UNIVERSIDADE ESTADUAL PAULISTA
"Júlio de Mesquita Filho"
Pós-Graduação em Ciência da Computação

BRUNO ELIAS PENTEADO

AUTENTICAÇÃO BIOMÉTRICA DE USUÁRIOS EM SISTEMAS DE E-LEARNING BASEADA EM RECONHECIMENTO DE FACES A PARTIR DE VÍDEO

UNESP

2009

BRUNO ELIAS PENTEADO

Autenticação Biométrica de Usuários em Sistemas de E-Learning Baseada em Reconhecimento de Faces a Partir de Vídeo

Dissertação apresentada para obtenção do título de Mestre em Ciência da Computação, área de Processamento de Imagens e Visão Computacional, junto ao Programa Pós-Graduação em Ciência da Computação do Instituto de Biociências, Letras e Ciências Exatas, da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto.

Orientador: Aparecido Nilceu Marana

Bauru, 27 de julho de 2009

Penteado, Bruno Elias.
Autenticação biométrica de usuários em sistemas de E-Learning baseada em reconhecimento de faces a partir de Vídeo / Bruno Elias Penteado. - São José do Rio Preto : [s.n.], 2009.
86 f. : il. ; 30 cm.

Orientador: Aparecido Nilceu Marana
Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas

1. Biometria. 2. E-Learning. 3. Autenticação web. 4. Reconhecimento de faces por vídeo. 5. Visão por computador. 6. Internet (Redes de computação) 7. Software - Arquitetura. I. Marana, Aparecido Nilceu. II. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. III. Título.

CDU - 57.087.1

Ficha catalográfica elaborada pela Biblioteca do IBILCE
Campus de São José do Rio Preto - UNESP

BRUNO ELIAS PENTEADO

Autenticação Biométrica de Usuários em Sistemas de E-Learning Baseada em Reconhecimento de Faces a Partir de Vídeo

Dissertação apresentada para obtenção do título de Mestre em Ciência da Computação, área de Processamento de Imagens e Visão Computacional, junto ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Biociências, Letras e Ciências Exatas, da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto.

BANCA EXAMINADORA

Prof. Dr. Aparecido Nilceu Marana
Professor Doutor
UNESP – Bauru
Orientador

Prof. Dr. Agma Juci Machado Traina
Professor Titular
Universidade de São Paulo – USP - São Carlos

Prof. Dr. Wilson Massashiro Yonezawa
Professor Doutor
UNESP – Bauru

Bauru, 27 de Julho de 2009

Dedico este trabalho:

aos meus pais, Waldemar e Maria;

Agradecimentos

Este trabalho é fruto de muita dedicação e apoio de muitas pessoas.

Gostaria de agradecer de todo o coração...

... à minha família, por minha formação e o apoio nesta conquista;

... aos meus colegas e amigos nesta jornada, Leandro, Evandro e Giovani, os quais compartilharam as alegrias e angústias desta fase;

... aos companheiros da República Gato Morto, com os quais morei e compartilhei minhas expectativas durante esta jornada;

... à MStech, pela oportunidade de conciliar os estudos com minha profissão, com horários flexíveis e ambiente voltado a inovação, e a seus funcionários que colaboraram para a construção da base de dados usada neste trabalho;

... à minha namorada, Julia, pela compreensão e apoio nos momentos mais difíceis;

... ao Prof. Nilceu, pela confiança depositada em meu trabalho, pela introdução às sutilezas do mundo acadêmico, pela orientação conduzida minuciosamente e paciência na exploração desta linha de pesquisa;

... a todos os outros que direta ou indiretamente contribuíram para o êxito deste trabalho.

"Se eu vi mais longe, foi por estar de pé sobre ombros de gigantes."

(Isaac Newton)

RESUMO

Nos últimos anos tem sido observado um crescimento exponencial na oferta de cursos a distância realizados pela Internet, decorrente de suas vantagens e características (menores custos de distribuição e atualização de conteúdo, gerenciamento de grandes turmas, aprendizado assíncrono e geograficamente independente, etc.), bem como de sua regulamentação e apoio governamental. Entretanto, a falta de mecanismos eficazes para assegurar a autenticação dos alunos neste tipo de ambiente é apontada como uma séria deficiência, tanto no acesso ao sistema quanto durante a participação do usuário nas atividades do curso. Atualmente, a autenticação baseada em senhas continua predominante. Porém, estudos têm sido conduzidos sobre possíveis aplicações da Biometria para autenticação em ambientes Web. Com a popularização e conseqüente barateamento de hardware habilitado para coleta biométrica (como *webcams*, microfone e leitores de impressão digital embutidos), a Biometria passa a ser considerada uma forma segura e viável de autenticação remota de indivíduos em aplicações Web. Baseado nisso, este trabalho propõe uma arquitetura distribuída para um ambiente de *e-Learning*, explorando as propriedades de um sistema Web para a autenticação biométrica tanto no acesso ao sistema quanto de forma contínua, durante a realização do curso. Para análise desta arquitetura, é avaliada a performance de técnicas de reconhecimento de faces a partir de vídeo capturadas on-line por uma *webcam* em um ambiente de Internet, simulando a interação natural de um indivíduo em um sistema de *e-Learning*. Para este fim, foi criada uma base de dados de vídeos própria, contando com 43 indivíduos navegando e interagindo com páginas Web. Os resultados obtidos mostram que os métodos analisados, consolidados na literatura, podem ser aplicados com sucesso nesse tipo de aplicação, com taxas de reconhecimento de até 97% em condições ideais, com baixos tempos de execução e com pequena quantidade de informação trafegada entre cliente e servidor, com *templates* em torno de 30KB.

Palavras-chave: Biometria, *e-Learning*, autenticação Web, reconhecimento de faces por vídeo, visão por computador, Internet (redes de computação), software - arquitetura.

ABSTRACT

In the last years it has been observed an exponential growth in the offering of Internet-enabled distance courses, due to its advantages and features (decreased distribution and content updates costs, management of large groups of students, asynchronous and geographically independent learning) as well as its regulation and governmental support. However, the lack of effective mechanisms that assure user authentication in this sort of environment has been pointed out as a serious deficiency, both in the system logon and during user attendance in the course assignments. Currently, password based authentication still prevails. Nevertheless, studies have been carried out about possible biometric applications for Web authentication. With the popularization and resultant decreasing costs of biometric enabled devices, such as webcams, microphones and embedded fingerprint sensors, Biometrics is reconsidered as a secure and viable form of remote authentication of individuals for Web applications. Based on that, this work presents a distributed architecture for an e-Learning environment, by exploring the properties of a Web system for biometric authentication both in the system logon and in continuous monitoring, during the course attendance. For the analysis of this architecture, the performance of techniques for face recognition from video, captured on-line by a webcam in an Internet environment, is evaluated, simulating the natural interaction of an individual in an e-Learning system. For that, a private database was created, with 43 individuals browsing and interacting with Web pages. The results show that the methods analyzed, though consolidated in the literature, can be successfully applied in this kind of application, with recognition rates up to 97% in ideal conditions, with low execution times and with short amount of information transmitted between client and server, with templates sizes of about 30KB.

Keywords: Biometrics, e-Learning, Web authentication, face recognition from video, computer vision, Internet (computing networks), software architecture.

SUMÁRIO

LISTA DE SIGLAS E ABREVIATURAS

ABED	Associação Brasileira de Educação a Distância
CAS	Central Authentication Service
EaD	Educação a Distância
EER	*Equal Error Rate*
FAR	*False Acceptance Rate*
FRR	*False Rejection Rate*
FTC	*Failure To Capture*
FTE	*Failure To Enroll*
HMM	*Hidden Markov Models*
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais
LMS	*Learning Management System*
MEC	Ministério da Educação
PCA	*Principal Component Analysis*
ROC	*Receiver Operating Curve*
UAB	Universidade Aberta do Brasil
UNIVESP	Universidade Virtual do Estado de São Paulo

LISTA DE FIGURAS

LISTA DE TABELAS

Capítulo 1 - Introdução

A popularização da Internet tem estimulado o desenvolvimento de tecnologias que possibilitam o trabalho colaborativo. Nos últimos anos tem-se observado um crescimento acelerado de uma destas tecnologias: os sistemas de *e-Learning*, que servem como apoio para a educação a distância por meio da Internet. Segundo dados da ABED (2007), a procura por cursos a distância no Brasil cresceu 315% entre 2003 e 2006, totalizando mais de 2,5 milhões de matrículas em programas de educação a distância. Seguindo essa tendência, surgem programas de expansão pública do ensino superior por meio da Internet como as Universidades Abertas, presentes em países como Brasil, Portugal, Índia, Itália, Inglaterra (REDDY, 2005).

A educação a distância surgiu muito antes da Internet[1]. Entretanto, a evolução das tecnologias de informação e comunicação tem proporcionado maior eficiência na distribuição de conteúdo para estudantes remotos. A proliferação dos sistemas de *e-Learning* é conseqüência das vantagens da educação baseada na Web (CANTONI; CELLARIO; PORTA, 2003), tais como: custos de distribuição diminuídos, aprendizado auto-dirigido, aprendizado geograficamente independente, atualização simples de materiais, gerenciamento simplificado de grandes grupos de estudantes e assim por diante.

No que tange à segurança em ambientes Web, estudos têm sido realizados sobre possíveis aplicações da biometria para autenticação de usuários nesses ambientes (ALMEIDA *et al*. 2006; MURAS *et al*. 2007; HERNÁNDEZ, 2008; AGULLA *et al*. 2008; dentre outros). Entretanto, a autenticação baseada em senhas ainda é predominante. Com a popularização e conseqüente barateamento dos preços de hardware habilitado para a aquisição de dados biométricos

[1] No Brasil, o Instituto Universal Brasileiro, por exemplo, oferece cursos a distância por correspondência desde 1941.

(*webcams*, microfones, leitores de impressões digitais), assim como a popularização do acesso à Internet de banda larga, a Biometria torna-se uma forma viável de autenticação remota de indivíduos em aplicações Web.

Dentre as tecnologias biométricas mais pesquisadas, o reconhecimento de faces destaca-se por sua aceitabilidade por parte das pessoas e facilidade de coleta, mesmo de forma não colaborativa (JAIN; ROSS; PRABHAKAR, 2004). As técnicas tradicionais de reconhecimento biométrico de faces fazem uso de imagens estáticas (TURK; PENTLAND, 1991; BELHUMEUR; HESPANHA; KRIEGMAN, 1997) como dados de entrada, em geral capturadas com a colaboração do usuário ou então quando uma imagem da face em boa qualidade (isto é, frontal) seja captada pela câmera, em modo não colaborativo. Porém, estudos no campo das Ciências Cognitivas apontam que as propriedades de temporalidade e movimento ajudam no processo de reconhecimento de faces por humanos (O'TOOLE; ROARK; ABDI, 2002; KNIGHT; JOHNSTON, 1997). Baseados neste fato, diversas pesquisas no campo da Biometria têm sido realizadas nos últimos anos tentando tirar proveito destas propriedades inerentes aos vídeos: sua temporalidade, a captura de movimentos dinâmicos e a abundância de informação presente nos quadros que constituem um vídeo.

1.1 Objetivos

O objetivo geral deste trabalho é propor uma arquitetura para sistemas de autenticação biométrica de usuários em ambientes de *e-Learning*, baseada em reconhecimento de faces a partir de vídeo e avaliar o desempenho desta tecnologia biométrica. Para isso são utilizadas técnicas estabelecidas na literatura, de autenticação remota de indivíduos em aplicações Web.

As tecnologias de identificação biométrica de pessoas têm mostrado grande potencial de aplicação devido à sua capacidade de autenticar unicamente os indivíduos. O uso de técnicas para identificação biométrica baseada em faces é justificado por sua fácil coletabilidade e aceitabilidade por parte das pessoas (JAIN; ROSS; PRABHAKAR, 2004). Na última década, o reconhecimento biométrico de faces a partir de vídeo tem chamado a atenção da comunidade científica devido às suas vantagens em relação à identificação baseada em imagens estáticas, tais como a continuidade temporal, a abundância de quadros, a grande quantidade de câmeras instaladas para monitoramento urbano e pela conseqüente oportunidade de pesquisas.

Nesse contexto, como objetivos específicos deste trabalho pode-se citar: (i) a exploração do vídeo como fonte de dados biométricos; (ii) a investigação da possibilidade do uso da interação natural do usuário com o sistema para a criação da base de dados; (iii) a avaliação da arquitetura proposta, por meio do desenvolvimento de um protótipo de sistema de autenticação biométrica baseado nos conceitos mencionados.

De acordo com a arquitetura proposta neste trabalho, o sistema de autenticação biométrica pode ser acoplado a qualquer ambiente de educação a distância, independentemente da linguagem na qual o ambiente tenha sido desenvolvido. A arquitetura proposta tem também como objetivo a exploração das características intrínsecas de aplicações para Internet, buscando eficiência no tráfego da rede, distribuição de carga entre clientes e servidores e independência de plataforma de *e-Learning* adotada.

Durante a revisão da literatura foram encontrados poucos trabalhos sobre a eficiência, em termos de taxa de identificação e tempo de processamento, de técnicas biométricas aplicadas a ambientes de educação a distância. Para avaliação do sistema proposto neste trabalho, foram coletadas amostras de usuários interagindo naturalmente com uma aplicação simulada de *e-Learning*. As performances das técnicas de localização e reconhecimento de faces, em termos de classificação correta, foram medidas e analisadas de modo a se investigar a sua aplicabilidade.

1.2 Estrutura da Dissertação

Nesta seção, ao apresentar a organização da dissertação¸ pretende-se orientar o leitor sobre as linhas seguidas ao longo do seu desenvolvimento. Esta dissertação é dividida em capítulos conforme apresentado a seguir.

Neste Capítulo, é introduzida a motivação e as justificativas do presente trabalho, e são descritos os objetivos gerais e específicos a serem atingidos com esta dissertação.

Após a presente introdução, o Capítulo 2 apresenta uma breve introdução à Biometria, descrevendo seus principais conceitos e características, as limitações e métricas de desempenho.

Também são abordados os componentes básicos dos sistemas biométricos, servindo como base para a implementação do método e do sistema propostos.

No Capítulo 3 é apresentada uma revisão da literatura sobre reconhecimento de imagens a partir de vídeo. Nele são estabelecidos os conceitos, as características e as propriedades desta forma de reconhecimento biométrico e que servem de base para a determinação do método proposto.

No Capítulo 4 são estabelecidos os conceitos básicos da educação a distância, o papel da tecnologia da informação como catalisadora desta modalidade educacional, a abrangência de seus programas no Brasil atualmente e a importância da verificação da identidade de indivíduos neste tipo de ambiente. Também é feita uma revisão da literatura sobre os trabalhos relacionados ao tema da autenticação biométrica para o uso em sistemas de *e-Learning*.

No Capítulo 5 são descritos os passos e as técnicas adotadas para o reconhecimento de faces a partir de vídeo bem como a arquitetura usada para a implementação do sistema proposto.

No Capítulo 6 são descritos detalhadamente o material, os resultados experimentais e as análises sobre os dados coletados na realização da avaliação do método proposto.

No Capítulo 7 são apresentadas a discussão e as conclusões sobre os resultados obtidos e as observações realizadas durante o processo de pesquisa, implementação e avaliação do método proposto. São discutidas também as limitações e dificuldades observadas no desenvolvimento do presente trabalho.

No Capítulo 8 são descritas possíveis pesquisas futuras que podem ser desenvolvidas a partir de melhorias ou pontos a serem aprofundados no método proposto neste trabalho. Também são apresentados os trabalhos publicados na área pelo autor desta dissertação, durante a realização do presente trabalho

No Apêndice A são mostradas as telas das páginas usadas para a coleta dos vídeos de treinamento utilizados neste trabalho.

Capítulo 2 – Identificação Biométrica

Este capítulo apresenta uma breve introdução sobre os conceitos relacionados às tecnologias biométricas para identificação pessoal bem como suas características, medidas de desempenho dos sistemas biométricos, tipos de erros, componentes, dentre outros. Os conceitos apresentados neste capítulo servem para contextualizar o desenvolvimento deste trabalho.

2.1 Identificação de Indivíduos

A identificação de indivíduos tem sido a base para possibilitar a interação de seres humanos em comunidade. Devido à complexidade e sofisticação desenvolvidas por estas comunidades, ao crescimento no número de indivíduos e à mobilidade geográfica entre as populações, surgiu a necessidade de se criar mecanismos para comprovar se uma pessoa realmente é quem alega ser. Estes mecanismos estão presentes atualmente em nossas vidas sob várias formas: documentos de identidade, cartões bancários, registros governamentais, entre outros.

O processo de associar uma identidade a uma pessoa é chamado *identificação pessoal* (JAIN; BOLLE; PANKANTI, 2002). Este processo permite que seja concedido ou negado o acesso a um determinado recurso (informações, documentos, locais reservados, etc.) com base na identidade associada ao indivíduo que a solicita. Neste contexto, a identidade pode ser entendida como a informação associada a uma pessoa em um determinado sistema.

Os avanços nos campos de comunicação, mobilidade e relacionamento apresentam situações em que a identificação do indivíduo é crucial. Verificar se uma pessoa pode freqüentar determinada instalação, se ela é procurada pelo governo, se lhe é permitido acessar dados em um computador ou sacar fundos em caixas eletrônicos são exemplos de aplicações atuais da identificação para segurança.

Entretanto, alguns problemas associados ao procedimento de identificação surgem, tais como: perda, deterioração, falsificação, empréstimo ou cópia de documento ou cartão de

identificação, esforço para memorização de códigos de acesso ou até mesmo portar documentos exigidos; grande variedade de códigos e cartões e assim por diante.

O processo de identificação pessoal baseia-se no uso de credenciais que podem ser classificadas como de três tipos (MILLER, 1994), conforme mostra a Figura 1:

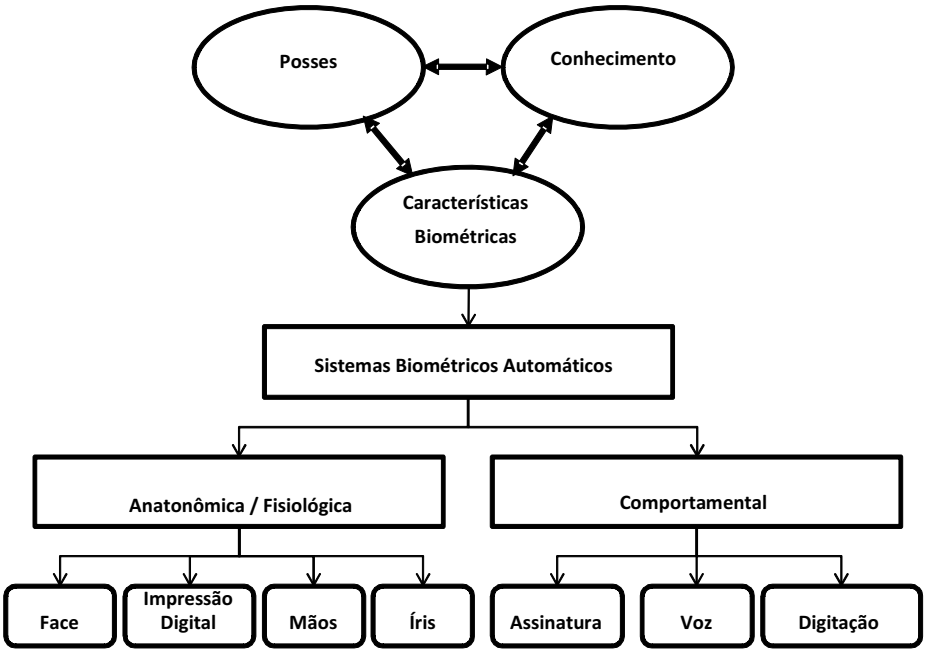


Figura 1. Tipos de métodos de identificação usados atualmente (MILLER, 1994).

- **Posse**: algo que o indivíduo possui (por exemplo: um cartão);
- **Conhecimento**: algo conhecido pelo indivíduo (por exemplo: um código);
- **Biometria**: característica anatômica (por exemplo: face, impressão digital e íris), fisiológica (por exemplo: padrão de calor do rosto, veias das mãos), ou comportamental (por exemplo: voz e assinatura) do indivíduo.

Em sistemas computacionais, a autenticação clássica de usuários é baseada em cartões (posse) ou senhas (conhecimento) que podem facilmente ser perdidos ou esquecidos. Esses

problemas podem ser minimizados pelo uso de autenticação biométrica (biometria), que faz com que a interação homem-máquina para autenticação de usuários seja mais transparente, aumentando sua usabilidade.

2.2 Biometria

Biometria pode ser definida como o ramo da ciência que estuda a mensuração dos seres vivos (HOLANDA, 2009). A palavra tem origem no grego *bios* (vida) e *metron* (mensuração). A Biometria explora o fato de que certas características físicas ou comportamentais diferenciam, de maneira confiável, uma pessoa da outra. Prabhakar, Pankanti e Jain (2003) definem biometria como sendo o "reconhecimento pessoal baseado em características comportamentais ou fisiológicas de um individuo".

Inicialmente, a identificação biométrica foi utilizada por especialistas em aplicações de alta segurança. Porém, atualmente, é comum encontrar sistemas automáticos que realizem essa funcionalidade em aplicações de nosso cotidiano, como em controle de acesso a instalações, autenticação em sistemas de votação, etc.

Como exemplo de características anatômicas e fisiológicas, podem-se citar: impressões digitais, aparência facial, padrão da íris, geometria das mãos, DNA, arcada dentária, padrão de veias das mãos, retina, padrão de temperatura da face, entre outros. Essas características são originadas pela herança genética do indivíduo e tendem a ser pouco alteradas ao longo do tempo. As características comportamentais, por outro lado, podem variar muito em intervalos de tempo relativamente curtos. Além disso, podem ser aprendidas ou treinadas ao longo do tempo. Entre as mais comuns, podem-se citar: padrão de voz, assinatura, dinâmica de digitação, modo de andar, movimentos labiais, entre outros.

Existem diversos fatores que determinam a adequação de um traço biométrico a uma determinada aplicação. Clarke (1994) e Newham (1995) apontam as principais propriedades desejáveis a uma determinada característica biométrica, que são:

- **Universalidade**: todas as pessoas devem possuir a característica;

- **Unicidade**: a característica deve ser única para cada pessoa;
- **Permanência**: a característica não deve se alterar ao longo do tempo;
- **Mensurabilidade**: a característica deve ser passível de ser coletada e medida;
- **Desempenho**: a acurácia e os recursos exigidos devem respeitar restrições para a aplicação;
- **Aceitabilidade**: os indivíduos a serem identificados devem aceitar fornecer a característica ao sistema;
- **Grau de Impostura**: a característica deve ser de difícil imitação.

Nenhuma característica biométrica possui todas estas propriedades. Portanto, não existe característica biométrica ótima.

Neste trabalho foi explorado o reconhecimento biométrico de pessoas baseados em suas faces. Dentre as vantagens do reconhecimento de pessoas pela face destacam-se: traço biométrico mais comumente usado por humanos no reconhecimento de pessoas; permite identificação à distância, sem contato com o sensor (câmera) e até mesmo sem conhecimento do indivíduo; pode ser capturada com boa qualidade mesmo por câmeras de baixo custo; existência de bases de dados legadas (passaportes, licenças de motoristas, etc.).

A Tabela 1 mostra a comparação entre as características biométricas mais populares em termos das propriedades desejáveis, ilustrando seus pontos fortes e suas limitações (JAIN; ROSS; PRABHAKAR, 2004):

Tabela 1. Comparativo entre as principais características biométricas (JAIN; ROSS; PRABHAKAR, 2004).

Biometria	Universalidade	Unicidade	Permanência	Coletabilidade	Desempenho	Aceitabilidade	Impostura
Face	Alta	Baixa	Média	Alta	Baixa	Alta	Baixa
Impressão Digital	Média	Alta	Alta	Média	Alta	Média	Alta
Geometria das mãos	Média	Média	Média	Alta	Média	Média	Média
Íris	Alta	Alta	Alta	Média	Alta	Baixa	Alta
Veias das mãos	Média	Média	Média	Média	Média	Média	Alta
Orelha	Média	Média	Alta	Média	Média	Alta	Média
Digitação	Média	Média	Baixa	Média	Baixa	Média	Média
Odor	Alta	Alta	Alta	Baixa	Baixa	Média	Baixa
DNA	Alta	Alta	Alta	Baixa	Alta	Baixa	Baixa
Termograma	Alta	Alta	Baixa	Alta	Média	Alta	Alta
Retina	Alta	Alta	Média	Baixa	Alta	Baixa	Alta
Assinatura	Baixa	Baixa	Baixa	Alta	Baixa	Alta	Baixa
Voz	Média	Baixa	Baixa	Média	Baixa	Alta	Baixa
Modo de andar	Média	Baixa	Baixa	Alta	Baixa	Alta	Média

2.3 Sistemas Biométricos

Um sistema biométrico é fundamentalmente um sistema de reconhecimento de padrões. Reconhecimento de padrões pode ser entendido como a capacidade de distinguir padrões e separá-los em diferentes categorias ou classes, levando em conta a distribuição de seus elementos

(HAYKEN, 2000). No contexto dos sistemas biométricos, um padrão pode ser uma impressão digital, uma imagem da íris, uma face, o modo que uma pessoa caminha, etc.

A classificação pode ser feita de duas maneiras:

- **Classificação supervisionada**: as classes são definidas a priori, pelo projetista do sistema;
- **Classificação não-supervisionada**: baseada no aprendizado do sistema, que, a partir de um conjunto de treinamento, delimita o espaço de características a que os padrões pertencem.

Nos sistemas biométricos, um ou mais *templates*[2] são associados a um determinado usuário. Podem-se armazenar estes *templates* em bases de dados centralizadas, cartões magnéticos, cartões de memória, entre outros, dependendo da aplicação e do tamanho dos *templates* obtidos (NEWHAN, 1995).

Na classificação em sistemas biométricos, existem dois casos:

- **Classificação entre duas categorias**: amostras da própria pessoa (genuíno) e amostras de outras pessoas (impostores);
- **Classificação entre várias categorias**: onde se têm *N* classes (*N* é o número de indivíduos diferentes cadastrados na base de dados), da qual se deseja categorizar dentre as *N* classes aquela classe à qual o indivíduo pertence – ou seja, qual indivíduo está sendo consultado.

[2] *Template* é definido como o conjunto de características que identificam um determinado indivíduo

2.3.1 Componentes dos Sistemas Biométricos

Em um sistema biométrico, é possível identificar os seguintes módulos (JAIN; ROSS; PRABHAKAR, 2004), ilustrados na Figura 2:

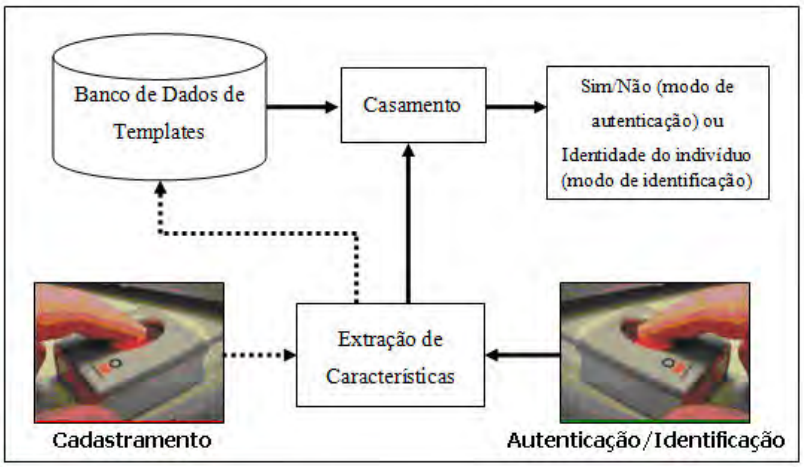


Figura 2. Interação entre os componentes de um sistema biométrico (JAIN; ROSS; PRABHAKAR, 2004).

- **Módulos de cadastramento e autenticação/identificação**: leitor ou *scanner* captura o dado biométrico do usuário. Eles definem a interface com o usuário e influenciam diretamente na qualidade do *template* extraído. No cadastramento do usuário, seus dados são armazenados pela primeira vez no sistema, enquanto nas operações de autenticação e identificação os dados coletados são comparados aos *templates* na base de dados.

- **Módulo de extração de características**: o dado extraído pelo sensor é avaliado, tipicamente utilizando-se de algoritmos de realce de sinal; em seguida, o módulo processa o dado, gerando um conjunto de dados que representa o traço capturado do indivíduo (*template*).

- **Módulo de casamento**: as características extraídas são comparadas com os *templates* armazenados (imagens de galeria), gerando uma pontuação (*match score*) que mede a similaridade entre eles; o módulo também apresenta a

capacidade de validar a identidade ou fornecer a classificação do indivíduo dentro dos *templates* cadastrados previamente.

- **Módulo de banco de dados de *templates***: repositório da informação biométrica extraída dos indivíduos.

A partir destes módulos deriva-se o seguinte fluxo de trabalho para a realização do procedimento de identificação utilizando sistemas biométricos (MATHYAS; STAPLETON, 2000):

- Captura dos dados biométrico, a partir de um sensor;
- Avaliação da qualidade da amostra biométrica capturada, e, se necessário, nova aquisição;
- Processamento dos dados biométricos capturados para a criação da amostra biométrica (*template*);
- Comparação (ou cadastro) do dado extraído com os *templates* armazenados no bando de dados (verificação ou identificação).

2.3.2 Modos de Operação

Os sistemas biométricos de identificação podem operar de duas formas distintas (BOLLE *et al*. 2004): por meio da *verificação* ou por meio da *identificação*.

A *verificação* (também chamado de *autenticação*) consiste no processo de confirmar a autenticidade do indivíduo, ou seja, verificar se ele é realmente quem alega ser. Este processo também é conhecido como *um para um*, devido ao seu funcionamento: compara-se apenas a credencial fornecida pelo indivíduo com a credencial armazenada por ele anteriormente no sistema.

A *identificação* (ou *reconhecimento*) refere-se ao processo de associar uma identidade à pessoa a partir de um conjunto de identidades conhecidas pelo sistema. É também conhecida

como *um para muitos*, pois, dada a credencial do indivíduo, busca-se entre todos os registros de indivíduos cadastrados no sistema qual é a identidade que corresponde às credenciais do requerente.

2.3.3 Desempenho dos Sistemas Biométricos

Sistemas biométricos apresentam certas limitações de desempenho relacionadas às amostras capturadas entre os indivíduos cadastrados. Dentre elas, podem-se destacar a **variabilidade intraclasse**: em que dados biométricos de uma mesma pessoa, coletados sob diferentes condições (pose, iluminação, escala, etc.), podem levar a grandes variações nos *templates* gerados para o mesmo indivíduo e a **similaridade interclasses**: em que amostras biométricas de diferentes pessoas podem apresentar variações muito pequenas entre si, como por exemplo, a face de irmãos gêmeos, em um sistema biométrico baseado na aparência facial.

Devido a variações na captura da amostra biométrica, tais como diferenças no posicionamento do traço biométrico, mudanças ambientais, deformações do traço biométrico na interação com o sensor, ruídos e má interação com o sensor, é praticamente impossível duas amostras de uma mesma característica biométrica coincidirem perfeitamente. Por este motivo, adota-se uma pontuação que quantifica o grau de semelhança existente entre o dado de entrada e o *template* armazenado. Quanto maior for a pontuação, maior é a certeza que ambos referem-se à mesma identidade (PRABHAKAR; PANKANTI; JAIN, 2003).

Sistemas biométricos são projetados para tomar decisões binárias – aceitar o indivíduo como genuíno e autenticá-lo ou considerá-lo impostor e, portanto, rejeitá-lo. A partir desta premissa, podem ocorrer dois tipos de erros no domínio do sistema: (i) *Falsa Aceitação* (FA): reconhecimento de uma amostra biométrica como sendo verdadeira, quando na verdade é falsa; (ii) *Falsa Rejeição* (FR): reconhecimento de uma amostra biométrica como sendo falsa, quando na verdade ela pertence ao indivíduo em questão.

As taxas de ocorrência de tais erros são críticas para se avaliar a eficiência do sistema. À vista disso, a avaliação de desempenho de um sistema biométrico leva em consideração os seguintes fatores (JAIN; ROSS; PRABHAKAR, 2004):

- **Taxa de falsa aceitação (*false acceptance rate* – FAR) ou taxa de falsos positivos**: probabilidade de que um indivíduo não autorizado seja autenticado. Pode ser estimada da seguinte maneira:

$$FAR = \frac{\text{número de falsas aceitações}}{\text{número de tentativas de impostores}}$$

- **Taxa de falsa rejeição (*false rejection rate* – FRR) ou taxa de falsos negativos**: probabilidade de que um indivíduo autorizado seja incorretamente rejeitado. Pode ser estimada da seguinte maneira:

$$FRR = \frac{\text{número de falsas rejeições}}{\text{número de tentativas de genuínos}}$$

Para classificar um indivíduo, é utilizado um limiar de casamento *t* (*matching threshold*) para determinar em qual categoria ele se enquadra.

A Figura 3 ilustra distribuições hipotéticas de probabilidades de duas categorias (genuínos e impostores), tendo como separador o limiar *t*. Pode ser observada a seguinte propriedade: ao mover o limiar *t* para a direita, diminui-se a taxa de falsa aceitação (FAR), ao passo que a taxa de falsa rejeição (FRR) aumenta e ao mover o limiar à esquerda, ocorre o contrário.

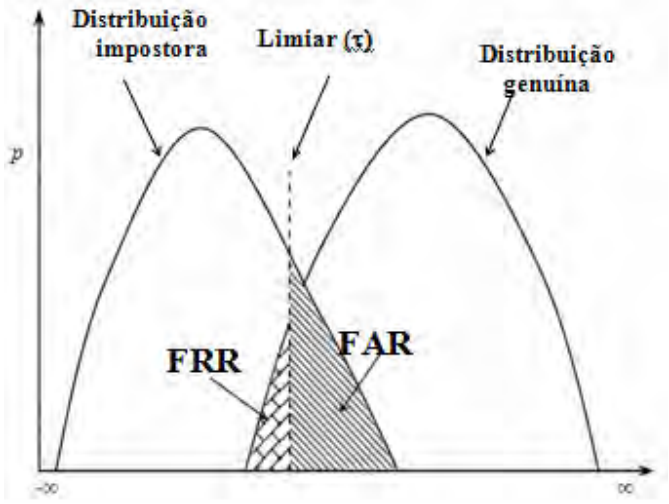


Figura 3. Distribuição das probabilidades de genuínos e impostores. (JAIN; ROSS; PRABHAKAR, 2004)

Existe um valor de limiar para o qual as taxas de FAR e FRR têm valores iguais. Essa taxa de erro é denominada *taxa de erro igual* (*equal error rate* - EER). Esta taxa de erro pode ser utilizada como uma medida comparativa estatística normalizada entre dois sistemas biométricos. Quanto menor for o EER de um sistema, melhor será seu desempenho.

O valor de EER pode ser calculado a partir da curva ROC (*receiver operating curve*), que representa a FAR em relação à FRR. A Figura 4 mostra um exemplo de curva ROC com as medidas de FAR e FRR para um sistema biométrico hipotético.

É importante ressaltar que se pode variar o limiar do sistema de identificação biométrica, priorizando-se a conveniência ou a segurança, dependendo do uso da aplicação. Por exemplo: para aplicações forenses (questões judiciais), a FRR é a maior preocupação, pois não se deseja perder um indivíduo genuíno, mesmo que seja necessária futura investigação entre vários suspeitos. Para aplicações de alta segurança, a intenção é evitar ao máximo que impostores sejam aceitos mesmo tendo como conseqüência o incômodo gerado pelas falsas rejeições. Aplicações comerciais tendem a balancear as falsas aceitações e rejeições, o que leva ao ajuste de seu limiar para um valor de próximo ao valor que gera o EER.

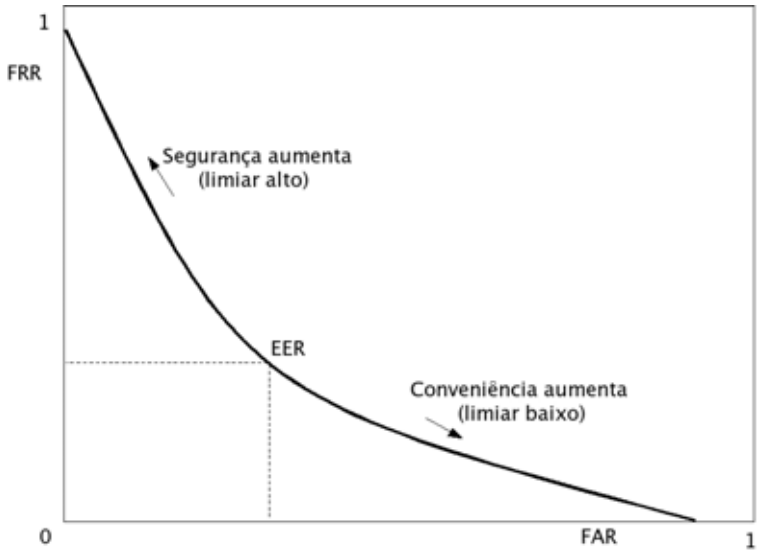


Figura 4. Curva ROC com as taxas de falsa aceitação (FAR) e falsa rejeição (FRR). O ponto de Taxa de Erro Igual (EER – *Equal Error Rate*) encontra-se a 45º a partir da origem, no ponto onde intercepta a curva (COSTA; OBELHEIRO; FRAGA, 2006).

Existem ainda dois outros tipos de erros resultantes dos sistemas biométricos: a chamada **falha de captura** (*failure to capture* – FTC), que expressa a porcentagem de vezes que o dispositivo sensor falha na tarefa de capturar automaticamente a amostra biométrica, e a **falha de cadastro** (*failure to enroll* - FTE), que denota a porcentagem de vezes que o usuário não consegue se registrar no sistema (JAIN; ROSS; PRABHAKAR, 2004).

2.4 Considerações Finais

A todo instante surgem novas aplicações para a identificação de pessoas, abrangendo desde acesso a instalações e a sistemas computacionais até vigilância de locais públicos.

Para o tratamento deste tipo de problema, a biometria tem se destacado por não exigir inúmeras memorizações ou acessórios extras para a autenticação de indivíduos; apenas as características (anatômicas, fisiológicas ou comportamentais) inerentes ao individuo são avaliadas.

Apesar de suas vantagens incontestáveis, os sistemas biométricos não são à prova de falhas, apresentando limitações desde a aquisição do traço biométrico até a classificação incorreta de indivíduos. Deste modo, certas propriedades devem ser observadas ao se projetar um sistema biométrico, levando em consideração o contexto de sua aplicação.

Neste capítulo foram introduzidos os conceitos da tecnologia biométrica para identificação humana, apresentando suas características e limitações. Foram apresentados também os componentes que formam um sistema completo, suas limitações e suas métricas de desempenho.

No Capítulo 3 é apresentada a motivação do uso do reconhecimento de faces a partir de vídeo como base de um sistema biométrico, apresentando as características, vantagens e limitações desta abordagem, bem como os principais métodos e técnicas presentes na literatura.

Capítulo 3 – Reconhecimento de Faces a Partir de Vídeo

Neste capítulo são apresentados conceitos e características do reconhecimento de faces baseado em seqüências de vídeo, suas vantagens e desvantagens em relação às técnicas tradicionais, baseadas em imagens estáticas. É apresentada também uma revisão da literatura sobre trabalhos desenvolvidos na área de reconhecimento biométrico a partir de vídeo.

3.1 Introdução

A percepção humana não usa somente a estrutura facial para reconhecer faces, mas também outros traços como cor, movimento facial, conhecimento contextual, entre outros. Estudos neuropsicológicos demonstram que o movimento facial ajuda no processo de reconhecimento das faces especialmente em ambientes degradados. De acordo com os estudos de O'Toole, Roark e Abdi (2002) e Knight e Johnston (1997):

- Tanto as informações estáticas da face quanto as dinâmicas são úteis para o processo de reconhecimento;
- As pessoas se baseiam principalmente em informação estática, pois a informação dinâmica fornece informação menos precisa do que a estrutura facial estática;
- A informação dinâmica contribui mais para o reconhecimento em condições desfavoráveis de visualização, como baixa iluminação, baixa resolução da imagem, reconhecimento a distância, etc.;
- Movimento facial contribui para o reconhecimento ao facilitar a percepção da estrutura tridimensional da face;

- Movimento facial é aprendido mais lentamente do que a estrutura facial estática;
- O reconhecimento de faces familiares é mais efetivo quando são mostradas como uma seqüência animada do que como um conjunto de múltiplos quadros sem animação. Entretanto, para faces não familiares, a seqüência animada não fornece mais informação útil do que múltiplas imagens estáticas.

Baseados nestas descobertas, pesquisadores têm tentado explorar a dinâmica facial para melhorar o desempenho dos sistemas biométricos no processo de reconhecimento de faces.

O reconhecimento de faces a partir de vídeo tem se tornado um tema popular em pesquisas relacionadas a tecnologias biométricas. Muitos locais públicos possuem câmeras de vigilância instaladas para a captura de vídeo e tais câmeras apresentam um enorme potencial para investigações forenses.

Uma definição do problema de reconhecimento facial pode ser estabelecida como: dadas imagens estáticas ou vídeos de uma determinada cena, identificar uma ou mais pessoas na cena usando uma base de dados de faces previamente cadastradas (CHELLAPPA; WILSON; SIROHEY, 1995). Tradicionalmente, as técnicas de reconhecimento de faces têm se limitado às imagens estáticas, as quais descartam a informação dinâmica temporal e a inerente estrutura tridimensional da face.

3.2 Propriedades das Seqüências de Vídeo

O reconhecimento de faces baseado em vídeo se originou a partir das técnicas baseadas em imagens estáticas. Uma das vantagens do uso do vídeo em relação às imagens estáticas é o fato de melhor se adaptar a ambientes não-colaborativos (em que o usuário não precisa fornecer voluntariamente sua face), sem o conhecimento do indivíduo. Esta característica torna esta modalidade interessante para aplicações de segurança, como vigilância e controle de fronteiras.

Contudo, ainda existem vários desafios para o reconhecimento de faces baseado em vídeo. A Figura 5 mostra imagens obtidas por câmera de vigilância no controle de acesso a um

determinado local. Podem-se notar questões relacionadas à iluminação, poses não-frontais, faces de baixa resolução, borramento devido ao movimento, etc.

Figura 5. Imagens de câmera de vigilância, apresentando diferentes condições de iluminações, poses, oclusão e ausência de poses frontais. (STALLKAMP; EKENEL; STIEFELHAGEN, 2007)

Dentre os principais desafios do reconhecimento de faces a partir de vídeo destacam-se (ZHAO *et al.* 2003):

- **Baixa qualidade do vídeo capturado**: as amostras de vídeo são adquiridas em ambientes com más condições para captura e sem cooperação dos usuários;
- **Imagens pequenas das faces**: devido às condições de aquisição, o tamanho das imagens pode ser bem menor que os das faces cadastradas na base de dados, o que pode impactar negativamente o desempenho, não só do reconhecimento em si, como também da detecção de pontos ou características usadas para a representação da face;
- **Variações significantes de pose, expressão e iluminação**: em uma mesma seqüência de vídeo um indivíduo pode haver variações significativas da pose e das expressões faciais dos indivíduos, dependendo do tipo de atividade que ele está desempenhando;

- **Poses não-fontais**: geralmente, as câmeras de vídeo são posicionadas de tal modo que os indivíduos não notem sua presença, para que o reconhecimento seja realizado sem seu conhecimento. Assim, dificilmente a imagem capturada da face terá qualidade adequada;
- **Incerteza na detecção**: devido ao movimento das pessoas em ambientes onde câmeras de vídeo estão instaladas, a detecção e o rastreamento das faces presentes na cena tornam-se um problema de estimativa dos movimentos, não tão eficientes quanto para imagens estáticas ou com colaboração dos usuários;
- **Mudanças severas de iluminação**: as câmeras que capturam os vídeos podem estar localizadas em ambientes internos, sujeitos à variação de iluminação bem como em ambientes externos com situações que alterem fortemente a iluminação, como farol de veículos, lanternas, etc.

Dado que o reconhecimento de faces pode ser efetuado tanto por imagens estáticas quanto por seqüências de vídeo, Zhou e Chellappa (2006) propuseram uma classificação baseada na natureza de composição da galeria e de consulta. A Tabela 2 mostra a relação de como é explorado o processo de reconhecimento de faces com as duas abordagens. Técnicas mais tradicionais empregam imagens tanto para a construção da base de dados quanto para consultar a base de dados (abordagem *imagem-para-imagem*). Técnicas mais recentes têm usado amostras de vídeo tanto para cadastro na base de dados quanto para verificar a identidade do indivíduo (abordagem *vídeo-para-vídeo*).

Tabela 2. Relação entre o uso de imagens estáticas e de seqüências de vídeo usadas como base de dados e como consulta (ZHOU; CHELLAPPA, 2006).

	Imagem (reconhecimento / teste)	**Vídeo (reconhecimento / teste)**
Imagem (base de dados/galeria)	**Reconhecimento baseado em imagens estáticas (imagem-para-imagem)**	**Reconhecimento imagem-para-vídeo**
Vídeo (base de dados/galeria)	-	**Reconhecimento baseado em vídeo (vídeo-para-vídeo)**

Uma forma de combinar as duas abordagens consiste em construir a base de dados com imagens estáticas dos indivíduos e efetuar o reconhecimento da identidade utilizando seqüências de vídeo. Outra forma de combinar é utilizar vídeos no cadastro do usuário e efetuar a consulta com imagens estáticas. Entretanto, esta combinação é pouco explorada na literatura.

Zhou e Chellappa (2006) propuseram também uma classificação levando em conta o processo de rastreamento da face na seqüência de vídeo e a continuidade temporal. A Tabela 3 mostra a relação entre a exploração da informação temporal nos processos de rastreamento e reconhecimento da face.

Tabela 3. Uso da informação temporal nos processos de rastreamento e reconhecimento de faces (ZHOU; CHELLAPPA, 2005).

Processo	Função	Uso de informação temporal
Rastreamento visual	Modelagem das diferenças entre os quadros	No rastreamento
Reconhecimento visual	Modelagem da diferença entre imagens de consulta e da base de dados	-
Rastreamento seguido por reconhecimento	Combinação de rastreamento e reconhecimento seqüencialmente	Somente no rastreamento
Rastreamento e reconhecimento	Unificação de rastreamento e reconhecimento em um único *framework*	Tanto no rastreamento quanto no reconhecimento

Uma seqüência de vídeo ou um conjunto de múltiplas imagens também podem ser considerados uma coleção de imagens estáticas (*quadros*). Deste modo, os algoritmos tradicionalmente usados em imagens estáticas podem ser aplicados quadro a quadro no vídeo.

Suponha a existência de um algoritmo de reconhecimento de faces A. É possível construir um algoritmo baseado nas múltiplas imagens estáticas ao combinar múltiplas instâncias do algoritmo base: A_i. Cada A_i tem como entrada um quadro i vindo da seqüência de vídeo e como resultado uma identidade. Como regra de combinação pode ser adotada a soma, o produto das pontuações individuais e assim por diante.

Tal abordagem omite algumas propriedades adicionais inerentes às seqüências de vídeo e que não estão presentes em imagens estáticas isoladas. Em particular, as três propriedades seguintes têm motivado várias pesquisas (ZHOU; CHELLAPPA; ZHAO, 2006):

- **Conjunto de observações**: os quadros são considerados como um conjunto de imagens (observações) de um mesmo indivíduo. Esta propriedade é diretamente utilizada pela fusão de algoritmos. Análises teóricas baseadas em um conjunto de observações podem ser derivadas, como por exemplo: sumarização de observações em matrizes, função de densidade de probabilidades, variantes (*manifolds*), etc. Logo, este conhecimento correspondente pode ser levado em conta no casamento de dois conjuntos;

- **Continuidade/dinâmica temporal**: a continuidade temporal pode atualizar as probabilidades posteriores usadas para identificar o indivíduo em uma imagem em particular, dada a identidade em uma imagem previamente encontrada. Faz uso de modelos que levam em conta a probabilidade posterior para descrever as variações da face e da cabeça (dinâmica) como um todo ao longo da seqüência. De acordo com esta propriedade, quadros sucessivos em uma seqüência de vídeo são contínuos na dimensão temporal que pode ser caracterizada por leis de cinemática.

- **Modelos 3-D**: reconstrução de modelos tridimensionais a partir de um conjunto de múltiplos quadros. Assim, o processo de reconhecimento pode ser baseado neste modelo tridimensional, o que proporciona invariância a pose e a iluminação.

Dentre as propriedades expostas acima, a primeira e a terceira são compartilhadas pelas seqüências de vídeo bem como por um conjunto de múltiplas imagens; porém a segunda propriedade é observada somente nas seqüências de vídeo. Outras propriedades importantes derivadas do uso do vídeo são: construção de imagens de super-resolução, que, de forma análoga ao modelo tridimensional, pode usar a informação presente na grande quantidade de quadros e montar uma única imagem de melhor resolução; aprendizado e atualização do modelo de

classificação, permitindo que o modelo de reconhecimento subjacente seja atualizado e se torne mais robusto a variações com o passar do tempo.

Zhao *et al.* (2003) apresentam uma revisão das técnicas de reconhecimento de faces a partir de vídeo e divide-as em três categorias:

- **Reconhecimento baseado em imagens estáticas aplicadas em quadros selecionados**: devido à abundância da quantidade de quadros em uma seqüência de vídeo, uma maneira de melhorar a taxa de reconhecimento é utilizar mecanismos de votação no resultado do reconhecimento aplicado a cada *frame*. Esta votação pode ser determinística ou probabilística. A votação probabilística tem, em geral, melhor desempenho, pois pode ponderar positivamente quadros de melhor qualidade;

- **Sistemas multimodais**: humanos usam múltiplas características para realizar o reconhecimento de pessoas. Assim, espera-se que sistemas que exploram outros traços além da face tenham melhor desempenho do que os que utilizam somente este traço. Em alguns cenários não-colaborativos, em que o usuário não deseja ou não precise fornecer explicitamente a face, ela talvez não seja a melhor característica para identificar um indivíduo. Assim, outros traços podem complementar a definição da identidade, como seu modo de caminhar ou sua voz.

- **Uso simultâneo de informação temporal e espacial**: informações como trajetória das características faciais (temporal) e a determinação destas características em cada quadro (espacial) que fazem uso de modelos de representação em um espaço de junção espacial e temporal para a identificação entre indivíduos;

O desempenho dos métodos tradicionais aplicados a imagens estáticas é fortemente influenciado por diversos fatores, como a variação de pose, iluminação, expressões faciais e o uso de acessórios (óculos, chapéu, entre outros). Conseqüentemente, para melhorar o desempenho, são necessárias certas transformações e correções, tanto geométricas (como rotações e registro da imagem da face) quanto fotométricas (normalização de iluminação).

Para o tratamento dos problemas de pose e variação a partir de seqüências de vídeo, duas técnicas são comumente aplicadas (PARK; JAIN; ROSS, 2007): *baseada na visão*, na qual várias imagens de um mesmo indivíduo com diferentes poses e condições de iluminação são registradas na base de dados; e *síntese de visão*, na qual são geradas visões sintéticas a partir de mapeamentos ou modelagem 3-D, tomando como base a imagem de entrada.

Nas técnicas baseada na visão, *M* diferentes conjuntos de *n* projeções são exigidas, uma para cada visão (pose). A abordagem de espaços de visão (*viewspaces*) corresponde a *M* subespaços independentes, cada um descrevendo uma região particular do espaço das faces (correspondendo a uma visão/pose particular de uma face).

Pentland, Moghaddam e Starner (1994) propuseram uma técnica de criação de múltiplos espaços – um para cada pose – em que, apresentada uma imagem de consulta, é calculada a chamada distância a partir do espaço de característica (DFFS – *Distance From Feature Space*), a qual estima a pose da face computando a sua distância (relativa ao erro de descrição residual) para cada um dos subespaços. Assim, a codificação é feita com a base de autovetores do subespaço escolhido e a comparação da face de consulta com a base de dados só é feita para a pose estimada, eliminando a busca nos outros *M*-1 subespaços. Trata-se de uma extensão da técnica de autofaces para múltiplos conjuntos de autovetores, um para cada orientação. A Figura 6 mostra alguns exemplos de variações em relação às poses testados em seu trabalho.

Figura 6. Exemplos de imagens usadas para teste de imagens de face sob diferentes orientações da cabeça – visões (PENTLAND; MOGHADDAM; STARNER, 1994).

Nas técnicas baseadas em síntese da visão, a face é reconstruída por meio da estimativa de fatores intrínsecos (expressão, maquiagem, óculos, cabelo, envelhecimento, etc.) e extrínsecos (intensidade de iluminação, direção da iluminação, uso de acessórios) à face capturada. Como exemplo, Blanz e Vetter (2003) combinam modelos deformáveis 3-D usando técnicas de computação gráfica para estimar parâmetros de projeção e iluminação. Deste modo, dada uma imagem de face de entrada, o algoritmo calcula automaticamente parâmetros tridimensionais de forma, textura e outros parâmetros relevantes da cena. Primeiro, um conjunto de 200 faces tridimensionais são capturadas e pré-processadas, analisando seu aspecto espacial (*x, y, z*) e sua textura (valores RGB). A partir dessas imagens são calculados os autovalores e autovetores para cada aspecto independentemente, formando uma nova base de vetores de face. Assim, uma nova imagem de entrada é transformada para este novo espaço e seus coeficientes de forma e textura são computados. A Figura 7 ilustra este procedimento. Por conseqüência, a partir da combinação linear destes coeficientes é feita a identificação do indivíduo em relação à base de dados.

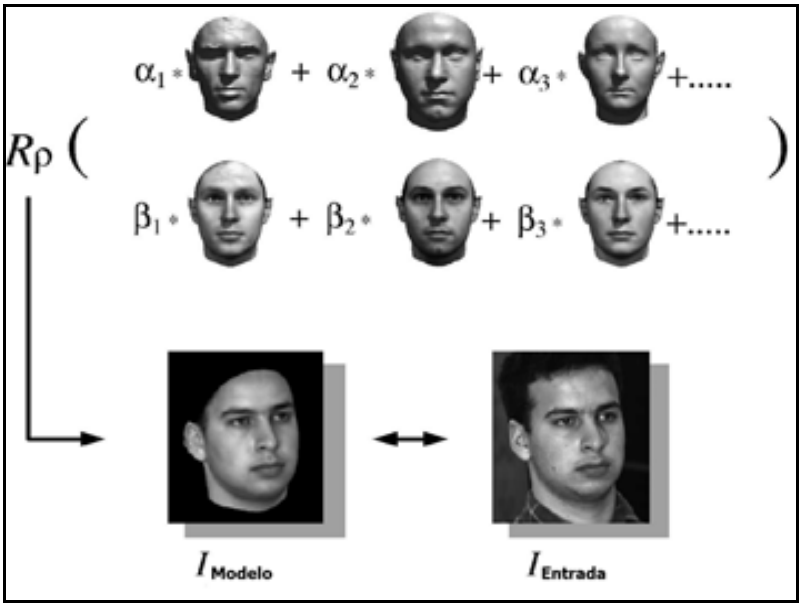


Figura 7. Estimativa dos coeficientes de forma e textura para descrição 3-D da face tal que R_p produza uma imagem I_{modelo} tão similar quanto possível à imagem de entrada $I_{entrada}$ (BLANZ; VETTER, 2003).

Na próxima seção são abordados alguns dos trabalhos mais relevantes na área de reconhecimento de faces baseado em vídeo.

3.3 Trabalhos Sobre Reconhecimento de Faces a Partir de Vídeo

As técnicas de reconhecimento de faces tradicionais se baseiam apenas em imagens estáticas, descartando a informação tridimensional inerente à face e à percepção humana, o que as tornam sensíveis à variação de pose. Uma maneira de superar este problema é a reconstrução de modelos 3-D a partir de várias imagens ou de seqüências de vídeo. Mesmo quando as resoluções das imagens ou do vídeo são altas, o que raramente acontece, o modelo da face gerado pelas técnicas conhecidas estão longe de serem perfeitas, o que não torna esta abordagem tridimensional viável em aplicações reais.

Estudos psicofísicos, como o de Bülthoff, Edelman e Tarr (1994), sugerem que um objeto 3-D é representado como um conjunto de imagens 2-D em nosso cérebro (multivisão). Assim, ao invés de trabalhar com complexos modelos tridimensionais, os métodos computacionais mais eficientes trabalham com imagens bidimensionais.

Yamaguchi, Fukui e Maeda (1998) propõem o uso do método de subespaço mútuo (MSM – *Mutual Subspace Method*), que considera o ângulo entre os subespaços de entrada e de referência formado pelas componentes principais da seqüência de imagens (não necessariamente ordenadas), como medida de similaridade. Esta abordagem desconsidera a coerência temporal inerente presente em uma face que pode ser crucial para o reconhecimento.

Liu e Chen (2003) usam as estatísticas temporais de uma face a partir do vídeo usando modelos ocultos de Markov (*Hidden Markov Models* - HMM) para efetuar reconhecimento vídeo-para-vídeo. Tais modelos são treinados e aprendidos para cada indivíduo em seqüências de autofaces, após redução de dimensionalidade. As pontuações de semelhança são comparadas e à maior é atribuída a identidade. Classificação de confiança é usada para adaptar estes modelos: caso a semelhança entre os modelos seja maior que um limiar *t*, o modelo do indivíduo é

atualizado levando em conta sua última amostra. A Figura 8 ilustra esquematicamente a projeção das autofaces e a modelagem dos estados do modelo HMM para um indivíduo variando sua pose.

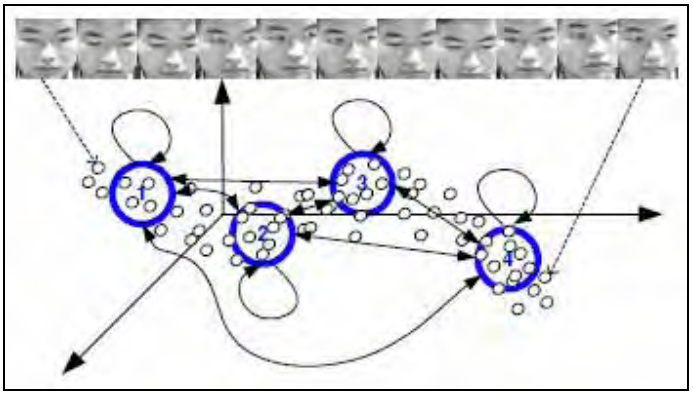


Figura 8. HMM temporal para modelagem de seqüências de faces (LIU; CHEN, 2003)

Zhou e Chellappa (2003) propõem uma estrutura probabilística tanto para rastreamento quanto para reconhecimento de faces ao estimar a distribuição de probabilidade posterior do vetor de movimento e da variável de identidade. Assim, uma variável de identidade é adicionada ao vetor de estado no método de amostragem por importância seqüencial (*Sequential Importance Sampling* - SIS). Um modelo de série temporal é usado, com o vetor de estado (n_t, θ_t), em que n_t é a variável de identidade e θ_t é o parâmetro de rastreamento, e a observação y_t (ou seja, o quadro do vídeo). O modelo de série temporal é especificado completamente pela probabilidade de transição de estado $p(n_t, \theta_t \mid n_{t-1}, \theta_{t-1})$ e a semelhança observacional $p(y_t \mid \theta_{t,}, n_t)$. Uma maneira alternativa de modelar as estruturas temporais é por meio do uso do algoritmo de propagação condicional de densidade (*CONditional DENSity PropagATION* - CONDENSATION). Este algoritmo tem sido aplicado com sucesso no rastreamento e reconhecimento de várias características espaço-temporais, como a face (ZHOU; KRUEGER; CHELLAPPA, 2002).

Aggarwal (2004) usa o modelo ARMA (Auto Regressive and Moving Average – média móvel e auto-regressiva) para o processo de reconhecimento da face. A seqüência de faces é tratada como um processo estocástico de média móvel e auto-regressiva de primeira ordem. A Equação 1 mostra o modelo ARMA de primeira ordem.

$$\theta_{t+1} = A\,\theta_t + v_t,\ y_t = C\,\theta_t + w_{t,} \qquad (1)$$

onde $v_t \sim N(0, Q)$ e $w_t \sim N(0, R)$ e θ_t é o vetor de estado que caracteriza a pose da face. A identificação consiste em estimar os parâmetros A, C, Q e R a partir das observações $\{y_1, y_2, ..., y_n\}$. Uma vez que os parâmetros estejam estimados, várias medidas de distância podem ser usadas. Bons resultados experimentais foram obtidos quando variações significantes de pose e expressão estão presentes na seqüência de vídeo.

Lee *et al.* (2003) representam a aparência de uma face por meio de variantes de poses (*pose manifolds*) que são conectadas por probabilidades de transição, que capturam diretamente a continuidade temporal. Cada pessoa é representada por uma variante de pose de baixa dimensionalidade, aproximado e comparados por partes de subespaços lineares. Em geral, as probabilidades de transição entre poses vizinhas são maiores que aquelas entre poses distantes. A Figura 9 ilustra estas variantes e como se dão as transições entre poses.

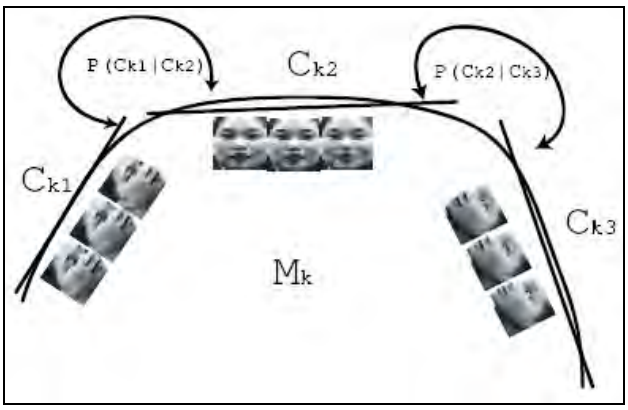


Figura 9. Dinâmica entre diferentes poses. A dinâmica entre as variantes de pose são aprendidas a partir de vídeos de treinamento que descrevem a probabilidade de se mover de uma variante para outra em qualquer instante de tempo (LEE *et al.* 2003).

Este método se mostrou um dos mais capazes de lidar com grandes rotações bi e tridimensionais. A Figura 10 mostra graficamente as probabilidades de transição entre diferentes poses. Neste exemplo, a variante de aparência é aproximada por cinco subespaços de pose. A probabilidade é representada pelo valor do brilho do elemento correspondente: quanto maior o brilho, maior a probabilidade de transição entre as poses consideradas. Pode-se notar que a pose

frontal (pose 1) tem maior probabilidade para mudar para qualquer outra pose; a pose para a direita (pose 2) tem probabilidade quase igual a zero de mudar diretamente para a pose para a esquerda (pose 5).

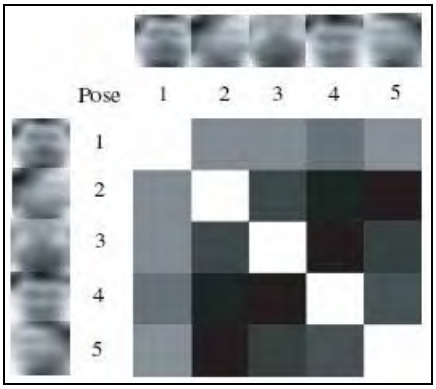


Figura 10. Representação gráfica de uma matriz de transição aprendida a partir de um vídeo de treinamento. Neste exemplo, quanto maior o brilho significa maior probabilidade de transição entre as poses. (LEE *et al.* 2003)

3.4 Considerações Finais

Estudos no campo da ciência cognitiva têm mostrado os benefícios da propriedade do movimento na melhoria do processo de reconhecimento de faces por humanos. Dessa forma, o reconhecimento baseado em vídeo tem ganhado importância na comunidade acadêmica nos últimos anos, devido às suas possíveis aplicações para fins de segurança, investigação forense, comerciais, etc.

A disponibilidade de seqüências de imagens e amostras de vídeos confere ao reconhecimento biométrico baseado em vídeo uma grande vantagem em relação ao reconhecimento baseado em imagens estáticas: a abundância de informação. Entretanto, as imagens presentes em amostras de vídeo apresentam um grande desafio: a perda de informação espacial (faces pequenas, baixa resolução, borrões devido ao movimento, etc.), como explicitado na seção 3.2. Assim, um bom modelo baseado em vídeo deve usar a informação temporal para compensar a informação espacial perdida.

Neste capítulo foram apresentadas as características intrínsecas ao reconhecimento de faces baseado em vídeo, suas vantagens e desvantagens. Os métodos mais citados na literatura foram brevemente explanados.

No Capítulo 4 é apresentada a motivação para a aplicação dos conceitos vistos até aqui em um contexto de sistemas de *e-Learning*. O uso de técnicas biométricas, mais precisamente do reconhecimento de faces a partir de vídeo, é considerado como complemento aos mecanismos de segurança usados atualmente para a autenticação de alunos que acessam os cursos remotamente.

Capítulo 4 – E-Learning

Neste capítulo são abordados os conceitos básicos que caracterizam um sistema de *e-Learning*, bem como os quesitos de segurança envolvidos. Também são listados os principais programas de educação a distância implantados no Brasil, dando uma dimensão do alcance de tais programas. É apresentada também uma revisão da literatura relacionada ao tratamento do problema de autenticação de usuários com sistemas biométricos em ambientes desta modalidade e são apontados métodos desenvolvidos para o problema.

4.1 Introdução

Uma nova forma de aprendizado, a chamada Educação a Distância (EaD), tem sido apresentada como complemento ou substituição à forma clássica presencial de ensino em escolas e ao aprendizado informal ocorrido fora do sistema escolar (RABUZIN; BACA; SAJKO, 2006). A definição para o termo educação a distância, utilizada pela legislação brasileira, é declarada como: "a modalidade educacional na qual a mediação didático-pedagógica nos processos de ensino e aprendizagem ocorre com as utilizações de meios e tecnologias de informação e comunicação, envolvendo estudantes e professores no desenvolvimento de atividades educativas em lugares ou tempos diversos" (BRASIL, 2005).

Esta forma de educação a distância tem se tornado viável devido aos avanços recentes das tecnologias de informação e de comunicação, trazendo vantagens tais como: custos de distribuição diminuídos, aprendizado auto-dirigido, aprendizado geograficamente independente, atualização simples de materiais, gerenciamento simplificado de grandes grupos de alunos, etc.

Embora a educação a distância não seja algo recente (por exemplo: cursos por correspondência são utilizados desde o século XIX), sua prática foi levada a um novo patamar

com o uso das redes de computadores e das tecnologias da informação, como CDs/DVDs, Internet, TV Digital (CANTONI; CELLARIO; PORTA, 2003).

Os sistemas de educação a distância que utilizam a Internet como meio de distribuição do conteúdo e como canal de interatividade estão inseridos na classe de sistemas conhecida como *e-Learning*. Na literatura não existe um consenso sobre este conceito, sendo que alguns autores consideram que o *e-Learning* inclui toda forma de comunicação e distribuição de conteúdo por meio digital, outros consideram apenas a Internet, outros autores consideram outras combinações entre diferentes tecnologias de comunicação e multimídia.

Neste trabalho, o conceito de *e-Learning* adotado envolve apenas os sistemas de aprendizagem baseados na Internet, como adotado, por exemplo, em Rosenberg (2002). Demais tecnologias digitais de EaD, anteriores à Web, não são consideradas, pois utilizam conteúdos seqüenciais e isentos de interatividade, resolvendo, de certa forma, apenas o problema da distância. Assim posto, *e-Learning* é concebido como uma forma de educação a distância, porém, nem toda educação a distância é necessariamente considerado *e-Learning*.

Os oponentes do *e-Learning* alegam que a principal desvantagem de tal meio de ensino é a conduta antiética neste tipo de ambiente (KENNEDY *et al.,* 2000; ROVE 2004; BARON; CROOKS, 2005). Em particular, argumentam que a incapacidade de autenticar os indivíduos que realizam as avaliações é um dos grandes desafios dos ambientes de *e-Learning*. Como resultado, algumas instituições tomam medidas como exigir que os estudantes façam as avaliações em centros presenciais ou chegam até mesmo abandonar a oferta deste tipo de curso (GUNASEKARAN; McNEIL; SHAUL, 2002). Os instrutores necessitam verificar que a submissão das avaliações é realmente realizada pelo estudante ao invés de alguma outra pessoa se passando por ele, de forma fraudulenta.

Os Sistemas de Gestão de Aprendizagem (*Learning Management Systems* - LMS), também conhecidos como Ambientes Virtuais de Aprendizagem (AVA) são as ferramentas voltadas à distribuição de conteúdo, registro de desempenho dos alunos, criação de cursos, gestão dos cursos a distância, entre outras funcionalidades. Em geral, este tipo de sistema adota como prática de segurança a utilização de mecanismos de autenticação por senhas. O uso deste tipo simples de autenticação aumenta a vulnerabilidade a fraudes, uma vez que a certeza da

autenticação do usuário é muito baixa. Assim, a não-presença dos alunos torna a fraude fácil e tentadora, pois outra pessoa pode substituir facilmente a pessoa que deveria ser avaliada.

Como a oferta dos cursos de *e-Learning* tende a aumentar, tanto em quantidade quanto em nível educacional (graduação, especialização, pós-graduação, etc.), poderá existir no futuro, por exemplo, um aluno de pós-graduação que jamais tenha assistido a uma disciplina e nunca tenha realizado de fato seus exames, e ainda assim tenha obtido o título. Isso reforça a idéia da necessidade de sistemas robustos de monitoramento e autenticação da identidade de usuários em sistemas de educação a distância.

4.2 Educação a Distância no Brasil

Segundo dados da ABED (Associação Brasileira de Educação a Distância) (ABED, 2007) e do Inep (Instituto Nacional de Estudos e Pesquisas Educacionais) (INEP, 2006), no ano de 2006 existiam no Brasil mais de 2,5 milhões de alunos matriculados em cursos de educação a distância, distribuídos em mais de 200 instituições oficialmente autorizadas a ministrar cursos dessa natureza. No período de 2003-2006 houve um crescimento de 571% na oferta e 315% no número de matrículas em cursos superiores na modalidade a distância, impulsionadas por programas de incentivo como a Universidade Aberta do Brasil[3], mostrando importância da EaD para o país.

A Tabela 4 mostra os principais programas educacionais realizados a distância no Brasil e o respectivo número de matrículas realizado até 2006 – ano do último levantamento de dados da ABED.

[3] Universidade Aberta do Brasil (UAB): http://www.uab.capes.gov.br/

Tabela 4. Número de matrículas nos maiores projetos de educação a distância no Brasil, até 2006.

Instituições	Cursos	Matrículas
Instituições credenciadas e cursos autorizados pelo Sistema de Ensino	EJA, Fundamental, Médio, Técnicos, Graduação, Pós-graduação.	778.458
Educação corporativa e treinamento em 27 instituições	Formação de funcionários, colaboradores, fornecedores e empreendedores.	306.858
Brasil Telecom	Formação de funcionários, colaboradores e fornecedores.	30.934
Vale do Rio Doce	Formação de funcionários, colaboradores e fornecedores.	12.726
Secretaria Especial de Educação a Distância do Ministério a Educação (Seed/MEC)	Formação pela Escola; Universidade Aberta do Brasil; Pró-Licenciatura; Mídias na Educação; Proformação; Proinfantil.	50.872
Governo do Estado de São Paulo	Ensinar Matemática nas Séries Iniciais; Práticas de Leitura e Escrita na Contemporaneidade; Gestão Escolar e Tecnologias; Aluno Monitor; Mídias na Educação; Interaction Teachers; Tecnologias na Educação; RIVED; Progestão.	85.470
Sebrae	Cursos para empreendedores: Análise e planejamento financeiro, Aprender a aprender, Como vender mais e melhor, De olho na qualidade, Iniciando um pequeno grande negócio e Desafio Sebrae.	300.000
Senac	Cursos de extensão e de formação inicial de trabalhadores.	73.000
Oi Futuro (Instituto Telemar)	Cursos de inclusão educacional e digital ministrados por meio do projeto Tonomundo.	515.000
CIEE	Cursos de iniciação profissional conceituais, técnicos e atitudinais.	33.771
Fundação Bradesco	Cursos para a comunidade de Tecnologia da Informação, para iniciantes em tecnologia e para educadores.	88.981
Fundação Roberto Marinho	Multicurso Ensino Médio – Matemática	3.000
Total		2.279.070

Fonte: Anuário Brasileiro Estatístico de Educação Aberta e a Distância, 2007, p. 150.

O assunto de educação a distância também é abordado pela legislação brasileira. Suas bases legais foram estabelecidas pela Lei de Diretrizes e Bases da Educação Nacional (BRASIL, 2005). Em 3 de abril de 2001, a Resolução nº. 1 do Conselho Nacional de Educação estabeleceu as normas para a pós-graduação *lato* e *stricto sensu* a distância.

Desde 2001, o Plano Nacional de Educação, instituído pela Lei federal n° 10.172, afirma em seu item 6.2: "É preciso ampliar o conceito de educação a distância para poder incorporar todas as possibilidades que as tecnologias de comunicação possam propiciar a todos os níveis e modalidades de educação, seja por meio de correspondência, transmissão radiofônica e televisiva, programas de computador, Internet, seja por meio dos mais recentes processos de utilização conjugada de meios como a telemática e a multimídia". O decreto 5.622, de 2005, que regulamenta o ensino a distância no Brasil, prevê, no Capítulo I, Artigo 1º, parágrafo 1º, a obrigatoriedade de momentos presenciais para:

- Avaliações dos estudantes
- Estágios obrigatórios
- Defesas de trabalhos de conclusão de curso
- Atividades relacionadas a laboratórios de ensino

Este mesmo decreto declara que os diplomas em cursos de graduação, na modalidade a distância, terão direito de diploma equivalente ao dos cursos de graduação presenciais, emitido pela instituição ofertante do curso.

Desde então, o *e-Learning* é utilizado pelo poder público como ferramenta de expansão da oferta de vagas principalmente para o ensino superior. No âmbito de programas governamentais, foram criados planos de expansão educacional, em todos os níveis, baseados nesta modalidade de ensino. Como exemplo, o governo do Estado de São Paulo criou em 2008, em conjunto com as

universidades estaduais paulistas – USP, UNESP e UNICAMP e com a FAPESP, o Programa Univesp[4], que tem por objetivo ampliar a oferta de vagas na formação de professores em áreas básicas. Em 2009, primeiro ano de implementação do programa, a meta é oferecer 6.600 vagas para os cursos de ensino superior em Pedagogia, Licenciatura em Biologia e Licenciatura em Ciências.

Para dar suporte ao programa Univesp, serão utilizados ambientes virtuais de aprendizagem, a fim de reproduzir e ampliar as possibilidades de ações pedagógicas normalmente usadas no ambiente de sala de aula. Tais sistemas serão baseados em experiências significativas desenvolvidas anteriormente, como o Teleduc[5], desenvolvido pela Unicamp, e o TIDIA[6], financiado pela FAPESP e desenvolvido em vários laboratórios de pesquisa do estado. No Univesp, o acompanhamento dos estudos e das atividades pedagógicas se dá de forma presencial (nos pólos presentes nas instituições participantes), por telefone e pela Internet. É usada também a mídia televisiva para a transmissão de programas-aula. As avaliações e aulas laboratoriais são realizadas de modo presencial.

Outro exemplo de grande escala de *e-Learning* no Brasil é a Universidade Aberta do Brasil (UAB). Este programa federal surgiu em 2005, criado pelo Ministério da Educação (MEC), e tem como objetivo expandir e interiorizar a oferta de cursos e programas de educação superior. Os eixos fundamentais do programa são a expansão pública da educação superior, a avaliação da educação a distância baseada na regulamentação do MEC e a contribuição para a investigação em educação superior a distância no país. Atualmente, a UAB conta com 562 pólos presenciais espalhados pelo interior do Brasil, que oferecem mais de 67 mil vagas. O objetivo é atingir 600 mil alunos até 2012.

---

[4] Programa Univesp - Universidade Virtual do Estado de São Paulo (http://www.ensinosuperior.sp.gov.br/portal.php/univesp)

[5] Teleduc – Educação a Distância (http://www.teleduc.org.br)

[6] TIDIA – Tecnologia da Informação no Desenvolvimento da Internet Avançada (http://www.tidia.fapesp.br)

O conceito de universidade aberta, para educação de grandes massas populacionais por meio do ensino a distância, está presente em mais de 60 países. Uma das pioneiras foi a universidade aberta do Reino Unido, criada há mais de 30 anos e que conta, atualmente, com mais de 200 mil alunos[7]. A maior delas está na Índia (Universidade Nacional a Distância Indira Gandhi)[8] e conta com aproximadamente 2 milhões de alunos.

Além do governo brasileiro, o governo norte-americano aprovou recentemente o *Higher Education Opportunity Act* (ESTADOS UNIDOS, 2008) para a legalização da oferta de cursos de ensino superior em geral. O decreto estabelece, dentre outras declarações, que as instituições de ensino superior que oferecem cursos a distância possuam um processo de estabelecimento da identidade do aluno, de modo que seja assegurado que o aluno matriculado para o curso é o mesmo aluno que participa, completa e recebe créditos para o curso.

Conclui-se, portanto, que esta modalidade de ensino tem se tornado cada vez mais abrangente. A educação a distância usufrui dos benefícios proporcionados pelas tecnologias da informação e comunicação e atinge várias camadas sociais e regionais em diversos países. Porém, não há em nenhum dos casos uma verificação eficaz da identidade dos estudantes para acesso aos sistemas, acompanhamento dos cursos e tampouco para as avaliações.

4.3 Segurança em Ambientes de e-Learning

Huang *et al*. (2004) criticam os sistemas de *e-Learning* proprietários existentes, até 2004, por não considerarem apropriadamente a questão de autenticação dos estudantes, em particular durante avaliações e questionários. Hugl (2005) menciona várias tecnologias relacionadas à segurança com potencial aplicação neste problema e que não são utilizadas nos sistemas de *e-*

[7] OU – The Open University (http://www.open.ac.uk/about/ou/)

[8] IGNOU – Indira Gandhi National Open University (http://www.ignou.ac.in/)

Learning. Uma das potenciais soluções sugeridas é a inclusão de tecnologias biométricas como parte integral dos sistemas LMS.

Eibl, Solms e Schubert (2006) citam como pilares da segurança da informação:

- **Confidencialidade**: os dados armazenados em uma base de dados e transmitidos sobre uma rede não podem ser capturados por terceiros não autorizados;
- **Integridade**: os dados armazenados na base de dados e os transmitidos pela a rede não podem ser modificados por terceiros não autorizados;
- **Disponibilidade**: os dados devem estar disponíveis às partes autorizadas em qualquer momento;
- **Identificação e autenticação**: um indivíduo deve ser devidamente identificado e verificado durante o processo de entrada no sistema;
- **Autorização (controle de acesso lógico)**: um indivíduo deve ter acesso somente aos dados que lhe dizem respeito.
- **Não-repúdio**: um usuário deve ser responsabilizado por quaisquer ações realizadas por ele dentro do sistema.

Segundo Weippl (2005) a segurança em sistemas LMS é uma questão relevante, pois:

- Sistemas LMS não são mais protótipos de pesquisa e sim sistemas em produção que precisam considerar seriamente a segurança de suas aplicações;
- Todo novo sistema eletrônico adiciona novas ameaças;
- Confiança em um sistema eletrônico é um pré-requisito para a aceitação do usuário.

A segurança nos ambientes LMS é um tema essencial por várias razões: realização de avaliações, acesso a informações confidenciais, comprometimento do sistema, etc. Como exemplo, as avaliações são componentes previstos pela maioria dos sistemas LMS e uma questão importante é como evitar fraudes, ou seja, como assegurar que a pessoa que está realizando o

exame seja realmente o aluno matriculado no curso. Comumente, a identidade do usuário em sistemas LMS é verificada por meio de senha. Porém, a possibilidade de fraude é alta e soluções mais robustas são necessárias. Por exemplo, um usuário pode se autenticar no sistema e depois permitir que outra pessoa responda em seu lugar às questões de um exame. Por este motivo, alguns autores alegam que aulas presenciais entre estudantes e seus tutores devem existir sempre que possível.

Outra questão importante é a segurança do próprio sistema LMS e sua estabilidade. A importância da estabilidade pode ser entendida, por exemplo, na seguinte situação: um estudante faz o exame on-line, mas não está preparado o suficiente. Em dado momento, ele pode se dar conta de que não vai passar no teste e, então, pode tentar pôr em risco a estabilidade do sistema de modo a evitar uma nota baixa. Assim, é de suma importância saber exatamente quem está autenticado no sistema e quem está autorizado a fazer quais ações, de modo que alguém que esteja executando uma ação em um computador não possa negar falsamente que realizou tal ação – princípio conhecido como não-repúdio (RABUZIN; BACA; SJAKO, 2006).

4.4 Biometria em Sistemas LMS

Sistemas de *e-Learning* representam uma nova forma de aprendizado e têm atingido um maior número de pessoas. Paralelamente, a biometria tem se tornado um método essencial de identificação e verificação e oferece grandes possibilidades. Como mencionado anteriormente, um dos problemas relativos aos sistemas de *e-Learning* é a autenticação remota e contínua dos alunos durante a realização do curso. De fato, muitos autores demonstram-se surpresos pelo fato que poucos recursos são investidos neste quesito nos sistemas LMS de apoio à educação a distância.

Rabuzin, Baca e Sjako (2006) argumentam ainda que as tecnologias biometria não são suficientemente utilizadas e que certos problemas relativos à verificação de usuários de sistemas de *e-Learning* podem ser resolvidos por meio de seu uso.

Levy e Ramim (2007) propõem uma abordagem teórica para a autenticação em avaliações eletrônicas. Nesta arquitetura, o estudante fornece sua impressão digital para se identificar no

servidor da aplicação de *e-Learning*, mantendo sua autenticação durante sua sessão. Durante as avaliações, a impressão digital de um dedo aleatório é pedida ao usuário em um curto intervalo de tempo, de modo a desencorajar que outra pessoa assuma sua posição depois de entrar no sistema.

Um protótipo de autenticação biométrica aplicada em cursos baseados na Web é apresentado por Auernheimer (2005), no contexto de uma universidade. Nesse protótipo é desenvolvido um módulo de autenticação biométrica (por impressão digital) que é usado somente para identificação do estudante, sendo que uma senha tradicional é usada para sua autenticação. Uma *webcam* é utilizada também para capturar periodicamente a imagem do aluno, para efeitos de auditoria quando da suspeita de fraude. O objetivo desta abordagem é manter-se compatível com a cultura da universidade, rompendo minimamente com os procedimentos existentes.

Almeida *et al.* (2006) apresentam um sistema de autenticação para ambientes LMS que explora a verificação biométrica baseada na dinâmica de digitação do usuário, ou seja, leva em conta fatores como intervalos de digitação entre as teclas. Essa abordagem tem a vantagem de não precisar de *hardware* especial para a coleta, uma vez que ela pode ser efetuada pelo teclado convencional e por ser um método de rápido processamento. Porém, como qualquer biometria comportamental, está sujeita a grande variação dada a condição do indivíduo.

No estudo realizado por Hernández *et al.* (2008), também foi adotada uma abordagem multimodal. Os alunos tiveram suas impressões digitais cadastradas no sistema para poder acessá-lo. O sistema também contava com uma *webcam*, que neste caso, não fazia o reconhecimento, somente verificava a presença de alguma pessoa no momento do exame. Ao final do experimento, os alunos responderam a um questionário sobre suas impressões em relação ao uso do sistema. Uma característica importante da pesquisa foi que a média das notas dos alunos que fizeram os testes a distância foi menor quando comparadas com um teste similar presencial. Os alunos sentiram-se incomodados pela presença da câmera e do tempo limitado, porém, levantaram os seguintes pontos positivos: sistema mais rápido, seguro e fácil de usar, com um índice de aprovação de 78% em um universo de 102 alunos. Cerca de 20% dos alunos tentaram alguma alternativa de burlar o sistema biométrico.

Uma plataforma para o problema de interoperabilidade de sistemas biométricos entre diferentes aplicações e dispositivos de diferentes fornecedores foi desenvolvida por Muras *et al.*

(2007), explorando a especificação de padronização de tecnologias biométricas BioAPI[9]. Nesta plataforma, foi desenvolvido um serviço de autenticação central (*Central Authentication Service* - CAS) responsável por prover a funcionalidade de *login* único (*single sign-on*) entre diferentes aplicações autenticadas pelo CAS. A proposta do trabalho foi a criação de uma plataforma genérica entre tecnologias biométricas (técnicas e sensores), sendo que nenhuma foi realmente implementada. A Figura 11 mostra a arquitetura de alto nível do sistema de Muras *et al.* (2007).

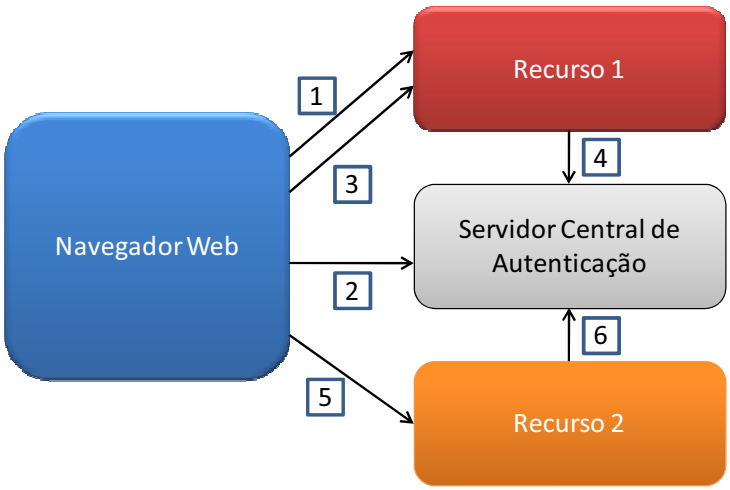


Figura 11. Arquitetura de *login* único usada por MURAS *et al.* (2007).

O protocolo de autenticação mostrado na Figura 11 descreve seu funcionamento: o usuário faz a requisição para o Recurso 1 (1). Caso ele ainda não tenha sido autenticado, ele será redirecionado para o servidor de autenticação (CAS) que fornece as credenciais e redireciona o usuário novamente para o recurso anterior (2). Em uma segunda tentativa, o navegador envia as credenciais (3) para o Recurso 1, que o valida junto ao CAS (4). O usuário autenticado tenta acessar outro recurso protegido pelo CAS (5). Se o login único o validou corretamente, nenhuma outra autenticação é necessária. O recurso valida novamente as credenciais passadas junto ao CAS (6).

[9] BioAPI Consortium (ANSI/INCITS 358-2002). http://www.bioapi.org/.

Agulla *et al.* (2008) desenvolveram um sistema, *BioTracker*, responsável por processar as imagens das faces de seus usuários. Nesta abordagem de arquitetura cliente/servidor, a carga de localização da face e seleção de imagens, baseada na estimativa de obtenção de uma boa imagem (face frontal), fica no lado cliente. O processamento no servidor encarrega-se de extrair as características da imagem da face e compará-las com as presentes na base de dados.

O sistema suporta tanto verificação colaborativa, na qual o aluno permite voluntariamente a captura de sua face, quanto não-colaborativa (modo coberto), em que o usuário não sabe que está sendo monitorado. A verificação colaborativa se dá durante o acesso do usuário ao sistema e quando não são coletadas amostras com qualidade suficiente por um determinado tempo. O modo coberto é ativado durante a sessão do usuário.

O sistema *BioTracker* teve como objetivos melhorar os processos de controle de acesso, rastrear a participação do aluno (quanto tempo ele realmente gastou no curso) e durante avaliações. Como apontado no trabalho, embora o sistema não assegure que os usuários não consigam burlar o sistema, ele pode garantir que o aluno está em frente ao computador e validar quanto tempo o usuário gastou navegando pelo curso no sistema LMS.

O *BioTracker* foi adaptado para os seguintes sistemas: ILIAS, Moodle e Claroline. A Figura 12 mostra a interface da aplicação *BioTracker* com o usuário.

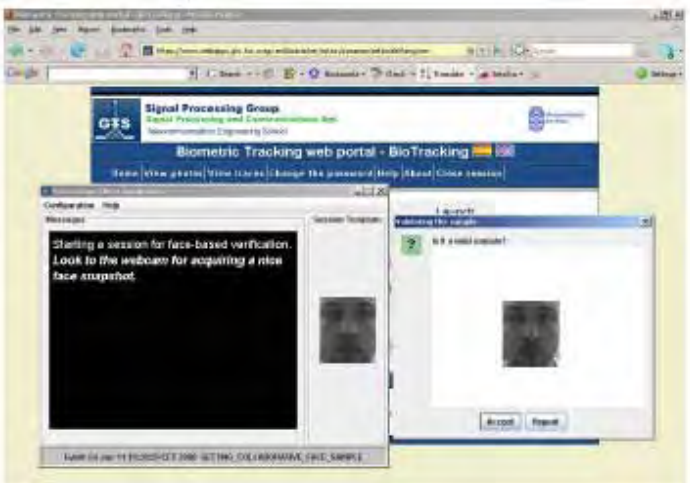


Figura 12. Interface da aplicação BioTracker (AGULLA *et al.* 2008).

Em seu trabalho, Asha e Chellappan (2008) combinam traços biométricos anatômicos e comportamentais em um esquema multimodal. São usadas as tecnologias de reconhecimento por impressão digital e por dinâmica de *mouse*, na qual direção, velocidade e tempo são usados como características. Todavia, nenhuma avaliação de acurácia foi realizada.

Rolim e Bezerra (2008) desenvolveram o SIAF (Sistema de Identificação Automática de Faces) e acoplaram-no ao ambiente virtual de aprendizagem Moodle. Este sistema faz uso de *webcams* para a captura periódica de imagens da face do aluno, tirada em intervalos de tempo predefinidos, e utiliza a transformada de Cosseno Discreta (DCT) para extrair as características da face presente. O SIAF não é utiliza técnicas para a localização e extração das faces obtidas, o que influencia negativamente o desempenho do sistema. A Figura 13 ilustra a arquitetura do sistema.

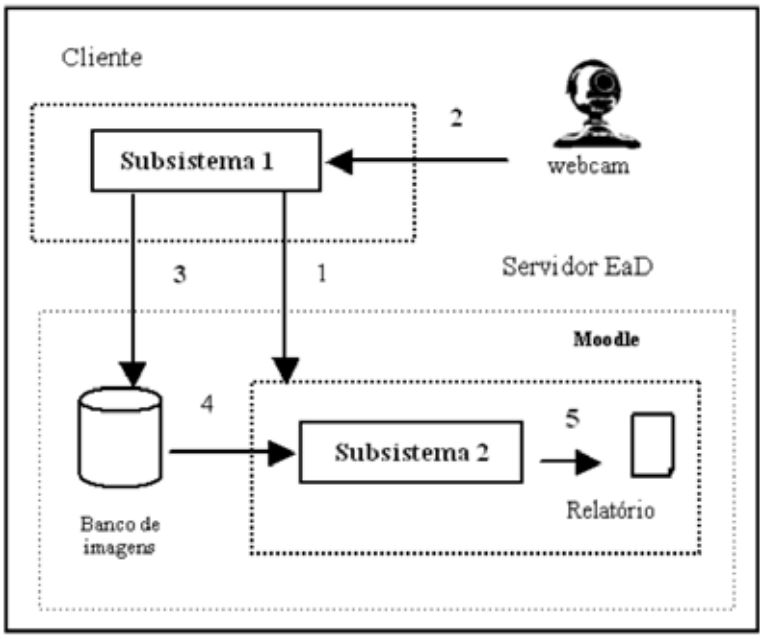


Figura 13. Arquitetura do SIAF (ROLIM e BEZERRA, 2008).

O SIAF é dividido em dois subsistemas. O Subsistema 1 é responsável por executar o navegador no ambiente Moodle (1). Em seguida, inicia o processo de captura de imagens em intervalos regulares a partir da *webcam* (2). As imagens capturadas são enviadas ao servidor de LMS (3) que realiza o reconhecimento das faces e as armazenam em um banco de imagens.

O subsistema 2 é uma extensão do Moodle que recupera as imagens (4) e as apresenta nos relatórios do próprio ambiente (5). A geração de relatórios tem a função de registro do resultado do conhecimento (afirmativo ou negativo) e trazem as imagens dos alunos durante sua presença no curso.

4.5 Considerações Finais

A autenticação de alunos têm papel importante em ambientes virtuais de aprendizagem, com dois propósitos principais: permitir que somente usuários autorizados possam acessar o sistema e garantir que a pessoa sendo avaliada é realmente quem ela diz ser.

Na literatura sobre o tema, a pouca preocupação com estes itens é apontada como séria deficiência. Segundo os autores, o processo de autenticação em sistemas de educação a distância deve ser contínuo, atuando não somente no acesso ao sistema, mas também durante todo o processo de aprendizado e avaliações.

O emprego de métodos biométricos para melhorar a autenticação dos usuários tem se apresentado como uma boa alternativa, uma vez que as informações usadas para verificação nesta abordagem não podem ser perdidas, roubadas ou adivinhadas, conferindo integridade ao processo de verificação de identidade.

Neste capítulo foram apresentados alguns casos de uso de tecnologias biométricas para a autenticação remota de indivíduos em sistemas de *e-Learning* de apoio à educação a distância. Dentre os exemplos apresentados, nenhum explora o vídeo para realizar a autenticação dos indivíduos e poucos fazem uma análise quantitativa do desempenho das técnicas utilizadas, como, por exemplo, taxa de reconhecimento, tempo de resposta, etc. Assim, esta dissertação se diferencia dos demais trabalhos mencionados neste capítulo pelo fato de levar em consideração esses aspectos.

No Capítulo 5 são mostradas técnicas e métodos selecionados para a aplicação de reconhecimento de faces a partir de vídeo neste trabalho, atendendo aos requisitos levantados.

Capítulo 5 – Sistema de Autenticação Biométrica

Neste capítulo são descritos a arquitetura distribuída proposta, levando em conta as características de aplicações de Internet, a implementação do sistema e o funcionamento dos módulos individuais adotados neste trabalho.

5.1 Introdução

Neste trabalho foram utilizadas as abordagens de reconhecimento de imagem-para-vídeo, quadro a quadro (conjunto de observações), com o método PCA explorando a grande quantidade de imagens presentes na amostra de vídeo e baseada em múltiplas visões dos indivíduos, conforme os conceitos levantados no Capítulo 3.

Para a construção do protótipo foram observadas as características discutidas na seção 2.3, que trata sobre as características e componentes de sistemas biométricos. A Figura 14, na seção 5.2, estende o funcionamento genérico descrito na Figura 2, exibindo os componentes básicos de um sistema biométrico.

As seções no restante deste Capítulo detalham a arquitetura geral proposta, os componentes e a sua distribuição entre cliente e servidor.

5.2 Arquitetura do Sistema

Para a avaliação da metodologia e das técnicas apresentadas foi projetado e implementado um sistema baseado na arquitetura cliente/servidor de modo a encapsular o acesso ao ambiente de

e-Learning, reforçando seu processo de autenticação (PENTEADO; MARANA, 2009). A Figura 14 mostra a disposição dos módulos, distribuídos entre cliente e servidor.

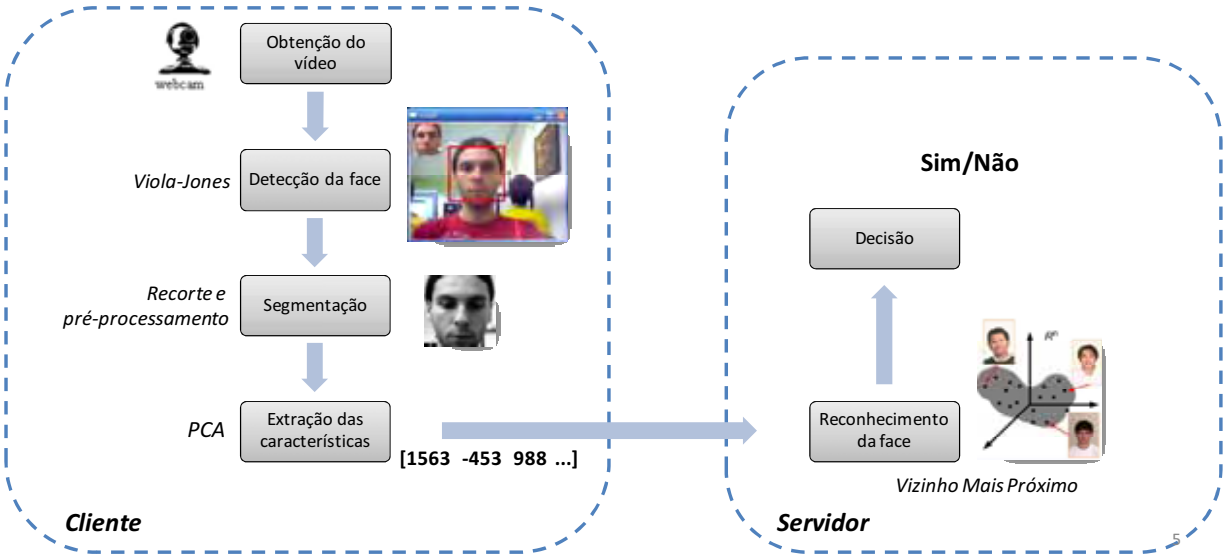


Figura 14. Arquitetura para o reconhecimento biométrico de faces adotado, aplicado em cada quadro do vídeo.

A captura dos dados biométricos do usuário a ser verificado é feito por uma *webcam*, que atua como sensor para a coleta das amostras de vídeo. Em seguida, para cada *frame* do vídeo, é detectada a face, que é segmentada do fundo e passa por pré-processamento de normalização de iluminação e escala. Na seqüência, são extraídos os vetores de características das faces pré-processadas, que serão usados para classificação com as faces previamente cadastradas na base de dados. Baseado na pontuação do casamento dos vetores de características da face de consulta e da face na base de dados o sistema emite sua decisão.

Devido à complexidade computacional dos algoritmos biométricos, alguns pontos devem ser levados em consideração ao se projetar um sistema para um grande número de usuários. A carga de processamento do sistema deve ser distribuída entre cliente e servidor para se chegar a um tempo de resposta razoável, bem como limitar a quantidade de informação trocada entre os clientes e o servidor. Assim, formas de distribuir o processamento e diminuir também o tráfego de dados devem ser observadas. Navegadores (*browsers*) de internet também não permitem

certos tipos de acesso a recursos no cliente, como acesso ao sistema de arquivos, a dispositivos periféricos, em favor da aplicação Web. Desta maneira, o modo como o vídeo deve ser coletado também deve ser levado em consideração.

Algoritmos de processamento de imagens e visão computacional apresentam, em geral, grande consumo de recursos computacionais, dada a natureza matricial das imagens e sua complexidade algorítmica. Para aplicações Web, em que vários usuários podem interagir simultaneamente com o sistema, esta deve ser uma preocupação, pois a experiência do usuário exige baixos tempos de resposta.

No sistema proposto, os seguintes subsistemas estão presentes:

- **Aplicação desktop**, que bloqueia ou permite o acesso ao curso, captura e extrai o vetor de características a ser enviado ao servidor. Neste módulo existe um componente de *browser* embutido, o qual é usado para interceptar as requisições ao servidor e suas respostas;
- **Servidor LMS**, que hospeda o sistema responsável pelo curso;
- **Serviço Web controlador**, responsável por gerenciar se uma página requisitada está marcada para ser verificada e processar o vetor de característica coletado; e
- **Base de dados biométrica**, usada para armazenar os *templates* dos usuários cadastrados.

A Figura 15 mostra o fluxo de controle do sistema proposto e seu fluxo de informação.

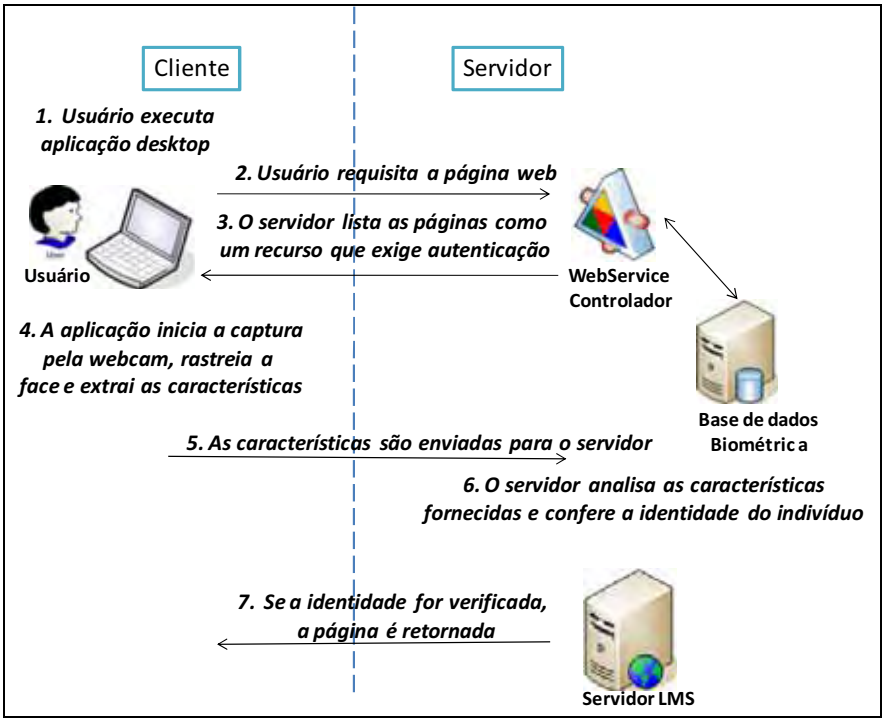


Figura 15. Interação entre os módulos localizados no cliente e no servidor.

O usuário, para poder freqüentar o curso, executa uma aplicação *desktop* em seu computador. Então, a aplicação requisita a página do curso para o servidor *Web*. Junto com a requisição, uma consulta é também enviada para verificar se a página Web sendo requisitada exige autenticação biométrica. Este poderia ser o caso para uma avaliação ou um conteúdo protegido, por exemplo. Se não for, a página é enviada de volta ao cliente. Caso contrário, a aplicação cliente começa a capturar o vídeo por meio da *webcam* do usuário. A aplicação processa o vídeo, detecta, pré-processa as faces presentes no quadros amostrados, concatena os vetores de características extraídos e envia-os para o servidor. O servidor consulta a base de dados pelo traço biométrico e retorna a resposta para o cliente *desktop*. A aplicação *desktop* então bloqueia ou retorna a página Web.

A Figura 16 mostra a interface do sistema criado, com um componente de navegador embutido na aplicação *desktop*.

Figura 16. Interface do sistema proposto.

Ao projetar a arquitetura do sistema, os seguintes itens foram considerados:

- **Redução do tráfego na rede**: transmite-se somente o vetor de características ao invés do *streaming* de vídeo, buscando reduzir o tráfego para somente o necessário à autenticação;
- **Balanceamento do processamento**: a aplicação cliente é responsável por rastrear e extrair as características das faces, deixando para o servidor a tarefa da classificação, visto que ele será usado por todos os clientes;
- **Permissões de segurança**: *browsers*, por padrão, não podem acessar recursos locais dos clientes como, por exemplo, dispositivos periféricos ou o sistema de arquivos. Ao permitir que a *webcam* seja capturada pela aplicação *desktop* no cliente, não existe a necessidade da instalação de controles adicionais no navegador, como *ActiveX* ou *Java Applets* que demandam permissões explícitas do usuário;
- **Portabilidade**: é possível a integração a qualquer sistema de gestão de aprendizado (LMS), independente da tecnologia na qual foi construído.

Dessa forma, o sistema apresenta tratamento para as questões levantadas para a arquitetura cliente/servidor, fazendo a distribuição dos procedimentos biométricos entre os envolvidos na interação.

5.3 Módulos

Nesta seção são descritos detalhadamente os métodos escolhidos em cada um dos módulos para a realização dos experimentos.

5.3.1 Detecção das Faces no Vídeo

Como primeiro passo em um sistema automático de reconhecimento de faces, deve-se localizar e segmentar a face presente na cena capturada. Uma localização imprecisa pode provocar grande degradação do desempenho do sistema.

O algoritmo utilizado neste trabalho para detectar as faces no vídeo foi o de Viola-Jones (VIOLA; JONES, 2001). Este algoritmo tenta encontrar em uma imagem características que codificam alguma informação do padrão a ser detectado. Para tal tarefa, são usadas as características de *Haar*, ilustradas na Figura 17, à esquerda, responsáveis por codificar, de modo multi-escala, informações sobre a existência de contrastes orientados entre regiões da imagem. No caso, são explorados os contrastes naturais da face, respeitando seus relacionamentos espaciais, como ilustra a Figura 17, à direita.

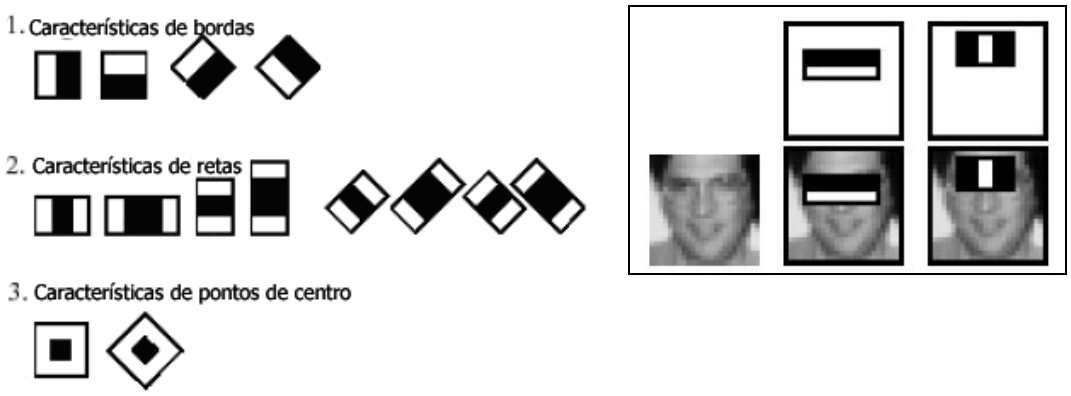


Figura 17. Esquerda: Características de *Haar* comumente utilizadas para detecção de faces. Direita: relacionamento entre as características de *Haar* e os contrastes específicos da face (VIOLA; JONES, 2001).

Tais características podem ser computadas eficientemente usando uma representação intermediária para a imagem original, denominada *imagem integral*. O ponto *ii(x, y)* na imagem integral corresponde à soma de todos os pixels acima e à esquerda de *ii(x,y)*, inclusive. Usando-se a recorrência da Equação 2:

$$\begin{aligned} s(x, y) &= s(x, y-1) + i(x, y) \\ ii(x, y) &= ii(x-1, y) + s(x, y) \end{aligned} \tag{2}$$

onde *s(x,y)* é a soma cumulativa da linha, $s(x, -1) = 0$ e $ii(-1, y) = 0$. A imagem integral pode ser calculada em um único passo sobre a imagem original.

A Figura 18a ilustra a representação da imagem integral. O ponto *ii(x,y)* ilustrado contém o valor da soma de todos os pixels da linha x e da coluna y, desde a origem $(x,0)$ e $(0,y)$, respectivamente, até o próprio ponto (x,y).

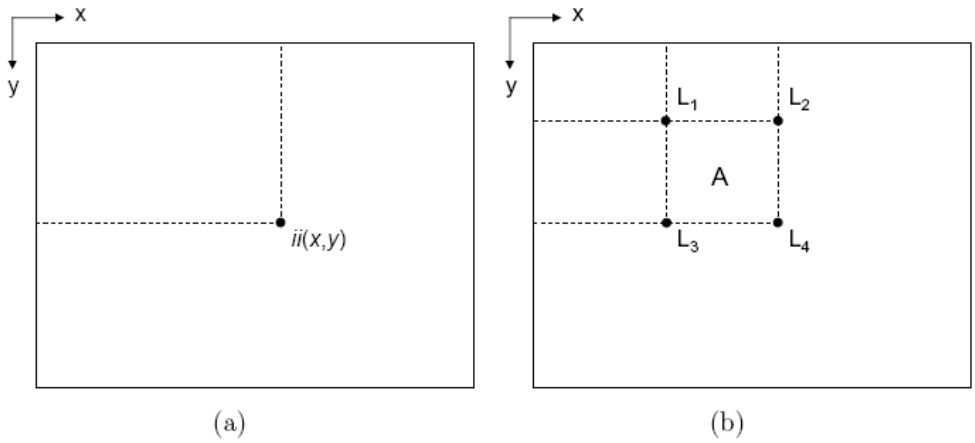


Figura 18. Representação da imagem integral: (a) valor do ponto *ii(x,y)* na imagem integral refere-se à soma dos valores de todos os pixels exatamente acima e à esquerda dele próprio (linha pontilhada); (b) região A pode ser eficientemente calculada usando apenas os valores dos pontos L4 + L1 – (L2 + L3), sem a necessidade de outros cálculos (VIOLA; JONES, 2001).

Uma vez calculada a imagem integral, qualquer das características de *Haar* pode ser computada em qualquer escala ou local em tempo constante, com simples somas e subtrações, como ilustrado na Figura 18b. Como cada ponto já contém o somatório das linhas e colunas até chegar a si próprio, é possível calcular o tamanho de qualquer região apenas com os valores de *pixels* dos cantos da região. Desta maneira, para detectar os padrões dentro da imagem o detector é redimensionado e não a imagem em si. Na seção 6.2 é realizado um experimento para analisar o impacto do tamanho das subjanelas sobre o tempo total do processamento da detecção da face.

Entretanto, dentro de uma subjanela de qualquer tamanho, o número de combinações das características de *Haar* possível, tanto em escala quanto em localização, é muito maior que o número de pixels dentro da subjanela. Para tornar o processo de classificação rápido, deve-se excluir a maioria das características disponíveis e manter somente um pequeno conjunto de características mais discriminantes. Para esta tarefa, é adotado o procedimento modificado do método *AdaBoost* (TIEU; VIOLA, 2000), que consiste no processo de seleção de características que constrói um classificador "forte" baseado na combinação de subclassificadores "fracos" que dependem de uma única característica.

Outro passo, que diminui significativamente o tempo de processamento deste algoritmo, refere-se à combinação desses classificadores fracos em uma estrutura de cascata ou uma árvore de decisão degenerada. Um resultado positivo do primeiro classificador dispara a avaliação de um segundo classificador que também foi ajustado para conseguir altas taxas de detecção. Um resultado positivo do segundo classificador dispara um terceiro classificador, e assim por diante. Uma resposta negativa em qualquer um dos classificadores leva à imediata rejeição da subjanela. Deste modo, a detecção do padrão em uma subjanela somente tem êxito caso seu resultado seja positivo em todos os classificadores. A Figura 19 ilustra o funcionamento da cascata de classificadores.

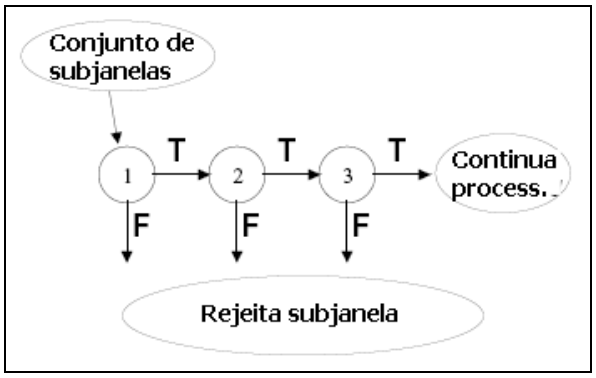


Figura 19. Funcionamento da cascata de classificadores do algoritmo Viola-Jones: a janela é rejeitada caso não passe por algum dos classificadores (VIOLA; JONES, 2000).

Tal estrutura se baseia no fato de a maioria das subjanelas serem classificadas como negativas. Os primeiros classificadores são mais simples, de cálculo mais rápido, o que concentra o esforço computacional em regiões promissoras na imagem. Nos resultados apontados em (VIOLA; JONES 2000), com o primeiro classificador usando apenas 2 características de *Haar*, é possível rejeitar 60% de padrões que não são faces, ao passo que detecta corretamente perto de 100% dos padrões de face.

A Figura 20 mostra o algoritmo Viola-Jones agindo sobre um quadro do vídeo e detectando a face em tempo real.

Figura 20. Algoritmo de detecção de face Viola-Jones aplicado em um vídeo.

A principal característica desta técnica diz respeito à rapidez com que as faces são detectadas, mesmo apresentando variações de escala. O conjunto de passos do algoritmo faz com que ele propicie uma taxa de detecção tão boa quanto outras apresentadas na literatura, porém, com um tempo de processamento consideravelmente menor. Por este motivo, esta técnica foi a selecionada para este trabalho, pois aplicações que tem resposta imediata ao usuário, como é o caso da autenticação em sistemas de *e-Learning*, demandam processamento perto de tempo real.

5.3.2 Segmentação e Pré-Processamento

As imagens obtidas a partir da *webcam* foram amostradas quadro a quadro e as faces foram detectadas usando o algoritmo Viola-Jones na implementação da biblioteca OpenCV (OPENCV 2008) discutido na seção 5.3.1.

Como etapa de pré-processamento, as imagens das faces foram extraídas e convertidas para escala de cinza, redimensionadas para um tamanho padrão (64x64 *pixels*, escolhido empiricamente), por interpolação bilinear, e foi aplicada equalização de histograma para ajustar seu contraste (Figura 21).

Figura 21. Etapa de pré-processamento: (a) imagem extraída e redimensionada, (b) convertida para escala de cinza e (c) depois da equalização de seu histograma.

5.3.3 Extração das Características

Pelo fato de o método da Análise das Componentes Principais (PCA – *Principal Component Analysis*) ser o primeiro método de reconhecimento de faces a ser aplicado com sucesso e ainda hoje ser usado como algoritmo base em comparações entre novas propostas de algoritmos de reconhecimento de faces, ele foi o método escolhido para o desenvolvimento deste trabalho.

Usado inicialmente nos trabalhos de Turk e Pentland (TURK; PENTLAND, 1991) e Kirby e Sirovich (KIRBY; SIROVICH, 1990) para análise, representação e reconhecimento facial, o método PCA tem servido como base nas comparações de diversos algoritmos de reconhecimento facial.

É conhecida a existência de redundâncias estatísticas significativas em imagens de faces (RUDERMAN, 1994). Este método atua na descorrelação dos dados presentes em uma imagem, por meio da análise da variância em cada uma de suas dimensões[10] (representada como um vetor), para minimizar o esforço computacional sem perder a discriminação entre as classes.

[10] O termo dimensão aqui se refere a um elemento do vetor, depois de a imagem ser transformada em um vetor coluna, concatenando cada coluna da imagem ao final do vetor.

Assim, a análise de componentes principais reduz a dimensionalidade dos dados iniciais, mantendo as dimensões de maior variância e descartando as demais. Desta forma, os dados são comprimidos sem grandes perdas de informação, por meio de uma transformação linear:

$$y = T(x): R^n \rightarrow R^m, \text{ com } m<<n \quad (3)$$

No caso de uma imagem de tamanho $n = h$ x w, uma imagem de face corresponde a um ponto no espaço vetorial R^n. Devido à alta correlação dos pixels de imagens de faces, os dados das faces correspondem a apenas um subespaço de menor dimensionalidade. A análise das componentes principais identifica e representa eficientemente este subespaço de m-dimensões (R^m). A Figura 22 ilustra a composição deste subespaço. A fase de treinamento deste algoritmo consiste em encontrar o conjunto de autovetores (bases) que formam este espaço de faces.

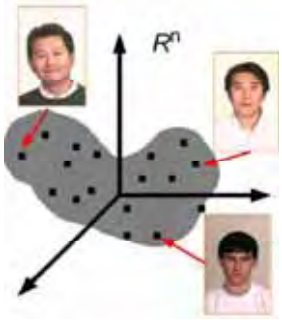


Figura 22. Representação do subespaço das faces no espaço das imagens de entrada na fase de treinamento (JAIN, 2004).

As bases deste novo espaço de m-dimensões são obtidas pela solução do seguinte problema:

$$\lambda = \Phi^T \sum \Phi \quad (4)$$

onde $\sum$ é a matriz de covariância entre as n dimensões, Φ é a matriz de autovetores de $\sum$ e λ é a matriz diagonal contendo os autovalores λ_i.

A Figura 23 ilustra os passos básicos para o reconhecimento de faces baseado no método de análise das componentes principais (PCA).

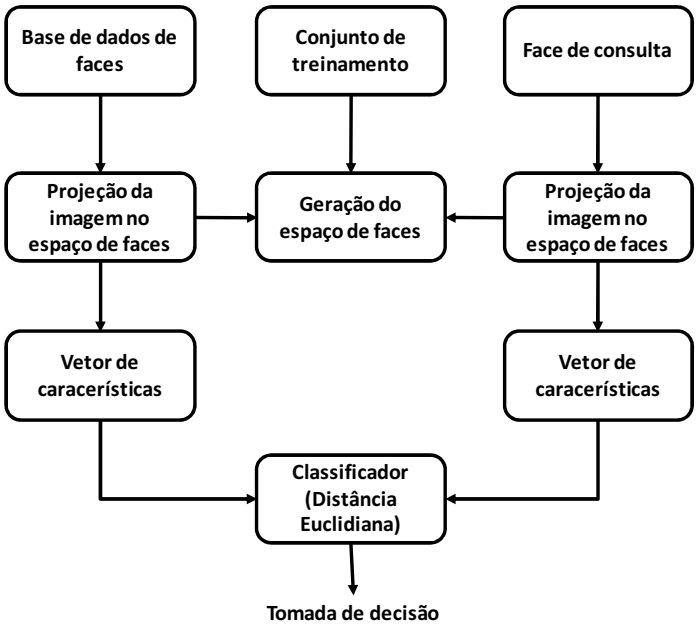


Figura 23. Diagrama do algoritmo PCA para o reconhecimento de faces

De início, um conjunto de imagens representando o domínio da aplicação é selecionado para o treinamento, ou seja, para a criação do espaço de faces. Em um segundo momento, é criada a base de dados e os usuários são cadastrados. As imagens de suas faces são coletadas e projetadas neste espaço, gerando o vetor de características que representa a face e estas são armazenadas na base de dados. Ao se fazer uma consulta usando uma face de entrada, esta mesma é projetada no mesmo espaço de faces, seu vetor de características é gerado e, deste modo, ele é comparado com os vetores armazenados, por meio de algum classificador, conforme detalhado na seção 5.3.4.

5.3.4 Reconhecimento da Face

Uma vez gerado o espaço de faces, as imagens de face a serem cadastradas na base de dados são nele projetadas. Após a projeção, tem-se a representação da face em termos da combinação linear da nova base dos *m* autovetores identificados previamente, na fase de treinamento. Os coeficientes obtidos são então usados como vetores de características representando as faces.

A distância entre os vetores de características da face projetada e das faces existentes da base de dados pode ser determinada por várias métricas. Neste trabalho, foi utilizada a distância Euclidiana:

$$d(p,q) = \sqrt{\sum_{i=1}^{M} (p_i - q_i)^2} \tag{5}$$

onde *d(p,q)* denota a distância entre os vetores de características *p*, de uma imagem cadastrada na base de dados e *q*, da imagem de consulta e p_i e q_i denotam os correspondentes elementos dos vetores de características das imagens *p* e *q* na dimensão *i*.

Pelo fato de a identidade atribuída à imagem de consulta ser a da amostra cadastrada na base de dados que tenha menor distância entre os respectivos vetores de características, é também chamado de *vizinho mais próximo*.

5.3.5 Decisão / Fusão dos Resultados

De modo a ter uma decisão única dentro do conjunto de quadros analisados, necessita-se de uma fusão dos resultados obtidos a cada quadro. Para fundir os resultados de cada quadro individualmente, foi usada a regra de maioria de votos. Nesta abordagem, a cada quadro, o classificador atribui uma identidade à face encontrada, evento representado como uma função binária:

$$d_{i,m} = \begin{cases} 1, se\ o\ resultado\ do\ i-ésimo\ quadro\ for\ a\ classe\ m \\ 0, caso\ contrário \end{cases}$$

Pela maioria dos votos, a identidade final é escolhida por:

$$ID^{(M)} = \arg\max_{i=1..M} \left(\sum_{i=1}^{N} d_{i,m} \right) \tag{6}$$

onde d_i representa a identidade da decisão do algoritmo para o frame *i* dentre os *N* existentes e *M* representa as identidades cadastradas.

Assim, é atribuída à face a identidade na qual a maioria das decisões estiver de acordo, tratando cada decisão como um evento independente. Como cada decisão é tomada pelo mesmo algoritmo, nenhuma outra forma de normalização ou ponderação é necessária.

5.5 Considerações Finais

Neste capítulo foram apresentados os métodos usados em cada um dos módulos da arquitetura proposta do sistema biométrico implementado para autenticação biométrica de usuários em sistemas de *e-Learning*.

Para tornar o processo de autenticação mais confortável e natural para o usuário, foi proposto um sistema que busca reproduzir seu comportamento ao usar um ambiente virtual de aprendizagem, de modo colaborativo e não-colaborativo.

No Capítulo 6 são apresentados os resultados experimentais e as análises sobre a aplicação das técnicas selecionadas sobre o material coletado.

Capítulo 6 - Resultados Experimentais

Neste capítulo são apresentados os resultados obtidos com a aplicação e avaliação do sistema e dos métodos descritos no Capítulo 5 e sobre o conjunto dos vídeos coletado conforme descrito na seção 5.2.1. São também apresentadas as análises dos dados levantados.

6.1 Material

Nesta seção são descritos o conjunto de vídeos utilizado neste trabalho bem como *software* e *hardware* usados na implementação dos métodos e na realização dos experimentos.

6.1.1 Conjunto de Vídeos

Os vídeos usados neste trabalho foram criados especificamente para os experimentos, na tentativa de replicar a interação natural de um indivíduo navegando por um curso de *e-Learning*. Os vídeos foram coletados em 2 sessões distintas, com poucas semanas de intervalo entre elas.

A coleta foi feita usando uma *webcam* modelo *Creative Webcam Pro eX PD 1050*. Ambos os conjuntos foram coletados sem grande variação de iluminação no ambiente. A Figura 24 mostra como o equipamento foi montado para o experimento.

Figura 24. Equipamento usado nos experimentos.

Na primeira sessão de coleta das amostras de vídeo, foi solicitado aos indivíduos que navegassem em um determinado site e respondessem a uma questão de texto livre. Os

participantes foram orientados a variar suas poses e expressões faciais durante a sessão. Foram coletadas 45 amostras (uma por pessoa) de, em média, 1 minuto e meio.

Na segunda sessão, os participantes foram orientados a realizar os seguintes passos, de modo a simular o comportamento em um sistema de *e-Learning*:

- Olhar frontalmente para a câmera durante alguns segundos, de modo a simular sua autenticação no sistema LMS, colaborativamente;
- Ler um texto de aproximadamente 500 caracteres, de modo a simular a leitura de um conteúdo do curso disponível no sistema LMS;
- Preencher um formulário com certos dados pessoais, de modo a simular o preenchimento de questionários e avaliações no sistema LMS.

As páginas com as instruções usadas para a coleta dos vídeos encontram-se no Apêndice A do presente trabalho.

Em ambas as sessões os vídeos foram coletados em ambiente interno, com iluminação controlada, apresentando pequenas variações decorrentes do período do dia em que foram gravados. Não foram impostas restrições quanto à pose ou ao uso de acessórios. A Figura 25 mostra alguns exemplos das amostras coletadas para o experimento.

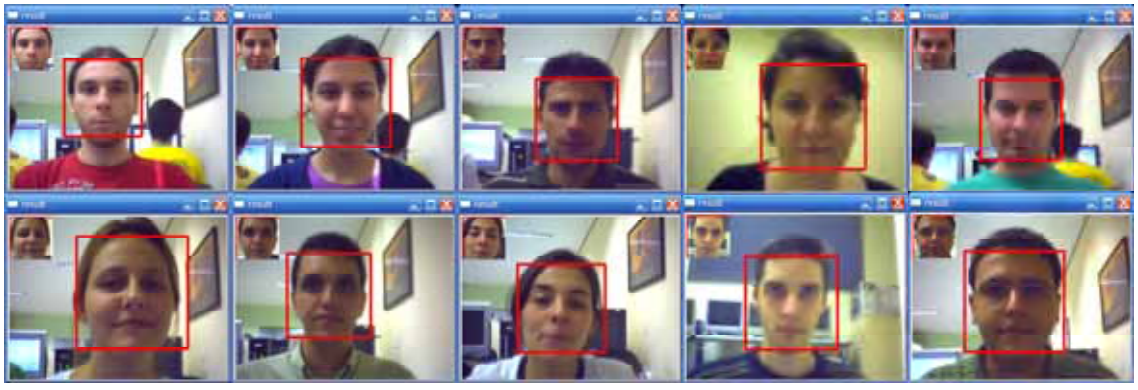


Figura 25. Exemplos dos vídeos coletados.

6.1.2 Software e Hardware

Para a captura dos vídeos, bem como para os métodos de detecção de faces (Viola-Jones) e de representação das faces (PCA) foram utilizadas as implementações padrão da biblioteca de código livre e multiplataforma OpenCV (OPENCV, 2008). Esta biblioteca contém vários algoritmos e métodos de visão computacional escritos nativamente em C e com rotinas otimizada para processadores da arquitetura *Intel*.

Como ambiente integrado de desenvolvimento do sistema implementado foi utilizada a IDE *Microsoft Visual Studio 2005*, com as linguagens C/C++ padrão, para extensões da biblioteca OpenCV, e C++.Net, com suas bibliotecas de acesso a dados, interface com o usuário e comunicação entre processos. O banco de dados escolhido para armazenar as faces foi o *Microsoft SQL Server 2000*.

Todos os experimentos foram realizados em um laptop modelo *Acer Aspire* 3624, composto por um processador Intel Celeron M 1.6 GHz e 1 GB de memória RAM.

6.2 Organização do Experimento

Para efetuar a coleta dos dados no contexto de uma aplicação Web, foi seguida a metodologia descrita na seção 5.2.1, com a divisão em 2 sessões.

Os quadros amostrados de cada um dos vídeos na primeira sessão foram usados para a geração do espaço de faces (treinamento do algoritmo PCA), para a criação do conjunto dos autovetores. A Figura 26 mostra a face média e os nove autovetores com maiores autovalores associados gerados a partir deste conjunto. Este conjunto de autovetores forma a base do novo espaço de faces no qual as imagens serão projetadas. O vetor de características de cada face projetada será composto pelos coeficientes de sua projeção no espaço de faces.

Figura 26. Face média (primeira à esquerda, na primeira linha) e os autovetores de maior variância obtidos a partir de conjunto de treinamento.

Como mencionado na seção 5.3, os *pixels* da imagem de uma face são altamente correlacionados entre si. Boa parte deles não contém informação discriminatória, ou seja, não servem para diferenciar os indivíduos entre si. Assim, conhecendo quais são as dimensões de maior variância, pode-se eliminar as que menos contribuem para a separabilidade dos dados.

A característica mais importante da análise das componentes principais (PCA) é a redução da dimensionalidade dos dados a serem analisados. Ou seja, este método permite que apenas as dimensões que contêm maior variância na distribuição dos dados sejam mantidas, e, por conseguinte, possibilitam maior discriminação entre as classes.

A Figura 27 mostra a quantidade de informação retida pelos autovetores associados aos maiores autovalores dentro do conjunto de treinamento coletado. Com isso, apenas uma pequena quantidade de autovetores pode representar eficientemente todo o subespaço das faces.

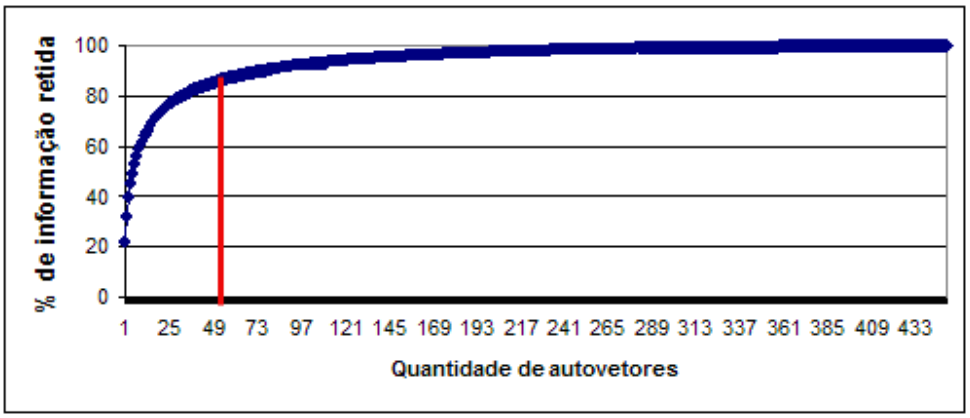


Figura 27. Relação entre os autovetores e sua representatividade

Neste trabalho, para a realização dos experimentos, apenas os 50 autovetores associados aos maiores autovalores foram utilizados, o que corresponde aproximadamente a 85% da informação de variação contida no conjunto de treinamento. Pode-se notar pela Figura 27 que, a partir de certo número de autovetores (em torno de 200), a quantidade de informação descorrelacionada é pouco relevante – menos de 3%. Este número é condizente com os valores apresentados no estudo de Moon e Phillips (2001) sobre os efeitos do número de autovetores no desempenho do método PCA para reconhecimento de faces.

Em posse dos vídeos coletados na segunda sessão, foram manualmente recortados trechos de 5 segundos para cada pessoa em cada pose. Para a construção da base de dados, foram escolhidos três quadros representativos para cada pose e indivíduo (total de nove para cada indivíduo). A Figura 28 mostra as imagens selecionadas para um determinado indivíduo.

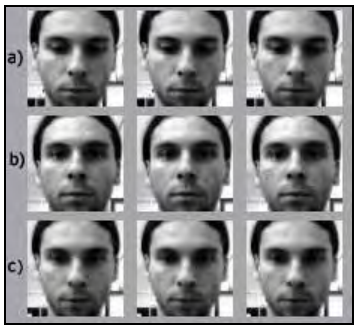


Figura 28. Imagens selecionadas de um indivíduo nas poses: (a) escrevendo, (b) frontal, (c) lendo.

Para tal escolha, foi utilizado o método proposto em (THOMAS; BOWYER, FLYNN 2007), que determina a variação de uma face em relação a um dado conjunto de faces. Neste método, a variação é definida pela distância entre as imagens projetadas no espaço de faces, consideradas como pontos em R^m. Uma série de quadros de um mesmo indivíduo é projetada no espaço de faces previamente treinado. Então, é aplicado o Co-seno de Mahalanobis como medida de distância entre as imagens projetadas, duas a duas. Deste modo, a primeira imagem selecionada é aquela cuja distância total para todas as outras é a maior. Neste trabalho foram selecionadas as faces que menos variam dentro do conjunto de faces da amostra de vídeo, pois as

de maior variação apresentaram ruídos (*outliers*) como bocejo e pequenas oclusões, o que acaba por prejudicar o processo de autenticação.

A Figura 29 mostra as imagens de face selecionadas para um indivíduo em sua pose frontal, conforme a aplicação da técnica descrita.

Figura 29. Imagens de face mais divergentes (acima) menos divergentes (abaixo) na pose frontal para o trecho de vídeo usado na fase de cadastro.

Além destes módulos, outro para o cadastro dos indivíduos na base de dados foi construído, baseado na mesma metodologia usada durante a fase de coleta das amostras de treinamento.

6.3 Resultados

Os dados obtidos aqui foram coletados e avaliados *off-line*, ou seja, antes da implementação do protótipo do sistema, com exceção da mensuração do tamanho do *template* gerado (PENTEADO; MARANA, 2008).

Para o problema de detecção de faces, duas medidas são importantes, conforme a seção 2.3.3: a taxa de falsos positivos, quando um padrão que não é uma face é aceito como uma face; e taxa de falsos negativos, quando uma imagem de face não é reconhecida. Um sistema de detecção eficiente deve apresentar taxas baixas tanto para falsos positivos quanto para falsos negativos, e, para aplicações Web em particular, deve também apresentar um tempo de resposta em tempo real.

De modo a medir o desempenho do sistema foram feitas as seguintes análises:

- **Eficiência do algoritmo detector de faces, em termos de taxas de falsa aceitação e falsa rejeição**: como as faces a serem reconhecidas devem ser extraídas do *frame* onde se encontram, o desempenho global do sistema tem grande dependência do resultado do passo de detecção e extração da face.

- **Número de autovetores selecionados e suas respectivas taxas de reconhecimento, tanto em nível de *frame* quanto em nível de amostra**: este é um dado importante, pois a quantidade de autovetores influencia o custo computacional despendido durante a autenticação. Como se trata de um sistema Web, sob a arquitetura cliente/servidor, é um requisito que a aplicação tenha um baixo tempo de resposta. A análise em nível de frame é relevante para se ter uma idéia do grau de acerto do algoritmo no nível mais baixo de classificação, antes da fusão dos resultados. Em nível de amostra (trechos individuais), já se leva em consideração a fusão por maioria de votos.

- **Taxa de reconhecimento em relação aos *top* N usuários retornados na busca à base de dados**: esta informação é especialmente útil no cenário de autenticação biométrica, em que o usuário diz quem é, e o sistema procura se, dentre as *top* N identidade retornadas, está a identidade alegada.

- **Tamanho, em bytes, do *template* enviado para verificação no servidor:** este dado impacta diretamente no tempo de resposta da aplicação, uma vez que os dados extraídos no cliente devem ser completamente enviados para o servidor para se dar o processo de autenticação. Quanto maior a quantidade de dados trafegada, maior a latência do sistema.

Para se avaliar o algoritmo de detecção e rastreamento, foram levados em consideração:

- Tamanho mínimo da subjanela a partir da qual a face será localizada;

- Número de falsos positivos,

- Número de falsos negativos e o

- Tempo de processamento das amostras.

A Figura 30 mostra o tempo de processamento gasto variando-se o tamanho da subjanela. Este tamanho serve como ponto de partida para a detecção de faces dentro do frame. Padrões de face menores que esta subjanela são ignorados.

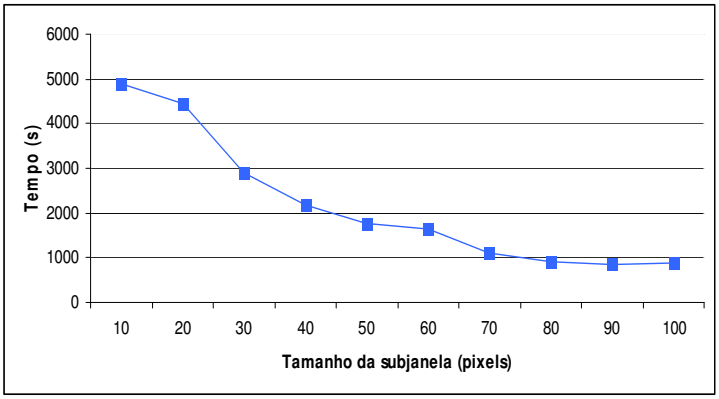


Figura 30. Tempo de processamento do conjunto de amostras em relação ao tamanho da janela de varredura.

A face ocupa, em média, por volta de 100x100 *pixels*, devido à proximidade do usuário ao computador. Os quadros capturados apresentam resolução de 320x240 *pixels*. Com uma subjanela de 100x100, o processamento é praticamente em tempo real. Isto traz também outras conseqüências: diminuição do número de falsos positivos capturados (não-faces aceitas) bem como aumento dos falsos negativos (rejeição de faces). A Figura 31 e 32 ilustram este fato.

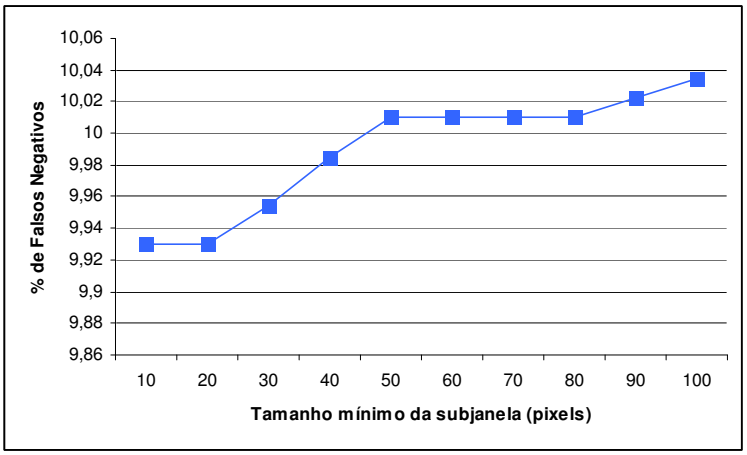


Figura 31. Porcentagem de faces não detectadas pelo algoritmo, considerando o conjunto de amostras.

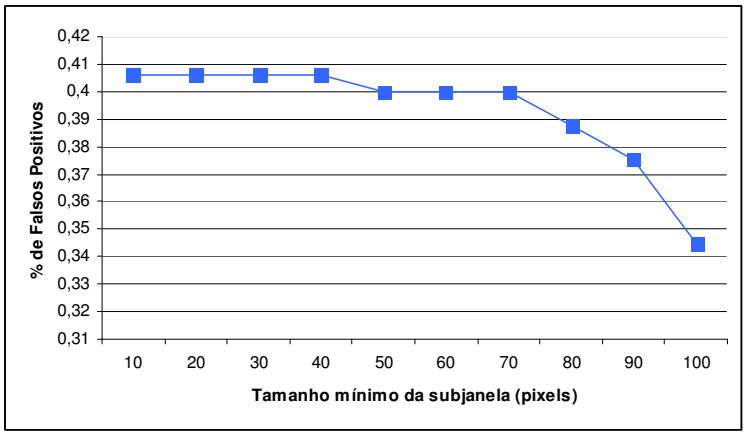


Figura 32. Porcentagem de padrões que não são faces e que foram detectados, considerando o conjunto de amostras.

A taxa dos falsos positivos influencia negativamente o desempenho do sistema, pois a tentativa de identificação de padrões que não são faces pode prejudicar o classificador de toda a amostra do vídeo. Deste modo, é desejável mantê-la a menor possível.

Quanto à taxa de falsos negativos, espera-se que não influencie significativamente o desempenho global do sistema. Isto devido à abundância de informação presente no vídeo: mesmo com 10% de rejeição, ainda tem-se cerca de 90% das faces detectadas, sendo suficiente para realizar o reconhecimento.

O uso de janelas de varredura de pequena dimensão faz com que exista um número muito maior comparações a serem feitas dentro de uma mesma imagem do que com grandes subjanelas, aumentando o tempo total de processamento. Assim, ao se trabalhar com janelas maiores, proporciona-se um menor tempo de resposta ao passo que se mantém a taxa de reconhecimento das faces.

Outro fator que influencia o tempo de processamento do sistema é a quantidade de autovetores utilizados para representar a face, o que impacta no tamanho de seu vetor de características. Como se pode notar na Figura 33, a partir de 25 autovetores, a taxa de reconhecimento tende a crescer pouco, mas ainda assim, mantendo um bom nível de reconhecimento (mais de 90%).

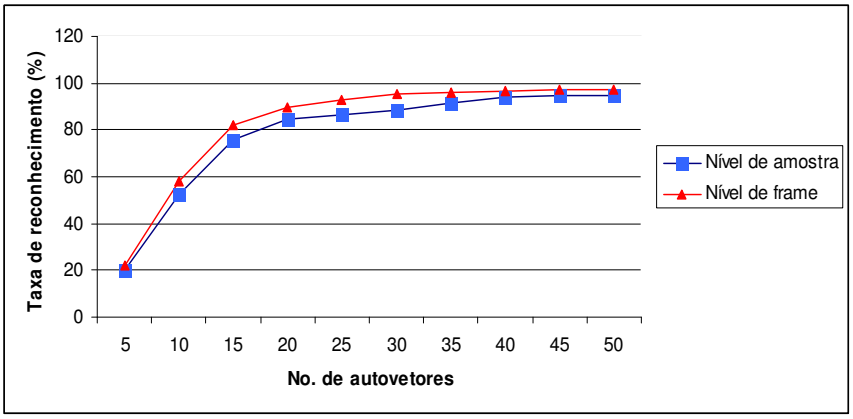


Figura 33. Taxa de reconhecimento em função da quantidade de autovetores considerada (usando top 1, nível de frame)

Para avaliar o impacto da fusão das identidades estabelecidas, foi aplicada a técnica de maioria de votos no conjunto de amostras. Ainda na Figura 33 é mostrada a taxa de reconhecimento em um nível mais alto, levando em conta o conjunto dos quadros, ou seja, a amostra de vídeo como um todo. Pode-se notar que as taxas são levemente inferiores às taxas de reconhecimento em nível de frame, o que deixa espaço para experimentos com novas técnicas de fusão mais elaboradas.

Outra medida importante é a quantidade de identidades necessárias para que o sistema possa autenticar corretamente o indivíduo (*top N*). A Figura 34 mostra que a primeira identidade retornada pelo método já apresenta índice de acerto de 97,2% em relação ao total de quadros.

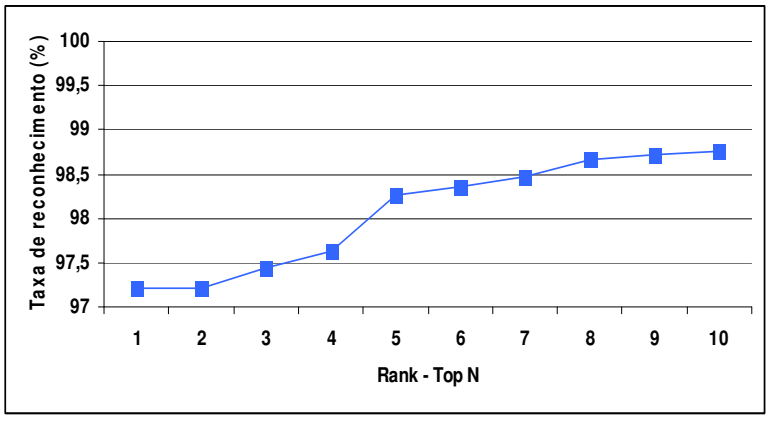


Figura 34. Taxa de reconhecimento em função das top 10 identidades retornadas por frame (usando 50 autovetores).

Com relação ao protótipo, também foi estimada a quantidade de informação trafegada pela rede para que o *template* do usuário, extraído e processado no cliente, seja enviado para o servidor para sua verificação. Tal medida também é importante para avaliar o tempo total de resposta da aplicação considerando o meio usado para a comunicação entre cliente e servidor. Na Figura 35 é apresentada uma variação linear do tamanho em *bytes* do *template*, conforme esperado. Como o tamanho da amostra de vídeo foi o mesmo para cada experimento, a diferença ficou por parte da quantidade de elementos nos autovetores, que também foi linearmente variada.

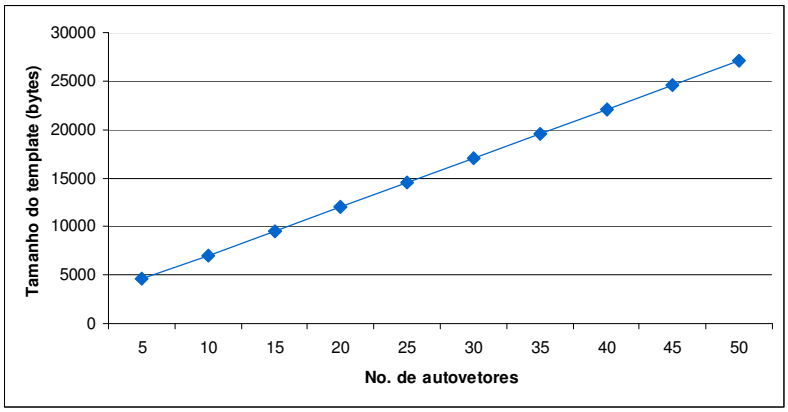


Figura 35. Quantidade de dados relativa ao *template* do usuário enviada para o servidor (em bytes).

Pode-se notar pelo gráfico que a quantidade de informação trocada com o servidor é relativamente pequena (por volta de 30 *Kilobytes* usando 50 autovetores). Levando em consideração a largura de banda disponível comercialmente atualmente, esta abordagem mostra-se factível do ponto de vista do tempo de resposta ao usuário.

Esta mesma abordagem, baseada em autovetores, pode ser usada como fonte de características para diferentes algoritmos de verificação, uma vez que *templates* baseados em PCA podem ser usados como entrada por outros tipos de classificadores, como redes neurais (OH, 2005) ou Máquinas de Vetor de Suporte - SVM (WANG; YANG; LIAO, 2007) entre outros, tornando desacoplados os módulos de extração de características, no cliente, e o classificador, no servidor mantendo-se a mesma estimativa de tamanho do *template*.

Capítulo 7 – Discussão e Conclusões

Neste Capítulo são apresentadas as conclusões e as contribuições deste trabalho. Também são mostradas as limitações da proposta que não foram tratadas neste trabalho e que foram identificadas durante o seu desenvolvimento.

7.1 Discussão

Este trabalho apresenta uma alternativa e/ou complemento baseado na Biometria aos métodos atuais de autenticação de indivíduos pela Internet, normalmente baseado em senhas. Os resultados indicam que as técnicas biométricas de reconhecimento de faces a partir de vídeo podem ser aplicadas com sucesso para este problema, ainda que com limitações. Aproveitando-se da interação natural com sistemas *Web*, a autenticação biométrica pode ser usada para garantir que o indivíduo genuíno está acessando o sistema remotamente, o que pode ser especialmente útil em um contexto de educação a distância (sistemas de *e-Learning*).

Este trabalho mostra a viabilidade do uso das técnicas biométricas de verificação de identidade para o problema da autenticação remota de indivíduos pela Internet, com aplicação em sistemas de *e-Learning*, ao propor uma arquitetura distribuída que faz uso do vídeo como fonte de dados e que faz balanço de processamento entre cliente e servidor com pequena quantidade de informação trafegada.

Para este fim, no Capítulo 2, foram levantados os componentes comuns a um sistema biométrico e seus indicadores de desempenho. Com base nestas informações foram definidas a arquitetura do sistema proposto e as métricas a serem analisadas nos experimentos.

No Capítulo 3 foram discutas os conceitos, peculiaridades e distinções do reconhecimento de faces a partir do vídeo, bem como o levantamento dos trabalhos mais relevantes neste tópico.

Com isso, contribuiu para a classificação deste trabalho dentre categorias levantadas, buscando tirar proveito de suas vantagens.

No Capítulo 4 foi abordada a contextualização do problema da autenticação remota de indivíduos em sistemas de *e-Learning*. Foi abordado o alcance deste tipo de educação, em termos de crescimento no número de matrículas e de cursos e foi verificada na literatura a crítica pela deficiência de mecanismos eficientes para a verificação confiável da identidade dos alunos. Este fato motivou o desenvolvimento deste trabalho, do mesmo modo que para a adaptação do sistema para a Web.

Pela natureza da arquitetura cliente/servidor dos sistemas *Web* de *e-Learning*, foi projetada uma arquitetura eficiente, de modo a se distribuir a carga computacional e o tráfego de informações entre as estações cliente e servidor, independentemente do sistema de gerenciamento de aprendizado utilizado. O sistema proposto tem um bom desempenho dentro das condições expostas, porém com espaço para melhorias. O algoritmo de detecção de faces mostrou-se eficiente tanto na taxa de acertos quanto no tempo de processamento. A taxa de reconhecimento também apresentou níveis aceitáveis.

Para o reconhecimento a partir do vídeo é recomendada uma maior investigação. Modelos que exploram melhor as características intrínsecas do vídeo devem ser investigados para avaliar sua efetividade nesta aplicação. Os algoritmos presentes na literatura podem, inclusive, gerar *templates* mais compactos que o vetor de autovetores utilizado neste trabalho, diminuindo assim o total de dados trafegado pela rede.

7.1.1 Contribuições

De um modo geral, o desenvolvimento deste trabalho contribui com os seguintes resultados:

- Pesquisa sobre arquitetura eficiente para a realização da autenticação biométrica em sistemas Web.

- Validação da eficiência, em termos de tempo de processamento e de taxa de erros, de métodos biométricos quando aplicados no contexto da análise de vídeos pela Internet;
- Exploração de uma abordagem tecnológica alternativa para o problema da autenticação remota de indivíduos no contexto de um ambiente de *e-Learning* na Internet;
- Pesquisa sobre método para verificação contínua da identidade do usuário realizando o curso a distância, em vez da autenticação única no acesso ao sistema;
- Estudo sobre as características e propriedades do reconhecimento de faces a partir do vídeo como fonte de dados para a autenticação biométrica e seu uso para o aperfeiçoamento da autenticação por reconhecimento de faces de maneira não intrusiva para o usuário;
- Pesquisa na elaboração de uma metodologia para construção de base de dados que possa representar os traços biométricos de uma pessoa de maneira eficiente e simples a partir de uma amostra de vídeo, ao simular sua interação com o sistema, a partir de uma *webcam* comum;

7.1.2 Limitações

Com a utilização do protótipo foram detectadas as seguintes questões que podem afetar significativamente o desempenho (em termos de taxas de reconhecimento e verificação) do sistema. São questões importantes que devem ser tratadas, mas que fogem ao escopo deste trabalho.

- **Variações de iluminação no ambiente**: o sistema apresentou bom desempenho em ambientes que apresentaram iluminação uniforme, bem distribuída e sem grandes variações durante a coleta. Porém, foi notado que a variação entre diferentes graus de iluminação pode ser maior que a variação

entre indivíduos, constituindo no fator que mais influenciou negativamente os testes realizados.

- **Oclusão parcial da face**: em certas situações o usuário pode não ter toda a sua face capturada, seja por inclinar demais a cabeça para digitar, pelo mau posicionamento da câmera, por apoiar a face sobre a mão, dentre outros fatores. O cansaço decorrente de certo tempo em frente ao computador é uma das principais razões para que isto aconteça.

- **Uso de acessórios**: o uso de certos acessórios, como óculos, pode causar oclusões e afetar o desempenho do sistema. No caso dos óculos, quando combinados com iluminação inadequada, reflexos da luz do ambiente podem ser refletidos nas lentes dos óculos, prejudicando a detecção e o reconhecimento da face.

- **Detecção de vida:** o uso de imagens e fotografias usadas para fraudar o sistema constitui um outro obstáculo para seu desempenho. Neste protótipo não é feito tratamento para a detecção de vida (*liveness detection*). Técnicas como as baseadas em detecção do piscar de olhos podem ser efetivamente aplicadas neste caso, já que se dispõe de trechos de vídeo para a verificação da identidade do indivíduo.

7.2 Conclusões

Este trabalho teve por objetivo mostrar, por meio da arquitetura proposta, a viabilidade do uso das técnicas biométricas de verificação de identidade, em termos de taxa de reconhecimento e de tempo de processamento, para o problema da autenticação remota de indivíduos pela Internet, com aplicação em sistemas de *e-Learning*. Para isto, foi abordada uma arquitetura distribuída que leva em conta as características de sistemas baseados na Internet. Foram usados métodos estabelecidos na literatura, de modo a se medir uma aplicabilidade básica para o assunto. Métodos com melhor desempenho devem trazer melhores resultados que os explicitados neste trabalho.

O uso do vídeo como fonte de dados biométricos mostrou-se, devido à quantidade de informação disponível e da dinâmica de movimentos capturada nas amostras, um caminho promissor para futuras pesquisas. Um modelo simples de representação de vídeo foi apresentado, levando em conta o comportamento de um usuário ao interagir com o sistema de *e-Learning*. Mesmo com algoritmos simples, tratando cada *frame* como uma imagem estática independente, o método trouxe um bom desempenho para o sistema, com taxa de reconhecimento de pouco mais 97% em uma base de dados de 43 pessoas, em condições ideais.

Trabalhos anteriores sobre Biometria aplicada em *e-Learning* exploraram várias técnicas entre elas: face, impressão digital, dinâmica de digitação e abordagens multimodais. Este trabalho estendeu os demais ao realizar medidas de desempenho dos componentes de um sistema de reconhecimento de faces, aplicando o vídeo como fonte dos dados biométricos.

A análise dos resultados experimentais mostrou que o algoritmo Viola-Jones apresentou bom desempenho, com alta taxa de acertos, baixo tempo de processamento e com taxas aceitáveis de falsos negativos e falsos positivos, levando em conta o tamanho das subjanelas respectivas à distância das faces para a câmera. Por este ser um passo que influencia o desempenho do sistema como um todo, algoritmos de rastreamento e de filtragem por cor da pele podem ser investigados para diminuir ainda mais essas taxas.

Os resultados também apontam que a decisão por maioria dos votos não é muito eficiente, pois as taxas de reconhecimento das amostras de vídeo unificadas por esta abordagem trouxeram um desempenho menor que a análise independente dos *frames*. Um caminho apontado na literatura é de atribuir pesos referentes à qualidade da face detectada em determinado quadro, de modo que faces de melhor qualidade tenham pesos maiores na decisão final.

Outro resultado importante diz respeito à quantidade de informação trafegada entre cliente e servidor. O método proposto mostrou-se viável, pois a quantidade de *bytes* usada na comunicação pode ser trafegada rapidamente em conexões de banda larga comuns atualmente.

A arquitetura proposta neste trabalho mostrou-se genérica a qualquer tipo de sistema Web, pois não requer nenhuma configuração adicional no sistema no qual ele será aplicado. A motivação inicial foi pela aplicação em sistemas de *e-Learning*, mas ela pode ser facilmente estendida a outros tipos de sistemas Web.

Apesar das limitações apontadas na seção 7.1.2, este contexto de aplicações Web também apresenta algumas vantagens como: faces com pouca variação de pose, expressão e próximas à câmera. Aliadas à quantidade de informação presente no vídeo, torna-se atrativa sua aplicação em ambientes deste tipo. Assim, de maneira geral, a aplicação da Biometria na autenticação remota de indivíduos mostrou-se viável para este cenário, deixando o processo de autenticação mais confiável que a abordagem atual baseada em senhas.

No entanto, o método abordado neste trabalho não apresenta uma solução definitiva para o problema. A questão da confirmação da identidade de pessoas ocorre mesmo em cursos e exames presenciais. Além disso, diferentes maneiras de burlar o sistema podem ser tentadas. Ao adicionar mais esta camada de segurança espera-se tornar menos freqüente a ocorrência de fraudes.

Capítulo 8 - Trabalhos Futuros

Este trabalho apresenta uma proposta inicial para o tratamento de uma questão importante e a cada dia mais presente em nosso cotidiano. Porém, existem várias dificuldades a serem tratadas. Uma das preocupações do projeto do protótipo desenvolvido foi a de ser facilmente extensível, ao permitir que as técnicas sejam testadas modularmente, sem modificações estruturais no sistema. Sugestões de trabalhos futuros abrangem diversas áreas de conhecimento, dentre as quais podemos destacar:

- Investigar algoritmos que explorem melhor a informação temporal a partir do vídeo, mais robustos a variações de iluminação, oclusão e que gerem um *template* mais compacto que o apresentado;
- Estudar um melhor esquema de fusão das decisões individuais dos quadros, utilizando ponderações em relação à qualidade das faces segmentadas dos quadros do vídeo;
- Normalizar a iluminação dos quadros obtidos, de modo que a verificação do indivíduo possa ser feita independentemente do ambiente no qual ele se encontra. Este item mostrou ser o que mais influenciou negativamente o desempenho do sistema, o que aponta para uma melhor investigação;
- Prover suporte a detecção de vida, de modo a evitar que usuários impostores possam se valer de fotografias, imagens digitais de um outro usuário genuíno para tentar fraudar o sistema;
- Explorar outras arquiteturas para o sistema proposto, comparando e verificando maneiras mais seguras e portáveis para a transferência de dados,

no sentido de modificar os algoritmos utilizados sem grandes modificações estruturais, como por exemplo, transmitir o streaming de vídeo para o servidor, centralizando nele as tarefas atualmente no cliente;

- Investigar a interação dos usuários com o sistema, através de estudos de caso, avaliando qual a reação dos usuários ao usar tal sistema, o grau de incômodo causado pelo monitoramento.

Trabalhos publicados

Os seguintes trabalhos correlatos à dissertação de mestrado foram publicados:

- PENTEADO, B. E.; MARANA, A. N. **A Video-Based Biometric Authentication for e-Learning Web Applications**. Proceedings of the 11th International Conference on Enterprise Information Systems, Milan, Italy. Lecture Notes on Business Information Processing. Springer/Verlag, vol. 24. p.770-779. 2009. ***Prêmio de Best Student Paper***.

- PENTEADO, B. E.; MARANA, A. N. **Autenticação Biométrica On-Line de Usuários em Aplicações Web de Ensino a Distância**. In: WebMedia - XIV Simpósio Brasileiro de Sistemas Multimídia e Web - Webmedia, Vila Velha - ES. Porto Alegre: Sociedade Brasileira de Computação, vol. 2. p.53-56. 2008.

- CHIACHIA, G.; PENTEADO, B. E.; MARANA, A. N. **Fusão de Métodos de Reconhecimento Facial através da Otimização por Enxame de Partículas**. In: IV Workshop de Visão Computacional (WVC), Canal6 Editora, Bauru, 2008. ISBN: 978-85-99728-33-8.

Referências Bibliográficas

ABED – Associação Brasileira de Educação a Distância. **Anuário Brasileiro Estatístico de Educação Aberta e a Distância**, Instituto Monitor, São Paulo, p.58-60. 2007.

AGGARWAL, G. **A System Identification Approach For Video-Based Face Recognition**. *Proceedings of the International Conference on Pattern Recognition*. p. 175-178. 2004.

AGULLA *et al*. **Is My Student at the Other Side? Applying Biometric Web Authentication to E-Learning Environments**. 8th IEEE International Conference on Advanced Learning Technologies (ICALT), p. 551-553, 2008.

ALMEIDA *et al.* 2006. **User Authentication in E-Learning Environments Using Keystroke Dynamic Analysis**. In: 22ª International Council for Distance Education, 2006, Rio de Janeiro. Proceedings of the 22ª International Council for Distance Education, 2006.

ASHA, S., CHELLAPPAN, C. **Authentication of E-Learners Using Multimodal Biometric Technology**. International Symposium on Biometrics and Security Technologies (ISBAST), p. 1-6, 2008.

ASHBOURN, J.; **The Biometric Whitepaper**. 2000. Disponível em: <http://www.jsoft.freeuk.com/whitepaper.htm>. Acessado em: 11 jan. 2009.

AUERNHEIMER, B. **Biometric Authentication for Web-Based Course Examinations**. Proceedings of the 38th Hawaii International Conference on Systems Science. p. 294b. 2005.

BELHUMEUR, P. N.; HESPANHA, J. P.; KRIEGMAN, D. J. **Eigenfaces vs. Fisherfaces: recognition using class specific linear projection**. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, n. 7, p. 711-720. 1997.

BARON, J.; CROOKS, S. M. **Academic Integrity In Web Based Distance Education**. TechTrends, Springer Boston, vol. 49, n. 2. p. 40-45. 2005.

BLANZ, V.; VETTER, T. **Face Recognition From One Example View**. IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 25. p. 1063-1074. 2003.

BOLLE *et al*. **Guide To Biometrics**. Springer Professional Computing, 1st edition. 2004.

BRASIL. **Decreto n. 5622**, 19 de dezembro de 2005. Regulamenta o art. 80 da Lei no 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional. 2005.

BÜLTHOFF, H. H.; EDELMAN, S. Y.; TARR, M. J. **How Are 3-Dimensional Objects Represented In The Brain**. Cerebral Cortex. vol. 5, n. 3. p. 247-260. 1994.

CANTONI, V.; CELLARIO, M.; PORTA, M. **Perspectives And Challenges In E-Learning: Towards Natural Interaction Paradigms**. Journal of Visual Languages and Computing, n. 15, p. 333-345. 2003.

CHELLAPPA, R.; WILSON, C.L.; SIROHEY, S. **Human And Machine Recognition: A Survey**, Proceedings of the IEEE, vol.83, n. 5, p.705-741. 1995.

CLARKE R. **Human Identification In Information Systems: Management Challenges And Public Policy Issues**. In: Information Technology & People. vol. 7, n. 4, p.6-37, 1994.

COSTA, L.; OBELHEIRO, R. R.; FRAGA, J. S. **Introdução À Biometria**. In: Jorge Nakahara Jr; Lau Cheuk Lung. (Org.). Livro texto dos Minicursos do VI Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg2006). SBC: Porto Alegre, v. 1, p. 103-151. 2006.

EIBL, C. J., SOLMS, B. S. H., SCHUBERT, S. **A Framework For Evaluating The Information Security Of E-Learning Systems**. Proceedings of the 2nd International Conference on Informatics in Secondary Schools Evolution and Perspectives (ISSEP), Vilnius, Lituânia, 2006.

ESTADOS UNIDOS. **Public Law 110-315**, 14 ago. 2008. U.S. Department of Education. HR 4137, Part H – Program Integrity, p. 248. 2008.

GUNASEKARAN, A.; MCNEIL, R. D.; SHAUL, D. **E-Learning: Research and Applications. Industrial and Commercial Training**. Emerald Group Publishing Limited. vol. 34, p. 44-53. 2002.

HAYKEN S. **Redes Neurais Artificiais: Princípios e Prática**. 2ª edição. Editora Bookman, 2001.

HERNÁNDEZ *et al*. **Biometrics In Online Assessments: A Study Case in High School Students**. 18th International Conference on Electronics, Communications and Computers (CONIELECOMP), p. 111-116, 2008.

HOLANDA, A. B. **Dicionário Aurélio Eletrônico**, Versão 5.0, 2009.

HUANG *et al*. **How to Compete in a Global Education Market Effectively: A Conceptual Framework for Designing a Next Generation e-Education System**. Journal of Global Information Management. vol. 12, p. 84-107. 2004.

HUGL, U. **Tech-Developments And Possible Influences On Learning Processes And Functioning In The Future**. Journal of American Academy of Business, vol.6, p.250-256. 2005.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, **Censo da Educação Superior**, Brasília, 2006.

JAIN, A.K.; BOLLE, R.; PANKANTI, S. **Biometrics: Personal Identification in Networked Society**. Kluwer Academic Publishers, 1999.

JAIN, A. K.; ROSS, A.; PRABHAKAR, S. **An Introduction to Biometric Recognition**, IEEE Transactions on Circuits and Systems for Video Technology Special Issue on Image and Video-Based Biometrics. v. 14, n. 1, p. 4-20, 2004.

JAIN, A. K. **Face Recognition**. Disponível em: <http://biometrics.cse.msu.edu>. Acessado em: 10 jun. 2009. 2004.

KENNEDY *et. al*. **Academic Dishonesty And Distance Learning: Student And Faculty Views**. The College Student Journal, vol. 34 n. 2. p. 309. 2000.

KIRBY, M.; SIROVICH, L. **Application Of The Karhunen-Loéve Procedure For The Characterization Of Human Faces.** IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12 n. 1:103-108, 1990.

KNIGHT, B.; JOHNSTON, A**. The Role Of Movement In Face Recognition**. Visual Cognition 4:265–274.

KUMAR *et al*. **Using Continuous Biometric Verification to Protect Interactive Login Sessions**. Proceedings of the 21st Annual Computer Security Applications Conference (ACSAC). 2005.

LEE *et al.* **Video-Based Face Recognition Using Probabilistic Appearance Manifolds**. Proceedings of the IEEE Computer Vision and Pattern Recognition. p. 313-320. 2003.

LEVY, Y.; RAMIM, M. M. **A Theoretical Approach For Biometrics Authentication of E-Exams**. The 2007 Chais Conference on Instructional Technologies Research, The Open University of Israel, Raanana, Israel. 2007.

LIU, X.; CHEN, T. **Video-Based Face Recognition Using Adaptive Hidden Markov Models**. Proceedings of the IEEE Computer Vision and Pattern Recognition. p. 340-345. 2003.

MARAIS, E.; ARGLES, D.; SOLMS, B. V. **Security Issues Specific To E-Assessments**, 8th Annual Conference on WWW Applications, Bloemfontein, 2006.

MATHYAS, S. M.; STAPLETON, J. **A Biometric Standard For Information Management And Security**. Computers & Security, vol.19, n. 5, pp. 428-441, 2000.

MILLER, B. **Vital Signs Of Identity**. *IEEE Spectrum*, vol. 31, n.2, p.22–30. 1994.

MOON, H. J.; PHILLIPS, P. J. **Computational and performance aspects of PCA-based face recognition algorithms**. Perception Journal, vol. 30, n. 3, p.303-321. 2001.

MURAS *et al*. **Biometrics for Web Authentication: an Open Source Java-Based Approach**. 1st Spanish Workshop on Biometrics (SWB), Girona, Espanha, 2007.

NEWHAM, E. **The Biometric Report**. http://www.sjb.com/: SJB Services, New York, 1995.

OH, B. J. **Face Recognition By Using Neural Network Classifiers Based On Pca And Lda,** IEEE International Conference on Systems, Man and Cybernetics. p. 1699-1703. 2005.

OPENCV - **Open Source Computer Vision Library**, Disponível em: <http://opencvlibrary.sourceforge.net/>. Acessado em: 15 fev. 2008.

O'TOOLE, A. J.; ROARK, D. A.; ABDI, H. **Recognizing Moving Faces: A Psychological And Neural Synthesis**. Trends in Cognitive Science, 6:261-266, 2002.

PARK, U.; JAIN, A. K.; ROSS, A. **Face Recognition in Video: Adaptive Fusion of Multiple Matchers**. IEEE Conference on Computer Vision and Pattern Recognition, EUA. p.1-8. 2007.

PENTEADO, B. E.; MARANA, A. N. **Autenticação Biométrica On-Line de Usuários em Aplicações Web de Ensino a Distância**. In: WebMedia - XIV Simpósio Brasileiro de Sistemas Multimídia e Web - Webmedia, Vila Velha - ES. Porto Alegre: Sociedade Brasileira de Computação, vol. 2. p.53-56. 2008.

PENTEADO, B. E.; MARANA, A. N. **A Video-Based Biometric Authentication for e-Learning Web Applications**. Proceedings of the 11th International Conference on Enterprise Information Systems, Milan, Italy. Lecture Notes on Business Information Processing. Springer/Verlag, vol. 24. p.770-779. 2009.

PENTLAND, A.; MOGHADDAM, B.; STARNER, T. **View-Based and Modular Eigenspace for Face Recognition**. In Proceedings of Computer Vision and Pattern Recognition. p.84-91. 1994.

PRABHAKAR, S.; PANKANTI, S.; JAIN, A. K., **Biometric Recognition: Security and Privacy Concerns**. IEEE Security & Privacy. p. 33-42. 2003.

RABUZIN, K.; BACA, M.; SJAKO M. **E-Learning: Biometrics as a Security Factor**. Proceedings of the International Multi-Conference on Computing in the Global Information Technology. p. 64-74. 2006.

REDDY, L. N. **An Analysis of E-Journals in Open and Distance Education from Mega Open Universities**. EURODL – European Journal of Open, Distance and e-Learning 2005. Disponível em <http://www.eurodl.org/materials/contrib/2005/Reddy.htm>. Acessado em: 25 maio 2009.

ROLIM, A. L.; BEZERRA, E. P. **Um Sistema De Identificação Automática De Faces Para Um Ambiente Virtual De Ensino E Aprendizagem**. In: XIV Simpósio Brasileiro de Sistemas Multimídia e Web, Vila Velha. 2008.

ROSENBERG, M. J. **E-Learning: Strategies for Delivering Knowledge in the Digital Age**. McGraw Hill, New York, NY, USA. 2002.

ROVE, N. C. **Cheating in Online Student Assessment: Beyond Plagiarism**. Online Journal of Distance Learning Administration, vol. 7, no. 2, 2004.

RUDERMAN, D. L. **The Statistics Of Natural Images.** Network: Computational Neural Systems, vol. 5, no. 4, p.517-548, 1994.

SIM *et al*. **Continuous Verification Using Multimodal Biometrics**. IEEE Transactions On Pattern Analysis and Machine Intelligence, vol. 29. no. 4, p. 687-700. 2007.

STALLKAMP, J.; EKENEL, H. K.; STIEFELHAGEN, R. **Video-based Face Recognition on Real-World Data**. IEEE International Conference on Computer Vision. p 1-8. 2007.

THOMAS, D.; BOWYER, K.; FLYNN, P. **Multi-Frame Approaches To Improve Face Recognition**, IEEE Workshop on Motion and Video Computing, p.19. 2007.

TIEU, K; VIOLA, P. **Boosting Image Retrieval**. International Journal of Computer Vision, vol.1, p. 228-235. 2000.

TURK, M.; PENTLAND, A. **Eigenfaces For Recognition**, Journal of Cognitive Neuroscience, vol. 3, no. 1, p.71-86. 1991.

VIOLA, P. A.; JONES, M. J. **Robust Real-Time Object Detection**, International Journal of Computer Vision, vol. 57, no.2, p. 137-154. 2001.

WANG, H.; YANG, S.; LIAO, W. **An Improved PCA Face Recognition Algorithm Based on the Discrete Wavelet Transform and the Support Vector Machines**, International Conference on Computational Intelligence and Security. p. 308-311. 2007.

WEIPPL, E. R., **Security In E-Learning**. eLearn Magazine. Disponível em: http://elearnmag.org/subpage.cfm?section=tutorials&article=19-1. Acessado em: 19 nov. 2008.

YAMAGUCHI, O.; FUKUI, K; MAEDA, K. **Face Recognition Using Temporal Image Sequence**. Proceedings of International Conference on Automatic Face and Gesture Recognition. p. 318-323. 1998.

ZHAO *et al*. **Face Recognition: A Literature Survey**. ACM Computing Surveys, vol. 35, no. 4, 2003.

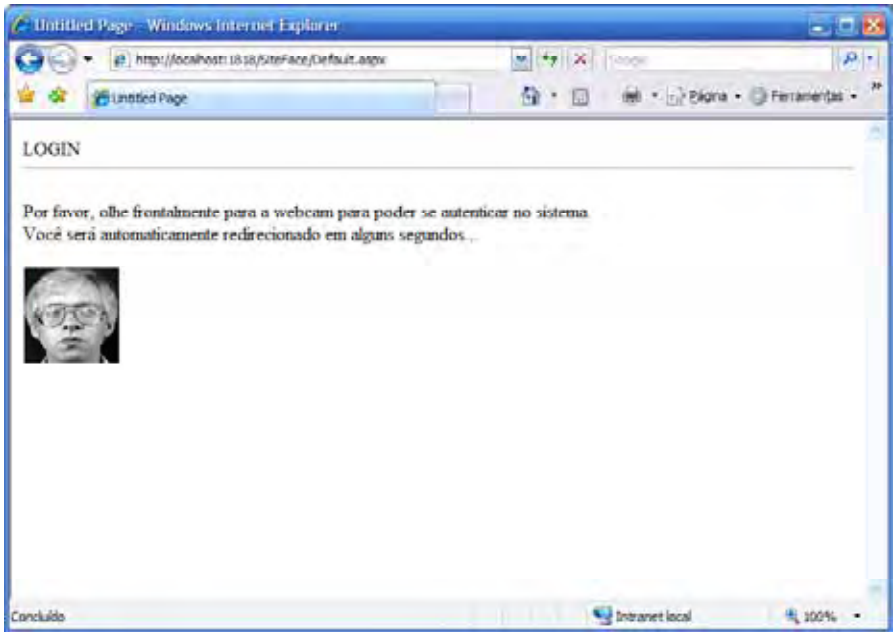
ZHOU, S. K.; CHELLAPPA, R. **Face Tracking And Recognition From Video**. In: Jain, A. K., Li, S. Handbook of Face Recognition. New York Springer Science+Business Media. p. 1-25. 2005.

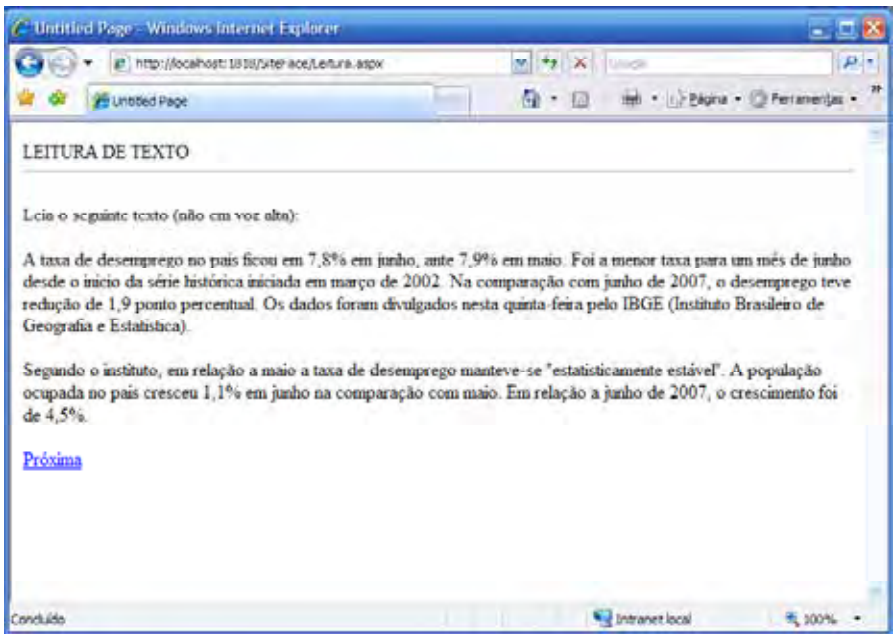
ZHOU, S. K.; CHELLAPPA, R.; ZHAO, W. **Unconstrained Face Recognition**. New York: Springer Science+Business Media. p. 31-39. 2006.

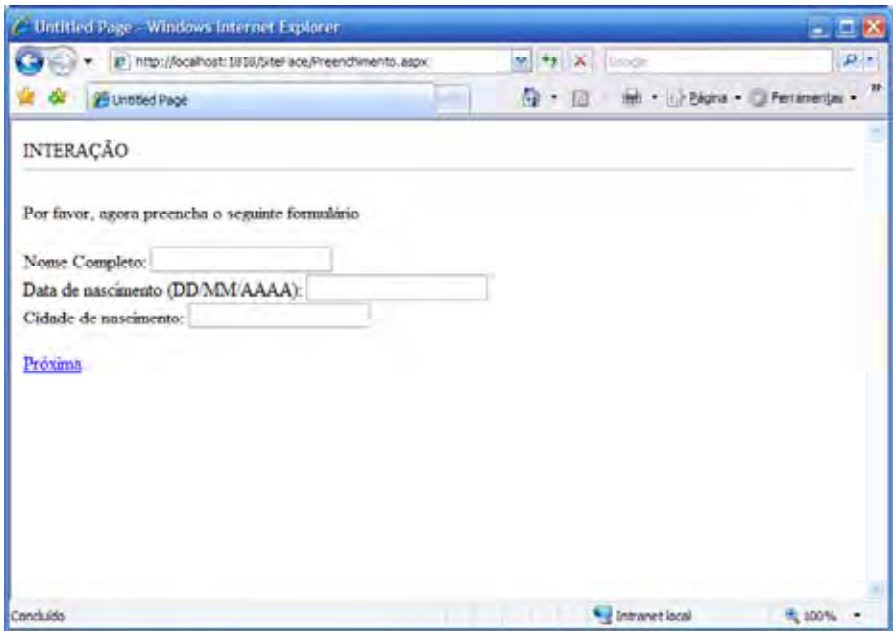
ZHOU, S. K.; KRUEGER, V.; CHELLAPPA, R. **Face Recognition From Video: A CONDENSATION Approach**. Proceedings of the International Conference on Automatic Face and Gesture Recognition. p. 221-228. 2002.

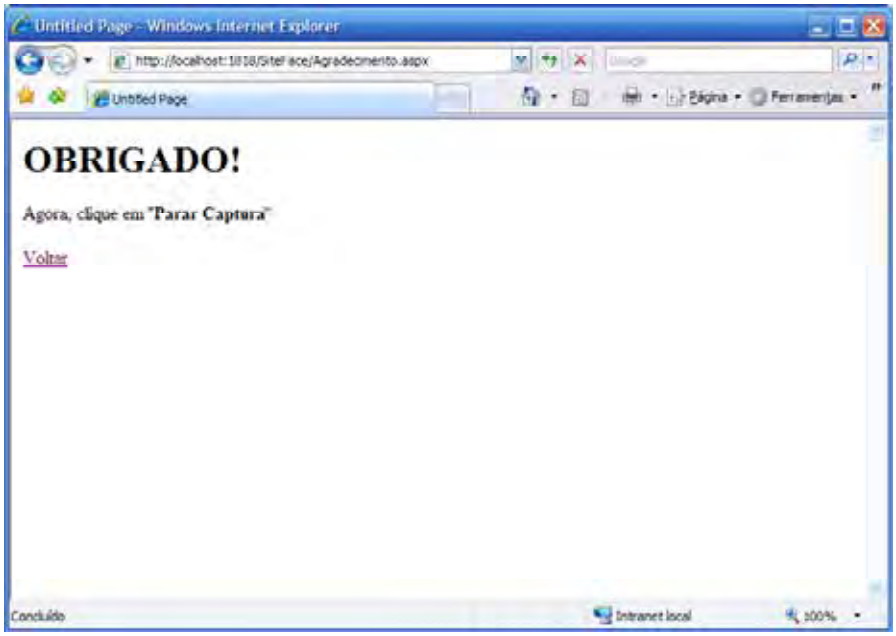
Apêndice A

Telas utilizadas na coleta das seqüências de vídeo para a construção da base de dados.






Untitled Page - Windows Internet Explorer
http://localhost:1018/siteFace/Preenchimento.aspx
Untitled Page
Página
Ferramentas
INTERAÇÃO
Por favor, agora preencha o seguinte formulário
Nome Completo:
Data de nascimento (DD/MM/AAAA):
Cidade de nascimento:
Próxima
Concluído
Intranet local
100%


Untitled Page - Windows Internet Explorer
http://localhost:1018/siteFace/Agradecimento.aspx
Untitled Page
Página
Ferramentas
OBRIGADO!
Agora, clique em "Parar Captura"
Voltar
Concluído
Intranet local
100%

Autorizo a reprodução xerográfica para fins de pesquisa.

Bauru, 27 / 07 / 2009

_____

Assinatura