunto PessoasCAM.

Tabela 12 – Taxas de *FAR* e *recall* geradas para cada limiar de decisão do conjunto PessoasCAM.

Limiar de Decisão	FAR	Recall
25,00%	70,00%	100,00%
40,00%	40,00%	94,00%
50,00%	10,00%	91,67%
75,00%	0,00%	86,11%

A tabela 13 mostra o resultado obtido com um limiar de decisão de 50%.

Tabela 13– Resultados para o conjunto PessoasCAM adotando um limiar de decisão de 50%.

Conjunto PessoasCAM		
Acessos legítimos	36	
Acessos Impostores	10	
Número de Acessos	46	
Verdadeiro Positivo	33	91,67%
Verdadeiro Negativo	9	90,00%
Falso Positivo	1	10,00%
Falso Negativo	3	8,33%
Total	46	

Como pode ser observado na tabela 12, para um limiar de decisão fixado em 50% o sistema seria capaz de autenticar corretamente 91,67% dos usuários, com um *FAR* de 10%. Apesar de alto, essa taxa de *FAR* representa apenas um usuário, como pode ser visto na tabela 13.

Os resultados obtidos neste conjunto de avaliação, apesar de estarem abaixo da média dos conjuntos anteriormente testados, indicam a viabilidade da utilização do sistema proposto a partir de imagens de vídeo.

Vale ressaltar que a webcam usada neste experimento é de baixa qualidade (300 *pixels*), o que gera muitas imperfeições na imagem após o processamento da mesma (recorte, rotação, escalonamento). Este fato acarretou o descarte de parte das imagens capturadas. O uso de uma filmadora com melhor resolução poderia trazer melhores resultados.

A figura 33 mostra a comparação entre as imagens obtidas pela webcam usada no experimento e uma webcam com resolução de 1.2 Megapixels.

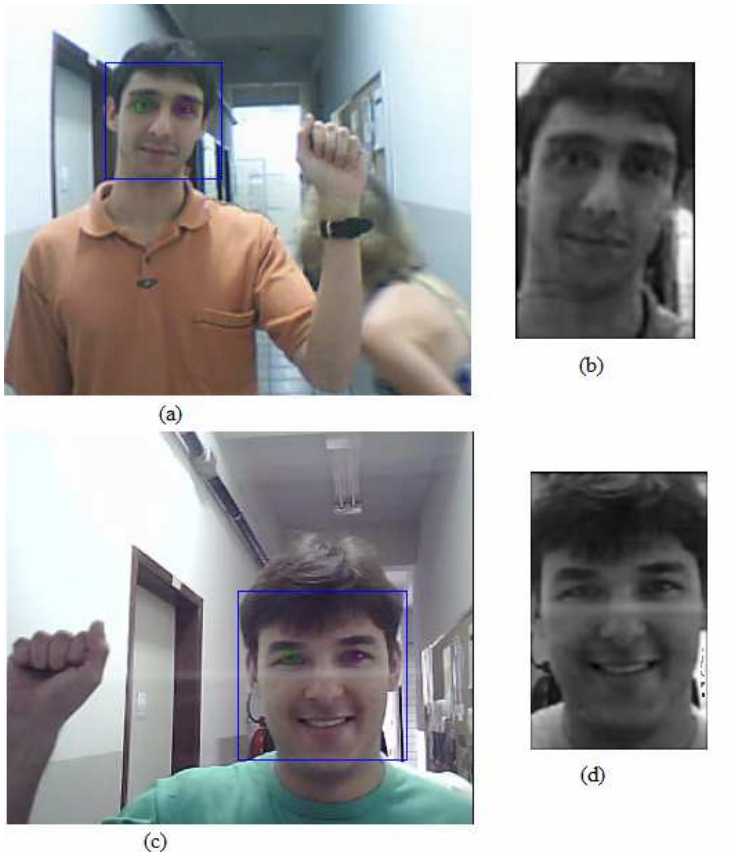

Figura 33 – (a) – Imagem captura pela webcam usada nos experimentos ; (b) - a mesma imagem após ser recortada e escalonada. (c) – Imagem captura por uma webcam com resolução de 1.2 *Megapixels* ; (d) - a mesma imagem após ser recortada e escalonada.

7 DISCUSSÃO

Este trabalho apresenta uma abordagem alternativa — ou completar — aos sistemas biométricos de controle de acesso existentes. Seu grande diferencial é usar unicamente as características biométricas da face para conceder ou não o acesso de um indivíduo a determinado recurso.

Apesar de resultados iniciais modestos, a abordagem proposta se mostrou um excelente foco para pesquisas futuras. Nas próximas seções, discutiremos os principais trabalhos correlatos, bem como as principais limitações da abordagem proposta.

7.1 Trabalhos Correlatos

Na área de pesquisa conhecida como Reconhecimento de Face (*Face Recognition*), existem excelentes trabalhos publicados, como os de [Anil04], [Tolba05], [Miller94], [Phillips00]. Porém, todos usam o reconhecimento facial para tratar do problema de verificação (*face verification*) — onde o sistema recebe algum código de identificação junto com os dados biométricos da face e reporta se a mesma pertence ou não ao código de identificação reclamado — ou do problema de identificação (*face identification*).

Apesar da similaridade, o problema de controle de acesso tem particularidades que envolvem tanto o problema de identificação quanto o de verificação, o que torna a abordagem única. O uso do reconhecimento facial especificamente nesta área de controle de acesso é ainda pouco abordado, não existindo avaliações específicas de performance de algoritmos de controle de acesso.

Na impossibilidade de comparações diretas com outras pesquisas, optamos por comparar a performance de nosso trabalho no quesito identificação (*face identification*), pois este é uma das etapas do processo de controle de acesso.
Para esta comparação, utilizamos os testes propostos pela FERET (*The Face Recognition Technology*). O FERET é um projeto do DARPA (*Defense Advanced Research Projets Agency*) do governo dos Estados Unidos, que tem como principais objetivos buscar o estado da arte em

sistemas de reconhecimento de face, identificar novas abordagens do problema e testar a performance de algoritmos que abordam o problema [Phillips00].

O trabalho do FERET foi escolhido como parâmetro por ser um dos mais reconhecidos na área e ter um protocolo de teste bem definido, indicando tamanho de conjuntos de teste e treino e características das imagens utilizadas em cada conjunto.

Além disso, pelo fato de utilizarmos as imagens do próprio FERET, podemos simular um ambiente de testes idêntico ou muito próximo do utilizado nos testes publicados, fornecendo assim métricas de performance equivalentes.

Este teste foi efetuado usando o conjunto de avaliação PessoasFB. Conforme definido na seção 6.1.1, ele é formado por 991 pessoas, sendo que os subconjuntos de treino e testes contém os mesmos indivíduos, porém com imagens diferentes.

Na fase de treino, o sistema “aprende” todas as imagens constantes no subconjunto de treino, onde cada imagem é relacionada com um ID único. Na fase de testes, é apresentada a imagem de uma pessoa anteriormente “aprendida” com o seu respectivo ID, e o sistema deve retornar uma lista com as imagens e respectivos ID’s que possuem maior similaridade com a pessoa apresentada.

O algoritmo perfeito retornaria a pessoa correta (que possui o mesmo ID da imagem de entrada) sempre na primeira posição da lista, para todas as imagens apresentadas. Os testes mostram a performance dos cinco melhores algoritmos para este conjunto de avaliação, que são:

- UMD 97 (University Maryland);
- UMD 96 (University Maryland);
- USC (University of South California);
- BaseLine EF; e,
- BaseLine Cor.

Dos algoritmos citados, apenas o USC é totalmente automático (detectando automaticamente a posição dos olhos). Nos demais, ela deve ser indicada manualmente.

A figura 34 mostra a performance dos cinco algoritmos

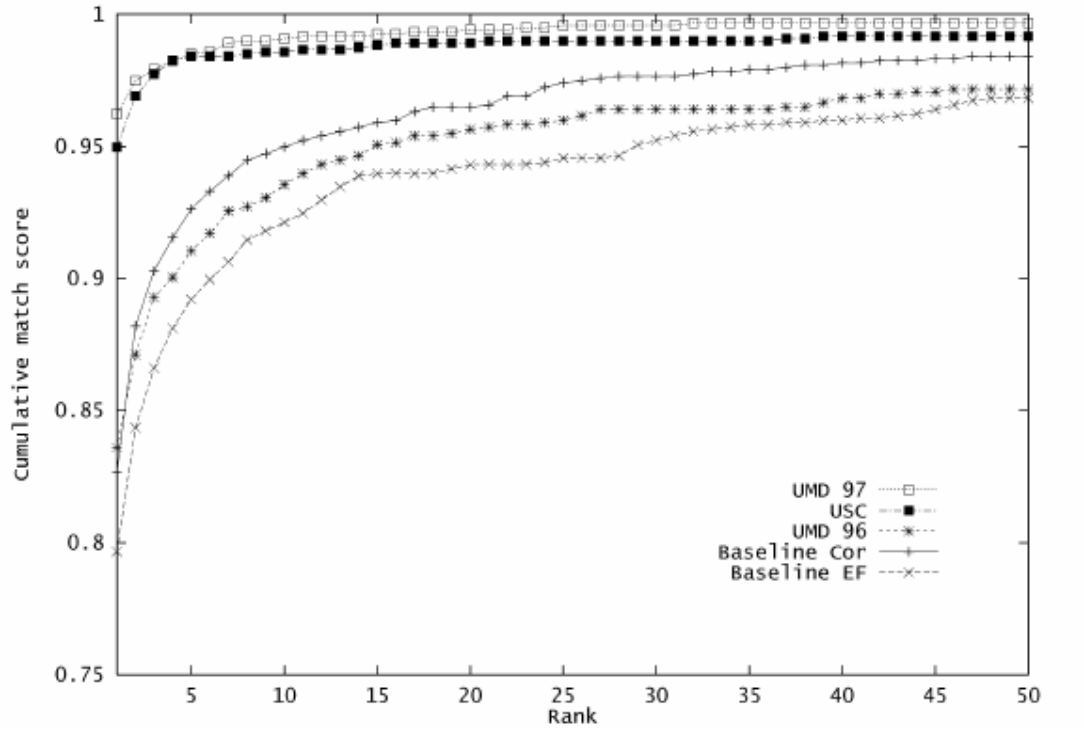

Figura 34 — Gráfico com a performance dos cinco melhores algoritmos para o problema de identificação face – Conjunto de avaliação PessoasFB
Fonte: Extraído de [Phillips00]

No eixo X, o gráfico mostra a posição ordinal do rank onde foi retornada a imagem da pessoa correta; o Y mostra a taxa de acerto acumulado para dada posição do ranking. O melhor algoritmo para esta base foi o UMD 97. Conseguiu retornar a pessoa correta na primeira posição do rank em aproximadamente de 96% dos casos.

Logo abaixo, aparece o USC, conseguindo retornar a pessoa correta na primeira posição do rank em 95% dos casos. Se considerarmos as cinco primeiras posições do rank, ambos possuem taxa de acerto de aproximadamente 97,5%.

A figura 35 mostra o gráfico com a performance alcançada neste trabalho para o mesmo conjunto de dados.

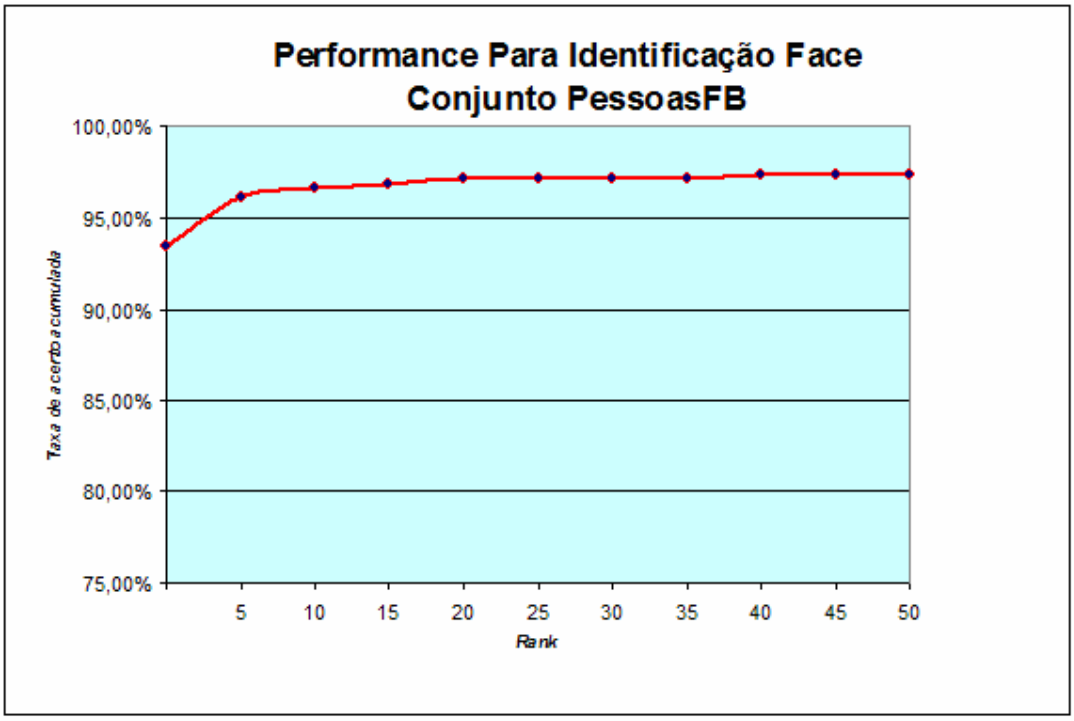

Figura 35 — Performance para o problema de identificação de face – Conjunto de avaliação PessoasFB.

No gráfico, podemos constatar que a solução implementada neste trabalho para a identificação da face retorna a pessoa correta na primeira posição do rank em aproximadamente 93,5% dos casos. Considerando-se as cinco primeiras posições do rank, o índice de acerto é de aproximadamente 96,5%.

A performance alcançada por nosso algoritmo se encontra dentro dos padrões dos algoritmos considerados o “estado da arte” para este problema, apresentando performance muito próxima aos dois melhores algoritmos testados no projeto FERET. Ressalte-se que apenas um dos algoritmos é completamente automático como o proposto neste trabalho.

O Projeto FERET prevê ainda uma série de outros testes — como, por exemplo, a comparação entre imagens com mais de um ano de diferença entre as seções de coleta. Estes não foram levados em consideração neste trabalho, por não fazer parte do escopo do mesmo.

7.2 Análise Critica Deste Trabalho de Pesquisa

Com desenvolvimento do protótipo funcional do sistema proposto — e as diversas avaliações efetuadas em conjuntos de dados de diversos tamanhos —, foram detectadas algumas limitações na técnica proposta, questões importantes que podem ser tratadas em trabalhos futuros.

a) O uso de óculos escuros pode deteriorar a performance do sistema. Em boa parte dos casos, apesar do uso de óculos, consegue detectar a posição correta dos olhos através da geometria facial. Porém, parte importante da face fica oclusa, o que dificulta o reconhecimento.

b) A abordagem proposta não fez nenhum tipo de teste de detecção de vida. De posse de uma foto de boa qualidade de um legitimo um “impostor” poderia ter o acesso a determinado recurso indevidamente liberado.

c) A taxa de *FAR* (*False Acceptance Rate*) em torno de 2% ainda é uma limitação para o uso em aplicações de segurança no mundo real.

8 CONCLUSÃO

Neste capítulo, um breve sumário do deste trabalho, suas principais conclusões e sugestões de trabalhos futuros.

8.1 Síntese

O avanço da tecnologia em várias áreas possibilitou à sociedade moderna oferecer as mais diversas facilidades aos seus indivíduos. Hoje, é possível efetuar transações financeiras sem sair de casa, fazer reuniões com pessoas que estão a milhares de quilômetros de distância, assistir aulas e palestras proferidas em outro país ou viajar de um continente a outro em poucas horas.

Porém, todas essas conveniências, e um número cada vez maior de pessoas usufruindo as mesmas, tornaram indispensável o uso de mecanismos de identificação pessoal, cada vez mais robustos, que possam comprovar que um indivíduo realmente é quem alega ser. Estes mecanismos, que se apresentam na forma de cartões magnéticos, senhas pessoais, cartões de identidade, passaporte etc.. , trazem também uma série de problemas associados, como perda, falsificação, empréstimo e dificuldade de memorização ou armazenamento de vários códigos, dentre outros.

O processo de identificação pessoal baseado em biometria tenta minimizar estes problemas, pois ela deixa de ser baseada em “algo que o individuo tem”, ou em “algo que o individuo sabe”, e passa a considerar o próprio indivíduo como código de identificação.

O projeto proposto e desenvolvido neste trabalho criou um sistema de controle de acesso baseado unicamente na biometria facial. Após o desenvolvimento de um protótipo funcional, que detecta, captura e processa a imagem de forma totalmente automática, sem o auxilio de um ser humano, o mesmo foi testado em vários conjuntos de avaliações que simulam um ambiente controlado com 50, 100, e 200 usuários.

Os resultados obtidos são animadores em ambiente de pesquisa. Com o conjunto de 200 usuários, o sistema conseguiu autenticar corretamente 93% dos usuários com um *FAR* (*False Acceptance Rate*) de apenas 0,77%; com o conjunto de 100 usuários o sistema conseguiu autenticar corretamente 90,25% dos usuários com um *FAR* de 1,79%; e com o conjunto de 50 usuários o sistema autenticou corretamente 93,11% dos usuários com um *FAR* de 4,76%.

Se comparados com formas de autenticação biométricas mais tradicionais — como a impressão digital —, os números parecem modestos, visto que um leitor de impressão digital de última geração consegue autenticar corretamente 99.9% com uma taxa de *FAR* de 0,001% . Porém, isso somente é possível utilizando hardware especifico para a coleta da digital.

Vale lembrar que, por volta de apenas cinco anos atrás, a taxa de autenticação oferecida pelos leitores de digitais não passava de 92% de autenticação com uma taxa de *FAR* em torno de 2% — performance abaixo da conseguida neste trabalho.

8.2 Conclusões

Este trabalho teve como objetivo testar a viabilidade de criar sistemas de controle de acesso baseados unicamente na biometria facial. Para isso foi criado um protótipo funcional de um sistema de reconhecimento de faces que detecta, captura, processa e tenta reconhecer uma face em uma imagem de entrada.

Para a tarefa de detecção de face, a abordagem proposta por Viola e Jones [Viola01] e usada neste trabalho se mostrou bastante adequada, atendendo as necessidades. Na detecção dos olhos esta técnica ainda precisa ser aperfeiçoada, tendo oferecido uma taxa média de acerto em torno de 85%, o que nos levou a usar geometria facial para encontrar a posição dos olhos em alguns casos.

Na tarefa de reconhecimento, foi usado Redes Neurais Sem Peso *VG-RAM* que também ofereceu bons resultados dentro dos limites à que foi submetida.

Apesar da performance de nosso sistema de controle de acesso ficar abaixo da performance dos sistemas biométricos comerciais baseados em impressão digital, do ponto de vista de pesquisa, os resultados finais obtidos são animadores, pois chegam a conseguir uma taxa de 93% de autenticações corretas, com um *FAR* (*False Acceptance Rate*) de 0,79%, sem usar nenhum tipo de coletor especifico para capturar as características biométricas.

Pelos resultados colhidos, o controle de acesso baseado somente em biometria facial se mostrou viável, dentro de alguns limites. Seu uso ainda não é indicado para aplicações que exijam um nível de segurança maior.

O controle de acesso baseado em características faciais se mostra uma área extremamente atrativa para aprofundar pesquisas e tentar novas abordagens, dado que apresenta demanda crescente e extremamente rica em termos de abordagens e técnicas a serem implementadas.

Conforme citado na seção anterior, há poucos anos atrás, os leitores de digitais ofereciam taxa de acerto menor que a conseguida neste trabalho. Agora, oferecem performance perto da perfeição. Se as técnicas de reconhecimento facial conseguirem avançar com esta mesma velocidade, provavelmente em poucos anos teremos o reconhecimento baseado em biometria facial como padrão de identificação.

8.3 Trabalhos futuros

Como sugestões de trabalhos futuros, podemos destacar:

- Implementar técnicas de detecção de vida (Liveness Detection), para que um usuário “impostor” não possa usar a foto de um usuário “legítimo” para ganhar acesso a determinado recurso;

- Aperfeiçoar a técnica de detecção dos olhos, sendo que a localização dos mesmos é ponto crucial para a correta identificação da face;

- Aumentar a robustez do sistema em relação à pose do usuário; e,

• Investigar o uso de ensemble no reconhecimento de faces, combinando, por exemplo, Redes Neurais Sem Peso com Modelos Escondidos de Markov (HMM).

9 REFERÊNCIAS BIBLIOGRÁFICAS

[Aleksander67] ALEKSANDER, I. e ALBROW, R. C.. *First National Symposium on Logic Design*. British Computer Society, Jul 1967.

[Bradley97] BRADLEY, A.P. The use of the area under the *ROC* curve in the evaluation of machine learning algorithms — Pattern Recognition. V. 30. Nº. 7. P. 1.145, 1.159 e 1.997.

[Brosso10] BROSSO, Maria Inês Lopes. *Autenticação continua de usuários em redes de computadores.* Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Computação e Sistemas Digitais. São Paulo, 2006

[Chellappa10] CHELLAPA R., SINHA, P., e PHILLIPS, J. *Face recognition by computers and humans verification*. Computer IEEE — Computer Society. V. 43. Nº 3. P. 46-55. Feb 2010.

[Souza08] SOUZA, Alberto F., BADUE, Claudine, PEDRONI, Felipe, DIAS, Stiven Schwanz, OLIVEIRA, Hallysson e SOUZA, Soterio Ferreira. *VG-RAM — Weightless Neural Networks for Face Recognition*. Disponível em http://sciyo.com/articles/show/title/vg-ram-weightless-neural-networks-for-face-recognition.

[Fawcett06] FAWCETT, T. *An introduction to ROC analysis — Pattern Recognition*. Letters 27, 861-874. Dec 2005.

[Gafurov06] GAFUROV, D., HELDALA, Kirsi E SANDOL, Torkjel. *Biometric gait authentication using accelerometer sensor*. Journal of Computers. V. 1. Nº. 7. Nov 2006.

[Haykin01] HAYKIN, S. *Redes neurais — Princípios e prática*. 2ª. São Paulo, 2001. Brookman.

[Hjelmas01] HJELMAS, E. *Face detection: a survey — Computer vision and image understanding*.

[Holanda09] HOLANDA, A. B. *Dicionário Aurélio Eletrônico* — Versão 5.0. 2009.

[Hong98] HONG, L. Hong e JAIN, Anil. *Integrating faces and fingerprints for personal identification*. IEEE Transactions on Pattern Analysis and Machine Intelligence. V. 20. Nº 12. P. 30-36. Dec 1998.

[Jain02] JAIN, A. K., BOOLE, R. e PANKANTI, S. *Biometrics — Personal identification in networked society*. New York, 2002. Kluwer Academic Publishers. Acesso em 1º/9/2010. Diponível em http://books.google.com.br/books?hl=pt-BR&lr=&id=XPC9-ucFbddsC&oi=fnd&pg=PR7&dq=biometrics:+Personal+identification&ots=zGgedT5bXT&sig=OfuxTNJZX30Aqsgp0QNkzSMg0Ko#v=onepage&q&f=false.

[Jain04] JAIN, A. K., ROSS, Araun e PRABHAKR, Salil. *An introduction to biometric recognition*. IEEE Transactions on Circuits and Systems for Video Technology — Special issue on image and video — Based Biometrics. V. 14. Nº. 1. Jan 2004.

[Jain07] JAIN, A. K., ROSS, Araun e PRABHAKR, Salil. *Human recognition using biometrics: an overview*. Appeared in Annals of Telecommunications. V. 62. P., 11-35. Jan 2007.

[Jonsson01] JONSSON, K., MATAS, J. e KITTLER, J. *Support Vector Machine for face authentication*. Image and Vision Computing 20. P. 369-375. Dec 2001.

[Kyong03] KYONG, C., BOWYER, Kevin W. e VICTOR, Barnabas. *Comparison and combination of ear and face images in appearance — Based Biometrics*. IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 25, Nº 9, September 2003.

[Magalhães02] MAGALHÃES, Marcos e PEDROSO, Antônio Carlos. *Noções de probabilidade e estatística*. 4ª Ed. São Paulo, 2002. Editora da Universidade de São Paulo.

[Rowley98] ROWLEY, H. *Neural network-based face detection.* IEEE Transaction on Pattern Analysis and Machine Intelligence. V. 20. Nº 1. Jan 1998.

[Mitchell98] MITCHELL, R. J., BISHOP, J. M., BOX, S. K. e HAWKER, J. F. *RAM based neural networks — Comparison of some methods for processing grey level data in weightless networks*. World Scientific. 1998.

[Miller94] MILLER, B. Vital signs of identity — Special Report: Biometrics. IEEE Spectrum. Feb 1994.

[Peng05] PENG, K. A robust algorithm for yes detection on gray intensity face without spectacles, Journal of Computer Science and Technology. V. 5, P. 127-132. Oct 2005.

[Phillips96] PHILLIPS, J., RAUSS, P. e FERET, S. Der. *Face recognition technology — Recognition algorithm development and test results*. U.S. Army Laboratory. 1996.

[Phillips98] PHILLIPS, J., SYED, R. e HYEONJOON, M. *The Feret verification testing protocol for face recognition algorithms* — Technical Report. Oct 1998.

[Phillips00] PHILLIPS, J., SYED, R. e HYEONJOON, M. e RAUSS, P. *The Feret evaluation methodology for face recognition algorithms*. IEEE Transaction on Pattern Analysis and Machine Intelligence. V. 22. N^{o} 10. Oct 2000

[Sakai72] SAKAI, T., NAGAO, M. e KANADE, T. *Computer analysis and classification of photographs of human faces*. First USA-Japan Computer Conference. 1972.

[Schapire00] SHAPIRE, E. e SINGER, Y. *BoosTexter: a boosting-based system for text categorization*. Machine Learning, 2000.

[Sung98] SUNG, K. Sung e POGGIO, T. *Example-based learning for view-based human face detection*. IEEE Transaction on Pattern Analysis and Machine Intelligence. V. 20. N^{o} 1. P. 30-32. Jan 1998

[Tolba05] TOLBA, A. S, EL-BAZ, A. H. e EL-HARBY, A. A. *Face recognition: a literature review*. International Journal of Signal Processing. V. 2. N^{o} 2. 2005.

[Viola01] VIOLA, P.a e JONES, M. *Robust real-time object detection*. Technical Report Series CRL 2001/01. Cambridge Research Laboratory/Compaq Computer Corporation. Cambridge, Massachusetts 02142, USA. Feb 2001.

[Viola01b] VIOLA, P. e JONES, M. *Robust real-time object detection*. Second International Workshop on Statistical and Computational Theories of Vision — Modeling, Learning, Computing and Sampling. Vancouver, Canada. Jul 2001.

[Vetter10] VETTER, R. *Authentication by biometric verification*. Computer, IEEE Computer Society. V. 43. N^{o} 3, P. 28-29. Feb 2010.

[Waring05] WARING, C. *Face detection using spectral histogram and SVM's*. IEEE — Transactions on Systems, Man and Cybernetics. Part B. V. 25. P. 467-476. Jun 2005.

[Zang06] ZANG, H., Gao, W., Chen, X. e ZHAO, D. Object detection using spatial histogram features, Image and Vision Computing. V. 20, p. 1-15. 2006

[Zhao03] ZHAO, W., CHELLAPPA, R., PHILLIPS, J. e ROSENFELD, A. *Face recognition: a literature survey*. ACM Computing Surveys. V. 35. N^{o}. 4. P. 399–458. Dec 2003.