UNIVERSIDADE FEDERAL DA BAHIA

DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

PROGRAMA DE GRADUAÇÃO

MATHEUS MAGALHÃES BATISTA DOS SANTOS

AUTENTICAÇÃO FACIAL CONTÍNUA USANDO IMAGÉNS DE INFRAVERMELHO

Salvador - BA, Brasil

31 de Maio de 2016

Matheus Magalhães Batista dos Santos

Autenticação Facial Contínua usando Imagens de Infravermelho

Monografia apresentada para obtenção do Grau de Bacharel em Ciência da Computação pela Universidade Federal da Bahia.

Universidade Federal da Bahia

Departamento de Ciência da Computação

Programa de Graduação

Orientador: Maurício Pamplona Segundo

Salvador - BA, Brasil

31 de Maio de 2016

Matheus Magalhães Batista dos Santos
Autenticação Facial Contínua usando Imagens de Infravermelho / Matheus Magalhães Batista dos Santos. – Salvador - BA, Brasil, 31 de Maio de 2016-
40 p. : il. (algumas color.) ; 30 cm.

Orientador: Maurício Pamplona Segundo

Monografia – Universidade Federal da Bahia
Departamento de Ciência da Computação
Programa de Graduação, 31 de Maio de 2016.

1. reconhecimento facial. 2. infravermelho. 2. padrões binários locais. 3. autenticação contínua. I. Maurício Pamplona Segundo. II. Universidade Federal da Bahia. III. Departamento de Ciência da Computação. IV. Autenticação Facial Contínua usando Imagens de Infravermelho

Matheus Magalhães Batista dos Santos

Autenticação Facial Contínua usando Imagens de Infravermelho

Monografia apresentada para obtenção do Grau de Bacharel em Ciência da Computação pela Universidade Federal da Bahia.

Trabalho aprovado. Salvador - BA, Brasil, 24 de novembro de 2012:

Maurício Pamplona Segundo
Orientador

Professor
Convidado 1

Professor
Convidado 2

Salvador - BA, Brasil
31 de Maio de 2016

Agradecimentos

Gostaria de agradecer aos meus pais José Batista e Maria Cristina, minha irmã Gabriela Magalhães e a toda a minha família pelo apoio, fé, confiança e incentivo durante toda a minha vida, incluindo na graduação.

A todos os meus amigos que sempre me apoiaram e estiveram comigo não só nos momentos bons, mas principalmente nos momentos mais difíceis. Existem pessoas que estão presentes em minha vida há muitos anos e outras não mais, mas cada uma tem/teve sua participação na construção da minha jornada e do meu ser, e por isso sou grato à elas. Em especial para Ricardo, Philippe, Leonardo, Brunos, Júnior, Mariana, Ágata, Emanuele e Águeda. Também gostaria de agradecer a todos os amigos que fiz na graduação pelas dicas, diversão, conselhos, experiências e ajuda na construção do conhecimento que certamente chegou a este trabalho. São muios, mas em especial para Anderson, Erik, Ester, Fernando, Fred, Ive, Nanci, Raniere, Rodrigo e Ygor.

A todos os professores que repassaram seu conhecimento, desde a infância até este momento.

Ao meu orientador Maurício Pamplona pela oportunidade de trabalhar nesse projeto, orientação e por ter me apresentado a área acadêmica que me motivou muito a concluir o curso.

A todos os membros do projeto de Autenticação Contínua pela ajuda na construção do projeto, solução de problemas, amizade e muitas risadas. Aprender foi muito divertido graças a todos eles.

"And now his watch is ended".
(The Night's Watch)

Resumo

A utilização da biometria para autenticação e identificação de indivíduos tem crescido muito nos últimos anos. Diversos sistemas de autenticação tem sido desenvolvidos utilizando reconhecimento facial devido a sua facilidade de aquisição, baixa rejeição e por não ser intrusivo. Porém, boa parte desses sistemas tem seu desempenho afetado por variações de iluminação ambiente e não garantem que apenas o usuário permitido continue a utilizar o sistema após sua autenticação. Para tratar esses problemas, este trabalho apresenta um sistema de autenticação facial contínua baseado em imagens de infravermelho. O objetivo foi o desenvolvimento de um sistema totalmente automático utilizando apenas imagens em infravermelho como entrada, visando alcançar invariância à iluminação ambiente durante toda a utilização. Até onde sabemos, este é o primeiro sistema que utiliza imagens em infravermelho para a autenticação contínua. Avaliamos nosso sistema utilizando quatro vídeos de 1000 quadros cada, simulando ataques de intrusos, e obtivemos como resultado uma taxa de reconhecimento de 93%.

Palavras-chave: reconhecimento facial; infravermelho; padrões binários locais; autenticação contínua.

Abstract

The use of biometrics for authentication and identification of individuals has grown tremendously in recent years. Several biometric authentication systems were developed using facial recognition due to its ease of acquisition, low rejection and non-intrusiveness. However, most of them have their performance affected by changes in environmental illumination and cannot ensure that only the allowed user accesses the system after his or her authentication. To overcome these problems, this work presents a continuous infrared-based face authentication system. The objective was to develop a fully automatic system using only infrared images as input, seeking to achieve illumination invariance during the entire access. To the best of our knowledge, this is the first system that uses infrared images for continuous authentication. We evaluated our system on four 1000 frames long videos, simulating attacks from intruders, and achieved a 93% recognition rate.

Keywords: face recognition; infrared; local binary pattern; continuous authentication.

Lista de ilustrações

Lista de tabelas

Lista de abreviaturas e siglas

NIR	Near Infrared
LBP	Local Binary Pattern
LBPH	Local Binary Pattern Histogram
LDA	Linear Discriminant Analysis
PCA	Principal Component Analysis
ROC	Receiver Operating Characteristic
TPR	True Positive Rate
FPR	False Positive Rate

Sumário

1 Introdução

Com o passar dos anos, métodos tradicionais de autenticação, como senhas ou cartões, se tornaram arriscados em ambientes que demandam um controle de segurança mais rígido. A biometria foi uma solução adotada em muitos sistemas para suprir essa demanda, e muitas pesquisas foram desenvolvidas nessa área recentemente (DUGELAY et al., 2002; KONG et al., 2004). Diversas características biométricas podem ser utilizadas, dentre elas: face (JANAKIRAMAN et al., 2005; NIINUMA; PARK; JAIN, 2010), impressão digital (SIM et al., 2007a), voz (DAMOUSIS; TZOVARAS; BEKIARIS., 2008), íris (MOCK et al., 2012), eletroencefalograma (NAKANISHI; BABA; MIYAMOTO, 2009) e eletrocardiograma (AGRAFIOTI; HATZINAKOS, 2009).

Os sistemas tradicionais de reconhecimento realizam a verificação da identidade do usuário apenas uma vez, o que não garante que um ataque de um usuário não autorizado, posterior a uma autenticação válida, possa ser feito. Para solucionar esse problema, a autenticação contínua realiza a verificação da identidade do usuário constantemente, garantindo assim que o usuário autorizado seja o mesmo durante toda a utilização do sistema, e, por isso, se tornou um alvo de pesquisas nos últimos anos. A autenticação contínua é bastante importante em ambientes de alto risco, onde o custo de um uso do sistema por alguém não autorizado é alto, como no controle de aviões, computadores de bancos, departamentos de defesa e outras aplicações que lidem com grandes quantidades de dinheiro ou que afetem a segurança de vidas humanas. Nesses casos, é desejável que o sistema se torne inoperante quando um usuário autorizado não possa ser autenticado (JANAKIRAMAN et al., 2005).

Diferentes características biométricas podem ser empregadas no contexto da autenticação contínua. A forma de digitar foi a característica pioneira utilizada para a realização de exames a distância, por exemplo. Estudantes poderiam ceder o acesso a outras pessoas para realização do exame, uma vez que os tradicionais métodos de usuário/senha não conseguem lidar com isso. Através de características como velocidade de digitação, letras por determinada unidade de tempo, e tempo entre o apertar e o soltar de teclas, esta característica é capaz de validar constantemente o usuário do sistema (FLIOR; KOWALSKI, 2010; GUNETTI; PICARDI, 2005a; GUNETTI; PICARDI, 2005b; LEGGETT et al., 1991). Apesar dos benefícios, é necessário muito tempo para se detectar um impostor, e comandos nocivos ao sistema podem ser digitados rapidamente antes que o sistema invalide o acesso.

Outras características também apresentam problemas para este fim. Sistemas baseados em impressão digital são inconvenientes, pois exigem que o usuário coopere

constantemente para a captura de imagens das digitais (SIM et al., 2007b). Eletrocardiogramas exigem que o usuário vista sensores, estão sujeitos a variações cardíacas causadas por fatores patológicos ou emocionais (AGRAFIOTI; HATZINAKOS, 2009), sendo assim intrusivos para o usuário. Já faces tem grande aceitação, não requerem cooperação do usuário na captura, e não são intrusivas (CHANG; BOWYER; FLYNN, 2003).

Nesse trabalho, é proposto um sistema de reconhecimento facial baseado em imagens de infravermelho invariante a iluminação ambiente, rápido e seguro para fins de autenticação contínua. O uso de espectro próximo ao infravermelho (NIR, *Near-Infrared*) evita variações de iluminação, permitindo um reconhecimento mais preciso. Câmeras NIR são acessíveis atualmente, custando menos de 70 reais. Em especial, o Kinect One utilizado neste trabalho permite capturar imagens em infravermelho com baixo nível de ruído, além da imagem em luz visível e de profundidade. Esta informação é importante, pois este trabalho está inserido em um projeto maior, cujo objetivo é combinar estas três imagens na autenticação facial contínua.

O restante deste trabalho está organizado da seguinte maneira:

- no Capítulo 2 apresentamos uma revisão dos trabalhos relacionados da literatura;
- no Capítulo 3 é apresentado o sistema de autenticação facial contínua com imagens em infravermelho;
- o Capítulo 4 mostra os resultados experimentais;
- o Capítulo 5 mostra as considerações finais.

2 Referencial Teórico

Reconhecimento facial é uma área de pesquisa que cresceu bastante nos últimos 30 anos. Isso se deve, principalmente, ao fato de que os métodos tradicionais de segurança, como cartões de identificação e senhas, não são seguros ou adequados o suficiente. A análise facial também permite interpretar expressões faciais, emoções humanas, intenções e comportamentos, peças chaves para sistemas de segurança cada vez mais inteligentes. Trata-se de uma maneira não incômoda e possivelmente mais natural de identificação, tanto que é a mais utilizada pelos seres humanos (KONG et al., 2005).

O reconhecimento facial é uma biometria de custo relativamente baixo e resultados satisfatórios, e pode ser utilizada na autenticação contínua sem a colaboração do usuário. Existem três propriedades que podem ser utilizadas no reconhecimento facial: textura, infravermelho e geometria. A textura é uma propriedade obtida em fotos comuns, através da captura da luz visível da cena. Infravermelho é obtida através de luz invisível aos olhos humanos, muito utilizada em câmeras de segurança, pois não dependem de iluminação ambiente. A geometria captura a profundidade dos objetos e da cena em relação ao sensor.

Diversos fatores podem interferir na segurança de cada propriedade, dentre eles a presença de óculos, variações de pose, expressões faciais, iluminação e/ou temperatura facial. O sistema baseado em imagens de textura é o mais amplamente empregado, mas é também o que sofre mais influência da variação de iluminação externa, mesmo em ambientes fechados. Com isso, a variação causada pela iluminação em faces da mesma pessoa pode ser maior do que a variação entre imagens de pessoas diferentes com condições de iluminação semelhantes (ADINI; MOSES; ULLMAN, 1997). Para sistemas baseados em infravermelho no espectro termal, a temperatura da face pode alterar devido a condições emocionais, estações do ano ou prática de atividade física antes da verificação. Por fim, os baseados em geometria necessitam de sensores mais caros, possuem mais ruído e exigem mais recursos computacionais.

Uma imagem facial possui uma quantidade muito grande de informações, portanto um dos objetivos de um sistema de reconhecimento facial é selecionar e extrair as características que mais discriminam as faces, reduzindo assim o custo computacional (LI; LIAO, 2003; WEI; ZHIHUA, 2011; LI et al., 2007; ZHENG, 2012; ZHIHUA; GUODONG, 2013). Existem três categorias principais de métodos que podem ser empregados para este fim: métodos baseados em características, métodos holísticos e métodos híbridos. Os baseados em características dependem de características da face, como olhos, nariz e boca, assim como as relações geométricas entre elas. Métodos holísticos levam em consideração a face como um todo. As técnicas pioneiras desta categoria foram a Análise de

Componentes Principais (PCA, *Principal Component Analysis*), popularmente conhecida como *eigenfaces* (MOGHADDAM; PENTLAND, 1998), e a Análise de Discriminantes Lineares (LDA, *Linear Discriminant Analysis*), popularmente conhecida como *fisherfaces* (BELHUMEUR; HESPANHA; KRIEGMAN, 1996). Ambas são baseadas em autovetores e autovalores e criam representações concisas da aparência global de imagens faciais, permitindo comparações bastante eficientes e precisas (FERNANDES; BALA, 2013; LU; VENETSANOPOULOS, 2003; H; PJ, 2001). Os métodos híbridos combinam métodos baseados em característica e holísticos (ZHAO; GRIGAT, 2005) para obter informações sobre as características específicas da face e sobre a mesma como um todo.

2.1 Luz Visível

A grande maioria dos sistemas de reconhecimento facial na literatura utiliza a luz visível (i.e. imagens 2D). Abordagens que utilizam a forma e a textura para reconhecer imagens faciais 2D foram desenvolvidas, baseadas em histogramas de Padrões Binários Locais (LBP, *Local Binary Pattern* (AHONEN; PIETIKäINEN, 2004)) e Modelos de Aparência Ativa (AAM, *Active Appearance Model* (COOTES; EDWARDS; TAYLOR, 2001)). Porém, sistemas baseados em luz visível tem seu desempenho severamente afetado pela variação de iluminação ambiente. Diversas pesquisas foram realizadas com o intuito de superar este problema (ADINI; ULLMAN, 1997), e uma alternativa possível é sintetizar uma imagem em infravermelho a partir de uma imagem 2D através de algoritmos que mensuram a relação linear entre dois conjuntos de dados. Os resultados das verificação em imagens em luz visível e a sintetização em infravermelho podem ser combinados com desempenho superior ao reconhecimento utilizando apenas luz visível (MAVADATI; KITTLER, 2010).

Na autenticação contínua, o uso de *fisherfaces* combinado com a utilização de classificadores Bayesiano foi empregado para realizar a autenticação contínua em computadores desktop usando imagens 2D (JANAKIRAMAN et al., 2005). O sistema calcula a probabilidade do usuário ainda estar presente e utilizando o computador, e, se essa probabilidade estiver abaixo de um limiar, o sistema operacional pode travar o acesso ou atrasar processos desse usuário. Outra proposta foi utilizar características simples, como a cor de pele da face do usuário combinada com a cor das suas vestes para autenticação contínua, buscando aumentar a usabilidade do sistema (NIINUMA; PARK; JAIN, 2010). Porém, essas abordagens não são robustas a ataques de invasores. Por fim, o uso do reconhecimento facial com luz visível também pode ser combinado com outras modalidades biométricas, como, por exemplo, o reconhecimento de impressões digitais (SIM et al., 2007a). Fusões biométricas para autenticação contínua tem desempenho potencialmente maior do que utilizando apenas uma modalidade, mas deve-se tomar cuidado para não impactar na usabilidade do sistema.

2.2 Infravermelho

Uma solução para o problema da variação de iluminação é a utilização de imagens em infravermelho. Diversos sistemas de reconhecimento facial que utilizam esta abordagem foram propostos e os resultados mostram um desempenho maior em relação aos que utilizam a luz visível (LI LUN ZHANG; HE, 2006; ZOU; KITTLER; MESSER, ; LI et al.,). O potencial para reconhecimento facial invariante a iluminação usando emissões termais também receberam destaque na literatura (BHOWMIK KANKAN SAHA; NASIPURI, 2007). Comparada com outras soluções, o reconhecimento facial com imagens NIR é mais prático em cenários reais por quatro razões. Primeira, com iluminação NIR através de diodos emissores de luz (LED, *Light Emitting Diode*) em intensidade suficiente, o reconhecimento facial se torna robusto a variações de iluminação ambiente. Segundo, comparada com o reconhecimento facial térmico, a imagem NIR é menos afetada pela temperatura ambiente, condições emocionais e de saúde do usuário. Terceiro, comparado com o reconhecimento facial utilizando 3D, o NIR possui um custo menor. Por último, o uso de NIR tem uma grande variedade de aplicações, pois pode funcionar tanto de dia quanto a noite (HUANG; WANG; WANG, 2007). Embora o reconhecimento facial em infravermelho tenha sido explorado nos últimos anos, ele não foi utilizado para a autenticação contínua.

2.3 Profundidade

A utilização de imagens de profundidade tem potencial para realizar o reconhecimento facial mesmo sob variações de pose, iluminação e expressão facial. Porém, a utilização desse tipo de imagem também cria novos problemas, como o alinhamento das malhas e a qualidade do sensor. Sensores de profundidade também podem ser afetados por fontes de luz intensas ou superfícies reflexivas, e a distância entre o usuário e a câmera altera a resolução da superfície da face capturada (ABATE et al., 2007). Extensões de *eigenfaces* e do Modelo Oculto de Markov (HMM, *Hidden Markov Model*), usados previamente no reconhecimento facial com luz visível, foram feitas para imagens em profundidade (ACHERMANN; JIANG; BUNKE, 1997).

Sistemas que utilizam imagens de profundidade também foram usadas na autenticação contínua. Através da informação 3D, o nível de cooperação do usuário foi reduzido, permitindo também realizar a autenticação de maneira robusta a variações de expressões faciais, pose e oclusões. Como variações de pose afetam apenas um lado da face, causando buracos e bastante ruído, técnicas de estimação da posição da cabeça e de divisão da face em regiões de interesse permitiram contornar esse problema através da seleção de regiões de acordo com a pose do usuário (PAMPLONA et al., 2013).

2.4 Híbridos

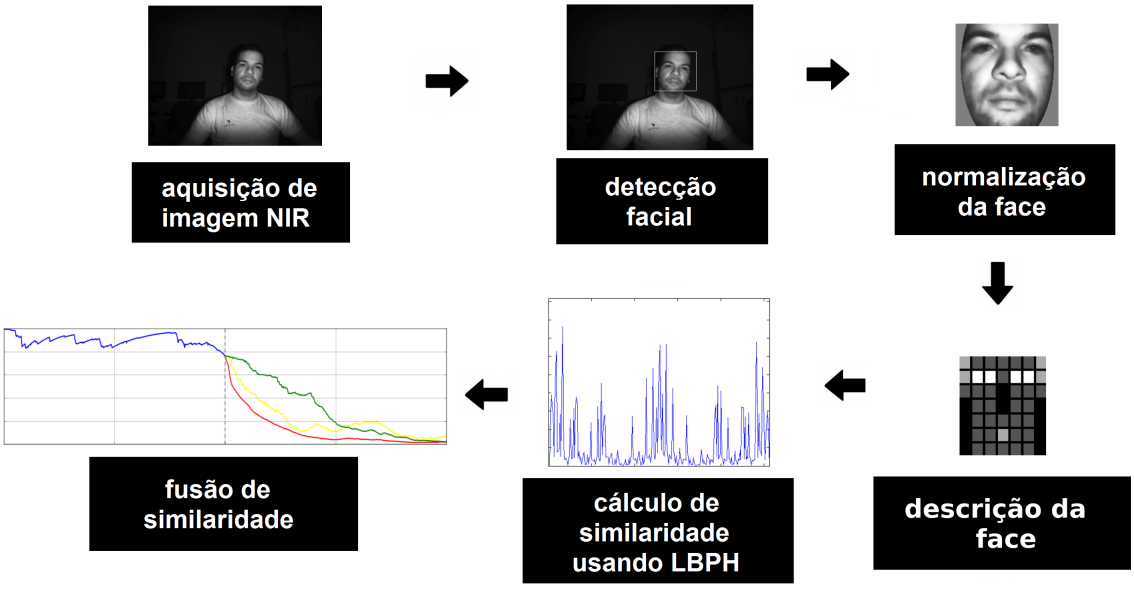
Sistemas que utilizam mais de uma modalidade biométrica também foram desenvolvidos, com foco em serem discretos aos usuários que são identificados, como por exemplo, combinado imagens faciais, voz e eletrocardiograma (DAMOUSIS; TZOVARAS; BEKIARIS., 2008). Porém, no contexto da autenticação contínua, combinar dois métodos através do tempo não é uma tarefa trivial, pois a fusão dos métodos tem que levar em conta que a precisão de cada método é diferente, além de casos onde uma das modalidades poder estar ausente em um determinado instante da utilização. Fusões holísticas integram as observações biométricas ao longo do tempo, calculando a probabilidade do sistema estar seguro. Definindo um limiar de segurança, é possível tomar alguma ação quando a probabilidade do sistema estar seguro for menor que esse limiar estabelecido. Sim et al. (2007a) utilizou essa abordagem para combinar reconhecimento facial com reconhecimento de impressões digitais.

Apesar de existirem sistemas de autenticação contínua híbridos, não existe na literatura um sistema de autenticação contínua baseado em múltiplas modalidades faciais. Devido ao desempenho superior dos sistemas híbridos, a ideia de um sistema de autenticação facial contínua multimodal é promissora, e é um objetivo futuro deste trabalho.

3 Sistema

O sistema proposto neste trabalho utiliza uma câmera NIR e implementa o reconhecimento facial com auxílio da biblioteca OpenCV[1], uma biblioteca de computação visual de licença aberta que possui um conjunto extenso de algoritmos da área de visão computacional. O objetivo é garantir que o usuário permitido é o mesmo durante toda a execução do sistema, podendo ser utilizado em computadores pessoais, terminais, caixas eletrônicos de banco, cabines de piloto ou qualquer equipamento que possa conter uma câmera em infravermelho que permita a captura contínua da face do usuário. O sistema é dividido em seis etapas, como mostrado na Figura 1: (1) aquisição de imagens, onde as imagens são capturadas pelo sensor; (2) detecção facial, em que a face é localizada na imagem capturada previamente; (3) normalização facial, onde a face passa por uma série de ajustes para remover ruídos e seguir um padrão; (4) descrição, onde as características mais discriminantes da face são extraídas; (5) cálculo da similaridade, onde é calculada o grau de semelhança entre a face atual e a face do usuário autenticado; e (6) fusão de similaridade, onde a similaridade mais recente é utilizada para atualizar a probabilidade do sistema estar seguro. Mais detalhes sobre cada etapa são dados nas seções posteriores.

Figura 1 – Diagrama de estágios do sistema proposto.



[1] http://opencv.org/

3.1 Aquisição de imagens

A aquisição de imagens é o processo de conversão de uma cena real em uma imagem digital através de um sensor. Neste trabalho, as imagens são capturadas utilizando o sensor Microsoft Kinect One[2]. O Microsoft Kinect One é um dispositivo com sensor de profundidade infravermelho, câmera colorida e um emissor de infravermelho, sendo que as imagens em infravermelho são capturadas com resolução de 512×424 pixels. Seus emissores são capazes de gerar uma boa iluminação frontal com baixíssimo ruído, funcionando até mesmo em completa escuridão, sendo esse um dos motivos para sua utilização neste trabalho. O infravermelho é uma radiação do espectro eletromagnético com comprimento de onda entre 0,75 μm a 100 μm, invisível aos olhos humanos e não causa nenhum tipo de dano aos olhos ou à pele, ao contrário da radiação ultravioleta.

As imagens foram capturadas pelo Kinect através da biblioteca Libfreenect2[3], um *driver* de licença aberta que dá suporte à transferência e registro de imagens em infravermelho, luz visível e profundidade.

3.2 Detecção facial

O primeiro passo em qualquer sistema de reconhecimento facial é a detecção da face em uma imagem. Esta tarefa, trivial para os seres humanos, representa um grande desafio para a área de visão computacional. A principal dificuldade da detecção se dá porque inicialmente não se sabe em que regiões da imagem podem existir faces e nem as suas dimensões. Além disso, objetos podem se assemelhar a faces quando analisados em baixa resolução e oclusões da face, variações entre pessoas, pose, iluminação e expressões faciais podem interferir no desempenho de detecção. Podemos resumir a detecção facial na habilidade para distinguir duas classes de imagens: faces, e tudo que não é uma face. Como neste trabalho a entrada do sistema é um vídeo, uma sequência de imagens em tempo real, é crítico para o sistema que a detecção seja realizada muito rapidamente.

As principais medidas para avaliar o desempenho de um algoritmo de detecção facial são: a quantidade de objetos que foram incorretamente classificados como faces (FDR, *False Discovery Rate*) e a quantidade de faces que não foram detectadas (FRR, *False Rejection Rate*). Quanto menores esses valores, melhor é o algoritmo.

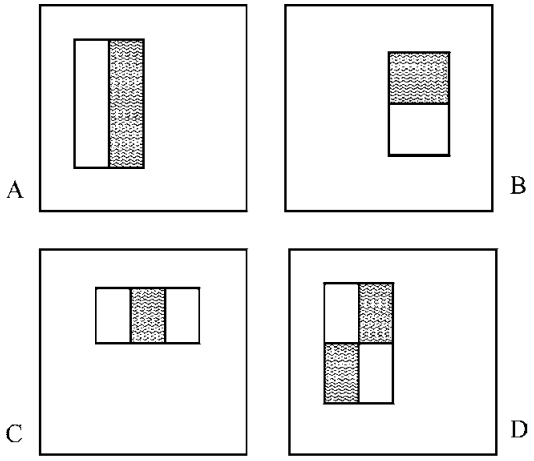
Neste trabalho foi utilizado um classificador em cascata de características de Haar (VIOLA; JONES, 2001), pois este permite detectar faces com grande velocidade e atingir altos índices de detecção, pois descartam rapidamente objetos de fundo e só reconhecem uma face quando a mesma é aceita por vários classificadores. O método de Viola e Jones (2001) e o treinamento do classificador em cascata estão detalhados nas próximas subseções.

[2] https://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx
[3] https://github.com/OpenKinect/libfreenect2

3.2.1 Método de Viola-Jones

Este método classifica imagens baseado em conjuntos de características simples, conhecidas como características de Haar. Essas características são basicamente máscaras retangulares, e o valor de uma característica é a diferença entre a soma interna dos pixels de duas áreas desta máscara, ilustradas em branco e cinza na Figura 2. Estas características representam uma diferença de intensidade luminosa entre áreas da imagem (PAPAGEORGIOU; OREN; POGGIO, 1998).

Figura 2 – Representação visual dos três tipos de características Haar. A soma dos pixels dos retângulos brancos são subtraídos da soma dos pixels dos retângulos cinzas. Características de dois retângulos são mostrados nas janelas (A) e (B). A janela (C) mostra uma característica de três retângulos e a (D) quatro retângulos (VIOLA; JONES, 2001).

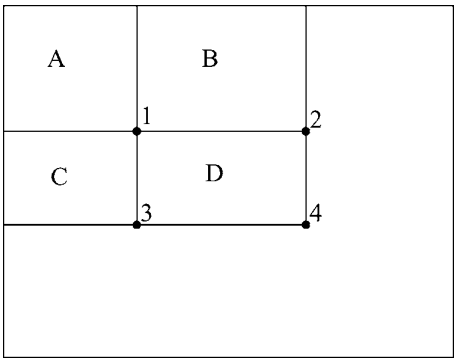


O detector facial proposto por Viola e Jones (2001) é capaz de processar imagens rapidamente enquanto atinge altas taxas de acerto baseado em três conceitos: uma nova representação de imagem, chamada de imagem integral, que permite que os valores das características de Haar sejam computados rapidamente; um classificador simples e eficiente, construído utilizando o algoritmo de aprendizagem Adaboost; e um método para classificação em cascata, que permite que as regiões de fundo da imagem sejam rapidamente descartadas, poupando recursos computacionais.

Os três tipos de características de Haar podem ser computados utilizando uma representação intermediária, chamada de imagem integral. Ela pode ser computada em uma única varredura pela imagem original, e um ponto (x,y) representa a soma de todos os pixels acima de y e à esquerda de x, incluindo x e y. Com isso, qualquer característica de Haar pode ser calculada em qualquer escala ou localização em tempo constante, pois a soma dos pixels de qualquer retângulo pode ser obtida com apenas quatro valores na imagem integral. Por exemplo, dada a região retangular D na Figura 3, a soma dos pixels

dessa área é dada pela soma dos pontos 1 e 4 subtraídos dos pontos 2 e 3 na imagem integral.

Figura 3 – Diferentes áreas retangulares em uma imagem. A soma dos pixels no retângulo A é o valor da imagem integral no ponto 1 (VIOLA; JONES, 2001).



Cada tipo de característica em uma determinada área de uma janela de detecção pode ajudar a reconhecer um determinado padrão. Por exemplo, a caraterística B da Figura 2 pode identificar uma área na imagem que possui uma diferença de intensidade significativa entre a parte superior e a inferior de uma região. Então ela pode indicar a presença de face, visto que frequentemente a região dos olhos (retângulo superior) é mais escura que a região das bochechas (retângulo inferior), como mostrado na imagem (b) da Figura 4. Porém, esse padrão pode ocorrer em regiões de uma imagem que não fazem parte de uma face, criando a necessidade de se combinar várias características para melhorar o poder discriminatório da busca. O problema é que, em uma resolução base da janela do detector de 24×24 pixels, existem 160.000 possíveis características Haar se considerarmos diferentes posições e tamanhos dentro dessa janela.

Para solucionar esse problema, é utilizado uma variação do algoritmo de aprendizagem Adaboost (FREUND; SCHAPIRE, 1997) para selecionar as características mais discriminantes. Para garantir uma classificação rápida, a ideia é construir classificadores fortes através da combinação linear de vários classificadores fracos, que são então organizados em cascata.

A detecção facial consiste em varrer a imagem, do início ao fim, utilizando janelas de tamanho determinado a cada iteração, como ilustrado na Figura 5. Em uma determinada janela de detecção, cada estágio da cascata aplica um classificador forte. Caso a janela passe em um estágio, ela será testada no próximo, mais complexo e mais específico que o anterior. Porém, se ocorre uma rejeição, a janela é descartada, evitando que os estágios posteriores sejam executados desnecessariamente. Só será considerada uma face a janela que passar por todos os estágios. A ideia é que classificadores menores e mais simples sejam utilizados nos primeiros estágios, pois são mais rápidos e permitem descartar rapidamente a maioria das regiões negativas, como cenários e objetos de fundo. A Figura 6 ilustra este

Figura 4 – Exemplos de características selecionadas pelo Adaboost que mais discriminam faces. Ambas se baseiam no fato de que a região dos olhos geralmente é mais escura que a região das bochechas e da base do nariz (VIOLA; JONES, 2001).

(a) face a ser analisada (b) região dos olhos e bochecha (c) região dos olhos e base do nariz

processo. Este processo é eficiente porque a grande maioria das janelas testadas em uma imagem não são faces, e são desconsideradas rapidamente nos primeiros estágios deste classificador por não apresentarem as características padrões de uma face.

Figura 5 – Varredura em uma imagem em busca de faces através das caraterísticas Haar em cascata.

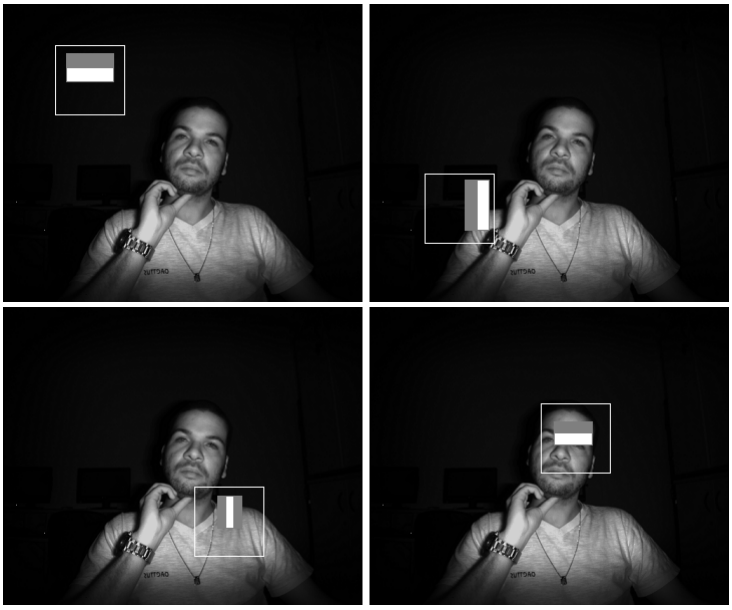
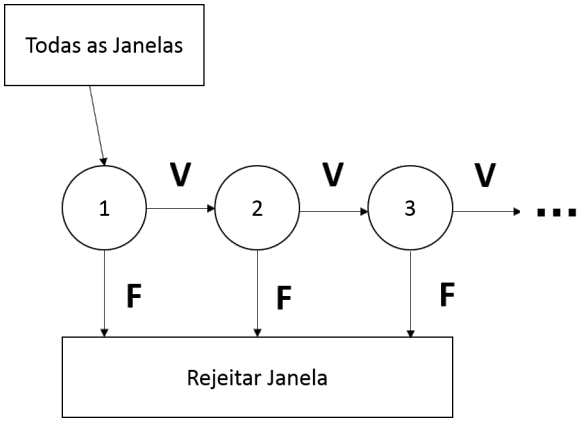


Figura 6 – Esquema de decisões da classificação em cascata



3.2.2 Treinamento de Classificador

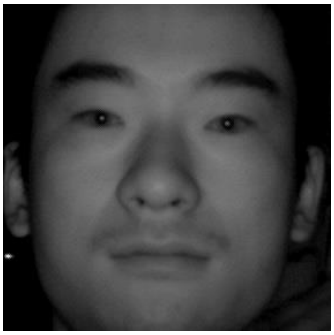
Existem muitos classificadores para imagens de textura, pois estas são as mais utilizadas pelos sistemas de detecção facial. Porém, o mesmo não ocorre para imagens em infravermelho. Por este motivo, um classificador foi criado neste trabalho utilizando funções da biblioteca OpenCV. Ela possui tanto implementações para o treinamento de classificadores quanto para a detecção facial baseados no trabalho de Viola e Jones (2001).

Para o treinamento, é preciso um conjunto de imagens positivas e negativas que permita o aprendizado do padrão desejado. Para isso, utilizamos as bases Surveillance Cameras Face Database (SCFACE) (GRGIC et al., 2011) e CASIA NIR Database (CASIA) (LI et al., 2007). A SCFACE contém 4160 imagens de 130 indivíduos, porém apenas 130 imagens estão em infravermelho. Já a CASIA possui 3940 faces de 197 pessoas em infravermelho. As imagens positivas são aquelas que contém apenas faces, e as negativas as que contém qualquer coisa exceto uma face. Para o conjunto de imagens negativas foram utilizadas as imagens destas bases com a região da face removida e também imagens aleatórias de paisagens e objetos capturados com câmeras de infravermelho, como mostrado na Figura 7.

Foram utilizadas no treinamento um total de 3986 imagens positivas e de 7645 imagens negativas. Realizamos o treino através das funções opencv_createsamples e opencv_traincascade com os parâmetros mostrados nas Tabelas 1 e 2, respectivamente. A primeira função cria novas imagens positivas aplicando transformações, distorções e rotações a partir de um conjunto de imagens positivas. Isso aumenta a variedade de imagens positivas, melhorando a capacidade de detecção do classificador. A segunda função é a responsável pela criação do classificador em cascata utilizando as imagens negativas e positivas baseado no método de Viola e Jones (2001).

Uma vez obtido o classificador, o método de detecção facial do OpenCV utiliza-o

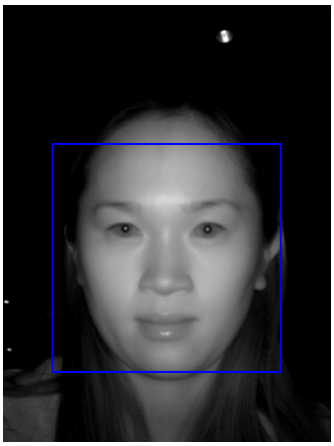
Figura 7 – (a)-(b) Exemplos de imagens para treino de um classificador e (c) resultado da detecção.



(a) Imagem positiva

(b) Imagem negativa

(c) Imagem negativa



(d) Face detectada

Tabela 1 – Tabela de parâmetros para a função opencv_createsamples

Parâmetro	**Valor**
bgcolor	0
bgthresh	0
maxxangle	1.1
maxyangle	1.1
maxidev	0.5
w	20
h	20

para localizar faces nas imagens adquiridas. Para evitar que indivíduos que transitem atrás do usuário tenham suas faces detectadas e computadas, o sistema mantém a posição da última face corretamente detectada. A cada nova detecção ele compara a distância euclidiana entre esta nova face e a última detectada. Se essa distância for maior que o limiar de 50 pixels previamente estabelecido, a face é descartada. Do contrário ela será normalizada e utilizada no reconhecimento.

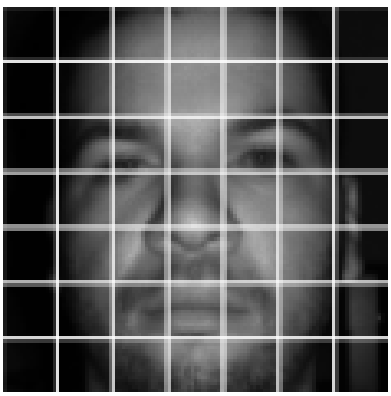
Tabela 2 – Tabela de parâmetros para a função opencv_traincascade

Parâmetro	**Valor**
data	classifier
numStages	20
minHitRate	0.999
maxFalseAlarmRate	0.5
numPos	1000
numNeg	600
w	20
h	20
mode	ALL
precalcValBufSize	1024
precalcIdxBufSize	1024

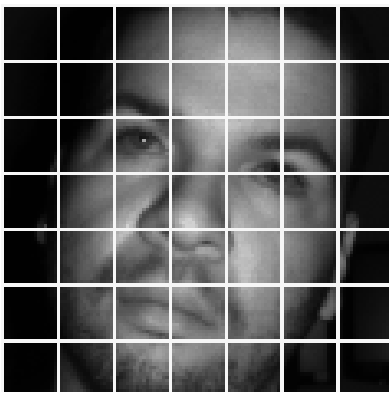
3.3 Normalização da face

É crucial para o reconhecimento facial que as faces utilizadas estejam seguindo algum padrão, pois, do contrário, o algoritmo de reconhecimento facial pode comparar partes diferentes entre as faces. Por exemplo, se uma face perfeitamente vertical for comparada com uma face com inclinação de 30 graus, uma mesma região nas duas imagens conterá características diferentes, como ilustrado na Figura 8. Para resolver esse problema, a etapa de normalização realiza uma série de transformações e processamentos de imagem em uma face detectada, como detalhado nas seções 3.3.1 e 3.3.2.

Figura 8 – Exemplo de comparação entre duas faces com variação de pose.



(a) Face detectada 1



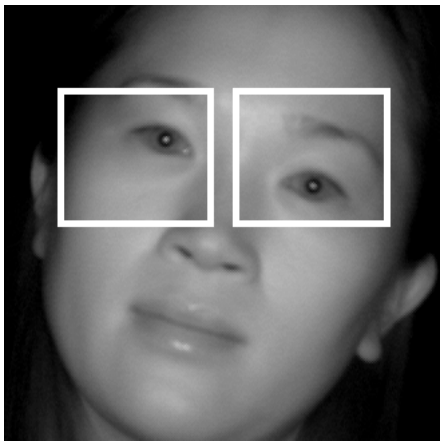
(b) Face detectada 2

3.3.1 Detecção dos olhos e normalização de pose e resolução

A primeira parte da etapa de normalização e pré-processamento da face é a detecção dos olhos. É possível assumir que os olhos estão em posição relativamente simétricas, além de possuírem uma posição e tamanho padrões em relação a face, independentemente da expressão facial. Para a detecção dos olhos, é importante limitar a busca a duas regiões

pequenas e retangulares da face onde muito provavelmente se encontram os olhos esquerdo e direito, como mostrado na Figura 9. Tentar detectar os olhos na face inteira é muito mais custoso e menos confiável, pois falsos positivos podem ser encontrados.

Figura 9 – Ilustração de duas regiões onde provavelmente se encontram os olhos.



Para a detecção, utilizamos três classificadores em cascata da biblioteca OpenCV, um específico para o olho esquerdo, um para o olho direito e um genérico para ambos. Para cada subregião da face onde se encontram os olhos, primeiro é utilizado o classificador específico daquela região. Caso não seja possível detectar o olho, o classificador mais genérico é utilizado. Quando não é possível detectar os dois olhos, suas respectivas posições são estimadas utilizando o filtro de Kalman (ZHU et al., 2002), um método para estimar posições a partir de um histórico de medições observadas. A cada novo quadro, a posição dos olhos é estimada e corrigida pelo filtro de Kalman. Para evitar o aumento no erro da normalização, esse recurso só é utilizado por um número fixo e pequeno de quadros (i.e. 5 quadros neste trabalho) em que não seja possível detectar os olhos.

Detectados os dois olhos, o ângulo entre eles é calculado para realizar uma rotação da face de maneira que eles fiquem alinhados horizontalmente. Depois a face é redimensionada para que a distância entre os olhos seja sempre a mesma, seguida de uma translação para que toda face tenha os olhos na mesma altura, como mostrado na Figura 10.

3.3.2 Normalização da iluminação e remoção de ruído

Nesta etapa, um filtro de equalização de histogramas e um filtro bilateral são aplicados. O primeiro melhora o contraste e o brilho da imagem, reduzindo a variação de iluminação decorrente da distância dos emissores de infravermelho. Porém, a equalização de histogramas pode aumentar o nível de ruído existente na imagem. O filtro bilateral é então aplicado para reduzir esses ruídos, pois ele suaviza regiões homogêneas da imagem enquanto mantém as bordas bem delimitadas. Por fim, uma máscara elíptica é aplicada

Figura 10 – Rotação, redimensionamento e translação da face.

(a) Olhos devidamente detectados

(b) Face normalizada em pose e resolução

para remover o máximo de informações irrelevantes que estão em volta da face, como cabelos, orelhas, mãos e objetos de fundo. Esse processo é mostrado na Figura 11.

Figura 11 – Etapas finais da normalização.

(a) Equalização de histograma

(b) Filtro de suavização

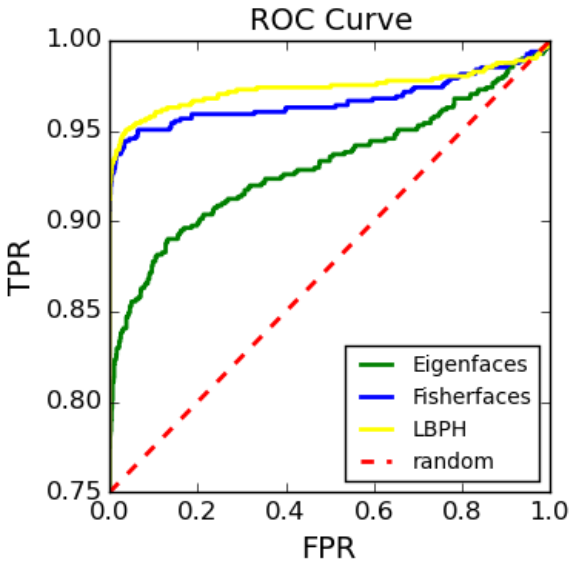
(c) Máscara elíptica

3.4 Descrição e cálculo da similaridade

Existem três métodos principais de para descrição e comparação de faces implementados na biblioteca OpenCV: Eigenfaces, Fisherfaces e Histogramas de LBP (LBPH, *Local Binary Pattern Histogram*). Cada um deles possui funções de treinamento e de predição. Para o treinamento, é preciso utilizar um conjunto de imagens faciais normalizadas, e quanto mais imagens por usuário, mais preciso é o reconhecimento, pois há mais informação sobre o padrão da face de cada usuário. Na autenticação contínua, quando o usuário faz o login, é treinado um modelo com algumas imagens de sua face. A função de predição permite que qualquer face possa ser comparada com um dado modelo, mensurando quão distante do modelo esta face está, ou seja, qual a similaridade entre a face analisada e a face do usuário que fez o login. A diferença entre os três métodos citados anteriormente

está em como eles criam um modelo e como eles calculam a similaridade. Realizamos uma experiência utilizando os três métodos para descobrir qual método possui maior capacidade de distinguir faces cadastradas de faces não cadastradas. Foram utilizadas, para cada método, 810 imagens de 90 pessoas diferentes para o treino de um modelo, 810 imagens para o teste com pessoas cadastradas e 864 imagens de 48 pessoas não cadastradas. Como mostrada na curva de característica de operação do receptor (ROC, *Receiver Operating Characteristic*) da Figura 12, o método LBPH se mostrou superior em acurácia em relação aos outros dois métodos e por isso foi utilizado neste trabalho.

Figura 12 – Curva ROC das similaridades para os métodos de reconhecimento testados.

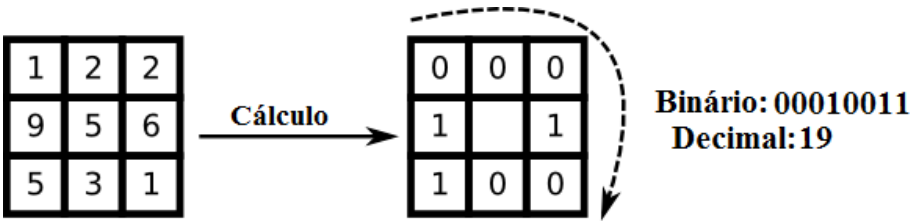


3.4.1 LBPH

Para representar a imagem de uma face, o método LBPH considera tanto a forma quanto as características específicas da face (AHONEN et al., 2004). A face é dividida em diversas regiões, e para cada região é calculado o LBP (OJALA; PIETIKAINEN; MAENPAA, 2002; OJALA; PIETIKäINEN; HARWOOD, 1996) de cada pixel. O LBP codifica cada pixel de uma imagem comparando-o com seus vizinhos. Caso o valor de intensidade do pixel for maior ou igual a intensidade do vizinho, atribuímos o valor 1 a esse vizinho, caso contrário atribuímos 0. No final do processo, cada pixel terá um valor numérico que é a concatenação dos valores atribuídos aos seus vizinhos, conforme mostrado na Figura 13.

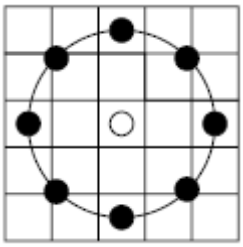
O LPB pode ser melhorado utilizando vizinhanças de diferentes tamanhos (AHONEN et al., 2004). A ideia é utilizar vizinhos em um círculo com um raio variável, como mostrado na Figura 14. Quando as coordenadas de um ponto do círculo não correspondem

Figura 13 – Operador básico LBP 3x3.



a uma coordenada da imagem, é utilizada uma interpolação linear. A quantidade de vizinhos também pode ser alterada, sendo os valores 8 e 16 os mais utilizados.

Figura 14 – Vizinhança circular (8,2).



Após calculado o LBP de cada pixel em uma região, eles são agrupados em um histograma local. Esse histograma local pode ser usado como descritor de textura. Isso porque esse histograma contém informações sobre a distribuição dos micro padrões locais, como bordas, borrões e regiões planas na região de interesse.

Os histogramas locais são então concatenados em um único histograma que representa a face como um todo. A textura das regiões faciais é codificada pelos padrões LBP, enquanto a divisão em regiões acaba por representar a forma da face. Porém, como a imagem é dividida em regiões, é esperado que algumas regiões contenham informações mais úteis que outras em termos de distinção entre pessoas. Em detrimento disso, pesos podem ser atribuídos para cada região, baseado na importância da informação que ela contém. Por exemplo, os olhos são regiões importantes para o reconhecimento de uma face humana (GONG; MCKENNA; PSARROU, 2000; ZHAO et al., 2003), e por isso recebem pesos maiores, como mostrado na Figura 15.

3.4.2 Cálculo de similaridade utilizando LBPH

Após o período de login, um modelo LPBH com raio 1 e 8 vizinhos é treinado para o usuário. Subsequentemente, para cada face detectada e normalizada com sucesso, extraímos o descritor LBPH e comparamos com o modelo treinado para gerar um valor de similaridade. A similaridade é calculada comparando o histograma da face a ser reconhecida

Figura 15 – (a) Um exemplo de imagem dividida em janelas 7×7 e os respetivos (b) pesos para cada região. Quadrados pretos indicam peso 0.0, cinza escuro 1.0, cinza claro 2.0 e branco 4.0 (AHONEN et al., 2004).

(a) (b)

com cada um dos histogramas do modelo através da fórmula Chi-Quadrado, muito utilizada para comparação de textura (PUZICHA; HOFMANN; BUHMANN, 1997). A menor das distâncias encontradas é então utilizada como medida de similaridade.

3.5 Fusão de similaridade

O sistema está seguro durante o período de login, então são capturados 5 quadros para criação do modelo do usuário. Após o login, cada quadro posterior tem sua similaridade calculada em relação ao modelo do usuário. Essa similaridade é utilizada para calcular a probabilidade do sistema ainda estar seguro. Se a probabilidade ficar abaixo de um limiar definido, o sistema é considerado inseguro e o usuário perde a autorização para utilizar o sistema.

Para cada quadro ao longo de um histórico de observações Z_t é calculada a probabilidade do sistema estar seguro, chamada de P_{seguro}, no momento t desse histórico. Cada observação $z_i \in Z_t$ corresponde a uma similaridade LBPH calculada entre a imagem atual e o modelo do usuário no instante i. Esta fusão de similaridades contínua é baseada no modelo proposto por Sim et al. (2007b), modificado por Pamplona et al. (2013). Para garantir que o usuário atual continua sendo o usuário permitido, antigas observações são esquecidas sem a necessidade de manter um histórico de observações, apenas as probabilidades da última observação e da atual são mantidas.

A probabilidade do sistema estar seguro é calculado a todo momento, com ou sem observações, de acordo com as Equações 3.1 a 3.4. O sistema deve assumir que está seguro no momento do login, então $\mathrm{P}(\text{seguro} \mid Z_0) = 1$ e $\mathrm{P}(\neg\text{seguro} \mid Z_0) = 0$, onde k é a taxa de decaimento que define quão rápido o sistema esquece antigas observações (*i.e.* P_{seguro} cai pela metade a cada k segundos sem observações), Δt é o tempo decorrido desde a última

observação z_t, X = {seguro, ¬seguro} e u é o tempo da última observação antes de t, z_u.

$$P_{seguro} = \frac{2^{\frac{-\Delta t}{k}} \times P(seguro \mid \ Z_t)}{\sum_{x \in X} P(x \mid Z_t)} \tag{3.1}$$

$$P(seguro \mid \ Z_t) = P(z_t \mid x) + 2^{\frac{(u-t)}{k}} \times P(x \mid \ Z_u) \tag{3.2}$$

$$P(z_i \mid seguro) = 1 - \frac{1}{2}\left[1 + erf\left(\frac{similaridade - \mu_{seguro}}{\sigma_{seguro} \times \sqrt{2}}\right)\right] \tag{3.3}$$

$$P(z_i \mid \neg seguro) = \frac{1}{2}\left[1 + erf\left(\frac{similaridade - \mu_{\neg seguro}}{\sigma_{\neg seguro} \times \sqrt{2}}\right)\right] \tag{3.4}$$

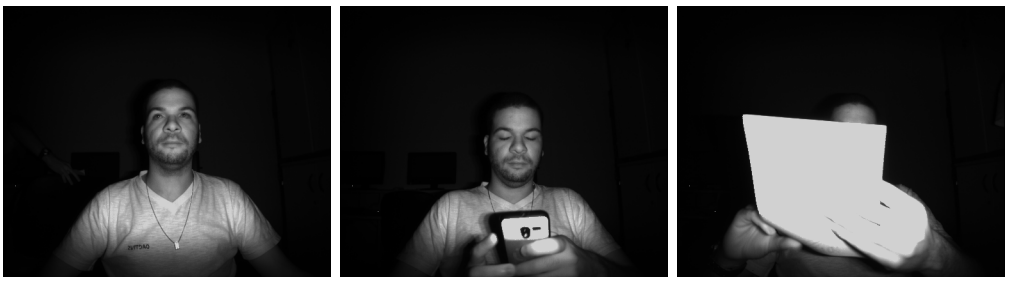
Os parâmetros (μ_{seguro}, $\mu_{\neg seguro}$, σ_{seguro}, $\sigma_{\neg seguro}$) são respectivamente (30.0509, 95.6884, 24.8666, 30.7036), e foram obtidos através de experimentos com 100 imagens faciais de 5 indivíduos, onde cada uma foi selecionada para servir como modelo e todas as outras eram comparadas com esta. Os valores de similaridades entre imagens da mesma pessoa pertencem ao grupo seguro, e entre imagens de pessoas diferentes, não-seguro. Desde modo, μ e σ são as respectivas média e desvio padrão das similaridades para os estados seguro e não-seguro. A Figura 16 mostra o sistema em execução e sua respectiva probabilidade de estar seguro.

Figura 16 – Execução do sistema. No topo é exibida a face normalizada e no canto inferior esquerdo a probabilidade do sistema estar seguro. Em cima da face é exibida sua similaridade.

4 Resultados obtidos

Para os experimentos foram gravados 4 vídeos através do Kinect. Cada um desses vídeos possui 1000 quadros em que um usuário realiza o login e utiliza o computador normalmente, sem nenhum tipo de restrição imposta aos usuários sobre como eles deveriam utilizar o computador e sobre seu comportamento, como mostrado na Figura 17.

Figura 17 – Exemplos de quadros em vídeo de teste mostrando: (a) utilização normal, (b) foco em outros objetos da cena e (c) oclusão.



(a) (b) (c)

Apenas no momento de login o usuário precisava olhar para a tela e permanecer imóvel. Esses vídeos foram utilizados como entrada para o sistema de autenticação contínua apenas com pessoas autorizadas. Para cada vídeo, foi concatenado ao seu final o início de cada outro vídeo gravado para simular uma situação de ataque e verificar se o sistema detecta intrusos. No total, foram 12 ataques realizados e os resultados são exibidos na Figura 18. Apesar de algumas sessões apresentarem uma queda no valor de P_{seguro} após o período de login, ele apresentou bom desempenho em simulações de ataque.

Como pode ser observado, na grande maioria dos quadros a probabilidade do sistema estar seguro com um usuário autorizado é maior do que a de um não autorizado. A curva ROC dos valores de P_{seguro} na Figura 19 mostra que o sistema consegue atualmente mais de 93% de acerto com uma taxa de alarme falso próxima de 7%. No eixo x da Figura 19 temos a taxa de positivos falsos (FPR, *False Positive Rate*), que representa a taxa em que um impostor foi reconhecido como o usuário permitido, e no eixo y temos a taxa de positivos verdadeiros (TPR, *True Positive Rate*), que mede o quanto os usuários permitidos foram corretamente reconhecidos.

Analisando os resultados, vemos que o momento de captura de quadros no login é crucial no desempenho do sistema. O terceiro vídeo tem uma queda no valor de P_{seguro} logo no início que se deve à diferença de pose entre o login e o resto do vídeo, que pode conter fala com pessoas da cena, oclusões e interação com outros objetos da cena além do computador.

Figura 18 – Cada gráfico é o resultado de um teste com um usuário. As linhas azuis representam os usuários autorizados nos primeiros 1000 quadros. As outras linhas representam os ataques de intrusos a partir do quadro 1000.

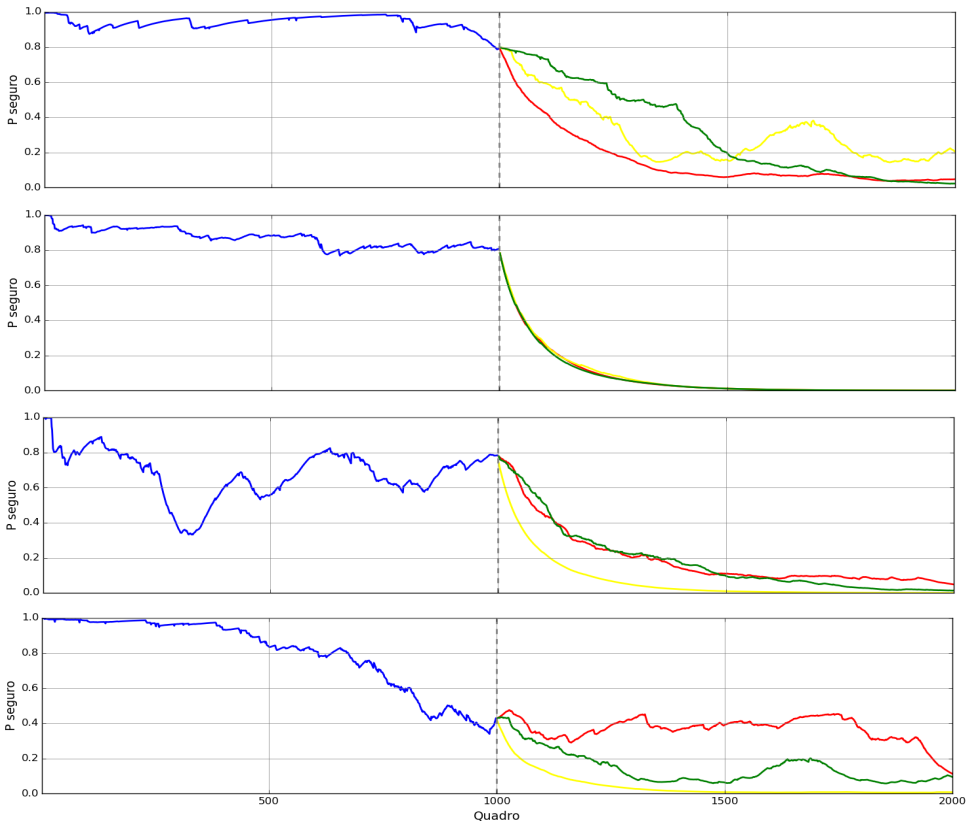
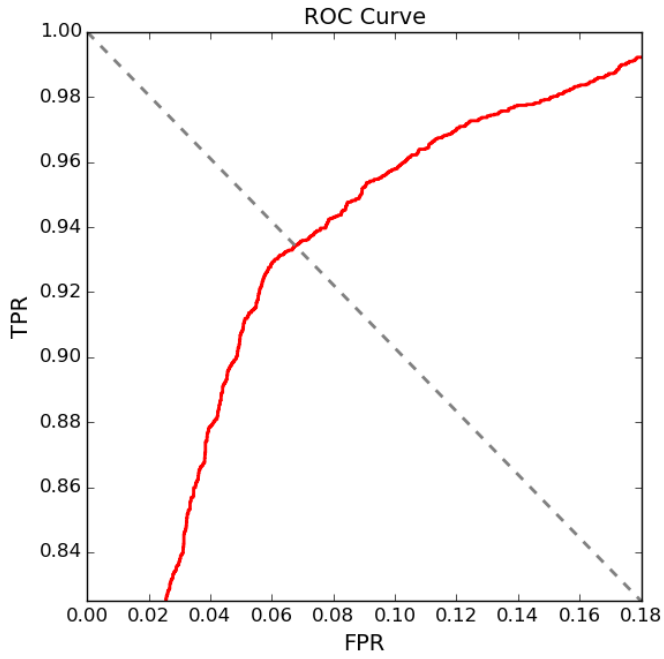


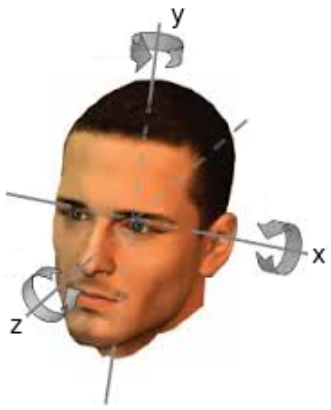
Figura 19 – Curva ROC dos valores de P_{seguro}.



5 Conclusão

Até onde sabemos, este é o primeiro sistema de autenticação contínua que utiliza faces em infravermelho para garantir que apenas o usuário permitido está utilizando o sistema. Apesar de usarmos o sensor Kinect One, o sistema pode ser utilizado com qualquer câmera que possua emissores de infravermelho e capture imagens em infravermelho. O sistema não precisa de cooperação do usuário, pois captura as imagens, detecta as faces, normaliza-as, e então calcula a similaridade e a probabilidade do sistema estar seguro, tudo de forma automática. Como as imagens em NIR são invariantes a iluminação ambiente, o sistema opera em qualquer condição de iluminação. Foram analisados um total de 4000 quadros de usuários autorizados e 12000 quadros de tentativas de invasão, alcançando 93% de TPR e 7% de FPR. É possível observar que o maior problema na autenticação facial contínua usando imagens de infravermelho é a variação de pose. A normalização é robusta no eixo z, mas não a rotações no eixo y e x, como ilustrado na Figura 20, e que causa quedas no valor de P_{seguro} mesmo com o usuário original logado.

Figura 20 – Variações de pose segundo os eixos x, y e z (LUZARDO et al., 2014).



Tendo em vista essas limitações do uso de faces em infravermelho, como trabalho futuro pretendemos utilizar o sistema desenvolvido neste trabalho para combinar os três tipos de autenticação facial contínua (i.g. luz visível, profundidade e infravermelho). O objetivo é aproveitar as vantagens de cada modalidade para formar um sistema mais seguro e robusto do que um sistema que utiliza apenas uma dessas modalidades.

Este trabalho foi publicado em SIBGRAPI, Conference on Graphics, Patterns and Images 2015, Workshop of Ungraduated Work e recebeu menção honrosa (MAGALHãES; PAMPLONA, 2015).

Referências

ABATE, A. F. et al. 2d and 3d face recognition: A survey. *Pattern Recognition Letters*, v. 28, n. 14, p. 1885 – 1906, 2007. ISSN 0167-8655. Image: Information and Control. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0167865507000189>. Citado na página 17.

ACHERMANN, B.; JIANG, X.; BUNKE, H. Face recognition using range images. In: *Virtual Systems and MultiMedia, 1997. VSMM '97. Proceedings., International Conference on*. [S.l.: s.n.], 1997. p. 129–136. Citado na página 17.

ADINI, Y.; MOSES, Y.; ULLMAN, S. Face recognition: The problem of compensating for changes in illumination direction. *IEEE Trans. Pattern Analysis and Machine Intelligence*, v. 19, n. 7, p. 721–732, July 1997. Citado na página 15.

ADINI, Y. M. Y.; ULLMAN, S. Face recognition: the problem of compensating for changes in illumination direction. *Pattern Analysis and Machine Intelligence*, v. 19, n. 7, p. 721 – 732, 1997. Citado na página 16.

AGRAFIOTI, F.; HATZINAKOS, D. Ecg biometric analysis in cardiac irregularity conditions. *Signal, Image and Video Processing*, p. 3(4):329–343, 2009. Citado 2 vezes nas páginas 13 e 14.

AHONEN, A. H. T.; PIETIKäINEN, M. Face recognition with local binary patterns. *Computer Vision - ECCV*, v. 3021, p. 469–481, 2004. Citado na página 16.

AHONEN, T. et al. Face recognition with local binary patterns. *Computer Vision - ECCV 2004*, v. 3021, p. 469–481, 2004. Citado 3 vezes nas páginas 8, 29 e 31.

BELHUMEUR, P.; HESPANHA, J.; KRIEGMAN, D. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. In: BUXTON, B.; CIPOLLA, R. (Ed.). *Computer Vision — ECCV '96*. Springer Berlin Heidelberg, 1996, (Lecture Notes in Computer Science, v. 1064). p. 43–58. ISBN 978-3-540-61122-6. Disponível em: <http://dx.doi.org/10.1007/BFb0015522>. Citado na página 16.

BHOWMIK KANKAN SAHA, S. M. G. M. A. S. A. N. S. D. B. D. K. B. M. K.; NASIPURI, M. Thermal infrared face recognition – a biometric identification technique for robust security system. *Reviews, Refinements and New Ideas in Face Recognition*, 2007. Citado na página 17.

CHANG, K. I.; BOWYER, K. W.; FLYNN, P. J. Face recognition using 2d and 3d facial data. In: *ACM Workshop on Multimodal User Authentication*. [S.l.: s.n.], 2003. p. 25–32. Citado na página 14.

COOTES, T.; EDWARDS, G.; TAYLOR, C. Active appearance models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 23, n. 6, p. 681–685, Jun 2001. ISSN 0162-8828. Citado na página 16.

DAMOUSIS, I. G.; TZOVARAS, D.; BEKIARIS., E. Unobtrusive multimodal biometric authentication: the humabio project concept. *EURASIP Journal on Advances in Signal Processing*, v. 2008, p. 110:1–110:11, 2008. Citado 2 vezes nas páginas 13 e 18.

DUGELAY, J. L. et al. Recent advances in biometric person authentication. In: *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*. [S.l.: s.n.], 2002. v. 4, p. IV–4060–IV–4063. Citado na página 13.

FERNANDES, S.; BALA, J. Performance analysis of pca-based and lda-based algorithms for face recognition. *International Journal of Signal Processing Systems*, v. 1, n. 1, p. 1–6, 2013. Citado na página 16.

FLIOR, E.; KOWALSKI, K. Continuous biometric user authentication in online examinations. *Seventh International Conference on Information Technology: New Generations,*, 2010. Citado na página 13.

FREUND, Y.; SCHAPIRE, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.*, Academic Press, Inc., Orlando, FL, USA, v. 55, n. 1, p. 119–139, ago. 1997. ISSN 0022-0000. Disponível em: <http://dx.doi.org/10.1006/jcss.1997.1504>. Citado na página 22.

GONG, S.; MCKENNA, S. J.; PSARROU, A. *Dynamic Vision: From Images to Face Recognition*. 1st. ed. London, UK, UK: Imperial College Press, 2000. Citado na página 30.

GRGIC, M. et al. Scface - surveillance cameras face database. *Multimedia Tools and Applications Journal*, v. 51, n. 3, p. 863–879, February 2011. Citado na página 24.

GUNETTI, D.; PICARDI, C. Keystroke analysis of free text. *ACM Trans. Inf. Syst. Secur*, 2005. Citado na página 13.

GUNETTI, D.; PICARDI, C. Keystroke analysis of free text. *ACM Trans. Info. and System Security*, p. 8(3):312–347, 2005. Citado na página 13.

H, M.; PJ, P. Computational and performance aspects of pca-based face-recognition algorithms. *Perception*, v. 30, p. 303 – 321, 2001. Citado na página 16.

HUANG, D.; WANG, Y.-H.; WANG, Y.-D. A robust infrared face recognition method based on adaboost gabor features. *Wavelet Analysis and Pattern Recognition, 2007. ICWAPR '07. International Conference on*, v. 3, p. 1114–1118, Nov 2007. Citado na página 17.

JANAKIRAMAN, R. et al. Using continuous face verification to improve desktop security. In: *Application of Computer Vision, 2005. WACV/MOTIONS '05 Volume 1. Seventh IEEE Workshops on*. [S.l.: s.n.], 2005. v. 1, p. 501–507. Citado 2 vezes nas páginas 13 e 16.

KONG, S. et al. Recent advances in visual and infrared face recognition—a review. *Computer Vision and Image Understanding*, v. 97, p. 103–135, January 2005. Issue 1. Citado na página 15.

KONG, S. G. et al. Recent advances in visual and infrared face recognition—a review. *Computer Vision and Image Understanding 97 (2005) 103–135*, v. 97, p. 103–135, April 2004. Issue 1. Citado na página 13.

LEGGETT, J. et al. Dynamic identity verification via keystroke characteristics. *Int'l Jrnl. of Man-Machine Studies*, p. 35(6):859–870, 1991. Citado na página 13.

LI, D.-Y.; LIAO, W.-H. Facial feature detection in near-infrared images. *Proc. of 5th International Conference on Computer Vision, Pattern Recognition and Image Processing*, 2003. Citado na página 15.

LI LUN ZHANG, S. L. X. Z. R. C. M. A. S. Z.; HE, R. A near-infrared image based face recognition system. *Automatic Face and Gesture Recognition*, p. 455 – 460, 2006. Citado na página 17.

LI, S. et al. Illumination invariant face recognition using near-infrared images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 29, n. 4, April 2007. Citado na página 24.

LI, S. Z. et al. *in Proc. IAPR Int. Conf. Biometrics*, p. 151–158. Citado na página 17.

LI, S. Z. et al. Illumination invariant face recognition using near-infrared images. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, v. 29, n. 4, p. 627 – 639, April 2007. Citado na página 15.

LU, K. N. P. J.; VENETSANOPOULOS, A. N. Face recognition using lda-based algorithms. *Neural Networks, IEEE Transactions*, v. 14, p. 195 – 200, 2003. Citado na página 16.

LUZARDO, M. et al. Estimating head pose and state of facial elements for sign language video. In: *Proceedings of 9th Language Resources and Evaluation Conference (LREC 2014), Reykjavik, Iceland*. [S.l.: s.n.], 2014. Citado 2 vezes nas páginas 9 e 35.

MAGALHãES, M.; PAMPLONA, M. Autenticação facial contínua usando imagens de infravermelho. *CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES, 28. (SIBGRAPI)*, 2015. Citado na página 35.

MAVADATI, M. T. S. S. M.; KITTLER, J. Fusion of visible and synthesised near infrared information for face authentication. *Image Processing (ICIP)*, p. 3801 – 3804, 2010. Citado na página 16.

MOCK, K. et al. Real-time continuous iris recognition for authentication using an eye tracker. In: *Proceedings of the 2012 ACM Conference on Computer and Communications Security*. New York, NY, USA: ACM, 2012. (CCS '12), p. 1007–1009. ISBN 978-1-4503-1651-4. Disponível em: <http://doi.acm.org/10.1145/2382196.2382307>. Citado na página 13.

MOGHADDAM, W. W. B.; PENTLAND, A. Beyond eigenfaces: Probabilistic matching for face recognition. *Automatic Face and Gesture Recognition*, p. 30 – 35, 1998. Citado na página 16.

NAKANISHI, I.; BABA, S.; MIYAMOTO, C. Eeg based biometric authentication using new spectral features. In: *Intelligent Signal Processing and Communication Systems, 2009. ISPACS 2009. International Symposium on*. [S.l.: s.n.], 2009. p. 651–654. Citado na página 13.

NIINUMA, K.; PARK, U.; JAIN, A. Soft biometric traits for continuous user authentication. *Information Forensics and Security, IEEE Transactions on*, v. 5, n. 4, p. 771–780, Dec 2010. Citado 2 vezes nas páginas 13 e 16.

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 24, n. 7, p. 971–987, Jul 2002. Citado na página 29.

OJALA, T.; PIETIKäINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, v. 29, n. 1, p. 51 – 59, 1996. Citado na página 29.

PAMPLONA, M. S. et al. Continuous 3d face authentication using rgb-d cameras. *Computer Vision and Pattern Recognition Workshops (CVPRW)*, p. 64 – 69, June 2013. Citado 2 vezes nas páginas 17 e 31.

PAPAGEORGIOU, C. P.; OREN, M.; POGGIO, T. A general framework for object detection. *Proceedings of the Sixth International Conference on Computer Vision*, p. 555–, 1998. Citado na página 21.

PUZICHA, J.; HOFMANN, T.; BUHMANN, J. Non-parametric similarity measures for unsupervised texture segmentation and image retrieval. In: *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. [S.l.: s.n.], 1997. p. 267–272. ISSN 1063-6919. Citado na página 31.

SIM, T. et al. Continuous verification using multimodal biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 29, n. 4, April 2007. Citado 3 vezes nas páginas 13, 16 e 18.

SIM, T. et al. Continuous verification using multimodal biometrics. *IEEE PAMI*, p. 9(4):687–700, 2007. Citado 2 vezes nas páginas 14 e 31.

VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. *Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 511–518, 2001. Citado 6 vezes nas páginas 8, 20, 21, 22, 23 e 24.

WEI, T.; ZHIHUA, X. Infrared face recognition based on local binary pattern and multi-objective genetic algorithm. *Proceeding of the IEEE International Conference on Information and Automation Shenzhen*, p. 359 – 362, June 2011. Citado na página 15.

ZHAO, S.; GRIGAT, R.-R. An automatic face recognition system in the near infrared spectrum. *Machine Learning and Data Mining in Pattern Recognition*, v. 3587, p. 437–444, 2005. Citado na página 16.

ZHAO, W. et al. Face recognition: A literature survey. *ACM Comput. Surv.*, ACM, New York, NY, USA, v. 35, n. 4, p. 399–458, dez. 2003. Citado na página 30.

ZHENG, Y. Near infrared face recognition using orientation-based face patterns. *Biometrics Special Interest Group (BIOSIG)*, p. 1 – 4, 2012. Citado na página 15.

ZHIHUA, X.; GUODONG, L. Weighted infrared face recognition in multiwavelet domain. *Imaging Systems and Techniques (IST)*, p. 70 – 74, 2013. Citado na página 15.

ZHU, Z. et al. Combining kalman filtering and mean shift for real time eye tracking under active ir illumination. In: *Pattern Recognition, 2002. Proceedings. 16th International Conference on*. [S.l.: s.n.], 2002. v. 4, p. 318–321 vol.4. ISSN 1051-4651. Citado na página 27.

ZOU, X.; KITTLER, J.; MESSER, K. *Face Recognition Using Active Near-IR Illumination.* Citado na página 17.