UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

PROPOSTA DE CRITÉRIOS PARA CÂMERAS DE VIGILÂNCIA EM APLICAÇÕES DE CFTV *INDOOR* PARA FINS DE IDENTIFICAÇÃO FORENSE DE SUSPEITOS

RODRIGO ALBERNAZ BEZERRA

ORIENTADOR: PROFº. DR. DÍBIO LEANDRO BORGES

DISSERTAÇÃO DE MESTRADO EM ENGENHARIA ELÉTRICA
ÁREA DE CONCENTRAÇÃO INFORMÁTICA FORENSE E SEGURANÇA DA INFORMAÇÃO

PUBLICAÇÃO: PPGENE.DM - XX A / XX
BRASÍLIA/DF: FEVEREIRO – 2012

UNIVERSIDADE DE BRASÍLIA

FACULDADE DE TECNOLOGIA

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

PROPOSTA DE CRITÉRIOS PARA CÂMERAS DE VIGILÂNCIA EM APLICAÇÕES DE CFTV *INDOOR* PARA FINS DE IDENTIFICAÇÃO FORENSE DE SUSPEITOS

RODRIGO ALBERNAZ BEZERRA

DISSERTAÇÃO DE MESTRADO SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA ELÉTRICA DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE PROFISSIONAL EM INFORMÁTICA FORENSE E SEGURANÇA DA INFORMAÇÃO.

APROVADA POR:

_____

Díbio Leandro Borges, PhD (CIC/UNB)
(Orientador)

_____

Anderson Clayton Alves Nascimento, PhD (ENE/FT)
(Examinador Interno)

_____

Flávio de Barros Vidal, Dr. (CIC/UNB)
(Examinador Externo)

BRASÍLIA/DF, 17 DE FEVEREIRO DE 2012.

FICHA CATALOGRÁFICA

BEZERRA, RODRIGO ALBERNAZ
Proposta de Critérios para Câmeras de Vigilância em Aplicações de CFTV Indoor para Fins de Identificação Forense de Suspeitos [Distrito Federal] 2012.
xvii, 84p., 297 mm (ENE/FT/UnB, Mestre, Engenharia Elétrica, 2012).

Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia. Departamento de Engenharia Elétrica.

1. Reconhecimento Facial de Suspeitos 2. Segurança Orgânica
3. Dimensionamento de Projetos de CFTV 4. Exames Periciais
I. ENE/FT/UnB II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

BEZERRA, R. A. (2012). Proposta de Critérios para Câmeras de Vigilância em Aplicações de CFTV Indoor para Fins de Identificação Forense de Suspeitos. Dissertação de Mestrado em Engenharia Elétrica, Publicação PPGENE.DM - XX A/XX, Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, DF, 84p.

CESSÃO DE DIREITOS

AUTOR: Rodrigo Albernaz Bezerra.

TÍTULO: Proposta de Critérios para Câmeras de Vigilância em Aplicações de CFTV Indoor para Fins de Identificação Forense de Suspeitos.

GRAU: Mestre ANO: 2012

_____

Rodrigo Albernaz Bezerra

Universidade de Brasília

Campus Universitário Darcy Ribeiro – Asa Norte

CEP 70.910-900 Brasília/DF/Brasil

Dedico esta dissertação aos meus pais, irmãs e à minha futura família.

AGRADECIMENTOS

Este é um momento único, singular, pois conseguir expressar meus agradecimentos a todos aqueles que me acompanharam durante esta penosa fase de minha vida não é uma tarefa fácil. Limitar-me-ei, portanto, a agradecer àqueles com quem convivo no momento e, desde já, peço desculpas aos que não foram citados.

Em primeiro lugar, gostaria de agradecer a Deus. Sem ele eu não estaria aqui. Mas meus agradecimentos iniciais vão além da minha própria existência. Quando eu não acreditava mais em mim, Ele continuou acreditando. E, Ele sabe que dei muitos motivos para desistir de mim, mas Ele continuou. Obrigado, Deus, por ser autor e guia da minha vida.

Agradeço aos meus pais, Ronaldo e Ilda, por me educarem e por não terem medido esforços para isso, sempre me apoiando e incentivando. Às minhas irmãs Hevelin e Leilane, maravilhosas pessoas que estão sempre acreditando no meu sucesso. Aos animais de estimação da família, fontes constantes de amor e alegria.

Agradeço também ao meu orientador, o Prof. Dr. Díbio, pela sua paciência e disponibilidade para me orientar, mesmo nos momentos mais críticos durante a fase de dissertação. Obrigado pela confiança a mim dispensada.

Ao meu ex-chefe de Setor, Daniel e ao meu atual chefe, Ricardo, por terem autorizado minha participação nesta pós-graduação. Obrigado por compreenderem que a atividade de um Perito Criminal Federal não é só laboral, mas sim, por se tratar de um trabalho científico, exige contínua capacitação.

Ao programa de Pós Graduação em Engenharia Elétrica da UNB, que me proporcionou a infraestrutura necessária para a realização deste trabalho.

Aos professores do curso de mestrado, por dividirem conhecimento e experiências.

Aos funcionários do programa de Mestrado em Informática Forense e Segurança da Informação, pela companhia e pelos divertidos momentos compartilhados durante esta etapa da minha vida.

Aos amigos de curso, pela troca de ideias, divertimentos e valiosa ajuda durante toda a fase de créditos do mestrado.

Aos colegas do mestrado da FGV, pela amizade, momentos de reflexão e descontração.

Gostaria de agradecer em especial aos parceiros de corrida de kart, válvula de escape em momentos tensos após exercícios e provas que exigiram grande esforço intelectual dos alunos.

Aos meus fiéis amigos, Luciano, Túlio e Victor, pela paciência, amizade e convivência durante todos esses anos que nos conhecemos.

Ao cunhado Bruno, pelo incentivo e amizade.

À Patrícia, por toda nossa história.

À Nathaly, uma pessoa especial neste momento da minha vida.

Ao Departamento de Polícia Federal – DPF e ao Ministério da Justiça, através dos recursos do PRONASCI – Programa Nacional de Segurança Pública com Cidadania, idealizadores e incentivadores deste curso de Pós Graduação, por promoverem a melhoria da Criminalística no Brasil.

Este trabalho não ficaria completo sem agradecer a todos os que me ajudaram a concretizá-lo. Fiz um grande esforço para lembrar-me das principais pessoas que contribuíram ao meu trabalho durante todo o mestrado. Para aqueles das quais não lembrei, estão aqui os meus sinceros agradecimentos.

RESUMO

PROPOSTA DE CRITÉRIOS PARA CÂMERAS DE VIGILÂNCIA EM APLICAÇÕES DE CFTV INDOOR PARA FINS DE IDENTIFICAÇÃO FORENSE DE SUSPEITOS.

Autor: Rodrigo Albernaz Bezerra
Orientador: Prof. Dr. Díbio Leandro Borges
Programa de Pós-graduação em Engenharia Elétrica
Brasília, Fevereiro de 2012

Os sistemas de CFTV utilizados em ambientes orgânicos, empresas privadas, órgãos do governo ou condomínios residenciais, apresentam certas deficiências técnicas que inviabilizam o reconhecimento facial de suspeitos quando da ocorrência de crimes. Esses sistemas falham em fornecer imagens com qualidade suficiente para que a face de uma pessoa possa ser reconhecida com segurança na esfera forense (identificação criminal). Dentre os principais problemas encontrados nesses sistemas podem ser enumerados o uso de equipamentos mal dimensionados, posicionamento inadequado das câmeras de vigilância, geração de imagens com baixa resolução e condições de iluminação ambientais não controladas. Não há no Brasil norma técnica específica que subsidie projetos e instalações de sistemas de CFTV. A atividade probatória criminal do Estado, materializada através dos exames periciais, busca auxiliar as denúncias direcionadas pelo Ministério Público e a atuação jurisdicional do Poder Judiciário, no combate às atividades criminosas. Entretanto, a utilização de imagens coletadas em sistemas de CFTV para fins probatórios tem pouco contribuído na solução dos crimes, em virtude das deficiências que esses sistemas apresentam. A proposta deste trabalho é propor critérios mínimos necessários para o dimensionamento de projetos de câmeras de vigilância de sistemas de CFTV, em aplicações *indoor*, de forma que as imagens geradas apresentem qualidade suficiente para permitir o reconhecimento facial de indivíduos. Foram tratadas como critérios as condições de posicionamento das câmeras (altura e ângulo), de iluminação ambiental, de resolução espacial das imagens e distância mínima para fins de reconhecimento facial. Adicionalmente foi tratada uma sistemática de auditoria para os arquivos gerados em sistemas de CFTV, de forma a robustecer a prova (manter a cadeia de custódia) gerada para adequada apreciação pelo Judiciário.

Palavras-chave: Projetos de CFTV. Câmeras de Vigilância. Reconhecimento Facial. Aplicações Forenses.

ABSTRACT

PROPOSAL OF CRITERIA FOR SURVEILLANCE CAMERAS IN INDOOR CCTV APPLICATIONS FOR FORENSIC IDENTIFICATION OF SUSPECTS.

Author: Rodrigo Albernaz Bezerra
Advisor: Prof. Dr. Díbio Leandro Borges
Programa de Pós-graduação em Engenharia Elétrica
Brasília, February of 2012

CCTV systems used in organic environments, private companies, government agencies or residential condominiums, present certain technical deficiencies that prevent the face recognition of suspects when a crime occurs. These systems fail to provide images with sufficient quality for the reliable face recognition in forensic sphere (criminal identification). Among the main problems encountered in these systems can be enumerated the wrong sized equipment, improper placement of surveillance cameras, imaging with low resolution and lighting conditions uncontrolled environmental. There is no technical standard in Brazil to subsidize specific projects and installations of CCTV systems. Evidential criminal justice, materialized through technical expertise, seeks to assist the work performed by prosecutors and court actions of the judiciary in combating criminal activities. However, the use of images collected from CCTV systems for evidential purposes has little contribution to solve the crimes, due to the CCTV systems technical deficiencies. The goal of this work is to propose minimum criteria for the design of projects of CCTV surveillance cameras equipment, in indoor applications, so they can provide face images of suspects of crimes with sufficient quality to enable face recognition activity. It will be discuss in these criteria the conditions for positioning of the cameras (height and angle), environmental lighting conditions, spatial resolution of images and minimum distance for face recognition. Additionally, a systematic audit files system for CCTV was treated, in order to strengthen the proof (to maintain chain of custody) generated for adequate consideration by the judiciary.

Keywords: CCTV design. Surveillance Cameras. Face Recognition. Forensic Applications.

SUMÁRIO

LISTA DE TABELAS

LISTA DE FIGURAS

LISTA DE GRÁFICOS

LISTA DE SÍMBOLOS, NOMENCLATURAS E ABREVIAÇÕES

AFIS - *Automated Fingerprint Identification System*

AFR - *Automatic Face Recognition*

AFV - *Face Recognition Vendor Test*

AGC - *Automatic Gain Control*

AIC - *Australian Institute of Criminology*

ATM - *Automatic Teller Machine*

ATW - *Automatic Tracing White Balance*

BLC - *Back light Compensation*

BNC - *Bayonet Neill-Concelman*

CCD - *Charged-Coupled Device*

CD - *Compact Disc*

CFTV - Circuito Fechado de Televisão

CMOS - *Complementary Metal-Oxide Semiconductor*

CPP - Código de Processo Penal

CRC - *Cyclic Redundancy Check* (Verificação cíclica de redundância)

CRT - *Cathode Ray Tube*

DCT - *Discrete Cosine Transform*

DITEC - Diretoria Técnico Científica da Polícia Federal

DPF - Departamento de Polícia Federal

DVD - *Digital Versatile Disc*

DVR - *Digital Video Recorder*

EGBM - *Elastic Graph Bunch Model*

FAR - *False Accept Rate*

FERET - *Face Recognition Technology*

FoV - *Field of View* (Campo de Visão)

FPS - *Frame per second*

FRGC - *Face Recognition Grand Challenge*

FRR - *False Reject Rate*

IAI - *International Association for Identification*

IBGE - Instituto Brasileiro de Geografia e Estatística

INMETRO - Instituto Nacional de Metrologia, Qualidade e Tecnologia

IP - *Internet Protocol*

IPD - *Interpuppilary distance* (distância entre as pupilas)

IR - *Infrared* (infravermelho)

JPEG - *Joint Photographic Expert Group* (padrão de compressão de imagens)

JPEG2000 - *Joint Photographic Experts Group 2000* - padrão de compressão de imagens.

LCD - *Liquid Crystal Display*

LDA - *Linear Discriminant Analysis*

LED - *Light-Emitting Diode*

M-JPEG2000 - *Motion-JPEG* 2000 (padrão de compressão de vídeo)

MPEG - *Moving Picture Expert Group* (padrão para compressão de áudio e vídeo)

NIST - *National Institute of Standars and Technology*

OPENCV - *Open Computer Vision Library*

PCA - *Principle Component Analysis*

POP - Procedimento Operacional Padrão

PRONASCI - Programa Nacional de Segurança Pública com Cidadania

RGB - *Red Green Blue* (Vermelho Verde Azul)

SHA - *Secure Hash Algorithm*

SKL - *The Swedish National Laboratory of Forensic Science*

SNR - *Signal-to-noise ratio*

SSD - *Solid-State Disk*

SWGIT - *Scientific Working Group on Imaging Technology*

TIFF - *Tagged Image File Format* (formato de arquivo de imagem)

UNB - Universidade de Brasília

TVL - Television lines (linhas de televisão horizontal)

VHS - *Video Home System* (Sistema Caseiro de Vídeo)

YIQ - *Luminance (Y), In-phase (I) e Quadrature (Q) color space*

1 – INTRODUÇÃO

1.1 – CONSIDERAÇÕES INICIAIS

O nível de criminalidade no país, bem como os episódios de terrorismo ocorridos em diversas partes do mundo, tem alertado a população para as questões de segurança. Todos os dias a mídia televisiva apresenta matérias relacionadas à violência, crime e operações policiais. A prevenção ao crime é um conceito que tem sido integrado ao cotidiano das pessoas (PEASE, 1997).

Esta demanda cada vez maior por segurança tem obrigado diversos setores da sociedade a tratarem a segurança pessoal e patrimonial como prioridade (VALERA E VELASTIN, 2005). E essas estratégias não são um privilégio da esfera privada, mas pertencem também ao cotidiano da atuação governamental, no enfrentamento ao crime (PEASE, 1997). E, de fato, as estratégias de prevenção ao crime têm sido objeto de ação governamental para a manutenção da segurança pública. Dentre as varias ações desenvolvidas está o uso de sistemas de CFTV como uma das estratégias de prevenção primária, método que tem mais tem se proliferado ao redor do mundo (REIS, 2006). Em Londres, por exemplo, os perímetros do metrô e do aeroporto de Heathrow possuem vigilância por mais de 10.000 câmeras (VALERA E VELASTIN, 2005).

É importante mencionar que importantes investimentos têm sido feitos pelo governo, seja da esfera federal ou estadual, em setores da segurança pública associados à atividade probatória criminal. A modernização dos setores de Perícia do Estado tem permitido evoluir a Criminalística nacional, nos últimos anos, em termos de processos e procedimentos, traduzindo estes investimentos em resultados que têm chamado a atenção da sociedade, como pôde ser observado na mídia televisiva, nas reportagens abordando os exames periciais e simulações realizados no Caso Isabella Nardoni, em São Paulo (SERIBELI, 2008).

Desde sua criação a partir do século XIX, as Ciências Forenses, em especial, referindo-se à Criminalística, têm ajudado a solucionar crimes e prender criminosos, utilizando métodos e técnicas científicas de áreas diversas. A Criminalística tem ajudado os órgãos integrantes

da persecução penal brasileira, procurando mostrar, através das provas periciais, a verdade real dos fatos, contribuindo assim, com as investigações do Poder Executivo (Polícia), com as denúncias do Ministério Público e com o convencimento dos juízes na formulação de suas sentenças (Poder Judiciário). E há interesse social para que os setores de Perícia dos Estados estejam preparados para o exercício de suas atividades, pois este preparo está diretamente ligado ao combate da impunidade de criminosos. A consequência deste fato pode auxiliar a condenar adequadamente os suspeitos culpados de crimes, bem como isentar da culpa aqueles suspeitos inocentes.

A instalação de sistemas de segurança associados ao uso de CFTV tem se tornando comum no país. Sistemas de CFTV têm sido encontrados inclusive em ambientes residenciais, cujos moradores buscam maior segurança para seus lares. Essa proliferação está associada ao custo benefício deste tipo de sistema de segurança, aliado ao baixo custo de instalação, manutenção e à eficiência de muitas soluções disponíveis no mercado (COSTA, 2010). Entretanto, a ausência de normativos pode levar clientes à adquirirem produtos de baixa qualidade, utilizando um computador, uma placa de captura de vídeo e uma câmera tipo *webcam*, que podem vir a não fornecer qualidade de imagem suficiente para aplicações de reconhecimento facial de suspeitos. O uso desses sistemas precários de CFTV tem sido encontrado inclusive em órgãos públicos, dificultando ainda mais o trabalho investigativo da polícia.

Arquivos de imagens ou vídeos podem se tornar importantes evidências a serem utilizadas pela polícia e pelos órgãos periciais na identificação de suspeitos de crimes. De acordo com Russ (2009), as imagens são responsáveis por um grande percentual de informações que as pessoas adquirem sobre o ambiente ao seu redor, tendo reservada aplicação na seara forense, como por exemplo, auxiliando na identificação da possível dinâmica e autoria de um crime (SAFERSTEIN, 2011).

É muito comum que o conteúdo de imagens e vídeo extraídos de sistemas de CFTV seja encaminhado para exames periciais, no intuito de ser melhorado, buscando trazer informações não vistas a olho nu por investigadores, tais como a identificação de uma placa veicular ou o reconhecimento facial de um suspeito. Entretanto, apesar dos esforços técnicos periciais, há pouca contribuição com a investigação quando a fonte material apresenta baixa qualidade. Em alguns momentos, esses exames podem auxiliar na

identificação de alguma peça de vestuário ou podem ajudar a isentar de culpa alguém detido como suspeito, mas raramente consegue-se extrair de imagens ou vídeos de baixa qualidade provas conclusivas para a identificação unívoca de um suspeito (RUSS, 2009). As Figuras 1.1 e 1.2 apresentam uma amostra de imagens típicas, extraídas de sistemas de CFTV em locais de crime no país, que são encaminhadas para a polícia e órgãos pericias investigarem.

Figura 1.1 − Imagem de CFTV de uma agência de Correios. Nos dois quadros de imagens captados, notam-se problemas na cadência de captação das imagens (efeito *blurring* nas faces).

Figura 1.2 − Imagem de CFTV de uma agência de Correios. Nota-se problemas nas condições de iluminação do ambiente (luz natural forte e luz artificial fraca), criando regiões de sombra na face dos indivíduos monitorados.

A Biometria é um termo aplicado a muitas maneiras às quais uma pessoa pode ser identificada através de aspectos do seu corpo (NIETO, JOHNSON-DODDS e SIMMONS, 2002). Formalmente, a Biometria pode ser conceituada como a ciência do reconhecimento

da identidade de uma pessoa por meio de suas características físicas ou comportamentais, como a face, as impressões digitais, a voz, a retina ou íris dos olhos, características únicas de movimento, DNA, utilizando para isso recursos eletrônicos e computacionais (GARCIA, 2009). A premissa fundamental da Biometria é a de que cada indivíduo possui características físicas e comportamentais únicas, capazes de diferenciá-los entre si (MACHADO *et al.*, 2010). Técnicas biométricas são utilizadas em documentos de identidade nacional, passaportes, controles de acesso físico ou de dados, controle de transações financeiras e na identificação criminal (ZHAO *et al.*, 2003). No Brasil, as impressões digitais são o tipo de identificador biométrico utilizado na identificação criminal. Há planos e projetos para estender os tipos de biometria utilizados na identificação criminal, como o uso de sistemas de banco de dados por DNA e por reconhecimento facial, mas não se trata de uma realidade encontrada na Criminalística brasileira. O Gráfico 1.1 apresenta as principais técnicas biométricas utilizadas na identificação de pessoas.

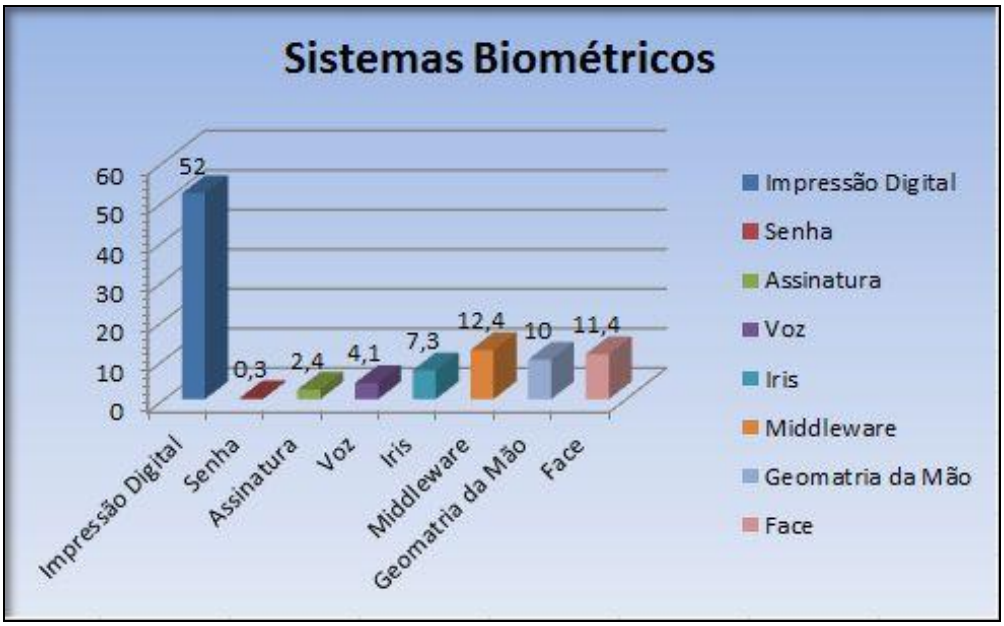


Gráfico 1.1 − Tipos mais populares de sistemas biométricos, em termos porcentuais (ABATE *et al.*, 2007).

O reconhecimento facial é uma área de pesquisa integrante das Ciências Biométricas (MACHADO *et al.*, 2010), que tem despertado interesse de pesquisadores de vários campos do conhecimento científico, como da Visão Computacional, do Reconhecimento de Padrões, da Inteligência Artificial, do Processamento Digital de Imagens (OLIVEIRA,

2006) e inclusive das Ciências Forenses. Esse interesse, muitas vezes econômicos, ocorre em virtude da grande gama de aplicações derivadas. Por exemplo, muitas pesquisas têm surgido no intuito de desenvolverem métodos para sistemas automatizados de reconhecimento facial com foco em aplicações de segurança. Esses sistemas buscam comparar faces de pessoas, extraídas de imagens estáticas ou de vídeos, com outras armazenadas em um banco de dados, gerando relatórios positivos, para os casos de identificação de faces convergentes entre si, ou negativos, nos casos contrários (RODRIGUES, 2007).

A utilização de sistemas de CFTV, bem como o aperfeiçoamento de técnicas biométricas, como as de reconhecimento facial computadorizado podem ser utilizadas em conjunto para facilitar a capacidade de aplicação da lei em identificar suspeitos de terrorismo ou criminosos (NIETO, JOHNSON-DODDS e SIMMONS, 2002). Essa aplicação legal pode ser conduzida através de normas que especifiquem critérios para sistemas de CFTV que forneçam imagens com detalhes para aplicações forenses e que vão ser utilizadas em atividades de investigação e exames periciais ou mesmo inseridas em banco de dados para a identificação criminal. O uso adequado da tecnologia pode ser um forte aliado da lei.

1.2 – DEFINIÇÃO DO PROBLEMA DE PESQUISA

A motivação para o desenvolvimento deste trabalho de pesquisa vem da frustração, durante muitos anos de trabalho para o Departamento de Polícia Federal, realizando exames periciais em arquivos de imagem estática e de vídeos, de câmeras de vigilância instaladas em órgãos do governo federal e em alguns casos, em empresas privadas, na qual não houve êxito em realizar o reconhecimento facial de suspeitos da realização de crimes, pela falta de qualidade das imagens geradas pelos equipamentos de segurança. Ademais, não existe normativo técnico no país que subsidie projetos de instalação de câmeras em sistemas de CFTV.

Kovesi (2009) relata a mesma problemática levantada no parágrafo anterior, ao tentar auxiliar instituições policiais, na tentativa de melhoria das imagens de baixa qualidade capturadas por sistemas de segurança CFTV. Segundo o autor, mesmo utilizando ferramentas que buscam melhorar a qualidade de imagem, nenhum ganho real foi obtido,

ou seja, as imagens resultantes não adicionaram melhorias suficientes a ponto de trazer informações extras, além daquelas já presentes na imagem de baixa qualidade.

Uma questão importante a ser ressaltada com relação a muitos experimentos publicados associados ao reconhecimento facial de pessoas está no uso de imagens com alta qualidade e resolução espacial, como por exemplo, em Bruce *et al.* (1999). A realidade de muitos sistemas de vigilância é outro. No Brasil, por exemplo, há sistemas de CFTV de pequena escala que ainda possuem base analógica e utilizam fitas de vídeo VHS para gravação das imagens monitoradas. Outros, apesar de também possuírem base analógica, procuraram aperfeiçoar o bloco de gravação de dados para a tecnologia digital, utilizando mídias óticas ou discos rígidos. Entretanto, a qualidade das imagens capturadas e gravadas continua sendo baixa, inadequadas para fins de reconhecimento facial de suspeitos.

Além deste fator, tais sistemas são frequentemente instalados com pouca atenção com relação às condições de iluminação ou ângulo de visão das câmeras se vigilância. Isto significa que quando uma sequência de imagens ou vídeo é necessária para ser utilizada como prova, por exemplo, na ocorrência de um crime, é quase sempre inviável confirmar, através de exames periciais ou mesmo através de sistemas de reconhecimento de faces, se a pessoa capturada nas imagens do dispositivo de segurança é a mesma pessoa acusada ou suspeita do crime (BURTON *et al.*, 1999).

É dentro do contexto doravante apresentado que o presente trabalho está inserido. Busca-se uma forma de auxiliar os trabalhos investigativos e periciais da polícia na solução da dinâmica e autoria de crimes.

1.3 – ESCOPO E OBJETIVOS DO TRABALHO

Este trabalho de mestrado é parte de um projeto de capacitação profissional idealizado pela Diretoria Técnico Científica da Polícia Federal (DITEC/DPF), na qual foi firmada uma parceria entre a Secretaria Nacional de Segurança Pública (SENASP/MJ) e a Universidade de Brasília (UNB), de forma a prover capacitação em nível de mestrado profissional para Peritos Criminais da Polícia Federal e das Polícias Civis e/ou Institutos de Criminalística Estaduais.

O intuito desse projeto foi o de subsidiar a capacitação de um grupo de Peritos Criminais, aprovados por meio de um edital de seleção, em nível de pós-graduação *strictu sensu*, na área de Informática Forense e Segurança Computacional, estreitar relações entre os profissionais selecionados de forma a desenvolver a área da Computação Forense em nível nacional, além de incentivar, através dos projetos de pesquisa a serem desenvolvidos na capacitação, o desenvolvimento e evolução da Computação Forense e da própria Criminalística.

O cerne do problema está na questão de que muitos sistemas de segurança em utilização são ineficazes no auxílio ao reconhecimento facial de suspeitos, principalmente quando o foco é o cenário forense. Não existe no país normativo técnico associado a projetos de sistemas de CFTV, que possa servir como regra para empresas privadas e órgãos públicos, durante a aquisição de soluções de segurança, no dimensionamento de seus projetos de CFTV.

Kovesi (2009) utiliza uma expressão muito feliz que caracteriza adequadamente a realidade de grande parte dos sistemas de CFTV no país: são *legalmente cegos*. Os fatores que contribuem para a ineficácia desses sistemas são diversos, dentre eles, pode-se destacar a ausência normas técnicas sobre o assunto, os baixos investimentos em tecnologia de segurança, equipamentos de má qualidade, posicionamento inadequado das câmeras de vigilância, condições de iluminação do ambiente insatisfatórias, imagens e vídeos captados com baixa resolução espacial, compressão de arquivos, dentre outros.

Diante desta lacuna normativa apresentada, o objetivo geral deste projeto de pesquisa é definir parâmetros para um conjunto de critérios para o dimensionamento de projetos de câmeras de vigilância utilizadas em sistemas de CFTV. O foco dessa parametrização é a capacidade desses sistemas auxiliarem no reconhecimento facial de pessoas, com o fim de padronizar os requisitos mínimos necessários para que esses sistemas possam colaborar com a seara forense na identificação de criminosos, de forma a auxiliar a Polícia, o Ministério Público e o Poder Judiciário, no combate à criminalidade.

Para viabilizar a realização da proposta, o escopo do projeto foi restringido e limitado, pois o grande número de variáveis a serem consideradas na proposta de critérios para projetos de CFTV inviabilizaria sua conclusão dentro de um prazo considerado apropriado para um

projeto de mestrado. Desta forma, este projeto aborda apenas alguns critérios dentre os necessários para o projeto de instalação de câmeras de vigilância em aplicações de sistemas de CFTV *indoor* (ambientes internos), quais sejam, o posicionamento das câmeras em relação a altura e ângulo de trabalho, as condições de iluminação do ambiente, resolução espacial das imagens geradas e a distância mínima para fins de reconhecimento facial. Os experimentos foram testados utilizando um sistema de reconhecimento facial baseado em imagens estáticas. Aspectos relacionados aos demais equipamentos integrantes de sistemas de CFTV, bem como o monitoramento de ambientes externos (abertos) e sistemas baseados em vídeo não foram abordados neste trabalho.

1.4 – PRINCIPAIS CONTRIBUIÇÕES

As principais contribuições esperadas com o presente trabalho são as seguintes:

- Desenvolvimento de um conjunto de critérios para o dimensionamento de projetos de instalação de câmeras de vigilância em aplicações CFTV *indoor*, para fins de identificação unívoca de suspeitos[1] de crimes;
- Alertar sobre as deficiências dos sistemas de segurança existentes no país, no que tange à materialização e identificação da autoria de delitos;
- Servir de guia para novos projetos ou aprimoramento de sistemas já implantados, seja no nível público ou no privado, de forma a fornecer subsídios para a definição das condições de projeto para a instalação de câmeras de vigilância;
- Auxiliar, de forma secundária, a Polícia (investigação policial e exames periciais), Ministério Público e o Judiciário, na condenação de criminosos, através do fornecimento de imagens de suspeitos com qualidade suficiente para realizar o reconhecimento facial.

[1] A expressão identificação de suspeitos utilizada no título desta dissertação e doravante em todo este documento tem o mesmo significado de reconhecimento facial, ou seja, que se possa distinguir uma pessoa de outra com confiança, utilizando características próprias da face.

1.5 – ORGANIZAÇÃO DA DISSERTAÇÃO

Esta dissertação foi estruturada em 8 capítulos, conforme apresentação a seguir.

No capítulo um é apresentada a introdução do trabalho. Foram discutidos neste capítulo a contextualização do trabalho de pesquisa, a definição do problema de pesquisa, os objetivos gerais, a delimitação do trabalho, bem como foi apresentado a estrutura da dissertação.

O capítulo dois discute brevemente sobre questões forenses. São abordados neste capítulo a admissão de provas periciais perante o Poder Judiciário, bem como alguns tipos de exames periciais que usufruem das imagens de sistemas de segurança, para a materialização do delito e da autoria de crimes.

O capítulo três apresenta conceitos práticos relacionados a sistemas de CFTV, importantes para a compreensão harmônica de elementos da proposta do trabalho.

O capítulo quatro discute aspectos sobre sistemas de reconhecimento facial e apresenta o estado da arte em técnicas de detecção facial e reconhecimento facial.

O capítulo cinco apresenta trabalhos correlacionados ao projeto de pesquisa, que auxiliaram na solução da proposta apresentada.

O capítulo seis trata da proposta do trabalho propriamente dito, apresentando os critérios analisados.

O capítulo sete detalha os experimentos e testes utilizados para avaliar os critérios propostos no capítulo anterior, discutindo seus resultados.

Por fim, o capítulo oito apresenta as conclusões finais e direciona sugestões de trabalhos futuros dentro da linha de pesquisa abordada.

2 – PROVAS PERICIAIS

Este capítulo trata de aspectos relacionados ao trabalho forense. Parte da motivação para o desenvolvimento deste trabalho foi derivada da experiência do autor na execução de suas atividades como perito criminal federal, sendo importante inserir o presente capítulo no trabalho, como forma de contextualização do trabalho de pesquisa do mestrado.

A Seção 2.1 apresenta os aspectos legais relacionados à formação probatória pela Criminalística e a Seção 2.2 trata de aspectos ligados à doutrina criminalística, bem como de alguns exames periciais relacionados ao reconhecimento facial de suspeitos de delitos.

2.1 – ASPECTOS LEGAIS

O ordenamento legal brasileiro prevê dois tipos de perícias: aquelas relacionadas à solução de conflitos de interesses (*perícias cíveis*) e aquelas relacionadas à pretensão punitiva do Estado (*perícias criminais*). O Código de Processo Penal[2] (CPP) é o dispositivo jurídico que rege as perícias criminais, possuindo, além de artigos esparsos em seu corpo legal, dois capítulos específicos que tratam do assunto no TÍTULO VII – DA PROVA: CAPÍTULO I – DISPOSIÇÕES GERAIS (arts. 155 a 157) e CAPÍTULO II - DO EXAME DO CORPO DE DELITO, E DAS PERÍCIAS EM GERAL (arts. 158 a 184).

No processo criminal, busca-se a verdade real dos fatos, procurando saber como eles se passaram na realidade, sem quaisquer presunções ou outros artifícios. A reprodução da verdade no processo penal deve ser feita através da busca das melhores provas em matéria criminal, sendo que o juiz não pode se contentar apenas com aquelas fornecidas pelas partes, salvo se forem efetivamente as melhores (SOARES, 2008). A prova, de acordo com a definição vernacular, significa aquilo que serve para estabelecer uma verdade por verificação ou demonstração (ESPINDULA, 2009). Desta forma, o juiz e as partes (autor e réu) têm ampla liberdade para provarem suas afirmações e, para isso, dispõem dos meios de prova (OLIVEIRA e SANCHEZ, 2008). Segundo Demercian e Maluly (2005), meio de prova é tudo o que possa ser utilizado para a demonstração da ocorrência dos fatos

[2] Código de Processo Penal (CPP) disponível em: http://www.planalto.gov.br/ccivil_03/decreto-lei/Del3689.htm . Acesso em 18/01/2012.

alegados e perseguidos no processo. São ainda, os instrumentos necessários para comprovar a existência ou não da verdade de um fato.

Os meios de prova podem ser classificados em objetivos (periciais e documentais) e subjetivos (testemunhos e confissões). A prova pericial ou científica, em especial, possui um revestimento próprio diferenciado, tendo em vista a sua natureza técnico científica. Ela é toda estudada com base em diversos ramos do conhecimento científico, sendo considerada, inegavelmente, como um dos meios de prova mais consistentes para robustecer um processo judicial. E essa preferência é facilmente explicável, pois, enquanto a prova pericial é produzida a partir de fundamentação científica dos elementos materiais deixados pela ação delituosa, nos meios de prova subjetivos pode ocorrer uma série de erros, desde a simples falta de capacidade da pessoa em relatar determinado fato ou a situação de má-fé, na qual existe a intenção particular de distorcer os fatos para não se chegar à verdade (ESPINDULA, 2009).

O sistema pericial criminal brasileiro encontra-se estruturado como um órgão de apoio ao sistema judiciário, na produção das provas periciais, que subsidiam tanto a investigação policial (fase pré-processual), quanto o processo criminal (fase processual penal), (ESPINDULA, 2009). O perito é um auxiliar do juízo, pois colabora com o juiz nos pontos em que este precisa de esclarecimentos ou de conhecimentos especializados (TÁVORA e ALENCAR, 2011). A Lei Federal nº 12.030[3], sancionada no ano de 2009, dispõe sobre as perícias oficiais, relacionado em seu artigo 5º, a definição dos servidores legalmente admitidos com a definição de peritos oficiais, no âmbito do processo penal, conforme texto legal apresentado a seguir:

> Art. 5º Observado o disposto na legislação específica de cada ente a que o perito se encontra vinculado, são peritos de natureza criminal os peritos criminais, peritos médico-legistas e peritos odontolegistas com formação superior específica detalhada em regulamento, de acordo com a necessidade de cada órgão e por área de atuação profissional (BRASIL, 2009).

[3] Lei nº 12.030 disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2009/Lei/L12030.htm . Acesso em 18/01/2012.

É função do Estado a execução das atividades de perícia criminal e o motivo de existência desta atividade está aposta no artigo 158 do CPP, que possui o seguinte conteúdo:

> Art. 158. Quando a infração deixar vestígios, será indispensável o exame de corpo de delito, direto ou indireto, não podendo supri-lo a confissão do acusado (BRASIL, 1941).

A interpretação do referido artigo é clara e define como regra, a indispensabilidade da realização do exame de corpo de delito quando a infração penal deixar vestígios. É necessário enfatizar que, apesar dos *exames de corpo de delito* estarem naturalmente associados a exames médico-legais em pessoas vivas ou mortas, a expressão na lei também abrange os vestígios relacionados a todas a modalidades de infrações penais (ARAÚJO, 2011).

O *caput* do artigo 159 do CPP ainda dispõe que:

> Art. 159. O exame de corpo de delito e outras perícias serão realizados por perito oficial, portador de diploma de curso superior (BRASIL, 1941).

A legislação processual penal brasileira é transparente em demonstrar que a função do perito oficial e a sua atividade em produzir provas periciais possui reservada importância no auxílio de toda a cadeia da persecução penal. Materializar os fatos ocorridos na infração penal constitui um dos objetivos do inquérito policial, pressuposto para o oferecimento da denúncia ou queixa, para a decretação da prisão preventiva, a decisão de pronúncia, a sentença absolutório-imprópria e a condenatória. A prova da existência dos fatos, via de regra, é pericial. Além disso, a prova pericial também é necessária à aferição de requisitos legais subjetivos indispensáveis à concessão de benefícios na execução penal (MANZANO, 2011).

Por fim, cabe ressaltar que apesar da notória relevância das provas no processo criminal, sejam periciais ou não, é necessário frisar que, segundo entendimento abstraído dos artigos 155, 200 e 381 do Código de Processo Penal, o sistema de apreciação de provas pátrio é misto, ou seja, o juiz formará sua convicção pela livre apreciação das provas produzidas

em contraditório judicial, devendo justificar formalmente no âmbito da sentença, os motivos de fato e de direito aos quais fundamentaram sua decisão, dentre eles, a preferência ao aceitar ou rejeitar uma prova (ESPINDULA, 2009; MANZANO, 2011).

> Art. 155. O juiz formará sua convicção pela livre apreciação da prova produzida em contraditório judicial, não podendo fundamentar sua decisão exclusivamente nos elementos informativos colhidos na investigação, ressalvadas as provas cautelares, não repetíveis e antecipadas.
>
> Art. 200. A confissão será divisível e retratável, sem prejuízo do livre convencimento do juiz, fundado no exame das provas em conjunto.
>
> Art. 381. A sentença conterá:
> I - os nomes das partes ou, quando não possível, as indicações necessárias para identificá-las;
> II - a exposição sucinta da acusação e da defesa;
> III - a indicação dos motivos de fato e de direito em que se fundar a decisão;
> IV - a indicação dos artigos de lei aplicados;
> V - o dispositivo;
> VI - a data e a assinatura do juiz (BRASIL, 1941).

2.1.1 – ADMISSIBILIDADE DAS PROVAS PERICIAIS

A despeito das regras previstas no CPP e na Legislação Penal Extravagante, não existe lei específica sobre o procedimento probatório a ser adotado em todas as perícias. Mesmo quanto se trata de perícias nominadas em lei, tais como a autópsia, a exumação e as perícias laboratoriais, o procedimento técnico envolvido não se encontra neles disciplinado (MANZANO, 2011).

A aceitação de uma prova pericial no âmbito processual é relegada ao juiz. Surge então, o problema da admissibilidade e assunção da prova pericial, referentes ao regime de aquisição do elemento ou fonte de prova. A fase de admissão é aquela em que o juiz, de ofício ou mediante requerimento da parte, provê a decisão sobre a validade da fonte ou meio de prova. Se o juízo for positivo, prossegue-se à fase de assunção, que consiste na

operação da concreta entrega do elemento de prova no processo, como dado utilizável pela sentença (MANZANO, 2011).

No campo da prova pericial, a assunção envolve o procedimento técnico realizado para evidenciar o elemento de prova, a partir da experiência do perito. O processo legislativo brasileiro dificilmente acompanharia o rápido avanço científico, pelo que seria impensável disciplinar todos os procedimentos técnicos envolvidos no procedimento probatório pericial. Por essa razão, a admissibilidade, para posterior utilização da prova pericial no processo demanda, via de regra, um juízo de admissibilidade dessa prova no processo, visando aferir a confiabilidade da prova pericial. (MANZANO, 2011).

2.2 – ASPECTOS DOUTRINÁRIOS

A Criminalística é um termo utilizando em língua portuguesa para fazer referência à Ciência Forense e à Perícia criminal e está relacionada aos métodos científicos utilizados pela polícia para a materialização de provas (ARAÚJO, 2011). Segundo Carvalho (2006), a Criminalística é uma ciência que, para atender aos seus objetivos, se utiliza de outras ciências, de algumas artes e disciplinas, e, sem se subordinar a nenhuma delas, lhes faz exigências nos procedimentos técnicos científicos, quando lhes oferece o mínimo de material para ser examinado e requer que os resultados extraídos possam ser novamente refeitos.

A execução das atividades da Criminalística no país pode ser dividida em dois segmentos: federal e estadual. No âmbito da União Federal, o Departamento de Polícia Federal, órgão de polícia judiciária, possui a atribuição de realização das perícias criminais, através dos policiais ocupantes do cargo de Perito Criminal Federal. No âmbito dos Estados e do Distrito Federal, as perícias criminais são realizadas também por Peritos, cujos cargos têm várias nomenclaturas. No segmento estadual, os Institutos de Criminalística nem sempre estão ligados a órgãos policiais, em muitos casos, são órgãos independentes inseridos dentro da estrutura da Secretaria Estadual de Segurança Pública.

A característica comum dos profissionais da Criminalística, independente da nomenclatura, está no fato de que realizam exames técnico científicos no interesse de processos criminais (ARAÚJO, 2011). Esses exames são materializados através do Laudo Pericial, documento

oficial que expõe em detalhes os procedimentos periciais realizados sobre o fato delituoso, bem como os resultados auferidos no trabalho. E, segundo Carvalho (2009), o trabalho do perito criminal é árduo e importante, pois é esse profissional o responsável pela produção da prova material, necessária à elucidação de casos onde a infração penal deixar vestígios, fornecendo à justiça elementos necessários que permitem a um juiz condenar ou absolver, com segurança, um acusado. O autor acrescenta ainda que a tarefa é árdua, pois impõe ao perito a busca de um aprendizado permanente e evolutivo das ciências, artes e disciplinas afins, bem como do conhecimento e operacionalidade de equipamentos de tecnologia sofisticada e, a aplicação, sem sofismas, deste aprendizado nas pesquisas técnico científicas.

A Criminalística, como Ciência, apresenta também doutrina própria, com postulados e princípios que subsidiam e auxiliam o trabalho do perito criminal. Cabe destacar neste momento, os principais postulados da Criminalística (DOREA, STUMVOLL e QUINTELA, 2010):

1) O conteúdo de um Laudo de Perícia Criminal é invariante em relação ao profissional Perito Criminal que o produziu, pois como os resultados de um trabalho pericial são invariavelmente baseados no método científico, seja qual for o perito que recorrer a ele para analisar um fenômeno criminalístico, o resultado não poderá depender dele, indivíduo.
2) As conclusões de um Laudo Pericial Criminal são independentes dos meios utilizados para alcançá-las, pois utilizando-se os meios adequados para se concluir a respeito do fenômeno criminalístico, esta conclusão durante a reprodução dos exames será constante.
3) Os resultados de um trabalho pericial são invariáveis com o tempo, ou seja, a partir da realização dos exames, os resultados aferidos no trabalho, por utilizarem o método científico, são imutáveis com o decorrer do tempo, desde que as evidências não sejam voláteis.

Já os princípios fundamentais da Perícia Criminalística referem-se à observação, à análise, à interpretação, à descrição e à documentação da prova, como mostrado abaixo (DOREA, STUMVOLL e QUINTELA, 2010):

1) *Princípio da Observação*: também chamado de Princípio da Troca de Locard, em virtude do famoso médico criminalístico, Edmond Locard, qualquer um ou qualquer coisa, que entra em um local de crime, leva consigo algo do local e deixa alguma coisa para trás quando parte.
2) *Princípio da Análise*: a análise pericial deve sempre seguir o método científico. A perícia científica visa a definir como o fato ocorreu (teoria), através de uma criteriosa coleta de dados (vestígios e indícios), que permitem estabelecer-se conjecturas sobre como se desenvolveu o fato, formulando-se hipóteses coerentes sobre ele. É esse o método científico que baseiam as condutas periciais.
3) *Princípio da Interpretação*: também chamado Princípio da Individualidade, preconiza que a identificação deve ser sempre enquadrada em três graus, ou sejam, a identificação genérica, a específica e a individual, sendo que os exames periciais deverão sempre alcançar este último grau.
4) *Princípio da Descrição*: os resultados dos exames periciais, baseados em princípios científicos, não podem variar pela passagem do tempo, ou seja, os resultados obtidos devem ser os mesmos; e, ainda, considerando que qualquer teoria científica deve gozar da propriedade da refutabilidade, os resultados da perícia, quando expostos através do Laudo, devem ser de uma forma bem claros, racionalmente dispostos, bem fundamentados e reproduzíveis.
5) *Princípio da Documentação*: este princípio, baseado na Cadeia de Custódia da prova material, visa a proteger, seguramente, a fidelidade da prova material, evitando a consideração de provas forjadas, incluídas no conjunto das demais, para provocar a incriminação ou a inocência de alguém. Assim, todo o material apreendido deve ser oficializado *in loco*. Ainda, todos aqueles que manipularem qualquer dos materiais apreendidos, devem documentar o feito, procurando manter a integridade das evidências.

Após apresentar brevemente a Criminalística neste trabalho, cabe por fim, neste capítulo, retratar alguns exames periciais que se relacionam com o tema deste trabalho e fizeram parte, de certa forma, da motivação para o desenvolvimento desta pesquisa. Serão apresentados dois exames periciais: exames de locais de crime e exames de reconhecimento facial.

2.2.1 – EXAMES DE LOCAL DE CRIME

O local de crime pode ser definido como sendo uma área física onde ocorreu um fato, não esclarecido até então, que apresente características ou configurações de um delito e por isso, reclama providências da polícia. Mais especificamente, local de crime é todo espaço físico onde ocorreu uma prática de infração penal, podendo ser uma área física externa, interna ou mista (ESPINDULA, 2009).

É consenso que todo o exame de local se transforme no ponto de partida para uma investigação criminal (DOREA, STUMVOLL e QUINTELA, 2010). O início de qualquer procedimento para o esclarecimento de um delito será o local onde o crime ocorreu. Nesse sentido, é necessário que a polícia tome conhecimento de imediato, a fim de tomar as providências necessárias de investigação daqueles fatos (ESPINDULA, 2009). Havendo o comunicado à instituição policial sobre locais desta natureza, os peritos criminais são requisitados a comparecer no local para a realização de exames periciais de local de crime, não havendo impedimento ou exclusão dos trabalhos investigativos, que podem ocorrer simultaneamente, através de outros policiais.

Os tipos de delitos que podem ocorrer nos locais de crime são diversos. No entanto, didaticamente, é feita a seguinte divisão básica para fins de organização dos exames periciais: perícias em crimes contra a pessoa, perícias em crimes de acidente de tráfego e perícias em crimes contra o patrimônio, sendo que este último agrupa uma grande quantidade de tipos de delitos, alguns dos quais já ganharam títulos próprios e técnicas específicas de realização dos exames perícias, como é o caso de perícias de incêndio e meio ambiente (ESPINDULA, 2009).

A casuística da Perícia da Polícia Federal neste tipo de exame é diversa e a depender da região federativa considerada, alguns exames são mais frequentes do que outros. A título de exemplo, pode-se listar que em muitos Estados do país, as perícias em crimes contra o patrimônio são mais frequentes, relacionados a crimes de roubo ou furto em agências bancárias da Caixa Econômica Federal ou em agências dos Correios. Em outros, crimes em áreas ambientais possuem grandes demandas para o Setor de Perícias.

Um dos grandes e graves problemas das perícias em locais onde ocorrem crimes é a quase inexistente preocupação das autoridades em isolar e preservar adequadamente um local de infração penal, de maneira a garantir as condições de se realizar um exame pericial da melhor forma possível. No país não há uma cultura e nem mesmo uma preocupação sistêmica, o correto isolamento do local do crime e por consequência, a respectiva preservação dos vestígios do delito (ESPINDULA, 2009). Desta forma, por interferências externas, quando os peritos chegam a local de crime, o cenário do crime pode ser outro diverso daquele esperado quando da ocorrência do crime, podendo haver a adulteração ou destruição de vestígios importantes para a materialização do fato criminoso.

Durante os procedimentos periciais de um local de crime, o perito criminal deve então, isolar o perímetro para evitar interferências externas que possam vir a contaminar o ambiente, caso isso já não tenha sido feita pelas autoridades que primeiro chegaram ao local, fazer levantamento fotográfico do local, bem como constatar a existência de vestígios atinentes ao crime ali cometido. Posteriormente, cada detalhe encontrado na cena de crime será devidamente interpretado no sentido de procurar estabelecer o que realmente ali aconteceu e quem foram os responsáveis causais do evento.

O elo de ligação do projeto de pesquisa deste trabalho com os exames de locais de crime são os possíveis vestígios encontrados na cena. Como já foi abordado, a Perícia da Polícia Federal atua em muitos delitos contra o patrimônio, principalmente em casos de roubo ou furto com arrombamento em agências bancárias da Caixa Econômica Federal e em agências dos Correios. Durante as perícias de local nestes casos, é comum se deparar com imagens ou vídeo de sistemas de CFTV que integram sistemas de segurança destes estabelecimentos públicos.

Na tentativa de coibir ou impedir crimes, os estabelecimentos públicos, principalmente aqueles que atuam com alguma atividade financeira e dispõem de cofres para resguardo de valores monetários, utilizam sistemas de segurança, como o CFTV, para evitar ações criminosas que possam obter a apropriação ilícitas de objetos ou bens de valor. Entretanto, muitos desses sistemas de CFTV não possuem configuração técnica suficiente para disponibilizar imagens dos suspeitos do crime, com o grau necessário para que seja realizado seu reconhecimento facial. Em algumas situações, os sistemas são desligados à noite ou quando funcionam, as câmeras não possuem configuração para modo noturno,

captando apenas imagens escuras. Técnicas de melhoria de contraste e redução de ruídos podem até serem aplicados nesses tipos de imagens, procurando trazer a tona novos detalhes, mas quase nunca são suficientes para fins de identificação de suspeitos (RUSS, 2009). Desta forma, é quase impossível sustentar por meio probatório que algum suspeito é o mesmo que foi flagrado pelas câmeras do sistema CFTV. Ainda sim, é um vestígio útil de ser coletado e analisado em locais de crime, permitindo, por exemplo, fazer a exclusão de suspeitos ou analisar como foi a dinâmica do crime.

As figuras a seguir apresentam imagens extraídas de casos reais de locais de crime em agências de Correios[4]. Como pode ser visualmente observado, as imagens de CFTV apresentam baixa resolução espacial e não permitem o reconhecimento facial dos suspeitos.

Figura 2.1 – Imagem extraída de sistema de CFTV de agência dos Correios mostrando suspeito de roubo (Resolução 352 X 288 e IPD (distância entre pupilas) = 10,0 pixels). O *motion blur* ocorreu nos quadros em que o suspeito foi flagrado frontalmente pela câmera, mas não está presente em todos os quadros do vídeo. Excesso de movimentos diante da taxa de disparos podem ocasionar efeitos de *blurring* na imagem, o que dificulta o reconhecimento facial.

[4] A localização das agências de Correios foi propositalmente ocultada para manter a segurança desses estabelecimentos.

Figura 2.2 – Imagem extraída de arquivo de vídeo de CFTV de agência dos Correios mostrando suspeito de roubo (Resolução 350 X 236 e IPD = 10,0 pixels).

2.2.2 – EXAMES DE RECONHECIMENTO FACIAL

Segundo Machado *et al.* (2010), o termo reconhecimento para o escopo pericial refere-se a um processo de identificação, de classificação, dos parâmetros faciais a fim de estabelecer a unicidade entre duas ou mais faces. Busca-se realizar um processo de classificação da face, levantando padrões, requisitos, estatísticos e morfológicos, de forma a estabelecer a comparação entre as faces padrão e questionada, para posteriormente concluir, baseado nas convergências ou divergências, se estas atendem a todos os critérios para o reconhecimento de um indivíduo.

Enfim, os exames periciais de reconhecimento facial consistem em um processo de captura e comparação entre as características biométricas faciais, de duas ou mais faces, utilizando métodos científicos como, por exemplo, processamento digital de imagens, fotogrametria facial (vide Figura 2.3) e antropometria craniofacial (vide Figura 2.4). Os exames de reconhecimento facial visam classificar características das faces e têm por função apontar as suas convergências ou divergências. Assim, os resultados somente serão conclusivos se as técnicas utilizadas apontarem convergências ou divergências suficientes, nos exames de verificação 1 x 1, para subsidiarem os peritos criminais a afirmar se tratar ou não do mesmo indivíduo (MACHADO *et al.*, 2010).

O reconhecimento facial é uma das áreas de estudo das Ciências Biométricas, como também o são os estudos de impressões papilares, da voz, da retina, da íris e da geometria da mão (MACHADO *et al.*, 2010). Algumas biometrias, como as citadas, permitem a identificação inequívoca das pessoas, por terem altíssimo grau de fidedignidade. Isso

ocorre em decorrência do fato de haver um número extremamente grande de variações nos detalhes dessas biometrias, o que torna possível afirmar que não há mais de uma pessoa com as biometrias apresentadas de forma idêntica. Assim, apenas uma pessoa possui uma determinada impressão digital, uma íris ou uma retina com exatamente o mesmo desenho, isto é, idênticas. No caso da biometria facial, ainda não há estudos estatísticos acerca da individualidade facial das pessoas, nem bancos de dados criminais que utilizem deste recurso, motivo pelo qual tem sido difícil a aplicação do reconhecimento facial automatizado de indivíduos em termos probatórios (seara da Criminalística). A identificação biométrica visual por face tem sido aplicada naqueles casos em que outras formas de identificação não foram utilizadas ou não há outro tipo de evidência a não ser imagens sobre os suspeitos quando da ocorrência do crime.

Na casuística criminal, o processo de reconhecimento facial é comumente realizado de forma indireta, por meio de registros em imagens estáticas, como fotografias ou imagens de CFTV, e imagens dinâmicas, como vídeos, com o objetivo determinar se as anatomias faciais dos indivíduos presentes nos registros visuais são compatíveis (MACHADO *et al.*, 2010). A aplicação desta técnica exige que sejam observados alguns critérios, como: condições adequadas de iluminação na obtenção das fotos, fotos obtidas em posições frontais e eretas. Além disso, é importante que as fotos questionadas e padrões sejam contemporâneos entre si e tenham sido obtidas na mesma posição (DPF, 2008).

Esses exames não são subjetivos, como ocorre, por exemplo, na realização de um retrato falado ou em um exame prosopográfico (descrição de uma pessoa), pelo contrário, possuem base científica e permitem determinar a identidade[5] de um indivíduo (MACHADO *et al.*, 2010). As Figuras 2.3 a 2.5 apresentam exemplo de uso de técnicas de reconhecimento facial para fins de classificação da face.

[5] O sentido de identidade apresentado neste parágrafo refere-se ao conjunto de medidas faciais utilizadas para individualizar uma pessoa, ou seja, fazer seu reconhecimento facial (Machado *et al.*, 2010).

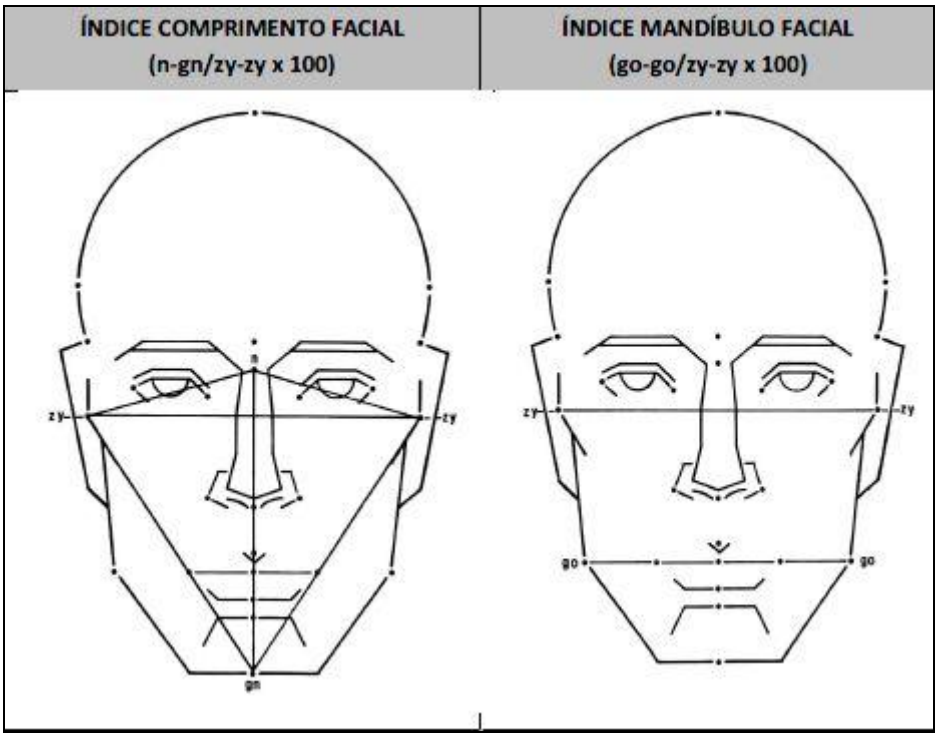


Figura 2.3 – Amostra de técnica utilizada para determinar características da face (GEORGE, 2007).

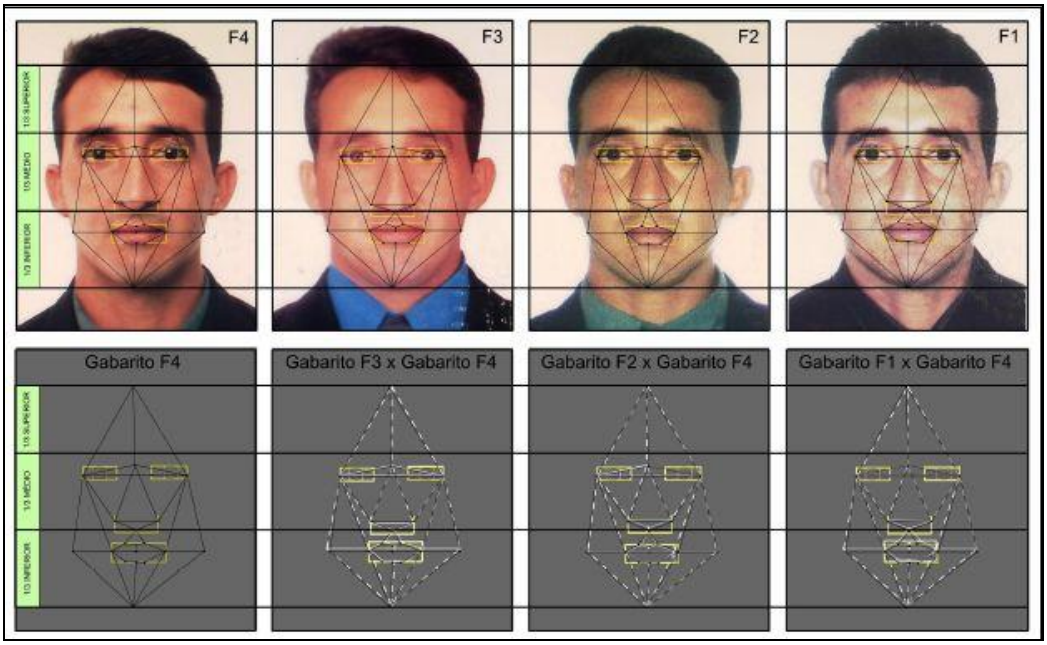


Figura 2.4 – Aplicação de técnicas de reconhecimento facial (análise cefalométrica de tecidos moles) em fotos de passaporte (DPF, 2010).

As imagens a seguir apresentam dois exemplos de casos reais de exames de reconhecimento facial realizados no âmbito do Departamento de Polícia Federal. Em ambos os casos, foram presos suspeitos de autoria de crime, fotografados em várias posições, ângulos e distâncias e foram feitas comparações, utilizando técnicas de reconhecimento facial, com outras imagens questionadas, extraídas de sistemas de CFTV. Nota-se nitidamente, nos dois casos, a baixa resolução das imagens de CFTV e o posicionamento não frontal das imagens. A consequência é que os exames podem se tornar inconclusivos, contribuindo para a impunidade.

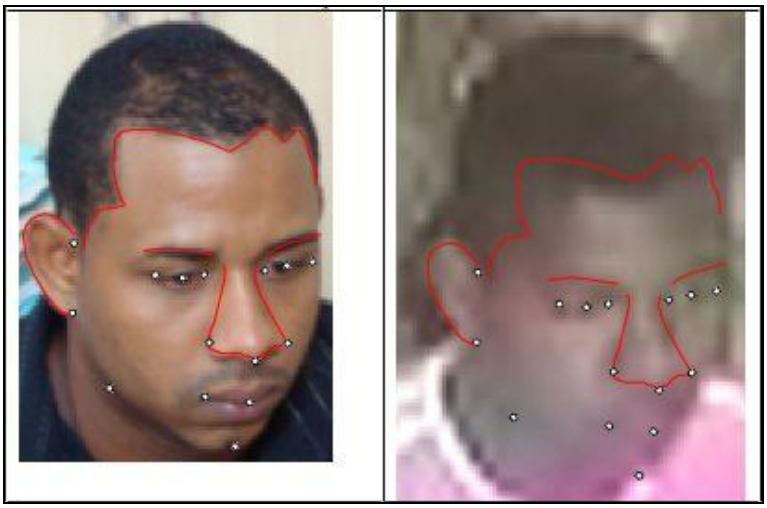


Figura 2.5 – Aplicação de técnicas de reconhecimento facial (pontos cefalométricos) em imagens de fotos e de CFTV (DPF_a, 2010).

Figura 2.6 – Comparação entre imagem de CFTV e fotos de suspeitos do crime. IPD de 6 pixels, 36 pixels e 36 pixels, respectivamente (DPF, 2011).

É importante ressaltar que experimentos importantes na área de reconhecimento facial, realizados por Bruce *et al.* (1999), Kemp, Towell e Pike (1997) e Yarmey (1979), utilizando rostos conhecidos e desconhecidos mostraram a dificuldade das pessoas em realizar o reconhecimento facial. Apesar de parecer contraditório, conseguimos identificar facilmente a face de pessoas conhecidas, como a nossa própria, a de familiares e amigos, bem como de pessoas famosas, mesmo se tratando de imagens de baixa resolução (BURTON *et al.*, 1999; SINHA *et al.*, 2006). Essa tarefa se torna árdua, entretanto, quando se trata de rostos de pessoas desconhecidas, casuística frequente na área forense. Bruce *et al.* (1999) realizou experimentos utilizando imagens de CFTV de boa qualidade, apresentando imagens de faces de pessoas desconhecidas a um grupo de pessoas, para que elas pudessem identificá-las em uma base de dados, conforme amostra apresentada na Figura 2.7. Os resultados mostraram que quando o rosto questionado estava presente no banco de dados, 30% das pessoas questionadas não o identificaram corretamente. No entanto, quando o rosto questionado era retirado do conjunto, 70% das pessoas ainda indicavam que o rosto estava presente no conjunto, corroborando com as afirmações do início do parágrafo.

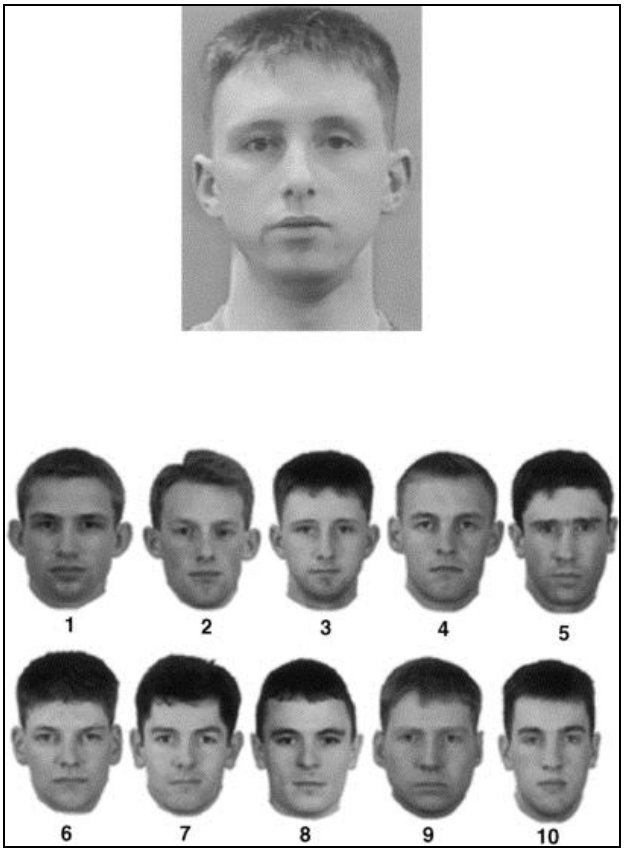


Figura 2.7 − Amostra do teste experimental (extraído de (BRUCE *et al.* (1999)).

Uma questão importante a ser ressaltada com relação aos experimentos realizados por Bruce *et al.* (1999) trata-se do uso de imagens de CFTV com alta qualidade e resolução espacial, discrepante com a realidade de muitos sistemas de CFTV existentes. Na maioria das situações encontradas, a qualidade das imagens capturadas e gravadas continua sendo baixa, inadequadas para fins de reconhecimento facial de suspeitos.

3 – SISTEMAS DE CIRCUITO FECHADO DE TELEVISÃO (CFTV)

Neste capítulo é apresentada uma breve apresentação sobre sistemas de CFTV, um dos tópicos sobre o qual está inserido o projeto de pesquisa deste trabalho. O problema de pesquisa abordado nesta dissertação tem um enfoque mais prático do que teórico, motivo pelo qual justificou a inserção deste capítulo no trabalho, ao invés de outro relacionado a fundamentação teórica, como é encontrado costumeiramente em dissertações de mestrado.

Buscou-se apresentar neste capítulo conceitos breves, mas importantes para o entendimento do trabalho. É relevante mencionar que não é objetivo desta dissertação, nem deste capítulo, esgotar o assunto sobre sistemas de CFTV, mas sim, apresentar os conceitos necessários que possibilitem uma leitura suave de todos os tópicos abordados nesta dissertação, bem como uma compreensão harmônica da proposta. A definição de outros conceitos citados neste capítulo e nos subsequentes foi omitida, pois extrapolava o objetivo de todo o trabalho. Caso o leitor tenha interesse nesses conceitos, deve buscar uma literatura especializada da área.

Desta forma, neste capítulo serão abordadas considerações iniciais sobre sistemas de CFTV (Seção 3.1), o princípio de funcionamento de um sistema de CFTV (Seção 3.2), serão apresentadas as principais tecnologias e componentes encontradas nestes sistemas (Seção 3.3) e por fim, serão discutidos aspectos sobre câmeras de segurança (Seção 3.4).

3.1 – CONSIDERAÇÕES INICIAIS SOBRE CFTV

Os sistemas de Circuito Fechado de Televisão (CFTV), terminologia derivada da expressão em língua inglesa *Closed Circuit Television* (CCTV), são basicamente sistemas de monitoramento, que distribuem sinais provenientes de câmeras de vigilância posicionadas em locais específicos, para um ponto central de supervisão (GOOLD, 2004).

A expressão "*circuito fechado*" refere-se ao fato que os sistemas de CFTV são restritos, no sentido de que os sinais transmitidos estão acessíveis apenas aos equipamentos contidos no sistema. A expressão serve de contraposição aos sistemas de TV aberta, na qual os sinais transmitidos pelas emissoras de TV são irradiados livremente e podem ser captados por qualquer um com equipamentos adequados de recepção (CIESZYNSKI, 2007).

O uso de equipamentos de gravação permite aos sistemas de CFTV gravar atividades para serem observadas posteriormente, auxiliando as equipes de segurança a recordar fatos que ocorreram e não foram detectados em tempo real. Além disso, as imagens e vídeos podem ser observados tanto no local de gravação, quanto remotamente, em sistemas mais modernos (POLICOM, 2009).

O sinal de vídeo distribuído pelas câmeras pode ser armazenado em fitas de vídeo analógicas ou como informações digitais. As câmeras de vigilância podem ser fixas ou móveis, abrangendo neste último caso, uma área maior de monitoramento (AIC, 2007). Ainda, as câmeras podem ter sua instalação aparente (*overt camera*) ou oculta (*covert camera*) dependendo da aplicação, sendo que as primeiras são utilizadas geralmente para coibir crimes e fornecer vigilância de zonas remotas, como áreas de estacionamentos, entrada e corredores de armazéns e áreas de produção. A exposição das câmeras de vigilância tem como premissa alertar quem quer que seja, funcionário ou cliente, de que determinado local está em constante monitoramento. Quando surge a necessidade, câmeras ocultas são utilizadas para detectar e observar atividades clandestinas. Enquanto as câmeras aparentes são grandes e de instalação fixa e permanente, as câmeras ocultas geralmente são pequenas, projetadas para serem escondidas em objetos no ambiente e utilizadas em curtos períodos de tempo, como algumas horas, dias ou semanas, para depois serem removidas do local de instalação. Normalmente as transmissões de sinais de vídeo ocorrem por meios guiados, através de cabeamento metálico, independente do tipo de câmera. Em algumas situações, como é o caso de algumas micro câmeras, normalmente utilizadas de forma oculta, as transmissões dos sinais de vídeo podem ocorrer através de tecnologia sem fio (KRUEGLE, 2007).

Os sistemas de CFTV foram desenvolvidos como parte de sistemas complexos de segurança, nos quais estão inseridos outros recursos, tais como muros, arames, cercas eletrificadas, vigilantes, cães de guarda, rádios comunicadores, alarmes, sensores de presença, sistemas de controle de acesso, dentre outros, com o intuito de auxiliar corporações industriais, comerciais e órgãos governamentais civis e militares a protegerem seus estabelecimentos contra ameaças internas e externas. A utilização de sistemas de CFTV, permitindo a vigilância remota, complementou e aperfeiçoou os serviços de segurança privada outrora realizados apenas por profissionais vigilantes (KRUEGLE, 2007).

Atualmente, os sistemas de CFTV são muito utilizados na vigilância de bens, pessoas e ambientes, como por exemplo, em ATM's, terminais de autoatendimento bancário (POLICOM, 2009). A utilização de sistemas de CFTV trouxe soluções efetivas para determinadas necessidades na área de segurança patrimonial, proporcionando monitorar múltiplos locais simultaneamente utilizando apenas um operador na central de supervisão, observar pessoas e veículos durante seu trajeto dentro de uma área monitorada, facilitando sua localização, para o caso de necessária interceptação, além de evitar rondas desnecessárias de vigilantes em áreas sem atividade (KRUEGLE, 2007).

Nos casos de combate ao crime, os sistemas de CFTV tem importante função preventiva, de forma que a ação criminosa pode ser inibida, quando é observado o uso de câmeras de vigilância aparentes no local, aumentando as chances do ofensor ser pego. Além da função preventiva, as evidências trazidas pelos registros gravados são importantes armas para a polícia durante a investigação de um crime, corroborando ou desqualificado a oitiva de suspeitos, ajudando a identificar testemunhas presentes no local de crime, além de auxiliar na possível identificação dos criminosos. Neste último cenário, a baixa qualidade das imagens gravadas em sistemas de CFTV representa um grave problema para as autoridades policiais e para o judiciário, no que tange à adequada identificação de um criminoso (RATCLIFFE, 2006).

Uma importante classificação de sistemas de CFTV refere-se à divisão do monitoramento em aplicações internas (*indoor*) e externas (*outdoor*). Esta divisão estabelece uma fronteira natural entre os tipos de equipamentos a serem utilizados em cada aplicação: aqueles adequados às condições controladas dos ambientes fechados e aqueles adequados às condições mais ríspidas de ambientes ao ar livre. Dois parâmetros permitem qualificar bem essa divisão com relação aos equipamentos: fatores ambientes e iluminação. Aplicações *indoor* requerem iluminação artificial, que podem ou não ser complementada pela iluminação natural do dia. Além disso, nestas aplicações, os principais problemas relacionados às condições de ambiente restringem-se a variações de umidade, poeira, sujeiras, fumaça, além de pequenas variações de temperatura. Em aplicações *outdoor*, diferentemente do anterior, os equipamentos devem suportar condições de temperaturas extremas, precipitações ambientais (nevoeiro, chuva e neve), vento, sujeira, poeira, areia, sal, fumaça, dentre outras condições encontradas ao ar livre. No quesito iluminação, os

equipamentos usufruem da iluminação natural durante o dia, podendo utilizar iluminação artificial, principalmente no período noturno (KRUEGLE, 2007).

Apesar de possuírem forte apelo para a área de segurança, os sistemas de CFTV não estão restritos somente a esse tipo de aplicação, tendo sido utilizados também no controle de tráfego, de multidões e de pacientes, em laboratórios de pesquisa, treinamentos, controle de produção em indústrias, avaliações de desempenho profissional, gerenciamento de informações, dentre outras inúmeras aplicações (CIESZYNSKI, 2007).

3.2 – FUNCIONAMENTO DE UM SISTEMA DE CFTV

Um sistema CFTV, basicamente, deve ser capaz de realizar cinco tarefas: capturar uma imagem visual de um ambiente, converter o sinal capturado para o formato de vídeo, transmitir o sinal para um receptor remoto, mostrar as imagens em um equipamento monitor e, gravar e imprimir para controle permanente. Adicionalmente, em sistemas mais complexos, uma nova tarefa também deve ser contemplada por esses sistemas, que é o processamento do vídeo dos sinais capturados (KRUEGLE, 2007). O diagrama de blocos a seguir apresenta o esquema de funcionamento de um sistema CFTV.

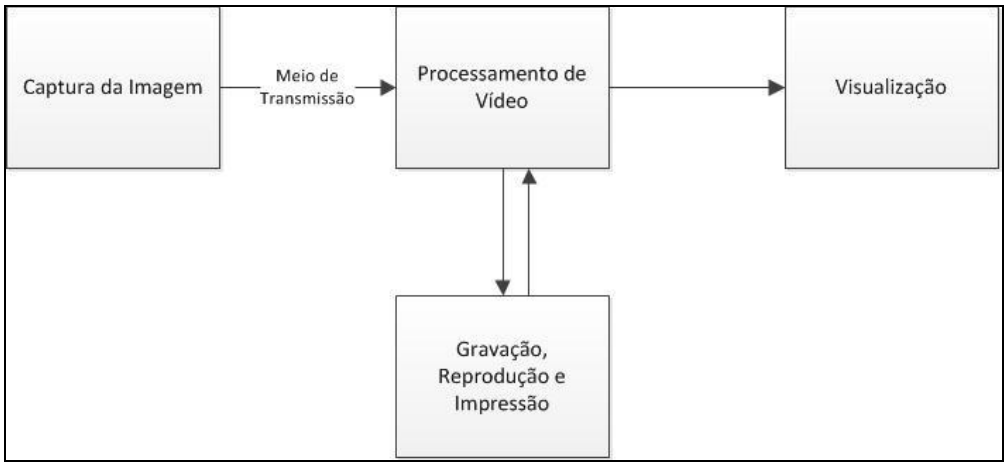


Figura 3.1 – Diagrama de blocos de funcionamento de um sistema de CFTV.

O bloco de captação de imagem é formado pelo conjunto lente e câmera e são os responsáveis pelas duas primeiras tarefas de um sistema CFTV. A luz da fonte de iluminação, seja natural ou artificial, reflete os objetos presentes na cena observada. As lentes do sistema, então, coletam a luz da cena visível, formando a imagem nos sensores da

câmera. Os sensores da câmera, por sua vez, convertem a imagem formada pelas lentes em sinais elétricos adequados para serem transmitidos aos demais equipamentos do sistema, tais como os monitores remotos, gravadores e impressoras (KRUEGLE, 2007).

As mídias ou meios de transmissão estão presentes em todo o sistema CFTV, conectando devidamente cada equipamento formador do sistema. São essas mídias que levam os sinais elétricos gerados na câmera para o restante do sistema. Dependendo da mídia utilizada para transmissão, os sinais podem ser analógicos ou digitais. Os tipos mais comuns de mídias de transmissão podem ser divididos em guiados, abrangendo o cabo coaxial, o cabo par trançado (UTP), o cabo de fibra ótica; e não guiados, abrangendo as transmissões por rádio frequência, transmissão por micro-ondas e as transmissões por infravermelho (IR), (KRUEGLE, 2007).

Em sistemas de CFTV simples, na qual há apenas o uso de uma câmera, a etapa de processamento de vídeo não é uma real necessidade. Quando o cenário muda e há o uso de múltiplas câmeras, equipamentos adicionais são necessários no sistema para realizarem as tarefas de selecionar diferentes câmeras para serem apresentadas em vídeo, além de permitirem que múltiplas imagens, advindas das câmeras, sejam apresentadas simultaneamente em um único ou em múltiplos monitores (KRUEGLE, 2007).

A próxima etapa do sistema refere-se à apresentação das imagens capturadas em um equipamento monitor. Esses equipamentos convertem o sinal elétrico recebido de volta às imagens visíveis capturadas pelas câmeras. Tipicamente são utilizados monitores analógicos (CRT) ou sistemas computacionais com telas de LCD, LED ou plasma (KRUEGLE, 2007).

A etapa final de um sistema genérico de CFTV abrange as tarefas de gravação e impressão. Na gravação, as imagens são permanentemente gravadas em um equipamento específico para esse fim. Os sistemas de CFTV utilizaram por muito tempo gravação em figuras VHS feitas por equipamentos *time-lapses*. Atualmente, os equipamentos mais modernos utilizam disco rígido magnético ou disco rígido SSD (estado sólido) para gravação, além de permitirem backup de cenas gravadas em mídias externas CD e DVD. Na tarefa de impressão, cenas específicas gravadas podem ser impressas, utilizando equipamentos de impressão com tecnologia a jato de tinta, a laser, dentre outras (KRUEGLE, 2007).

3.3 – TECNOLOGIAS E COMPONENTES BÁSICOS DE UM SISTEMA CFTV

A segurança de um ambiente pode abranger o uso de sistemas de CFTV simples ou complexos, de tecnologia exclusivamente analógica, exclusivamente digital ou híbrida, utilizando uma combinação das tecnologias analógica e digital. Os primeiros sistemas de CFTV foram implantados durante a década de 1960, utilizando equipamentos analógicos, tecnologia esta que vigorou até o início dos anos 90. A partir desta época, um período de transição foi iniciado, com o amadurecimento do legado analógico para a tecnologia digital, inclusive tornando os equipamentos CFTV compatíveis com computadores e as tecnologias de rede vigentes (KRUEGLE, 2007). Atualmente, os sistemas de CFTV comumente encontrados no mercado abrangem tecnologias híbridas, dependendo da necessidade e da aplicação ao qual é destinado o sistema.

A seguir são apresentados os equipamentos tipicamente encontrados em soluções de CFTV convencionais (sistemas analógicos), (DAMJANOVSKI, 2005):

- ***Câmeras Analógicas***: tipo de câmeras mais frequentes em sistemas de CFTV. Podem apresentar sensores CCD ou CMOS, modo de cores monocromático ou colorido.

- ***Sequenciais de Vídeo (video sequencial switchers):*** é o dispositivo eletrônico destinado a combinar o sinal de múltiplas câmeras, apresentando uma de cada vez na tela do monitor. O equipamento sequencial de vídeo é a forma mais econômica de visualização de câmeras de uma solução de CFTV, possibilitando que de um mesmo monitor ou TV, todas as câmeras sejam visualizadas. Utilizam tipicamente conectores BNC e possuem, normalmente, entradas para 4 ou 8 câmeras.

- ***Quad Splitter ou Quad compressor***: equipamento eletrônico usado quando múltiplas câmeras necessitam ser apresentadas simultaneamente em um único monitor. O equipamento é interposto entre as câmeras e os monitores para permitir a visualização simultânea de imagens provenientes de quatro câmeras, dividindo a imagem apresentada no monitor em quatro quadrantes iguais. Há equipamentos *Duo-Quad*, oriundos da união de um *quad splitter* e um sequencial, sendo possível o sequenciamento de 8 câmeras, 4 x 4. Foram fabricados também duplicadores ou

quadruplicadores *quad* que expandiram o número simultâneo de câmeras para 8 ou 16.

- ***Multiplexadores (multiplexers)***: equipamento eletrônico que permite a divisão da tela do monitor ou televisão envolvida na solução de circuito fechado de televisão em até dezesseis quadros, possibilitando que de um mesmo monitor ou TV, até dezesseis câmeras sejam visualizadas simultaneamente. Além da função *multi-screen*, permitem a gravação de multi-câmeras simultaneamente em um mesmo *time-lapse*. Ainda, as imagens são gravadas em quadro completo, permitindo que apenas um sinal de câmera seja apresentado em tela cheia, como se fosse o único gravado. Durante muitos anos, estes equipamentos foram a melhor alternativa em termos de equipamentos para gerenciamento avançado de CFTV. Podem ser encontrados multiplexadores simplex ou duplex, de acordo com a capacidade do multiplexador em realizar as tarefas de gravação e reprodução, simultaneamente.

- ***Matriz de Vídeo (vídeo matrix switchers – VMSs)***: equipamento eletrônico para controlar grandes soluções de sistemas de CFTV, ou seja, quando o número de monitores e câmeras a serem controladas extrapolam aquelas quantidades normalmente encontradas em equipamentos menos sofisticados de processamento de vídeo. Oferece, geralmente, funções avançadas de monitoramento e distribuição de vídeo.

- ***Time-Lapses***: equipamentos de gravação em fitas de vídeo, similar aos videocassetes, entretanto, programados para gravações de longa duração. Permitem gravação entre 24 horas e 960 horas em uma fita VHS normal de 120 minutos, modificando a taxa de quadros gravados por segundo. Em equipamentos videocassetes, essa taxa é de 30 quadros/s.

- ***Impressoras de vídeo (vídeo printers)***: equipamento encontrado normalmente em grandes sistemas de CFTV no qual é necessária uma cópia física de imagens gravadas para utilização como evidência. Geralmente encontrado nos formatos monocromático ou colorido.

- ***Monitores de vídeo***: normalmente de tecnologia CRT (tecnologia de tubos de raios catódicos). São fabricados para funcionamento 24/7 (24 horas 7 dias na semana), diferentemente de equipamentos televisores, que possuem uso diário restrito. Possuem diferentes tamanhos e podem ser monocromáticos ou coloridos.

- ***Cabeamento***: as soluções tipicamente analógicas utilizam principalmente cabeamento coaxial interligando os vários equipamentos componentes do sistema de CFTV. O cabo coaxial mais utilizado no mercado é o RG59/U, possuindo impedância de 75 Ω, podendo alcançar poucas centenas de metros de transmissão de vídeo, sem perder qualidade. Normalmente, cabeamento coaxial utiliza conectores BNC.

Com a evolução da tecnologia, novos equipamentos passaram a complementar as tarefas básicas dos sistemas de CFTV, permitindo maior flexibilidade, rapidez e facilidade de instalação, uso e administração desses sistemas. Segundo Portela (2007), os sistemas digitais diminuem os custos de operação de CFTV, além de proporcionarem melhor custo benefício. Isto é traduzido pela maior qualidade de gravação de imagens, acesso imediato às imagens ao vivo e gravadas, as formas de armazenamento são mais simples, além dos sistemas digitais proporcionarem maior autonomia de gravação.

Em sistemas tipicamente digitais, puderam ser observadas algumas mudanças com relação aos sistemas de CFTV tradicionais (PORTELA, 2007):

- ***Câmeras***: o uso de câmeras digitais trouxe maior qualidade para a captação de imagens. Com o uso destes equipamentos, o sinal analógico convertido pelo sensor da câmera é processado de forma digital, ou seja, é convertido para digital, analisado, comparado e amplificado digitalmente.

- ***Meios de transmissão***: os meios de transmissão analógicos perduraram em sistemas de CFTV por muito tempo, mesmo durante a migração de muitos equipamentos analógicos para seus equivalentes digitais. Entretanto, o uso de cabos par trançado UTP e de fibra ótica, que possibilitam a transmissão de sinais de vídeo de forma digital trouxeram novos limites de distância e qualidade aos sistemas de CFTV.

- ***Processamento de Vídeo***: o uso de multiplexadores digitais, com mais recursos impulsionou fortes mudanças em outros tipos de equipamentos utilizados em sistemas de CFTV, como naqueles de gravação de dados, que acabaram incorporando as funções de multiplexadores, sequencias e dos time-lapses. Os equipamentos de processamento de vídeo e de gravação, com a evolução dos equipamentos de CFTV acabaram sendo mesclados.

- ***Gravação***: a necessidade de maior autonomia dos sistemas de gravação e a disponibilidade de mais recursos nesta etapa do processo acabou impulsionando a criação de gravadores digitais de vídeo (DVRs – *Digital Video Recorders*). Os equipamentos de processamento de vídeo foram incorporados aos equipamentos de gravação e s fitas VHS, anteriormente utilizadas em conjunto com os equipamentos *time-lapses*, foram substituídas por discos rígidos (*hard disks*). A gravação por detecção de movimento, ferramentas de backups em mídias óticas (CD's e DVD's), a regravação automática do disco rígido, a facilidade de operação desses sistemas, maior autonomia de gravação (maior capacidade e tempo de gravação disponível), maior resolução suportada, além de permitir acesso remoto, foram algumas das transformações proporcionadas pela migração para equipamentos digitais de gravação.

- ***Visualização e controle***: os grandes monitores analógicos monocromáticos foram substituídos por telas de LCD e plasma coloridos. A integração com sistemas computacionais automatizou as funções básicas do sistema e facilitou sua operação.

Outra notória mudança que tem ocorrido nos sistemas de CFTV atuais é a sua integração com redes locais computacionais ou mesmo à Internet (vide Figura 3.4), agregando o acesso remoto como forma de controle do sistema de CFTV, permitindo assim, uma supervisão em níveis cada vez maiores e mais completos. Novos equipamentos também foram desenvolvidos para trabalhar nesta nova topologia, como câmeras IP, equipamentos hub's ou switches de rede, além de servidores de vídeo (*vídeo web servers*) que passaram a controlar o processamento de vídeo e a gravação dos dados.

3.4 – CÂMERAS DE CFTV

A função primária de todo sistema de vigilância é prover "olhos remotos" para a equipe de segurança localizada na central de controle local ou remota, função desempenhada pelo equipamento de câmera de CFTV. Elas realizam a captação do ambiente, convertendo os níveis de iluminação e cor da cena, através de sensores especiais, em sinais elétricos que vão ser distribuídos para os demais equipamentos do sistema. Nos tópicos a seguir são apresentados alguns detalhes relacionados à configuração de câmeras de CFTV.

3.4.1 – ESPECIFICAÇÕES TÉCNICAS DE CÂMERAS DE CFTV

A seguir são apresentadas algumas das características normalmente encontradas em câmeras de sistemas de CFTV, nas tecnologias analógica ou digital.

3.4.1.1 – Sensor de Imagem

É o dispositivo eletrônico encontrado no interior da câmera contendo elementos sensíveis às variações de iluminação do ambiente (luz). É este dispositivo que tem a função de converter a imagem visual da cena observada pela lente da câmera em sinais elétricos. Atualmente, há dois tipos principais de sensores de imagem encontrados em câmeras de CFTV: CMOS (*Complementary Metal Oxide Semiconductor*) e CCD (*Charged Coupled Device*). Os sensores CMOS normalmente produzem imagens com qualidade inferior àquelas produzidas por um CCD do mesmo nível e, por isto, geralmente são encontrados em dispositivos de micro câmeras de baixo custo e *webcams* para usuários domésticos. Em CFTV, grande parte das câmeras utilizam sensores CCD. São comuns os formatos de área do CCD de ½'', ⅓'' e ¼'' de polegada (na diagonal). Os CCD's de ¼" são os mais modernos atualmente. O tamanho do sensor tem influência direta na distância focal das lentes das câmeras (SWGIT, 2004).

3.4.1.2 – Resolução

O termo resolução possui tantos conceitos e nenhum é suficiente para definir corretamente todas as situações (Toshiba, 2011). Em se tratando de aplicações CFTV isso não é diferente, a terminologia resolução também confunde bastante seus usuários. De forma

geral, resolução pode ser definida como a clareza de detalhes em que uma imagem pode ser distinguida. Assim, a resolução governa a qualidade global de uma imagem de uma câmera de vídeo.

Há diferentes métodos e medidas para descrever a resolução de imagem de uma câmera (PELCO, 2011). A resolução espacial é um deles, sendo exatamente a razão entre o número de pixels obtidos na imagem e o tamanho da imagem real (SCURI, 2002). As unidades de medida mais comuns para resolução espacial são os pares de linhas por unidade de medida, número total de pixels ou os pixels por unidade de medida (GONZALEZ e WOODS, 2010). Quando a resolução espacial é expressa em número de pares de linhas de TV, não se tem definido o número de pixels da imagem, mas sim o número de linhas discerníveis horizontalmente ou verticalmente da tela (PELCO, 2011). Esta é a forma clássica de expressar a resolução de um monitor e é definida como o número máximo de linhas pretas e brancas discerníveis na tela (BOSCH, 2006).

Como já mencionado, é popularmente comum definir a resolução da imagem através da medida do número total de pixels da imagem. O termo pixel é derivado da expressão inglesa *PICture ELement* e são os menores detalhes discerníveis em uma imagem. Eles são reunidos horizontalmente e verticalmente e a resolução pode ser expressa ao multiplicar o número de pixels encontrados nas linhas pelo das colunas. Por exemplo, uma imagem que possui 2.048 pixels de largura e 1.536 pixels de altura tem um total de 2.048 X 1.536 = 3.145.728 pixels ou 3.1 megapixels (PELCO, 2011).

Há outras medidas também encontradas no mercado que são utilizadas para descrever a resolução de uma imagem, como o número de pixels por unidade de comprimento (por exemplo, pixels por polegada) ou o número de pixels por área (por exemplo, pixels por polegada quadrada) (PELCO, 2011).

Câmeras de CFTV analógicas de boa qualidade possuem tipicamente resoluções horizontais de 330 TVL (*Television lines*) até 470 TVL, no caso de câmeras coloridas. Câmeras coloridas digitais de alta resolução possuem normalmente entre 450 e 480 TVL. Comparativamente, as micro câmeras apresentam normalmente resolução média de 300 a 330 TVL (PORTELA, 2007). As tabelas apresentadsa no Apêndice D relacionam os principais tipos de resoluções de câmeras encontradas no mercado e o Apêndice E

apresenta recomendações de câmeras do SWGIT (*Scientific Working Group on Imaging Technology*) para sistemas de CFTV com propósitos de reconhecimento facial.

3.4.1.3 – Iluminação da Câmera

Característica intrínseca de cada câmera e refere-se às condições de trabalho da câmera em relação à iluminação ambiente. É medida em lux e esse parâmetro irá depender de cada modelo e fabricante de câmeras. Normalmente possuem valores entre 0,01 e 0,5 lux para câmeras monocromáticas e 0,7 a 3 lux para câmeras coloridas. Em condições de baixa iluminação no ambiente, quanto menor o parâmetro de lux da câmera, melhor será a imagem registrada. Esse parâmetro é influenciado pela abertura de lente escolhida e o mínimo de luz incidente sobre o sensor de imagem da câmera (SWGIT, 2004).

3.4.1.4 – AGC (Automatic Gain Control)

O controle automático de ganho é um recurso efetuado por circuitos da câmera que permite um ajuste automático do nível de sinal do vídeo, ou seja, o intuito é fazer com que esse permaneça de forma aproximadamente constante independente das variações do ambiente registrado pela câmera. A utilização de AGC ou ganho não é desejável, é melhor que o ambiente forneça luz suficiente para a captação de imagens com o mínimo de qualidade. O uso de ganho eletrônico em cenas com pouca iluminação natural resulta em imagens de baixa qualidade (SWGIT, 2004).

3.4.1.5 – Electronic Shutter (Obturador Eletrônico)

Este recurso, também denominado íris eletrônica, *shutter speed*, EI, ES, AEI, determina a velocidade de leitura dos pixels em uma imagem, podendo ser ajustado para compensar a variação de iluminação em um ambiente. O obturador eletrônico é um controle do tempo de exposição de cada quadro de vídeo. Enquanto a íris aumenta ou diminui a abertura da lente, regulando assim a quantidade de luz que entra na câmera, o obturador eletrônico funciona como uma camada (hélice) à frente da lente que bloqueia parcialmente a entrada da luz durante parte do ciclo de captação. A principal finalidade deste recurso é a de reduzir o *motion blur* (borrado de movimento) nas cenas de objetos com movimento rápido (SWGIT, 2004).

3.4.1.6 – BLC (Back light compensation)

A compensação de luz de fundo é um recurso que proporciona às câmeras um modo de compensar a situação de iluminação ambiente, sejam situações com iluminação muito intensa em um ambiente com imagens saturadas ou áreas com sombreamento indesejáveis. Pode ser analógico ou digital. O BLC é um recurso desejável em câmeras de CFTV, pois funcionam como um atenuador da iluminação do ambiente, melhorando a definição da imagem captada (PORTELA, 2007).

3.4.1.7 – ATW (Automatic Tracing White Balance)

O recurso de balanço automático do nível de branco visa ajustar automaticamente os pontos de cores de um ambiente em relação aos diferentes pontos brancos da imagem. Busca-se evitar o brilho ou a reflexão excessiva dos pontos claros da imagem. O ATW busca corrigir automaticamente o balanço de branco (SWGIT, 2004).

3.4.2 − TIPOS DE CÂMERAS

Há vários tipos de câmeras disponíveis no mercado e que podem ser utilizadas em sistemas de CFTV (PORTELA, 2007):

- **Micro câmeras**: câmeras de pequeno porte e baixo custo. Podem ser coloridas ou monocromáticas. Não apresentam muitos recursos disponíveis e geralmente apresentam baixa resolução espacial. As lentes neste tipo de câmera são fixas. Há um tipo especial de micro câmera denominado *pin hole*. Este tipo de câmera geralmente é utilizado quando se deseja ocultar o dispositivo de terceiros e captar imagens de orifícios pequenos.

- **Mini câmeras**: apresentam mais recursos que as micro câmeras, além de permitirem encaixe de lentes. Possuem custo intermediário entre as micro câmeras e as câmeras profissionais.

- **Câmeras profissionais**: possuem recursos avançados e geralmente fornecem melhor qualidade de imagem (melhores opções de resolução espacial). Um tipo especial de câmera profissional são as câmeras *speed dome*, que permitem movimentar as câmeras (giros horizontais de 360º e giro vertical de 90º) de acordo com a cena desejada a ser observada.

3.4.3 – LENTES

As lentes em uma câmera têm por função direcionar a luz refletida do ambiente registrado para sensor da câmera. Os tipos de lentes encontradas no mercado podem ser divididos em (SWGIT, 2004):

- **Íris fixa**: é o mais simples tipo de lente, possuindo somente ajuste de foco. Pode ser utilizada para ambientes em que a iluminação não se altere de forma excessiva. Normalmente utilizada em micro câmeras. Não é recomendável para sistemas de CFTV, cujo fim é a segurança orgânica.

- **Íris manual**: permite ajuste da luz refletida, permitindo direcionar a quantidade ideal de iluminação a ser captada pelo sensor de imagem. Pode ser utilizada para ambientes com iluminações muito intensas ou pouco intensas.

- **Auto íris**: permitem o ajuste automático da íris da lente de acordo com o nível de iluminação do ambiente. Possuem motores e sistemas de verificação que definem quando a íris deve ser aberta ou fechada.
- **Varifocais**: são lentes que permitem o ajuste de sua distância focal[6] (zoom manual). Podem apresentar íris manual ou automática. Recomendável para sistemas de CFTV, pela flexibilidade de ajustes para o dimensionamento do projeto.

[6] Normalmente medido em mm (milímetros), este parâmetro está diretamente relacionado ao ângulo de visão que será obtido da cena. Distâncias focais longas captam ângulos de visão da cena menores e distâncias focais curtas captam ângulos de visão da cena maiores (abertos). Os valores mais comuns são 4, 6, 8, 12, 16, 25, 35, 50 e 75 mm.

- **Lentes zoom**: possuem funções de zoom, foco e/ou auto íris, que são controlados por um operador remoto.

Outras características encontradas em lentes de câmeras para CFTV diz respeitos ao padrão de montagem da câmera. Os tipos mais comuns encontrados no mercado de CFTV são (PORTELA, 2007):

- **C-Mount**: montagem por rosca, apresentam 1” de diâmetro com 32 TPI (threads per inch - linhas por polegadas) e 17,5 mm de distância para a flange posterior da câmera. São os tipos de montagem mais comuns.

- **CS-Mount**: montagem por rosca, apresentam 1” de diâmetro com 32 TPI (threads per inch - linhas por polegadas) e 12,5 mm de distância para a flange posterior da câmera. São os tipos de montagem mais comuns.

- **Fixo**: não permitem a retirada da lente da câmera.

Por fim, um parâmetro que deve ser destacado com relação às configurações de lentes é a velocidade ótica. Ele indica a quantidade de luz que passa através da lente e é definido pelo índice *f-number*, calculado pela razão entre a distância focal e a abertura efetiva da lente. Lentes normalmente são caracterizadas neste parâmetro pelos valores de máxima e mínima abertura.

No apêndice B são apresentadas tabelas para escolha da lente adequada para determinada aplicação, relacionando campo de visão e sensor de imagem.

4 – RECONHECIMENTO FACIAL

Neste capítulo, serão discutidos sobre sistemas de reconhecimento facial. O conteúdo do capítulo abrange algumas discussões conceituais sobre reconhecimento facial, mas apresenta principalmente, os trabalhos práticos divulgados nesta área de conhecimento. Na abordagem dos trabalhos práticos, são apresentados em conjunto, os principais problemas encontrados no desenvolvimento de sistemas de reconhecimento facial.

4.1 – CONSIDERAÇÕES INICIAIS

Sistemas de reconhecimento facial registram características distintivas da face para desempenhar sua função. Há vários métodos disponíveis desenvolvidos para serem utilizados nesses sistemas, no entanto, todos envolvem o uso de medidas de características do rosto de um indivíduo.

Embora o conceito de reconhecer alguém através de sua face, como normalmente fazemos com amigos e familiares, seja intuitivo, o reconhecimento facial, em termos biométricos, trata-se de um processo computadorizado e automatizado que procura simular o mesmo processo realizado rotineiramente pelas pessoas em seu dia-a-dia. As pesquisas apresentadas na área mostraram que este é um problema difícil de ser resolvido. A utilidade do reconhecimento facial, frente às demais formas de sistemas biométricos utilizados para identificação de indivíduos, está na sua capacidade de uso para fins de vigilância, ou seja, podem ser utilizadas para passivamente localizar e reconhecer indivíduos que são monitorados pelas câmeras do sistema.

O uso do reconhecimento facial como técnica de identificação da identidade de uma pessoa oferece algumas vantagens: não exige contato físico com o indivíduo (técnica passiva), pois os sistemas de câmeras capturam imagens dos indivíduos circulando em seu campo de visão a partir de certa distância e podem ser dissimulados, evitando que as pessoas saibam que estão sendo monitoradas (WOODWARD *et. al*, 2003).

Para qualquer sistema biométrico poder operar, ele deve ter registros em um banco de dados, que serão comparados com aqueles informados como entrada do sistema. Os

sistemas de reconhecimento facial podem ser utilizados de muitas formas, por exemplo, recebendo imagens de suspeitos presos, feitas em alta resolução, para serem analisadas e comparadas no banco de dados disponível, ou mesmo utilizando imagens de baixa resolução, coletadas de sistemas de CFTV.

Sistemas de reconhecimento facial são sistemas biométricos utilizados para verificar a identidade de uma pessoa. Independentemente do método de reconhecimento utilizado, ele pode ser dividido em cinco etapas (ROSA, 2012):

- Em primeiro lugar, uma imagem da face é recebida pelo sistema. Esta aquisição pode ser realizada por digitalização de uma foto física ou utilizando imagens extraídas de câmeras de CFTV, por exemplo. Arquivos de vídeo podem ser utilizados como fonte de imagens para o sistema, pois tratam-se de sequências cadenciadas de imagens estáticas.

- Em segundo lugar, o software de reconhecimento facial é empregado para detectar a localização de quaisquer faces na imagem recebida. Esta não é uma tarefa fácil, os padrões utilizados no mercado procuram generalizar um padrão do que se parece uma face de um indivíduo (dois olhos e uma boca definida em uma forma oval), para escolher as áreas de interesse na imagem. Em alguns sistemas, são realizados pré-processamentos nesta imagem da face detectada com intuito de melhoria da performance do reconhecimento facial.

- Uma vez que o software de detecção facial conseguiu localizar uma face na imagem recebida, ela pode ser então analisada em relação à geometria espacial das características distintivas do rosto. Um dos métodos mais populares de análise de faces é o PCA (*Principle Component Analysis*) (KIRBY e SIROVICH, 1990), que é comumente referido como o método de *eigenfaces*. O método PCA pode ser combinado com o uso de redes neurais e análise de características da face (*local feature analysis*) como forma de melhorar seu desempenho. A geração do modelo (*template generation*) é o resultado desta etapa de extração de características da face. O modelo é reduzido a um conjunto reduzido de dados que representam unicamente a identidade de um indivíduo. É importante notar que nestes sistemas

não são usados detalhes externos ao rosto de um indivíduo para classificá-lo, como por exemplo, o cabelo, forma de penteado, dentre outros (ROSA, 2012).

- O quarto passo compara o modelo gerado para a imagem de entrada, com aquelas presentes em um banco de dados. Em um software de reconhecimento facial, este processo gera pontuações que indicam o quão perto o modelo gerado corresponde àqueles encontrados no banco de dados.

- O passo final procura determinar se as pontuações produzidas na imagem da face questionada são suficientes para declarar a identidade de uma pessoa dentre aquelas presentes no banco de dados.

O diagrama a seguir, apresentado na Figura 1.3, ilustra o processo apresentado.

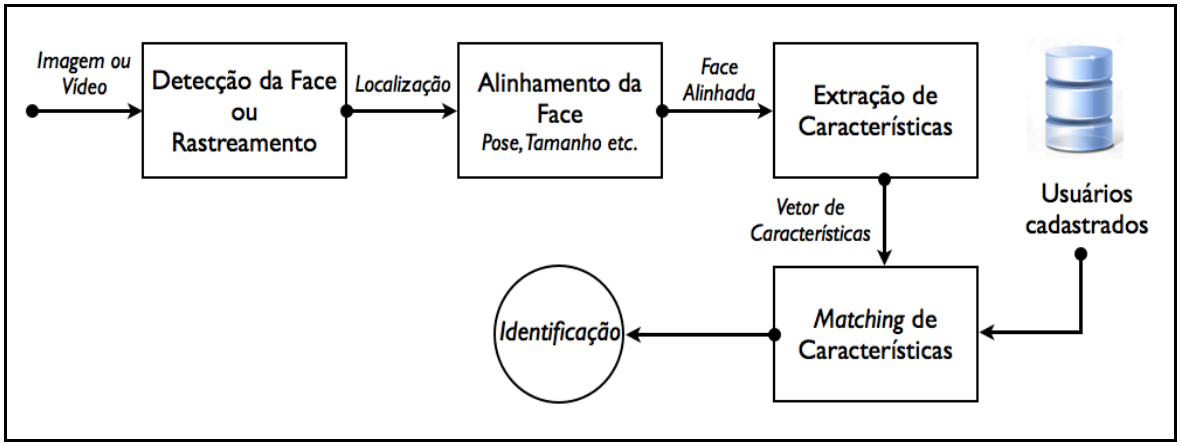


Figura 4.1− Diagrama do processo geral de reconhecimento facial automatizado (extraído de (MACHADO et al., 2010) citando (LI e JAIN, 2005; ZHAO *et al.*, 2003)).

É importante ressaltar que o desempenho de sistemas de reconhecimento facial sofre influência do posicionamento da face capturada (ângulo e altura da câmera), da iluminação ambiente, da resolução da imagem, de forma que este trabalho procurará tratar estes critérios de forma a auxiliar o dimensionamento de projetos de CFTV, para fins de reconhecimentos de suspeitos.

4.2 – DETECÇÃO FACIAL

No desenvolvimento dos sistemas de reconhecimento facial automatizados, os cientistas se depararam, inicialmente, com a problemática em realizar a detecção de faces em imagens estáticas (YANG e HUANG, 1993; YANG, KRIEGMAN e AHUJA, 2002). A sistematização dessa fase consiste em localizar uma ou mais faces dentro do cenário complexo de uma imagem, extrair a área específica das faces identificadas, eliminar os demais elementos da imagem e apresentar ao sistema uma janela contendo apenas as faces extraídas (QUINTILIANO e ROSA, 2006). As abordagens mais conhecidas aplicadas a este problema são baseadas em: cores (*color based*), moldes (*template based*) e características faciais (*feature based*), (PARK, 2009).

Na abordagem baseada em cores (*color based*), a cor da pele humana é utilizada para separar a face de outros objetos presentes na imagem. Um trabalho relevante utilizando este método foi desenvolvido por Bhuiyan *et al.* (2003). No método de Bhuiyan *et al.* (2003), as cores relevantes e dominantes da pele são extraídas da imagem no modelo de cor RGB. A imagem é, então, transformada para o espaço de cores YIQ (*Luminance (Y), In-phase (I) e Quadrature (Q) color space)*. Baseado no fato de que a cor da pele tende a se aglomerar em uma região do espaço de cores, o método utiliza um parâmetro limítrofe para detectar pixels de pele (*skin pixels*) na imagem transformada. Para localizar a face, toma-se então, uma imagem com região maior conectada de pixels de pele. Outros trabalhos importantes utilizando a abordagem de cores foram desenvolvidos por Sobottka e Pittas (1996), Saber e Tekalp (1998), Yang e Ahuja (1998), Cai e Goshtasby (1999), Feris *et al.* (2000), Hsu *et al.* (2002) e Jin *et al.* (2007).

Na abordagem baseada em moldes (*template based*), o objeto é procurado em uma imagem, utilizando um molde padrão que pode ser definido manualmente ou parametrizado por uma função. É comum o uso de componentes geométricos básicos, como círculos, quadrados ou triângulos para descrever o molde de um objeto. A detecção do objeto consiste então, em localizar a melhor simetria entre o objeto presente na imagem e o seu molde (*template*), (LOPES, 2005; YANG, KRIEGMAN e AHUJA, 2002). Utilizando esta abordagem, merece destaque o trabalho desenvolvido por Craw, Tock e Bennett (1992). O método foi aplicado somente a imagens em escala de cinza e consistia de duas etapas. Na primeira, o sistema era treinado utilizando análise estatística aplicado

em vários exemplos de imagens, para extrair as bordas e contrastes relacionados a contornos e características da face, tais como olhos, nariz, boca e silhueta do rosto. Em seguida, ao analisar uma imagem real sobre confronto, o sistema estimava regiões candidatas a conter uma face e fazia previsões sobre onde encontrar os demais moldes treinados pelo sistema. Quando estas previsões incidiam sobre regiões reais de face na imagem, casando a máscara modelada pelo sistema com uma face real na imagem, a região selecionada apresentava uma identificação positiva de uma face (CRAW, TOCK e BENNET, 1992).

O estado da arte em métodos para detectar faces em imagem é baseado em características faciais (*feature based*). Nesta categoria, os métodos utilizam classificadores binários para a detecção de faces, separando os elementos de uma imagem que correspondem ou não a uma face. Foram desenvolvidos muitos algoritmos utilizando esta abordagem, dentre eles, relacionam-se os trabalhos de Rowley, Baluja e Kanade (1998), Heisele *et al.* (2001), Levi e Weiss (2002), Viola e Jones (2004) e Wu *et al.* (2008). O trabalho de Viola e Jones (2004) está entre os mais citados na literatura, utilizando esta abordagem, por tratar imagens em tempo real, pela alta precisão na separação de faces da imagem, superando os métodos desenvolvidos anteriormente, bem como pelo projeto estar disponível para uso na biblioteca de programação OpenCV (*Open Computer Vision Library*) (OpenCV, 2012). O método introduziu um novo arcabouço de trabalho (*framework*) em detecção facial, baseado em três contribuições. A primeira relaciona-se à uma nova representação da imagem, denominado imagem integral (*integral image*) que permite que as características utilizadas pelo algoritmo detector sejam processadas rapidamente. A segunda contribuição refere-se ao uso de um novo método classificador, baseado no algoritmo de aprendizagem *AdaBoost* (Freund e Schapire, 1995), que seleciona um pequeno número de características visuais críticas de um grande conjunto de potenciais características. A última contribuição do método consiste em combinar classificador em cascata, permitindo descartar regiões de fundo da imagem de forma rápida, focando a varredura de faces em pequenas áreas de interesse.

4.3 – RECONHECIMENTO FACIAL

As principais abordagens em sistemas de reconhecimento facial utilizando imagens 2D são baseadas em características da face (*feature based*) e em fotometria (*view based*). Muitos

algoritmos foram desenvolvidos buscando resolver o problema de reconhecimento facial, entretanto, três deles se destacam: o algoritmo PCA (*Principle Component Analysis*) (KIRBY e SIROVICH, 1990), o algoritmo LDA (*Linear Discriminant Analysis*) (FISHER, 1936) e o algoritmo EGBM (*Elastic Graph Bunch Model*) (WISKOTT *et al*, 1997), (NTSC, 2006).

O algoritmo PCA, também conhecido como método *eingenface*, consiste em linearmente projetar o espaço de imagens em um espaço de características com dimensões reduzidas, fazendo o uso da análise de componentes principais (PCA), (KINUTA *et al.*, 2006). Essa redução de dimensões remove as informações que não são úteis ao método e decompõe precisamente as estruturas da face em componentes ortogonais denominados *eigenfaces*. Cada imagem facial é representada como um vetor de características (*feature vector*) de *eigenfaces*, armazenados em vetores unidimensionais. As comparações do método são realizadas entre a imagem questionada e o banco de dados, através das medidas de distância entre os vetores de características de cada imagem. Esta abordagem exige que as imagens faciais sejam frontais (NTSC, 2006).

A Figura 4.2 apresenta amostra de aplicação do método *eigenfaces*.

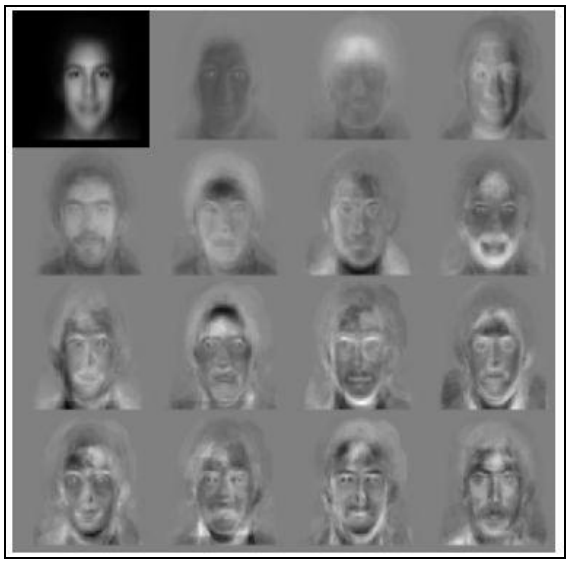


Figura 4.2 – Vetor de características derivadas do uso do método *eigenfaces* (MIT, 2002).

A discriminante linear de Fisher (LDA), também conhecida como método *fisherfaces*, foi desenvolvida por Fisher (1936). É um método específico a classes, pois trabalha com o uso de rótulos para reconhecimento facial. O conjunto de faces de uma mesma pessoa é

rotulada em classes específicas associadas àquela pessoa. O algoritmo trabalha obtendo matrizes de dispersão entre classes distintas (interclasse) e dentro da mesma classe (intraclasse). As projeções são feitas maximizando a dispersão interclasse e minimizando a intraclasse, visando melhorar o desempenho da classificação (BELHUMEUR, HESPANHA e KRIEGMAN, 1997). A Figura 4.3 apresenta um conjunto de classes utilizadas para a aplicação do método *fisherfaces*.

Figura 4.3 – Exemplo de classes utilizadas pelo método LDA (LU, PLATANIOTIS, VENETSANOPOULOS, 2008).

O algoritmo EBGM (Combinação de Grafos em Grupos de Elásticos) foi proposto por Wiskott *et al.* (1997) e é fundamentado no conceito de que as imagens faciais apresentam muitos aspectos não lineares, tais como aqueles produzidos por variações das fontes de iluminação, posicionamento e expressões da face. Nesta abordagem, um filtro de Gabor (*Gabor wavelet*), transformada janelada de Fourier com uma janela gaussiana (GABOR, 1946), é utilizado para projetar a face em uma grade elástica, formada por nós (*Garbor jets*), que descrevem o comportamento da imagem em um dado pixel. O processo de reconhecimento facial baseia-se na similaridade dos nós modelados pelo filtro de Gabor sobre cada imagem. A Figura 4.4 apresenta o processo de modelamento da face pelo método EBGM (NTSC, 2006).

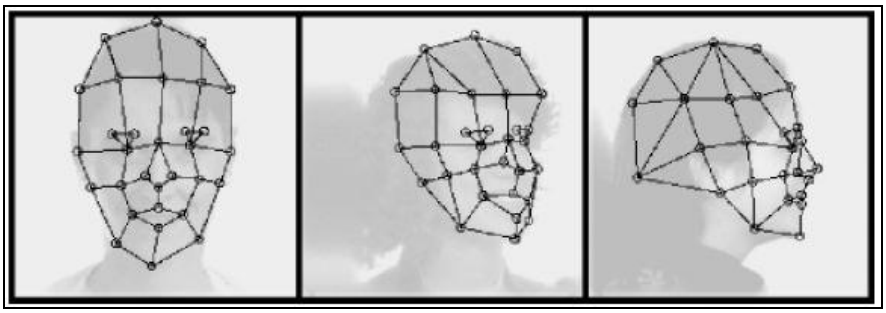


Figura 4.4 – Aplicação do método EBGM em uma imagem facial (NTSC, 2006).

O maior problema dos algoritmos mencionados refere-se ao fato de que o desempenho apresentado por eles, na maioria dos casos, utilizou imagens em condições ideais, na qual houve colaboração do sujeito para capturar uma imagem frontal de sua face, as condições de iluminação eram controladas e a imagem de fundo era simples. Na prática, principalmente em se tratando de sistemas de vigilância (CFTV), muitos fatores impactam no baixo desempenho dos sistemas desenvolvidos, conforme apresentado nas próximas seções (PARK, 2009).

4.2.1 – VARIAÇÕES DE POSICIONAMENTO DA FACE

O posicionamento facial é um dos principais fatores impactantes na degradação do desempenho de sistemas de reconhecimento facial (LI e JAIN, 2005). Quando a captura de uma face monitorada não é frontal, os sistemas de reconhecimento facial têm dificuldade em realizar sua tarefa. Métodos que utilizam a abordagem baseada em características (*feature based*) procuram minimizar o impacto deste fator, como apresentado nos trabalhos de Arca, Campadelli e Lanzarotti (2003), Hadid, Ahonen e Pietikainen (2004) e Zhang *et al.* (2005) (PARK, 2009).

4.2.2 – VARIAÇÕES DE ILUMINAÇÃO

Zhao e Chellappa (1999) mostraram que grandes variações de iluminação sobre imagens faciais de uma mesma pessoa são mais significativas do que as diferenças faciais entre pessoas diferentes. Como o rosto é um objeto 3D, fontes de iluminação distintas provocam variações nas condições de iluminação e sombra da face. Alguns trabalhos científicos têm focado o desenvolvimento de métodos robustos frente às variações de iluminação ambiente, em sistemas de reconhecimento facial. Destacam-se neste campo, os trabalhos de Zhao e Chellappa (2003) e Qing *et al.* (2006). Esses métodos procuram normalizar as condições de iluminação sobre a face. Os resultados obtidos auxiliaram a melhorar o desempenho dos sistemas de reconhecimento facial nos quais o método foi aplicado (PARK, 2009).

4.2.3 – EXPRESSÕES FACIAIS

Outro problema impactante em sistemas de reconhecimento facial refere-se às variações da expressão facial de uma mesma pessoa, dificultando sua identificação. Métodos baseados em características (*feature based*), como o desenvolvido por Martinez (2002); e em modelos 3D da face, como os apresentados por Kakadiaris *et al.* (2007) e Lu e Jain (2008) procuram minimizar esse fator (PARK, 2009).

4.2.4 – OCLUSÃO DA FACE

Não é incomum encontrar objetos cobrindo regiões da face, como óculos, bonés, chapéus e lenços, especialmente quando se trata de sistemas de vigilância, no qual o volume de pessoas monitoradas é grande. Muitos sistemas comerciais de reconhecimento facial simplesmente rejeitam imagens nas quais a região dos olhos não pode ser detectada. Métodos baseados em características (*feature based*), como os apresentados por e Martinez (2002) e Hotta (2008), foram propostos para contornar este problema (PARK, 2009).

4.2.5 – VARIAÇÕES FACIAIS COM A IDADE

O desempenho de sistemas de reconhecimento facial associados a problemas de variações da face com a idade tem sido pouco estudado no meio científico. A ausência de bancos de dados de imagens para este fim e a presença de outros fatores críticos que impactam negativamente no desempenho dos sistemas tem contribuído para este quadro. Os efeitos na idade na face advêm principalmente das rugas, manchas, ganho ou perda de peso, além de mudanças em características do rosto, como boca, bochecha e recaimento dos olhos. Todas essas mudanças influenciam no desempenho de sistemas de reconhecimento facial. Alguns métodos de reconhecimento facial poderiam ser treinados para retirar essas variações da imagem, de forma a melhorar o desempenho dos sistemas associados (PARK, 2009).

5 – TRABALHOS CORRELATOS

Neste capítulo serão apresentados trabalhos correlatos ao desenvolvimento do presente trabalho de pesquisa. O conteúdo do capítulo abrange principalmente os trabalhos práticos divulgados nesta área de pesquisa.

Na seção 5.1 serão apresentados os projetos do NIST fomentando o desenvolvimento de sistemas de reconhecimento facial automático. Na seção 5.2 são apresentadas as recomendações do SWGIT (*Scientific Working Group on Imaging Technology*), grupo de trabalho internacional de pesquisa em assuntos relacionados a imagens. Na seção 5.3 são apresentadas as recomendações de várias instituições suecas que formaram um grupo de trabalho europeu para definir recomendações para sistemas de CFTV. Por fim, na seção 5.4 é apresentado o trabalho de um pesquisador australiano referente à qualidade de imagens em sistemas de CFTV, necessárias ao reconhecimento facial de pessoas.

5.1 – NIST E SEUS PROJETOS

Nos Estados Unidos, o desempenho de sistemas de reconhecimento facial automatizado tem sido testado nas últimas décadas. O *National Institute of Standards and Technology* (NIST), órgão norte-americano, equivalente ao INMETRO brasileiro, iniciou um projeto, no ano de 1993, denominado *Face Recognition Technology* (FERET), tendo funcionado até o ano 1997. O objetivo do FERET visava desenvolver sistemas de reconhecimento faciais automatizados que pudessem assistir aos órgãos de segurança, de inteligência, às forças policiais e ao judiciário, em seus novos projetos de segurança. Os participantes deste projeto eram formados principalmente por entidades acadêmicas, aos quais eram aplicados vários testes relacionados à automatização de processos de reconhecimento facial, até então realizados de forma manual e consideravelmente lenta. A gama de testes conduzidos no projeto FERET se subdividiam em duas vertentes principais: *Automatic Face Verification* (AFV) e *Automatic Face Recognition* (AFR), (CAMASTRA e VINCIARELLI, 2008; NIST, 2011).

Alguns anos após o encerramento do projeto FERET, o NIST iniciou outros dois projetos na área de reconhecimento facial automatizado, envolvendo desta vez, além de instituições acadêmicas, empresas especializadas da área, cujo objetivo era pesquisar os problemas não

resolvidos pelo projeto FERET. Os novos projetos foram denominados *Face Recognition Vendor Test* (FRVT), operacionalizado nos anos 2000, 2002 e 2006; e o *Face Recognition Grand Challenge* (FRGC), cuja duração ocorreu entre os anos 2004 e 2006 (NIST_a, 2011; NIST_b, 2011). Foi utilizado nos testes realizados durante o projeto FRVT de 2002, um banco de dados contendo 121.589 imagens de 37.437 pessoas distintas. O desempenho nos testes de reconhecimento facial obteve um êxito aproximado de 70% das comparações (PHILLIPS *et al.*, 2003). Na última série de experimentos do projeto FRVT, em 2006, envolvendo testes de verificação, o desempenho do melhor sistema apresentado obteve como resultados uma percentagem de falso positivo ou *False Reject Rate* (FRR) de 0,01 e uma percentagem de falso negativo ou *False Accept Rate* (FAR) de 0,001, utilizando imagens 2D de alta resolução com distância entre as pupilas ou *interpupillary distance* (IPD) de 400 pixels, além de imagens 3D, como aquelas apresentadas na Figura 5.1. O Gráfico 5.1 apresenta a evolução dos sistemas de reconhecimento faciais automatizados desenvolvidos nos EUA (PHILLIPS *et al.*, 2007).

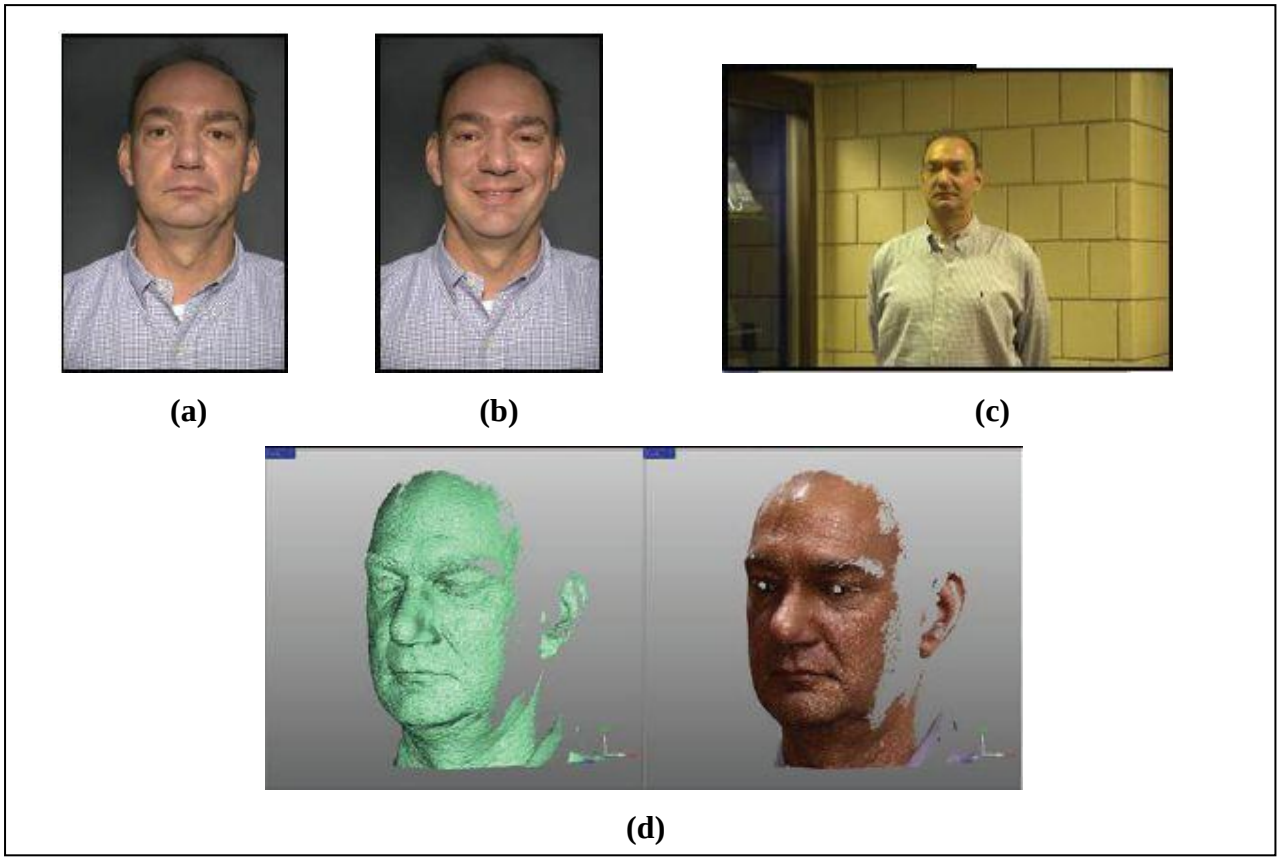


Figura 5.1 − Exemplo de imagens utilizadas nos testes do FRVT 2006. (a) Iluminação controlada e expressão facial neutra (IPD = 400 pixels), (b) Iluminação controlada e expressão de sorriso, (c) Sem controle de iluminação e expressão de sorriso (IPD = 190 pixels) e (d) canal de formato e textura de uma imagem 3D da face (extraído de (PHILLIPS *et al.*, 2007)).

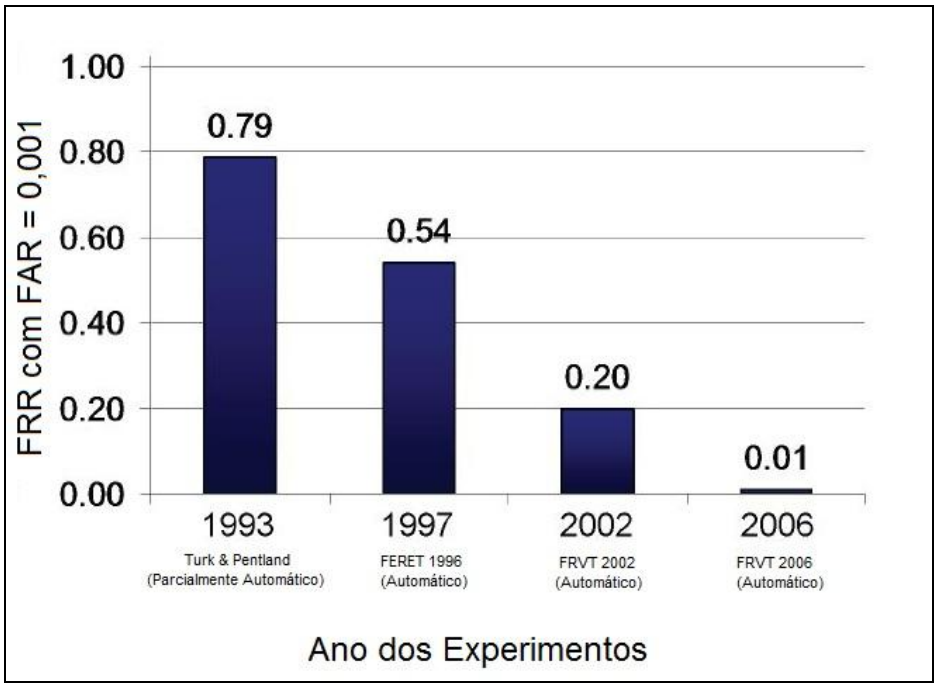


Gráfico 5.1 − Evolução da taxa de erros nos Sistemas de Reconhecimento Facial (extraído de (PHILLIPS *et al.*, 2007)).

Apesar da baixa taxa de erros apresentada nos sistemas apresentados ao NIST, deve-se ressaltar que o desempenho dos testes foi baseado em condições controladas, não refletindo a realidade operacional de muitos desses sistemas de segurança, principalmente no que tange às condições de iluminação ambiente e ao posicionamento da face do sujeito. Em situações práticas, o desempenho dos sistemas testados tornava-se insatisfatório. Ainda assim, esses sistemas desempenhavam importantes papéis na verificação e reconhecimento de suspeitos, naquelas aplicações onde poderiam ser controladas as variáveis internas e externas do sistema, como ocorre no registro de documentação pessoal (passaporte, identidade e na licença para dirigir), (JAIN, KLARE e PARK, 2002).

5.2 – SWGIT/IAI

A *International Association for Identification* (IAI) é uma organização com objetivos forenses, criada em 1915 na cidade de Oakland, Estado da Califórnia/EUA, com o intuito de promover pesquisas na área de identificação de pessoas. A IAI, hoje composta por cerca de 6.700 membros espalhados em todo o mundo, se esforça para ser a principal associação profissional envolvida em pesquisas relacionadas à identificação forense, investigação e análise pericial de provas físicas. Desde o ano de 1930, a IAI possui estreito

relacionamento com o *Federal Bureau of investigation* (FBI) e sua atuação tem como base as normas do sistema judiciário norte americano (IAI, 2011).

Dentre as diversas frentes de trabalho do IAI está o *Scientific Working Group on Imaging Technology* (SWGIT), um grupo de trabalho que foi criado com a missão de facilitar a integração de sistemas e tecnologias de imagens apresentadas por órgãos governamentais e empresas em geral, com o sistema de justiça criminal, fornecendo definições e recomendações para a captura, armazenamento, processamento, análise, transmissão e produção de imagens. A meta do SWGIT é envolver todas as autoridades e instituições legais no desenvolvimento dessas diretrizes (SWGIT, 2010).

Atualmente o SWGIT possui 22 seções, que correspondem a 22 publicações em temas relacionados às metas do grupo. Para o interesse deste trabalho de pesquisa, destacam-se as publicações das seguintes seções:

- ***seção 1***: *Overview of SWGIT and the Use of Imaging Technology in the Criminal Justice System* – Visão Geral do SWGIT e o uso de tecnologia de imagem no sistema de justiça criminal;
- ***seção 4***: *Recommendations and Guidelines for Using Closed-Circuit Television Security Systems in Commercial Institutions* – Recomendações e Orientações para o uso de circuito fechado de televisão em sistemas de segurança de instituições comerciais;
- ***seção 13***: *Best Practices for Maintaining the Integrity of Digital Images and Digital Video* – Melhores práticas para manter a integridade de imagens e vídeos digitais;
- ***seção 14***: *Best Practices for Image Authentication* – Melhores práticas para autenticação de imagens.

A seção 1 do SWGIT apresenta o próprio grupo de trabalho SWGIT e estabelece as diretrizes básicas sobre manipulação de imagens para fins forenses. O documento está na versão 3.3 e sua última versão foi publicada em 11/06/2010. Destacam-se nesta publicação a admissibilidade de imagens pela Justiça, questões sobre a captura e integridade de imagens. Com relação ao primeiro tópico, as imagens digitais têm sido aceitas na esfera judiciária, desde que apropriadamente autenticadas e quando alteradas, no intuito de

melhorá-las, que essas alterações sejam devidamente explicadas pelos peritos (SWGIT, 2010).

No tópico de captura de imagens, o documento relata que quando as imagens são produzidas por profissionais autorizados, como peritos criminais, é possível controlar os equipamentos, métodos e técnicas aplicadas à captura da imagem. Esse controle não é possível quando as imagens são geradas por pessoas não autorizadas ou qualificadas, prática que deve ser evitada (SWGIT, 2010).

Já em relação à manutenção da integridade da imagem, é requisito para a sua admissibilidade que a imagem possa ser autenticada. Uma das formas listadas é o uso de uma testemunha para certificar que a imagem é autêntica. Caso o conteúdo da imagem seja contestado pelo réu, o autor no processo judicial deve estar preparado para mostrar que a imagem não foi alterada (SWGIT, 2010).

Na seção 4 são apresentadas as recomendações para o uso de sistemas de CFTV com fins de segurança de estabelecimentos comerciais. O documento está na versão 2.1 e sua última atualização foi feita em 22/07/2004. O intuito destas recomendações é aperfeiçoar a qualidade das imagens geradas por estes sistemas de forma a facilitar o reconhecimento facial de pessoas desconhecidas e de objetos capturados pelas câmeras do sistema. São abordados neste documento detalhes de projeto do sistema de CFTV, equipamentos de gravação, câmeras, mídias de armazenamento de dados, manutenção do sistema, retenção das gravações e manipulação de evidências (SWGIT, 2004).

A proposta da seção 4 visa aumentar a probabilidade de que imagens recuperadas de sistemas de CFTV tenham qualidade suficiente para permitir que as autoridades legais possam identificar pessoas e objetos através das imagens coletadas. No caso do reconhecimento facial de pessoas, as imagens devem fornecer pequenos detalhes de pessoas, suficientes para poderem ser observadas marcas, como tatuagens e cicatrizes de pele. Na Figura 5.2, o SWGIT apresenta um exemplo do que é esperado dos sistemas de CFTV recomendados por esta seção (SWGIT, 2004).

Figura 5.2 – (a) Imagem extraída de sistema de CFTV com características adequadas para reconhecimento facial. (b) Imagem extraída de sistema de CFTV sem condições adequadas para o reconhecimento facial. (c) Imagem da figura (a) processada para focar apenas a face da pessoa na imagem. (d) Imagem da figura (b) processada para focar apenas a face da pessoa na imagem (extraído de (SWGIT, 2004)).

Na seção 13 são apresentadas recomendações métodos para manter e demonstrar a integridade de imagens, desde sua aquisição até sua apresentação perante um tribunal. A versão 1.0 é a última publicada e data de 04/06/2007. Manter a integridade de uma imagem requer a segurança dos arquivos durante seu transporte e armazenamento. Demonstrar integridade refere-se aos métodos utilizados para mostrar que o arquivo não foi alterado (SWGIT_a, 2007).

Dentre os métodos utilizados para manter a integridade, destacam-se (SWGIT_a, 2007):

- *Documentação escrita*: POP (Procedimento Operacional Padrão) documentando os passos requeridos para manter apropriadamente a segurança do arquivo. Este método deve incluir a cadeia de custódia, no caso de instituições policiais.
- *Ambiente de segurança física*: sistemas mecânicos ou físicos para evitar acesso não autorizado ou perda de dados, tais como fechaduras em portas, seguranças, sistemas de controle de entrada de pessoas, sistemas de combate a incêndio, computadores isolados de redes, dentre outros.
- *Cópias físicas redundantes*: cópias duplicadas dos arquivos mantidas em localizações distintas para evitar a perda de dados no caso de ocorrência de desastres.
- *Segurança lógica*: sistemas operacionais ou programas destinados a evitar acesso aos arquivos, tais como proteções por senhas, *firewalls*, dentre outros.
- *Empresas terceirizadas*: neste caso os arquivos são repassados a uma empresa terceirizada que fica responsável pela segurança dos arquivos. Embora esta situação pode não ser apropriada em determinadas circunstâncias, a agência deve ter um método viável para demonstrar a integridade independente do fornecedor, além de haver entre as partes um contrato adequado que esclareça as obrigações de cada parte.

Dentre os métodos utilizados para demonstrar a integridade, destacam-se (SWGIT_a, 2007):

- *Funções de hash*: cálculo matemático que gera um valor numérico de acordo com a entrada de dados utilizada. Este valor numérico é o *hash*. Valores de *hash* são calculados utilizando fórmulas complexas e seus resultados são sensíveis a quaisquer pequenas mudanças de dados dos valores de entrada.
- *Verificação visual*: o processo de confirmação da precisão de uma imagem através de inspeção visual.
- *Assinatura digital*: este é um processo utilizado junto com o processo de *hash*. O resultado do cálculo de *hash* é criptografado com a utilização de uma chave privada. A integridade do arquivo pode ser valorada através de novo cálculo de *hash*, que é comparado com o valor criptografado armazenado. A vantagem do uso

de assinatura digital é que a fonte do arquivo pode ser atribuída a determinada pessoa (autoria).

- *Documentação escrita*: notas escritas que documentem a sequência de passos do arquivo.
- *Verificação de redundância cíclica (Checksum/CRC)*: normalmente utilizados para verificar se a transferência de um arquivo foi bem sucedida. Podem ser tão poderosos quanto as funções de hash. Aconselha-se utilizar verificações CRC com outros métodos, de acordo com a segurança requerida.
- *Criptografia*: este processo altera o conteúdo do arquivo e não demonstra que o arquivo não foi alterado. A criptografia pode ser usada em conjunto com outros processos para então, demonstrar a garantia da integridade do arquivo.
- *Marca d'água*: este método modifica o conteúdo do arquivo, podendo então persistir como parte do arquivo. Não é um método recomendado.
- *Métodos proprietários*: métodos oferecidos mediante pagamento ou licença, que não podem ser verificados de forma independente. Não é um método recomendado.

Na seção 14 são apresentadas recomendações práticas para realizar a autenticação de imagens, como parte das práticas de análise forense de imagens. A última versão do documento (1.0) data de 04/06/2007. Deve-se lembrar de que o processo de autenticação de imagens não pode ser confundido com os requerimentos de autenticação de uma evidência, como exigidos para a admissibilidade perante um tribunal. O processo de autenticação de uma imagem pode envolver muitas tarefas, que incluem a avaliação da estrutura da imagem e de seu conteúdo. A avaliação da estrutura da imagem inclui a descoberta de artefatos consistentes com a manipulação da imagem ou sua degradação, análise dos metadados e indicações de sua procedência. A avaliação do conteúdo da imagem inclui questões de continuidade do conteúdo mostrado, manipulação deste conteúdo, encenação e a presença de anacronismo (SWGIT, 2007).

5.3 – O GRUPO DE TRABALHO DAS INSTITUIÇÕES SUECAS

Foi emitido no ano de 2005 um conjunto de diretrizes para sistemas de CFTV, produzidos pelas instituições suecas *The Swedish National Laboratory of Forensic Science* – Laboratório Nacional Sueco de Ciências Forenses, *The Swedish Police* – A Polícia Sueca,

The Swedish Bankers' Association – A Associação Sueca de Bancos e *The Swedish Federation of Trade and Services* – Federação Sueca de Comércio e Serviços (BERGSTRÖM, 2005).

A fim de proporcionar uma oportunidade a fabricantes, instaladores, usuários e autoridades em discutir os problemas relacionados a sistemas de CFTV, já que nem todos os sistemas podem fornecer imagens com qualidade suficiente para prover o tipo de proteção esperada pelo usuário, o Laboratório Nacional Sueco de Ciências Forenses (SKL) tomou a iniciativa de criar fóruns para debater o assunto. A primeira conferência, ocorrida em setembro de 2003, teve como foco criar recomendações para evitar os problemas conhecidos de sistemas de CFTV, bem como realizar um levantamento da capacidade de polícia sueca em analisar diferentes tipos de mídias e formatos de imagens. A segunda conferência, ocorrida em março de 2005, continuou os trabalhos iniciados na primeira conferência, tendo como resultado, em um esforço das instituições acima mencionadas, bem como de fabricantes e engenheiros que trabalham nessa área, um conjunto de diretrizes envolvendo sistemas de CFTV, de forma que as imagens produzidas por estes sistemas possam ser utilizadas pela Polícia e pelos demais órgãos do sistema legal, no combate à criminalidade. As recomendações se aplicam à qualidade da imagem fornecida (qualidade e conteúdo) e à manipulação dessas imagens (acessibilidade, segurança e documentação), (BERGSTRÖM, 2005). A seguir é apresentada uma síntese das recomendações relacionadas à qualidade da imagem e à sua acessibilidade.

Os sistemas de CFTV são utilizados geralmente para aplicações de segurança orgânica. O usuário desse sistema deve definir adequadamente qual é essa necessidade de segurança e direcionar assim, a proposta do sistema, ou seja, o que é esperado das imagens do sistema. Foram definidos quatro níveis de segurança com base nessas necessidades (BERGSTRÖM, 2005):

- *Chain of events* – Cadeia de eventos (A): o primeiro nível de segurança tem como objetivo fornecer imagens com grande amplitude de visualização.
- *Characteristic features* – Características (B): neste nível, as imagens devem fornecer foco em locais importantes do perímetro a ser vigiado.
- *Identification* – Identificação (C): o terceiro nível de segurança, as imagens devem fornecer *close-up*s a ponto de permitirem o reconhecimento facial de pessoas.

- *Biometry* – Biometria (D): o último nível deve ser capaz de fornecer ao usuário meios de reconstruir a cadeia de eventos e o reconhecimento facial de pessoas, em detalhes.

Em relação às recomendações de qualidade de imagem gerada pelo sistema de CFTV, há três níveis de qualidade: cadeia de eventos, características e identificação. O Laboratório Nacional Sueco de Ciências Forenses produziu uma tabela de teste, que é uma versão modificada dos testes de acuidade visual aplicados na oftalmologia, conforme mostrado na Figura 5.3 (a versão da tabela em tamanho normal é apresentada no apêndice A). A tabela colocada sobre a área de visão de uma câmera de vigilância permite que seja testada a capacidade visual da câmera em reproduzir detalhes, conforme a linha da tabela que possa ser lida (BERGSTRÖM, 2005).

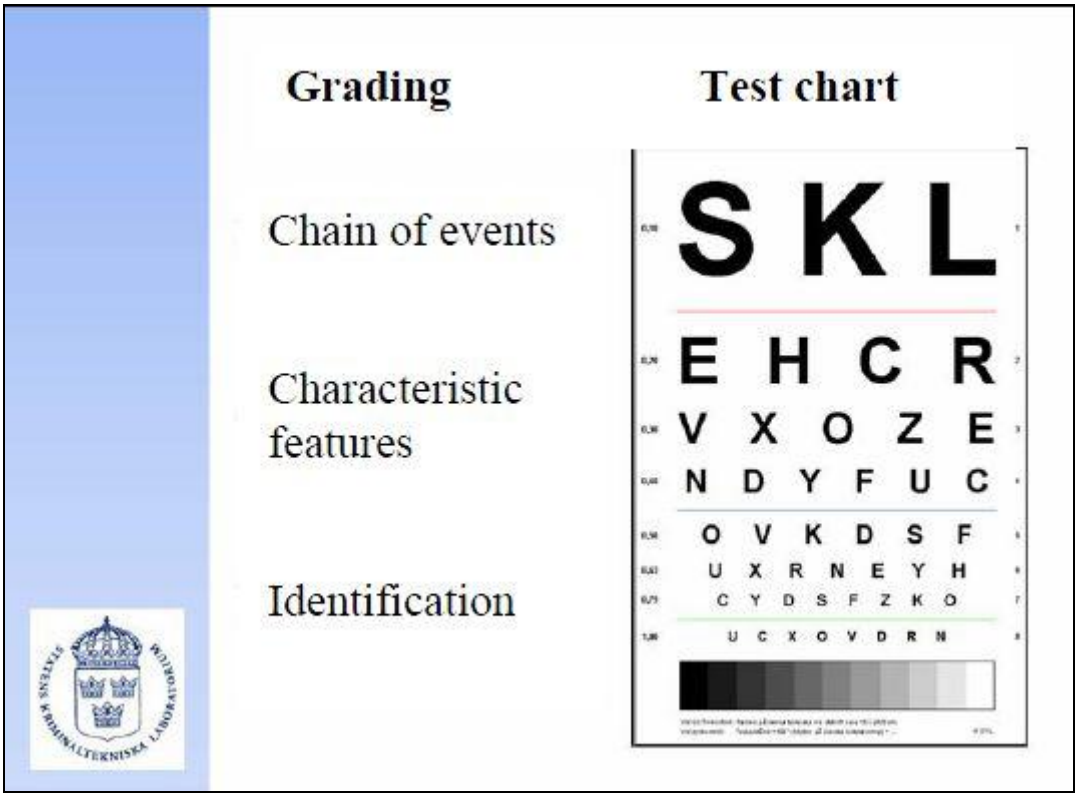


Figura 5.3 − Tabela de teste visual produzido pelo SKL apresentado em escala reduzida (extraído de (BERGSTRÖM, 2005)).

A Tabela 5.1 indica o nível de qualidade exigido de acordo com os níveis de segurança exigidos pelo usuário do sistema de CFTV. Os requerimentos de velocidade de captura da imagem se aplicam em incidentes que estão ocorrendo e indicam os requisitos mínimos exigidos (BERGSTRÖM, 2005).

Tabela 5.1 − Requerimentos de qualidade da imagem (extraído de (BERGSTRÖM, 2005)).

Nível de qualidade da imagem	Linha que deve ser lida da tabela SKL	Velocidade de captura de quadros (quadros/s)
Cadeia de eventos (A)	1	1
Características (B)	2-5	3
Identificação (C)	6-8	5
Biometria (D)	2-5	5

Em relação às recomendações de acessibilidade, o sistema de CFTV deve permitir que as imagens gravadas estejam facilmente e rapidamente acessíveis para quem necessitar de seu uso, inclusive a polícia. Os requisitos de acessibilidade envolvem a remoção do equipamento de gravação, o destinatário deve ser capaz de lidar com o formato de imagem gravado e ter o reprodutor necessário para ter acesso ao conteúdo da imagem. Destacam-se nos requisitos de acessibilidade os formatos de imagens recomendados pelo grupo de trabalho, sendo eles os padrões TIFF (*Tagged Image File Format)* e JPEG2000 (*Joint Photographic Experts Group 2000*), como os padrões de compressão de imagens com menos perdas na qualidade da imagem. Além disso, o padrão JPEG2000 também de uma versão para imagens sequenciais (vídeo), através do padrão M-JPEG2000 (*Motion-JPEG 2000)* (BERGSTRÖM, 2005).

5.4 – O TRABALHO DE KOVESI

Kovesi (2009) publicou um artigo científico relatando experimentos envolvendo câmeras de segurança de sistemas de CFTV. A motivação do autor no trabalho envolvia a frustração ao tentar ajudar a polícia a melhorar imagens extraídas de sistemas de CFTV, cujos resultados pouco ou nada acrescentavam à investigação. O trabalho de Kovesi procurou determinar qual a qualidade mínima necessária para que as imagens de CFTV permitissem o reconhecimento facial de suspeitos e envolveu quatro variáveis no processo: resolução espacial necessária para reconhecimento facial, compressão, luminância e reprodução de cores.

Baseado em estudos de outros pesquisadores, Kovesi observou que as frequências espaciais mais importantes para o reconhecimento facial humano estão na faixa de 5 ciclos

por largura da face a até 20 ciclos por largura da face. Segundo relato de Kovesi (2009), as frequências espaciais acima dessa faixa revelam detalhes como pequenas manchas de pele que não são consideradas significantes para o reconhecimento facial. Da mesma forma, as frequências espaciais abaixo da faixa mencionada estão mais ligadas às variações de iluminação do que de detalhes da face. O pesquisador concluiu, com base nessas pesquisas que, a faixa de frequências espaciais superiores são mais importantes para o reconhecimento facial e que, para reconhecer faces de pessoas desconhecidas com confiança, é preciso resolver as frequências espaciais entre 10 e 20 ciclos por largura da face.

Com base nessa premissa e de que a média de largura de um rosto é aproximadamente 160 mm (16 cm), Kovesi observou que as frequências espaciais de 10 e 20 ciclos por largura de face correspondem aos comprimentos de onda espaciais de 16 mm e 8 mm, respectivamente, conforme apresentado na Figura 5.4. A partir de então, utilizou duas ferramentas para verificar a eficiência das câmeras de vigilância em resolver as frequências espaciais necessárias ao reconhecimento facial: o teste de resolução gráfica USAF 1951 (padrão MIL-STD-150A) (Silverfast, 2012) e o gráfico optométrico logMAR (Bailey e Lovie, 1976).

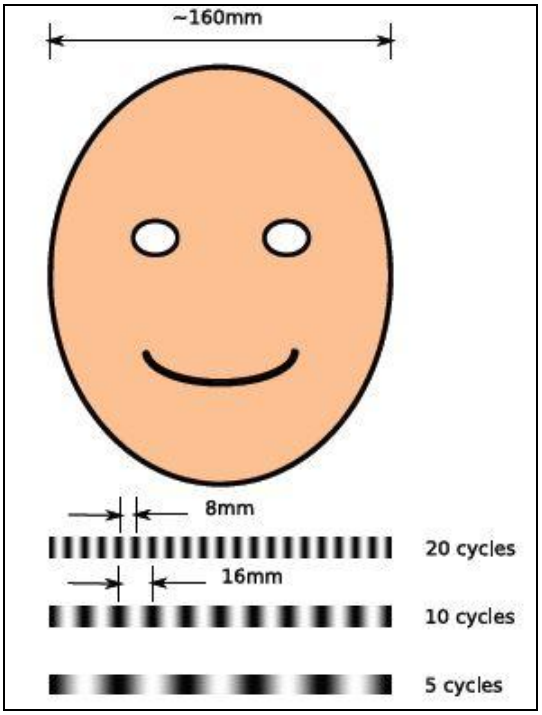


Figura 5.4 − Frequências espaciais importantes para o reconhecimento facial humano (extraído de (KOVESI, 2009)).

O teste de resolução gráfica USAF 1951 é composto por três conjuntos de padrões, cada um contendo seis grupos de pares de barras geometricamente desenhados em escala. Cada grupo corresponde em tamanho à metade do grupo anterior, conforme mostrado na Figura 5.5. O conjunto de barras do topo à direita (comprimento de onda de 8 mm) e da base à direita (comprimento de onda de 16 mm) do gráfico estão em compasso com as frequências espaciais importantes para o reconhecimento facial (KOVESI, 2009).

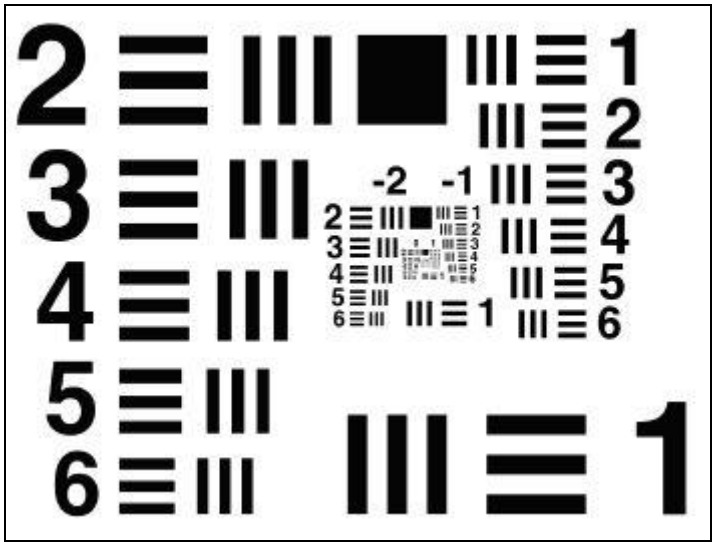


Figura 5.5 – Teste de Resolução Gráfica USAF 1951 (extraído de (KOVESI, 2009)).

A segunda ferramenta utilizada por Kovesi procurou realizar uma espécie de exame de vista nas câmeras de vigilância, determinando sua acuidade visual através do teste gráfico logMAR. A avaliação da acuidade visual de um paciente humano trata-se de um dos procedimentos clínicos mais comuns utilizados na oftalmologia e se refere a um importante parâmetro para medir o desempenho funcional de um sistema visual, compreendendo a capacidade da visão em discriminar formas (BICAS, 2002). As avaliações de acuidade visual geralmente utilizam optótipos padrões (tipos de letras) que devem ser reconhecidos a certa distância, permitindo inferir, através da capacidade visual do paciente, a relação de sua visão à visão normal esperada (EVANS, 2006).

O teste realizado por Kovesi foi desenvolvido pelos oftalmologistas Ian Bailey e Jan Lovie, em 1976 e trata-se de uma tabela composta por 5 (cinco) caracteres por linha, com espaçamento entre as letras e linhas igual ao tamanho do espaço ocupado pelo caractere, conforme mostrado na Figura 5.6. A tabela apresenta um formato de triângulo invertido e

os caracteres utilizados em cada linha mudam de acordo com a progressão logarítmica do tamanho das letras (SOARES, 2009). A acuidade visual de uma pessoa é frequentemente descrita em função da fração de Snellen, que relaciona a distância na qual uma pessoa consegue ler determinada linha do teste de triagem, com a distância que uma visão normal deveria conseguir ler àquela mesma distância. A expressão "visão 20/20" (em pés) ou "visão 6/6" (em metros) está associada a uma pessoa com visão normal. Assim, caso a acuidade de uma pessoa resulte em uma visão 6/12, quer dizer que a pessoa consegue ler/enxergar caracteres no teste de triagem à 6 metros de distância, aqueles caracteres que deveriam ser lidos/vistos por uma visão normal a 12 metros de distância.

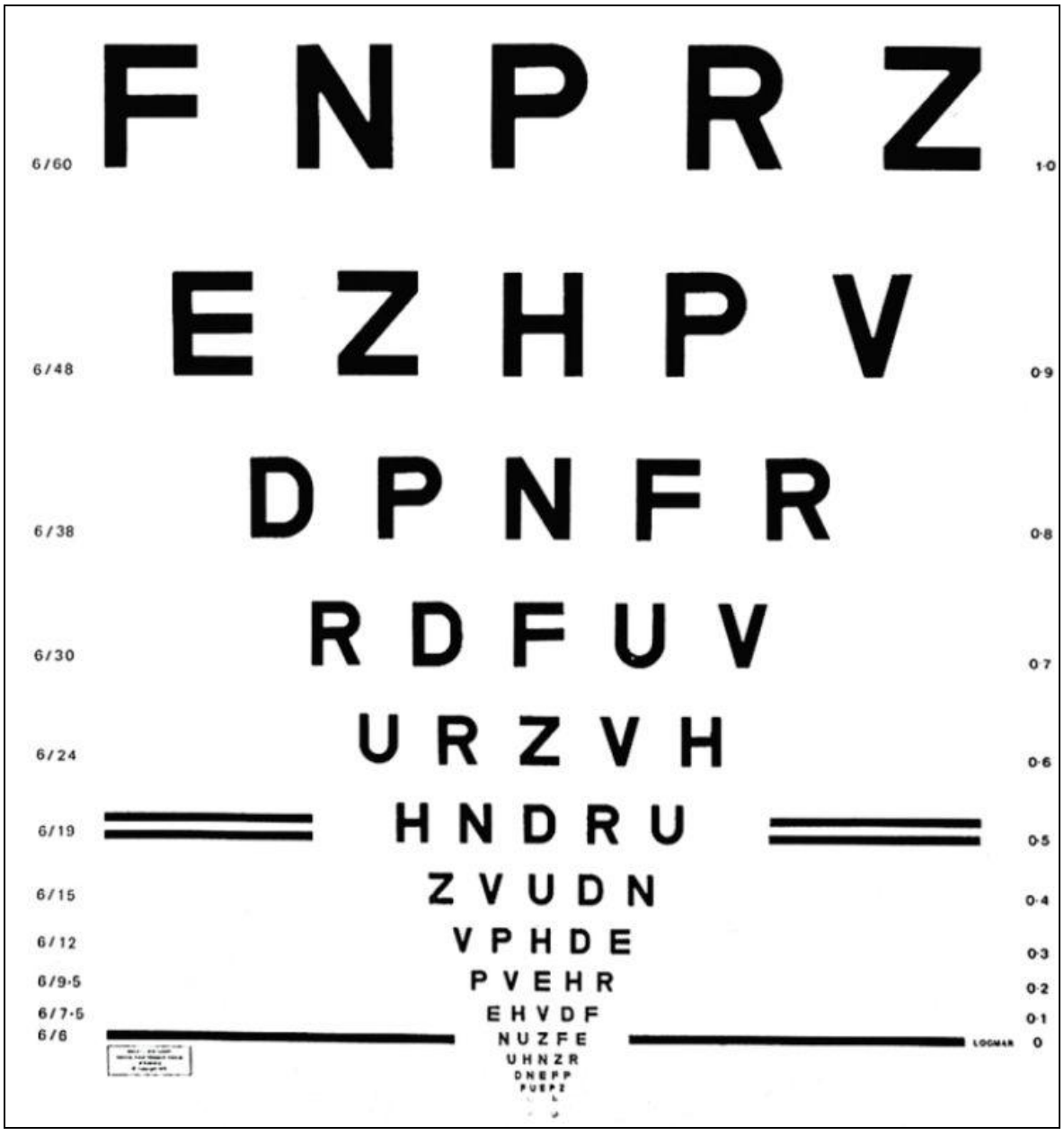


Figura 5.6 – Tabela de teste logMAR (extraído de (KOVESI, 2009)).

Utilizando as ferramentas descritas, Kovesi procurou determinar se uma câmera de vigilância, a uma certa distância de operação, conseguiria resolver as frequências espaciais importantes para o reconhecimento facial. Foi utilizada nos testes uma câmera da marca Pulnix, modelo TM6CN, com sensor CCD de 1/2", utilizando lentes C-mount, cuja distância focal permitia variação de 4mm a 16mm, posicionada a 6 metros dos testes gráficos e de quatro imagens de faces em tamanho real, conforme mostrado na Figura 5.7. As imagens capturadas pela câmera foram digitalizadas através de uma placa de captura da marca Data Translation, modelo 3155.

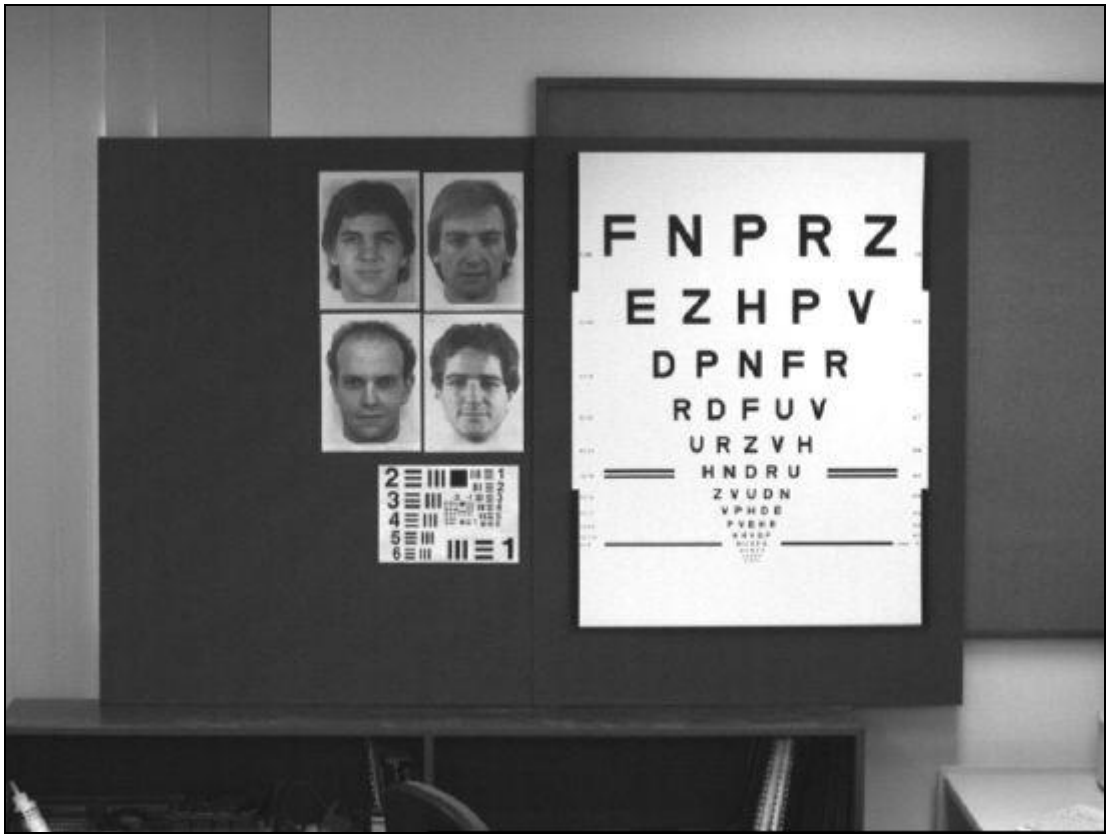


Figura 5.7 – Imagem extraída dos experimentos, utilizando lente com distância focal de 16mm (extraído de (KOVESI, 2009)).

Os resultados dos experimentos de Kovesi, em termos de resolução das frequências espaciais para o reconhecimento facial, demonstraram que o tamanho da face deve possuir pelo menos 40 pixels de largura na imagem, desde que não haja outras deficiências na imagem, tais como a utilização de compressão, para que seja possível reconhecer uma face humana com confiabilidade (KOVESI, 2009). Kovesi utilizou como média antropométrica, as medidas de 160 mm para largura da face e de 63 mm para a distância entre as pupilas.

Desta forma, conclui-se que o IPD mínimo para reconhecimento facial é de aproximadamente 16 pixels.

Com relação à variável compressão, Kovesi chegou à conclusão de que se trata de um dos grandes problemas causadores da degradação de imagens de sistemas de CFTV. Esses sistemas costumam utilizar métodos de compressão agressivos em imagens, no intuito de otimizar a capacidade de armazenamento das mídias computacionais utilizadas. Kovesi, observando experimentos próprios e de outros pesquisadores, constatou que em imagens com largura de face de 38 pixels estão apenas um pouco abaixo do limite de Nyquist de fornecer as informações de resolução espaciais necessárias de até 20 ciclos por largura de face. O maior problema do uso da compressão está nas interações dos blocos 8x8 da transformada do cosseno discreto (DCT) em imagens no padrão JPEG ou vídeos no padrão MPEG. Há aproximadamente 5 desses blocos DCT abrangendo cada face (KOVESI, 2009).

A quantização dos componentes de frequência aplicadas pela compressão em cada um destes blocos mostrou que as informações de frequência espacial exigidas para o reconhecimento facial, ou seja, a partir de 5 ciclos por largura de face até 20 ciclos por largura de face, podem ter sido corrompidas. Desta forma, em imagens típicas de sistemas de CFTV, com baixa resolução, a aplicação de compressão tem efeitos significantes no desempenho do reconhecimento facial de pessoas. Comparativamente, as imagens de baixa resolução utilizadas nos testes do FRVT 2006, de largura de face de 190 pixels, explicam os bons desempenhos dos sistemas de reconhecimento facial automatizados em taxas de compressão de 20:1. Neste caso, havia até 23 blocos de matrizes 8x8 DCT cobrindo cada face da imagem. Em casos reais, a largura de face possui níveis de detalhe inferiores daqueles utilizados no FRVT 2006 (KOVESI, 2009).

Por fim, Kovesi observou que as variáveis luminância e a reprodução de cores têm efeitos sobre a forma e a pigmentação das faces observadas, influenciando diretamente na correta diferenciação das faces de pessoas, citando experimentos de Russel *et al.* (2007), conforme apresentado na Figura 5.8.

Figura 5.8 – Rostos na linha inferior são imagens digitalizadas de faces diferenciadas em forma e pigmentação. Rostos na linha do meio só diferem em pigmentação, não em forma. Rostos na linha superior diferem apenas em forma, não em pigmentação (extraído de (KOVESI, 2009) citando (RUSSEL *et al.*, 2007)).

6 – PROPOSTA DO TRABALHO

Esse capítulo apresentará em detalhes os critérios propostos para modelar projetos e instalações de câmeras de segurança de sistemas de CFTV, em ambientes *indoor*. A escolha dos critérios teve como base a experiência adquirida durante a fase de pesquisa bibliográfica, a própria experiência do autor do texto em sua atuação profissional, a possível utilização, bem como o foco forense de utilização das imagens coletadas desses sistemas de segurança, em um tribunal judicial, como evidência de um crime.

Desta forma, este capítulo foi estruturado em 3 seções. Na seção 6.1 será abordada a metodologia de pesquisa utilizada para chegar à proposta do trabalho. A seção 6.2 retoma brevemente o problema de pesquisa. São descritos, por fim, na seção 6.3, os critérios propostos e suas características.

6.1 – METODOLOGIA DE PESQUISA

Para atingir os objetivos propostos para esta dissertação, após a definição do tema e a identificação do problema de pesquisa, etapa realizada junto ao professor orientador deste trabalho, optou-se por uma abordagem teórico-empírica, explicada a seguir, para estruturar todo o planejamento e execução da pesquisa.

A abordagem teórica consistiu na etapa de levantamento bibliográfico e revisão de literatura. Foram consultadas literaturas gerais relacionadas a Processamento Digital de Sinais, Processamento Digital de Imagens, Telecomunicações, Oftalmologia, Direito e Criminalística; artigos científicos relacionados ao problema de pesquisa, disponíveis na Internet, além de dissertações de mestrado e teses de doutorados, contendo referências ao tema pesquisado. A abordagem teórica, apesar de ter sido executada desde a fase inicial, continuou ao longo de todo o desenvolvimento da pesquisa. A importância desta etapa não se restringiu apenas às considerações referentes ao embasamento teórico do trabalho, mas auxiliou principalmente na busca de uma solução inspiradora que pudesse resolver o problema de pesquisa.

A abordagem empírica foi dividida em duas etapas: a primeira foi marcada pela definição de uma proposta de critérios para a instalação de câmeras de segurança em projetos de

CFTV e a segunda, pela realização de experimentos que permitiram validar parâmetros para os critérios propostos. A etapa experimental será tratada no próximo capítulo.

6.2 – RETOMANDO O PROBLEMA DE PESQUISA

Os sistemas de CFTV têm sido largamente utilizados em ambientes corporativos e órgãos de governo, como parte das soluções dos sistemas de segurança que integram estes ambientes, monitorando espaços físicos onde há acesso de pessoas. A proliferação do uso desses sistemas, em qualquer que seja o ambiente instalado, tem contribuído relativamente com o combate à criminalidade, no sentido que as imagens captadas por esses sistemas podem fornecer evidências importantes de materialidade e autoria de crimes, permitindo aos órgãos integrantes da persecução penal, maior efetividade na condenação de criminosos.

Apesar da disponibilidade cada vez mais frequente de sistemas de CFTV ser positiva como ferramenta de combate ao crime, a eficiência do uso desses sistemas neste aspecto ainda tem muito que evoluir. Não é difícil encontrar problemas associados à instalação de câmeras de segurança de sistemas de CFTV, seja ao adentrar órgãos públicos, seja em estabelecimentos comerciais. Esses sistemas são mal dimensionados e falham em fornecer imagens com qualidade suficiente, para que seja possível a realização do reconhecimento facial de suspeitos. A realidade brasileira é crítica neste ponto. Laudos periciais emitidos pela Polícia Federal já constataram o posicionamento inadequado de câmeras estratégicas para a segurança de órgãos públicos, no qual o campo de visão da câmera era obstruído por pilastras do interior do ambiente. A título de exemplo, para aqueles que já assistiram ao filme brasileiro “Assalto ao Banco Central”, obra fictícia que procurou retratar furto milionário ocorrido no país ao Banco Central do Brasil, na cidade de Fortaleza, Estado do Ceará, no ano de 2005 e investigado pela Polícia Federal; no filme é retratado esse problema dos projetos de CFTV. Independentemente de ser verídico ou não, os fatos mostrados no filme em relação ao sistema de CFTV do Banco, servem de ilustração ressaltar o erro grave de posicionamento das câmeras de segurança dentro do cofre principal do estabelecimento, no qual uma empilhadeira cobria parte do campo de visão das câmeras, permitindo que pessoas pudessem ali passar, sem serem vistas nas telas dos monitores da central de controle (vide Figura 6.1).

Figura 6.1 – Quadro extraído do filme "Assalto ao Banco Central". Na figura é mostrado um ambiente fictício da central de controle do sistema de CFTV do Banco Central. No monitor do centro mostrado na figura, uma empilhadeira e outros objetos obstruem boa parte do campo de visão de uma das câmeras que monitoram o cofre do Banco.

No capítulo introdutório deste trabalho já foram apresentados alguns dos fatores que contribuem para agravar o problema do dimensionamento inadequado dos projetos de CFTV. Em tempo, podem ser listados os baixos investimentos na aquisição de soluções de CFTV, utilização de equipamentos de baixa qualidade, ausência de amparo profissional especializado na especificação da melhor solução a ser adquirida, posicionamento inadequado das câmeras de segurança, condições de iluminação ambiente não controladas e geração de imagens com baixa resolução espacial (WILLIANS, 2012) (COSTA, 2010).

Park (2009) aborda alguns desses fatores, destacando algumas características normalmente encontradas em sistemas de CFTV:

- *Posicionamento da face*: a posição de instalação das câmeras normalmente não facilita a captura de posicionamento frontal da face das pessoas sendo monitoradas. Não é esperada também a cooperação das pessoas sendo monitoradas, em virtude das aplicações desses sistemas, normalmente associadas à segurança de ambientes. Por outro lado, a captura da face é passiva e não depende da vontade de quem está sendo monitorado.

- *Condições de iluminação do ambiente*: câmeras de vigilância instaladas em ambientes externos estão condicionadas às variações de iluminação natural (dias ensolarados e dias nublados). Essas variações nas condições de iluminação impactam negativamente no desempenho de sistemas de reconhecimento facial.
- *Baixa resolução espacial*: os sistemas de segurança geralmente utilizam um campo de visão amplo para cobrir a maior área possível do perímetro de segurança. Entretanto, o tamanho das faces gravadas nos quadros de vídeo é pequeno, com distância entre os olhos em torno de 10 pixels (IPD ≈ 10), encontradas em imagens com baixa resolução. As imagens apresentadas na seção 2.1.1 são uma amostra da qualidade das imagens encontradas em casos reais abordados pela Polícia Federal.

É importante frisar outro fator importante que contribui para agravar o problema aqui discutido. Não há na esfera legislativa brasileira, normativos técnicos, seja especificado pela ABNT, pelo CREA ou por lei federal, que possam ser utilizados para subsidiar o projeto de sistemas de CFTV em ambiente orgânicos. Os profissionais que lidam com projetos e serviços de instalação de CFTV dimensionam esses sistemas sem esse amparo técnico normativo, deixando uma grande margem para projetos mal dimensionados, que não suprem a necessidade do cliente, nem a demanda forense de reconhecimento facial de suspeitos e posterior uso deste material como evidência de um crime. Cabe ressaltar que, em muitas ocasiões, o amparo de profissionais técnicos da área nem é utilizado, podendo tornar ainda mais deficiente o sistema de CFTV.

Apesar da ausência da lacuna normativa específica para sistemas de CFTV, há outras normas que que influenciam, mesmo que de forma acessória, em projetos de CFTV. Dentre estes normativos, podem ser citados:

- NBR 5410 - Execução de instalações elétricas de baixa tensão;
- NBR 5419 - Proteção Contra Descargas Atmosféricas;
- NBR 5474 - Eletrotécnica e Eletrônica - conectores elétricos;
- NBR 5471 - Condutores elétricos;
- NBR-5382 Verificação de Iluminação de Interiores;
- NBR 5413 - lluminância de interiores;
- NBR 6854 Aparelhos de iluminação para interiores;

- NBR 6533 - Estabelecimentos dos Efeitos da Corrente Elétrica do Corpo Humano;
- EIA/TIA 606 - *Administration Standard for de Telecommunications Infrastructure of Commercial Buildings*;
- EIA/TIA 607 - *Grounding and Bonding Requirements for Telecommunications In Commercial Building*;
- EIA/TIA TSB-67 - *Transmission Performance Specification for Field Tests*;
- EIA/TIA 568 - *Commercial Building Telecommunications Cabling Standard*;
- NBR 14565 - Procedimentos básicos para elaboração de projetos de cabeamento de telecomunicações para rede interna estruturada;
- NBR 13531 - Elaboração de projetos de edificações - atividades técnicas;
- NBR 10582 - Apresentação da folha para desenho técnico – procedimentos;
- NBR 10068 - Folha de desenho - leiaute e dimensões – padronização;
- NR 10 - Segurança em Instalações e Serviços em Eletricidade - Apresentação do novo texto;
- Práticas SEAP - Governo Federal.

Apresentado o problema de pesquisa, resta dizer que a intenção deste trabalho é minimizar os problemas de instalação de câmeras de segurança encontrados em muitos sistemas de CFTV instalados no país, fornecendo alguns critérios que vão subsidiar a instalação desse equipamento. Ainda, este trabalho tem também o foco forense. Busca-se como resultado da parametrização destes critérios, que muitas imagens ou vídeos coletados ou enviados para a Polícia ou órgãos periciais, derivadas de sistemas de CFTV, tenham condições mínimas para a realização do reconhecimento facial de suspeitos, pois muitos dos casos reais analisados não auxiliam nos processos investigativos, tampouco na realização de exames probatórios que possam ser utilizados como provas no Judiciário.

6.3 – CRITÉRIOS PROPOSTOS

Sistemas de reconhecimento facial podem atingir marcas de 95% de acurácia, utilizando-se imagens de poses frontais com alta resolução espacial, como aquelas encontradas em passaportes. No entanto, o desempenho desses sistemas diminui para uma acurácia entre 10% e 20% em situações reais, utilizando, por exemplo, imagens de sistemas de CFTV, que apresentam deficiência técnicas com relação às condições de posicionamento da

câmera, iluminação ambiente, expressões faciais, oclusão da face, dentre outros (SHAN *et al.*, 2003).

Os critérios propostos para especificações de câmeras de vigilância de sistemas de CFTV, em ambientes *indoor*, com o finco de permitirem o reconhecimento facial de suspeitos foram divididos em três módulos básicos e um módulo especial, conforme mostrado na Figura 6.2. Buscou-se dimensionar critérios que auxiliassem a captar as melhores imagens de indivíduos circulando pelo perímetro de um estabelecimento.

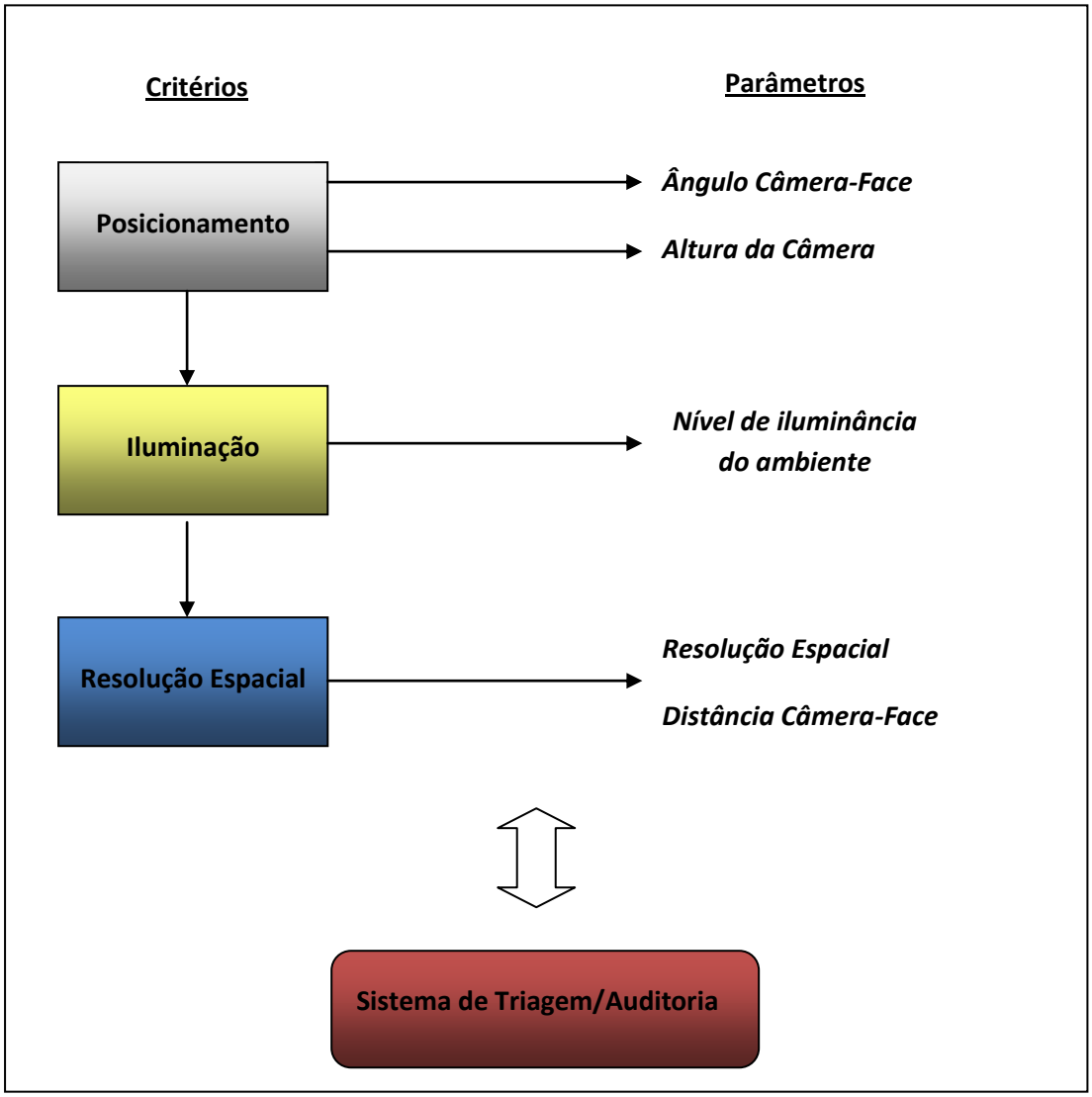


Figura 6.2 – Critérios propostos.

6.3.1 – POSICIONAMENTO DAS CÂMERAS

O critério posicionamento está ligado à posição que uma câmera de vigilância deve ser instalada no ambiente interno, de forma a capturar imagens das pessoas circulando pelo local. Deve-se ter em mente que o posicionamento da câmera deve ser feito buscando adquirir imagens frontais da face (NISSENBAUM e INTRONA, 2010), tanto quanto for possível, pois esta é a posição ideal para realização de reconhecimento facial. Além disso, espera-se que o posicionamento das câmeras capte um volume de espaço mais voltado para a aquisição da face da pessoa do que do ambiente ao redor.

Foram avaliados dois parâmetros no critério posicionamento: altura da câmera de vigilância e ângulo de trabalho da câmera para a face capturada. Os experimentos vão procurar mostrar os limites necessários para o correto dimensionamento desses parâmetros.

Esses parâmetros influenciam no campo de visão da câmera (*FoV*). Recomenda-se que a altura mínima de instalação das câmeras tenha como base as variadas estaturas que de uma pessoa[7] e os posicionamentos típicos esperados na circulação de pessoas em determinado ambiente, andando, correndo, em pé e estáticas ou sentadas e estáticas. O campo de visão das câmeras também não deve ser obstruído por qualquer tipo de objeto ou estrutura do ambiente.

6.3.2 – ILUMINAÇÃO

Este critério está relacionado às condições de iluminação do ambiente monitorado e sua importância reside no fato de que essas condições irão impactar na correta extração das características faciais de um indivíduo, ao ser flagrado por uma câmera de segurança (vide Figura 6.3). O importante é que a imagem esteja adequada para a consecução dos exames. Ambientes com iluminação inadequada, com pouca iluminação (sombreamento) ou com

[7] Segundo dados do IBGE, do ano de 2009, a estatura média dos jovens brasileiros urbanos entre 20 e 24 anos é de 173,1 cm para o homem e de 161,1 cm para a mulher; a média de altura para o homem adulto urbano é de 170,0 cm e para a mulher urbana é de 161,1 cm e a média de altura para o homem rural é de 167,0 cm e para a mulher rural é de 156,0 cm (IBGE, 2009).

excesso de luz (podendo causar saturação), podem degradar a qualidade das imagens captadas pelas câmeras e dificultar o processo de reconhecimento facial de um indivíduo.

Figura 6.3 – Exemplos da influência de variações de iluminância do ambiente na face.

Recomenda-se o uso de um mesmo tipo de lâmpada para adequar a iluminação de um ambiente. A utilização de diferentes tipos de lâmpadas pode causar diferenças nas temperaturas de cor, podendo causar degradação da qualidade da imagem captada. O cenário ideal é aquele em que há iluminação balanceada e que os contrastes faciais fiquem perceptíveis, ou seja, que não haja áreas de sombras na face.

Caso não seja possível melhorar as condições de iluminação do ambiente, recomenda-se utilizar câmeras com recursos especiais, que permitam compensar as condições de iluminação, atuando tanto em condições de baixa luminosidade ou de saturação. Outro recurso disponível em câmeras de vigilância é a utilização de lentes auto íris, que regulam o foco da imagem automaticamente, de acordo com as condições de iluminação do ambiente. O uso deste tipo de lente é adequado para ambientes internos que possuem contato com iluminação natural externa, provocando assim, variações nas condições de iluminação do local. Outra opção para compensar as deficiências da iluminação ambiental é o uso de luz infravermelha (principalmente em câmeras monocromáticas) ou em câmeras que irão atuar dia/noite (*day/night*).

Foram avaliados neste critério, os parâmetros de iluminância do ambiente e tipo de lâmpada a ser utilizado, procurando determinar as condições satisfatórias de trabalho para o funcionamento de aplicativos de reconhecimento facial de indivíduos.

As tabelas apresentadas no Apêndice C mostram os níveis de luz recomendados para determinados ambientes, bem como características de tipos de lâmpadas e informações de luminosidade de acordo com as condições ambientes.

6.3.3 – RESOLUÇÃO ESPACIAL

O critério de resolução espacial está relacionado ao nível de detalhes discerníveis na imagem que irão favorecer o reconhecimento facial de indivíduos. Abaixo de determinados limites em termos de resolução espacial, os sistemas de reconhecimento facial irão falhar em realizar sua função. Quanto maior a resolução espacial, melhor será a percepção da qualidade da imagem e dos detalhes discerníveis da face de um indivíduo.

Neste critério estará sendo avaliado o tamanho da face em relação aos parâmetros de resolução de imagem e da distância entre a posição da câmera e a face capturada. Neste critério em específico, foram utilizados os resultados científicos apresentados na seção 5.4. Segundo o trabalho citado, desde que não haja quaisquer outras deficiências na imagem, para que o reconhecimento facial seja funcional, deve haver pelo menos 40 pixels na largura da face capturada.

As tabelas apresentada no Apêndice D relacionam os principais tipos de resoluções de câmeras analógicas e digitais encontradas no mercado.

6.3.4 – SISTEMA DE TRIAGEM/AUDITORIA

Não há lei ou jurisprudência no Brasil que versa sobre a admissibilidade de arquivos de imagens ou vídeos perante um tribunal criminal. A admissibilidade de provas desta natureza segue o mesmo rito explicado na seção 2.1.1. Desta forma, há algumas exigências formais para a produção de provas criminais, apostas no Código de Processo Penal, como por exemplo, que ela seja coletada e analisada por um perito oficial, servidor público do Estado. Além disso, a assunção dessa prova vai depender de crivo do juiz criminal, podendo aceitá-la ou não, independentemente se foram seguidos os procedimentos técnicos e científicos adequados para a formação probatória, desde que seja fundamentado legalmente porque determinada prova foi recusada.

O objetivo do sistema de triagem/auditoria é fornecer um meio para a garantia da integridade dos arquivos gerados pelo sistema de CFTV, de forma que os arquivos analisados tratem-se realmente daqueles que foram originalmente gerados pelo sistema de CFTV questionado. Para que isso fosse realmente possível, o equipamento de filmagem

deveria gerar um código de autenticação de arquivos na forma criptográfica (*mac – message autenthication code*) intrinsicamente, entretanto, nesse tipo de equipamento para CFTV, não trata-se de uma função disponibilizada pelos fabricantes de câmeras. De forma paliativa, procurando amenizar esse problema, sugere-se uma sistemática envolvendo duas etapas,: uma realizada a cargo dos responsáveis pelo sistema de CFTV e outra realizada pelo servidor do Estado, perito oficial, que estará analisando o material coletado. Essa sistemática não garante a integridade dos arquivos gerados de forma inequívoca, mas adiciona maior confiança ao processo.

Assim, no âmbito do ambiente organizacional é recomendado que a gravação de dados do CFTV estivesse associada a algum tipo de sistema que gere resumos criptográficos[8] dos arquivos gerados (função de *hash*), armazenando os resultados em múltiplos locais ou em um único local seguro. Desta forma, os dados gerados poderiam ser comparados posteriormente, com seus resumos criptográficos previamente calculados, buscando identificar se houve adulterações intencionais. No final de cada dia, deve-se gerar também um resumo criptográfico do grupo de arquivos produzidos naquele dia. Na ocorrência de incidentes criminais, as autoridades policiais devem ser rapidamente acionadas, evitando-se maiores chances de alterações, e a mídia original dos dados deve ser mantida em segurança, seja fisicamente ou logicamente. Após a chegada de peritos no local para a coleta dos dados, a mídia original deve ser recolhida ou espelhada *in loco*, documentada formalmente (apreensão) e então, encaminhada para laboratório de análise pericial, buscando manter em todas as etapas, a cadeia de custódia da prova. É importante consignar os modelos/marcas dos equipamentos que geraram os dados de imagem ou vídeo, pois na ausência deles em procedimentos laboratoriais, deve ser possível identificar de outra forma se os arquivos produzidos/analisados coletados da mídia original poderiam ter sido gerados pelos equipamentos documentados.

A escolha de qual função de *hash* a ser utilizada é facultativa, entretanto, o método utilizado deve ser reproduzível, ou seja, deve ser utilizada uma função pública conhecida

[8] Resumo criptográfico ou *hash* trata-se de uma função que recebe um conteúdo de tamanho variável como entrada e gera um valor de tamanho fixo como saída. Os cálculos envolvidos no *hash* são complexos e muito sensíveis a qualquer variação de bit do conteúdo de entrada. O uso de resumos criptográficos podem ser utilizados para a identificação de um arquivo univocamente.

(critério de reprodutibilidade). Recomenda-se também o uso de funções de *hash* com blocos de saída superiores a 160 bits, pois ainda não foram encontrados no meio científico, métodos para quebrar funções de *hash* desta natureza. Funções de *hash* populares que atendem a estes critérios são aqueles da família *SHA-2* (*Secure Hash Algorithm* – versão 2), como a *SHA-256* (bloco de saída de 256 bits) e *SHA-512* (bloco de saída de 512 bits) (NIST, 2012). A Figura 6.4 apresenta a metodologia de garantia de integridade sugerida.

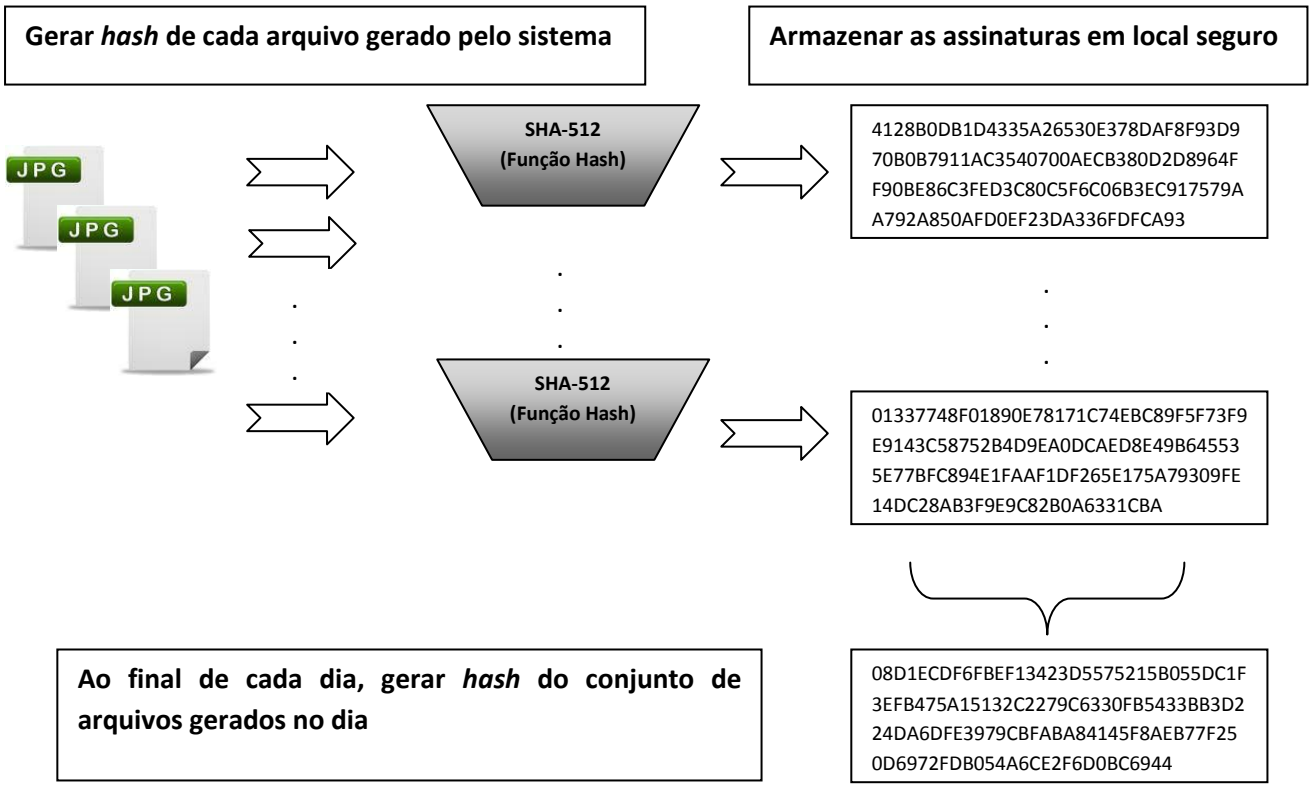


Figura 6.4 – Modelo de sistemática para garantia da integridade dos arquivos gerados em sistemas de CFTV.

No âmbito do laboratório forense, o perito deve realizar novo procedimento para verificação de integridade dos arquivos coletados, ou seja, de cálculo do resumo criptográfico e comparação com aqueles previamente calculados existentes na mídia original, procurando assim, identificar se há qualquer inconsistência nos arquivos coletados. Após este procedimento, os exames são realizados e, caso haja necessidade de tratamento das imagens, como o realce de determinadas regiões, retirada de ruído, ampliações ou reduções, é fundamental a utilização de um software que grave o histórico procedimental realizado, para que possa ser posteriormente reproduzido, como por

exemplo, o Adobe® Photoshop®[9], que permite registro do histórico nos próprios metadados do arquivo ou em arquivo de texto separado. A Figura 6.5 apresenta um exemplo de registro do aplicativo Adobe Photoshop CS 5 de gravação de históricos de edição de uma imagem.

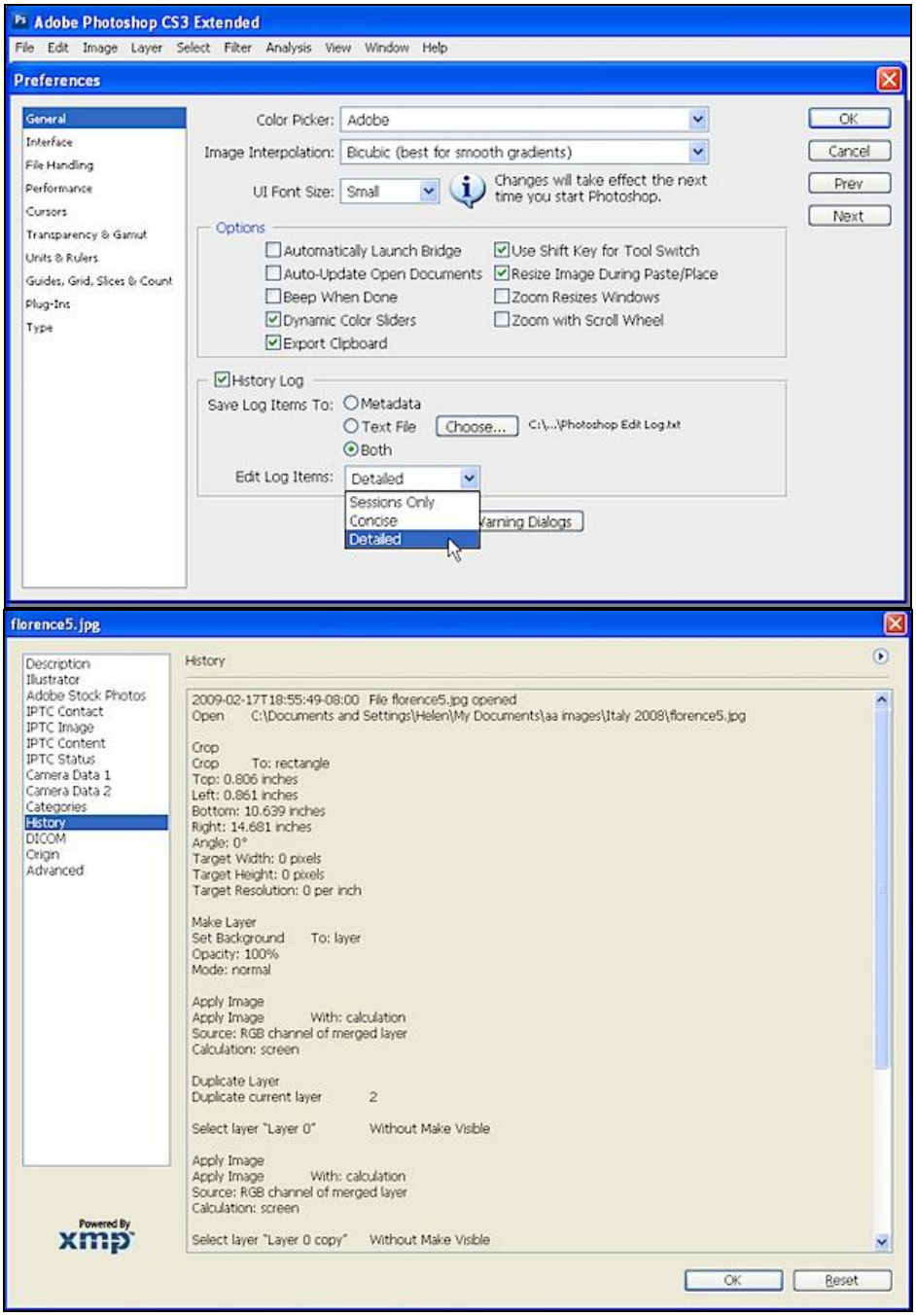


Figura 6.5 – Opção de software de edição de imagem com gravação de histórico de alterações do arquivo (BRADLEY, 2012).

[9] Disponível através da fabricante de software ADOBE (http://www.adobe.com/br).

Soluções alternativas para garantir a integridade de arquivos de imagens envolvem o uso de marcas d'água ou a ausência de formato de arquivo (formato *raw*), mas a disponibilidade destes recursos vai depender do fabricante do equipamento. Alguns fabricantes de câmeras profissionais fornecem como função para seus equipamentos a possibilidade de inserção nativa de marcas d'água nos arquivos de imagens ou vídeos gerados, de forma que caso haja adulterações nos arquivos, essas alterações podem ser detectadas através da análise da marca d'água inserida no arquivo. Outros fabricantes disponibilizam função para seus equipamentos gerarem imagens na forma *raw*, ou seja, sem um formato definido de arquivo. O benefício da utilização desta sistemática é que eles são virtualmente seguros contra alterações, ou seja, qualquer tipo de interferência na cadeia de bits do arquivo é facilmente detectável por especialistas. O inconveniente em trabalhar com esse tipo de sistema é que é necessário um pré-processamento do arquivo para que ele seja apresentado em tela e qualquer tratamento a ser realizado na imagem, não será possível manter essas alterações no próprio arquivo *raw*. Será necessário gerar novos arquivos formatados com algum padrão de compressão de imagem. O arquivo *raw* original permanecerá inalterado (BRADLEY, 2012).

7 – EXPERIMENTOS

Este capítulo descreve os experimentos realizados neste trabalho de pesquisa. Os resultados obtidos são discutidos.

Na seção 7.1 são apresentadas as especificações técnicas em termos de *hardware* e *software*, bem como as ferramentas de reconhecimento facial utilizadas para validar a proposta. Na seção 7.2 são apresentados os experimentos relacionados ao critério de posicionamento. Na seção 7.3 são apresentados os experimentos relacionados ao critério de iluminação. Na seção 7.4 são apresentados os experimentos relacionados ao critério de resolução. Por fim, na seção 7.5 são discutidos os resultados.

7.1 – CARACTERIZAÇÃO DO AMBIENTE

Neste tópico serão apresentados os recursos utilizados para a realização dos experimentos deste trabalho de pesquisa.

7.1.1 – EQUIPAMENTOS

A Tabela 7.1 apresenta uma breve descrição dos principais equipamentos utilizados nos experimentos deste trabalho. No Apêndice F, estão consignadas as especificações técnicas completas destes equipamentos. Foram utilizadas as configurações padrões de fábrica dos equipamentos durante os experimentos.

Tabela 7.1 – Descrição sucinta dos equipamentos utilizados.

Item	Descrição
Sistema 1	Notebook, marca LeNovo, modelo T60.
	Função:
	Gerenciamento do Sistema de CFTV.
	Acessórios:
	Placa de captura de áudio e vídeo, marca Conexant, modelo 3104; Aplicativo SuperDVR.

Sistema 2	Notebook da marca Sony, modelo Vaio VPCF1.
	Função:
	Realização das avaliações de medição dos experimentos
	Acessórios:
	Aplicativo Matlab versão R2011a (*Student Version*).
Câmera Digital	Câmera da marca Sony, modelo DSC-TX100V.
Câmera Profissional CFTV	Câmera da marca Intelbras, modelo 480 S.
Luxímetro	Luxímetro da marca Minipa, modelo MLM-1011.
Lâmpadas	Fluorescentes, cor branca, de diversas potências.

7.1.2 – SOFTWARE PARA RECONHECIMENTO FACIAL

O algoritmo utilizado nos testes foi baseado na técnica PCA (*Principal Component Analysis*), também conhecida como *eigenface*. O algoritmo tem obtido bom desempenho na avaliação de teste de reconhecimento facial utilizando bancos de dados públicos. Os resultados têm variado de acordo com as condições experimentais dos testes.

O funcionamento do método *eigenfaces* é detalhado a seguir (CAMPOS, 2000):

> Considerando $M = h \times w$, o "espaço de imagens", sendo h o número de linhas de uma imagem e w o número de colunas, com um conjunto de n imagens utilizadas para o treinamento do método, Γ, é criada uma matriz contendo n cópias da média das imagens contidas em Γ:

$$\Psi_{[M\times n]} = \frac{1}{n}\sum_{i-1}^{n}\Gamma_{i,[M\times 1]} \tag{7.1}$$

1) A diferença de cada face adicionada Γ_i e a média é gravada em Φ_i:

$$\Phi_{i,[M\times n]} = \Gamma_{i,[M\times n]} - \Psi_{[M\times n]} \tag{7.2}$$

sendo $$A_{[M\times n]} = [\Phi_1, \Phi_2, ..., \Phi_n] \tag{7.3}$$

2) Então, a matriz de covariância das imagens de treinamento é montada com os dados extraídos das imagens anteriormente. Os *eigenfaces* são extraídos desta matriz, **C**, que é construída como se segue:

$$C_{[M\times M]} = \mathrm{A}_{[M\times n]}\mathrm{A}^{\mathrm{T}}_{[n\times M]} \tag{7.4}$$

3) Realizar o cálculo dos auto vetores da matriz $[M \times M]$ de modo simplificado (Turk, 1991). Considere a matriz $L_{[n\times n]} = \mathrm{A}^{\mathrm{T}}_{[n\times M]}\mathrm{A}_{[M\times n]}$. Os primeiros *n* auto vetores de C podem ser expressos como uma combinação linear entre os auto vetores de **L** (denotados por **V**) e as imagens contidas em **A** (Wei, 1998), como mostrado a seguir:

$$U[M \times n] = \mathrm{A}_{[M\times n]}V_{[n\times n]} \tag{7.5}$$

4) Os *eigenfaces* são armazenados na matriz **U**. O reconhecimento de faces por *eigenfaces* é composto por duas fases: uma para o treinamento e outra para o reconhecimento (Wei, 1998). Na fase de treinamento, a face de cada indivíduo é projetada no espaço de imagens, gerando a matriz **Ω**:

$$\Omega_{[n\times n]} = U^{\mathrm{T}}_{[n\times n]}\left(\Gamma_{[M\times n]} - \Psi_{[M\times n]}\right) \tag{7.6}$$

5) A matriz **Ω** será aquela utilizada para a realização das comparações na hora de executar o reconhecimento facial. A dimensão desta matriz é menor do que a das imagens de entrada. Este é o objetivo principal da utilização dos *eigenfaces*.

6) Para a posterior classificação (reconhecimento) das imagens de entrada no sistema de reconhecimento facial, será preciso verificar se as imagens inseridas são realmente imagens de face de pessoas e, em caso positivo, identificar se a face inserida está presente no banco de dados. No método *eigenfaces* são utilizadas duas métricas para determinação desta classificação: a métrica do espaço de faces e a métrica do espaço de classes. Para determinar se uma

imagem de entrada é uma imagem de face, é utilizada a métrica de espaço de faces para verificar o quanto perto a imagem de entrada está do espaço de faces. A distância entre a imagem de entrada e o espaço de faces pode ser calculada utilizando as seguintes equações (TAN, 2005):

$$\in_F^2 = \left\| \Phi - \Phi_f \right\|^2, \tag{7.7}$$

onde $$\left\| \left(\Phi - \Phi_f \right) \right\|^2 = \left(\Phi - \Phi_f \right)^T \left(\Phi - \Phi_f \right), \tag{7.8}$$

com $$\Phi = \Gamma - \Psi \tag{7.9}$$

e $$\Phi_f = \sum_{i=1}^{M'} \omega_i e_i \tag{7.10}$$

O valor de Φ é simplesmente a imagem de entrada média ajustada e o valor de Φ_f é a imagem reconstruída (retirando a face média) da projeção da entrada imagem no espaço de faces gerados pelos vetores (*eigenvectors*). A classificação de uma imagem da face em sua classe é então realizada através da comparação dos pesos de cada elemento do banco de dados de conjunto de treinamento ao conjunto de imagens de teste. O método mais simples para averiguar essa comparação é achar a classe de face *k* que minimiza a distância Euclidiana para o espaço de faces. Antes dessa distância pode ser calculada, a imagem de entrada Γ tem que ser projetada em espaço de faces através da equação (7.11), obtendo-se o conjunto de pesos $\mathbf{\Omega}$. A métrica da classe de faces é definida como:

$$\in_{C,k}^2 = \left\| \Omega - \Omega_k \right\|^2, \tag{7.11}$$

onde $$\left\|\left(\Omega - \Omega_k\right)\right\|^2 = \left(\Omega - \Omega_k\right)^T \left(\Omega - \Omega_k\right), \quad (7.12)$$

e $$1 \leq k \leq M \quad (7.13)$$

onde $\mathbf{\Omega}$ é a matriz de espaço de faces e $\mathbf{\Omega}_k$ é um vetor descrevendo cada classe de face. Assim, dois valores limiares arbitrários, Θ_c e Θ_F são escolhidos para definir a distância máxima permitida a partir de qualquer classe de face (de acordo com a equação 7.11) e a máxima distância permitida a partir do espaço de faces (de acordo com a equação 7.7). Esses valores limiares são calculados para cada imagem de entrada e comparados com os valores limiares escolhidos para o sistema em uso. Há quatro resultados possíveis que indicam o quão perto uma imagem de entrada corresponde a um rosto no conjunto de treinamento (TAN, 2005):

1. Próximo ao espaço de faces e próximo à classe de faces
2. Próximo ao espaço de faces e distante da classe de faces
3. Distante do espaço de faces e próximo à classe de faces
4. Distante do espaço de faces e distante da classe de faces

Para os dois primeiros resultados, a imagem de entrada é reconhecida como uma imagem de face de um indivíduo. No entanto, no primeiro caso, a imagem de entrada é identificada, no segundo, uma face não cadastrada no sistema está presente. Os dois últimos resultados indicam que a imagem de entrada não é uma imagem de face. O terceiro caso trata-se de um falso positivo normalmente encontrado em sistemas de reconhecimento facial.

Especificamente neste trabalho, em virtude do número baixo de imagens utilizado para treinamento e teste, e em virtude do uso de imagens de teste da mesma classe das presentes no treinamento, não foram calculados os *thresholds* para a realização da classificação, logo, as faces são classificadas como pertencendo a uma das classes treinadas, de forma que não há rejeição por parte do sistema.

As Figuras 7.1 e 7.2 apresentam a rotina de funcionamento do método *eingenfaces* e um exemplo de espaço de faces do método para 5 indivíduos, respectivamente.

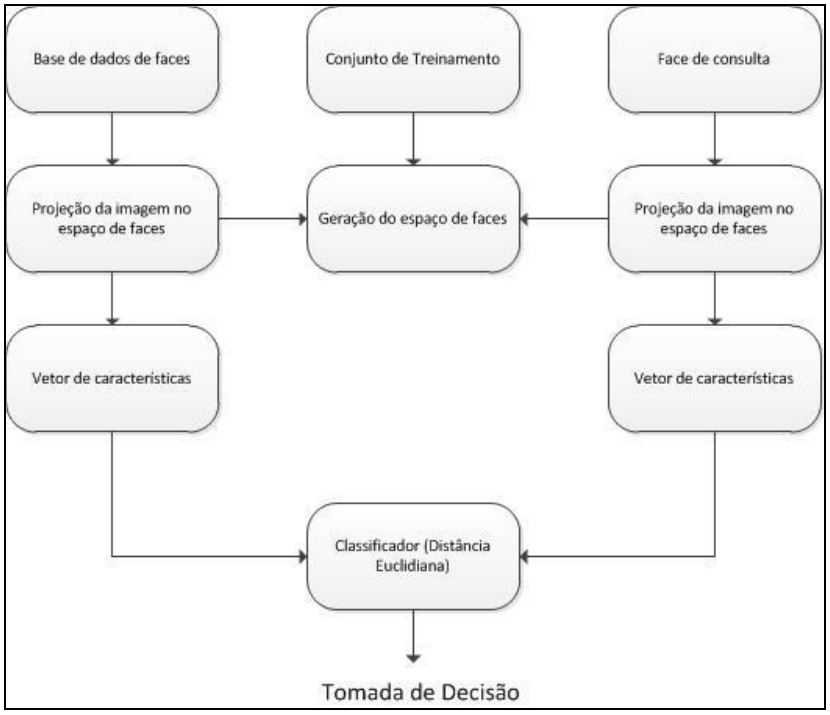


Figura 7.1 – Diagrama do algoritmo PCA para reconhecimento facial (extraído de (PENTEADO, 2009)).

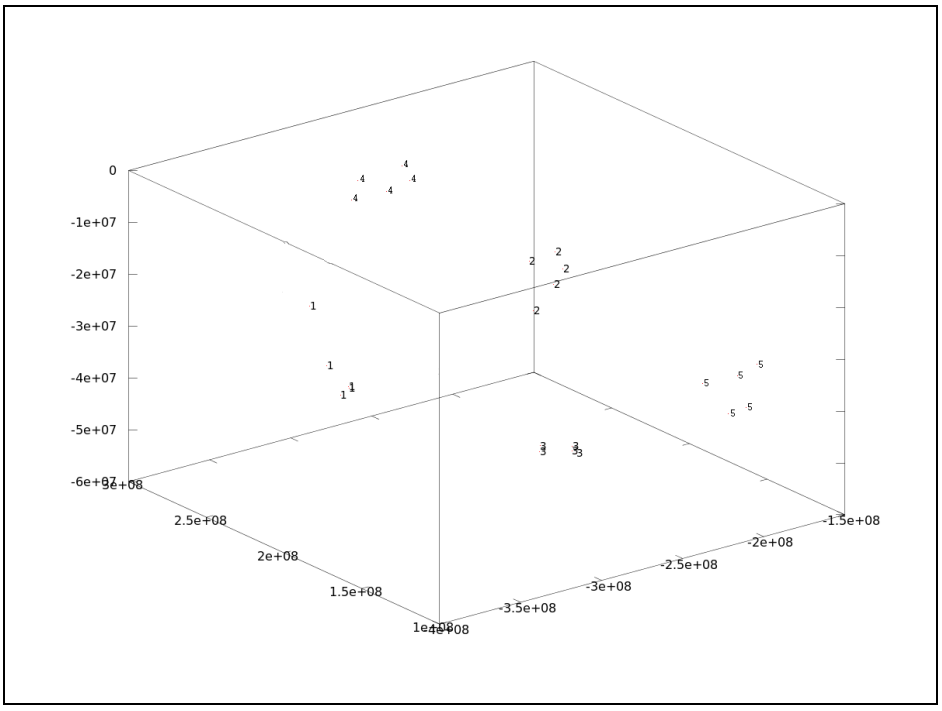


Figura 7.2 – Exemplo de Espaço de faces para 5 indivíduos cadastrados no sistema.

Para a realização dos testes operacionais deste trabalho, foi utilizado um pacote sem restrição de uso experimental[10], baseado no método *eigenfaces*, utilizando o algoritmo de *Karhunen-Loeve*. O código fonte do pacote é executado a partir do aplicativo Matlab. A técnica utiliza um banco de dados de imagens para o treinamento do sistema e, a partir de então, pode-se testá-lo, através do uso de imagens inseridas como entrada do sistema. O código fonte utilizado nos experimentos encontra-se disponível no Apêndice G.

Neste trabalho, não foram utilizados softwares comerciais destinados ao reconhecimento facial, pois, além dos resultados ficarem dependentes da ferramenta proprietária testada, a maioria dos fabricantes de softwares desta natureza realiza pré-processamento das imagens de entrada, seja para treinamento, seja para confronto, buscando normalizar posicionamento, condições de iluminação ambiente, dentre outros fatores, que influenciam no desempenho efetivo do sistema. Desta forma, essas soluções não permitiriam visualizar os reais problemas orgânicos necessários para o dimensionamento de projetos de CFTV, já que mascarariam esses problemas, com esta etapa de pré-processamento.

7.1.3 – BANCOS DE DADOS DE TREINAMENTO

Foram utilizadas diversas fontes de imagens para a realização dos testes experimentais. Para os casos em que havia um banco de dados público de faces disponível, este foi utilizado, não tendo sido produzidas imagens extras. Para os demais casos, foi gerado um banco de dados específico para o experimento.

Antes de descrever cada banco de dados criado ou utilizado, deve-se ressaltar que no Brasil não há banco de dados criminal baseado em faces. A prática comum da identificação criminal é realizada por meio das impressões digitais dos suspeitos presos. Na Polícia Federal, o banco de dados de impressões digitais é organizado em torno do sistema AFIS (*Automated Fingerprint Identification System*). Não há protocolo para inserção de fotos dos suspeitos, em alguns momentos ele é realizado, em outros não; quando é feito, as imagens são realizadas sem critérios específicos. Atualmente estas imagens servem de

[10] O autor do pacote de reconhecimento facial para Matlab é Luigi Rosa. Outras informações podem ser obtidas a partir do sítio WEB http://www.advancedsourcecode.com/face.asp.

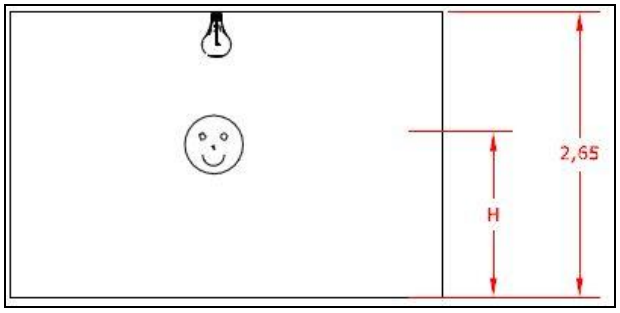
auxílio à polícia como registro individual para identificação dos criminosos para vítimas ou investigadores. Futuramente, desde que haja critérios para geração das imagens, elas podem vir a ser utilizadas para formação de um banco de dados criminal baseado em faces.

Foram criados neste trabalho, 2 (dois) bancos de dados próprios, doravante denominados como a seguir:

- **Passaporte**
- **Customizado 1**

O banco de dados ***Passaporte*** foi criado buscando simular o uso de banco de dados de passaportes para o reconhecimento facial de suspeitos. Foi utilizado uma câmera digital compacta, da marca Sony, modelo DSC-TX100V, sobre um suporte, à uma altura de 1,65 metros, posicionada à 1,70 metros de distância dos indivíduos, utilizando o recurso de *zoom* para focalizar a região dos ombros à face, sem o uso de *flash*. A resolução máxima da câmera foi utilizada nas fotos (16,2 *Megapixels*). As imagens geradas apresentavam alta relação sinal/ruído (SNR), contraste apropriado e alto nível de detalhes. As paredes do ambiente eram brancas. O cenário montado está apresentado nas Figuras 7.3 e 7.4.

Vista frontal



Vista Lateral

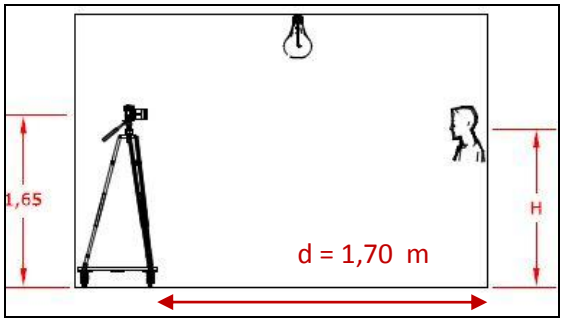


Figura 7.3 – Cenário montado para o banco de dados Passaporte em 2D.

Figura 7.4 – Cenário montado para o banco de dados Passaporte em 3D.

Neste banco de dados foram utilizadas uma única imagem frontal, colorida, de 5 (cinco) pessoas, cujas informações de altura e idade estão apostas nos Gráficos 7.1 e 7.2. Como exigência para utilização do método *eigenfaces* e por consequência, do pacote para Matlab, foi necessário realizar um pré-processamento em cada imagem gerada. As imagens para treinamento e confronto devem ter o mesmo tamanho. Como foi utilizado um banco de dados de faces público da AT&T para complementar as imagens gravadas no banco de dados Customizado 1, cujas imagens possuem dimensões 92 x 112 pixels, todas as imagens geradas nos demais banco de dados foram redimensionadas para este tamanho. Ademais, como o módulo para Matlab não realiza a etapa de detecção facial, mas apenas a etapa específica de reconhecimento facial, cada imagem gerada foi manualmente cortada para que a imagem final contivesse apenas a região de contorno da face do indivíduo, além de terem sido convertidas para a escala de cinza. As Figuras 7.5 e 7.6 apresentam as imagens originais geradas e as imagens pré-processadas, respectivamente.

Neste trabalho não foram implementadas as etapas de pré-reconhecimento facial, normalmente aplicadas a este tipo de pesquisa, como a calibração da câmera e a detecção facial em imagens, pela diversidade de fontes utilizadas nos testes e pela complexidade e tempo gasto a ser envolvido na implementação de tal procedimento, normalmente utilizando a biblioteca OpenCV e o método de detecção de faces viola-jones.

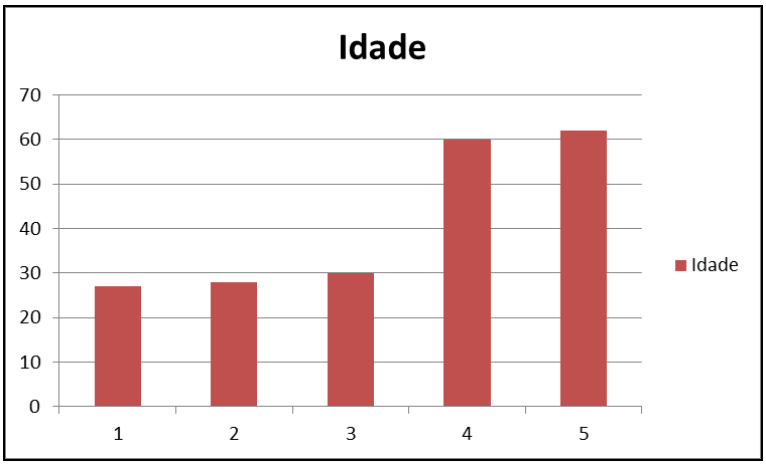


Gráfico 7.1 – Informações de idade dos indivíduos.

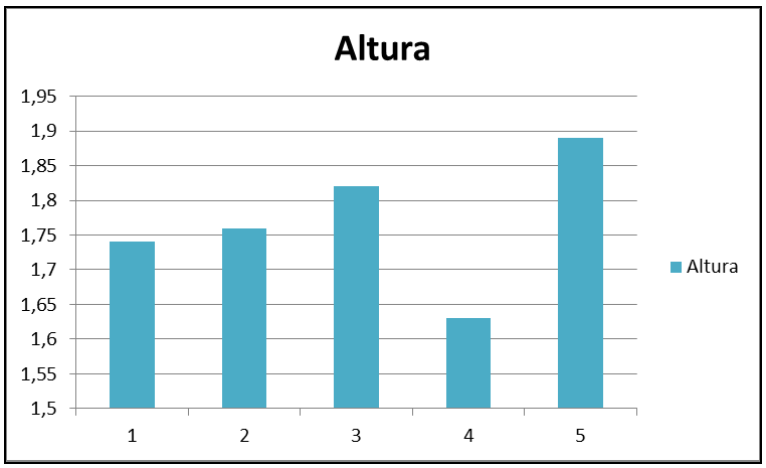


Gráfico 7.2 – Informações de altura dos indivíduos.

Figura 7.5 – Imagens originais utilizadas para compor o banco de dados Passaporte.

Figura 7.6 – Imagens do banco de dados Passaporte pré-processadas manualmente.

O banco de dados ***Customizado 1*** foi criado buscando simular um banco de dados criminal contendo muitas imagens de suspeitos. Foram utilizados os mesmos recursos e o mesmo ambiente do banco de dados anterior. Entretanto, desta vez, foram selecionadas 10 (dez) imagens para cada um dos cinco indivíduos, variando posicionamento da face e expressão facial. Para compor o banco de dados, foi utilizado também o banco de dados público de faces da AT&T[11], composto por 400 (quatrocentas) imagens de 40 (quarenta) indivíduos, sendo 10 (dez) imagens de cada indivíduo em diferentes ângulos de posicionamento da face. A Figura 7.7 apresenta as imagens do banco de dados da AT&T.

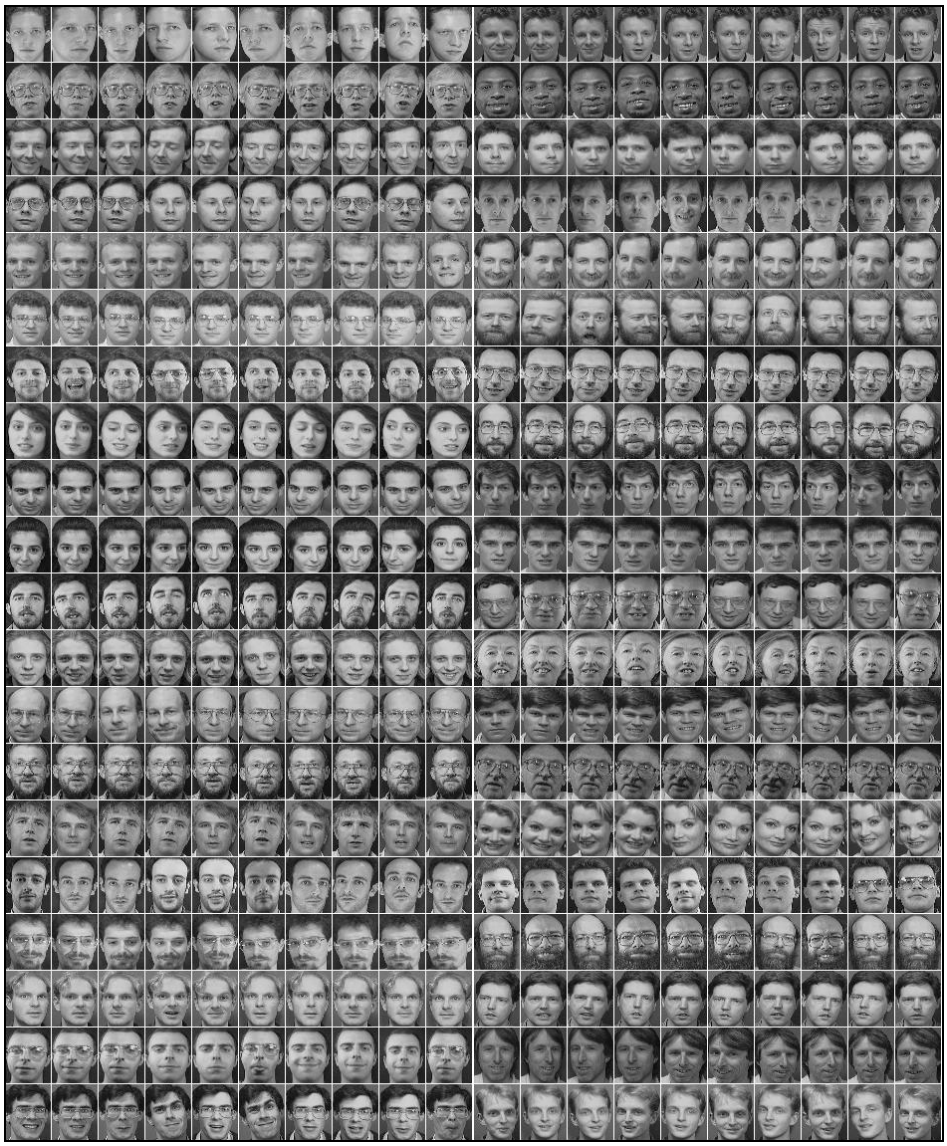


Figura 7.7 – Banco de dados de faces da AT&T.

[11] Disponível em http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

Por fim, foi utilizado também o banco de dados de imagens da face público *ICPR Workshop benchmark database*[12], publicado em 2004, que disponibiliza imagens de 15 (quinze) pessoas, em 93 (noventa e três) diferentes poses da face, dentre eles diferentes ângulos de posicionamento facial, especificando qual o ângulo utilizado em cada imagem. A Figura 7.8 apresenta imagens de um dos indivíduos presentes no banco de dados do ICPR Workshop.

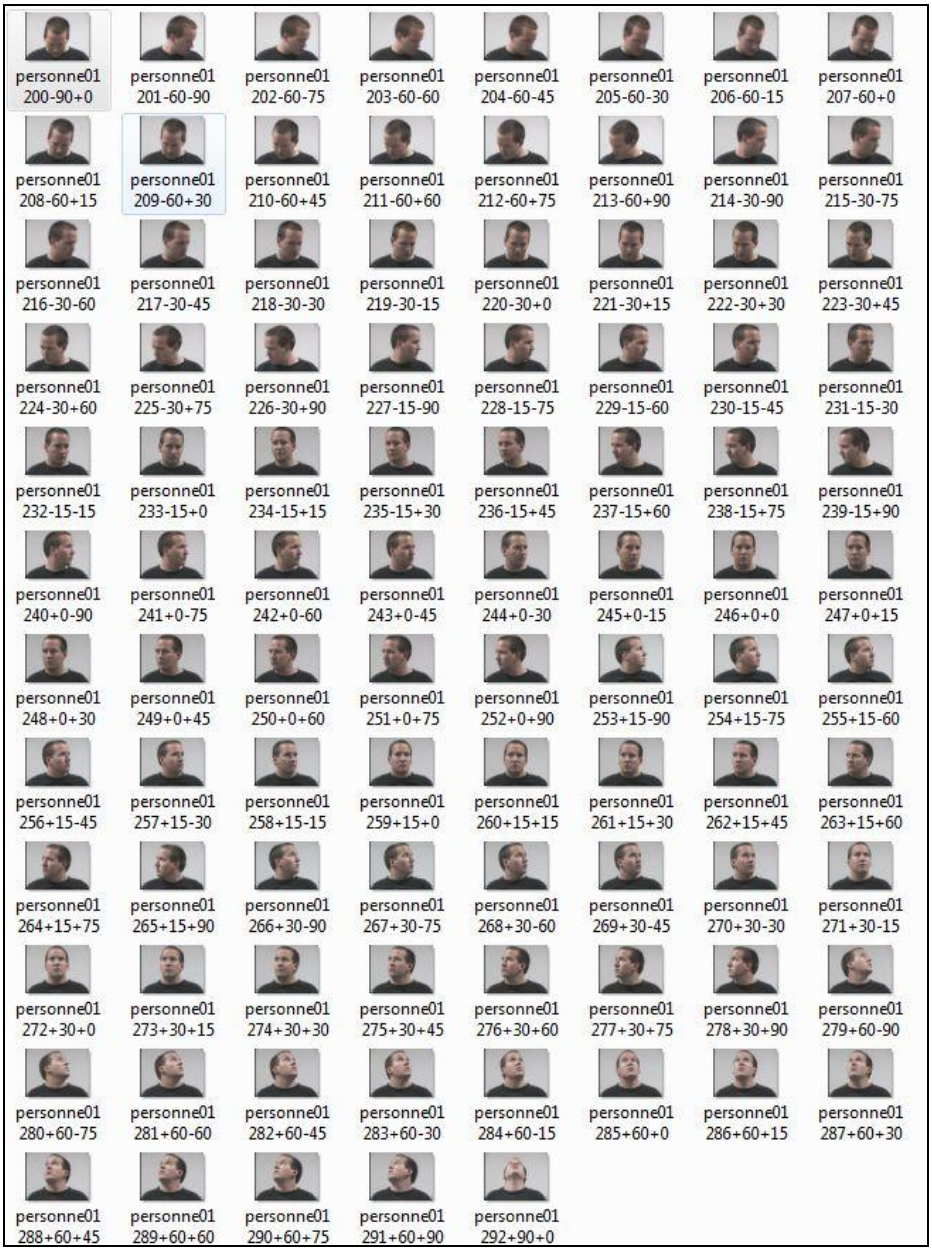


Figura 7.8 – Imagens extraídas de um indivíduo presente no Banco de dados de faces do ICPR.

[12] Disponível em http://www-prima.inrialpes.fr/Pointing04/data-face.html.

7.1.4 – MÉTRICAS DE AVALIAÇÃO

Há muitas formas de calcular o desempenho de sistemas biométricos. No caso de sistemas biométricos de identificação, intuito dos testes de reconhecimento facial a serem experimentados neste trabalho, o problema é de comparação "um para muitos (1:N)". Assim, a face submetida à apreciação é comparada com todas as outras presentes em um banco de dados.

A avaliação de sistemas biométricos de identificação pode ser feita através da taxa de reconhecimento ou de identificação, método conhecido como *Rank-One* (BOLLE *et. al*, 2003). Esta taxa compreende o porcentual de submissões feitas ao sistema biométrico que são corretamente identificados (ZHAO et al., 2003). O método funciona basicamente da seguinte forma:

- A característica biométrica (face) submetida ao sistema é comparada com todos os itens cadastrados no banco de dados do sistema;

- O sistema seleciona a combinação com maior similaridade, de acordo com o técnica aplicada (neste caso será a técnica *PCA*);

- Se a face submetida a teste pertence à mesma pessoa identificada no banco de dados, o sistema identificou corretamente a característica biométrica;

- A taxa de reconhecimento *Rank-one* é o porcentual das características biométricas que o sistema definiu como uma correspondência correta.

A equação da taxa de reconhecimento pode ser expressa por:

$$Rank - One = \frac{N_C}{N_S} \tag{7.14}$$

Onde, N_C corresponde ao número de submissões corretamente identificadas e N_S corresponde ao número total de submissões. O valor da taxa de reconhecimento varia no intervalo de [0, 1], podendo ser expresso percentualmente.

7.2 – EXPERIMENTOS: CRITÉRIO DE POSICIONAMENTO

Foram avaliados três parâmetros no critério posicionamento: ângulo de trabalho da câmera para a face capturada, altura da câmera de vigilância e a cadência de captura de imagens pela câmera. Os experimentos procuram demonstrar os limites necessários para o correto dimensionamento desses parâmetros.

O parâmetro de cadência de captura de imagens foi avaliado inicialmente antes do início dos demais experimentos. A necessidade desta informação reside no ponto de que as câmeras de CFTV devem capturar de maneira adequada os movimentos gestuais realizados por um indivíduo. Em geral, sistemas de DVR apresentam uma taxa de captura de quadros que são divididos dentre o número de câmeras do sistema. A título de exemplo, sistemas DVR com taxa de captura de 100 quadros/segundo (ou *frame per second* – fps) e capacidade de quatro câmeras de segurança permite dividir dentre as câmeras uma taxa de captura de 25 quadros/segundo.

Utilizando o sistema de CFTV montado para os experimentos deste trabalho, foi realizado filmagens de pessoas circulando por um corredor, mas variando a taxa de captura de quadros para a câmera de CFTV em 1 fps, 3 fps, 5 fps, 10 fps, 15 fps, 20 fps e 30 fps. Os resultados experimentais mostraram que até 3 fps, os movimentos gestuais normais de uma pessoa andando por um corredor foram capturados de forma robotizada e não completa, ou seja, alguns dos movimentos realizados pela pessoa monitorada não eram gravados. A partir de 5 fps, foi observado que as imagens gravadas capturavam a totalidade dos movimentos de uma pessoa andando, sem que tenham sido realizados movimentos gestuais bruscos. Para capturar movimentos gestuais bruscos, recomenda-se que a taxa de captura de quadros seja superior a 5 fps, evitando-se também desta forma, o efeito *motion blur*, que dificulta o reconhecimento facial.

Para a avaliação dos demais critérios de posicionamento, foi utilizado o banco de dados público *ICPR Workshop benchmark database*. O protocolo de geração das imagens para este banco de dados pode ser encontrado em seu próprio sítio na Internet (ICPR, 2004). Foram utilizadas imagens de 5 indivíduos variando o posicionamento da face em relação aos ângulos vertical e horizontal. As imagens apresentam todas resolução de 384 x 288 pixels e são coloridas. Foi necessária a realização de conversão das imagens para escala de

cinza, para utilização do aplicativo de testes de reconhecimento facial. Foi utilizado uma imagem frontal de 15 (quinze) indivíduos para compor o banco de dados de treinamento, dentre eles estão os 5 (cinco) indivíduos que serão utilizados para os testes comparativos. Para a avaliação do critério angular, foram utilizadas 9 (nove) imagens de cada indivíduo selecionado, variando posicionamento da face em relação ao ângulo vertical, nos ângulos de -90º, -60º, -30º, -15º, 0º, +15º, +30º, +60º e +90º e 13 (treze) imagens de cada indivíduo variando posicionamento da face em relação ao ângulo horizontal, no ângulos -90º, -75º, -60º, -45º, -30º, -15º, 0º, +15º, +30º, +45º, +60º, +75º e +90º. As Figuras 7.9 e 7.10 abaixo apresentam uma amostra das imagens utilizadas no experimento angular.

Figura 7.9 – Imagens frontais utilizadas para o banco de dados de treinamento, extraídas da base de dados do ICPR (ICPR, 2004).

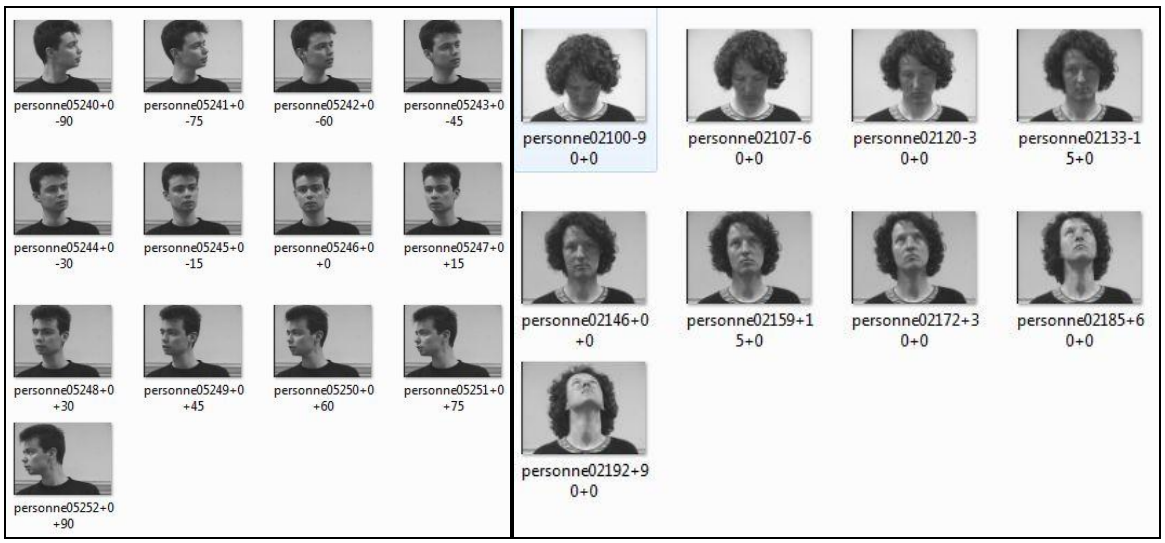


Figura 7.10 – Amostra de imagens angulares verticais e horizontals da face utilizadas para o banco de dados de confronto, extraídas da base de dados do ICPR (ICPR, 2004).

Os Gráficos 7.3 e 7.4 a seguir apresentam os resultados *Rank-One* para os parâmetros avaliados. Foram utilizados como limiar os seguintes parâmetros:

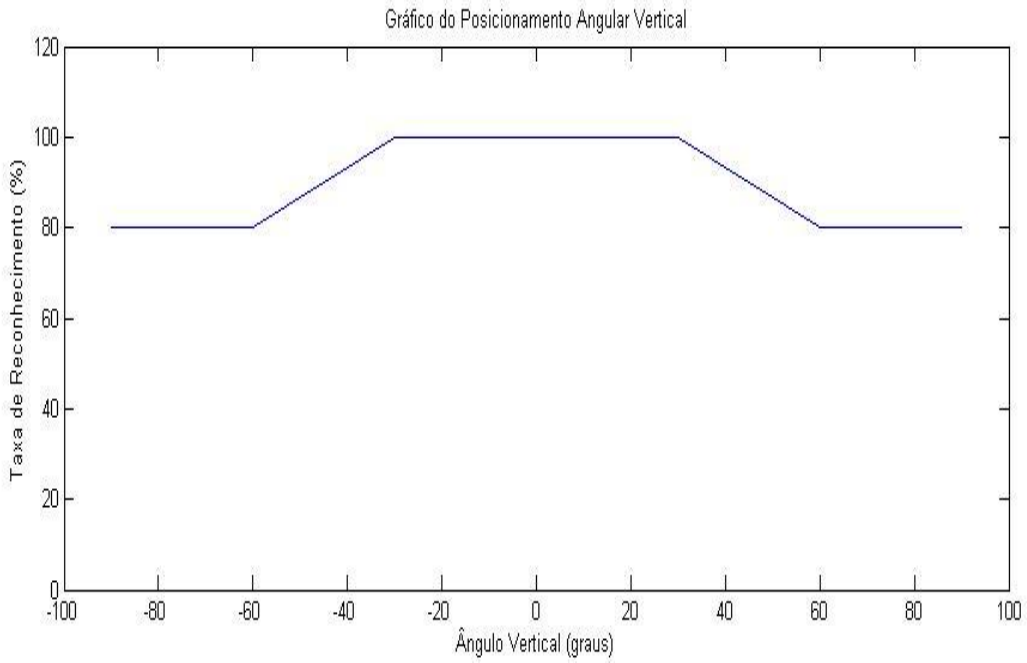


Gráfico 7.3 – Gráfico gerado durante os testes do posicionamento angular vertical.

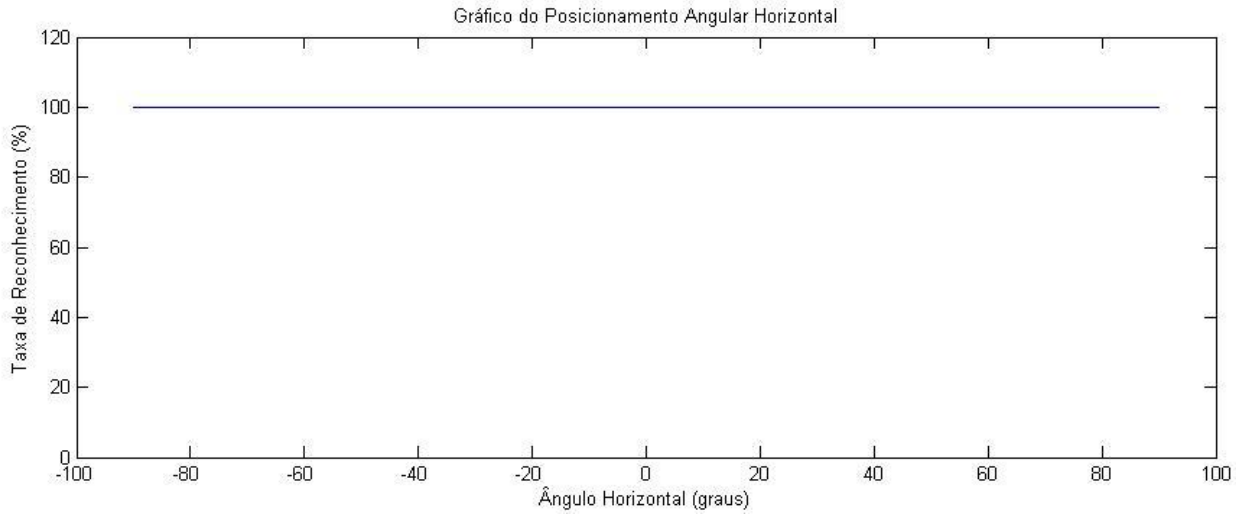


Gráfico 7.4 – Gráfico gerado durante os testes do posicionamento angular horizontal.

O parâmetro de altura da câmera foi avaliado na seção 7.4 deste trabalho.

7.3 – EXPERIMENTOS: CRITÉRIO DE ILUMINAÇÃO

Neste critério, foram avaliados os parâmetros de iluminância do ambiente, procurando determinar os limites críticos que dificultam o reconhecimento facial de um indivíduo, e o tipo de lâmpada a ser utilizado. Desta forma, para a avaliação destes parâmetros, foram capturadas imagens da face de um indivíduo dentre aqueles cadastrados no banco de dados de treinamento, variando-se os níveis de iluminância (em lux) do ambiente. As imagens foram geradas utilizando uma câmera analógica profissional para CFTV, da marca Intelbras, modelo 480 S, posicionada à 1,65 metros de altura, entre 2 a 3 metros do indivíduo, com resolução de 640 x 480 pixels. O ambiente utilizado possuía paredes brancas e dimensões de 3,60 metros de comprimento, 2,40 metros de largura e 2,50 metros de altura (pé direito). Os experimentos foram realizados com os indivíduos sentados (para estabilização da imagem, posicionamento das faces abaixo da perpendicular da câmera e controle dos parâmetros avaliados), buscando simular diferentes posicionamentos faciais encontrados em um cenário real de CFTV, à diferentes distâncias. Para controlar o parâmetro de iluminância do ambiente, foram utilizadas lâmpadas, na cor branca, em diversas quantidades e diferentes potências, apostas em um suporte próximo à face do indivíduo capturado. Foi utilizado um luxímetro para identificar o nível de iluminância próximo à face. O cenário montado é semelhante ao mostrado na Figura 7.11.

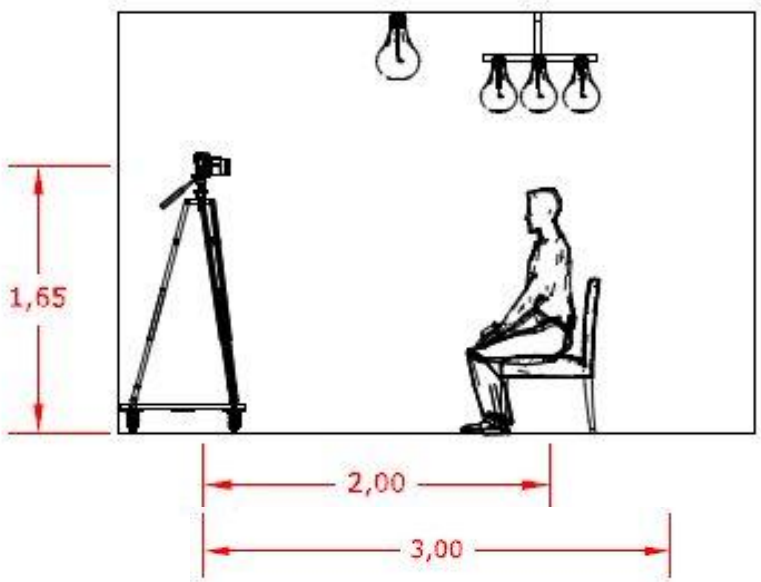


Figura 7.11 – Cenário montado para testes do critério de iluminação. Buscou-se variar a iluminância do ambiente, capturando fotos da face do indivíduo em diferentes níveis de iluminância próximo à face, à diferentes distâncias da câmera, além da utilização de diferentes tipos de lâmpadas.

São mostradas na Figura 7.12 seguir algumas das imagens usadas nos testes, gravadas durante a realização dos experimentos. Os arquivos foram originais foram cortados nas dimensões de 92 x 112 pixels, para manter a compatibilidade com as demais imagens utilizadas neste experimento.

Figura 7.12 – Face capturada com diferentes níveis de iluminância (1 lux / 10 lux / 110 lux / 600 lux / 7100 lux) .

O gráfico a seguir apresenta os resultados *Rank-One* para o parâmetro testado.

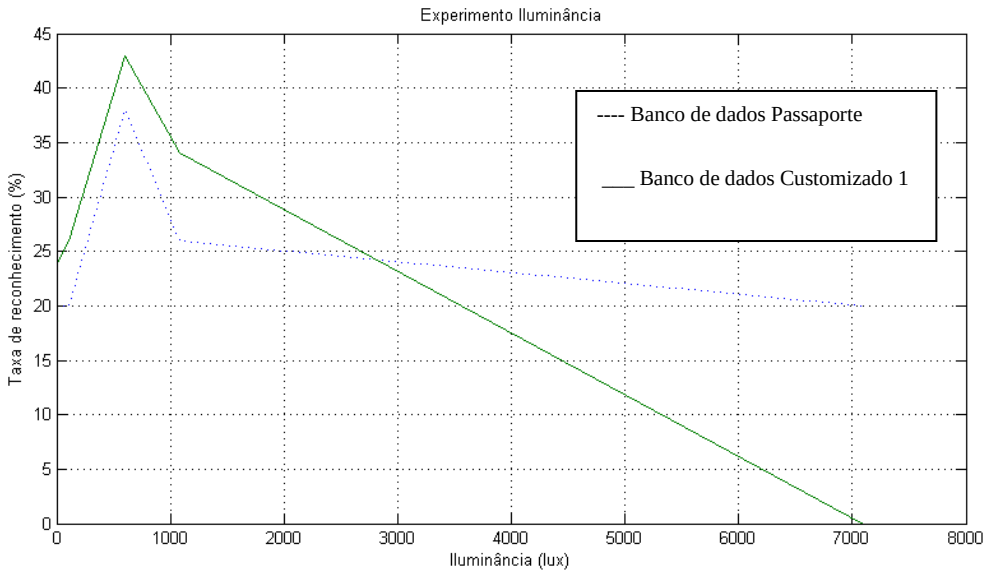


Gráfico 7.5 – Gráfico gerado durante os testes dos níveis de iluminância do ambiente.

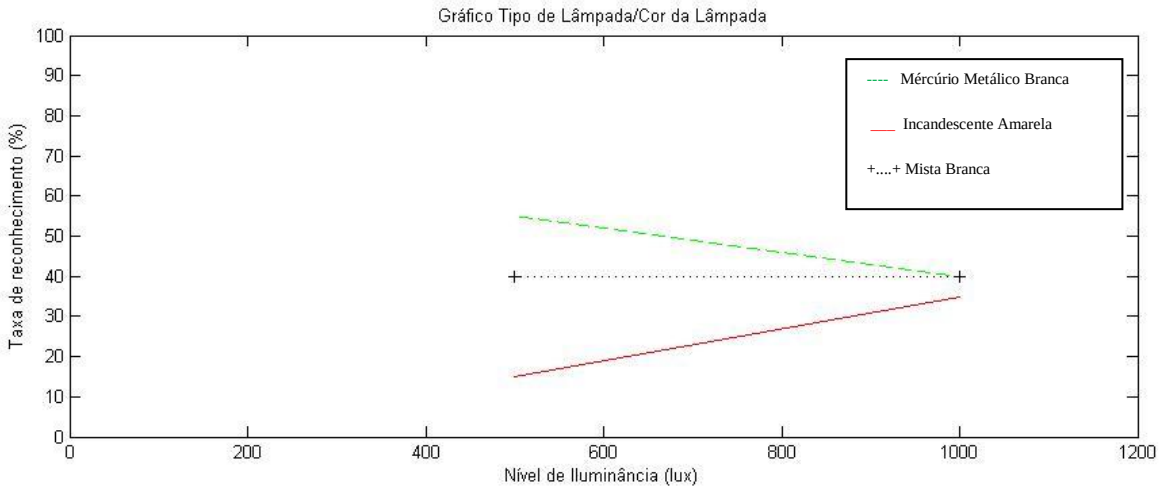


Gráfico 7.6 – Gráfico gerado durante os testes do tipo de lâmpada a ser utilizada no ambiente.

7.4 – EXPERIMENTOS: CRITÉRIO DE RESOLUÇÃO ESPACIAL

Neste critério estará sendo avaliado o aspecto de tamanho da face mínimo necessário para aplicações de reconhecimento facial, em relação aos parâmetros de resolução da câmera/equipamento de gravação e da distância máxima horizontal permitida entre a câmera e a face capturada. A Figura 7.13 ilustra os parâmetros avaliados neste experimento.

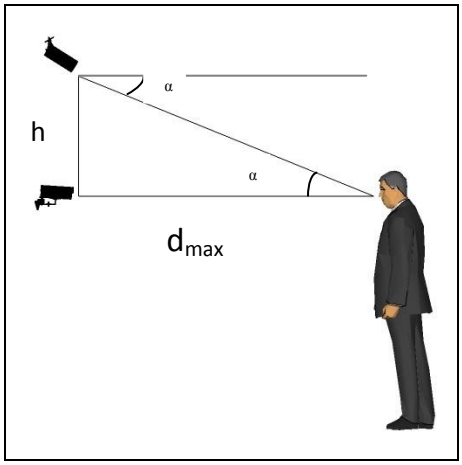


Figura 7.13 – Parâmetros sob avaliação.

Os resultados calculados neste critério utilizaram como base o trabalho científico publicado por Peter Kovesi (KOVESI, 2009). Segundo o autor, em aplicações de

reconhecimento facial, a largura da face na imagem deve ter pelo menos **40 pixels** em nível de detalhe da imagem, para que seja possível o procedimento de reconhecimento facial de suspeitos, considerando que não haja outros problemas na imagem, como ruídos, baixo contraste, *motion blur*, dentre outros.

Para a realização deste experimento, o sistema montado neste projeto foi utilizado, posicionando-se a câmera de CFTV à 0º em relação à face capturada (mesmo eixo horizontal e vertical), utilizando distância focal de 6 mm e sensor de imagem de 1/3", medindo-se então, o tamanho da face da imagem capturada à determinadas distâncias de posicionamento da face do indivíduo. Apesar de a câmera suportar maiores resoluções, por limitações da placa de captura de vídeo, foram analisadas duas resoluções espaciais: 320 x 240 e 640 x 480 (vide Figuras 7.14 e 7.15).

Figura 7.14 – Imagens da face na resolução de 320 x 240 a diferentes distâncias de captura da face (1m / 3 m / 5m / 7m). Todas foram redimensionadas para percepção do nível de pixels na largura da face.

Figura 7.15 – Imagens da face na resolução de 640 x 480 a diferentes distâncias de captura da face (1m / 3 m / 5m / 7m). Todas foram redimensionadas para percepção do nível de pixels na largura da face.

Tabela 7.2 – Tabela das distâncias e largura da face calculadas na resolução de 320 x 240 da imagem gravada.

Distância (em metros)	Largura da Face observada (em pixels)
7,0	10
6,5	11
6,0	12
5,5	13
5,0	14
4,5	15
4,0	16
3,5	19
3,0	24
2,5	27
2,0	30
1,5	40
1,0	50
0,5	90
0,0	180

Observe na Tabela 7.2 que, para que a largura da face tenha 40 pixels em uma imagem com resolução de 320 x 240, 1/8 do número de pixels horizontais da resolução da imagem será ocupado com a largura da face observada. A distância máxima permitida para que uma pessoa possa ser reconhecida, utilizando-se câmeras com essa resolução é de 1,5 metros de distância horizontal entre câmera-face.

Tabela 7.3 – Tabela das distâncias e largura da face calculadas na resolução de 640 x 480 da imagem gravada.

Distância (em metros)	Largura da Face observada (em pixels)
7,0	19
6,5	20
6,0	21

5,5	22
5,0	23
4,5	29
4,0	32
3,5	35
3,0	40
2,5	50
2,0	60
1,5	80
1,0	100
0,5	160
0,0	340

Observe na Tabela 7.3 que, para que a largura da face tenha 40 pixels em uma imagem com resolução de 640 x 480, apenas 1/16 do número de pixels horizontais da resolução da imagem será ocupado com a largura da face observada, aumentando o campo de trabalho de câmeras com esta resolução para 3 metros de distância horizontal entre câmera-face.

Foi observada uma relação linear entre a resolução da imagem gravada e a distância máxima horizontal entre câmera-face permitida para aplicações de reconhecimento facial. Essa relação é válida para as duas resoluções espaciais testadas. A fórmula abaixo apresenta a relação estabelecida.

$$D_{MAX} = \left(\frac{\mathrm{Re}\, s_{HORIZONTAL}}{320} \right) \times 1{,}5 \qquad (7.15)$$

A altura a ser posicionada a câmera de CFTV pode ser calculada através do ângulo de posicionamento e da distância máxima horizontal câmera-face já apresentada. O parâmetro h mostrado na Figura 7.13 indica a distância entre a face da pessoa e a posição da câmera, verticalmente. Assim, a altura H da câmera é:

$$H = h + h' \tag{7.16}$$

onde, H é a altura da câmera, h é a altura câmera-face e h' é a altura média considerada para as pessoas sendo monitoradas. O parâmetro h pode ser calculado por:

$$\tan\alpha = \frac{h}{D_{MAX}} \tag{7.17}$$

As tabelas a seguir sintetizam os possíveis valores de h e H, considerando os limites de D_{MAX} para câmeras de resolução 320 x 240 e 640 x 480. Dados do IBGE de 2009[13] foram utilizados para o cálculo da altura média do brasileiro.

Tabela 7.4 – Tabela para cálculo da altura da câmera de CFTV para resoluções de 320 x 240, na distância limite D_{max} da resolução.

α (graus)	**h (m)**	**H (m)**
10º	0,26	≈ 2,00
20º	0,55	≈ 2,30
30º	0,87	≈ 2,60
45º	1,50	≈ 3,23
60º	2,59	≈ 4,33

Tabela 7.5 – Tabela para cálculo da altura da câmera de CFTV para resoluções de 640 x 480, na distância limite D_{max} da resolução.

α (graus)	**h (m)**	**H (m)**
10º	0,53	≈ 2,30
20º	1,10	≈ 2,80
30º	1,73	≈ 3,45
45º	3,00	≈ 4,70
60º	5,20	≈ 6,90

[13] Altura média do homem urbano em 2009 é de 1,73 metros (IBGE, 2009).

Com base nos resultados acima apurados, a resolução espacial mínima adequada para utilização de câmeras de CFTV para aplicações de reconhecimento facial, considerando-se as resoluções espaciais padrões encontradas neste mercado, é de 640 x 480 pixels. Resoluções espaciais abaixo deste limiar exigem que a face capturada esteja muito próxima da câmera, ocupando grande parte do quadro de imagem.

A Figura 7.16 apresenta um exemplo de posicionamento da câmera, para as configurações de câmera: distância focal de 6 mm, resolução de 640 x 480 pixels, sensor de imagem de 1/3'' e altura de instalação de 2,8 metros.

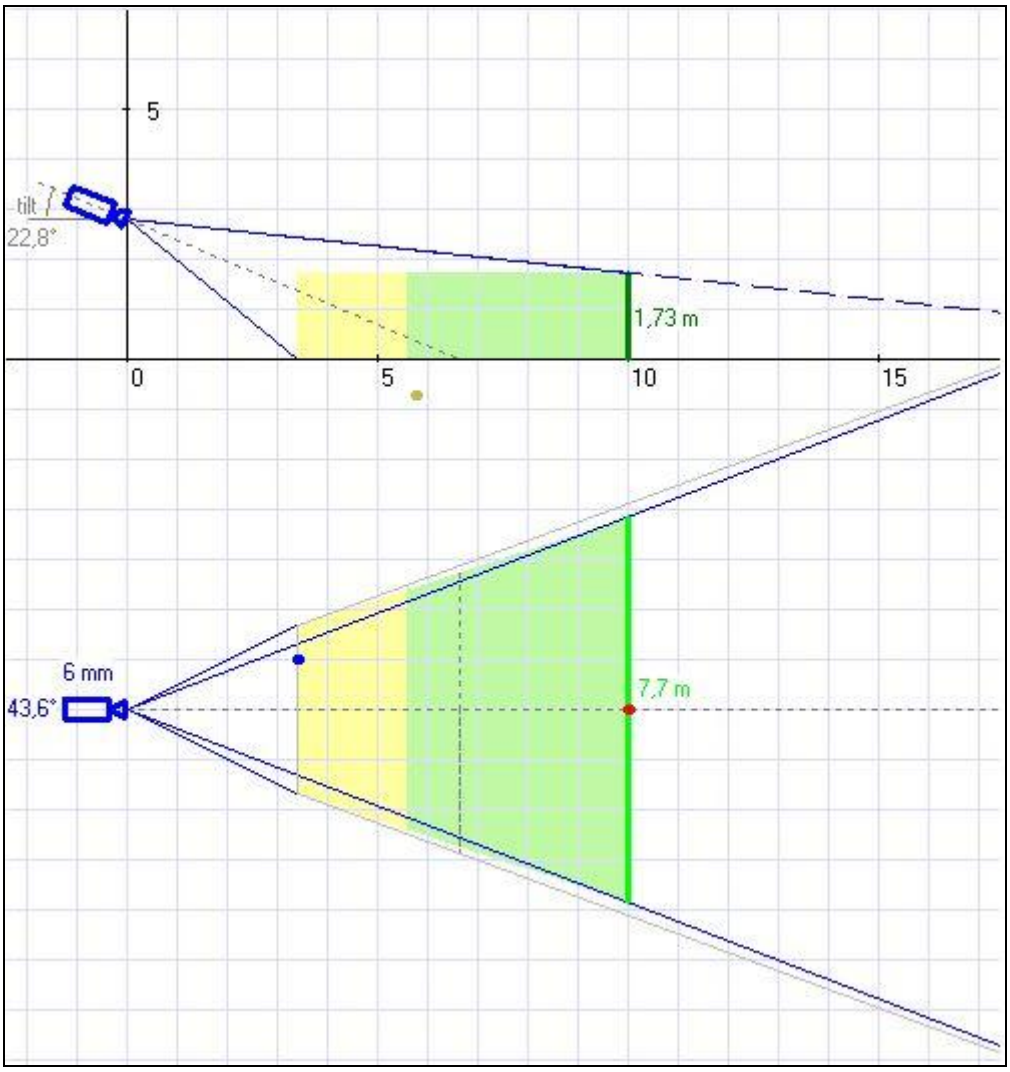


Figura 7.16 – Exemplo de posicionamento da câmera de segurança em projetos de CFTV.

7.5 – DISCUSSÃO DOS RESULTADOS

Os resultados obtidos nos experimentos são discutidos, com base nos critérios definidos na proposta. Procurou-se ao máximo nos experimentos, simular as condições reais encontradas em sistemas de CFTV, bem como no cadastramento de imagens em bancos de dados criminais.

No critério de iluminação, procurou-se avaliar os níveis limites de iluminância do ambiente que permitissem níveis satisfatórios de desempenho em sistema de reconhecimento facial, bem como o tipo adequado de lâmpada a ser utilizada. Foram realizados experimentos sob condições de iluminância do ambiente de 0 lux a 7100 lux, medidos na área próxima à face do indivíduo. Os resultados mostraram que a 0 lux, o sistema de reconhecimento facial não funciona, por não encontrar face nas imagens. Esse resultado já era esperado, pois a câmera possui um limite interno de funcionamento para determinado nível de iluminância. À 0 lux, a câmera utilizada nos experimentos simplesmente não capta imagens além de níveis de preto. À 7100 lux, as imagens apresentam saturação ao nível das faces capturadas, o que também dificulta a aplicação do sistema de reconhecimento facial. O sistema, baseado no método PCA (*eigenface*) funcionou adequadamente para intensidades de iluminação variando entre 110 lux a 1080 lux, parâmetros limites para captação de imagens funcionais aos sistemas de reconhecimento facial.

No critério de posicionamento, foram avaliados a cadência de captura de quadros e os ângulos de trabalho da câmera de segurança. Os resultados mostraram que, para a captura completa dos movimentos de pessoas andando por ambientes internos, deve-se utilizar taxas de pelo menos 5 fps por câmera. Abaixo desse limite, os movimentos não eram capturados por completo (movimentos robotizados). Com relação ao parâmetro ângulo vertical, o sistema de reconhecimento facial testado neste trabalho funcionou bem capturando imagens com ângulo vertical entre face-câmera variando -60º e 60º. Em se tratando do ângulo horizontal, os resultados apurados apontaram ser adequada a utilização de imagens com ângulo horizontal face-câmera variando entre -90º a 90º.

Os parâmetros encontrados na definição do critério de resolução espacial demonstram que, utilizando câmeras e gravadores DVR ou placas de captura nas resoluções de 320 x 240 pixels e 640 x 480 pixels, para que os sistemas de CFTV sejam capazes de fornecer

imagens com condições mínimas para o reconhecimento facial de suspeitos, a distância horizontal entre o suspeito e a câmera é restrita, à respectivamente, 1,5 metros e 3,0 metros. Estas distâncias são aplicadas para situações em que não há outras deficiências nas imagens geradas, como ruídos, baixa taxa de contraste e taxa de sinal ruído, dentre outros, que, caso existam, reduzirão a distância apresentada, para que o sistema mantenha-se funcional. Além disso, os resultados apresentados tomaram por base o uso de lentes com distância focal de 6 mm e sensor de imagem de 1/3". Variando a distância focal da lente, outros resultados podem ser alcançados. Quanto maior o valor da distância focal, menor a área visualizada pela câmera, entretanto, o objeto alvo é ampliado, permitindo ampliação da distância de trabalho da câmera para captura da face de pessoas.

A Tabela 7.15 a seguir sintetiza os resultados encontrados e podem ser utilizados para dimensionar ambientes internos onde se queira posicionar uma câmera de CFTV para a captura de faces de suspeitos, por exemplo, em uma sala de identificação criminal ou em um corredor onde se está monitorando a passagem de pessoas.

Tabela 7.6 – Quadro geral sobre os critérios criados para o dimensionamento de câmeras de segurança em aplicações de CFTV.

Critérios*	**Resultados**
Posicionamento	
Cadência	Pelo menos 5 fps
Ângulo Vertical Câmera-Face	Entre -60º e 60º
Ângulo Horizontal Câmera-Face	Entre -90º e 90º
Altura da Câmera	Mínimo de 2,30 metros para câmeras com resolução espacial de 640 x 480 pixels.
Distância mínima câmera-face	Mínimo de 3,0 metros para câmeras com resolução espacial de 640 x 480 pixels.
Iluminação	
Tipo de lâmpada / luz	Descarga de Mercúrio metálico (cor branca)

Iluminância do ambiente	Entre 110 lux e 1080 lux na área da face
Resolução Espacial *Resolução espacial mínima*	 640 x 480 pixels

* Os resultados apurados nos experimentos realizados neste trabalho utilizaram ambientes com paredes na cor branca.

A aplicação típica e prática para captura facial trata-se de posicionar uma câmera de CFTV para monitorar indivíduos percorrendo um corredor, como por exemplo, um hall de entrada de um estabelecimento, posicionando a câmera de forma a capturar imagens da face (rosto) o mais frontal possível, pois os algoritmos de reconhecimento facial apresentam melhor desempenho nesta condição ao extrair as características da face. Recomenda-se, desta forma, com base no que foi apurado nos experimentos realizados neste trabalho, baseado no método PCA (*eigenface*) e também, com base nos principais trabalhos encontrados na literatura técnica (Axis_a, 2012) (BERGSTRÖM, 2005) (KOVESI, 2009) (SWGIT, 2004), o dimensionamento do cenário (ambiente de captura de imagens) conforme apresentado na Tabela 7.7, como um padrão mínimo para a realização de reconhecimento facial de suspeitos.

Tabela 7.7 – Padrão mínimo para o reconhecimento facial de suspeitos.

Padrão	**Parâmetro**
Posicionamento da Face capturada na imagem	O mais frontal possível
Iluminação	Entre 500 e 800 lux na área da face
Resolução Espacial	640 x 480 pixels
Tamanho da Face proporcional à imagem	40 pixels na largura do rosto ou 1/16 de largura da imagem na resolução espacial de 640 x 480 pixels para a largura do rosto ou ainda pelo menos 1% da área da imagem para enquadramento da área da face na imagem

8 – CONCLUSÃO E TRABALHOS FUTUROS

Este capítulo apresenta as principais conclusões do presente trabalho, as limitações dos resultados experimentados, bem como o direcionamento de trabalhos futuros.

8.1 – CONCLUSÃO

A ausência de normativos técnicos para subsidiar projetos de sistemas de CFTV têm sido um dos grandes problemas encontrados pelas forças policiais brasileiras no que se refere à utilização das imagens extraídas de sistemas de CFTV para fins de reconhecimento facial de suspeitos de delitos. O projeto e implantação desses sistemas geralmente são mal dimensionados e falham em fornecer imagens com qualidade mínima para procedimentos de reconhecimento facial.

A lacuna normativa dá liberdade para a instalação de sistemas de CFTV sem critérios. Desta forma, os baixos investimentos realizados na aquisição e instalação de soluções de CFTV, a ausência de amparo profissional especializado na especificação da melhor solução a ser adquirida, aquisição de equipamentos inadequados, instalação da infraestrutura de forma deficitária, posicionamento inadequado das câmeras de segurança e da iluminação ambiente, imagens capturadas com baixa resolução espacial e a utilização de técnicas de compressão indevidas como forma de redução de espaço no armazenamento de arquivos são alguns dos fatores que impactam na dificuldade de processamento das evidências extraídas de sistemas de CFTV, no que tange ao reconhecimento facial de suspeitos.

Assim, o presente trabalho procurou propor critérios para que fossem minimizados esses problemas encontrados no projeto de sistemas de CFTV. O foco do trabalho foi na determinação de critérios para o projeto e instalação de câmeras de segurança, em ambientes *indoor* (internos), com base em sistemas de gravação em formato de imagem, podendo ser aplicado à sistemas que gerem arquivos em formato de vídeo, necessitando, entretanto, um pré-processamento anterior dos arquivos de vídeo para extração de imagens.

Neste trabalho, foi apresentada uma proposta com quatro critérios para projeto de câmeras em aplicações de CFTV: posicionamento, iluminação, resolução espacial e auditoria. No

critério de posicionamento, foram avaliados os parâmetros ângulo câmera-face, altura de instalação da câmera e cadência de captura das imagens. No critério de iluminação, foram definidos o tipo de lâmpada mais adequada e os limites necessários para os níveis de iluminância do ambiente, necessários para permitir o reconhecimento facial de suspeitos. No critério de resolução espacial, foi abordado o parâmetro de tamanho da face em relação à resolução espacial da imagem e da distância entre a posição da câmera e a face capturada, que possibilitem o reconhecimento facial de suspeitos. O critério de auditoria é um critério especial, em comparação aos demais, pois não é comum que fabricantes de equipamentos de CFTV disponibilizem recursos que auxiliem na garantia da integridade das imagens/vídeos gerados por seus equipamentos. Neste critério, portanto, foi definida uma sistemática para que as imagens coletadas de sistemas de CFTV possam ser verificadas contra adulterações intencionais, aumentando a confiabilidade da prova em um tribunal, à medida que a cadeia de custódia é resguardada. A solução proposta não é definitiva, pois não garante a integridade das imagens produzidas e/ou analisadas, mas adiciona uma camada de confiança ao processo.

A partir da proposta especificada, foram realizados experimentos, utilizando equipamentos de CFTV, de forma a simular aquelas condições encontradas em um cenário real de utilização de câmeras de CFTV, capturando imagens de faces de pessoas circulando pelo ambiente. Foram capturados imagens com alta resolução espacial para alimentarem dois bancos de dados criados neste trabalho, procurando simular bancos de dados criminais autênticos. Foram utilizados também bancos de dados de imagens de faces públicos para a realização dos experimentos. A partir de então, as imagens capturadas no cenário de CFTV montado foram extraídas e processadas para confronto com os bancos de dados criados. Estas imagens apresentam resolução espacial inferior àquelas que alimentaram os bancos de dados criminais simulados, procedimento este adotado em virtude da casuística real aplicada em reconhecimento facial de suspeitos. Foi utilizado um pacote para reconhecimento facial, baseado no método *eigenfaces*, executado a partir do aplicativo Matlab, para apuração dos resultados. Este método foi escolhido, pois, apesar de apresentar boa performance para imagens frontais, apresenta algumas deficiências de desempenho, quando comparado com outros métodos preparados para agir sobre determinadas situações, como diferentes expressões faciais, diferentes níveis de iluminância, diferentes poses faciais, dentre outros. Desta forma, os resultados obtidos com o método estariam condizentes com a proposta do trabalho, de apresentar limites satisfatórios para projetos de

câmeras de CFTV, que possibilitem o uso das imagens geradas para fins de reconhecimento facial de suspeitos.

Os resultados obtidos e apresentados no capitulo anterior buscaram estabelecer um padrão mínimo a ser utilizado em sistemas de CFTV e poderão auxiliar novos projetos ou projetos existentes, de forma que as imagens capturadas pelo sistema de CFTV poderão ser utilizadas posteriormente por autoridades policiais e pela perícia oficial, no auxílio ao reconhecimento facial de suspeitos e combate à criminalidade.

8.2 – LIMITAÇÕES

As principais limitações associadas à solução proposta foram:

- Foi utilizado um ambiente simulado para a realização dos experimentos, com boa resolução espacial (resolução de 640 x 480 pixels), alta taxa de sinal ruído e bom contraste nas imagens utilizadas para confronto com o banco de dado. Em ambientes reais, as imagens capturadas geralmente apresentam uma série de deficiências, como baixo sinal ruído, baixa resolução espacial, problemas com a configuração da cadência de quadros, gerando problemas de movimentação da face na imagem, *motion blur*, dentre outros.

- A eficácia dos resultados pode variar de acordo com a técnica utilizada para realizar o processo completo de reconhecimento facial, desde a detecção facial até a comparação de resultados. Foi utilizada apenas uma técnica para realização dos experimentos, baseada no método *eigenfaces*. Entretanto, o método escolhido foi usado justamente por apresentar deficiências de performance quando comparado com os métodos que evoluíram a partir desta técnica, para permitir melhor performance em cenários que apresentam condições distintas daquelas ideais. Desta forma, os resultados aplicados ao método *eigenfaces* servirão como base limítrofe, caso sejam utilizados outros métodos que possibilitem melhor desempenho.

8.3 – TRABALHOS FUTUROS

Dado o fato de que reconhecimento facial é um campo de pesquisa emergente e que os aplicativos que desempenham esta função ainda tem muito que amadurecer, alguns trabalhos são sugeridos como complemento às pesquisas realizadas nesta dissertação. As propostas são as seguintes:

- Realização de um estudo comparativo com outros métodos de reconhecimento facial, de forma a avaliar os resultados extraídos neste trabalho como proposta de critérios para projetos de câmeras de segurança de CFTV;
- Propor critérios para câmeras de CFTV em aplicações externas (*outdoor*), ampliando a aplicação dos critérios definidos neste trabalho;
- Propor critérios para câmeras de CFTV, para aplicações de reconhecimento de placas de veículos;
- Propor a geração de códigos criptográficos de autenticação de mensagens diretamente em equipamentos de câmeras de segurança, de forma a obter uma solução definitiva para garantir a integridade das imagens geradas por determinada câmera de vídeo CFTV;
- Formatar os critérios propostos e os resultados extraídos deste trabalho no padrão das normas técnicas da ABNT, para a aplicação dos critérios definidos em projetos de CFTV.

REFERÊNCIAS BIBLIOGRÁFICAS

ABATE, Andrea F.; NAPPI, Michele; RICCIO, Daniel; SABATINO, Gabriele. 2D and 3D face recognition: A survey. *Pattern Recognition Letters*, *28*(14), 1885-1906. doi: 10.1016/j.patrec.2006.12.018, 2007.

ABNT. NBR 5413. Associação Brasileira de Normas Técnicas (ABNT). 1991.

AIC. *CCTV as a crime prevention measure; What is CCTV?* Tip Sheet 5. Australian Institute of Criminology (AIC): Canberra, 2007.

ARAÚJO, Rommel Ferreira de. Agrupamento de documentos forenses utilizando redes neurais ART1. Dissertação de Mestrado em Engenharia Elétrica, ENE/FT/UNB. Brasília, 2011.

ARCA, S., CAMPADELLI, P. & LANZAROTTI, R. A face recognition system based on local feature analysis. *In Proc. Audio- and Video-Based Biometric Person Authentication*, pages 182-189, 2003.

AXIS. Resoluções. Axis Communications, 2012. Disponível em: <http://www.axis.com/pt/products/video/about_networkvideo/resolution.htm>. Acesso em: 10/01/2012.

AXIS_a. Resolução Necessária. Axis Communications, 2012. Disponível em: <http://www.axis.com/pt/edu/identification/resolution.htm>. Acesso em: 10/01/2012.

BAILEY, I., LOVIE, I. New Design Principles for Visual Acuity Letter Charts. *Am J Optom & Physiol Opt* 53:740-745, 1976.

BELHUMEUR, P. N., HESPANHA, J. P. & KRIEGMAN, D. J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.

BERGSTRÖM, Peter. Camera Surveillance – Test your system before a criminal does! Swedish National Laboratory of Forensic Science, Linköping, 2005.

BOLLE, R., CONNEL, J., PANKANTI, S., RATHA, N. & SENIOR, A. Guide to Biometrics. Springer Verlag, 2003.

BOSCH. Understanding Image Resolution. Bosch: New York, 2006.

BRADLEY, Hellen. How to keep a log of your work in Photoshop. 2012. Disponível em: <http://www.digital-photography-school.com/how-to-keep-a-log-of-your-work-in-photo shop>. Acesso em: 08/01/2012.

BRASIL. Código de Processo Penal. Decreto-Lei nº 3.689/1941. Presidência da República, Casa Civil, Subchefia para Assuntos Jurídicos. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto-lei/Del3689.htm>. Acesso em: 18/01/2012.

BRASIL. Lei nº 12.030/2009. Presidência da República, Casa Civil, Subchefia para Assuntos Jurídicos. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2009/Lei/L12030.htm>. Acesso em: 18/01/2012.

BRUCE, Vicki; HENDERSON, Zo; GREENWOOD, Karen; HANCOCK, Peter J. B.; BURTON, A. Mike & MILLER, Paul. Verification of face identities from images captured on video. *Journal of Experimental Psychology: Applied*. Vol 5(4), Dec 1999.

BURTON, A.M; WILSON, S.; COWAN, M. & BRUCE, V. Face recognition in poor quality video: evidence from security surveillance. *Psychological Science*, 10, 243-248, 1999.

CAI, J. & GOSHTASBY, A. Detecting human faces in color images. *In Image and Vision Computing*, vol. 18, pp. 63-75. Elsevier, 1999.

CAMASTRA, F. & VINCIARELLI, A. Machine Learning for Audio, Image and Video Analysis: Theory and Applications. Springer, 2008.

CAMPOS, T. E. Técnicas de Seleção de Atributos e de Classificação para Reconhecimento de Faces. Dissertação de Mestrado, Departamento de Ciência da Computação, IME – USP, 2000.

CARVALHO, João Luiz de. Fundamentos da perícia criminal. 1a. ed. Campinas: Bookseller, 2006.

CARVALHO, Suzana Papile; DA SILVA, Ricardo Henrique Alves; JÚNIOR, César Lopes & PERES, Arsenio Sales. A utilização de imagens na identificação humana em odontologia legal. Revista Radiologia Brasileira. Bauru/SP, Vol. 42, N.2, Mar/Abr, 2009. Disponível na URL <http://www.rb.org.br/detalhe_artigo.asp?id=838>. Acesso em: 09/07/2011.

CIESZYNSKI, Joe. Closed Circuit Television. 3a. ed. Newnes: Burlington, 2007.

COSTA, P. Crescimento e modernização do mercado CFTV. 2010. Disponível em: <http://www.segs.com.br/index.php?option=com_content&view=article&id=11000:crescimento-e-modernizacao-do-mercado-cftv-&catid=50:cat-demais&Itemid=331>. Acesso em: 13/01/2012.

CRAW, I., TOCK, D. & BENNETT, A. Finding face features. *In Proceedings of European Conference on Computer Vision*, pp. 92-96, 1992.

DAMJANOVSKI, Vlado. CCTV Networking and Digital technology. 2a. ed. Elsevier: Burlington, 2005.

DEMERCIAN, Pedro Henrique & Maluly, Jorge Assaf. Curso de processo penal. 3a. ed. Rio de Janeiro: Forense, 2005.

DOREA, Luiz Eduardo Carvalho; Stumvoll, Victor Paulo & Quintela, Victor. Criminalística. 4ª. ed. Campinas: Millennium Editora, 2010.

DPF. Departamento de Polícia Federal. Laudo nº 3412/2008 – NUCRIM/SETEC/SR/DPF/SP – Laudo de Exame Biométrico (Reconhecimento Facial). 2008.

_____. Departamento de Polícia Federal. Laudo nº 698/2010 – SETEC/SR/DPF/GO – Laudo de Exame Biométrico (Reconhecimento Facial). 2010.

_____. Departamento de Polícia Federal. Laudo nº 1114/2011 – SETEC/SR/DPF/GO – Laudo de Perícia Criminal Federal (Registros de Aúdio e Imagens). 2011.

DPF_a. Departamento de Polícia Federal. Laudo nº 1218/2010 – SETEC/SR/DPF/GO – Laudo de Exame Biométrico (Reconhecimento Facial). 2010.

ESPINDULA, Alberi. Perícia Criminal e Cível. 3ª. ed. Campinas: Millennium Editora, 2009.

FERIS, R. S., CAMPOS, T. E. & JUNIOR, R. M. C. Detection and Tracking of Facial Features in Video Sequences. MICAI 2000: Advances in Artificial Intelligence, pp. 127-135, 2000.

FREUND, Y., SCHAPIRE, R. A decision-theoretic generalization of on-line learning and na application to boosting. *In European conference on computational learning theory*, p. 23-37, 1995.

GABOR, D. Theory of communication. Journal of the Institute of Electrical Engineers, 93, 429-457. 1946.

GARCIA, Iberê Anselmo. A segurança na identificação: a biometria da íris e da retina. Dissertação de Mestrado, Faculdade de Direito, Universidade de São Paulo, São Paulo/SP, 2009.

GEORGE, Robert M. Facial Geometry: Graphic Facial Analysis for Forensic Artists. Springfield: Thomas Books, 2007.

GOOLD, B. J. CCTV and Policing: Public Area Surveillance and Police Practices in Britain. Oxford: Oxford University Press, 2004.

GONZALEZ, Rafael C. & WOODS, Richard E. Processamento Digital de Imagens. 3a. ed. Pearson: São Paulo, 2010.

HADID, A., AHONEN, T. & PIETIKAINEN, M. Face recognition with local binary patterns. *In Proc. European Conference on Computer Vision*, pages 469-481, 2004.

HEISELE, B., HO, P., WU, J. & POGGIO, T. Face recognition: component-based versus global approaches. *In Computer Vision and Image Understanding*, 91(1), pp. 6-21, 2003.

HSU, R.-L., ABDEL-MOTTALEB, M. & JAIN, A. K. Face Detection in Color Images. *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), pp. 696-706, 2002.

IAI. International Association for Identification. 2011. Disponível em: <http://www.theiai.org/>. Acesso em: 30/01/2012.

IBGE. Pesquisa de Orçamentos Familiares 2008-2009 Antropometria e Estado Nutricional de Crianças, Adolescentes e Adultos no Brasil. Instituto Brasileiro de Geografia e Estatística – IBGE: Rio de Janeiro, 2010.

ICPR. ICPR Workshop benchmark database. 2004. Disponível em: <http://www-prima.inrialpes.fr/Pointing04/data-face.html>. Acesso em: 03/03/2012.

INTRONA, L. & NISSENBAUM, H. Facial Recognition Technology A Survey of Policy and Implementation Issues. Organisation, Work and Technology Working Paper Series, Lancaster University: The Department of Organisation, Work and Technology, 2010.

JAIN, Anil K., KLARE, Brendan & PARK, Unsang. Face Recognition: Some Challenges in Forensics. Departamento de Ciência da Computação e Engenharia, Universidade do Estado de Michigan, East Lansing, MI, U.S.A, 2002.

JIN, Z., LOU, Z., YANG, J. & SUN, Q. Face detection using template matching and skin-color information. Neurocomputing, Vol. 70, pp. 794-800, 2007.

KEMP, R., TOWELL, N. & PIKE, G. When seeing should not be believing: Photographs, credit cards and fraud. *In Applied Cognitive Psychology*, 11, 211-222, 1997.

KAKADIARIS, I. A., PASSALIS, G., TODERICI, G., MURTUZA, N., LU, Y., KARAMPATZI-AKIS, N. & THEOHARIS, T. Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach. *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):640-649, 2007.

KING, Joel W. Cisco IP Video Surveillance Design Guide. Cisco, 2009.

KINUTA, C., MOLINA, D., DORNELES, E. G., GRECCHI, F. S., DIAS, G. T., SANTANA, J. & FERNANDES JUNIOR, O. O. Estudo Comparativo de Algoritmos para Reconhecimento Facial. Universidade IMES, São Caetano do Sul, São Paulo, 2006.

KIRBY, M. & SIROVICH, L. Application of the Karhunen-Loeve procedure for the characterization of human faces. *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1), pp. 103-108, 1990.

KOVESI, Peter. Video Surveillance: Legally Blind? Digital Image Computing: Techniques and Applications, 204-211. IEEE. doi: 10.1109/DICTA.2009.41, 2009.

KRUEGLE, Herman. CCTV Surveillance Video Practices and Technology. 2a. ed. Butterworth-Heinemann; Burlington, 2007.

LEVI, K. & WEISS, Y. Learning object detection from a small number of examples: the importance of good features. *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 53-60, 2002.

LI, Stan Z. & JAIN, Anil K. Handbook of Face Recognition. New York: Springer, 1a ed., 2005.

LOPES, E. C. Detecção de Faces e Características Faciais. Relatório Técnico No. 45, Pontífica Universidade Católica do Rio Grande do Sul, PUCRS, 2005.

LU, X. & JAIN, A. K. Deformation modeling for robust 3d face matching. *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1346-1357, 2008.

LU, J., PATANIOTIS, K. N. & VENETSANOPOULOS, A. N. Boosting Linear Discriminant Analysis for Facial Recognition. Bell Canada Multimedia Laboratory, The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Canada, 2002.

MACHADO, Carlos E. P., FILHO, Eduardo M. de Lacerda, REIS, Paulo M. G. I., ROCHA, Anderson de Rezende, GRECO, Leonardo G. & GOLDENSTEIN, Siome Klein. Reconhecimento Facial. Brasília: Academia Nacional de Polícia (ANP), 2010.

MANZANO, Luís Fernando de Moraes. Prova Pericial: Admissibilidade e assunção da prova científica e técnica no processo brasileiro. São Paulo/SP: Atlas, 2010.

MARTINEZ, A. M. Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(6):748-763, 2002.

MIT. Photobook/Eigenfaces Demo. 2002. Disponível em: <http://vismod.media.mit.edu/vismod/demos/facerec/basic.html>, acessado em 26/08/2011.

NIETO, Marcus, Johnson-Dodds, Kimberly & Simmons, Charlene W. Public and Private Applications of Video Surveillance and Biometric Technologies. California Research Bureau, 2002.

NIST. The Facial Recognition Technology (FERET) Database. Disponivel em: <http://www.itl.nist.gov/iad/humanid/feret/feret_master.html>. Acesso em: 22/08/2011.

NIST_a. Face Recognition Vendor Test (FRVT) Documents. Disponivel em: <http://www.nist.gov/itl/iad/ig/frvt-docs.cfm>. Acesso em: 22/08/2011.

NIST_b. Face Recognition Grand Challenge (FRGC). Disponivel em: <http://www.nist.gov/itl/iad/ig/frgc.cfm>. Acesso em: 22/08/2011.

NIST. Secure Hash Standard (SHS) – FIPS PUB 180-4. Disponível em: < http://csrc.nist.gov/publications/fips/fips180-4/fips-180-4.pdf>. Acesso em: 05/03/2012.

NTSC. Face Recognition. National Science and Technology Council, Committee on Technology, Committee on Homeland and National Security, Subcommittee on Biometrics, 2006.

OLIVEIRA, Douglas Rodrigues. Reconhecimento de faces usando Redes Neurais e Biometria. Dissertação de Mestrado, Instituto Nacional de Pesquisas Espaciais-INPE, São José dos Campos/SP, 2006.

OLIVEIRA, Talita Gouvea de & SANCHEZ, Cláudio José Palma. Os meios de prova e suas limitações no processo penal. ETIC – Encontro de Iniação Científica. ISSN 21-76-8498. Vol. 4, N. 4. Presidente Prudente/SP, 2008.

PARK, Unsang. Face Recognition: face in video, age invariance, and facial marks. Tese de Doutorado, Universidade do Estado de Michigan – EUA, 2009.

PEASE, K. Crime Prevention. *In The Oxford Handbook of Criminology*. 2ª. ed. Nova York: Oxford University Press, 1997.

PELCO. Image Resolution. Digital Video Systems Design for D.U.M.I.E.S. SecuritySales & Integration, 2011.

PENTEADO, Bruno E. Autenticação biométrica de usuários em sistemas de e-learning baseada em reconhecimento de faces a partir de video. Dissertação de mestrado, UNESP: Bauru, 2009.

PHILLIPS, P. J., GROTHER, P., MICHEALS, R. J., BLACKBURN, D. M., TABASSI, E. & BONE, M. Face Recognition Vendor Test 2002: Evaluation Report. Tech. Report NISTIR 6965, NIST, 2003.

PHILLIPS, P. J., SCRUGGS, W. T., O'TOOLE, A. J., FLYNN, P. J., BOWYER, K. W., SCHOTT, C. L. & SHARPE, M. Face Recognition Vendor Test 2006: FRVT 2006 and ICE 2006 Large-Scale Results. Tech. Report NISTIR 7408, NIST, 2007.

POLICOM. Soluções em Segurança e CFTV. 1a. ed. Grupo Policom, 2009.
Portela, Gutemberg. CFTV Treinamento Básico. HOT SAT, 2007.

QING, L., SHAN, S., CHEN, X. & GAO, W. Face recognition under varying light-ing based on the probabilistic model of gabor phase. *In Proc. International Conference on Pattern Recognition*, pages 1139-1142, 2006.

QUINTILIANO, Paulo & Rosa, Antonio. Face Recognition applied to computer forensics. *In The International Journal of Forensic Computer Science* (IJoFCS), Vol. 1, pp. 19-27, 2006.

RATCLIFFE, Jerry. Video Surveillance of Public Places. *In Problem-Oriented Guides for Police Response Guides Series*. No. 4. COPS Community Oriented Policing Services U.S. Department of Justice. Washington, 2006.

REIS, Fabio A. S. O monitoramento visual via circuito fechado de TV como estratégia de prevenção ao crime: breve revisão bibliográfica internacional. *In Revista de Direito das Novas Tecnologias*. IBDI, ISSN:1981-1020, v. 1, n. 2. 2006.

RODRIGUES, Vinícius Lopes. Reconhecimento Facial usando SVM. Departamento de Informática, Pontífica Universidade Católica do Rio de Janeiro (PUC-RIO). Rio de Janeiro, 2007.

ROSA, Luigi. Face Recognition Technology. 2012. Disponível em: <http://www.facerecognition.it>. Acesso em: 01/02/2012.

ROWLEY, H. A., BALUJA, S. & KANADE, T. Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20(1), pp. 23-38, 1998.

RUSS, John C. Image Processing and Measurement – Examples. NCSU Forensic Science Symposium, 2009.

RUSSEL, R., SINHA, P., BIEDERMAN, I. & NEDERHOUSER, M. The utility of surface reflectance for the recognition of upright and inverted faces. Vision Research vol. 47, issue 2, pg. 157-165, 2007.

SABER, E. & TEKALP, A. M. Frontal-view Face Detection and Facial Feature Extraction Using Color Shape and Symmetry Based Cost Functions. Pattern Recognition Letters, Vol. 19, pp. 669-680, 1998.

SAFERSTEIN, R. Criminalistics - An Introduction to Forensic Science. Pearson: Upper Saddle River, 2011.

SCURI, Antonio Escaño. Fundamentos da Imagem Digital. Tecgraf/PUC-RIO: Rio de Janeiro, 2002.

SERIBELI, Nathália Hernandes. Violência doméstica contra criança e adolescente "um crime entre quatro paredes". *In IV Encontro de Iniciação Científica e III Encontro de extensão universitária*. Presidente Prudente/SP, Vol. 4, N. 4, 2008.

SHAN, T., LOVELL, B. C., CHEN, S. & BIGDELI, A. Reliable Face Recognition for Intelligent CCTV. Intelligent Real-Time Imaging and Sensing Group / National ICT Australia – EMI, Schooel of ITEE, The University of Queensland, 2003.

SILVERFAST. Resolution Target (USAF 1951). 2012. Disponível em: < http://www.silverfast.com/PDF/resolution-target/Resolution-Target_long_en.pdf>. Acesso em: 15/01/2012.

SINHA, Pawan, BALAS, Benjamin, OSTROVSKY, YURI & RUSSELL, Richard. Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About. *Proceedings of the IEEE*. Vol. 94, No. 11, November 2006.

SOARES, Clara Dias. A verdade no processo penal brasileiro. Jus Navigandi, Teresina, ano 13, n. 1749, 15 abr. 2008. Disponível em: <http://jus.com.br/revista/texto/11160>. Acesso em: 19/01/2012.

SOBOTTKA, J. & PITTAS, I. Segmentation and tracking of faces in color images. *In Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, pp. 236-241, 1996.

SWGIT. Best Practices for Image Authentication (section 14). IAI, 2007. Disponível em: <http://www.theiai.org/guidelines/swgit/guidelines/section_14_v1-0.pdf>. Acesso em: 30/01/2012.

_____. Recommendations and Guidelines for Using Closed-Circuit Television Security Systems in Commercial Institutions (section 4). IAI, 2004. Disponível em: <http://www.theiai.org/guidelines/swgit/guidelines/section_4_v2-1.pdf>. Acesso em: 30/01/2012.

_____. Overview of SWGIT and the Use of Imaging Technology in the Criminal Justice System (section 1). IAI, 2010. Disponível em: <http://www.theiai.org/guidelines/ swgit/guidelines/section_1_v3-3.pdf>. Acesso em: 30/01/2012.

SWGIT_a. Best Practices for Maintaining the Integrity of Digital Images and Digital Video (section 13). IAI, 2007. Disponível em: <http://www.theiai.org/guidelines/swgit/guidelines/ section_134_v1-0.pdf>. Acesso em: 30/01/2012.

TOSHIBA. Designing an IP Camera Project: Is pixels per square foot an important measurement ? TOSHIBA: Irvine Boulevard, 2011.

TURK, M. A. & PENTLAND, A. P. Face Recognition Using Eigenfaces. IEEE CVPR, 1991.

UNICAMP. Iluminância e cálculo luminotécnico. Laboratório de Iluminação – Unicamp. Disponível em: <http://www.iar.unicamp.br/lab/luz/ld/Arquitetural/tabelas/luminotecnica.pdf>. Acesso em: 06/01/2012.

VALERA, M. & VELASTIN, S. Intelligent distributed surveillance systems: a review. *In IEEE Proc. – Vis., Image and Signal Process*, vol. 152, 2005.

VIOLA, P. A. & JONES, M. J. Robust real-time face detection. *In International Journal of Computer Vision*, Vol. 57(2), pp. 137-154, 2004.

WEI, S. A Shape Analysis in Computer Vision Final project report: Face recognition. Department of Electrical Engineering. McGill University: Canada, 1998.

WILLIANS, S. How to design the perfect CCTV system. 2012. Disponível em: < http://www.reeltech.net/wp-content/uploads/2012/02/cctvreport.pdf>. Acesso em: 02/02/2012.

WISKOTT, L., FELLOUS, J. M., KRÜGER, N. & MALSBURG, C. Face recognition and gender determination. *In International Conference on Automatic Face and Gesture Recognition*, pp. 92-97, 1997.

WOODWARD, J., HORN, C., GATUNE, J. & THOMAS, A. Biometrics: A look at facial recognition. RAND DRR-3465, ISBN: 0-8330-3302-6, 2003.

WU, J., BRUBAKER, S. C., MULLIN, M. D. & REHG, J. M. Fast asymmetric learning for cascade face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30(3), pp. 369-382, 2008.

YANG, G. & HUANG, T. S. Human face detection in a scene. *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 453-458, 1993.

YANG, M.-H. & AHUJA, N. Detection human faces in color images. *In International Conference on Image Processing*, Vol. 1, pp. 127-130, 1998.

YANG, M.-H., KRIEGMAN, D. & AHUJA, N. Detecting faces in images: A survey. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1): 34-58, 2002.

YARMEY, A.D. The psychology of eye-witness testimony. New York: The Free Press, 1979.

ZHANG, W., SHAN, S., GAO, W., CHEN, X. & ZHANG, H. Local gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition. *In Proc. IEEE International Conference on Computer Vision*, pages 786-791, 2005.

ZHAO, W. & CHELLAPPA, R. Robust face recognition using symmetric shape-from-shading. Technical Report, Center for Automation Research, University of Maryland, 1999.

ZHAO, W.; CHELLAPPA, R.; ROSENFELD, A. & PHILLIPS, P.J. Face Recognition: A Literature Survey. *ACM Computing Surveys*, Vol. 35 (4), pp. 399-458, December 2003.

APÊNDICES

A − SKL TEST CHART

O propósito do teste gráfico SKL test chart, produzido pelo Laboratório Nacional Sueco de Ciências Forenses (SLK), é forneceer um meio de medir a acuidade visual de uma câmera, indicando a qualidade de imagem necessária exigida por um sistema de CFTV (BERGSTRÖM, 2005).

O procedimento é sistematizado a seguir:

1) Imprimir a tabela de teste, apresentada no fim deste apêndice. Ela está apresentada na escala de 1:1. Na versão impressa, a moldura do quadro ao redor da tabela deve conter as medidas 16 x 24,5 cm.

2) Coloque a tabela de teste na posição onde a qualidade de imagem é para ser medida. Uma pessoa deve se posicionar também ao lado do gráfico.

3) Em seguida, verifique nas imagens produzidas pela câmera testada quais linhas na tabela de teste pode ser lidas. Uma linha é considerada como legível se mais do que 60% das letras podem ser lidas corretamente. Deve haver o cuidado na avaliação deste resultado, se um mesmo avaliador realiza várias medições, em virtude da memorização das combinações.

4) Verifique na tabela a seguir a qualidade esperada de acordo com o uso do sistema de CFTV.

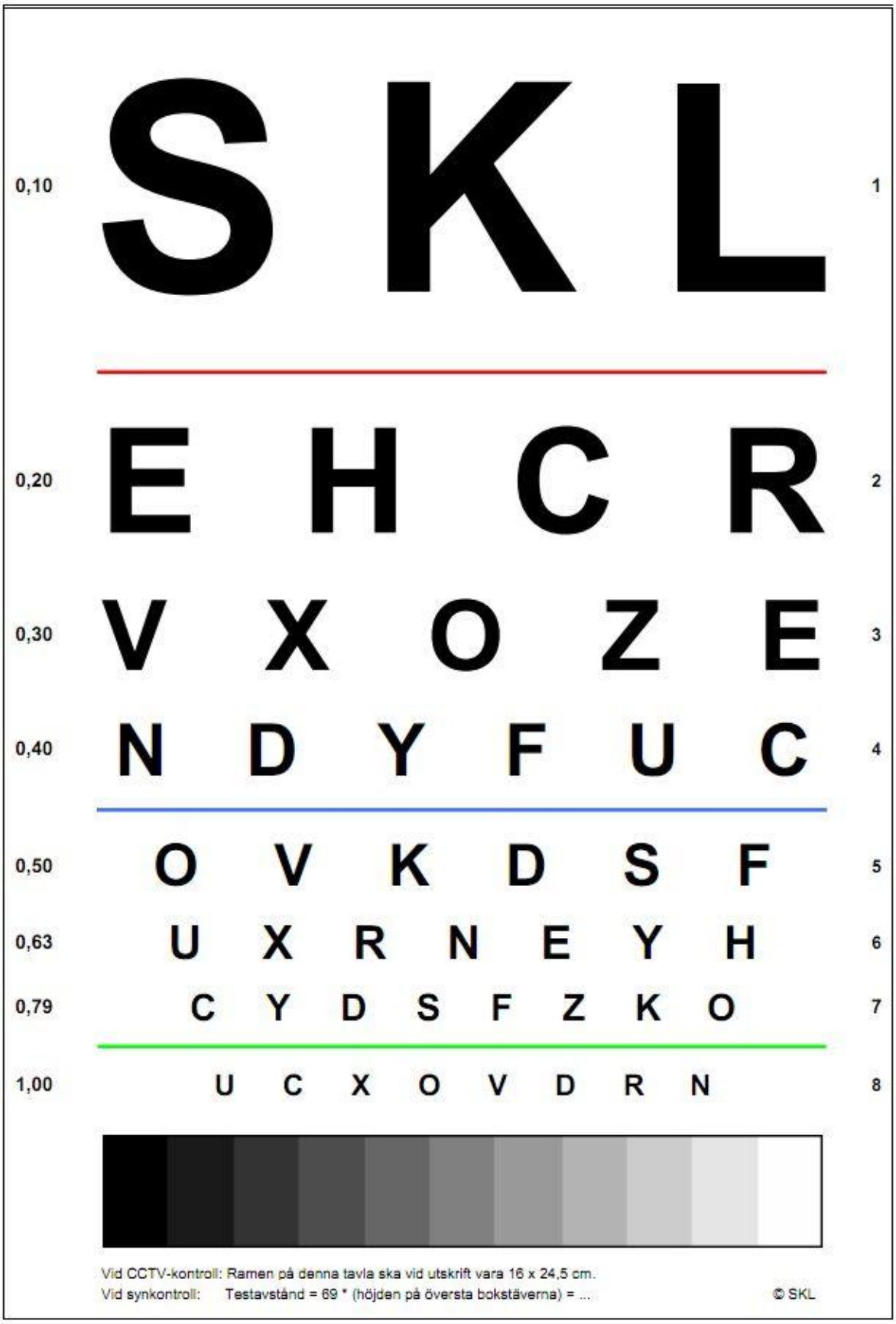
Tabela A.1 − Requerimentos de qualidade da imagem (BERGSTRÖM, 2005).

Nível de qualidade da imagem	Linha que deve ser lida da tabela SKL	Velocidade de captura de quadros (quadros/s)
Cadeia de eventos (A)	1	1
Características (B)	2-5	3
Identificação (C)	6-8	5
Biometria (D)	2-5	5

As imagens geradas por uma câmera grande angular permitem ser utilizadas para formar impressões sobre cadeias de eventos. Isto pode ser feito em diferentes níveis. Aqui, o nível é definido de forma de que a qualidade da imagem permita discernir o tipo de objeto que uma pessoa está segurando, tal como uma bolsa ou uma arma. Para que isso seja possível, a linha 1 da tabela de teste deve ser lida no área principal da câmera de vigilância.

Se o limite para o que pode ser lido pelas imagens são as linhas 2, 3 ou 4, o sistema de CFTV pode ser usado para descrever características das pessoas monitoradas. Quanto maior o número da linha legível, maior for a probabilidade de conseguir mais detalhes.

Se as linhas 5, 6 ou 7 podem ser lidas, é provável que uma pessoa suspeita pode ser reconhecida. Mais uma vez, quanto maior o número da linha legível, maior será essa probabilidade. Para o reconhecimento facial, as imagens devem permitir diferenciar entre a maioria dos níveis de escala de cinza apresentados na tabela de teste.


0,10 S K L 1
0,20 E H C R 2
0,30 V X O Z E 3
0,40 N D Y F U C 4
0,50 O V K D S F 5
0,63 U X R N E Y H 6
0,79 C Y D S F Z K O 7
1,00 U C X O V D R N 8
Vid CCTV-kontroll: Ramen på denna tavla ska vid utskrift vara 16 x 24,5 cm.
Vid synkontroll: Testavstånd = 69 * (höjden på översta bokstäverna) = ...
© SKL

B − DEFINIÇÃO DO CAMPO DE VISÃO (FOV) E LENTE

As tabelas a seguir apresentam as referências para determinar o tamanho da cena a ser monitorada para os três tipos de sensores de imagens mais comuns, como função da distância da câmera para o objeto e da distância focal da lente. Maiores detalhes sobre os parâmetros na Figura B.1.

Figura B.1 − Tabela guia para sensores de 1/4" (KRUEGLE, 2007).

DISTÂNCIA FOCAL (mm)	FoV ANGULAR: H X V (GRAUS)	GUIA PARA SENSORES DE 1/4" DISTÂNCIA DA CAMERA PARA A CENA OBSERVADA (D) EM PÉS LARGURA E ALTURA DA CENA (W X H) EM PÉS							
		5	10	20	30	40	50	75	100
		W×H	W×H	W×H	W×H	W×H	W×H	W×H	W×H
2.1	81.2×60.9	8.6×6.4	17×12.9	34×26	51×39	69×51	86×64	129×96	171×129
2.2	78.6×59.0	8.2×6.1	16.4×12.2	33×25	49×37	65×49	82×63	123×92	164×123
2.3	76.1×57.1	7.8×5.9	15.6×11.8	31×23	47×35	62×47	78×59	117×86	157×117
2.6	69.4×52	6.9×5.2	13.9×10.4	28×21	42×31	55×42	69×52	104×78	138×104
3.0	61.9×46.4	6.0×4.5	12×9	24×18	36×27	48×36	60×45	90×68	120×90
3.6	53.1×39.8	5.0×3.8	10×7.5	20×15	30×23	40×30	50×38	75×57	100×76
3.8	50.7×38.0	4.7×3.6	9.5×7.1	19×14	28×21	38×28	47×36	71×54	94×72
4.0	48.5×36.4	4.5×3.4	9×6.8	18×14	27×20	36×27	45×34	68×51	90×68
4.3	45.4×34.1	4.2×3.1	8.4×6.3	16.7×12.5	25×19	33×25	42×31	63×47	84×62
6.0	33.4×25.0	3×2.3	6×4.5	12×9	18×13.5	24×18	30×23	45×35	60×46
8.0	25.4×19.0	2.3×1.7	4.5×3.4	9×6.8	13.5×10.1	18×13.5	23×17	35×26	46×34
12.0	17.1×12.8	1.5×1.1	3×2.2	6×4.4	9.0×6.8	12×9	15×11	23×17	30×23
16.0	12.8×9.6	1.1×.8	2.3×1.7	4.5×3.4	6.8×5.1	9×6.8	11.2×8.4	17×13	22×17
25.0	8.2×6.2	.72×.54	1.4×1.1	2.9×2.1	4.3×3.2	5.8×4.3	7.2×5.4	10.8×8.1	14.4×10.8

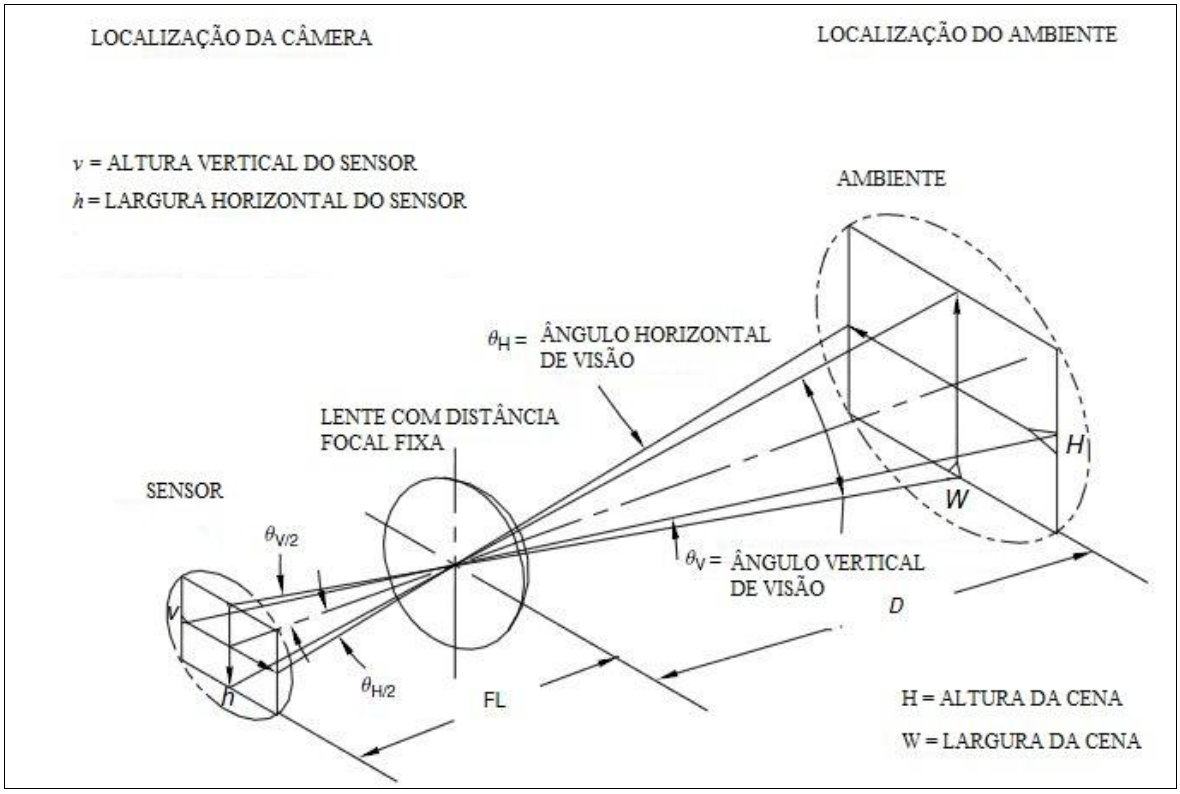
Figura B.2 − Tabela guia para sensores de 1/3" (KRUEGLE, 2007).

DISTÂNCIA FOCAL (mm)	FoV ANGULAR: H X V (GRAUS)	GUIA PARA SENSORES DE 1/3" DISTÂNCIA DA CAMERA PARA A CENA OBSERVADA (D) EM PÉS LARGURA E ALTURA DA CENA (W X H) EM PÉS							
		5	10	20	30	40	50	75	100
		W×H	W×H	W×H	W×H	W×H	W×H	W×H	W×H
2.3	92.4×69.3	10.4×7.8	20.8×15.6	41.6×31.2	63×47	83×62	104×78	156×117	208×156
2.6	85.4×64.1	9.2×6.9	18.5×13.8	36.8×27.6	55×41	77×58	92×69	138×104	184×138
2.8	81.2×60.9	8.6×6.5	17.2×13	34.4×26	51×39	69×52	86×65	129×98	172×130
3.6	67.4×50.5	6.7×5.0	13.3×10	26.7×20	40×30	53×40	67×50	101×75	134×100
3.8	64.6×48.4	6.3×4.7	12.6×9.5	25×18.9	37.9×28.4	50.5×37.9	63×47	95×71	123×92
4.0	61.9×46.4	6.0×4.5	12×9	24×18	36×27	48×36	60×45	90×68	120×90
4.5	56.1×42.1	5.3×4.0	10.6×8	21.2×15.9	31.8×23.9	42.4×31.8	53×40	80×60	106×80
6.0	43.6×32.7	4.0×3.0	8.0×6	16×12	24×18	32×24	40×30	60×45	80×60
8.0	33.4×25.0	3.0×2.3	6×4.5	12×9	18×13.5	24×18	30×22.5	45×34	60×45
12.0	26.6×20.0	2.0×1.5	4.0×3.0	8.0×6.0	12.0×9.0	16×12	20×15	30×23	40×30
16.0	17.1×12.8	1.5×1.2	3.0×2.3	6.0×4.5	9.0×6.8	12.0×9.0	15.0×11.3	23×17	30×22.5
25.0	11.0×8.2	.96×.72	1.9×1.4	3.8×2.9	5.8×4.4	7.7×5.8	9.6×7.2	14.4×10.8	19.2×14.4
50.0	5.5×4.1	.48×.36	.96×.72	1.9×1.4	2.9×2.2	3.8×2.8	4.8×3.6	7.2×5.4	9.6×7.2
75.0	3.7×2.8	.32×.24	.64×.50	1.3×.96	1.9×1.4	2.6×1.9	3.2×2.4	4.8×3.6	6.4×4.8

Figura B.3 − Tabela guia para sensores de 1/2" (KRUEGLE, 2007).

GUIA PARA SENSORES DE 1/2"									
DISTÂNCIA FOCAL (mm)	FoV ANGULAR: H X V (GRAUS)	DISTÂNCIA DA CAMERA PARA A CENA OBSERVADA (D) EM PÉS LARGURA E ALTURA DA CENA (W X H) EM PÉS							
		5	10	20	30	40	50	75	100
		W×H	W×H	W×H	W×H	W×H	W×H	W×H	W×H
1.4	133×100	23×17	46×34	91×69	137×103	183×137	228×171	342×257	457×348
2.6	101.8×76.4	12.3×9.2	24.6×18	49×37	74×55	98×74	123×92	185×138	246×184
3.5	84.9×63.7	9.1×6.9	18.2×13.8	37×28	55×41	73×55	91×69	137×104	182×138
3.6	83.3×62.5	8.9×6.7	17.8×13.4	36×27	53×40	71×53	89×67	134×101	178×134
3.7	81.7×61.3	8.6×6.5	17.2×13.0	35×26	52×39	69×52	86×65	129×98	172×130
4.0	77.3×58.0	8.0×6.0	16.0×12.0	32×24	48×36	64×24	80×60	120×90	160×120
4.2	74.6×56.0	7.6×5.7	15.2×11.4	30×23	48×34	61×46	76×57	114×86	156×114
4.5	70.8×53.1	7.1×5.3	14.2×10.6	28×21	43×32	57×43	71×53	107×80	142×107
4.8	67.4×50.5	6.7×5.0	13.4×10.0	27×20	40×30	53×40	67×50	101×75	134×100
6.0	56.1×42.1	5.3×4.0	10.6×8.0	21×16	32×24	43×32	53×40	80×60	106×80
7.5	46.2×34.7	4.3×3.2	8.6×6.4	17.1×12.8	26×19	34×26	43×32	65×48	86×64
8.0	43.6×32.7	4.0×3.0	8.0×6.0	16×12	24×18	32×24	40×30	60×45	80×60
12.0	29.9×22.4	2.7×2.0	5.3×4.0	10.7×8	16×12	21.3×16	27×20	41×30	53×40
16.0	22.6×17.0	2.0×1.5	4.0×1.5	8×6	12×9	16×12	20×15	30×23	40×30
25.0	14.6×10.9	1.3×1.0	2.6×2.0	5.1×3.8	7.7×5.8	10.2×7.7	12.8×9.6	19×14	25.6×19.2
50.0	7.3×5.5	.64×.48	1.3×1.0	2.6×1.9	3.8×2.9	5.1×3.8	6.4×4.8	6.5×4.8	12.8×9.6
75.0	4.9×3.7	.43×.32	.85×.64	1.7×1.3	2.6×1.9	3.4×2.6	4.3×3.2	3.2×2.4	8.6×6.4
150.0	2.4×1.8	.21×.16	.43×.32	.85×.64	1.3×.96	1.7×1.3	2.1×1.6	9.6×7.2	4.3×3.2

Figura B.4− Parâmetros analisados nas tabelas (KRUEGLE, 2007).



C − PARÂMETROS DE LUMINÂNCIA

As tabelas a seguir apresentam os padrões de iluminância por ambientes exigidos pela norma técnica NBR5413, uma referência sobre o nível de iluminação encontrado em determinadas condições ambientes e os fluxos luminosos de diversos tipos de lâmpadas encontradas no mercado.

Figura C.1 − Iluminâncias (em lux) para cada classe de ambiente de tarefas visuais (ABNT, 1991).

Classe de Ambiente	**Iluminância (em lux)**	**Tipo de Ambiente / Atividade**
CLASSE A (áreas de uso contínuo e/ou execução de tarefas simples)	20 – 30 – 50	Ruas públicas e estacionamentos
	50 – 75 – 100	Ambientes de permanência curta
	100 – 150 – 200	Recintos não usados para trabalhos contíuos, depósitos
CLASSE B (áreas de trabalho em geral)	200 – 300 – 500	Trabalhos brutos e auditórios
	500 – 750 – 1000	Trabalhos normais: escritórios e fábricas
	1000 – 1500 – 2000	Trabalhos especiais: gravação, inspeção e indústria de tecidos
CLASSE C (áreas com tarefas visuais minuciosas)	1000 – 1500 – 2000	Trabalho contínuo e exato: eletrônica
	1000 – 1500 – 2000	Trabalho que exige muita exatidão: placas microeletrônicas
	1000 – 1500 – 2000	Trabalho minucioso especial: cirurgia

Figura C.2 − Níveis de iluminância de acordo com as condições ambientes (PORTELA, 2007).

Condição do Ambiente	**Nível de Iluminação (em lux)**
Dia claro	10.000
Dia escuro	100
Entardecer	10
Anoitecer	1
Noite de lua cheia	0,1
Noite com lua minguante	0,01

Figura C.3 − Fluxo luminoso de acordo com o tipo de lâmpada (UNICAMP, 2012).

Tipo de Lâmpada	**Potência (W)**	**Fluxo Luminoso (lm)**	**Índice de Reprodução de cor (IRC)**
Incandescente (127 V)	60	864	100
Incandescente (127 V)	100	1620	100
Fluorescente TLT	20	1100	70
Fluorescente TLD	32	2350	66
Fluorescente TLD/840	36	3350	85
Fluorescente TLT	40	2600	70
Fluorescente TLD/840	58	5200	85
Fluorescente TLT	65	4400	70
Fluorescente TLT	110	7600	70
Vapor de mercúrio	80	3700	48
Vapor de mercúrio	125	6200	46
Vapor de mercúrio	250	12700	40
Vapor de mercúrio	400	22000	40
Vapor metálico	256	19000	69
Vapor metálico	390	35000	69
Vapor metálico	400	35000	69
Vapor metálico	985	85000	65
Mista	165	3150	61
Mista	260	5500	63
Vapor de sódio	70	6600	25
Vapor de sódio	150	17500	25
Vapor de sódio	250	33200	25
Vapor de sódio	400	56500	25
Vapor de sódio	600	90000	25

D − TABELA DE RESOLUÇÕES PARA CÂMERAS E MONITORES

Soluções de CFTV utilizazm padrões para resolução em vídeo e estão limitados às capacidades tanto da câmera quanto do dispositivo de gravação utilizado. As tabelas a seguir apresentam os principais tipos de resoluções encontrados no mercado, para câmeras analógicas e digitais, com base nos padrões de transmissão de vídeo mais comuns: o NTSC (*National Television System Committee*), utilizado principalmente nos Estados Unidos, Canadá e alguns países da América do Sul e o PAL (*Phase Alternating Line*), utilizado principalmente na Europa, China, Austrália e Brasil.

Tabela D.1 − Resoluções de video analógicas (em pixels), (KING, 2009).

Formato	NTSC	PAL
QCIF	176 x 120	176 x 144
CIF	352 x 240	352 x 288
2CIF	704 x 240	704 x 288
4CIF	704 x 480	704 x 576
D1	720 x 480	720 x 576

Tabela D.2 − Resoluções de video digital (KING, 2009; AXIS, 2012).

Formato	Pixels
QQVGA	160 x 120
QVGA	320 x 240
VGA	640 x 480
SVGA	800 x 600
XVGA	1024 x 768
4x VGA	1280 x 960
SXGA	1280 x 1024
UXGA	1600 x 1200
WUXGA	1920 x 1200

E − RECOMENDAÇÕES DE ESPECIFICAÇÕES TÉCNICAS PARA CÂMERAS DE CFTV

As especificações de câmeras apresentadas a seguir seguem as recomendações do SWGIT (SWGIT, 2004):

a) Câmeras de vídeo fixas:

- policromáticas, padrão NTSC ou PAL-M;
- tecnologia de processamento digital do sinal (DSP);
- elemento captador de imagem do tipo CCD 1/3” ou 1/4”;
- número de pixels: 768(H)×494(V);
- resolução mínima vertical, em cores, de 520 linhas de vídeo;
- capacidade para operação com iluminação mínima de 0,3 Lux (colorida);
- sistema de sincronismo interno;
- obturador eletrônico automático, de 1/60 a 1/100.000 seg.;
- compensação de luz de fundo (*backlight*) automático;
- balanço do nível de branco automático (ATW);
- controle automático de ganho (AGC);
- relação sinal/ruído de, no mínimo, 50 dB (com AGC desligado);

b) Câmeras de vídeo fixas *day/night*:

- policromáticas, padrão NTSC;
- tecnologia de processamento digital do sinal (DSP);
- elemento captador de imagem do tipo CCD 1/3” ou 1/4”;
- número mínimo de pixels: 768(H)×494(V);
- resolução mínima vertical, em cores, de 520 linhas de vídeo;
- capacidade para operação com iluminação mínima de 0,2 Lux (colorida) e 0,01 Lux (P/B);
- sistema de sincronismo interno;
- obturador eletrônico automático, de 1/60 a 1/100.000 seg.;
- compensação de luz de fundo (*backlight*) automático;
- balanço do nível de branco automático (ATW);

- controle automático de ganho (AGC);
- relação sinal/ruído de, no mínimo, 50 dB (com AGC desligado);
- sensor *day/night*;

c) Câmeras dome para sistemas DVR:

- policromática, padrão NTSC ou PAL-M;
- elemento captador de imagem do tipo CCD 1/4";
- número mínimo de pixels: 768(H) x 494(V);
- zoom de no mínimo 22x (óptico) e 10x (digital);
- lente de 3,9 a 80mm, ou superior;
- resolução mínima vertical, em cores, de 520 linhas de vídeo;
- capacidade para operação com iluminação mínima de 0,2 Lux (colorida) e 0,01 lux no modo *day/night*;
- sistema de sincronismo interno;
- dispositivo "auto focus" com possibilidade de ajuste manual;
- função auto-íris;
- compensação de luz de fundo (BLC) automático;
- balanço do nível de branco automático (ATW);
- controle automático de ganho (AGC);
- relação sinal/ruído de no mínimo 50dB (com AGC desligado);
- sensor *day/night*;
- protetor em material de alta resistência, escurecido para operaçao discreta;
- sistema anti-embaçante incorporado.

F − ESPECIFICAÇÕES TÉCNICAS DOS EQUIPAMENTOS

Nas tabelas abaixo são apresentadas as especificações técnicas dos equipamentos utilizados durante os experimentos[14].

Tabela F.1 – Configuração do Sistema 1.

Sistema 1 – Notebook com placa de captura de áudio/vídeo usb	
Função	Gerenciamento do Sistema de CFTV. Os equipamentos de CFTV dos experimentos foram conectados a este equipamento: uma placa de captura de vídeo USB e uma câmera profissional analógica para CFTV.
Hardware (Configuração)	Notebook, marca LeNovo, modelo T60, contendo principalmente os seguintes componentes: • Processador Intel Core 2 Duo T7600 2,33 GHz; • 2 GB de memória RAM DDR2; • Placa de vídeo ATI MOBILITY Radeon X1400, 128 MB de RAM; • HD HITACHI de 100GB SATA; • Drive de CD/DVD genérico.
Sistema Operacional	Windows XP Professional com Service Pack 3 (32 bits).[15]
Acessório extra	Placa de captura de áudio e vídeo, marca Conexant, modelo 3104, com 4 saídas de vídeo. Suporta gravações somente em formato de vídeo, nas resoluções de 320x240 e 640x480. Foi instalado junto com esta placa, o programa SuperDVR para monitoramento das câmeras em utilização.

[14] A necessidade de utilização de dois computadores ao invés de apenas um ocorreu em virtude da incompatibilidade de determinados hardwares e softwares utilizados na sistemática dos experimentos, com os equipamentos e softwares inicialmente utilizados nos experimentos. A título de exemplo, a placa de captura utilizada para gravar as imagens geradas pela câmera só suportava sistemas operacionais Windows XP 32 bits.

[15] As configurações de sistema operacional utilizadas foram escolhidas de acordo com as limitações de compatibilidade da placa de captura USB. Ela não possuía *drivers* para quaisquer outros sistemas, mesmo da Microsoft, nem para versões de 64 bits.

Tabela F.2 – Configuração do Sistema 2.

Sistema 2 - Notebook	
Função	Realização das avaliações de medição dos experimentos.
Hardware (Configuração)	Notebook, marca Sony, modelo Vaio VPCF1, contendo principalmente os seguintes componentes: • Processador Intel i7 740 QM, 1,73 GHz; • 6 GB de memória RAM DDR3; • Placa de vídeo nVidia GeForce 330M, 1GB de RAM; • HD TOSHIBA de 500GB SATA; • Drive de CD/DVD genérico.
Sistema Operacional	Microsoft Windows 7 Home Premium com Service Pack 1 (64 bits).
Software extra	Matlab® versão R2011a (*Student Version*), 32 bits, contendo *image processing toolbox*. Este aplicativo foi utilizado para a execução dos processos de reconhecimento facial testados neste trabalho.

Tabela F.3 – Configuração da Câmera Digital Compacta.

Câmera digital compacta	
Marca	Sony
Modelo	DSC-TX100V (Linha CyberShot)
Lente	Zoom óptico 4x
CCD	Tipo CCD Sensor CMOS Exmor R™
Resolução	16,2 Mega Pixels (pixels efetivos)/ SteadyShot
Foco	Sistema de Foco Automático
Flash	Automático/Ligado/Sincronização lenta/Desligado
LCD	8,8 cm (tipo 3.5)
Memória externa	Suporte de gravação Memory Stick Duo™. Memory Stick PRO Duo™ (Mark II apenas para filmes). Memory Stick PRO Duo™ de alta velocidade, Memory Stick PRO-HG Duo™.

Tabela F.4 – Configuração da Câmera Profissional para CFTV.

Câmera Profissional para CFTV	
Marca	Intelbras
Modelo	480 S (Analógica)
Compensação da luz de fundo (BLC)	On/Off selecionável
Consumo	150 mA – 2 W (máximo)
Controle auto-íris	Vídeo/DC
Controle automático de ganho (AGC)	On/Off selecionável
Dimensões do produto (L x A x P)	61 x 58 x 104 mm
Lente	Montagem C/CS
Obturador eletrônico(auto)	1/60 a 1/100.000
Peso bruto	0,43 kg
Peso líquido	0,36 kg
Pixels efetivos	768 (H) x 494 (H)
Relação sinal/ruído	>48 dB
Resolução horizontal	480 linhas de TV
Saída de vídeo	1,0 Vpp – 75 ohms – conector BNC
Sem oscilação (Fickerless FLC)	On/Off selecionável
Sensibilidade	0,18 lux / F1.4
Sensor de imagem	CCD Sony 1/3” Super Had
Sistema de sincronização	Sincronização interna
Temperatura de armazenamento	-20ºC a 50ºC
Temperatura de operação	-10ºC a 50ºC (recomendado: -5ºC a 40ºC)
Tensão de alimentação	12 Vdc / 24 Vac (automático)
Varredura	NTSC

Tabela F.5 – Configuração do luxímetro.

Luxímetro digital minipa	
Marca	Minipa
Modelo	MLM-1011
Display	Display: LCD 3 ½ dígitos com leitura máxima de 1999, indicação x10 e x100
Ambiente de Operação	-10°C a 40°C (32°F a 104°F) com umidade relativa < 70%
Faixas	2000, 20000 (leitura x10), 100000 Lux (leitura x100)
Precisão	0 ~ 10000 lux ± (4%Leit.+0.5% f.s)
Resolução	1Lux, 10Lux, 100Lux

G − CÓDIGO FONTE MÉTODO EIGENFACES

O código fonte do pacote para Matlab utilizado para a realização dos testes experimentais está apresentado a seguir. O código foi desenvolvido por Luigi Rosa e outras informações podem ser obtidas através do sítio[16] WEB do autor.

```
%FACE RECOGNITION SYSTEM
%
% Face recognition system based on EigenFaces Method.
% The system functions by projecting face images onto a feature space
% that spans the significant variations among known face images. The
% significant features are known as "eigenfaces" because they are the
% eigenvectors (principal components) of the set of faces.
%
% Face images must be collected into sets: every set (called "class") should
% include a number of images for each person, with some variations in
% expression and in the lighting. When a new input image is read and added
% to the training database, the number of class is required. Otherwise, a new
% input image can be processed and confronted with all classes present in database.
% We choose a number of eigenvectors M' equal to the number of classes (see
% algorithmic details in the cited references). Before start image
% processing first select input image. This image can be succesively added to
% database (training) or, if a database is already present, matched with
% known faces.
%
% The images included are taken from AT&T Laboratories Cambridge's
% Face DataBase. See the cited references for more informations.
%
%
```

[16] http://www.advancedsourcecode.com/face.asp

```
% FUNCTIONS
%
% Select image:                          read the input image
%
% Add selected image to database:        the input image is added to database and will be used for
training
%
% Database Info:                         show informations about the images present in database. Images must
%                                        have the same size. If this is not true you have to resize them.
%
% Face Recognition:                      face matching. The selected input image is processed
%
% Delete Database:                       remove database from the current directory
%
% Info:                                  show informations about this software
%
% Visualization tool:                    visualization utility
%
% Source code for Face Recognition System:    how to obtain the complete source code
%
% Exit:                                  quit program
%
%
%  References:
%  "Eigenfaces for Recognition", Matthew Turk and Alex Pentlad
%  Journal of Cognitive Neuroscience pp.71-86, March 1991
%  Vision and Modeling Group, The Media Laboratory
%  Massachusetts Institute of Technology.
%  This paper is available at http://www.cs.ucsb.edu/~mturk/Papers/jcn.pdf
%  See also Matthew Turk's homepage http://www.cs.ucsb.edu/~mturk/research.htm
%
%  AT&T Laboratories Cambridge. The ORL face database, Olivetti Research Laboratory available at
%  http://www.uk.research.att.com/pub/data/att_faces.zip
```

```
%  or http://www.uk.research.att.com/pub/data/att_faces.tar.Z
%
%
%
%Please contribute if you find this software useful.
%Report bugs to luigi.rosa@tiscali.it
%
%
%*********************************************************************
% Luigi Rosa
% Via Centrale 35
% 67042 Civita di Bagno
% L'Aquila --- ITALY
% email  luigi.rosa@tiscali.it
% mobile +39 320 7214179
% website http://www.advancedsourcecode.com
%*********************************************************************
%
%


%--------------------------------------------------------------------
clear;
clc;
chos=0;
possibility=9;


messaggio='Insert the number of set: each set determins a class. This set should include a number of images for each person, with some variations in expression and in the lighting.';


while chos~=possibility,
   chos=menu('Face Recognition System','Select image','Add selected image to database','Database Info','Face Recognition','Delete Database','Info',...
       'Visualization tool','Source code for Face Recognition System','Exit');
```

```
%----------------
if chos==1,
   clc;
   [namefile,pathname]=uigetfile('*.*','Select image');
   if namefile~=0
      [img,map]=imread(strcat(pathname,namefile));
      imshow(img);
      dimensioni = size(img);
      disp('Input image has been selected.');
      disp('Now press on "Add selected image to database" button to add this image to database or,');
      disp('press on "Face Recognition" button to start face matching.');
   else
      warndlg('Input image must be selected.',' Warning ')
   end
end
%----------------
if chos==2,
   clc;
   if exist('img')
      if (exist('face_database.dat')==2)
         load('face_database.dat','-mat');
         face_number=face_number+1;
         data{face_number,1}=img(:);
         prompt={strcat(messaggio,'Class number must be a positive integer <= ',num2str(max_class))};
         title='Class number';
         lines=1;
         def={'1'};
         answer=inputdlg(prompt,title,lines,def);
         zparameter=double(str2num(char(answer)));
         if size(zparameter,1)~=0
            class_number=zparameter(1);
```

```
            if
(class_number<=0)||(class_number>max_class)||(floor(class_number)~=class_number)||(~isa(class_number,'
double'))||(any(any(imag(class_number))))

                warndlg(strcat('Class number must be a positive integer <= ',num2str(max_class)),' Warning ')

            else

                if class_number==max_class;

                   max_class=class_number+1;

                end

                data{face_number,2}=class_number;

                save('face_database.dat','data','face_number','max_class','-append');

                msgbox(strcat('Database already exists: image succesfully added to class number
',num2str(class_number)),'Database result','help');

                close all;

                clear('img')

            end

        else

            warndlg(strcat('Class number must be a positive integer <= ',num2str(max_class)),' Warning ')

        end

    else

        face_number=1;

        max_class=1;

        data{face_number,1}=img(:);

        prompt={strcat(messaggio,'Class number must be a positive integer <= ',num2str(max_class))};

        title='Class number';

        lines=1;

        def={'1'};

        answer=inputdlg(prompt,title,lines,def);

        zparameter=double(str2num(char(answer)));

        if size(zparameter,1)~=0

            class_number=zparameter(1);

            if
(class_number<=0)||(class_number>max_class)||(floor(class_number)~=class_number)||(~isa(class_number,'
double'))||(any(any(imag(class_number))))

                warndlg(strcat('Class number must be a positive integer <= ',num2str(max_class)),' Warning ')

            else
```

```
                max_class=2;
                data{face_number,2}=class_number;
                save('face_database.dat','data','face_number','max_class','dimensioni');
                msgbox(strcat('Database was empty. Database has just been created. Image succesfully added to class number ',num2str(class_number)),'Database result','help');
                close all;
                clear('img')
            end
        else
            warndlg(strcat('Class number must be a positive integer <= ',num2str(max_class)),' Warning ')
        end

    end
else
    errordlg('No image has been selected.','File Error');
end
end
%----------------
if chos==3,
    clc;
    close all;
    clear('img');
    if (exist('face_database.dat')==2)
        load('face_database.dat','-mat');
        msgbox(strcat('Database has ',num2str(face_number),' image(s). There are',num2str(max_class-1),' class(es). Input images must have the same size.'),'Database result','help');
    else
        msgbox('Database is empty.','Database result','help');
    end
end
%----------------
if chos==4,
    clc;
```

```
close all;
if exist('img')
   ingresso=double(img(:));
   if (exist('face_database.dat')==2)
      load('face_database.dat','-mat');
      % face_number is equal to "M" of Turk's paper
      % i.e. the number of faces present in the database.
      % These image are grouped into classes. Every class (or set) should include
      % a number of images for each person, with some variations in expression and in the
      % lighting.
      matrice=zeros(size(data{1,1},1),face_number);
      for ii=1:face_number
         matrice(:,ii)=double(data{ii,1});
      end
      somma=sum(matrice,2);
      media=somma/face_number;
      for ii=1:face_number
         matrice(:,ii)=matrice(:,ii)-media;
      end
      matrice=matrice/sqrt(face_number);
      % up to now matrix "matrice" is matrix "A" of Turk's paper
      elle=matrice'*matrice;
      % matrix "elle" is matrix "L" of Turk's paper

      % eigenvalues and eigenvectors of the "reduced" matrix  A'*A
      [V,D] = eig(elle);
      % the following multiplication is performed to obtain the
      % eigenvectors of the original matrix A*A' (see Turk's paper)
      % See also Karhunen-Loeve algorithm, for face recognition
      if det(D)~=0
         Vtrue=matrice*V*(abs(D))^-0.5;
      else
         Vtrue=matrice*V;
```

```
end
%Vtrue=matrice*V;
Dtrue=diag(D);

% the eigenvalues are sorted by order and only M' of them
% are taken. We impose M' equal to the number of classes
% (max_class-1)
[Dtrue,ordine]=sort(Dtrue);
Dtrue=flipud(Dtrue);
ordine=flipud(ordine);
Vtrue(:,1:face_number)=Vtrue(:,ordine);

Vtrue=Vtrue(:,1:max_class-1);
Dtrue=Dtrue(1:max_class-1);

% we calculate the eigenface components of
% the normalized input (mean-adjusted). I.e. the input
% image is projected into "face-space"
pesi=Vtrue'*(ingresso-media);

pesi_database         = zeros(max_class-1,max_class-1);
pesi_database_mediati  = zeros(max_class-1,max_class-1);

numero_elementi_classe=zeros(max_class-1,1);
for ii=1:face_number
   ingresso_database=double(data{ii,1});
   classe_database=data{ii,2};
   pesi_correnti=Vtrue'*(ingresso_database-media);
   pesi_database(:,classe_database)=pesi_database(:,classe_database)+pesi_correnti;
   numero_elementi_classe(classe_database)=numero_elementi_classe(classe_database)+1;
end
for ii=1:(max_class-1)
   pesi_database_mediati(:,ii)=pesi_database(:,ii)/numero_elementi_classe(ii);
```

```
end
% pesi_database_mediati is a matrix with the averaged eigenface components of the images
% present in database. Each class has its averaged eigenface.
% We want to find the nearest (in norm) vector to the input
% eigenface components.

distanze_pesi=zeros(max_class-1,1);
for ii=1:(max_class-1)
   % distanze_pesi(ii)=norm(pesi-pesi_database_mediati(:,ii));
   distanze_pesi(ii) = sum((abs(pesi-pesi_database_mediati(:,ii))));
end

[minimo_pesi,posizione_minimo_pesi]=min(distanze_pesi);

% now we are evaluating the distance of the mean-normalized
% input face from the "space-face" in order to detrmine if
% the input image is a face or not.
proiezione=zeros(size(data{1,1},1),1);
for ii=1:(max_class-1)
   proiezione=proiezione+pesi(ii)*Vtrue(:,ii);
end
distanza_spazio_facce=norm((ingresso-media)-proiezione);
```

```
messaggio1='See Matlab Command Window to see matching result. The program has just calculated the minimal distance from classes and the distance from Face Space. ';

messaggio2='You now should fix the two threshold-values to determine if this mathing is correct. If both distances are below the threshold values it means that the input ';

messaggio3='face was correctly matched with a known face. If the distance from Face Space is below the threshold value but the minimal distance from classes is above the other threshold value, ';

messaggio4=' it means that the input image is an unknown face. See the cited article for more informations.';
```

```
            msgbox(strcat(messaggio1,messaggio2,messaggio3,messaggio4),'Matching result','help');


            disp('The nearest class is number ');
            disp(posizione_minimo_pesi);
            disp('with a distance equal to ');
            disp(minimo_pesi);
            disp('The distance from Face Space is ');
            disp(distanza_spazio_facce);




        else
            warndlg('No image processing is possible. Database is empty.',' Warning ')
        end
    else
        warndlg('Input image must be selected.',' Warning ')
    end
end
%----------------
if chos==5,
    clc;
    close all;
    if (exist('face_database.dat')==2)
        button = questdlg('Do you really want to remove the Database?');
        if strcmp(button,'Yes')
            delete('face_database.dat');
            msgbox('Database was succesfully removed from the current directory.','Database removed','help');
        end
    else
        warndlg('Database is empty.',' Warning ')
    end
end
```

```
%----------------
if chos==6,
   clc;
   close all;
   helpwin facerecexplanation;
end
%----------------
if chos==7,
   clc;
   close all;
   if (exist('face_database.dat')==2)
      load('face_database.dat','-mat');
      disp('Insert 0 to visualize total mean face');
      disp('Insert 1 to visualize class mean face');
      disp('Insert 2 to visualize the projection of input image onto face-space');
      scelta = input('Insert your choice: ');
      if scelta == 0
         clc;
         matrice=zeros(size(data{1,1},1),face_number);
         for ii=1:face_number
            matrice(:,ii)=double(data{ii,1});
         end
         somma=sum(matrice,2);
         media=somma/face_number;
         figure('Name','Total mean face');
         imshow(uint8(reshape(media,dimensioni)));
      end
      if scelta == 1
         clc;
         classescelta = input('Insert class number:');
         if classescelta <= max_class-1
            somma     = zeros(size(data{1,1},1),1);
            contatore = 0;
```

```
        for ii=1:face_number
            if data{ii,2}==classescelta
                somma = somma + double(data{ii,1});
                contatore = contatore+1;
            end
        end
        somma = somma/contatore;
        figure('Name','Class mean face');
        imshow(uint8(reshape(somma,dimensioni)));
    else
        warndlg('Class number is uncorrect',' Warning ');
    end
end
if scelta == 2
    clc;
    [namefile,pathname]=uigetfile('*.*','Select image');
    if namefile~=0
        [img,map]=imread(strcat(pathname,namefile));
        imshow(img);
        dimensioni = size(img);
    else
        warndlg('Input image must be selected.',' Warning ')
    end
    ingresso=double(img(:));
    % face_number is equal to "M" of Turk's paper
    % i.e. the number of faces present in the database.
    % These image are grouped into classes. Every class (or set) should include
    % a number of images for each person, with some variations in expression and in the
    % lighting.
    matrice=zeros(size(data{1,1},1),face_number);
    for ii=1:face_number
        matrice(:,ii)=double(data{ii,1});
    end
```

```
somma=sum(matrice,2);
media=somma/face_number;
for ii=1:face_number
   matrice(:,ii)=matrice(:,ii)-media;
end
matrice=matrice/sqrt(face_number);
% up to now matrix "matrice" is matrix "A" of Turk's paper
elle=matrice'*matrice;
% matrix "elle" is matrix "L" of Turk's paper

% eigenvalues and eigenvectors of the "reduced" matrix  A'*A
[V,D] = eig(elle);
% the following multiplication is performed to obtain the
% eigenvectors of the original matrix A*A' (see Turk's paper)
% See also Karhunen-Loeve algorithm, for face recognition
Vtrue=matrice*V*(abs(D))^-0.5;
%Vtrue=matrice*V;
Dtrue=diag(D);

% the eigenvalues are sorted by order and only M' of them
% are taken. We impose M' equal to the number of classes
% (max_class-1)
[Dtrue,ordine]=sort(Dtrue);
Dtrue=flipud(Dtrue);
ordine=flipud(ordine);
Vtrue(:,1:face_number)=Vtrue(:,ordine);

Vtrue=Vtrue(:,1:max_class-1);
Dtrue=Dtrue(1:max_class-1);

% we calculate the eigenface components of
% the normalized input (mean-adjusted). I.e. the input
% image is projected into "face-space"
```

```
            pesi=Vtrue'*(ingresso-media);



            figure('Name','Projection of inpu image onto face-space');
            imshow(uint8(reshape(Vtrue*pesi+media,dimensioni)))
         end
      else
         warndlg('Database is empty.',' Warning ');
      end
   end
   %----------------
   if chos==8,
      clc;
      close all;
      web http://utenti.lycos.it/matlab/face.htm
      helpwin sourcecode;
   end
end
```