



ANAIS DA CONFERÊNCIA INTERNACIONAL DE CIÊNCIAS FORENSES EM MULTIMÍDIA E SEGURANÇA ELETRÔNICA

**18 a 21 de Setembro de 2012
Brasília/DF, Brasil**

ANAIS DA CONFERÊNCIA INTERNACIONAL
DE CIÊNCIAS FORENSES EM MULTIMÍDIA
E SEGURANÇA ELETRÔNICA

18 a 21 Setembro de 2012
Brasília-DF

Sumário

3	Apresentação
6	Tutorial: Perspectivas da Fonética Forense num Cenário de Quebra do Dogma da Unicidade
28	Artigos
29	Aplicação de Engenharia Reversa em Exames de Verificação de Edições de Vídeos Digitais
34	Identificação de Origem e Verificação de Edições em Filmes Fotográficos: Análise de Um Caso Real.
40	Detecção de Nudez Utilizando Redes Neurais Artificiais
44	Método para Compatibilidade entre Impressões Digitais com e sem Contato
49	Estudo de Performance de Algoritmos de Verificação de Impressões Digitais Aplicáveis a Smart-Card
57	Avaliação de Técnicas de Seleção de Quadros Chave na Recuperação de Informação por Conteúdo Visual para Identificar Crimes Cibernéticos em Conteúdo Multimídia
64	Reconstrução de Trajetória de Voo para Simulações 3D em Sinistros Aeronáuticos
72	FI2 - Uma Proposta de Ferramenta para Investigação Forense em Imagens
78	Exames de Reconhecimento Facial na Polícia Federal
86	Estimativa de Altura de Indivíduos em Imagens de Perspectiva Cônica com Distorções Esféricas (Caso Prático)
94	On the Practical Aspects of Applying the PRNU Approach to Device Identification Tasks
100	Avaliação de Classificadores Haar Projetados para Detecção Facial
107	Identificação de Câmeras de Celulares Usando Sensor Pattern Noise

Apresentação

É com muita satisfação que reunimos nesta obra os artigos que foram submetidos à I Conferência Internacional de Ciência Forenses em Multimídia e Segurança Eletrônica e selecionados por sua relevância no desenvolvimento e na divulgação do trabalho pericial correlatado à análise forense de evidências multimídia e à implementação de sistemas de segurança eletrônica.

Fruto do trabalho de Peritos Criminais e pesquisadores, os artigos selecionados são reflexo de uma iniciativa de grandioso valor que busca a disseminação e uniformização do conhecimento na realização de exames periciais, a integração entre os especialistas forenses e o desenvolvimento de técnicas, métodos e doutrinas relacionados a prática pericial na realização de exames que envolvam a análise de evidências multimídia e sistemas de segurança eletrônica.

Sendo assim, podemos destacar alguns estudos de caso onde Peritos Criminais relatam a solução de problemas importantes na prática da atividade pericial, envolvendo a análise de conteúdo de evidências multimídia, como vídeos, imagens e registros de áudio, bem como a busca de indícios de edição. A realização de tais estudos serve não somente para a divulgação dos trabalhos, mas também para a discussão de outras

abordagens e soluções, bem como para a uniformização de procedimentos, sempre observando as melhores práticas.

Também são apresentados trabalhos de pesquisa envolvendo proposição de soluções para diferentes problemas, com suas vantagens e desvantagens, assim como novas abordagens para problemas em aberto. Nesse contexto, são apresentados trabalhos relacionados à utilização e ao desempenho de sistemas de biometria por impressões digitais, biometria facial, detecção de nudez e identificação de origem de imagens estáticas.

Alguns trabalhos relacionados à metodologia empregada em exames periciais envolvendo a análise de evidências multimídia foram realizados, funcionando como catalizadores para uma evolução sistemática e paulatina das práticas periciais, visando um melhor aproveitamento da prova material.

Este compêndio traz ainda um excelente tutorial relacionado a metodologia empregada em exames na área de fonética forense. O tutorial intitulado *Perspectivas da Fonética Forense num Cenário de Quebra do Dogma da Unicidade*, aborda um tema de extrema importância para a Criminalística como um todo, e de grande atualidade.

Gostaríamos, por último, de agradecer a todos os autores que, com sua participação, contribuíram para a produção de um evento de qualidade.

Comissão Técnica da ICMedia 2012

Tutorial

Perspectivas da fonética forense num cenário de quebra do dogma da unicidade

Charles Rodrigues Valente

SEPAEL, Instituto Nacional de Criminalística

Brasília, DF, CEP 70.610-902, Brasil

charles.crv@dpf.gov.br

RESUMO – Na perícia oficial brasileira ainda há pouca discussão a respeito da chamada “mudança de paradigma” na Criminalística, com sua crítica ao dogma da unicidade. Este tutorial apresenta os principais argumentos que embasam as propostas recentes de uma nova abordagem para a comparação forense de locutor, além de também discutir a utilização de razões de verossimilhança (LR, *Likelihood Ratios*) nesse tipo de exame.

1. INTRODUÇÃO

Muitas áreas da Criminalística envolvem o que chamaremos aqui de problema da determinação da fonte ou origem: dado um vestígio – o material questionado –, uma pessoa ou objeto suspeitos de o terem produzido (ou terem sido utilizados na sua produção) são apresentados à perícia, para obtenção de material de mesma natureza do material questionado, o qual constituirá o material padrão. Em alguns casos, o material padrão pode ter sido obtido por outras instâncias que não a pericial, sendo posteriormente encaminhado aos peritos.

O que se pergunta à perícia é: a pessoa ou objeto suspeitos produziram o vestígio?

São exemplos desse tipo de exame:

- um confronto microbalístico no qual se investiga se um projétil foi disparado por uma determinada arma de fogo;
- um exame grafoscópico (ou grafotécnico) que avalia se lançamentos manuscritos questionados foram produzidos por uma determinada pessoa;
- um confronto datiloscópico que procura atribuir (ou não) a produção de um fragmento de impressão digital a uma determinada pessoa;

- um exame de DNA forense que investiga se um resíduo biológico foi produzido por um determinado suspeito;
- um exame de reconhecimento de padrão que procura determinar se uma marca de solado provém de determinado calçado;
- um exame de reconhecimento facial em que se compara a imagem da face de uma pessoa em um vídeo de um sistema de vigilância com a fotografia de um suspeito obtida em situação controlada;
- um exame de fonética forense em que se avalia se as falas em uma gravação questionada foram emitidas por uma determinada pessoa.

Os exames relacionados à determinação da fonte são tradicionalmente abordados a partir de uma metodologia de comparação de características que tem os seguintes passos gerais:

- 1) análise prévia do material questionado a fim de verificar se ele contém informação suficiente para possibilitar a comparação;
- 2) obtenção de material padrão apropriado, ou análise do material enviado à perícia a título de padrão para verificar se é utilizável. Em geral, esse estágio leva em conta requisitos referidos nos textos de Criminalística como “adequabilidade”, “quantidade”, “contemporaneidade”, “autenticidade” e “espontaneidade” (por exemplo, ver: [1], p. 203-204; [2], p. 300-314; [3], p. 286-287);

- 3) extração ou levantamento de um determinado conjunto de características julgadas discriminadoras do material padrão e do material questionado;
- 4) comparação das características levantadas e conclusão a respeito da coincidência ou não da origem do material padrão e do material questionados.

A respeito dessa metodologia geral, há duas questões subjacentes de grande importância e que estão inter-relacionadas:

Questão 1: que embasamento científico os respectivos ramos da Criminalística apresentam para concluir sobre a questão da determinação da fonte ou origem a partir dessa metodologia geral?

Questão 2: que tipo de conclusão é possível apresentar? Quantitativa, qualitativa, categórica, escala de probabilidade?

2. IDENTIDADE QUALITATIVA E IDENTIDADE NUMÉRICA

Para tentar responder às questões propostas ao final da introdução, é necessário primeiramente distinguir duas operações lógico-conceituais que, normalmente, na determinação forense da origem de um vestígio, tanto nos laudos periciais quanto em muitos livros e periódicos da área, são confundidas.

2.1. IDENTIFICAÇÃO E INDIVIDUALIZAÇÃO

Usando a nomenclatura encontrada em [4], p. 206, essas duas operações são denominadas “identificação” (ou “identificação qualitativa”, como veremos a seguir) e “individualização” (ou “identificação numérica”, como também veremos a seguir).

A identificação seria o agrupamento de um conjunto de seres em uma classe com base no levantamento de características comuns entre esses seres. Já a individualização seria a operação de destaque de um único ser.

Deve-se notar que, na língua portuguesa, no seu uso comum, “identificação”, bem como cognatos como “identidade”, são termos ambíguos, que podem ter, conforme o contexto, qualquer um dos dois significados. Quando,

por exemplo, alguém diz “eu me identifico com *Fulano de Tal*”, quer dizer que vê entre si e *Fulano de Tal* um conjunto de características comuns (preferências, visão de mundo, atitudes etc.) que permitem colocar a si e *Fulano de Tal* em uma mesma classe, e esse significado de “identificação” está relacionado ao primeiro conceito definido no parágrafo anterior. Já quando, por exemplo, no noticiário é dito que “o suspeito do crime foi identificado”, o significado de identificação é (ou deveria ser) o que se denominou no parágrafo anterior de “individualização”.

Se a análise pericial demonstrar que propriedades do material questionado e do padrão indicam que o ente (pessoa ou objeto) que os produziu tem um certo conjunto de características comuns, terá, a rigor, determinado uma classe de seres que poderiam dar origem ao vestígio, e realizado uma identificação, não uma individualização. Essa identificação está relacionada, conforme [4], p. 206, ao que também é denominado identidade qualitativa dos membros da classe.

Por outro lado, se a análise da origem apontar uma pessoa ou objeto específico como aquele que originou o vestígio, o que se realizou foi uma individualização, a qual está relacionada ao que se denomina identidade numérica daquela pessoa ou objeto específico.

Pela própria natureza das duas operações, vê-se que a identidade qualitativa, no caso da investigação de origem, pode ser estabelecida a partir do levantamento de características comuns entre dois entes ou, ainda, pelas características comuns entre traços ou marcas por eles produzidas. Já a identidade numérica só pode ser estabelecida, a rigor, pela garantia da continuidade no tempo, no caso, entre o instante em que o ser produziu o vestígio e o instante em que a sua identidade foi atestada.

Portanto, da identidade qualitativa não segue, por dedução, a identidade numérica.

2.2. ESCOLHA DE CARACTERÍSTICAS DEFINIDORAS DA CLASSE

Deve-se notar que o estabelecimento da identidade qualitativa – processo que chamare-

mos aqui de “identificação qualitativa” – está umbilical e necessariamente associado à escolha de um modelo e/ou de um conjunto de características as quais são usadas para definir a classe de seres.

Essa escolha pode implicar, dependendo do tipo das características classificadoras, em ter outro nível de escolha embutido: a determinação, qualitativa ou quantitativa, de quando se considera a característica presente ou não, ou a adoção de uma escala de gradação que vai da ausência ou presença, em grau máximo, da característica.

No caso da perícia de determinação da fonte, o que se quer dizer é que o resultado da perícia é diretamente dependente dos parâmetros de comparação e da métrica escolhida para mensurar a distância entre o material questionado e o material padrão.

Esse processo de escolha não é pressuposto lógico necessário no estabelecimento da identidade – ou identificação – numérica.

2.3. CONTINUIDADE NO TEMPO

A continuidade no tempo não é uma questão alheia ao contexto jurídico-policial e pericial. Por exemplo, ela é a preocupação central quando se fala em cadeia de custódia ([4], p. 207), ou seja, quando se implementam mecanismos que busquem garantir que o vestígio colhido na cena de crime ou apresentado como prova permaneça o mesmo ao longo de toda a cadeia de persecução penal.

E ainda que muitos manuais tradicionais de Criminalística não explicitem a diferença lógica entre a identificação qualitativa e a identificação numérica, ela está implícita em relação à continuidade no tempo em um dos requisitos desejáveis do material padrão: a contemporaneidade em relação ao material questionado (ver, por exemplo: [2], p. 300 e 308; [1], p. 203-204; [3], p. 286 e 292).

Exceto em exames associados a características tidas como perenes ou de alteração apenas em eventos drásticos – como é o caso do material genético e dos desenhos papilares, respectivamente – o requisito da contemporaneidade

reconhece que a comparação de características, que é o cerne da metodologia geral apresentada na Introdução, não acompanha no tempo o objeto ou indivíduo que produziram o vestígio.

2.4. HIERARQUIA DE PROPOSIÇÕES

Como visto, a análise pericial da questão da origem feita com base na metodologia geral apresentada na Introdução não é capaz, por uma limitação lógica, de realizar a identificação numérica. Porém, a identificação numérica é um pressuposto, ao menos teórico, no qual se fundamenta a imputação penal. Essa constatação poderia levar a concluir que a perícia tem uma contribuição pequena para a Justiça, o que é obviamente incorreto.

Em primeiro lugar, o que se deve ter em mente é que a perícia, apesar de ser um elemento importante na cadeia de persecução penal, não retira do órgão julgador a responsabilidade última pela decisão de mérito, em particular sobre a fixação penal da autoria do delito.

Além disso, deve-se ter em mente que a perícia não se restringe à aplicação da metodologia básica de investigação da origem.

Uma maneira de situar o trabalho da perícia na cadeia de persecução penal é a proposta de hierarquia de proposições contida em [5], que é comentada, por exemplo, em [6], p. 118-120.

Com base nesse artigo, adaptando a abordagem para o enfoque aqui desenvolvido, pode-se dizer que as hipóteses ou proposições passíveis de serem confrontadas pela perícia são classificáveis em três níveis:

- nível I ou nível da fonte (*source*): é aquele em que as proposições ou hipóteses se referem à determinação da fonte ou origem do material questionado. O exame é centrado na análise do vestígio e do material padrão;
- nível II ou nível da atividade (*activity*): é aquele em que as proposições se referem à efetiva presença do suspeito no tempo, lugar e situação em que se deu a ação ou omissão objeto de questionamento;

- nível III ou nível do delito (*offence*): é aquele em que a perícia trata de hipóteses referentes à prática da ação ou omissão pelo suspeito.

Do nível I para o nível III cresce a aproximação entre as proposições e o estabelecimento da autoria delitiva.

É importante salientar que essa hierarquia não é um esquema rígido (ver [5], p. 233), podendo haver superposição de níveis. Além disso, o objetivo final dos autores daquele artigo é possibilitar uma análise baseada em razões de verossimilhança (LR, *likelihood ratios*), que veremos mais à frente. Por esse motivo, as hipóteses ou proposições, em cada nível, são sempre analisadas em pares antagônicos, sendo também importante distinguir entre proposições e explicações (ver [7]).

Seja, como exemplo, um caso de furto qualificado, em que são encontradas marcas de calçado no local de crime. A conclusão de um exame de reconhecimento de padrão que contrasta essas marcas questionadas com aquelas de um calçado do suspeito será de nível I.

O nível I é o campo por excelência da perícia. No caso da investigação da origem, esse é o nível em que se situa a metodologia geral descrita na Introdução e mesmo qualquer outra metodologia que tenha o mesmo objetivo, como as baseadas em razões de verossimilhança.

A rigor, no nível I, não está em questão a continuidade no tempo. Portanto, no nosso exemplo, mesmo que houvesse sido estabelecido que as marcas encontradas no local de crime vieram do calçado do suspeito (veremos mais a frente qual o “grau de certeza” possível de se ter numa afirmação desse tipo), isso não garantiria que foi o suspeito quem adentrou o local do furto, pois, por exemplo, outra pessoa poderia estar usando aquele calçado, emprestado pelo, ou subtraído do, suspeito.

Dessa forma, a perícia não pode afirmar, com base apenas no vestígio em questão e no material padrão, que efetivamente o suspeito estava no local, tempo e circunstância em que ocorreu o furto. Para isso, seria necessária in-

formação de natureza circunstancial, ou seja, informação que agregasse elementos de continuidade no tempo à análise. Proposições desse tipo já seriam do nível II na hierarquia proposta em [5].

Progredindo na hierarquia, no nível III tem-se que as hipóteses a serem confrontadas dizem respeito ao cometimento, pelo suspeito, da ação ou omissão descritas no tipo penal. O trabalho da perícia nesse nível tangencia o domínio que é, por excelência, do julgador. Acima do nível III estão as questões jurídico-penais propriamente ditas: intenção (dolo e culpa), excludentes de ilicitude etc.

No sistema processual brasileiro, e na própria cultura da sua perícia criminal oficial, o perito não é incentivado a incluir na sua análise elementos que vão além dos vestígios propriamente ditos, e, principalmente, informação contextual, e portanto extracientífica, que, apesar de não necessariamente incorreta, depende de outros tipos de prova (por exemplo, a testemunhal) ou mesmo da assunção de elementos circunstanciais.

O fato de [5] incluir a participação da perícia em considerações de nível II e, principalmente, nível III que não seriam da alçada pericial na tradição brasileira talvez se explique porque, na tradição processual dos autores daquele artigo – o Reino Unido – os peritos sejam costumemente chamados para testemunhar em juízo e opinar sobre esses níveis, sendo esse testemunho parte integrante do trabalho pericial. No Brasil, a perícia se manifesta fundamentalmente por meio do laudo pericial, podendo ser ouvida em audiência apenas para esclarecer sobre o exame (art. 159, § 5º, inciso I e art. 400, *caput* e § 2º, todos do Código de Processo Penal – CPP, [8]).

Além disso, como fica evidenciado em [7], mesmo no nível III a perícia avalia pares de hipóteses passíveis de possibilitar uma análise estatística, sendo isso distinto da mera emissão de uma opinião.

Em todo caso, talvez seja prudente considerar que conclusões de nível II e III, a rigor, só serão da alçada pericial quando os elementos

circunstanciais forem colhidos pela própria perícia, que, munida de sua fé pública, pode agregá-los nas considerações a serem feitas no laudo pericial.

No exemplo da marca de calçado, a existência de prova testemunhal que relacione o suspeito ao local e ao horário provável de ocorrência do furto é um elemento circunstancial relacionado ao nível II, porém não de natureza assimilável a prova pericial.

Nada impede, entretanto, que essa informação seja indiretamente incluída na análise pericial por meio, por exemplo, de um quesito que pede um estudo de plausibilidade de uma determinada dinâmica de eventos que inclui o suspeito como o agente delituoso.

Em contraste, um exemplo de proposição efetivamente pericial de nível II seria a que concluisse, num exame de local em suspeita de crime de telecomunicação, por meio de constatação *in loco* realizada pela equipe de peritos, que uma estação clandestina operada por um determinado indivíduo transmitia informação por processo eletromagnético.

2.5. RESUMO DAS CARACTERÍSTICAS DOS DOIS TIPOS DE IDENTIFICAÇÃO

A Tabela 1 contrapõe a identificação qualitativa e a identificação numérica tendo em vista o que foi apresentado até aqui.

Os exames periciais de determinação da fonte se baseiam, como regra geral, em análise de material questionado e padrão, circunscrevendo-se ao nível I (nível da fonte) da hierarquia de proposições e constituindo uma atividade de identificação qualitativa.

Esse é o caso de todos os exemplos dados na Introdução quando a análise se basear em metodologias de comparação entre vestígio e material padrão, sem preocupação direta com questões de persistência e transferência de marcas e traços.

Em particular, é o caso da comparação forense de locutor (último exemplo dado na Introdução).

TABELA 1
IDENTIFICAÇÃO QUALITATIVA VERSUS IDENTIFICAÇÃO NUMÉRICA

Tipo de identificação	Qualitativa	Numérica
Resultado	O resultado é uma classificação (a fonte do vestígio é – ou não – da mesma classe da fonte do padrão).	O resultado é uma individualização.
Número de seres discriminados	Variável conforme o caso, podendo ser qualquer inteiro maior ou igual a 1.	1 (um).
Dependência em relação à escolha de parâmetros de classificação/comparação	Dependente intrinsecamente.	Não utiliza parâmetros de classificação.
Relação com a continuidade no tempo	Não pressupõe mecanismo de garantia de continuidade. O resultado pode variar se a fonte do material questionado se alterar.	A continuidade no tempo é pressuposto essencial.
Correlação com a hierarquia de proposições	Relacionada a proposições de nível I (fonte).	Relacionada a proposições de nível II (atividade) e III (delito).

Deve ficar claro que as questões de persistência e, em particular, transferência podem ser, e normalmente são de fato, pericialmente importantes na aceitação, como vestígio, do áudio questionado. Porém, isso não é parte do exame de comparação forense de locutor, mas do exame de verificação de edição (ver [9], p. 299; [3], p. 284-285).

3. O “DOGMA DA UNICIDADE”

A diferença entre identificação qualitativa e identificação numérica teria pouca relevância prática na perícia se fosse possível demonstrar – para os exames de determinação da fonte circunscritos ao nível I da hierarquia de proposições – que a metodologia de análise, ainda que em condições específicas, seria capaz de determinar classes que contivessem um único elemento.

Ou, equivalentemente, se fosse possível demonstrar o que chamaremos aqui de unicidade: apenas um único ente (pessoa ou objeto) poderia produzir o vestígio com as características levantadas pela análise pericial.

Juntando-se à unicidade o requisito da contemporaneidade do material padrão, que garan-

tiria a continuidade no tempo, chegar-se-ia à conclusão de que o exame de investigação da fonte realizado conforme a metodologia de comparação delineada na Introdução, que é uma identificação qualitativa, faria as vezes, para os fins práticos, de uma individualização.

3.1. A “FALÁCIA DA INDIVIDUALIZAÇÃO”

Não se analisarão aqui as questões relativas ao requisito da contemporaneidade, que dependem de considerações particulares para cada tipo de vestígio.

No que se refere ao pressuposto da unicidade, verifica-se que, em geral, na Criminalística ele não somente é assumido como hipótese de trabalho provisória, a ser colocado sob teste empiricamente, mas como verdade óbvia e evidente (ver [10] e [11]).

O que se observa em muitos textos devotados ao tema da Criminalística - e não somente no Brasil - bem como na prática diária de várias áreas da perícia, é a assunção da unicidade como um verdadeiro dogma.

Sua base consiste de um raciocínio indutivo que iguala infrequência a unicidade, fundamentado no acúmulo de experiência pessoal não sistematizada, reforçado pelas afirmações de estudiosos vistos como baluartes nas suas respectivas áreas (argumento de autoridade).

Além disso, como afirmado em [4], p. 208, a hipótese da unicidade da fonte não só é assumida como verdade mas é também transferida, sem questionamento, para o vestígio produzido.

Configura-se, assim, o que [11] chama de “falácia da individualização”.

Por exemplo, para justificar a capacidade de determinação individualizadora da perícia grafoscópica, um texto da área afirma:

A escrita é individual. A escrita é resultante de estímulos cerebrais que determinam movimentos e estes criam as formas gráficas.

Muito embora os cérebros de todos sejam anatomicamente iguais, a sua função varia de pessoa para pessoa. O mesmo ocorre com o sistema somático. Vale dizer, portanto, que ambos tendem **vari**ar (sic) **ao infinito**. Como a escrita resulta do concurso desses dois

sistemas, evidentemente ela também **varia ao infinito**.

Se assim não fosse, a perícia grafotécnica, que é aceita universalmente, não teria o menor valor.

([1], p. 29; grifos em negrito meus).

De modo semelhante, mas em relação aos exames comparativos na balística, outro texto afirma:

(...) assim como é **pacífico** que duas impressões digitais cujos desenhos coincidam exatamente só podem corresponder a um mesmo dedo, assim também pode-se ter como **indiscutível** que a presença de deformações normais convergentes, em dois ou mais projetis, significa que foram todos expelidos por **um só e mesmo** cano raiado.

([2], p. 316; grifos meus).

Para exemplos semelhantes na literatura em inglês, ver [11], p. 203-206.

3.2. PREMISSAS DA “FALÁCIA DA INDIVIDUALIZAÇÃO”

Como discutido em [10] e [11], há diversas razões, inclusive de cunho sociológico, que podem explicar a presença razoavelmente disseminada da “falácia da individualização”, ou seja, da afirmação categórica da capacidade individualizadora dos exames baseados em comparação de características, na Criminalística.

Destacam-se a seguir duas das justificativas levantadas em [11] e que podem ser observadas nas duas citações diretas feitas acima.

3.2.1. O NÚMERO DE COMBINAÇÕES POSSÍVEIS

A primeira diz respeito ao número de combinações possíveis de determinadas características. Ainda que esse número seja grande, indicando que a possibilidade de coincidências é baixa, a matemática e a estatística não permitem concluir que ela pode ser assumida, mesmo para os fins práticos, como sendo nula.

O conhecido “paradoxo do dia de aniversário” (ver, por exemplo, [12], p. 64) mostra que o cálculo de probabilidades apresenta muitas vezes resultados que não são intuitivos. Assim é que, apesar de haver, desconsiderando-se os bissextos, 365 dias em um ano, basta tomar 23 pessoas ao acaso para que a probabilidade de que pelo menos duas delas façam aniversário no mesmo dia seja maior do que 50% (na verdade, essa probabilidade será de aproximada-

mente 50,73%). Com 57 pessoas, a probabilidade de ter pelo menos uma coincidência ultrapassa 99%.

Esses números – 23 e 57 – são, respectivamente, duas e uma ordem de grandeza menores do que o universo de possibilidades, 365.

No caso geral, havendo um universo de k possibilidades equiprováveis de combinações de um conjunto de características e n seres que apresentam uma dessas combinações, a probabilidade p de que pelo menos dois desses seres tenham as mesmas características é:

$$p = 1 - p_2$$

onde p_2 é a probabilidade de que não haja nenhuma coincidência, ou seja,

$$p_2 = \frac{k}{k} \cdot \frac{k-1}{k} \cdot \dots \cdot \frac{k-(n-1)}{k} = \frac{k!}{(k-n)! \cdot k^n}$$

e, portanto,

$$p = 1 - \frac{k!}{(k-n)! \cdot k^n} \quad (1)$$

Pode-se argumentar que, na prática, dada a enorme quantidade de possibilidades de combinações de um determinado conjunto de características – por exemplo, os pontos característicos em um fragmento de impressão digital –, o cálculo de probabilidade de coincidências ao acaso utilizando (1) (ou qualquer expressão semelhante) resultaria em valores ínfimos para p , que permitiriam, mesmo considerando toda a população de fontes possíveis no planeta, fazer uma identificação individualizadora categórica.

Entretanto, muitas características envolvidas em problemas de determinação da fonte comuns na Criminalística não têm o seu mecanismo satisfatoriamente descrito pela ciência.

Cálculos de número de combinações possíveis via de regra fazem suposições como a de independência de variáveis e distribuição uniforme de possibilidades. Porém, se o mecanismo físico de determinação dessas características não foi cientificamente levantado de modo satisfatoriamente completo, pode haver, na realidade, variáveis ainda não descobertas que fazem determinadas combinações mais prováveis do que outras.

Outros processos – como por exemplo, a aquisição de marcas de uso em solados de calçados e de marcas de disparo em projetis de armas de fogo de alma raiada – muitas vezes carecem de modelamentos e levantamentos estatísticos que permitam apresentar conclusões em exames periciais de modo mais objetivo.

Isso leva a uma consequência, que é parte da boa prática científica, mas, surpreendentemente, tem ainda pouca ressonância na cultura de muitas áreas periciais: a necessidade de realizar levantamentos empíricos e validações estatísticas, ao invés de se assumir a unicidade como uma premissa inquestionável.

Mesmo quando os fundamentos básicos do mecanismo sob análise são bem estabelecidos, como no caso do DNA, a estatística, vale lembrar, não leva à individualização. Por sua própria natureza, o que ela permite é a quantificação objetiva de uma inferência. O exame pericial de investigação da fonte continua sendo uma identificação qualitativa ([13], p. 199; [4], p. 209). A rigor, somente a verificação de todo o espaço amostral poderia levar à individualização ([11], p. 7).

Apesar da simplicidade da consideração embutida na equação (1), é importante reforçá-la. Muitas áreas da Criminalística, ainda hoje, desenvolvem raciocínios equivocados que pretendem embasar conclusões individualizadoras para seus respectivos exames a partir da mera determinação do número possível de combinações de um conjunto de características (para um exemplo desse equívoco, ver [14]).

3.2.2. CONHECIMENTO “PACÍFICO” E “INDISCUTÍVEL”

Outra justificativa subjacente à defesa apriorística da unicidade e da capacidade individualizadora dos exames de investigação da origem, mesmo no nível I da hierarquia de proposições, é a inocorrência de coincidências, ou seja, de seres diferentes produzindo vestígios indistinguíveis, observada pelos peritos ao longo de sua experiência profissional.

Como destacado em [11], p. 212-213 e em [10], há vários problemas nesse raciocínio.

Um deles é lógico: a indução pode – e é – usada cientificamente na formulação de hipóteses, porém o acúmulo de observações não prova a hipótese. Na realidade, basta uma observação em discordância com a hipótese para colocá-la em cheque (conceito da falseabilidade das teorias científicas, proposto por Karl Popper).

Afirmativas sobre a capacidade de individualização de exames criminalísticos deveriam, ao menos, ser mais modestas, sob o risco de pretenderem colocar a Criminalística “acima” das próprias Ciências Naturais.

Outro problema é que, no cenário de diversas áreas das chamadas ciências forenses, vê-se poucos estudos dedicados à verificação, embasada no método científico, da hipótese da individualização. Baixo nível de pesquisas formais e mesmo de esforço concentrado em busca de falsos positivos, além da ausência de levantamentos extensos e bancos de dados, parecem ser mais a regra do que a exceção, incluindo na área de comparação forense de locutor.

Epistemologicamente, todas as Ciências Naturais têm no raciocínio indutivo um dos seus pilares e, do ponto de vista sociológico, os cientistas, na prática do dia a dia, também podem tender a apropriar como dogma as hipóteses e teorias de ampla aceitação na comunidade científica em dado momento histórico, até por uma questão de eficiência (ver [15] ou [16]).

Porém, a cultura científica, ao mesmo tempo, também destaca a necessidade de testes sistematizados e objetivos no processo de estabelecimento de hipóteses.

Considerando, além disso, que em muitas áreas, incluindo a comparação de falantes, como veremos mais adiante, ainda não há parâmetros objetivos quantitativos para todos os aspectos levados em conta na análise, seria desejável implantar uma cultura de incentivo a testes cegos de proficiência dos grupos periciais ([10]; [11], p. 201-202).

4. RAZÃO DE VEROSSIMILHANÇA

Abandonado o dogma da unicidade, a lógica inferencial de base estatística se apresenta como a ferramenta preferencial para a solução,

em bases cientificamente mais sólidas, do problema da determinação da fonte.

Considerando o nível I da hierarquia de proposições, a metodologia tradicional de comparação de características entre os materiais questionado e padrão deverá ser complementada pela avaliação estatística da distribuição dessas mesmas características no restante da população de entes que, potencialmente, poderiam ter produzido ou ter sido utilizados na produção do vestígio.

O ideal, portanto, para a abordagem do problema da determinação da origem, não é apenas estabelecer o grau de compatibilidade entre as marcas ou traços do vestígio e as marcas ou traços do material padrão – ou seja, o seu grau de similaridade –, mas também avaliar estatisticamente se as características observadas no vestígio poderiam ter-se originado de outros seres de uma população de referência, ou seja, o grau de tipicidade do conjunto de características do vestígio.

Por exemplo, um alto grau de similaridade entre vestígio e material padrão é, em princípio, favorável a uma conclusão positiva a respeito da identidade qualitativa entre a fonte do material questionado e do material padrão. Porém, se o conjunto de características do vestígio for muito comum, ou seja, seu grau de tipicidade for grande, isso será desfavorável à conclusão positiva a respeito da identidade qualitativa.

Considera-se atualmente que a metodologia mais adequada para realizar os exames de determinação da fonte segundo esses princípios é a que utiliza razões de verossimilhança no contexto da análise bayesiana de dados.

4.1. ANÁLISE BAYESIANA DE DADOS

A abordagem, teoria ou interpretação bayesiana da probabilidade, bem como sua faceta operacional, a chamada análise bayesiana de dados, é uma das teorias matemáticas que tratam do tema da probabilidade. Na realidade, não é totalmente correto falar em uma única “interpretação bayesiana da probabilidade”, pois pode haver pequenas diferenças de visão

sobre aspectos específicos da probabilidade de acordo com o autor consultado.

Neste tutorial, adota-se em linha geral, chamando de “abordagem, teoria ou interpretação bayesiana”, a encontrada em [17] e [18].

A teoria bayesiana pretende ser um sistema lógico que permite atribuir um peso ou grau à plausibilidade de eventos incertos. A incerteza deriva da ausência de informação que permita determinar por dedução e seguramente se o evento ocorre ou não. A probabilidade, na visão bayesiana, é, portanto, um estado de conhecimento, não uma entidade do mundo físico, externa ao observador ([19], p. 10-11).

Apesar de haver um foco no observador, o que faz com que em geral se classifique a abordagem bayesiana como uma teoria subjetiva da probabilidade, ela permite atingir resultados objetivos definidos, ou seja, inferências de caráter conclusivo obtidas por dedução, a partir de inferências tomadas como hipóteses de trabalho indutivas, as quais são também informadas pelos dados observados. Vê-se, desde logo, o papel importante que a probabilidade condicional desempenha na análise bayesiana.

Para uma discussão aprofundada da teoria bayesiana, ver [18].

A teoria bayesiana se contrapõe, tanto do ponto de vista filosófico quanto do ponto de vista operacional, a muitos dos aspectos da chamada teoria frequentista da probabilidade, que, historicamente, norteou os estudos da probabilidade na maior parte das gerações de matemáticos nos séculos XIX e XX.

A abordagem bayesiana é centrada no teorema de Bayes, que recebe seu nome do matemático e reverendo inglês Thomas Bayes, o qual, em um artigo publicado postumamente em 1763, apresenta expressões e raciocínios relacionados à probabilidade condicional.

Entretanto, na realidade, a formulação definitiva, e utilizada atualmente, do chamado teorema de Bayes deve-se a Laplace (1749-1827), que, em 1812, publicou um estudo no qual desenvolvia vários aspectos de uma teoria da probabilidade propriamente bayesiana.

Nas décadas seguintes, entretanto, os princípios do trabalho de Laplace foram colocados em xeque, sob o argumento de que careciam de rigor matemático, por estarem baseados em uma noção intuitiva e subjetivista da probabilidade.

Prevaleceu, a partir daí, a chamada teoria frequentista, onde a probabilidade de um evento incerto X , aqui indicada por $p(X)$, é (por exemplo, ver [20]):

$$p(X) \equiv \lim_{n \rightarrow \infty} \frac{n(X)}{n} \quad (2)$$

onde:

n é o número de repetições, reais ou imaginadas hipoteticamente, da situação ou experimento em que X pode aparecer;

$n(X)$ é o número de vezes em que se observa X .

A teoria frequentista é também conhecida como a teoria tradicional ou ortodoxa da probabilidade. Ela é fortemente dependente do conceito de aleatoriedade (*randomness*), enquanto a teoria bayesiana, por outro lado, é centrada no conceito de incerteza.

Como a noção de inferência é inerente à definição bayesiana de probabilidade, essa teoria transita naturalmente para o campo da inferência estatística, ao passo que, no contexto da abordagem tradicional da probabilidade, a inferência estatística é considerada uma disciplina distinta, ainda que correlacionada com a probabilidade.

Outra diferença é que a abordagem tradicional, por sua definição de probabilidade, leva diretamente à regra da probabilidade conjunta de eventos, podendo-se dizer então que essa regra se sobressai nessa teoria. Na abordagem bayesiana, a regra básica é a da probabilidade condicional, que veremos mais adiante. Operacionalmente, ela permite obter resultados a partir da assunção de um grau de plausibilidade *a priori*. Ao mesmo tempo, essa regra seria a expressão da própria validade da abordagem bayesiana.

A Tabela 2 mostra, de maneira resumida, as principais diferenças entre a interpretação frequentista e a bayesiana. Alguns aspectos pre-

sentes na tabela serão abordados no próximo item.

TABELA 2
INTERPRETAÇÃO FREQUENTISTA VERSUS INTERPRETAÇÃO BAYESIANA DA
PROBABILIDADE

Interpretação:	Frequentista	Bayesiana
Significado de probabilidade	Frequência relativa	Grau de plausibilidade
Natureza da probabilidade	Propriedade da situação sob análise	Estado de conhecimento sobre a situação sob análise
Na relação observador/objeto é focada no ...	...objeto	...observador
Dedução x indução	Se pretende uma teoria puramente dedutivo-matemática	Pode-se dizer que é indutivo-dedutiva, se aproximando da física-matemática
Regra mais “característica”	Probabilidade conjunta	Probabilidade condicional
Conceito importante	Aleatoriedade	Incerteza
Admissão na teoria matemática de probabilidades apriorísticas	Não admite, mas as tem de certa forma embutidas na escolha do modelo de inferência estatística, que é disciplina a parte	Admite
Inferência estatística	É uma disciplina à parte	É conceitualmente parte da teoria
Papel do teorema de Bayes	Regra de aplicação no cálculo de probabilidade condicional	É um método geral de análise
Dependência em relação à disseminação de ferramentas computacionais	Já era dominante antes da disseminação dessas ferramentas	Deve parte de sua atual popularidade à disseminação do microcomputador a partir da década de 1980

A teoria bayesiana não é a única abordagem alternativa à frequentista. O sistema axiomático de Kolmogorov, por exemplo, utiliza teoria dos conjuntos para fundamentar o estudo da probabilidade (muitos manuais de probabilidade, como [21] e [22], usam essa abordagem). Porém, o sistema de Kolmogorov é mais afim com a teoria frequentista do que com a teoria bayesiana.

Para outras teorias da probabilidade, inclusive que se afirmam bayesianas mas diferem do entendimento desse termo aqui delineado, ver [18], apêndice A.

4.2. O TEOREMA DE BAYES

Um dos marcos na retomada do interesse pela teoria bayesiana foi a publicação do livro da referência [17], cuja primeira edição é de 1939.

Entretanto, do ponto de vista dos fundamentos da teoria, [23] tem relevância proeminente por tentar deduzir matematicamente, a partir de princípios elementares, as regras básicas da inferência lógica. Em outras teorias, como a frequentista e a de Kolmogorov, essa dificuldade não existe. Na teoria frequentista, essas regras básicas decorrem facilmente da definição de probabilidade. Na teoria de Kolmogorov, elas são admitidas como axiomas ou são postas como definições (ver [22], p. 20-21 e 27-28).

Em [23], as regras deduzidas são:

- probabilidade condicional:

$$p(X, Y|I) = p(X|Y, I) \cdot p(Y|I) \quad (3)$$

- probabilidade do complementar (ou regra da soma):

$$p(X|I) + p(\bar{X}|I) = 1 \quad (4)$$

onde:

- a vírgula indica a operação “E” lógica;
- a barra vertical “|” significa “dado” ou “supondo”, ou seja, $p(A|B)$ deve ser lido como “probabilidade de A dado B”, ou, equivalentemente, “ $p(A)$ supondo que $p(B)=1$ ”;
- $\bar{X}$ significa “X é falso” ou “X não ocorre”;
- I é a informação ou evento de referência (*background information*), que destaca o fato de que toda probabilidade depende de pressupostos ([19], p. 5).

Nas expressões acima, como I é pressuposto para todos os eventos ou informações, ele pode, por economia na notação, ser omitido, desde que nunca seja esquecido que é condição prévia para todo o restante.

No caso pericial, podem-se citar como exemplos de pressupostos I em um exame de determinação da fonte no nível I (um) da hierarquia de proposições: a integridade da cadeia de custódia da prova, de modo que o vestígio apresentado a exame é aquele coletado na cena do crime; a validade da teoria científica que

subjaz ao exame, por exemplo, os princípios da genética em um exame de DNA.

Dados dois eventos X e Y , a partir da regra da probabilidade condicional temos que:

$$\begin{cases} p(X, Y|I) = p(X|Y, I) \cdot p(Y|I) \\ p(Y, X|I) = p(Y|X, I) \cdot p(X|I) \end{cases}$$

mas:

$$p(X, Y|I) = p(Y, X|I)$$

pois falar no evento “ X e Y ” é o mesmo que falar no evento “ Y e X ”. Das expressões acima obtém-se, então:

$$\begin{aligned} p(X|Y, I) \cdot p(Y|I) &= p(Y|X, I) \cdot p(X|I) \Rightarrow \\ \Rightarrow p(X|Y, I) &= \frac{p(Y|X, I) \cdot p(X|I)}{p(Y|I)} \quad (5) \end{aligned}$$

A expressão (5) é conhecida como teorema de Bayes, e pode, como fica evidente acima, ser obtida e utilizada no contexto da teoria frequentista e de outras teorias da probabilidade, já que, basicamente, é um corolário de (3) (ver, por exemplo, [20], p. 36, e [22], p. 30-31).

O que distingue seu papel na probabilidade bayesiana é que, nela, o teorema de Bayes é a ferramenta de análise principal. Para apreciar sua importância, consideremos, na expressão (5), que X é uma hipótese H cuja probabilidade se deseja calcular e Y é a informação ou evidência E obtida dos dados empíricos. Dessa forma:

$$p(H|E, I) = \frac{p(E|H, I) \cdot p(H|I)}{p(E|I)} \quad (6)$$

A expressão (6) mostra que o teorema de Bayes é uma maneira de calcular a probabilidade da hipótese à luz dos dados empíricos – a chamada probabilidade *a posteriori*, $p(H|E, I)$ –, partindo de duas informações: uma estimativa inicial do valor dessa probabilidade – a probabilidade *a priori*, $p(H|I)$ – e o cálculo da probabilidade de que os dados empíricos seriam realmente aqueles observados caso a hipótese fosse verdadeira – a chamada função de verossimilhança, $p(E|H, I)$.

O termo $p(E|I)$ é uma espécie de “constante de proporcionalidade” que, em muitos proble-

mas práticos, incluindo no uso forense das razões de verossimilhança, não é relevante.

Ao contrário do que a nomenclatura dos termos pode talvez sugerir, não é apenas a função de verossimilhança que pode ser uma “função”. Qualquer dos termos do teorema de Bayes pode ser uma função densidade de probabilidade discreta ou contínua, ou um valor de probabilidade de um evento específico.

Não há em princípio restrições para o valor ou função admitida para a probabilidade *a priori*, mas o valor de convergência da probabilidade *a posteriori*, se houver um, poderá ser alcançado com maior ou menor quantidade de dados dependendo da função *a priori* utilizada.

Esse comentário introduz outra característica da análise bayesiana de dados, que é o fato de demandar mais esforço computacional para determinados problemas do que se utilizada uma abordagem de natureza frequentista. Não é coincidência, portanto, que a popularização do microcomputador a partir da década de 1980 tenha coincidido com o aumento significativo de interesse pela análise bayesiana de dados.

4.3. APLICAÇÃO FORENSE DO TEOREMA DE BAYES: LR

Para equacionar o problema pericial da determinação da fonte, seja H_{so} a hipótese de que vestígio e material padrão tenham vindo da mesma fonte (*so: same origin*), no caso a pessoa ou objeto suspeitos.

De (6), tem-se:

$$p(H_{so}|E, I) = \frac{p(E|H_{so}, I) \cdot p(H_{so}|I)}{p(E|I)} \quad (7)$$

De igual forma, seja H_{do} a hipótese de que a fonte não seja o suspeito (*do: different origin*). De (6):

$$p(H_{do}|E, I) = \frac{p(E|H_{do}, I) \cdot p(H_{do}|I)}{p(E|I)} \quad (8)$$

Pode-se dizer que (7) é uma medida da similaridade entre vestígio e material padrão, enquanto (8) é uma medida da tipicidade do vestígio, pois quanto maior $p(H_{do}|E, I)$, mais típico é o vestígio, pois maior é a probabilidade de

que tenha se originado de outra fonte que não a suspeita.

A razão entre as expressões (7) e (8), portanto, permitirá avaliar conjuntamente esses dois aspectos:

$$\underbrace{\frac{p(H_{so}|E, I)}{p(H_{do}|E, I)}}_{\text{Razão de probabilidade a posteriori}} = \underbrace{\frac{p(E|H_{so}, I)}{p(E|H_{do}, I)}}_{\text{Razão de verossimilhança}} \cdot \underbrace{\frac{p(H_{so}|I)}{p(H_{do}|I)}}_{\text{Razão de probabilidade a priori}} \quad (9)$$

Como mostrado acima, cada um dos termos da expressão (9) possui, a exemplo do que se tem no teorema de Bayes, um nome específico.

A razão à esquerda da igualdade é chamada “razão de probabilidade a posteriori”.

À direita da igualdade, o primeiro termo é a “razão de verossimilhança” ou LR (*Likelihood Ratio*). O segundo é a “razão de probabilidade a priori”.

O termo $p(E|I)$ acaba se cancelando quando passamos de (7) e (8) para (9).

Tendo em vista que a perícia realiza a análise científica do vestígio e não deve, nem pretende, substituir o julgador, a expressão (9) fornece uma base metodológica bastante adequada para os exames de determinação da fonte.

O objetivo dos exames passa a ser o fornecimento da razão de verossimilhança, LR, obtida a partir da análise do vestígio e do material padrão, bem como da amostra de uma população de referência. Dessa forma, a perícia se atém à evidência material.

Um LR maior do que um indica que é maior a probabilidade da evidência considerando que o suspeito é a fonte; um LR menor do que um significa que é maior a probabilidade da evidência supondo que sua fonte seja diversa do suspeito; um valor igual a 1 significa que ambas as probabilidades são iguais.

É muito importante, entretanto, notar que, na abordagem via LR, o exame pericial **não** fornece valores ou razões de probabilidade *a posteriori*. Mas é em geral a razão de probabilidade

de *a posteriori* o que interessa diretamente ao julgador.

Para chegar à razão de probabilidade *a posteriori*, é necessário, como mostra a expressão (9), ter um valor de razão de probabilidade *a priori*, a qual representará o aporte, por parte do julgador, de outras considerações pertinentes ao caso.

5. LR NA COMPARAÇÃO FORENSE DE LOCUTOR

A genética forense foi a primeira área da Criminalística a adotar de forma generalizada uma metodologia baseada em razões de verossimilhança (ver, por exemplo, [10]). Seu papel pioneiro demonstra que a sistemática é passível de implementação.

Na fonética forense, os chamados “sistemas de reconhecimento automático de locutor” (RAL) ou sistemas ASR (*Automatic Speaker Recognition*) que incorporam a abordagem por LR – e não meramente apresentam, como é o caso, em geral, das aplicações de controle de acesso, um resultado do tipo “positivo” ou “negativo” – são uma ferramenta que também demonstra a aplicabilidade da metodologia na comparação forense de locutor.

Entretanto, é importante ter em mente que a abordagem por LR não é uma panaceia, nem pode ser aplicada sem ter em vista suas principais características.

5.1. CARACTERÍSTICAS E LIMITES DA ABORDAGEM POR LR

5.1.1. A EVIDÊNCIA NÃO É O VESTÍGIO

O emprego da análise por LR não altera a natureza do exame de determinação da fonte: ele continua sendo uma identificação qualitativa, não uma individualização.

Valem, portanto, as observações feitas no item 2.2.: o resultado do exame é diretamente dependente das características e da métrica utilizados na comparação.

Na expressão da razão de verossimilhança,

$$LR = \frac{p(E|H_{so}, I)}{p(E|H_{do}, I)} \quad (10)$$

a evidência *E* não é o vestígio, mas o resultado da análise comparativa entre o vestígio e o material padrão, ou seja, a evidência é construída durante a análise, estando nela embutida a escolha de parâmetros a serem extraídos.

No caso do DNA, a técnica atual obtém um perfil genético que é distinto do sequenciamento completo do genoma, baseando-se, na realidade, no tamanho (massa) de segmentos específicos do DNA (para detalhes, ver [24], p. 249-264).

No caso da fala, por outro lado, a distância entre vestígio e evidência é ainda maior do que na genética forense, onde há um acentuado isomorfismo entre ambos. Na comparação forense de locutor, a evidência não é a gravação questionada, mas o resultado da análise comparativa entre essa gravação e a gravação padrão, que dependerá das metodologias – manuais e/ou automáticas; acústicas e/ou articulatórias – empregadas no exame.

Enquanto o vestígio genético é do tipo discreto, sem intravariabilidade – excetuando-se os casos raríssimos de quimerismo e mosaicism – , sem possibilidade de ser disfarçado e praticamente imutável no tempo, o vestígio na comparação forense de locutor tem características muito mais variáveis (ver [6], p. 7-9; [9], p. 303; [25], p. 148-151), pois:

- há intravariabilidade intrínseca na fala, em todos os seus níveis;
- os parâmetros (tanto acústicos quanto articulatórios) são contínuos;
- pode ocorrer tentativa de disfarce;
- a gravação do áudio depende das características do canal e do ambiente, como ruído de fundo e reverberação.

Pode-se dizer, então, que os resultados de uma análise por LR na comparação forense de locutor será intrinsecamente mais dependente da sistemática de obtenção da evidência a partir do vestígio do que no caso do exame de DNA.

5.1.2. “SISTEMAS DE RECONHECIMENTO AUTOMÁTICO DE LOCUTOR”

Os sistemas ASR forenses (ou FSAR, *Forensic ASR*), por exemplo, fazem análise predominantemente de parâmetros acústicos frequenciais e, na prática, costumam usar os coeficientes cepstrais, filtrados e transpostos segundo a escala Mel ([26], p. 262-265; [27], p. 432-433), recebendo a denominação de “coeficientes mel-cepstrais” (MFCC, *Mel Frequency Cepstral Coefficients*).

Outras tecnologias específicas são também empregadas nesses sistemas, sendo muito comum o modelamento estatístico por mistura de gaussianas (GMM, *Gaussian Mixture Model*) com modelo de locutor de referência único (UBM, *Universal Background Model*).

Obviamente, isso não significa que essas escolhas de implementação foram arbitrárias. Testes objetivos e controlados – e, portanto, cientificamente consistentes – foram realizados mostrando, por exemplo, que a abordagem por GMM era mais eficiente do que HMM (*Hidden Markov Model*) para fins forenses (ver, por exemplo, [28]).

Porém, tendo em vista a diferença entre vestígio e evidência, e as condições de contorno do vestígio forense de voz, pode-se afirmar que, embora os atuais sistemas FASR sejam uma ferramenta importante de análise, eles não são definitivos ([29], seção 3.3.2). Por exemplo, ainda há um grande campo de pesquisa aberto a outras abordagens acústicas, automáticas ou semiautomáticas, como a análise de ritmo ([29], seção 3.4.3 e p. 59).

A discussão sobre o chamado “*UK position statement*” ([30]; [25]; [9], p. 305-306; e [31]) mostra que a área de comparação forense de locutor está distante da fase de sedimentação de metodologias. Levantamento recente (ver [32]) sugere que ainda não há uniformidade, em nível mundial, nos métodos de execução desse exame e na forma de expressar seu resultado.

Deve-se ter em mente que a comparação forense de locutor envolve situações específicas (como a possibilidade de disfarce da voz; ambiente com ruídos específicos e reverberação), não levadas em conta em outras aplicações, como o controle de acesso, e que tendem a di-

minuir drasticamente o índice de validade e confiabilidade dos sistemas ASR.

De particular relevância, tanto no alerta que faz, válido para todas as metodologias de realização do exame de comparação forense de locutor, quanto pela experiência dos seus autores, alguns deles com contribuição científica decisiva para o próprio avanço dos sistemas ASR, é o artigo da referência [33].

5.1.3. POPULAÇÃO DE REFERÊNCIA

A abordagem por LR, sendo um método estatístico, depende da amostragem de uma população de referência.

Tem-se dois conjuntos de questões envolvidos, os quais, imbricados com a escolha das características de comparação (item 5.1.1.), também irão influenciar diretamente no resultado da análise: primeiramente, as questões pertinentes à definição da população de referência; e, em segundo lugar, as relacionadas à definição da amostra.

Essas questões são relevantes e ainda não estão totalmente resolvidas não somente nas áreas onde a abordagem estatística é relativamente nova, como na fonética forense, mas também em áreas “tradicionais” como a genética forense.

As questões passíveis de serem levantadas e suas soluções podem variar de caso a caso e refletir-se não somente no cálculo da LR, mas também na informação de referência *I* e nas probabilidades *a priori*. Para um panorama desses problemas, ver [34].

Do ponto de vista prático, para o desenvolvimento da comparação forense de locutor na perícia oficial brasileira, é importante que esse tema seja amplamente e continuamente debatido, gerando procedimentos de construção e utilização de bancos de amostras para o português brasileiro, preferencialmente comuns e compartilhados por todos os órgãos de perícia oficial. É importante lembrar que essa discussão não se aplica somente aos bancos de amostras necessários para os sistemas FASR do tipo MFCC/GMM/UBM, mas para toda a gama de possíveis parâmetros, acústicos e articulatórios,

que venham a ser abordados por meio de razões de verossimilhança.

5.1.4. QUANTITATIVO VERSUS QUALITATIVO

A expressão (9) e mesmo toda a análise bayesiana não necessariamente dependem de dados explicitamente colhidos, em base quantitativa, de um levantamento estatístico explícito. As inferências podem ter por base estimativas do observador, no caso o perito.

A referência [9], por exemplo, admite a possibilidade de atribuição de valores para as probabilidades envolvidas no cálculo da LR na comparação forense de locutor a partir da experiência do perito. Porém, como destaca na p. 306, apenas se essa experiência foi submetida a testes cientificamente controlados de confiabilidade.

Esse requisito de estimativa do desempenho de peritos ou de grupos de perícia, “em áreas onde o método é basicamente o convencimento do examinador” ([10], p. 894), entretanto, está longe de ser alcançado na prática.

Como discute [10], e já foi comentado no item 3.2.2., programas continuados de aferição da proficiência de peritos e grupos de perícia em condições de teste cego, com dados realísticos, são praticamente inexistentes na maioria dos ramos da Criminalística. A fonética forense brasileira não é exceção.

5.1.5. LR E OS NÍVEIS DA HIERARQUIA DE PROPOSIÇÕES

Embora, como ilustram os exemplos dados na Introdução, quando se fala em exames de determinação da fonte basicamente se está referenciando exames no nível I da hierarquia de proposições discutida no item 2.4., o uso de razões de verossimilhança não se restringe a esse nível.

Pelo contrário, como já apontado naquele item, os proponentes dessa hierarquia a definiram com base na possibilidade de realizar a análise pericial em termos de LR em qualquer nível.

As proposições, portanto, não são explicações ou hipóteses quaisquer, mas apenas aquelas passíveis de algum tipo de análise estatística.

ca por LR, onde o perito possa, com seu conhecimento científico, auxiliar a Justiça.

Para ilustrar essa diferença, [7] oferece alguns exemplos. Um deles (p. 4) diz respeito a uma mancha de sangue, encontrada em local de crime, cujo perfil genético é concordante com o do suspeito. Um conjunto de hipóteses aparentemente possível em nível I seria:

- hipótese da acusação (H_1): a mancha de sangue é do suspeito;
- hipótese da defesa (H_2): a mancha de sangue é de outra pessoa que, por acaso, tem o mesmo perfil genético do suspeito.

Apesar de serem explicações possíveis, o par de proposições H_1 e H_2 não é passível de análise pericial. Dadas as características da evidência genética já discutidas em 5.1.2., $p(E|H_1, I)$ é 1, como é em geral nos exames de DNA, onde o numerador da LR é ou 0, ou 1.

Porém, $p(E|H_2, I)$, pelo modo como H_2 foi formulada, também é 1, o que daria

$$LR = \frac{p(E|H_1, I)}{p(E|H_2, I)} = 1 \quad .$$

Para esse exemplo, um par de proposições passível de análise pela perícia seria, na verdade:

- hipótese da acusação (H_{so}): a mancha de sangue é do suspeito;
- hipótese da defesa (H_{do}): a mancha de sangue foi deixada por uma outra pessoa que não o suspeito.

A medida que se avança para os níveis II e III da hierarquia, com o aporte de informação contextual, fica mais crítica não somente a obtenção e análise de dados estatísticos, mas também a percepção pelo perito dos pares de proposições adequados para a análise por LR.

Entretanto, é possível, em teoria, como mostrado em [7], essa análise, de maneira que a perícia possa trazer informação estatística de base científica a fim de auxiliar o julgador, seja na constituição da autoria, seja na constituição da materialidade delitiva.

Na análise forense de áudio e vídeo, os exames de verificação de edição são uma das áreas que podem se beneficiar do aprofundamento dos estudos da aplicação de LR em níveis acima do nível I da hierarquia de proposições.

5.2. A METODOLOGIA E A CONCLUSÃO DO LAUDO DE COMPARAÇÃO FORENSE DE LOCUTOR

Embora desejável do ponto de vista científico, a introdução da análise por LR em todos os aspectos hoje utilizados nos exames de comparação forense de locutor é bastante problemática, como alerta [31].

O cenário atual, aparentemente (ver [32]), ainda é de predomínio, em nível mundial, da análise do tipo impressionística baseada na comparação de características acústicas e articulatórias.

Como discutido nos itens 5.1.2. e 5.1.3., entretanto, um cenário na verdade muito pior, em termos da qualidade dos resultados fornecidos pela perícia de comparação forense de locutor, seria aquele em que sistemas FASR do tipo MFCC/GMM/UBM fossem usados com exclusividade sem a compreensão das suas limitações e dos cuidados de implementação.

Admitindo, portanto, que a melhor solução no momento não seja abandonar as metodologias tradicionais, mas buscar agregar a elas, paulatinamente, cada vez mais elementos baseados na abordagem por LR, são apresentadas a seguir algumas considerações sobre as conclusões dos exames.

5.2.1. CONCLUSÕES CATEGÓRICAS ATRIBUTIVAS DE IDENTIDADE

Independente do uso ou não da abordagem bayesiana com razões de verossimilhança, a análise lógica apresentada no item 2. e as consideração feita no item 3. mostram que, mesmo quando peritos emitem um laudo com conclusão categórica atribuindo a identidade entre a fonte do vestígio (material questionado) e a fonte do material padrão, com base exclusivamente na análise desses materiais, essa conclusão deve ser interpretada como uma afirmação do grau (máximo) de convicção dos peritos, não como uma identificação numérica (individualização).

O mais adequado, portanto, seria não utilizar conclusões categóricas de atribuição de identidade, a fim de não dar ao exame a aparência de um peso científico que ele não possui.

5.2.2. CONCLUSÕES CATEGÓRICAS DE EXCLUSÃO DE IDENTIDADE

Outra questão diz respeito à possibilidade ou não de conclusões categóricas de exclusão (negativas de identidade), quando a análise aponta flagrantes incompatibilidades entre o vestígio e o material padrão.

De acordo com [25], item 3.3.1, que se atém ao nível I da hierarquia de proposições, esse tipo de conclusão é possível no caso de análise de características discretas, como é o caso do DNA.

De fato, se admitidos como condição de contorno *I* na expressão (9) os princípios da genética, a não ocorrência de erros no procedimento de exame (por exemplo, contaminação) e a não ocorrência de quimerismo ou mosaicismos, então, se não houve correspondência entre o perfil genético do vestígio e do suspeito,

$$p(E|H_{so}, I) = 0 \quad \text{e} \quad LR = 0.$$

Independente da razão de probabilidade *a priori*, a razão de probabilidade *a posteriori* também será nula, e é possível uma afirmação categórica, em nível I da hierarquia de proposições.

Entretanto, a natureza distinta da fala humana, conforme comentado em 5.1.2., não permitiria esse tipo de raciocínio na comparação de locutor, o que impediria o uso de conclusões categóricas de exclusão.

[31], por outro lado, reafirma a defesa da possibilidade desse tipo de conclusão categórica, lembrando, além de outras considerações, que há um limitante lógico, inescapável, na análise de gravações de voz no contexto dos exames de comparação forense de locutor, inclusive se todo o exame for realizado por meio de razões de verossimilhança, de modo manual, semiautomático ou automático (mesmo se empregados sistemas de diarização; ver, por exemplo, [35]).

Esse limitante são as decisões atributivas/não atributivas de identidade por meio de avaliação localizada, seja de trechos menores ou maiores da gravação, as quais fundamentam a própria delimitação do que constitui a fala questionada.

A necessidade dessa seleção fica evidente quando a gravação questionada “claramente” – por uma avaliação lógico-indutiva, não exclusivamente lógico-dedutiva – contém falas de mais de um locutor, mas essa atividade de seleção existe mesmo quando há, “claramente”, apenas falas de um locutor.

Pode-se dizer que a admissão de que essa seleção é correta faz parte da hipótese *I* da expressão (9).

Esse tema é pertinente, também, às questões da continuidade no tempo e da hierarquia de proposições, já que o limite entre o que é aceitável de intravariabilidade de um locutor, bem como o limite de distinção, via intervariabilidade, entre locutores distintos, quando se analisa um determinado conjunto de gravações questionadas, mesmo que a seleção seja feita por sistemas automáticos ou semiautomáticos de diarização, calibrados a partir de experimentos onde a origem das falas é garantida por meio da garantia de continuidade no tempo, sempre será, no final das contas, um limite estatístico.

Esse limite é, sob outro ponto de vista, o limite inerente à lógica indutiva que fundamenta o método científico.

Para [31], o peso lógico-indutivo de uma afirmação que exclui a identidade é o mesmo das avaliações localizadas que embasam todo o exame de comparação forense de locutor. Por esse motivo, esse artigo admite conclusões categóricas de exclusão.

Na verdade, essa admissão, implicitamente, considera suficiente o julgamento impressionístico do perito.

A análise da questão em [25], entretanto, não levantou esse ponto especificamente, embora seja forçoso admitir que, para atingir higidez lógica na sua defesa do uso incondicional da abordagem por LR, o referido artigo [25] deve-

ria defender também que a capacidade de realização dessas atribuições/não atribuições localizadas se subsumisse à sistemática de testes de calibração/proficiência segundo o método científico, preferivelmente sendo realizada por métodos automáticos de diarização.

De qualquer forma, na prática, pelo menos no estado atual das metodologias, talvez seja mais prudente seguir, por exemplo, o que apresenta [29], seção 3.8.1, e também não utilizar conclusões categóricas de exclusão.

5.2.3. CONCLUSÕES EM SITUAÇÕES DE CONJUNTO FECHADO

Outra questão que diz respeito ao uso ou não de conclusões categóricas refere-se aos casos em que, ao invés de um conjunto aberto de possíveis fontes para o vestígio, tem-se um conjunto fechado de suspeitos.

[30], item 5, fala, por exemplo, de situações em que se tem, por gravações de vídeo, a imagem dos falantes presentes.

Essa questão extrapola o nível I da hierarquia de proposições, pois inclui questões relativas à continuidade no tempo.

Pode-se dizer que, se as filmagens incluem o áudio ambiente e as imagens mostram claramente os movimentos de articulação da fala em cada locutor mostrado, o exame de comparação de locutor se torna desnecessário, podendo o problema ser abordado do ponto de vista da verificação de edição e do reconhecimento facial.

Se, de qualquer modo, nessa ou em outras situações, um exame de comparação de locutor em situação de conjunto fechado é solicitado, a metodologia básica a ser seguida não deveria diferir daquela empregada nos exames de conjunto aberto (ver [25], item 3.3.2), talvez apenas se considerando um menor número de passos de análise ou espaço amostral, para se chegar a uma conclusão, a qual seria, ainda, não categórica.

5.2.4. O FORNECIMENTO DE VALORES DE LR NA CONCLUSÃO DO LAUDO E AS ESCALAS DE PROBABILIDADE

Em princípio, se fosse – ou for algum dia – possível, nos exames de comparação forense de locutor, fornecer um valor de LR englobando todas as características relevantes para a

análise, conforme lista não exaustiva levantada em [31], nada impediria que esse número constituísse, por si só, a conclusão do laudo.

De fato, essa é a sistemática adotada nos exames de DNA.

Porém, na prática, essa questão não é tão simples. A própria experiência na área de genética forense mostra que a percepção do real significado de uma LR no contexto da análise bayesiana de dados é em si um desafio.

Uma das maiores dificuldades é separar o que é a razão de probabilidade *a posteriori* dos demais termos à direita da igualdade na expressão (9). Esses conceitos – razões de probabilidade *a priori* e *a posteriori* e razão de verossimilhança –, embora sejam matematicamente simples, não são intuitivos.

A situação se complica ainda mais porque há uma dificuldade em compreender os próprios limites lógico-científicos da atuação da perícia. Como se viu na discussão do item 3., essa observação se aplica até mesmo aos próprios peritos.

Por esse motivo, a experiência com o DNA mostra que contar com um razoável grau de entendimento dos termos da expressão (9) por parte dos demais atores da cadeia de persecução penal pode levar a profundas frustrações (ver, por exemplo, [36]).

Além disso, um valor numérico de LR pode ser de pouco esclarecimento e utilidade para o julgador, como apontado por [29], p. 61-62.

E, não raro, o valor de LR é assumido, para os fins práticos, por alguns ou mesmo todos os demais atores da cadeia de persecução penal, como um valor de razão de probabilidade *a posteriori*, absolutizando o resultado do exame de determinação da fonte, ainda que ele seja, como vimos, uma identificação qualitativa de nível I.

Quando a LR é muito maior do que 1, a acusação pode lançar mão do seu valor e, mediante a chamada “falácia da acusação”, tomá-la como razão de probabilidade *a posteriori* e afirmar que a probabilidade do suspeito ser

culpado é igualmente enorme (ou seja, para os fins práticos, o suspeito é culpado).

Da mesma forma, a defesa, lançando mão de um raciocínio conhecido como “falácia da defesa”, pode chegar a um valor de razão de probabilidade *a posteriori* muito baixo admitindo uma razão de probabilidade *a priori* não realística – por exemplo, igual ao inverso de toda a população de um país. A partir da razão de probabilidade *a posteriori* obtida, a defesa argumenta que, na verdade, é muito maior a probabilidade do suspeito ser inocente do que culpado, e portanto, havendo dúvida de tal monta, ele deve ser absolvido.

Em ambas as falácias, a falha é não se levar em conta as outras provas e circunstâncias pertinentes ao caso.

No caso da comparação forense de locutor, onde não há, no horizonte de médio prazo – e talvez no de longo prazo –, perspectiva de chegar sequer a uma situação de uso integral da abordagem por LR, talvez seja mais interessante, no atual momento, englobar os diversos tipos de análises (impresscionística/semiautomática/automática) em um esquema de conclusão por escalas verbais de probabilidade.

A referência [29], p. 60, traz um exemplo de escala, a qual, frise-se, não faz, mesmo nos graus máximo (atribuição) e mínimo (exclusão), conclusões categóricas.

Deve-se estar ciente que qualquer tentativa de uso de uma escala qualitativa – ainda que a partir da correspondência objetiva entre, por exemplo, faixas de valores numéricos de LR e expressões verbais qualitativas – sempre estará sujeita a críticas, como a ocorrência do efeito do valor de limiar (*cliff-edge effect*). Para detalhes, ver [25], p. 151-155 e [6], p. 133-136.

5.2.5. O EMPREGO DO TERMO “COMPATÍVEL” E OUTROS SEMELHANTES

Uma questão relacionada ao tema da conclusão do laudo, não somente na comparação forense de locutor, mas em outros tipos de exame de determinação da fonte, é o uso da palavra “compatível” e semelhantes.

Dependendo do tipo de vestígio (ver 5.2.1.), a Criminalística pode fornecer conclusões categóricas de exclusão. A questão se complica, entretanto, quando não há possibilidade de afirmar a exclusão ([6], p. 95).

No caso do exame de determinação da fonte, quando há algum grau de convergência entre o vestígio e o material padrão, não é incomum o emprego da expressão “há compatibilidade entre” (em inglês, é comum o uso do termo “*is consistent with*”; ver [11], p. 216) para expressar essa convergência.

Infelizmente, esse termo, por sua amplitude, não fornece nenhuma pista a respeito do grau de “compatibilidade” ([6], p. 96), nem tampouco fornece subsídios para uma conclusão sobre probabilidades *a posteriori*, como alerta [25], seção 3.4.3.

A luz dessas ponderações, não seria indicado o uso de conclusões que gravitam em torno do termo “compatível” e semelhantes, a menos que seja dado também o grau dessa compatibilidade.

O mesmo tipo de observação é feito por [6], p. 96, a respeito de expressões como “a evidência não exclui determinada hipótese”.

6. NOMENCLATURA DO EXAME

Neste tutorial, adotamos para o exame forense de determinação da fonte em que a fala é o vestígio de interesse a denominação “comparação forense de locutor”, ou simplesmente “comparação de locutor”.

Na perícia criminal oficial brasileira, a denominação atualmente mais encontrada para o exame é “verificação de locutor”. Esse nome é o utilizado na primeira publicação de fôlego a respeito desse exame no Brasil ([37], de 1999; a segunda edição, [38], de 2003, mantém a mesma nomenclatura) e foi adotado, por exemplo, no Departamento de Polícia Federal e, por meio de programas de capacitação, difundiu-se para outros órgãos de perícia criminal oficial.

Deve-se observar que, apesar da operação lógica de verificação 1-para-1 presente no exame de verificação de locutor encontrado na

área de controle de acesso ser a mesma do exame forense, como nota [39], p. 9, as nomenclaturas predominantes em inglês, internacionalmente, empregaram termos como “*recognition*” e “*identification*”, e não “*verification*”, por entender que na área forense havia condições de contorno distintas, como falta de controle da qualidade do material questionado e possibilidade de não cooperação por parte dos fornecedores das amostras de fala.

Para detalhes, ver, além de [39], p. 5-10, a discussão mais aprofundada encontrada em [26], capítulo 5. Em linhas gerais, a mesma classificação taxonômica dos exames da área de processamento da fala, mas com viés acústico, focada em ferramentas computacionais automatizadas, pode ser encontrada em [40].

O fato da fala ser efetivamente uma biometria – ou seja, ela tem valor discriminador – não necessariamente significa que ela seja individualizadora. Como se afirma na referência [4], p. 208, apenas a verificação empírica pode dar maior ou menor “grau de verossimilitude” a afirmações de que determinada biometria atende o critério de unidade.

Naquela referência, por exemplo, se admite que a afirmação da unicidade da decadactilar, colhida em condições ideais, atinge um grau de verossimilitude suficiente para que seja admitida na prática.

Outra questão, diferente, é a da ocorrência ou não de erros na coleta da decadactilar, o que é passível de acontecer, já que estamos falando de uma atividade humana, ou mesmo do desempenho de determinado algoritmo do tipo AFIS (*Automatic Fingerprint Identification System*) no que concerne ao armazenamento ou recuperação das informações baseadas nas imagens da decadactilar.

Outra questão, ainda, é a da transferência de informação de uma digital para um fragmento de digital em um local de crime. Um caso famoso de erro cometido pelo FBI, que havia afirmado ter uma identificação com “100% de probabilidade” em caso de grande repercussão (ver [41] para uma análise oficial sobre o

caso), é um dos exemplos que se relacionam aos problemas discutidos no item 3..

Já para a fala, a própria referência [4] entende não haver, atualmente (no caso, o ano de 2006, mas essa afirmação continua válida), suporte científico que dê a afirmações de unicidade grau de verossimilitude suficiente.

Esse fato já era comentado, por exemplo, em [39], ao tratar da multidimensionalidade da fala e dos limites de intravariabilidade e intervariabilidade.

Essas ponderações, em face também do que foi discutido nos itens 3. e 4., têm levado vários pesquisadores da área de fonética forense a preferir nomenclaturas que se afastem da ideia de individualização.

Ao invés de “reconhecimento”, “identificação”, “verificação” ou “discriminação”, o termo julgado mais adequado tem sido o de “comparação” ([9], p. 300-301; [25], item 3.2; [30]; [29]).

O presente tutorial emprega, além do termo “comparação”, a expressão “comparação de locutor”, o invés, por exemplo, de “comparação de voz” (como preferem, por exemplo, [9] e [25]) ou “comparação de fala” (como é usado em [29]).

Essa escolha foi motivada pela experiência de trabalho do autor e pelas argumentações apresentadas em [29], que distingue voz de fala. Essencialmente, o mesmo argumento de [29] aparece em [31], p. 146, mas, ao lembrar que também características extralinguísticas são relevantes no exame, o termo “locutor” parece mais adequado do que “fala”.

Deve-se notar que as denominações usadas tradicionalmente já estão amplamente difundidas, não sendo possível afirmar, no atual momento, se, na literatura técnica da área, elas irão perdurar ou se serão substituídas.

7. CONCLUSÃO

O dogma da unicidade, subjacente às metodologias tradicionalmente empregadas nos exames de diversas áreas da Criminalística dedicadas ao problema da determinação da fonte, foi

alvo de críticas contundentes principalmente a partir do final da última década do século passado.

Essas críticas motivaram uma mudança de entendimento sobre o papel e o alcance das perícias que envolvem a identificação de uma pessoa ou objeto.

Neste tutorial foi feito um apanhado dos principais conceitos envolvidos nessa mudança, com foco nos impactos sobre a área de comparação forense de locutor, mostrando também as dificuldades para que se tenha a implementação integral de uma abordagem baseada em razões de verossimilhança nessa área.

REFERÊNCIAS

- [1] MENDES, Lamartine Bizarro. *Documentoscopia*. 2. ed. Campinas, SP: Millennium, 2003. ISBN 85-86833-93-2.
- [2] TOCCHETTO, Domingos. *Balística forense: aspectos técnicos e jurídicos*. 4. ed. Campinas, SP: Millennium, 2005. ISBN 85-7625-065-9.
- [3] TOCCHETTO, Domingos; ESPINDULA, Alberi (Coord.). *Criminalística: procedimentos e metodologias*. 2. ed. Porto Alegre, RS: [s.n.], 2009. ISBN 978-85-99541-01-2.
- [4] MEUWLY, Didier. Forensic individualization from biometric data. *Science & Justice*, v. 46, n. 4, 2006, p. 205-213.
- [5] COOK, R.; EVETT, I. W.; JACKSON, G.; JONES, P. J.; LAMBERT, J. A.. A hierarchy of propositions: deciding which level to address in casework. *Science & Justice*, v. 34, n. 4, 1998, p. 231-239.
- [6] LUCY, David. *Introductions to statistics for forensic scientists*. West Sussex, Inglaterra: John Wiley & Sons, Ltd, 2005. ISBN 978-0470022016.
- [7] EVETT, I. W.; JACKSON, G.; LAMBERT, J. A.. More on the hierarchy of propositions: exploring the distinction between explanations and propositions. *Science & Justice*, v. 40, n. 1, 2000, p. 3-10.
- [8] BRASIL. *Código de Processo Penal, decreto-lei nº 3.689 de 3 de outubro de 1941*. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto-lei/del3689.htm>. Acesso em: 01.ago.2012.
- [9] MORRISON, Geoffrey Stewart. Forensic voice comparison and the paradigm shift. , v. 49, 2009, p. 298-308.
- [10] SAKS, Michael J., KOEHLER, Jonathan J.. The coming paradigm shift in forensic science. *Science*, v. 309, 2005, p. 892-895.
- [11] SAKS, Michael J., KOEHLER, Jonathan J.. The individualization fallacy in forensic science evidence. *Vanderbilt Law Review*, v. 61, n. 1, 2008, p. 199-219.
- [12] MLODINOW, Leonard. *The drunkard's walk: how randomness rules our lives*. 1. ed. USA: Vintage, 2009. ISBN 978-0307275172.
- [13] STONEY, David A.. What made us ever think we could individualize using statistics?. *Journal of the Forensic Science Society*, v. 31, n. 2, 1991, p. 197-199.
- [14] MALLMITH, Décio de Moura. A respeito do número de pontos característicos necessários à afirmativa de identidade. *Revista do IGP/RS*, n. 4, jan.2011, p. 27-29. Disponível em: <<http://www.igp.rs.gov.br/revistas/revista4.pdf>>. Acesso em: 01.ago.2012
- [15] KUHN, Thomas S.. The function of dogma in scientific research. In: CROMBIE, A. C. (Ed.). *Scientific change*. New York, USA: Basic Books, 1963, p. 347-369.
- [16] KUHN, T. S.. A função do dogma na investigação científica. In: DEUS, Jorge Dias de (Org.). *A crítica da ciência*. Rio de Janeiro, RJ: Zahar, 1979, p. 53-80.
- [17] JEFFREYS, Harold. *Theory of probability*. 3. ed. Oxford, UK: Oxford University Press, 1998. ISBN: 978-0198503682.
- [18] JAYNES, Edwin Thompson. *Probability theory: the logic of science*. 1. ed. Cambridge, UK: Cambridge University Press, 2003. ISBN: 978-0521592710.
- [19] SIVIA, D. S.; SKILLING, J.. *Data analysis: a bayesian tutorial*. 2. ed. Oxford, UK: Oxford University Press, 2006. ISBN: 978-0198568322.
- [20] ROZANOV, Y. A.. *Probability theory: a concise course*. New York, USA: Dover Publications, Inc., 1977. ISBN: 978-0486635446.
- [21] FELLER, William. *An introduction to probability theory and its applications*. 3. ed. USA: Wiley & Sons, Inc., 1968. ISBN: 978-0471257080.
- [22] PAPOULIS, Athanasios. *Probability, random variables, and stochastic processes*. 3. ed. New York, USA: McGraw-Hill, 1991. ISBN: 0-07-048477-5.
- [23] COX, Richard T.. Probability, frequency and reasonable expectation. *American Journal of Physics*, v. 14, n. 1, 1946, p. 1-13.
- [24] FIGINI, Adriano R. L. et al.. *Identificação humana*. 2. ed. Campinas, SP: Millennium, 2003. ISBN: 978-85-7625-002-9.
- [25] ROSE, Philip; MORRISON, Geoffrey S.. A response to the UK Position Statement on forensic speaker comparison. *The International Journal of Speech, Language and the Law*, v. 16, n. 1, 2009, p. 139-163.
- [26] ROSE, Philip. *Forensic speaker identification*. 1. ed. London, UK: Taylor & Francis, 2002. ISBN: 0-415-27182-7.
- [27] BIMBOT, Frédéric et al.. A tutorial on text-independent speaker verification. *EURASIP Journal fon Applied Signal Processing*, n. 4, 2004, p. 430-451.
- [28] REYNOLDS, D. A.; QUATIERI, T. F.; DUNN, R. B.. Speaker verification using adapted Gaussian Mixture Models. *Digital Signal Processing*, v. 10, 2000, p. 19-41.
- [29] ERIKSSON, Anders. Aural/acoustic vs. automatic methods in forensic phonetic case work. In: NEUSTEIN, A.; PATIL, H. A. (Org.). *Forensic speaker recognition: a law enforcement and counter-terrorism*. New York, USA: Springer-Verlag, 2011, p. 41-69. ISBN: 978-1-4614-0262-6
- [30] FRENCH, J. Peter; HARRISON, Philip. Position statement concerning use of impressionistic likelihood terms in forensic speaker comparison cases. *The International Journal of Speech, Language and the Law*, v. 14, n. 1, 2007, p. 137-144.
- [31] FRENCH, J. Peter et al.. The UK position statement on forensic speaker comparison: a rejoinder to Rose and Morrison. *The International Journal of Speech, Language and the Law*, v. 17, n. 1, 2010, p. 143-152.
- [32] GOLD, Erica; FRENCH, Peter. International practices in forensic speaker comparison. *The International Journal of Speech, Language and the Law*, v. 18, n. 2, 2011, p. 293-307.
- [33] BONASTRE, Jean-François et al.. Person authentication by voice: a need for caution. *Proceedings of Eurospeech 2003, Genebra, Suíça*, , 2003, p. 33-36.
- [34] AITKEN, Colin G. G.. Populations and samples. In: AITKEN, C. G. G.; STONEY, D. A.. *The use of statistics in forensic science*. : Taylor & Francis e-Library, 2002, 1991, p. 51-82. ISBN: 0-13-933748-2 (Print Edition)
- [35] BONASTRE, Jean-François et al.. ALIZE/SpkDet: a state-of-the-art open source software for speaker recognition. *Odyssey 2008: The Speaker and Language Recognition Workshop, Stellenbosch, África do Sul*, , jan.2008, p. .
- [36] DONNELLY, Peter. Appealing statistics. *Significance*, v. 2, n. 1, 2005, p. 46-48.
- [37] BRAID, Antonio César Morant. *Fonética forense*. 1. ed. Porto Alegre, RS: Sagra Luzzatto, 1999. ISBN 85-241-0606-9.
- [38] BRAID, Antonio César Morant. *Fonética forense*. 2. ed. Campinas, SP: Millennium, 2003. ISBN 85-86833-91-6.
- [39] FRANCIS, Nolan. *The phonetic bases of speaker recognition*. 1. ed. : Cambridge University Press, 1983. ISBN: 978-0521108270.
- [40] CAMPBELL, Jr., Joseph P.. Speaker recognition: a tutorial. *Proceedings of the IEEE*, v. 85, n. 9, 1997, p. 1437-1462.

- [41] USA Department of Justice. *A review of the FBI's handling of the Brandon Mayfield case*. Disponível em: <http://www.justice.gov/oig/special/s0601/PDF_list.htm>. Acesso em: 01.ago.2012.

Artigos

Aplicação de Engenharia Reversa em Exames de Verificação de Edições de Vídeos Digitais

Bruno Gomes de Andrade

Instituto Nacional de Criminalística, Polícia Federal, Brasil
SPO – Lote 7 – Setores Complementares – 70.610-902, Brasília/DF - Brasil
andrade.bga@dpf.gov.br

RESUMO – Nos exames de verificação de edições, um dos grandes desafios encarados pelos Peritos Criminais é a busca por elementos materiais que garantam, de forma objetiva, a continuidade do registro sob análise. Em se tratando de vídeo digital, os dispositivos modernos que efetuam a captura do registro original geralmente adicionam nos cabeçalhos dos arquivos dados de controle como abertura do diafragma, distância focal, data/hora etc., que geralmente não são disponibilizados de forma direta ao usuário. Neste artigo será discutido estudo de caso em que foi possível, através de engenharia reversa, a exposição de parte da formatação do cabeçalho do arquivo digital e a consequente extração de elementos de data/hora codificados que se mostraram relevantes indícios de ausência de edição no vídeo.

1. INTRODUÇÃO

O exame de verificação de edições visa a procurar elementos materiais indicativos de edição ou montagem que alterem o conteúdo original da gravação de modo a apresentar, de forma propositalmente de difícil ou impossível percepção ao observador comum, fatos ou eventos de modo distinto daquele em que ocorreram na realidade. As edições podem consistir em inserção, supressão, superposição ou substituição de registros de áudio/vídeo.

As técnicas empregadas no exame variam conforme a natureza do material questionado. Via de regra, sempre são realizadas análises perceptivas e de coerência contextual sobre o conteúdo da gravação. Além disso, vários tipos de verificações podem ser realizadas enfocando aspectos como: o suporte físico da gravação; os sinais de áudio e vídeo nela presentes; a coerência entre a gravação questionada e o suposto equipamento utilizado para sua obtenção; e também a estrutura dos arquivos no caso de gravações digitais.

Neste artigo será discutido estudo de caso em que a análise detalhada da estrutura de arquivos revelou dados codificados no cabeçalho

dos arquivos de vídeo digitais questionados que se tornaram importantes indícios de ausência de edição do material questionado.

2. MATERIAL QUESTIONADO

Os arquivos objeto da análise continham registros de áudio e vídeo com extensão *mpg*. Todos esses vídeos estavam codificados com *codec* MPEG-2 (MPEG2) e possuíam imagens com resolução de 720 *pixels* de largura por 480 *pixels* de altura, entrelaçada, em cores, padrão NTSC, com sistema 4:2:0 de sub-amostragem de cor. A taxa de quadros por segundo era, em todos os vídeos, de 29,970 quadros por segundo. O áudio utilizava *codec* Audio Coding 3 (AC-3), com taxa de amostragem de 48kHz, resolução de 16 bits por amostra, estéreo e taxa de *bits* total constante de 256Kbps.

3. EXAMES

3.1. ANÁLISE DO FLUXO DE DADOS

A análise com editor hexadecimal e analisador de *streams* (fluxo de dados) MPEG dos arquivos questionados mostrou que eles obedecem à estrutura de dados do padrão MPEG-2 [1].

O GOP (*Group of Pictures*, Grupo de Imagens), que se repete sucessivamente no *stream* de vídeo, corresponde ao intervalo de tempo de aproximadamente 0,5 s, e é constituído de 15 quadros ou imagens, com a estrutura “IBBPBBPBBPBBPBB”, sendo que “I” representa os *I-frame* ou *key frame* (*Intra-coded picture*), “P” os *P-frame* (*Predicted picture*) e “B” os *B-frame* (*Bi-predictive picture*).

3.2. PRIVATE STREAMS

Em todos os vídeos examinados verificou-se a presença, em vários trechos dos arquivos, de *streams* MPEG (PES, *Packetized Elementary Stream*) do tipo *private* (privado), com cabeçalho trazendo o identificador (*stream ID*) 0xBD (o prefixo “0x” indica que o valor que o sucede está representado em formato hexadecimal).

Streams do tipo *private* são uma maneira de introduzir nas transmissões e arquivos MPEG-2 metadados e mesmo informações de áudio e vídeo não especificados no padrão MPEG-2. É exatamente nesses *streams* em que os fabricantes de equipamentos audiovisual podem introduzir todo tipo de informação adicional que desejarem, como marca e modelo do equipamento gravador, data/hora do registro de vídeo, informações técnicas como abertura do diafragma e distância focal, dentre outras.

3.3. EQUIPAMENTO DE GRAVAÇÃO

Em cada arquivo questionado, alguns dos *private streams* traziam sequências de bytes com, dentre outros, os caracteres em ASCII “SONY” e “DCR-SR68”, sugerindo que essas eram, respectivamente, a marca e modelo da filmadora utilizada. A figura 1 traz, a título de exemplo, o trecho inicial de um dos arquivos questionados, analisado em editor de arquivos hexadecimal, destacando essa sequência ASCII.

Na figura 1, a parte central mostra os *bytes* do arquivo em hexadecimal (20 *bytes* por linha, agrupados de 4 em 4 *bytes*); a parte da esquerda mostra o endereço do *byte* inicial da linha em relação ao início do arquivo (*displace-*



Figura 2: Câmera de teste Sony DCR-SR68



Figura 3: Detalhe da Câmera Sony DCR-SR68

ment), também em hexadecimal; e a parte da direita mostra caracteres ASCII correspondentes a cada *byte* (quando houver um caractere correspondente).

Conforme manual do fabricante [3], a filmadora DCR-SR68 utiliza o sistema de gravação NTSC (colorido), formato de vídeo MPEG-PS e áudio DOLBY DIGITAL (outro nome para o formato AC3). Essas características se mostraram compatíveis com as presentes nos arquivos analisados.

00000000	00 00 01 BA	44 00 04 00	04 01 01 89	C3 F8 00 00	01 BB 00 12	00 00 00 00
00000014	80 C4 E1 04	E1 FF B9 E0	E8 B8 C0 20	BD E0 3A BF	E0 02 00 00	00 00 00 00
00000028	01 BD 07 D4	81 00 00 FF	4D 56 49 5F	44 41 54 41	01 00 FF FF	00 00 00 00
0000003C	53 4F 4E 59	00 44 43 52	2D 53 52 36	38 00 00 00	00 03 00 00	00 00 00 00
00000050	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00
00000064	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00
00000078	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00
0000008C	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00
000000A0	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00
000000B4	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00	00 00 00 00
000000C8	00 00 00 00	00 00 FF FF	01 FF FF FF	FF FF FF FF	FF FF FF FF	FF FF FF FF
000000DC	FF FF FF FF	FF FF FF FF	FF FF FF FF	FF FF FF FF	FF FF FF FF	FF FF FF FF
000000F0	FF FF FF FF	FF FF FF FF	FF FF FF FF	FF FF FF FF	FF FF FF FF	FF FF FF FF
00000104	FF FF FF FF	FF FF FF FF	FF FF FF FF	00 00 02 10	00 00 01 79	00 00 00 00
00000118	00 00 00 DC	00 00 00 00	FF FF FF FF	FF FF FF FF	FF FF FF FF	FF FF FF FF

00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F 10 11 12 13 14 15 16 17 18 19 1A 1B 1C 1D 1E 1F 20 21 22 23 24 25 26 27 28 29 2A 2B 2C 2D 2E 2F 30 31 32 33 34 35 36 37 38 39 3A 3B 3C 3D 3E 3F 40 41 42 43 44 45 46 47 48 49 4A 4B 4C 4D 4E 4F 50 51 52 53 54 55 56 57 58 59 5A 5B 5C 5D 5E 5F 60 61 62 63 64 65 66 67 68 69 6A 6B 6C 6D 6E 6F 70 71 72 73 74 75 76 77 78 79 7A 7B 7C 7D 7E 7F 80 81 82 83 84 85 86 87 88 89 8A 8B 8C 8D 8E 8F 90 91 92 93 94 95 96 97 98 99 A0 A1 A2 A3 A4 A5 A6 A7 A8 A9 AA AB AC AD AE AF B0 B1 B2 B3 B4 B5 B6 B7 B8 B9 BA BB BC BD BE BF C0 C1 C2 C3 C4 C5 C6 C7 C8 C9 CA CB CC CD CE CF D0 D1 D2 D3 D4 D5 D6 D7 D8 D9 DA DB DC DD DE DF E0 E1 E2 E3 E4 E5 E6 E7 E8 E9 EA EB EC ED EE EF F0 F1 F2 F3 F4 F5 F6 F7 F8 F9 FA FB FC FD FE FF

SONY DCR-SR68

Figura 1: Sequência ASCII no início do arquivo questionado sugestiva da marca e modelo da filmadora utilizada (destacada em vermelho). Em verde, encontra-se destacada a sequência de três *bytes* que indica início de PES, 0x000001, seguida do *stream ID* 0xBD (*private stream*).

Para fins de comparação, foram realizadas, durante os exames, gravações de teste utilizando uma câmera dessa marca e modelo. Essa câmera de teste é mostrada nas figuras 2 e 3.

Nas gravações de teste, verificou-se que a filmadora Sony DCR-SR68 registra continuamente no arquivo da filmagem a data e hora em que ela se realiza, tendo como referência a data/hora inicialmente ajustada na câmera.

Essa informação podia ser exibida no *display* da filmadora, na forma de caracteres sobrepostos em uma lateral das imagens, quando o arquivo era nela reproduzido. Porém, essa informação não era visível quando o mesmo arquivo era reproduzido em programas do tipo *media player* convencionais.

A análise das gravações de teste mostrou que essa informação de data/hora somente poderia estar registrada em posição e formato específicos, dentro de um pacote PES de *stream* do tipo *private*, localizado antes de cada *I-frame*. Entretanto esses dados se mostraram codificados de forma que a análise direta dos *streams* em programa editor hexadecimal não foi suficiente para revelar a informação de data/hora.

3.4. VERIFICAÇÃO E EXTRAÇÃO DOS RÓTULOS DE DATA/HORA DOS ARQUIVOS DE VÍDEO DIGITAIS

3.4.1. ENGENHARIA REVERSA DO PROGRAMA INFOMPG.EXE

Com a informação de que os dados de data/hora estariam codificados dentro dos *private streams* dos arquivos sob análise, iniciou-se a busca por informações sobre a formatação das informações técnicas presentes no cabeçalho dos arquivos questionados.

A busca por documentação sobre o conteúdo dos *private streams* da câmera em questão se mostrou infrutífera. Entretanto, um programa intitulado *infompg.exe* encontrado na internet prometia a obtenção de diversas informações técnicas de arquivos MPEG-2 registradas pelo equipamento de gravação. A idoneidade do arquivo não pode ser comprovada, sendo certo que não se tratava de programa comercial proveniente do fabricante.

Os dados extraídos pela utilização do programa se mostraram coerentes, entretanto a uti-

lização do programa sem a devida comprovação dos processos envolvidos na extração dos dados se mostrou temerária.

Dessa forma, com a finalidade de comprovar a idoneidade dos dados fornecida pelo programa e também de adquirir a informação sobre a formatação e codificação dos dados nos *private streams* dos arquivos sob análise procedeu-se a análise do código do programa *infompg.exe*. A figura 4, mostra a tela do mencionado programa.

A análise do código binário do programa foi realizada através de técnicas de engenharia reversa utilizando-se o programa OllyDbg [2] que é um analisador de linguagem *assembler* (*debugger*) para família de processadores x86.

3.4.2. DETALHES DA CODIFICAÇÃO DOS PRIVATE STREAMS

O estudo dos cabeçalhos dos arquivos questionados mostrou que eles também embutiam, na mesma posição e formato, dados que sugeriam, pelos seus valores, consistir em rótulos de data/hora.

A título de exemplo, a figura 5 mostra a posição e os valores referentes ao primeiro rótulo de data/hora, que traz os dados 12/05/2011 (dia/mês/ano) e 06:47:21 (h:min:s), contidos na sequência de cinco bytes 0x1F 6D 58 6B D5.

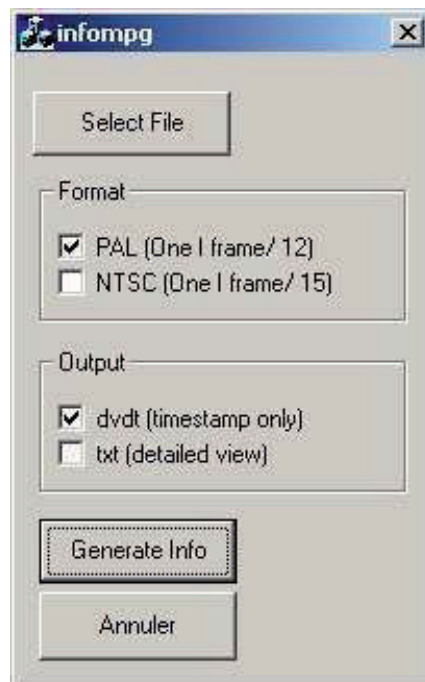


Figura 4: Tela do programa INFOMPG.EXE

O cabeçalho do pacote PES é destacado em verde na figura, e contém, como já visto na discussão da figura 1, o *start code* (sequência de início) 0x000001 e o *stream ID* 0xBD. Em seguida, ainda fazendo parte do cabeçalho, há dois *bytes* que dão o tamanho, em *bytes*, do restante do pacote PES. No exemplo, o tamanho do restante do pacote PES do tipo *private* é de 0x07EC, ou seja, 2.028 *bytes*.

A análise do código revelou que o programa *infompg.exe* realiza uma busca por todo arquivo de vídeo em busca do *start code* 0x000001 seguido pelo *stream ID* 0xBD. Ao encontrar o identificador do *private stream*, o programa localiza o *byte* de número 167 do *payload* do *private stream*.

Os cinco *bytes* destacados em vermelho têm início na 167ª posição após o cabeçalho do pacote PES. Nessa posição, em todos os pacotes PES *private* que antecedem imediatamente cada *I-frame* de todos os arquivos analisados, tanto os questionados quanto os de teste, encontravam-se os respectivos rótulos de data/hora.

Na sequência de cinco *bytes*, foi verificado que:

- os 14 primeiros *bits* representam o ano;

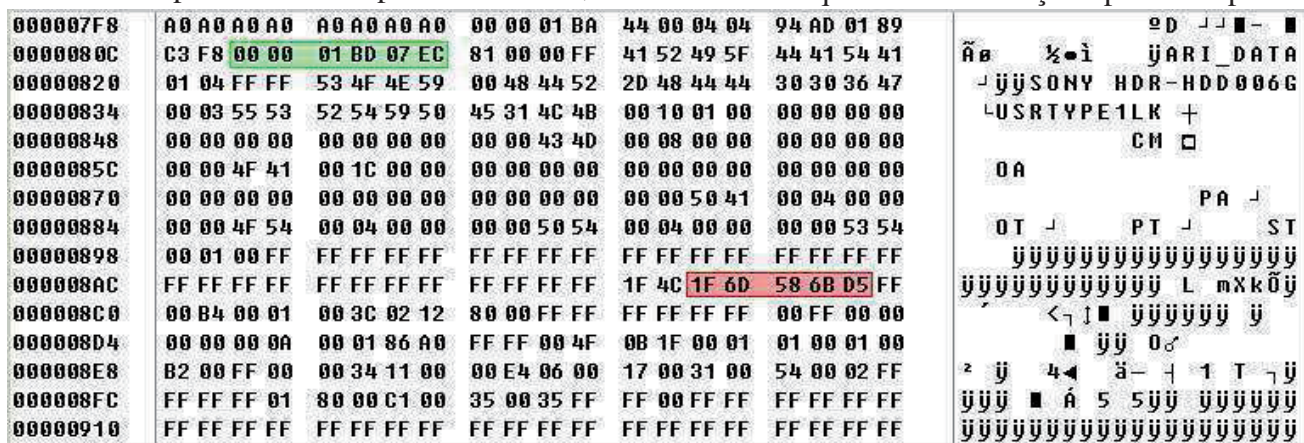


Figura 5: Primeiro rótulo de data/hora do arquivo questionado M2U00354.MPG, destacado em vermelho. Em verde encontra-se destacado o cabeçalho do pacote PES do *private stream*.

- os 4 *bits* seguintes, o mês;
- os 5 *bits* seguintes, o dia;
- os 5 *bits* seguintes, a hora;
- os 6 *bits* seguintes, os minutos;
- e os 6 *bits* restantes, os segundos.

Para o exemplo mostrado na figura 5, a obtenção dos valores de data/hora a partir dos cinco *bytes* destacados em vermelho é detalhado na tabela 1. A última linha da tabela traz, em decimal, o correspondente aos valores em binário de cada campo da linha anterior.

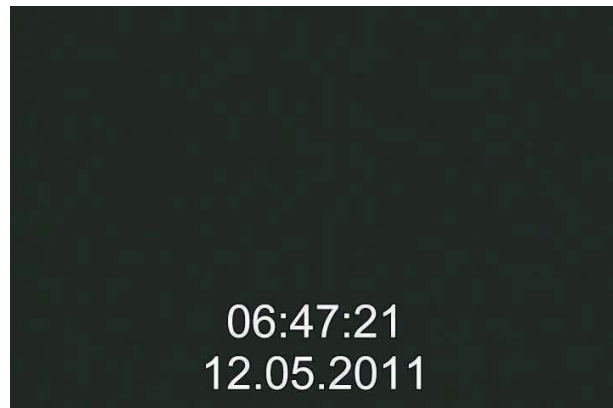


Figura 6: Exemplo de vídeo com legenda data/hora embutida

Por fim extraíram-se dos arquivos de vídeo questionados esses rótulos de data/hora e a partir dessas informações geraram-se novos arquivos em que essa informação aparece explicita-

TABELA 1
VALORES DE DATA/HORA PARA O EXEMPLO DA FIGURA 4

Bytes	0x1F				0x6D				0x58				0x6B				0xD5			
Bits	0	0	0	1	1	1	1	1	0	1	1	0	1	1	0	1	0	1	0	1
Campo	ano				mês				dia				hora				minuto			
Valores no exemplo	2011				5				12				6				47			

mente nas imagens (figura 6), na forma de legendas de contagem de tempo.

No caso em questão constatou-se que as informações de data/hora extraídas dos arquivos evoluíam de forma linear e contínua não apresentando incoerências como repetição, interrupção ou qualquer outra descontinuidade.

Conforme já explanado, é importante ressaltar que, na filmadora, a fonte dessa informação de tempo é dependente da data/hora inicialmente ajustada no equipamento. Além do mais, trata-se de informação que pode ser escrita diretamente nos arquivos por quem conheça a estrutura de dados exposta.

4. CONCLUSÃO

A utilização da técnica de engenharia reversa foi demonstrada para a extração e validação de dados de data/hora em arquivos digitais de vídeo para auxiliar nos Exames de Verificação de Edição.

Apesar de a situação apresentada ser o melhor caso possível, pois havia disponível tanto o equipamento gravador quanto um programa para a localização dos dados de data/hora no cabeçalho, a ideia principal da metodologia é possibilitar o estudo e catalogação da formatação dos cabeçalhos utilizadas pelos diversos fabricantes para que, em uma situação não ideal, onde não se tenha acesso ao equipamento gravador, seja possível acessar os rótulos de data/hora registrados pelo equipamento gravador.

REFERÊNCIAS

- [1] J. Keith, "MPEG 2", in *Video demystified: a handbook for the digital engineer*, 3rd ed. 1995.
- [2] OllyDbg. (2011) Disponível em: <http://www.ollydbg.de>.
- [3] Sony Corporation, DCR-SR68 Handycam Handbook, 2010.

Identificação de origem e verificação de edições em filmes fotográficos: Análise de um caso real.

Rodrigo Gurgel Fernandes Távora, Lucas de Melo Jorge Barbosa

Departamento de Polícia Federal, Brasil

tavora.rgft@dpf.gov.br, lucas.lmjb@dpf.gov.br

RESUMO – Este trabalho contesta a noção de originalidade comumente atribuída a mídias analógicas, na medida em que apresenta um caso real de disputa de propriedade intelectual de imagem comercial, envolvendo filmes fotográficos negativos e diapositivos, além de uma câmera, encaminhados pelas partes litigantes. A identificação da compatibilidade entre a imagem comercial e os filmes, bem como a identificação de origem das imagens, foram realizadas com base na análise de distorções e desenquadramento causados pelo emprego de lentes, além da análise de padrões fixos nas bordas dos filmes, causados por imperfeições da moldura de exposição da câmera empregada no registro.

1. INTRODUÇÃO

Apesar dos recentes esforços, ainda sem sucesso, para desenvolver um sistema de verificação de autenticidade e origem de imagens, por fabricantes como Nikon e Cannon [1], a incerteza de autenticidade ainda é inerente a imagens digitais, o que decorre da facilidade de edição. Por outro lado, negativos e slides consistem em um tipo de mídia analógica tipicamente empregada como mídia original, no primeiro registro da imagem, pelo fato de a gravação da imagem depender da exposição à luz. Por esses motivos, apesar da obsolescência dos filmes fotográficos, até pouco tempo apenas mídias analógicas eram aceitas como provas válidas em algumas cortes e protocolos de serviços de perícias, como no Instituto Nacional de Pesquisa e Polícia Científica do Japão-NRIPS. Entretanto, este entendimento não é correto. O processo de duplicação de filmes é possível e pode ser tanto analógico, por exposição do filme, usando uma montagem com câmera escura ou câmera, quanto digital, usando um esquema híbrido de

conversão A/D por sensor CCD (*charge-coupled device*) e D/A com raios catódicos. Na conversão por exposição, há inclusive disponibilidade de filmes específicos para esta finalidade, com ou sem reversão de cores, para a geração de cópias de segurança de negativos.

O presente artigo contesta esta noção de originalidade, na medida em que apresenta um caso real ocorrido em 2011 de disputa de propriedade intelectual de imagem usada em um cartão postal, onde foram realizados exames de identificação de origem e de verificação de edições em imagens gravadas em filmes fotográficos e em papel impresso.

2. MATERIAL E QUESITOS APRESENTADOS PELAS PARTES

Foram encaminhados pela Parte A, autora da ação, cinco filmes diapositivos de 35mm, com indicações de posições 10, 13, 14, 16 e 17, e pela Parte B um filme negativo de 35mm, como mostra o diagrama na figura 1, que lista ainda as máquinas fotográficas e os quesitos encaminhados. No primeiro momento não foram fornecidas pelas partes a câmera empregada no registro da imagem, que foi então solicitada para permitir uma análise conclusiva de identificação de origem das imagens. A Parte B enviou uma câmera MINOLTA, modelo MAXXUM 300si, e a Parte A informou ser impossível encaminhar a câmera, pois esta fora vendida.

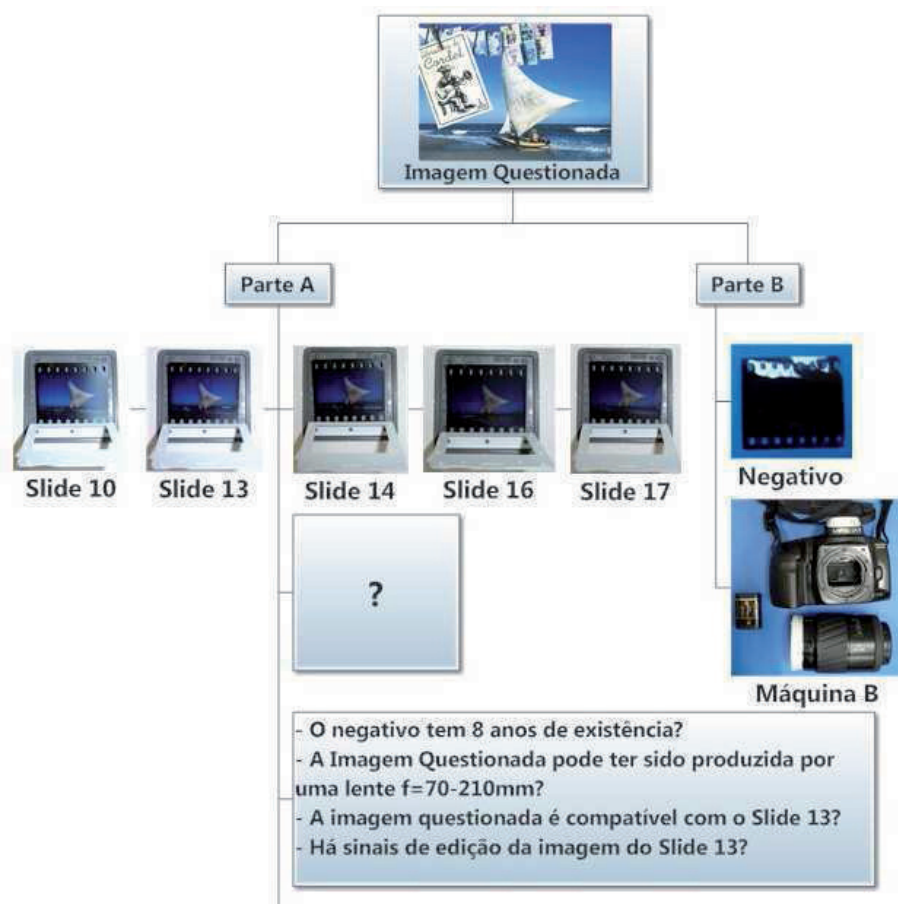


Figura 1: Diagrama com descrição das mídias encaminhadas (acima), da câmera (ao centro), e dos quesitos formulados (abaixo) pelas partes.

Os quesitos formulados pela parte A, tentavam claramente sugerir uma metodologia de análise, questionando a idade dos filmes e a compatibilidade dos elementos da imagem com os parâmetros ópticos da câmera da parte B.

3. METODOLOGIA APLICADA

Preliminarmente à descrição das análises realizadas, cabe ressaltar que as análises sugeridas pela Parte A não são conclusivas. Foram observadas manchas no negativo questionado, mas estas não eram características do efeito da “Síndrome do Vinagre”, podendo ser atribuídas ao atrito do suporte, ou à ação de fungos. Mesmo com uma análise química não seria possível estimar precisamente a idade do negativo, pois a velocidade do processo de hidrólise depende das condições de armazenamento, que não são conhecidas. Ademais, a incompatibilidade do emprego de uma câmera MINOLTA com a lente 70-210mm com a obtenção da imagem do cartão postal não pode ser confirmada, pois para uma estimativa do ângulo de visão seria necessária

a presença de medidas ou distâncias de referência, o que não ocorre na imagem. Além disso, a análise de profundidade de campo não é conclusiva, já que o efeito observado na imagem pode ser obtido de distintas formas e não é específico de uma configuração especial de lente.

A metodologia aplicada foi dividida em diversas análises. Todo o material foi digitalizado com resolução de 4800dpi e gravado em formato TIFF sem compressão. Primeiramente foi realizada a análise de compatibilidade de forma e pertinência das supostas mídias originais e a imagem questionada, onde se observou uma incompatibilidade do negativo com o cartão postal, como descrito no subitem 3.1. Em seguida foram analisados vestígios de edição do negativo, descritos no subitem 3.2, e os vestígios de edição do cartão postal, descritos no subitem 3.3. Por fim, no subitem 3.4, é feita uma análise de compatibilidade de origem dos slides, com base nas minúcias observadas nas bordas da imagem.

3.1. COMPATIBILIDADE DE FORMA E PERTINÊNCIA DAS IMAGENS

A imagem do negativo, mostrada na Figura 2, foi analisada e comparada à imagem do cartão postal. Para permitir a comparação, a imagem do negativo foi sobreposta à imagem do cartão postal, após a aplicação de detecção de bordas, limiarização e inversão horizontal, vide Figura 3. A imagem foi escalada e rotacionada na tentativa de sobrepor os elementos convergentes, entretanto não se conseguiu um acoplamento perfeito dos elementos das imagens. Com o alinhamento dos elementos à esquerda, os elementos à direita e ao centro ficaram desalinhados, vide Figura 4.



Figura 2: Imagem do negativo.

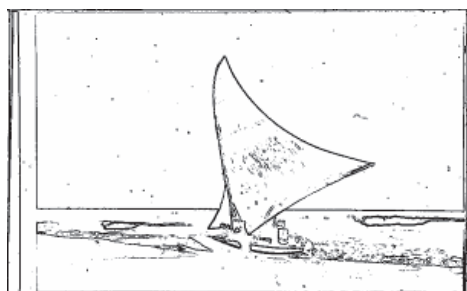


Figura 3: Resultado da detecção de bordas, limiarização e inversão horizontal da imagem da figura anterior.



Figura 4: Tentativa de acoplamento do negativo ao cartão postal, com elementos à esquerda alinhados e os demais desalinhados.

A Figura 4 mostra a não pertinência da imagem do cartão postal na imagem do negativo, já que parte das bordas da imagem estão fora dos limites da imagem do negativo. Portanto, descarta-se totalmente a hipótese do negativo ser a imagem original usada na elaboração do cartão postal.

3.2. VESTÍGIOS DE EDIÇÃO DO NEGATIVO

O emprego de algoritmos automáticos específicos para imagem digital [2], pela análise estatística do padrão de ruído, do método de interpolação para detecção de reamostragem ou dupla-compressão JPEG, ou mesmo de detecção de réplicas parciais dentro da imagem, não se aplicam para a imagem dos filmes fotográficos, por se tratar de registro analógico. Mesmo que o registro original corresponda a uma imagem digital, em uma conversão para mídia analógica, a relação entre os pixels e sua distribuição estatística é mascarada pelo ruído intrínseco à conversão. Dessa forma, na verificação de edições do negativo empregou-se a análise perceptiva dos elementos da imagem, buscando a identificação de texturas repetidas ou de elementos incoerentes.

A análise perceptiva revela que as imagens do negativo e do cartão postal registram uma mesma cena, desconsiderando-se a inserção das imagens dos livros de cordel. A natureza efêmera dos elementos observados, como ondas e homens empurrando uma jangada, descarta a hipótese de as imagens terem sido capturadas em momentos distintos, ou seja, indica que o negativo e o cartão postal foram obtidos a partir de outra imagem original. O descasamento observado nos contornos dos elementos das imagens sobrepostas indica ainda que uma das imagens sofreu distorção de sua forma. A imagem do negativo apresenta uma expansão das bordas da imagem, em relação à imagem do cartão postal, distorção semelhante ao efeito gerado em fotografia pelo emprego de lentes, como ilustrado na Figura 5, com um reticulado distorcido pela lente.

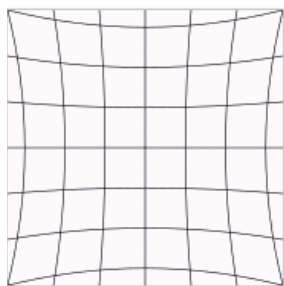


Figura 5: Distorção radial tipo “pincushion” (travesseiro).

Esta observação levanta a hipótese de o negativo ter sido obtido através do registro fotográfico, com emprego de lentes, da imagem de fundo do cartão postal (sem as inserções das imagens dos livros de cordel). Pelo exame das bordas do filme negativo, vide Figura 6, é possível observar, indicadas em verde e ampliadas, faixas brancas em forma de cunha no canto inferior esquerdo, na vertical, e no canto inferior direito, na horizontal, incompatíveis com o cenário natural da imagem, ou com efeitos típicos de fotografia, o que reforça a suspeita quanto à autenticidade da imagem. A presença dessas cunhas com ângulo no mesmo sentido sugere que o negativo foi gerado fotografando-se uma imagem impressa em papel branco, ou posicionada sobre fundo branco, com o enquadramento da câmera desalinhado com os eixos da imagem, como exemplificado na Figura 7.

A imagem parcial da posição seguinte do filme, indicada em vermelho na Figura 6, onde também há uma faixa branca, sugere que com o mesmo filme foi tirada mais de uma fotografia da imagem da paisagem do cartão postal com fundo branco, na tentativa de converter a imagem para o filme fotográfico. Isto é corroborado também pela compatibilidade das faixas verticais, indicadas em amarelo, situadas no canto esquerdo das duas posições consecutivas do filme, ambas apresentando variação da tonalidade na escala de azul.



Figura 6: Exame de incoerências nas bordas da imagem do negativo.



Figura 7: Ilustração de hipótese de desenquadramento de fotografia da imagem para gerar o negativo como causa das cunhas observadas.

3.3. VESTÍGIOS DE EDIÇÃO DO CARTÃO POSTAL

Pela mesma razão anteriormente exposta, algoritmos automáticos não são aplicáveis na verificação de edições do cartão postal. Assim, na comparação do cartão postal com o slide 13, que contém imagem semelhante, empregou-se a análise perceptiva. Para facilitar a comparação, a imagem do cartão postal foi realçada com detecção de bordas, limiarização e inversão horizontal, vide Figura 8, a qual foi sobreposta, após os ajustes de rotação e escala, à imagem do slide 13, vide Figura 9. Pode-se observar a pertinência de quase toda imagem do cartão postal na imagem do slide 13, com exceção de uma estreita faixa com tonalidade constante na parte superior. Também se observa o acoplamento perfeito dos elementos presentes nas duas imagens, com exceção da presença, na imagem do cartão postal, de livros de cordel, notadamente gerados a partir de uma editoração de arte gráfica, e da ausência de alguns elementos originalmente presentes na imagem do slide 13, destacados em vermelho na Figura 9.

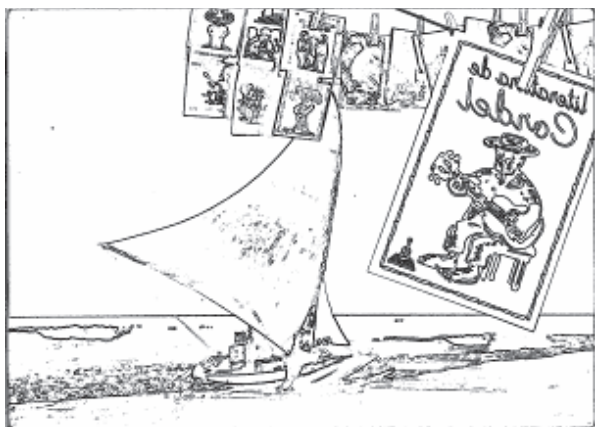


Figura 8: Resultado da detecção de bordas, limiarização e inversão horizontal da imagem do cartão postal.



Figura 9: Sobreposição da imagem processada do cartão postal à imagem do slide 13, após ajustes de escala e rotação



Figura 10: Detalhe da parte inferior da imagem do cartão postal, sobreposta de partes da imagem do slide 13 com indicação dos vestígios de edição da imagem.

A ausência desses elementos não configura uma divergência na origem das imagens, pois uma análise minuciosa do cartão postal revela vestígios destes elementos ainda remanescentes na imagem do cartão postal, vide Figura 10, como 1) a réplica de textura indicando cópia e colagem de imagem de uma marola para substituição da imagem removida de um pescador; 2) a imagem remanescente do reflexo na areia molhada, com cores vermelha e azul, do calção do pescador, cuja imagem estava presente na imagem do slide 13; 3) o resquício de parte da imagem do boné do pescador, cuja imagem estava presente na imagem do slide 13, mas foi superposta por imagem do mar no cartão postal, e 4) a imagem de banhista, borrada, mas ainda identificável, que no slide 13 aparecia de forma mais nítida.

O tipo de edição observada no cartão postal denota uma intenção de ajuste estético

da paisagem, aparentemente sem fins fraudulentos. A estreita faixa de tonalidade constante na parte superior do cartão postal, fora das margens do slide 13, pode ter sido gerada pela ampliação da área da imagem por colagem de uma faixa de céu, possivelmente objetivando o ajuste estético da imagem para a inclusão dos elementos de cordel. Dessa forma, pode-se concluir que há compatibilidade entre a imagem do slide 13 e a imagem empregada na produção do cartão postal.

A natureza efêmera dos elementos da imagem, como ondas e a presença de homens empurrando uma jangada, e a compatibilidade das imagens descartam a hipótese de as imagens terem sido capturadas em momentos distintos, e reforçam a hipótese de o cartão postal ter sido gerado a partir do slide 13.

3.4. COMPATIBILIDADE DE ORIGEM DOS SLIDES

A identificação da câmera fotográfica analógica fonte pela imagem do filme é possível através do método proposto em [3], de análise de um padrão fixo de minúcias presentes nas bordas da imagem dos filmes usados na mesma câmera, consequência de imperfeições na moldura plástica da abertura de exposição nas máquinas fotográficas. Estas irregularidades, decorrentes do processo de fabricação, apresentam aleatoriedade na forma e na posição, e podem, portanto, ser empregadas como uma assinatura para cada câmera.

Como a câmera da Parte A não foi encaminhada para análise, e como constatou-se que a máquina da Parte B estava inoperante, não foi possível realizar uma identificação de origem conclusiva. A análise das irregularidades das bordas ficou restrita à verificação de compatibilidade de origem dos slides encaminhados.

Por esta análise, observou-se que as bordas inferior, direita, esquerda e superior de todos os slides apresentavam minúcias convergentes indicadas por setas vermelhas, decorrentes do efeito de obstrução e difração da luz em torno das irregularidades da moldura, mesmo com variação de enquadramento das imagens. A Figura 11 e a Figura 12, mostram as bordas inferior e esquerda dos slides. Dessa forma, pode-se concluir que os slides foram registrados com o emprego de uma mesma câmera.

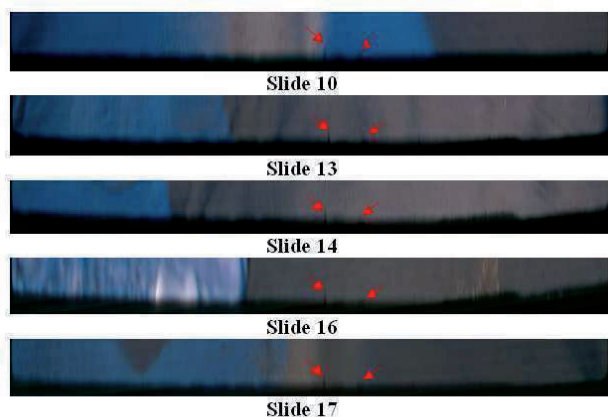


Figura 11: Detalhe da borda inferior dos slides, com minúcias equivalentes.

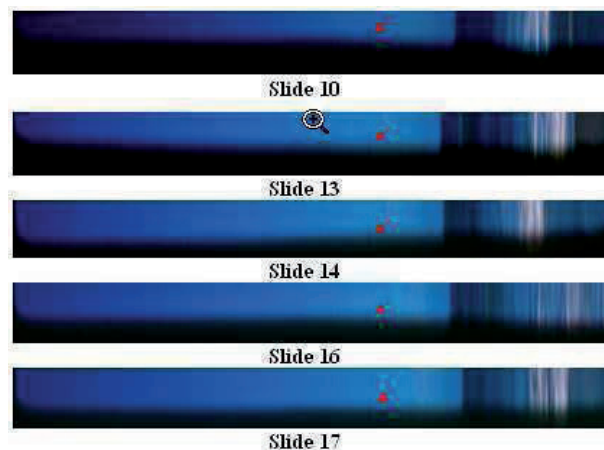


Figura 12: Detalhe da borda esquerda dos slides, com minúcias equivalentes.

A verificação de edição nos slides 10, 14, 16 e 17 foi feita através da análise perceptiva, pelos mesmos motivos anteriormente expostos, e nenhuma edição foi observada. Também não foram observados elementos sugestivos de duplicação analógica, como erros de enquadramento nas bordas. Além disso, a presença de elementos semelhantes, como pescadores e a mesma jangada, e a numeração sequencial dos slides sugerem que as imagens foram capturadas em um intervalo curto de tempo.

4. CONCLUSÃO

A análise apresentada permitiu, observando os efeitos de distorções da imagem do negativo e o desenquadramento das bordas da imagem sob fundo branco, indicar que o negativo foi forjado como mídia original, possivelmente empregando uma montagem de exposição com lentes. Apesar da ausência da câmera fotográfica, que viabilizaria uma análise conclusiva, a posse dos cinco slides registrados de uma mesma câmera e sem sinais de edição reforçam a hipótese de que o slide 3 da Parte A, que pelas análises mostrou-se compatível com a imagem do cartão postal, corresponda à imagem original usada na confecção desta.

REFERÊNCIAS

- [1] <http://www.canon.com.au/Support-Services/Support-News/OSK-DVK-Phenomenon>
- [2] P. Popescu, H. Farid. Statistical Tools for Digital Forensics, In 6th International Workshop on Information Hiding, 2004.
- [3] Shirase, Yonemura. Identification of camera using film images, Jasti Abstracts, p.162, 2001

Detecção de Nudez Utilizando Redes Neurais Artificiais

Pedro S. Franco, Daniel A. M. Sandoval, Alexandre Zaghetto, Célia Ghedini Ralha

Resumo—O presente artigo propõe um método baseado em redes neurais artificiais para detecção de nudez em imagens digitais a ser incorporado ao MADIK (*Multi-agent Digital Investigation ToolKit*) em um agente de processamento de imagens. O objetivo é utilizá-lo como ferramenta de auxílio a peritos em investigações que envolvem crimes como a pedofilia. O método proposto pode ser dividido em 5 etapas: (a) segmentação por cor de pele; (b) realização de operações morfológicas; (c) rotulação e contagem de elementos conectados; (d) medição da área do maior elemento conectado; (e) classificação por meio de uma rede neural artificial do tipo *feedforward*, treinada a partir do algoritmo Levenberg-Marquardt. Nosso conjunto de testes é formado por 1000 imagens, 500 contendo nudez e 500 não. Em nossos experimentos, foram treinadas 10 vezes cada uma das redes neurais com 1 a 15 neurônios na camada escondida, totalizando 150 redes. São calculadas as taxas de verdadeiros e falsos positivos e negativos e a F-measure é utilizada como métrica para seleção da rede que apresenta o melhor desempenho. A F-measure observada no melhor caso foi de 0.7967, para uma rede com 5 neurônios na camada escondida.

Index Terms—Pedofilia, Detecção de nudez, Redes Neurais Artificiais, *Multi-Agent Digital Investigation ToolKit*, Investigação forense.

I. INTRODUÇÃO

Com o avanço tecnológico e a popularização de máquinas fotográficas digitais e dispositivos de armazenamento, vastas bibliotecas digitais de imagens digitais tornaram-se comuns. No combate a crimes como a pedofilia, a inspeção humana dessas vastas bibliotecas torna-se muito difícil, desgastante e por vezes inviável.

O processo investigativo de sistemas computacionais resulta em uma tarefa complexa, envolvendo a participação e cooperação de diversos especialistas forenses. No entanto, a excelência do processo pode ser limitada devido à falta de recursos, sejam eles técnicos, científicos ou computacionais. Visando uma maior agilidade e eficiência dos processos investigativos foi desenvolvido o MADIK (*Multi-Agent Digital Investigation ToolKit*) [1], [2]. Trata-se de uma ferramenta baseada na abordagem de sistema multiagente distribuído, o qual se revelou apropriado, principalmente devido à capacidade de cooperação de agentes autônomos especializados [3]. No MADIK foram implementados diferentes agentes, cada qual baseado em uma área específica de conhecimento técnico utilizando uma arquitetura composta de quatro camadas, tal como uma metáfora à hierarquia das organizações: estratégica, tática e operacional. Na camada operacional foram definidos cinco agentes especialistas:

HashSetAgent, *TimeLineAgent*, *FileSignatureAgent*, *FilePathAgent* e *KeywordAgent*. Os agentes especialistas interagem durante o processo investigatório, correlacionando achados sobre os arquivos em análise através do uso de um *blackboard* colaborativo, o qual permite uma avaliação mais detalhada de cada arquivo sob investigação.

Tendo em vista a importância de vestígios elucidatórios em conteúdos digitais de imagens, este trabalho visa a extensão do MADIK com o desenvolvimento do agente especialista em processamento de imagens de alta resolução, denominado *ImageAgent*. De forma geral, o *ImageAgent* deve ser capaz de analisar o conteúdo de arquivos de tipos diversos de imagens, para permitir a identificação, extração e classificação de características de interesse de uma investigação criminal, sendo que o método proposto está mais especificamente ligado à detecção de imagens de nudez, com a finalidade de auxiliar no combate a pedofilia.

II. MÉTODO PROPOSTO

O método proposto pode ser dividido em cinco etapas: (i) segmentação por cor de pele; (ii) realização de operações morfológicas; (iii) rotulação e contagem de elementos conectados; (iv) medição da área do maior elemento conectado; e (v) classificação por meio de uma rede neural artificial [4] do tipo *feedforward*, treinada a partir do algoritmo Levenberg-Marquardt [5]. Cada uma dessas etapas será detalha a seguir.

A. Segmentação por Cor de Pele

Dentre os espaços de cores com os quais se pode operar, o HSV (*Hue, Saturation, Value*) é indicado pela literatura como um dos mais adequados para efeito de segmentação por cor de pele [6], [7], sendo o plano determinado por *H* e *S* utilizado para tal finalidade. Uma vez que a componente *V* não é levada em consideração, a interferência da luminância no processo de identificação das regiões de interesse é minimizada [8].

As cores típicas para pele humana foram primeiro estudadas por Felix Ritter Von Luschan [9], que desenvolveu uma escala de classificação denominada Escala de Von Luschan. A Figura 1 mostra um instrumento antropométrico que utiliza a classificação proposta por Von Luschan [10]. Em um cenário mais realista uma imagem composta por segmentos de fotografias digitais de pele humana por ser utilizada como referência na avaliação da escala proposta por Von Luschan. Um exemplo de imagem, bem como o plano $H \times S$ correspondente, são mostrados na Figura 2. A partir da análise

E-mail: {pedro, daniel}@loopec.com.br, {alexandre, ghedini}@cic.unb.br.
Departamento de Ciência da Computação, Universidade de Brasília

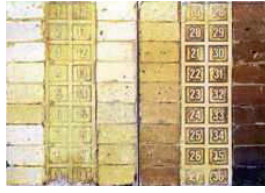
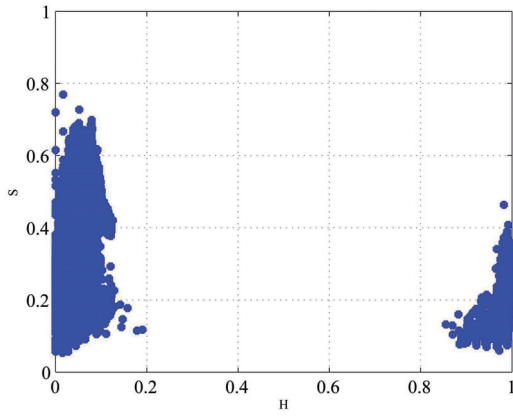


Fig. 1. Instrumento antropométrico construído a partir da Escala de Von Luschan¹.



(a)



(b)

Fig. 2. (a) Recortes de fotografias de pele; e (b) localização das cores no plano $H \times S$.

de gráficos dessa natureza, pode-se estabelecer os intervalos de interesse $H_i = [0.0.11]$ e $S_i = [0.2.0.7]$ para as componentes H e S , respectivamente [11]. *Pixels* cujos valores de H e S estiverem contidos nesses intervalos são marcados em uma máscara binária com o valor 1. Os demais são marcados com 0. Trata-se de um método de classificação trivial que pode em trabalhos futuros ser substituído por métodos mais robustos.

Por fim, dado que a complexidade computacional é um problema a ser considerado, a máscara binária resultante é subamostrada por um fator de 0.5 antes de ser submetida às etapas seguintes de processamento.

B. Realização de Operações Morfológicas

Após a binarização e a subamostragem descritas na etapa anterior, a imagem passa pelos processamentos morfológicos de abertura e fechamento [12]. O objetivo é diminuir a quantidade de regiões que erroneamente foram classificadas como regiões de interesse e conectar outras que por



(a)



(b)



(c)

Fig. 3. Binarização e operações morfológicas: exemplo de imagem original (a), e suas versões binarizadas antes (b) e depois (c) do processamento sugerido.

ventura tenham sido desconectadas pela binarização. A Figura 3 mostra um exemplo de imagem original (a), e suas versões binarizadas antes (b) e depois (c) do processamento morfológico sugerido.

C. Rotulação e Contagem de Elementos Conectados

Considerando que as imagens binárias são representadas pelos bits 1 (branco) e 0 (preto), a metodologia adotada para se rotular os elementos de interesse está resumida no algoritmo *ROTULA*, onde M é a imagem binária e R é a matriz de rótulos.

ROTULA(M)

```

1   $R(i, j) \leftarrow 0$ 
2  for each  $M(i, j) = 0$  do
3      if  $M(i, j - 1) = 1$  and  $M(i - 1, j) = 1$ 
4          then  $Indice = Indice + 1$ ;
5               $R(i, j) \leftarrow Indice$ ;
6      elseif  $M(i, j - 1) = 0$  and  $M(i - 1, j) = 1$ 
7          then  $R(i, j) \leftarrow R(i, j - 1)$ ;
8      else  $R(i, j) \leftarrow R(i - 1, j)$ ;
9  return R
```

Em uma segunda passagem, estabelece-se a equivalência entre rótulos, quando regiões vizinhas possuem rótulos diferentes. Após essa etapa, tem-se a informação de quantos elementos conectados há na imagem. A Figura 4 ilustra o resultado do processo de rotulação. Nesse exemplo foram rotuladas 4 regiões: o fundo (preto) mais três regiões de cor de pele (demais tons de cinza).

D. Medição da Área do Maior Elemento Conectado

Nessa etapa, calcula-se a área em *pixels* para cada elemento conectado identificado na etapa anterior e, em seguida,

¹Museo di Storia Naturale, Università degli Studi di Firenze - <http://www.msn.unifi.it/CMpro-v-p-69.html>.



Fig. 4. Exemplo de rotulação: foram rotuladas 4 regiões: o fundo mais três regiões de cor de pele.



Fig. 5. Destaque do elemento conectado de maior área: aproximadamente 23% da área total da imagem.

procura-se pelo maior valor observado. É importante ressaltar que são considerados como candidatos nessa etapa apenas os elementos conectados que encontram-se em sua maioria dentro da área classificada como cor de pele na etapa de segmentação. Dessa forma, evita-se contar o fundo como um possível elemento conectado. Por fim, calcula-se a razão entre a área do maior elemento conectado e a área total da imagem. A Figura 5 destaca a região de maior área para o exemplo em questão. Esse elemento conectado corresponde a aproximadamente 23% da área total da imagem.

E. Classificação Utilizando Redes Neurais Artificiais

O objetivo do classificador proposto é identificar imagens candidatas a conterem nudez. Optou-se por utilizar uma rede neural artificial [4] do tipo *feedforward*, treinada a partir do algoritmo Levenberg-Marquardt [5]. Foi proposta uma arquitetura de 3 camadas: entrada, saída e oculta. Na camada de entrada os neurônios não realizam qualquer computação. Já nas camadas de saída e oculta, a função de ativação tangente hiperbólica é empregada. Um único neurônio compõe a camada de saída, enquanto o número de neurônios na camada oculta variou de 1 a 15 durante o treinamento.

O vetor de entrada P para o classificador é composto por dois parâmetros: (i) o número de elementos conectados encontrados na imagem; e (ii) a área relativa do maior elementado conecatado. Esses dois valores são extraídos nas etapas descritas nas Subseções II-C e II-D, respectivamente.

A rede é treinada para responder com o valor $T = 1$, se o vetor de entrada P corresponde a uma imagem candidata a conter nudez, e com $T = -1$, caso contrário. É importante ressaltar que na operação os neurônios de saída não respondem exatamente com 1's ou ou -1's, mas apenas com valores aproximados. Por isso, é necessário incluir uma etapa de limiarização. Caso a resposta do neurônio de saída seja maior que 0, a imagem é marcada como contendo nudez. Por outro lado, se a resposta for menor ou igual 0, a imagem não é tida como suspeita.

A métrica de desempenho utilizada foi a *F-measure* [13], definida pela Equ. 1,

$$F\text{-measure} = 2 \cdot \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \quad (1)$$

onde *precision* e *recall* são definidas por,

$$\text{precision} = \frac{tp}{(tp + fp)} \text{ e } \text{recall} = \frac{tp}{(tp + fn)},$$

e tp , fp e fn são as taxas de verdadeiros positivos (*true positive*), falsos positivos (*false positive*) e falsos negativos (*false negative*), respectivamente. De uma maneira geral, a medida *precision* de um sistema de reconhecimento de padrões representa a relação entre a quantidade de classificações verdadeiramente positivas (tp) retornadas pelo sistema e a quantidade total de classificações retornadas como positivas ($tp + fp$). Em outras palavras, representa a probabilidade de uma classificação retornada pelo sistema como positiva ser uma ocorrência verdadeiramente positiva. Já a medida *recall* representa a relação entre a quantidade de classificações verdadeiramente positivas retornadas pelo sistema (tp) e a quantidade de classificações que deveriam ter sido retornadas como positivas ($tp + fn$). Ou seja, representa a probabilidade de uma ocorrência verdadeiramente positiva ser classificada como tal pelo sistema. A *F-measure* é, dessa forma, a média harmonica entre *precision* e *recall*.

III. RESULTADOS

Nossa base de testes é formada por 1000 imagens baixadas da Internet, 500 contendo nudez e 500 não. A Figura 6 mostra a distribuição dos vetores de entrada P , definidos a partir dessas imagens. O eixo vertical representa o número de elementos conectado e o eixo horizontal representa a área relativa do maior elemento conectado. Para o treinamento da rede, as duas classes de imagens, nudez e não-nudez, são divididas em três subconjuntos: treinamento, validação e teste, cada qual contendo aproximadamente 1/3 das imagens. O conjunto de treinamento é utilizado para se computar o gradiente e atualizar os pesos e *bias* da rede. O conjunto de validação é utilizado para interromper o processo de aprendizando assim que a rede começa a apresentar super-aprendizado (*overfitting*). O conjunto de teste é utilizando na avaliação do desempenho final da rede.

Durante o treinamento, variou-se a quantidade de neurônios na camada escondida de 1 a 15 e cada rede foi treinada 10 vezes, com seus pesos iniciados aleatoriamente a cada treinamento. O melhor desempenho observado ($F\text{-measure} = 0.7967$) foi obtido com 5 neurônios da camada escondida. As

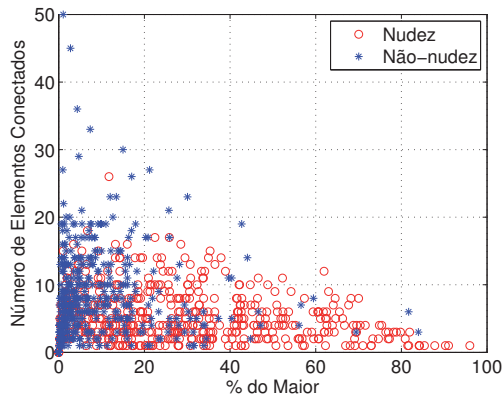
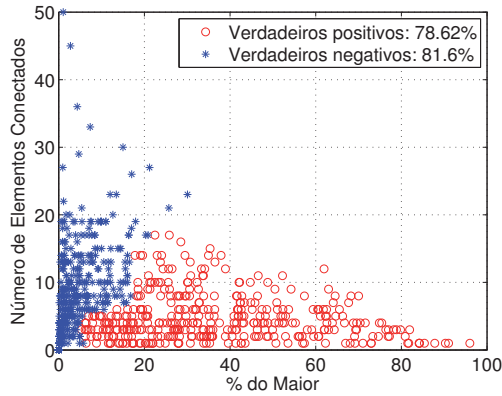
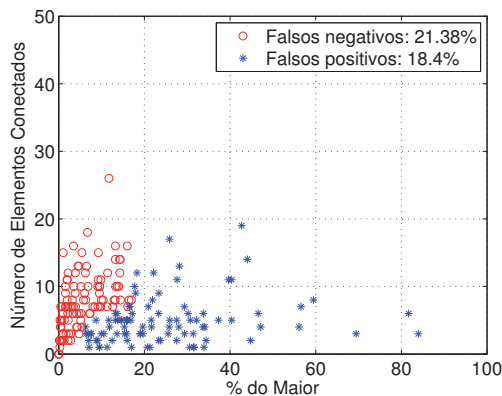


Fig. 6. Distribuição dos vetores de entrada P , definidos a partir das imagens da base de teste. O eixo vertical representa o número de elementos conectados e o eixo horizontal representa a área relativa do maior elemento conectado.



(a)



(b)

Fig. 7. (a) Verdadeiros positivos e negativos; e (b) Falsos positivos e negativos. $F\text{-measure} = 0.7967$.

Figuras 7 (a) e (b) mostram a distribuição de verdadeiros positivos e negativos; e falsos positivos e negativos, respectivamente.

IV. CONCLUSÕES

Foi apresentado um método baseado em redes neurais artificiais para detecção de nudez em imagens digitais. Pretende-se, com isso, auxiliar na investigação de crimes de pedofilia. Os resultados mostraram a viabilidade do método proposto, no entanto sugerem a necessidade de se incluir descritores de imagem mais adequados, em substituição ou além daqueles que foram utilizados no presente trabalho, que são: o número de elementos conectados e a área do maior elemento conectado. Em trabalhos futuros pretende-se, ainda, incluir detectores de face, de formas corporais e órgãos genitais. A versão final do método será incorporada à ferramenta de investigação forense MADIK em um agente de processamento de imagens.

REFERÊNCIAS

- [1] Bruno W. P. Hoelz, Célia Ghedini Ralha, Rajiv Geeverghese, and Hugo C. Junior, "A cooperative multi-agent approach to computer forensics," in *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Sydney, NSW, Australia, December 9-12, 2008*, 2008, pp. 477–483, IEEE Computer Society.
- [2] Bruno W. P. Hoelz, Célia Ghedini Ralha, Rajiv Geeverghese, and Hugo C. Junior, "Madik: A collaborative multi-agent toolkit to computer forensics," in *On the Move to Meaningful Internet Systems: OTM 2008 Workshops, OTM Confederated International Workshops and Posters, ADI, AWeSoMe, COMBEK, EI2N, IWSSA, MONET, OnToContent + QSI, ORM, PerSys, RDDS, SEMELS, and SWWS 2008, Monterrey, Mexico, November 9-14, 2008. Proceedings*, 2008, vol. 5333 of *Lecture Notes in Computer Science*, pp. 20–21, Springer-Verlag.
- [3] Bruno W. P. Hoelz, Célia Ghedini Ralha, and Rajiv Geeverghese, "Artificial intelligence applied to computer forensics," in *Proceedings of the 2009 ACM Symposium on Applied Computing (SAC), Honolulu, Hawaii, USA, March 9-12, 2009*, 2009, pp. 883–888, ACM.
- [4] Simon Haykin, *Neural Networks: A Comprehensive Foundation (2nd Edition)*, Prentice Hall, 2 edition, July 1998.
- [5] M. T. Hagan and M. Menhaj, "Training feedforward networks with the marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, pp. 989–993, 1994.
- [6] Benjamin D. Zarit, Boaz J. Super, and Francis K. H. Quek, "Comparison of five color models in skin pixel classification," in *In ICCV 99 Int. Workshop on*, 1999, pp. 58–63.
- [7] Juwei Lu, Qian Gu, K. N. Plataniotis, and Jie Wang, "A comparative study of skin-color models," in *Proceedings of the Second international conference on Image Analysis and Recognition*, 2005, ICIAR'05, pp. 729–736.
- [8] Vladimir Vezhnevets, Vassili Sazonov, and Alla Andreeva, "A survey on pixel-based skin color detection techniques," in *IN PROC. GRAPHICON-2003*, 2003, pp. 85–92.
- [9] F. von Luschan, *Völker, Rassen, Sprachen : Anthropologische Betrachtungen*, Deutsche Buchgemeinschaft, Berlin, 1927.
- [10] Università degli Studi di Firenze, "Strumenti antropometrici," 2011.
- [11] Anjali A. Shejul and Umesh L. Kulkarni, "A Secure Skin Tone based Steganography Using Wavelet Transform," *International Journal of Computer Theory and Engineering*, vol. 3, no. 1, pp. 16–22, February 2011.
- [12] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, EUA, 2nd edition, 2002.
- [13] C. J. van Rijsbergen, *Information Retrieval*, Butterworth-Heinemann, 2 edition, Mar. 1979.

Método para Compatibilidade entre Impressões Digitais com e sem Contato

Alexandre Zaghetto¹, Pedro Salum Franco¹ e Daniel Assad Maia Sandoval¹

Resumo—Soluções em biometria por impressões digitais sem contato têm sido propostas com o objetivo de se superar os problemas intrínsecos relacionados a tecnologias que exigem o contato. Para desfrutarmos das vantagens operacionais de dispositivos sem contato, as imagens resultantes desse tipo de captura precisam ser processadas para que se tornem similares àquelas obtidas nos digitalizadores convencionais. O presente artigo propõe uma solução em duas etapas para esse problema: a primeira procura reproduzir a textura de impressões digitais tradicionais; a segunda procura fazer refletir o processo de rolagem unha-a-unha. Dois experimentos foram propostos, um para avaliar a qualidade das imagens processadas, e outro para estabelecer a EER (equal error rate) para um conjunto de 200 impressões digitais (100 dedos, 2 imagens por dedo). Observou-se que 90% das imagens apresentaram índices bom, muito bom ou excelente, e a EER ficou em 4%, demonstrando a viabilidade do método proposto.

I. INTRODUÇÃO

A autenticação biométrica, ou simplesmente biometria, pode ser definida como a verificação ou o reconhecimento automático da identidade de um indivíduo a partir de características fisiológicas ou comportamentais [1]. Impressões digitais, geometria da mão, voz, íris, face, caligrafia e dinâmica de digitação são exemplos dessas características. Em geral, diferentes sistemas biométricos requerem tecnologia específica, dependendo da característica fisiológica/comportamental considerada. No entanto, independentemente das especificidades dos diversos sistemas biométricos, há problemas que precisam ser tratados pela maioria deles. Como exemplo pode-se citar a interoperabilidade entre os dispositivos que operam a partir de uma determinada biometria. O presente artigo trata da interoperabilidade entre dispositivos biométricos baseados em impressões digitais adquiridas com e sem contato.

A. Impressões Digitais com Contato

Não é tarefa fácil determinar qual característica biométrica é a mais adequada, considerando a diversidade de aplicações e condições de operação. No entanto, em aplicações forenses, civis e comerciais [2], impressões digitais têm se tornado largamente utilizadas. A entrada de um sistema de autenticação por impressões digitais é uma imagem digital que representa a estrutura de vales e franjas de um dedo real. Isso torna o processo de aquisição uma etapa crítica, que afeta significativamente o desempenho geral do sistema. A maior parte das

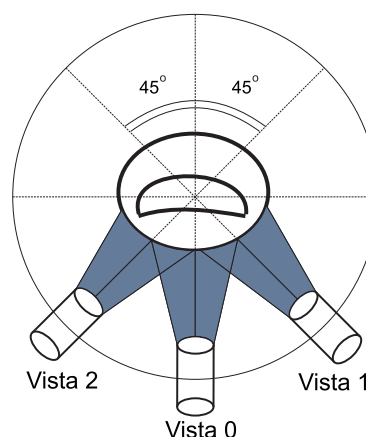


Fig. 1. Ilustração esquemática de um sistema de biometria por impressão digital multivista sem contato.

atuais tecnologias que lidam com impressões digitais é baseada em contato, ou seja, requerem que o usuário pressione os dedos contra uma superfície de aquisição. O maior problema advindo dessa tecnologia são as distorções e inconsistências que podem ser introduzidas devido à elasticidade da pele. A qualidade das impressões digitais pode também ser seriamente afetada pelo contato não ideal causado por sujeira, suor, oleosidade, excesso de secura ou humidade, temperatura e impressões latentes. Além disso, partes diferentes do mesmo dedo podem ser capturadas a cada vez que o mesmo é apresentado ao sensor, resultando em amostras irreprodutíveis. Em alguns cenários, tais desvantagens podem impor a necessidade de um operador qualificado durante a fase de registro e, em muitos casos, múltiplas tentativas são efetuadas por dedo até que sejam produzidas amostras com qualidade suficiente. Em aplicações governamentais como, por exemplo, o uso de biometria nas urnas eletrônicas empregadas pelo sistema eleitoral brasileiro¹, o processo de registro da população votante pode revelar-se extremamente lento. Nesse exemplo específico, a expectativa do Tribunal Superior Eleitoral (TSE) é de que até 2018 todos os eleitores possam votar utilizando-se desse recurso². Apesar de nos últimos anos alguns algoritmos terem sido propostos com o objetivo de se compensar as limitações dos dispositivos com contato, esse paradigma de captura pode representar um gargalo no que tange o apromiramento da qualidade de imagens de impressões digitais.

¹Departamento de Ciência da Computação, Universidade de Brasília, e-mail: zaghetto@image.unb.br, pedro@loopeccom.com.br, daniel@loopeccom.com.br.

¹<http://www.tse.jus.br/eleicoes/biometria-e-urna-eletronica>

²https://www.youtube.com/watch?v=A_ITOLoRtL4

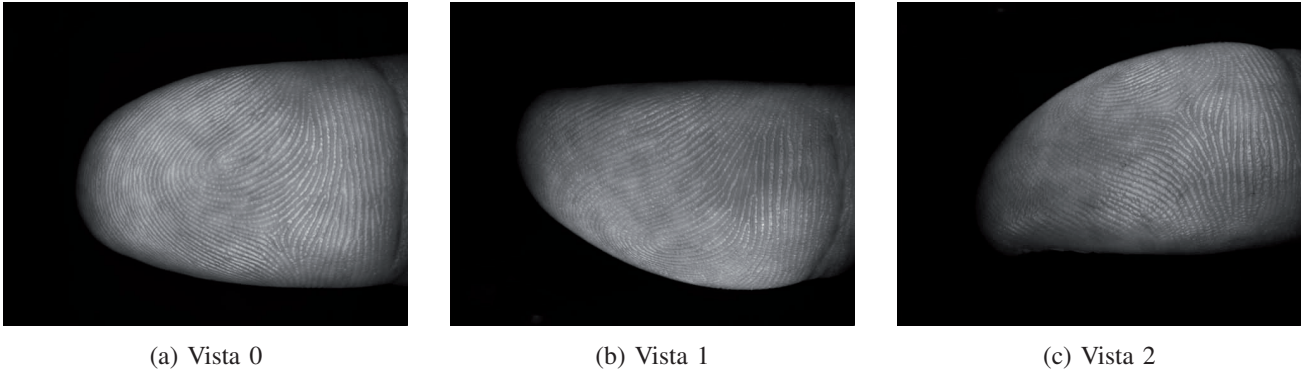


Fig. 2. Exemplos de vistas (1024×1280 pixels, 8 bits/pixels) adquiridas por meio de um dispositivo sem contato multivista.

B. Impressões Digitais sem Contato

Soluções em biometria por impressões digitais sem contato têm sido propostas [3], [4], [5] com o objetivo de se superar os problemas intrinsecamente relacionados a tecnologias que exigem o contato. Tais soluções procuram atacar o problema da qualidade da impressão digital em seu nível mais fundamental. Uma vez que sistemas sem contato não obrigam o usuário a pressionar os dedos em uma superfície, tais sistemas não requerem, por exemplo, algoritmos que procuram compensar os artefatos decorrentes das propriedades elásticas da pele.

Dentre as propostas de soluções sem contato, os dispositivos desenvolvidos pela empresa *Touchless Biometric Systems* (TBS)³ utilizam uma abordagem bastante interessante e promissora [6]. Esses dispositivos combinam o imageamento baseado em reflexão com um sistema multivista construído a partir de três câmeras. Uma câmera é posicionada de maneira a capturar a parte do dedo em que normalmente núcleo e deltas estão localizados e, tomando-se essa câmera central como referência, as outras duas são posicionadas a 45 graus, uma no sentido horário e a outra no sentido anti-horário. Se a sobreposição dos seus campos de visão é garantida, pode-se não apenas registrar fotograficamente uma área ampliada, como também gerar, por meio de algoritmos de costura [7], uma impressão digital equivalente a uma obtida por processo de rolagem (unha-a-unha). As Figuras 1 e 2 mostram uma ilustração esquemática de um dispositivo de captura sem contato multivista, bem como exemplos de vistas capturadas, respectivamente.

C. Compatibilidade

As imagens capturadas por meio da tecnologia sem contato em geral não são compatíveis com as imagens obtidas em digitalizadores ópticos, principalmente aqueles em que as impressões digitais, para serem obtidas, requerem a rolagem do dedo sobre uma superfície plana, transparente e retro-iluminada. A Figura 3 mostra um exemplo.

Para desfrutarmos das vantagens operacionais dos dispositivos sem contato, as imagens resultantes desse tipo de captura precisam ser processadas para que se tornem similares àquelas

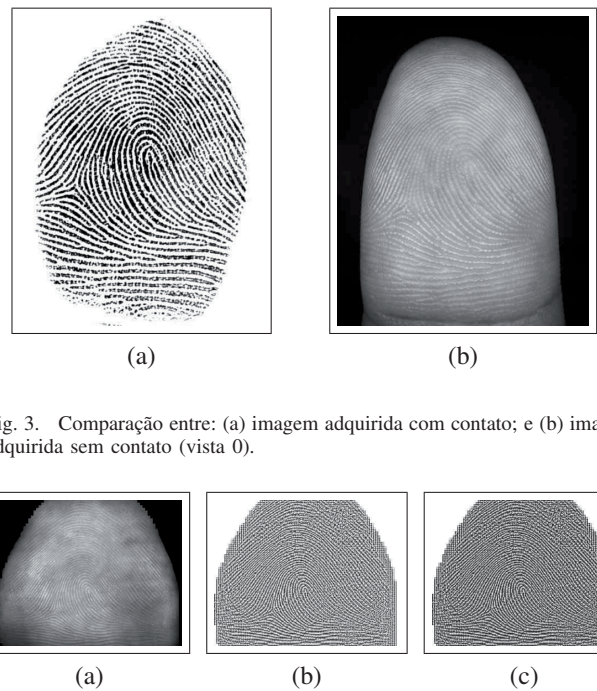


Fig. 3. Comparação entre: (a) imagem adquirida com contato; e (b) imagem adquirida sem contato (vista 0).

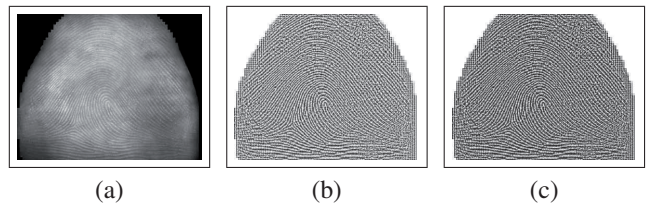


Fig. 4. Equalização local de histograma e transformação gama: (a) imagem original, I (712×560 pixels, 256 níveis de cinza); (b) image equalizada, I_h ; e (c) resultado da transformação gama, $I_{h\gamma}$.

obtidas nos digitalizadores convencionais. O presente artigo propõe uma solução para esse problema.

II. MÉTODO PROPOSTO

Inicialmente realizamos uma equalização local de histograma [8] a partir de blocos de 8×8 pixels. A Figura 4 mostra: (a) uma imagem original I , obtida por meio do dispositivo de captura sem contato aqui estudado⁴; e (b) o resultado da equalização local de histograma, I_h . É importante ressaltar que a imagem da Figura 4(a) é uma imagem obtida a partir da costura das três vistas fornecidas pelo dispositivo, ou seja, representa uma captura unha-a-unha.

³<http://www.tbsinc.com/>

⁴TBSGuard 3D.

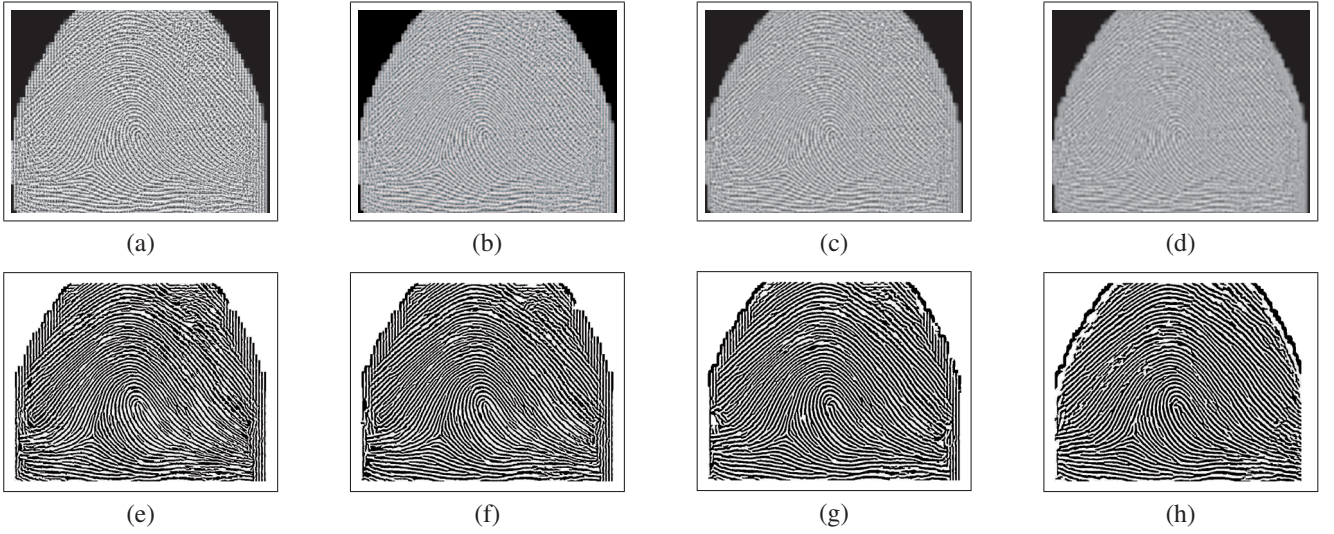


Fig. 5. Imagens I_{fi} resultantes do processo de filtragem: (a) a (d). Imagens I_{bi} resultantes do processo binarização: (e) a (h). Na filtragem foram utilizados filtros gaussianos de tamanho $h_i = [3 \ 5 \ 7 \ 9]$ e respectivos desvios padrões dados pela Equação 2. A binarização foi realizada utilizando-se o software *mindtct*.

O segundo estágio do algoritmo proposto consiste em submeter a imagem I_h a uma transformação gama [8], definida pela equação:

$$I_{h\gamma} = c \cdot I_h^\gamma \quad (1)$$

onde c e γ são constantes positivas. No presente artigo consideramos $c = 1$ e $\gamma = 1.5$. O resultado da transformação em níveis de cinza sugerida pode ser visualizado na Figura 4(c).

O passo seguinte consiste em se filtrar a imagem $I_{h\gamma}$ com 4 filtros gaussianos de tamanhos $h_i = [3 \ 5 \ 7 \ 9]$. Considerando que esses tamanhos de filtro representam as larguras a meia altura (*full width at half maximum* ou FWHM) de Gaussianas [9], os respectivos desvios padrões σ_i são definidos a partir da equação,

$$\sigma_i = \frac{h_i}{2\sqrt{2 \ln 2}}. \quad (2)$$

As 4 novas imagens, I_{fi} , $i = 1..4$, resultantes do processo de filtragem anteriormente citado, são então binarizadas por meio do software *mindtct* oferecido pelo NIST (*National Institute of Standards and Technology*) no pacote *NIST Biometric Image Software* (NBIS) [10]. O resultado são as imagens binárias I_{bi} , $i = 1..4$. As imagens filtradas e binarizadas podem ser visualizadas na Figura 5.

Gera-se, então, uma imagem I_M a partir da média das imagens I_{bi} , conforme a Equação 3, e realiza-se um recorte exatamente nos limites da impressão digital. O resultado pode ser visualizado na Figura 6. Pode-se observar que o aspecto de I_M começa a se aproximar ao de uma imagem adquirida por meio de um dispositivo com contato. No entanto, a imagem apresenta-se afunilada em sua parte superior. Geralmente, em imagens adquiridas por meio de rolagem unha-a-unha, a diferença entre a largura da parte superior e a largura da parte inferior não é tão acentuada. Por isso, faz-se necessário um

processamento cujo resultado é o alargamento da parte superior da impressão digital. O algoritmo proposto procura estimar por meio de uma regressão polinomial qual seria a função que aproxima as bordas direita e esquerda da impressão digital. Estamos assumindo que a distorção observada é simétrica. Uma vez encontrado o polinômio $p_1(x)$ que descreve a borda esquerda, por exemplo, assumimos que esse mesmo polinômio pode ser utilizado como referência para compensar também a borda direita, a partir de um polinômio simétrico $p_2(x)$. A correção geométrica proposta nada mais faz do que interpolar cada linha da impressão digital nas direções indicadas pela ilustração da Figura 7 (a). Os eixos foram rotacionados para acompanharem a orientação de captura da imagem. A Figura 7 (b) mostra a imagem I_D resultante da correção geométrica. Utilizaram-se polinômios de grau 5.

$$I_M = \frac{1}{4} \sum_{i=1}^4 I_{bi} \quad (3)$$

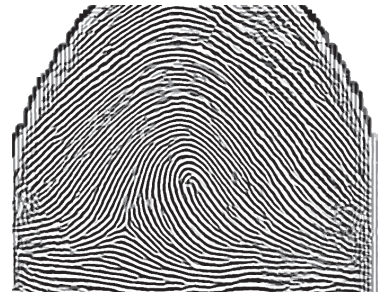


Fig. 6. Imagem I_M , gerada a partir da média das imagens I_{bi} .

A penúltima etapa de processamento consiste em se aplicar uma textura sobre a imagem I_D . Para isso, projetam-se sinteticamente [11] elipses sobre I_D com diagonais maiores (entre 1 e 5 *pixels*), diagonais menores (entre 1 e 3 *pixels*), ângulos

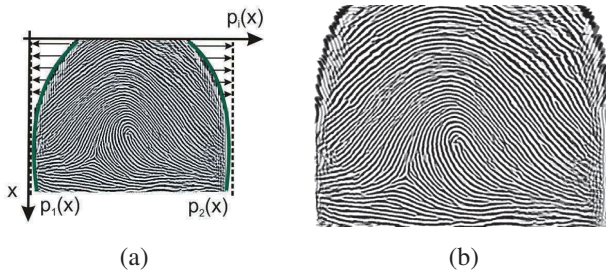


Fig. 7. Correção geométrica proposta: (a) cada linha da impressão digital é interpolada nas direções indicadas. O polinômio $p_1(x)$ é obtido a partir de regressão polinomial e descreve a borda esquerda da impressão digital. O polinômio $p_2(x)$ é obtido a partir da reflexão e deslocamento de $p_1(x)$; e (b) resultado da correção geométrica aplicada à imagem da Figura 6. Utilizaram-se polinômios de grau 5.

(entre 0 e 180°) e níveis de cinza (entre 0 e 255) gerados aleatoriamente. O resultado é o observado na Figura 8 (a). Com o objetivo de se descartar as margens da impressão digital por apresentarem ruídos causados pela baixa frequência de franjas, realiza-se um esmaecimento nas bordas da imagem. A Figura 8 (b) mostra o resultado do esmaecimento que, também, representa o passo final do método proposto.



Fig. 8. Aplicação de textura (a) e esmaecimento das bordas (b).

III. RESULTADOS EXPERIMENTAIS

Nosso conjunto de testes é formado por 200 imagens adquiridas a partir de um dispositivo sem contato multivista. As impressões digitais foram capturadas da forma mais natural possível, simulando um processo de aquisição real. Em outras palavras, procurou-se minimizar ao máximo a influência de um operador. O procedimento de captura foi descrito ao usuário e o operador apenas interfeuiu indicando ao usuário o momento em que o dedo deveria ser posicionado no dispositivo. Além disso, não foram descartadas quaisquer imagens, mesmo aquelas que pudessem ser classificadas como sendo de baixa qualidade a partir de uma avaliação prévia do operador. As imagens representam 100 dedos diferentes, capturados duas vezes cada um. Todas as impressões digitais foram processadas pelo algoritmo descrito na Seção II. Em seguida, computou-se para cada impressão digital um índice de qualidade a partir do *software* NFIQ (NIST Fingerprint Image Quality) [12]. Trata-se de um *software* baseado em redes neurais artificiais [13] que retorna um dos seguintes índices de qualidade: (1) Excelente; (2) Muito bom; (3) Bom; (4) Regular; e (5) Ruim. Detalhes a respeito de como o NFIQ estima a qualidade de imagens de impressões digitais podem ser consultados no documento

proposto por Tabassi e colaboradores [12]. Porém, de uma maneira geral, o estimador baseia-se em: (a) um mapa de qualidade onde são integradas informações de baixo contraste, baixo fluxo de franjas, alta curvatura; (b) área da impressão digital; (c) no número total de minúcias; (d) número de minúcias com qualidades superiores a 0.5, 0.6, 0.75, 0.8 e 0.9. A qualidade de uma minúcia é calculada a partir de dois fatores: (1) sua posição no mapa de qualidade; e (2) média e desvio padrão dos *pixels* que constituem sua vizinhança imediata.

A Figura 9 (a) mostra os resultados gerados pelo NFIQ para o conjunto de testes considerado nesse trabalho. Aproximadamente 90% das imagens receberam índices bom, muito bom ou excelente.

Em um segundo experimento, submetemos as impressões digitais processadas ao *software* de casamento *bozorth3* [10], com o objetivo de determinarmos a *equal error rate* (EER). Para avaliarmos a taxa de falsa rejeição, casamos uma das impressões digitais de cada dedo com a outra impressão do mesmo dedo. A taxa de falsa aceitação foi computada casando-se as duas impressões digitais de cada dedo com todas as impressões digitais dos demais dedos. Ou seja, as duas impressões digitais do dedo 1 foram casadas com todas as impressões digitais dos dedos 2 a 100 (396 casamentos). As duas impressões digitais do dedo 2 foram casadas com todas as impressões digitais dos dedos 3 a 100 (392 casamentos). E assim por diante, até que as duas impressões digitais do dedo 99 foram casadas com todas as impressões digitais do dedo 100 (4 casamentos). Foram no total 100 casamentos para se determinar a taxa de falsa rejeição e 19800 casamentos para se determinar a taxa de falsa aceitação. A definição dessas taxas depende também de um limiar estabelecido para o índice de similaridade retornado pelo *bozorth3*. A Figura 9 (b) mostra as curvas de falsa aceitação e falsa rejeição para limiares entre 0 e 150. A EER ficou por volta de 4% para o limiar 14.

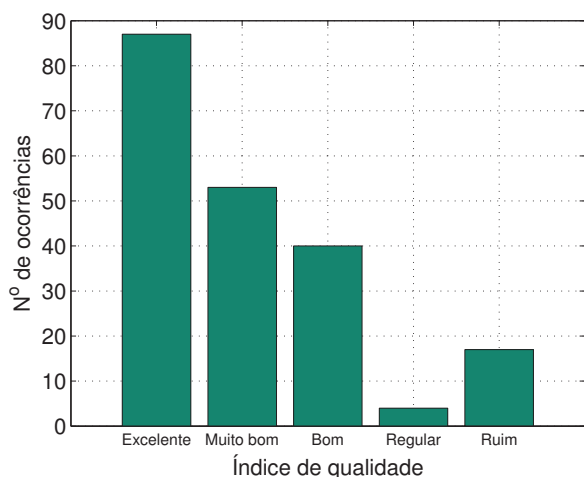
IV. CONCLUSÕES

Foi proposto um método para processamento de impressões digitais adquiridas com dispositivos de captura multivista sem contato, com o objetivo de torná-las semelhantes às adquiridas com dispositivos onde o contato é necessário. Com isso, procurou-se compatibilizar as duas tecnologias.

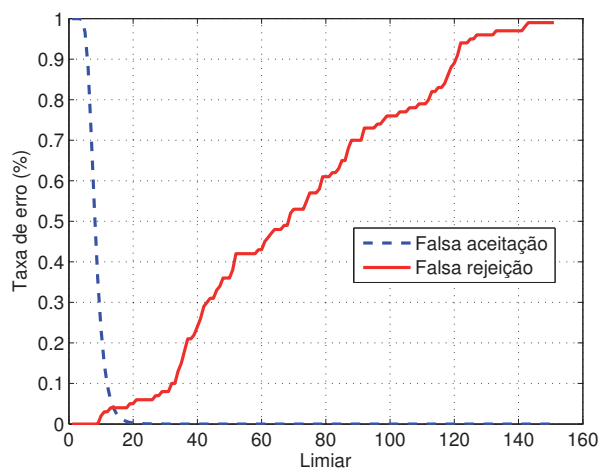
As 200 imagens adquiridas por meio de um dispositivo que dispensa o contato passaram por duas etapas principais de processamento. A primeira procura reproduzir a textura de impressões digitais tradicionais, e a segunda procura fazer refletir o processo de rolagem unha-a-unha. Dois experimentos foram propostos, um para avaliar a qualidade das imagens processadas, e outro para estabelecer a EER. Observou-se que 90% das impressões digitais apresentaram índices bom, muito bom ou excelente, e a ERR ficou em 4%, demonstrando a viabilidade do método proposto.

REFERENCES

- [1] J. Wayman, "A definition of biometrics," in *National Biometric Test Center Collected Works 1997-2000*. San Jose State University, 2000.



(a)



(b)

Fig. 9. (a) Índices de qualidade gerados pelo NFIQ para as 200 imagens do conjunto de testes: 90% das imagens receberam índices bom, muito bom ou excelente; (b) taxas de falsa aceitação e falsa rejeição geradas a partir do software de casamento de impressões digitais *bozorth3*: a equal error rate ficou em 4%.

- [7] D.L. Milgram, "Computer methods for creating photomosaics," *IEEE Transactions on Computers*, vol. C-24, no. 11, pp. 1113–1119, Nov. 1975.
- [8] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Prentice Hall.
- [9] Eric W. Weisstein, "Gaussian Function," <http://mathworld.wolfram.com/GaussianFunction.html>, [Online; Acessado em 21-Junho-2012].
- [10] C. Watson, M. Garriss, E. Tabassi, C. Wilson, R. M. McCabe and S. Janet, and K. Ko, *User's Guide to NIST Biometric Image Software*, NIST, 2008.
- [11] R. Cappelli, A. Erol, D. Maio, and D. Maltoni, "Synthetic fingerprint-image generation," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, 2000, vol. 3, pp. 471–474 vol.3.
- [12] E. Tabassi, C. L. Wilson, and C. I. Watson, *Fingerprint Image Quality*, NIST, 2004.
- [13] Simon Haykin, *Redes Neurais, Princípios e Prática*, Pearson Education, second edition, 1999.

- [2] R. Allen, P. Sankar, and S. Prabhakar, "Fingerprint identification technology," in *Biometric Systems: Technology, Design and Performance Evaluation*, J. Wayman, A. Jain, D. Maltoni, and D. Maio, Eds., chapter 2. Springer, London, 2005.
- [3] R.D. Labati, V. Piuri, and F. Scotti, "Neural-based quality measurement of fingerprint images in contactless biometric systems," in *Neural Networks (IJCNN), The 2010 International Joint Conference on*, July 2010, pp. 1–8.
- [4] Ruggero Donida Labati, Vincenzo Piuri, and Fabio Scotti, "A neural-based minutiae pair identification method for touch-less fingerprint images," in *Computational Intelligence in Biometrics and Identity Management (CIBIM), 2011 IEEE Workshop on*, April 2011, pp. 96–102.
- [5] V. Piuri and F. Scotti, "Fingerprint biometrics via low-cost sensors and webcams," in *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, 29 Oct–1 Nov 2008, pp. 1–6.
- [6] G. Parziale, "Touchless fingerprinting technology," in *Advances in Biometrics: Sensors, Algorithms and Systems*, Nalini K. Ratha and Venu Govindaraju, Eds., chapter 2. Springer, London, 2008.

Estudo de Performance de algoritmos de Verificação de Impressões Digitais Aplicáveis a *Smart-Card*

Ramysés de Macedo Rodrigues¹, Rafael Oliveira Ribeiro²

¹ Departamento de Polícia Federal, Setor Técnico-Científico da Superintendência Regional do DPF no Piauí
Teresina/Piauí, Brasil

² Departamento de Polícia Federal, Instituto Nacional de Criminalística,
Brasília/Distrito Federal, Brasil

ramyses.rmr@dpf.gov.br, rafael.ror@dpf.gov.br

RESUMO – Este artigo apresenta um estudo comparativo de algoritmos de verificação de impressões digitais propostos na literatura e aplicáveis a sistemas *Match-on-Card* (MOC). Serão avaliadas as contribuições relativas de dois desses algoritmos ante diferentes combinações de descritores e formatos de *templates* conforme as definições da ISO 19.749-2, no formato compacto e normal, propostas na literatura, selecionados levando-se em consideração a sua utilização em dispositivos de memória e capacidade de processamento limitados. Os resultados demonstram que é recomendável manter os descritores padrão no formato compacto especificados pela ISO/IEC 19.749-2, mesmo não se utilizando todos os seus elementos.

1. INTRODUÇÃO

A necessidade de se implementar formas seguras de identificação de pessoas faz o reconhecimento automático de características biométricas ser objeto de constantes estudos. As impressões digitais, nesse contexto, são as mais utilizadas, tendo em vista sua facilidade de aquisição e processamento quando comparada a outros identificadores, como íris ou DNA [1].

As impressões digitais já são, há muito, utilizadas para identificação de pessoas [1]. Em grandes bases de dados, como cadastros de identidade civil, elas são o elemento chave na individualização de um registro.

A Tabela 1, extraída de [1], relaciona de forma resumida as propriedades de diversos identificadores biométricos, com destaque para a impressão digital.

Pode-se verificar que esse identificador biométrico oferece média facilidade de aquisição (coletabilidade) com alta confiabilidade na

individualização de uma pessoa (universalidade, unicidade, permanência).

TABELA 1
PROPRIEDADES DE IDENTIFICADORES BIOMÉTRICOS

Identificador Biométrico	1	2	3	4	5	6	7
DNA	H	H	H	L	H	L	L
Iris	H	H	H	M	H	L	L
Retina	H	H	M	L	H	L	L
Orelha	M	M	H	M	M	H	M
Face	H	L	M	H	L	H	H
Impressão digital	M	H	H	M	H	M	M
Geometria da mão	M	M	M	H	M	M	M
Escrita	L	L	L	H	L	H	H
Voz	M	L	L	M	L	H	H

1: Universalidade; 2: Unicidade; 3: Permanência; 4: Coletabilidade; 5: Desempenho; 6: Aceitabilidade; 7: Suscetibilidade; L: Baixo (Low); M: Médio (medium); H: Alto (high)

O aumento da capacidade de processamento e de armazenamento de *chips* que podem ser embarcados em cartões de identificação, *tokens* USB, passaportes eletrônicos, entre outros, torna possível que a comparação biométrica seja realizada dentro do próprio *smart-card*.

A principal vantagem desta abordagem diz respeito ao aumento da segurança dos dados biométricos e da privacidade do usuário, uma vez que a biometria de referência, armazenada no *chip*, não é acessada por sistemas externos, sendo necessário apenas que o *smart-card* forneça uma resposta do tipo positivo/negativo ou um grau de similaridade entre a biometria de referência e aquela submetida à comparação.

O conteúdo deste artigo é organizado da seguinte forma: na seção 2 serão discutidos os conceitos e estrutura dos sistemas de verificação *Match-on-Card*. Na seção 3 serão demonstrados os formatos de armazenamento de minúcias de acordo com a norma ISO/IEC 19749-2. Na seção 4 serão descritos algoritmos comparadores compatíveis com esses sistemas e encontrados na literatura. Na seção 5 serão demonstrados os resultados de testes realizados com dois comparadores com diferentes descritores e formatos de *templates*. Finalmente, nas seções 6 e 7 serão expostas as conclusões e trabalhos futuros.

2. SISTEMAS MATCH-ON-CARD DE VERIFICAÇÃO BIOMÉTRICA

De acordo com [1] e [2], um sistema de verificação de impressões digitais convencional pode ser descrito pelo diagrama em blocos apresentado na Figura 1.

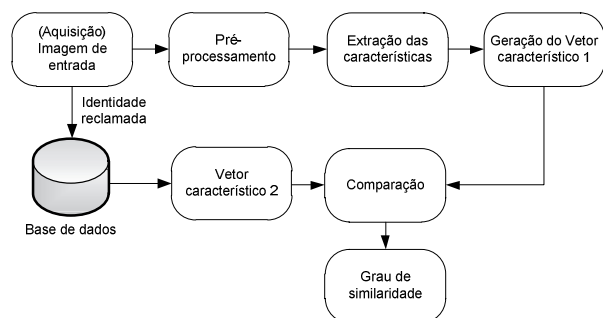


Figura 1: Sistema de verificação de impressões digitais.

A impressão digital é introduzida no sistema através dispositivos ópticos. Em seguida, ela é realçada para ter extraídos os seus elementos individualizadores, sendo as minúcias (Figura 2) e os pontos singulares os elementos mais comuns [1] – [3] e [6]. As primeiras podem ser divididas, para fins práticos, em terminações e bifurcações. Já os pontos singulares, em núcleo e delta.

As minúcias são descritas basicamente por quatro elementos: coordenadas x e y (em relação aos eixos coordenados da imagem), direção (em relação ao eixo horizontal) e tipo (terminação ou bifurcação). O conjunto de N minúcias detectadas na impressão digital será representado

pelo vetor básico, ou *template*, [1] e [6] $[x, y, \text{ângulo}, \text{tipo}]_i$, sendo $1 \leq i \leq N$. Uma imagem de impressão digital plana de boa qualidade geralmente possui entre 40 e 80 minúcias [1].

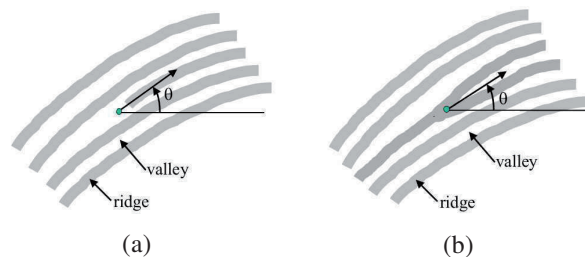


Figura 2: Minúcias e ângulos: (a) Terminações (b) Bifurcações.

O uso desses descritores, no entanto, não permite uma comparação direta, ponto a ponto, entre dois distintos *templates*, tendo em vista as inevitáveis distorções introduzidas no processo de aquisição, como rotações, translações e outras distorções não lineares.

Por isso, diversos algoritmos de verificação de impressões digitais transformam a representação básica em outra que seja imune à rotação e ao deslocamento, gerando um vetor característico [1] – [3], [5] – [8], [9], [14] – [15]. Isso pode ser feito definindo-se referências no conjunto *template*, em relação às quais todas as minúcias serão representadas.

Os algoritmos de comparação responderão à solicitação de verificação de dois *templates* através de um grau de similaridade, geralmente variando no intervalo $[0,1]$, o qual está relacionado com a probabilidade dessas duas impressões digitais serem provenientes de um mesmo dedo [1], [10].

Para plataformas computacionais baseadas em PC, o tamanho dessa representação não é uma preocupação imediata do projetista, tendo em vista a disponibilidade de recursos de memória e processamento.

Nos sistemas *Match-on-Card* (MOC), por outro lado, todo o processo de verificação é realizado no interior do *chip*, instalado, em geral, no corpo de um cartão plástico [7], [9] e [11]. O *template* que representa a impressão digital do proprietário do cartão de identidade é

armazenado em regiões de memória protegidas contra acessos não autorizados. O *template* do portador do cartão é extraído da imagem adquirida por sensores em geral ópticos ou capacitivos. Um sistema de autenticação biométrico baseado nesse conceito tem a estrutura mostrada na Figura 3.

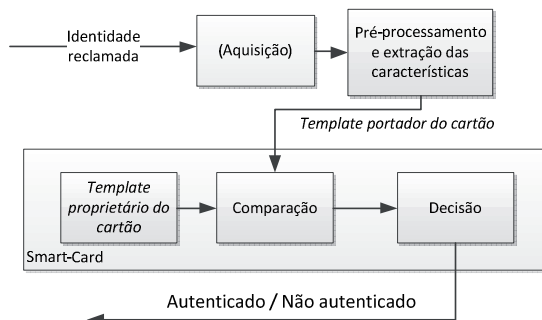


Figura 3: Sistema *Match on Card* de verificação de impressões digitais.

A aquisição, extração e geração de *templates* são implementadas em dispositivos de hardware de maior poder de processamento do que os *chips* usualmente empregados nos *smart-cards* [4], [7], [12] e [13].

O desempenho dos algoritmos de comparação aplicáveis a sistemas MOC tem melhorado à medida que os *smart-cards* têm aumentado sua capacidade de processamento e memória, assemelhando-se àqueles projetados para arquiteturas baseadas em PC [14].

3. ISO/IEC 19.749-2 – FORMATO PADRÃO DE INTERCÂMBIO DE INFORMAÇÕES BIOMÉTRICAS: IMPRESSÕES DIGITAIS

A norma ISO/IEC 19.749 define o formato de armazenamento de diversas características biométricas de modo a permitir a interoperabilidade de sistemas baseados em biometria. A parte 2 dessa norma é especializada em impressões digitais [6], estabelecendo, entre outros aspectos, o formato de armazenamento de suas minúcias.

O formato é composto por um cabeçalho geral (*Record Header* – com 24 bytes) e um bloco de informações – *Single Finger Record* – para cada dedo do indivíduo cadastrado. Em cada *Single Finger Record* são armazenados, além do seu

próprio cabeçalho, os descritores das minúcias (tipo, coordenadas, ângulo e qualidade), além de outros dados em um bloco denominado *extended data* [5], como coordenadas do núcleo e delta, quantidade de linhas entre minúcias, etc. São também definidos dois formatos de armazenamento: normal e compacto.

Para o formato normal, as informações relativas a cada minúcia são armazenadas em blocos de 6 bytes, com o alinhamento mostrado na Figura 4.

tipo	x	reservado	y	ângulo	qualidade
2 bits	14 bits	2 bits	14 bits	[0~360]	[0~100]
2 bytes		2 bytes		1 byte	1 byte

Figura 4: Formato das minúcias representadas no modo normal [15].

O formato compacto, por outro lado, utiliza apenas 3 bytes por minúcia, com alinhamento mostrado na Figura 5.

x	y	tipo	ângulo
1 byte	1 byte	2 bits	6 bits

Figura 5: Formato das minúcias representadas no modo compacto [15].

O formato normal representa uma coordenada espacial utilizando 14 bits (2^{14} níveis de quantização) e ângulos com passos de $360 / 2^8$ graus. Se uma impressão digital tiver 40 minúcias, serão necessários $40 * 6 = 240$ bytes de memória apenas para armazenar as minúcias.

No modo compacto, as coordenadas serão representadas em $2^8 = 256$ níveis. Os ângulos em passos de $360 / 2^6$ graus. Em termos de utilização de memória, o espaço ocupado pelas mesmas 40 minúcias será reduzido para $40 * 3 = 120$ bytes. Um ganho considerável se observarmos que trabalhamos com limitações de memória.

A desvantagem óbvia da utilização do modo compacto é a perda de resolução dos descritores da minúcia.

O MINEXII (*Minutiae Exchange Interoperability Test*) [13] comparou algoritmos de verificação de impressões digitais, fornecidos por fabricantes diversos, utilizando os dois formatos de armazenamento, concluindo que a

perda de desempenho global é aceitável para o formato compacto e, portanto, adequado para operações em sistemas *Match-On-Card*.

4. ALGORITMOS DE COMPARAÇÃO DE IMPRESSÕES DIGITAIS

O reconhecimento de impressões digitais através das minúcias é a metodologia mais popular entre os algoritmos comparadores. Aliado ao fato de ser semelhante à técnica utilizada pelo comparador humano, o trabalho com essas características transforma o processo de verificação de duas impressões digitais em um problema de combinação de padrões de pontos, ou *point pattern matching* [1] e [7].

Como discorrido na seção 2, as distorções introduzidas na aquisição não permitem a comparação ponto a ponto entre dois *templates* quando representadas por seus descritores básicos [x, y, ângulo, tipo].

As distorções do tipo rotação e translação tem relação direta com a área da superfície do sensor e o posicionamento do dedo do usuário durante a aquisição. As distorções não lineares são causadas principalmente pela transformação do espaço 3D do dedo do usuário para o espaço 2D da superfície do sensor da imagem [1]. Além disso, os erros de detecção de minúcias são consequências, em geral, da maior ou menor pressão com que o usuário coloca seu dedo no sensor [2].

Desse modo, um alinhamento entre os dois conjuntos de minúcias, ou a transformação de sua representação para outra imune à rotação e translação (vetor característico), é um passo mandatório [1].

Após o alinhamento, duas minúcias p e q , pertencentes a *templates* distintos e representadas por seus descritores básicos serão consideradas similares se satisfizerem, ao mesmo tempo, as seguintes equações [1]:

$$\text{Distância_Euclidiana}(p, q) \leq t_0 \quad (1)$$

$$\text{Diferença_angular}(p, q) \leq t_1 \quad (2)$$

$$\text{tipo}(p) = \text{tipo}(q) \quad (3)$$

Onde t_0 e t_1 são tolerâncias, utilizadas para considerar as distorções não lineares e podem ser determinadas dinamicamente em função da distância considerada, conforme [5].

Essa é a proposta do algoritmo [15] o qual foi desenvolvido para ser utilizado em *smart card's*. Ele aplica sucessivamente a função de transformação dada pela Equação 4 para alinhar os dois conjuntos de minúcias. Observa-se que rotacionar ou transladar um conjunto de minúcias não altera o descritor de tipo.

$$F_{\theta, \Delta x, \Delta y} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \quad (4)$$

Os algoritmos apresentados em [4] e [14] utilizam as coordenadas do núcleo como referência de um sistema de coordenadas polares, em relação ao qual todas as minúcias são codificadas, gerando o vetor característico (Figura 6). A nova representação, desse modo, absorve automaticamente as distorções de rotação e translação.

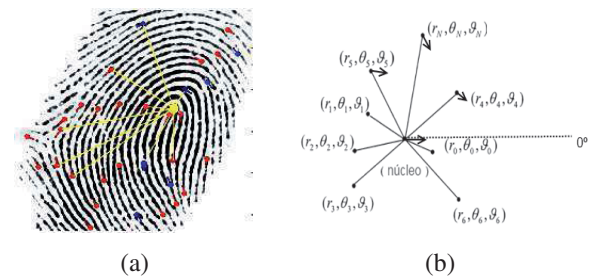


Figura 6: Codificação das minúcias e sua vizinhança segundo [14]. (a) Minúcias em relação ao núcleo. (b) Codificação da estrutura em coordenadas polares (r = distância radial; θ = distância angular; $\leq$ = diferença direcional)

Em [5], o algoritmo toma cada uma das minúcias do *template* e a relaciona com sua vizinhança mais próxima (Figura 7). A grande vantagem dessa abordagem é a não necessidade de se detectar o núcleo da impressão digital.

Outros algoritmos, como [2] e [15], geram grafos planares através da triangulação dos descritores espaciais das minúcias. Em [2], conforme se observa na Figura 8, o vetor característico é criado relacionando-se cada minúcia com sua vizinhança estabelecida pelas

arestas do grafo. Em [15], cada triângulo gerado é codificado em função de seus lados, ângulos internos e externos, e outras características associadas. O trabalho de [15] implementou esse algoritmo em um *smart card*.

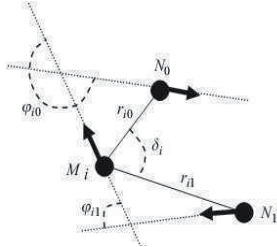


Figura 7: Codificação da minúcia M_i e suas duas vizinhas mais próximas, segundo [5].

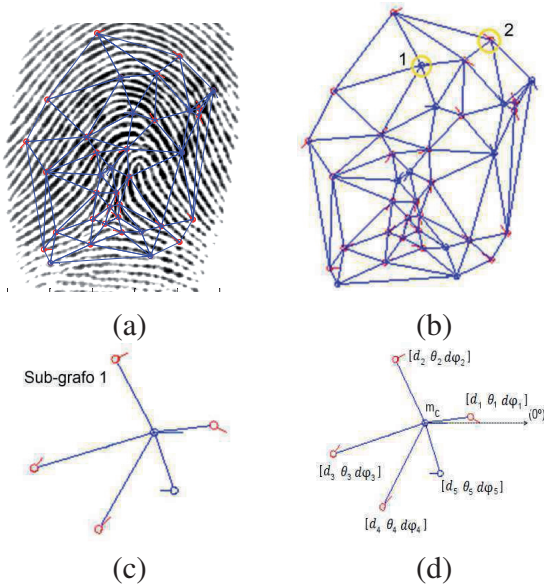


Figura 8: Codificação das minúcias e sua vizinhança segundo [2]. (a) Triangulação das minúcias. (b) e (c) identificação da vizinhança gerada pelo grafo e estrutura formada (sub-grafo). (d) Codificação da estrutura em coordenadas polares (d = distância radial; θ = distância angular; $d\phi$ = diferença direcional)

BOZORTH3 [16] é o nome dado ao algoritmo comparador de minúcias utilizado em [17]. Nessa metodologia, todas as minúcias do conjunto são agrupadas aos pares, codificando-se as distâncias, ângulos e diferenças direcionais entre extremos, de acordo com a Figura 9.

Outros algoritmos não baseados em minúcias já estão sendo aplicados no reconhecimento de impressões digitais em sistemas *MOC*, como o proposto em [3], que se utiliza do denominado *FingerCode* para o reconhecimento rápido de impressões digitais em *smart cards*.

Verifica-se que a maioria dos algoritmos baseados em minúcias utilizam apenas os descritores relacionados às coordenadas espaciais e ao ângulo para gerar o vetor característico. Em tese, o descritor de tipo aumenta o grau de individualização de uma minúcia e ocupa, nos formatos normal e compacto, apenas dois bits por elemento no *template*.

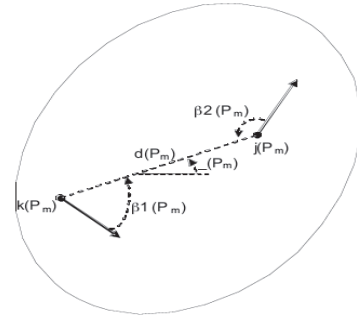


Figura 9: Codificação de pares de minúcias de acordo com [16].

A principal razão para o não uso desse descritor é sua sensibilidade aos erros de aquisição, onde, por exemplo, uma terminação pode ser computada como bifurcação se o usuário pressionar demasiadamente seu dedo no sensor [1] e [5].

Se comparada à economia de memória proporcionada pela utilização do formato compacto da ISO/IEC 19.794-2, a não utilização desse descritor geraria uma economia de apenas dois bits por minúcia. No exemplo da seção 3, em conjunto de 40 pontos a economia em espaço de memória seria de $40 * 2 \text{ bits} = 80 \text{ bits}$, ou 10 bytes.

Seria então adequado utilizar esses dois bits para aumentar a resolução do descritor de ângulo no modo compacto, o qual é representado por apenas 6 bits? Qual seria a perda de performance dos algoritmos comparadores quando utilizamos ou desprezamos esse descritor? É o que se pretende responder na próxima seção.

5. RESULTADOS EXPERIMENTAIS

Para se conferir a performance dos algoritmos ante a utilização dos formatos de armazenamento normal e compacto, com e sem a utilização de

todos os seus descritores, foram selecionados os algoritmos de comparação propostos em [5] e [2], implementados em MATLAB, e a base de dados DB1 da FVC2002 [10], disponibilizada em [1]. A seleção desses algoritmos se deve principalmente ao esforço de implementação já desenvolvido em [2]. Os valores de t_0 e t_1 são selecionados adaptativamente em função da distância, conforme a metodologia apresentada em [5]. Na prática, esses valores permanecem no intervalo de 10 a 25 *pixels* para t_0 e de 10 a 30 graus para t_1 .

A base de dados contém oito amostras de cem indivíduos, totalizando oitocentas imagens de impressão digital. Os usuários foram orientados a colocar seus dedos no sensor com maior ou menor pressão, sem preocupação com rotações ou translações. Isso produziu uma base de dados com variabilidade intraclasse (variações entre amostras de um mesmo dedo) bastante compatível com o mundo real [10].

Foram utilizadas como figuras de mérito a AUC (Area Under Curve – Área sobre a curva) de uma curva ROC (Receiver Operating Characteristic) (Figura 10), com eixo das abscissas representando o FAR (False Acceptance Rate) e o eixo ordenado o CAR (Correct Acceptance Rate). Essa curva aponta a precisão do comparador, ou seja, a probabilidade do sistema classificar corretamente um dado padrão de entrada. O protocolo de avaliação também foi baseado em [10].

Ambos os algoritmos [5] e [2] geram vetores característicos relacionando cada minúcia do *template* à suas vizinhas. O primeiro, às mais próximas. O segundo, àquela estabelecida por uma triangulação.

No primeiro teste, os vetores característicos dos dois algoritmos foram originados dos descritores [x, y, ângulo]. No segundo, dos descritores [x, y, ângulo, tipo] e no terceiro, de uma combinação de [x, y, ângulo] para a minúcia central e [x, y, ângulo, tipo] para a vizinhança, todos esses descritores representados no modo normal da ISO 19.794-2. As curvas ROC exibidas nas Figuras 10 a 13 mostram os

resultados dos valores do parâmetro AUC para diversas configurações de descritores e formatos de *template*.

Em seguida, foram repetidas todas as configurações de descritores dos testes anteriores, utilizando-se desta vez *templates* no modo compacto, obtidos pela requantização dos descritores anteriores. Os parâmetros x e y passam a ser representados com 8 bits, o ângulo com 6 bits e o tipo com 2 bits. A Figura 13 mostra o detalhe da curva ROC resultante quando utilizamos os descritores [x, y, ângulo] no modo compacto e a Figura 14 mostra a curva quando são utilizados os descritores [x, y, ângulo], sendo o ângulo representado dessa vez por 8 bits (adicionando os dois bits do descritor de tipo ao descritor de ângulo: modo compacto modificado).

As Tabelas 2 e 3 listam os resultados gerais obtidos.

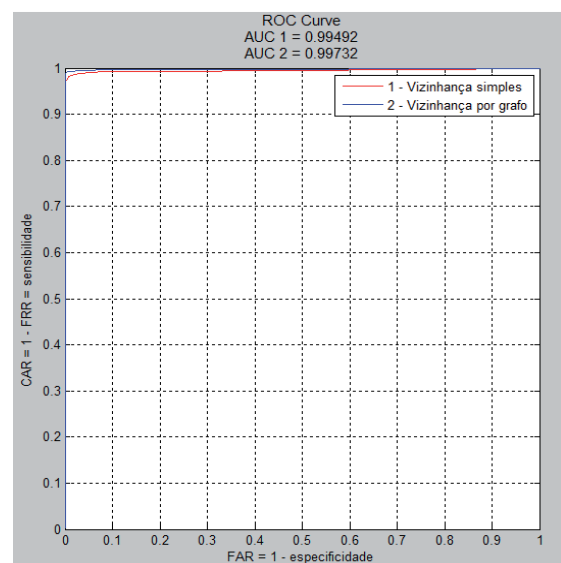


Figura 10: Curva ROC para os algoritmos de comparação por vizinhança simples [5] e vizinhança por grafo [2], com vetor característico originado dos descritores [x, y, ângulo], utilizando-se seis bytes por minúcia (modo normal).

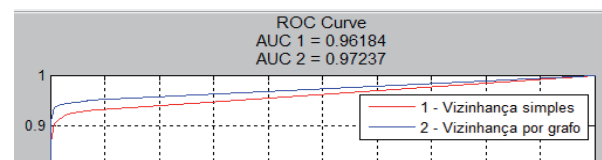


Figura 11: Detalhe da curva ROC para os algoritmos de comparação por vizinhança simples [5] e vizinhança por grafo [2], com vetor característico originado dos descritores [x, y, ângulo, tipo], utilizando-se seis bytes por minúcia (modo normal).

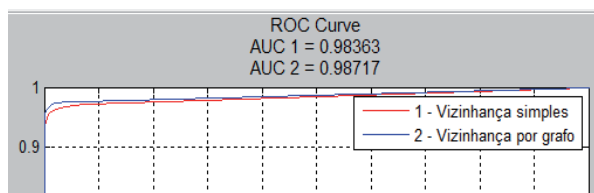


Figura 12: Detalhe da curva ROC para os algoritmos de comparação por vizinhança simples [5] e vizinhança por grafo [2], com vetor característico originado dos descritores [x, y, ângulo] para a minúcia central e [x, y, ângulo, tipo] para as vizinhas, utilizando-se seis bytes por minúcia (modo normal).

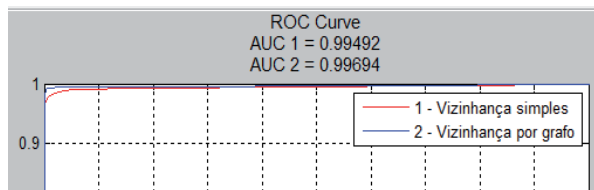


Figura 13: Detalhe da curva ROC para os algoritmos de comparação por vizinhança simples [5] e vizinhança por grafo [2], com vetor característico originado dos descritores [x, y, ângulo], utilizando-se três bytes por minúcia (modo compacto), ângulo representado por seis bits.

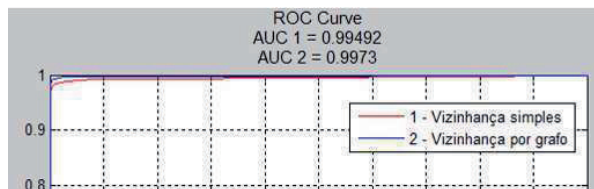


Figura 14: Detalhe da curva ROC para os algoritmos de comparação por vizinhança simples [5] e vizinhança por grafo [2], com vetor característico originado dos descritores [x, y, ângulo], utilizando-se três bytes por minúcia e ângulo representado por oito bits (modo compacto modificado).

TABELA 2

RESULTADOS GERAIS DOS TESTES PARA O FORMATO NORMAL

Configuração/ algoritmo	AUC	Perda de desempenho
A-1	0,99492	0,00%
A-2	0,99732	0,00%
B-1	0,96184	-3,32%
B-2	0,97327	-3,56%
C-1	0,98363	-1,13%
C-2	0,98717	-1,37%

A: algoritmo utiliza os descritores [x, y, ângulo]

B: algoritmo utiliza os descritores [x, y, ângulo, tipo]

C: algoritmo utiliza a combinação dos descritores [x, y, ângulo] e [x, y, ângulo, tipo]

1: algoritmo de vizinhança simples [5]

2: algoritmo de vizinhança por grafo [2]

6. CONCLUSÕES

A representação do *template* no modo compacto insere nos algoritmos comparadores avaliados perdas de performance compatíveis

com os resultados apresentados no MINEXII [13], sendo essa queda proporcionada, principalmente, pela menor resolução do descritor de ângulo (que passa a ser representado por apenas 6 bits). O ganho em espaço de memória com esse formato compensa as perdas em precisão [13], uma vez que viabiliza a utilização desses algoritmos em sistemas de baixa capacidade de memória.

TABELA 3

RESULTADOS GERAIS DOS TESTES PARA O FORMATO COMPACTO

Configuração/ algoritmo	AUC	Perda de desempenho
A-1	0,99492	0,00%
A-2	0,99694	0,00%
B-1	0,9614	-3,37%
B-2	0,97	-3,56%
C-1	0,9835	-1,15%
C-2	0,9862	-1,35%

TABELA 4

COMPARAÇÃO DOS RESULTADOS GERAIS DOS TESTES PARA O FORMATO NORMAL E COMPACTO

Configuração/ algoritmo	Perda de desempenho Normal/Compacto
A-1	0,00%
A-2	-0,04%
B-1	-0,05%
B-2	-0,34%
C-1	-0,01%
C-2	-0,10%

O emprego do descritor de tipo nos vetores característicos, nos formatos normal ou compacto, degrada o desempenho dos algoritmos de comparação. A razão para isso é a alta sensibilidade desse parâmetro aos erros de detecção de minúcias.

Desprezar esse descritor proporciona baixa economia de memória (apenas dois bits por minúcia). Adicionar os seus dois bits ao descritor de ângulo, no modo compacto, fazendo aumentar a sua resolução (que passa agora a ser representado por oito bits), aproxima o desempenho dos algoritmos de comparação

àqueles que se utilizam do modo normal e descritores [x, y, ângulo].

Entretanto, considerando a pequena diferença de desempenho entre os modos normal e compacto e os constantes estudos e possível desenvolvimento de algoritmos detectores de minúcias menos suscetíveis a erros, a manutenção do descritor de tipo no *template* poderá vir a ser útil no futuro.

7. TRABALHOS FUTUROS

Os testes experimentais realizados na seção 5 deverão ser repetidos com os algoritmos implementados em *smart-cards* a fim de se avaliar a compatibilidade dos resultados obtidos, medindo-se inclusive o tempo de resposta às solicitações de verificação.

8. AGRADECIMENTOS

Ao SETEC/PI e ao SEPAEL pelo apoio e incentivo.

REFERÊNCIAS

- [1] D. Maltoni, D. Maio, A.K. Jain, and S. Prabhakar, *Handbook of Fingerprint Recognition*, Springer 2005.
- [2] Rodrigues, R.M., Filho, C.F.F.C. ; Costa, M.G.F; "Fingerprint Verification Using Planar Graphs"; Page(s): 87 - 92 IVMSWP Workshop, 2011 IEEE 10th, 16-17 June 2011.
- [3] Kumar P. Y. and Ganesh T. S. "Integration of Smart Card and Gabor Filter Method based Fingerprint Matching for Faster Verification", INDICON, 2005 Annual IEEE, Page(s): 526 – 529.
- [4] Mimura, M., Ishida, S.; Seto, Y., "Fingerprint Verification System on Smart Card" Consumer Electronics, 2002. ICCE. 2002 Digest of Technical Papers. International Conference on, Page(s): 182 – 183.
- [5] T. Y. Jea, V. Govindaraju, "A Minutia-based partial fingerprint recognition system", Pattern Recognition 38 (2005) pp. 1672-1684, March 2005, published by Elsevier Ltd.
- [6] ISO/IEC 19749-2 Biometric Data Interchange Formats – Part 2: Finger Minutiae Data.
- [7] Moon, Y., Ho, H., Ng, K., Wan, S., Wong, S.: Collaborative Fingerprint Authentication by Smart Card and a Trusted Host.
- [8] Sanchez-Reillo, R., Sanchez-Avila, C.: Fingerprint Verification using Smart Cards for Access Control Systems. Proc. of Int. Carnahan Conf. (2001) 250-253.
- [9] R. Sanchez-Reillo, L. Mengibar-Pozo, C. Sanchez-Avila, "Microprocessor Smart Cards with Fingerprint User Authentication". IEEE Aerospace and Electronic Systems Magazine; Vol. 18; 22-24 ; Mar 2003
- [10] FVC2002, "Fingerprint Verification Competition", 2002, <http://bias.csr.unibo.it/fvc2002/>
- [11] Hendry, M., Prideaux, J., *Smart Card Security and Applications*, Artech House, Inc., 1997.
- [12] Struif, B.; Scheuermann, D., "Smartcards with biometric user verification", Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference on, 2002, Page(s): 589 - 592 vol.2.
- [13] MINEXII, *Minutiae Exchange Interoperability Text*, <http://www.nist.gov/itl/iad/ig/minex.cfm>
- [14] K. C. Chan, Y. S. Moon, P.S. Cheng, "Fast Fingerprint Verification using Sub-regions of Fingerprint Images", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, Issue 1, pp. 95-101, Jan. 2004.
- [15] Barral, C. *Biometrics & Security: Combining Fingerprints, Smart Cards and Cryptography*, Tese de Doutorado, 2010, Ecole Polytechnique Fédérale de Lausanne.
- [16] User's Guide to NBIS-EC – Nist Biometric Image Software (http://www.nist.gov/customcf/get_pdf.cfm?pub_id=51096)
- [17] NIST/NBIS, <http://www.nist.gov/itl/iad/ig/nigos.cfm>.

Avaliação de técnicas de seleção de quadros-chave na recuperação de informação por conteúdo visual para identificar crimes cibernéticos em conteúdo multimídia.

Shênia Salvador de Pinho¹, Kleber J. F. Souza²

Instituto de Ciências Exatas e Informática, PUC Minas
Guanhães, MG, Brasil

shenia.salvador@gmail.com, kleberjfsouza@gmail.com

RESUMO – A recuperação de informação de conteúdo visual de vídeos digitais é uma forma de gerenciamento e busca de dados multimídias. Todavia, sem utilizar técnicas de seleção de quadros-chave no processo todos os quadros do vídeo são utilizados independente de sua relevância gerando excesso de informações para armazenar. Neste trabalho é avaliado o comportamento de três métodos de seleção de quadro-chave: seleção de 1 em 1 segundo, 3 em 3 segundos e por *pixel-a-pixel* aplicado na recuperação de informação por conteúdo visual. Os resultados obtidos indicam que a relevância dos quadros-chave selecionados pode ser determinante na obtenção de melhores resultados.

1. INTRODUÇÃO

Hoje em dia com as inovações tecnológicas que baratearam os preços dos dispositivos e da comunicação de dados, houve a popularização do compartilhamento de conteúdo visual e sua disponibilização em meio digitais, como exemplo o site Youtube que possibilita seus usuários compartilhem dados multimídia.

Nestas demandas de conteúdos publicados há o risco de encontrar alguém que aproveita da velocidade e disseminação de conteúdo multimídia na Internet para cometer crimes cibernéticos, ou seja, infrações referentes ao conteúdo do vídeo, como exemplo a pornografia infantil, racismo e atentado a propriedade intelectual.

Os usuários ao publicarem seus vídeos na Internet, geralmente descrevem o assunto do vídeo através de descrição textual, ou seja, o próprio usuário retrata ao publicar seu vídeo o contexto que ele se refere. De acordo com o trabalho de Araújo et al [1], a descrição do conteúdo do vídeo feita pelo usuário não é suficiente para realizar um controle eficaz de dados multimídias publicados. O motivo é que

nem sempre o conteúdo original do vídeo condiz com o que o usuário descreveu, sendo que alguns dos conteúdos podem ser crimes cibernéticos. Por isso surge a necessidade de se ter um controle destes dados publicados [2].

Outro cenário que deve ser levado em consideração está relacionado às diretrizes de sites referentes a publicação de vídeos que contém cenas de violência, encorajam atividades perigosas e ilegais, afetam direitos autorais, contendo crianças sendo sexualmente sugestivos ou violentos entre outras [3].

Para que estes conteúdos impróprios sejam publicados apenas com o consentimento do administrador do sistema [2], é muito importante ter um gerenciamento do conteúdo visual publicado [4] e esta automação pode ser realizada através de técnicas de recuperação de informação por conteúdo visual.

Sabendo disso, o presente trabalho visa avaliar o comportamento de diferentes técnicas de seleção de quadros-chave aplicada na recuperação de informação por conteúdo visual.

Segundo Gomes [5], no processo inicial da recuperação de informação pelo conteúdo visual é realizada a coleta de todos os quadros do vídeo. Entretanto os vídeos são compostos por inúmeros quadros, sendo necessário selecionar como base para a recuperação da informação os quadros que melhor representem o vídeo [6].

A não utilização de técnicas de seleção de quadros-chave tem como consequência o

excesso de informações armazenadas, pois todos os quadros são utilizados no processo de recuperação, independentemente de serem relevantes ou não, sendo elas: quadros negros, quadros repetidos, quadros sem relevância, entre outros.

Trabalhos que apresentaram ganhos com a utilização de técnicas de seleção de quadros-chave foi realizado por Pimentel, Saibel e Buck [7] e Bueno et al [8] onde foi feito comparações entre técnicas de seleção de quadros-chave e obteve diminuição do custo computacional e armazenamento.

Outro trabalho que apresentou benefícios com a utilização de técnicas de seleção de quadros-chave foi o Nascimento, Manzato e Goularte [9] o mesmo comparou seis técnicas com diferentes intervalos de tempo de seleção de quadros-chave sendo 0.5, 1.0, 1.5, 2.0, 2.5 e 3.0 segundos, pode-se observar que a técnica três segundos proporcionou melhora no desempenho e sabendo disso faz-se necessário investigar os ganhos que se pode obter com a utilização de seleção de quadros-chave no processo de recuperação de informação pelo conteúdo visual, utilizando não só a técnica do Nascimento, Manzato e Goularte [9], mas também outras técnicas de seleção de quadros-chave.

O trabalho está organizado da seguinte forma: Primeiramente são apresentadas as técnicas de seleção de quadros-chave que foram utilizadas no trabalho. Logo após, discute o teste dos algoritmos, o experimento que foi realizado e análise dos resultados. Para finalizar a conclusão da pesquisa e as referências utilizadas.

2. TÉCNICAS DE SELEÇÃO DE QUADRO-CHAVE

A partir dos estudos realizados foi possível definir as técnicas de seleção de quadros-chave que foram implementadas e analisadas no desenvolvimento deste trabalho. As técnicas selecionadas foram: comparação de quadro pixel-a-pixel, seleção de quadros-chave a cada três segundos do vídeo e seleção de quadros-chave a cada um segundo do vídeo.

A técnica de seleção de quadros-chave por comparação pixel-a-pixel foi selecionada

devido ser uma abordagem mais simples de comparar os quadros do vídeo [7].

O algoritmo consiste em selecionar o primeiro quadro de um vídeo como quadro-chave a partir deste, calcular a distância Euclidiana entre todos os quadros seguintes. A distância entre o quadro-chave selecionado e o quadro que está sendo analisado é comparada com o limiar 80. O valor do limiar foi definido experimentalmente observando as diferenças visuais dos quadros-chave selecionados, ou seja, o que obteve entre os experimentos a menor taxa de quadros repetidos e a menor taxa de perda de quadros relevantes do vídeo. Quando o valor do limiar na comparação da técnica de pixel-a-pixel é alcançado, o quadro que está sendo comparado é selecionado como o novo quadro-chave do vídeo e a partir deste momento, ele passa a ser a referência para as próximas comparações até a seleção do próximo quadro-chave.

Devido à técnica de seleção de quadros-chave a cada três segundos do vídeo ter reduzido o custo computacional e proporcionou melhora no desempenho no trabalho de Nascimento, Manzato e Goularte [9] em comparação com seis técnicas de intervalos diferentes do trabalho do mesmo. Foi definido que seria a segunda técnica para ser utilizada no trabalho. A técnica baseia-se em selecionar um quadro-chave a cada intervalo de tempo de três em três segundos do vídeo.

Por fim, a última técnica foi selecionada devido já ter sido utilizada no trabalho de Silva [11] o trabalho pelo qual foi utilizado seu processo de recuperação da informação por conteúdo visual neste artigo. Sendo o método de seleção de quadros-chave a cada um segundo do vídeo [10]. O funcionamento da técnica consiste em selecionar um quadro-chave a cada um segundo do vídeo.

Podemos observar que as técnicas de seleção de quadro-chave a cada três segundos e um segundo do vídeo é feito a seleção dos quadros-chave de acordo com intervalos de tempo, todavia o método de seleção pixel-a-pixel foi realizado etapa de processamento de imagem nos quadros do vídeo.

Os quadros-chave selecionados em cada técnica de seleção foram utilizados no processo de recuperação de informação por conteúdo visual no trabalho de Silva [11].

O trabalho de Silva [11] consiste em realizar comparações entre as métricas de indexação e recuperação de vídeos. Além disso, mostrou a quantidade de informações para se descrever uma imagem e como estas características influenciam diretamente na recuperação de informação por conteúdo visual.

3. TESTE DOS ALGORITMOS

Primeiramente foi realizado o teste dos algoritmos, com a finalidade de avaliar o funcionamento das técnicas de seleção de quadros-chave. Para o teste foram utilizados dois vídeos do projeto Open Video¹. O Open Video é um projeto que visa à disponibilização gratuita de vídeos digitais para pesquisa de recuperação de multimídia.

Os vídeos aplicados no teste dos algoritmos são apresentados na Tabela 1. São documentários da University of Maryland do ano de 1998, e foram obtidos no projeto Open Video. Os vídeos utilizados possuem som e cor.

TABELA 1
VÍDEOS DE TESTE

	Nome do Vídeo	Duração	Total de quadros
V1	New Indians, Segment 08	00:23	707
V2	New Indians, Segment 09	00:16	503

As três técnicas implementadas foram executadas em ambos os vídeos, com o objetivo de verificar a quantidade de quadros-chave que os algoritmos selecionavam. A Tabela 2 apresenta a quantidade de quadros-chave selecionados por cada técnica em cada vídeo.

Para verificar se tais resultados estavam corretos, na técnica de seleção 3 em 3 segundos e 1 em 1 segundo foi realizado uma contagem manual da taxa de quadros por segundo do vídeo pela quantidade total dos

quadros. Sendo que, em ambos os vídeos esta taxa é de 29 quadros por segundo, ou seja, a cada 29 quadros na técnica de 1 em 1 segundo o quadro é selecionado como quadro-chave e cada 87 quadros para a técnica de 3 em 3 segundos. Na técnica pixel-a-pixel o resultado foi verificado pelas diferenças visuais entre os quadros selecionados. Interessante ressaltar que os quadros-chave selecionados pela técnica tiveram como resultado, quadros de tomadas distintas. Esta observação também foi verificada manualmente, apenas com a observação das características visuais dos quadros e tomadas.

TABELA 2
RESULTADO DO TESTE

	Pixel-a-Pixel	Seleção 3 seg.	Seleção 1 seg.
V1	5	8	24
V2	6	5	17

4. EXPERIMENTO

Depois da etapa de teste dos algoritmos, foi realizada uma análise das técnicas implementadas, utilizando a mesma abordagem de avaliação do trabalho de Silva [11]. A partir da base de quadros-chave selecionados foram analisados os acertos da busca de um objeto no vídeo. Vale ressaltar que o trabalho [11] teve como objetivo principal a análise comparativa do uso de estruturas métricas para a indexação e a recuperação de vídeos utilizando vocabulário visual. O autor comparou três estruturas: M-Tree, Slim-Tree e D-Index.

Como o objetivo deste artigo não é avaliar tais estruturas métricas, uma destas estruturas foi escolhida para realização do experimento deste trabalho: a estrutura D-Index é uma estrutura de indexação baseada em hashing. Seu objetivo é minimizar o número de acessos a disco e número de cálculos de distâncias [11].

As técnicas de seleção de quadros-chave foram implementadas em C# no ambiente de programação Visual Studio 2010. Na realização dos experimentos foi utilizado o mesmo vídeo utilizado por Silva [11]: The Big Bang Theory da CBS (Columbia Broadcasting System). Este vídeo tem duração de 19 minutos, 56 segundos e possui 28.682 quadros.

Para o processo de recuperação de

¹ Open Video é um projeto que disponibiliza gratuitamente vídeos digitais para fins de Pesquisa de recuperação de multimídia. Mais informações: <http://www.open-video.org>.

informação por conteúdo visual em vídeo, foram selecionados aleatoriamente 12 objetos do vídeo como se pode observar na Figura 1. Para fazer a busca pelo seu conteúdo visual, através de sua imagem e não de sua descrição textual,



FIGURA 1 – OBJETOS PARA CONSULTA

Inicialmente foram executadas as 3 técnicas de seleção de quadros-chave no vídeo para selecionar as imagens que mais representavam o conteúdo. Sendo no processo de seleção de 1 em 1 segundo e 3 em 3 segundos foi feito com base na taxa de 23 quadros por segundo. Na Tabela 3 é apresentada a quantidade de quadros-chave selecionados por técnica:

TABELA 3
RESULTADO DO EXPERIMENTO

Pixel-a-Pixel	Seleção 3 seg.	Seleção 1 seg.
440	415	1247

Os quadros-chave selecionados pelas técnicas foram utilizados no processo de recuperação da informação por conteúdo visual no trabalho de Silva [11]. Ou seja, foram aplicadas as 3 técnicas implementadas e o restante do processo de recuperação de informação por conteúdo visual é do trabalho [11] como se pode observar na Figura 2.

O objetivo do processo de recuperação é retornar as imagens que possuem os objetos apresentados na Figura 1 independente de suas posições.

A Figura 3 apresenta alguns exemplos de quadros que foram utilizados na base para

consulta dos objetos.

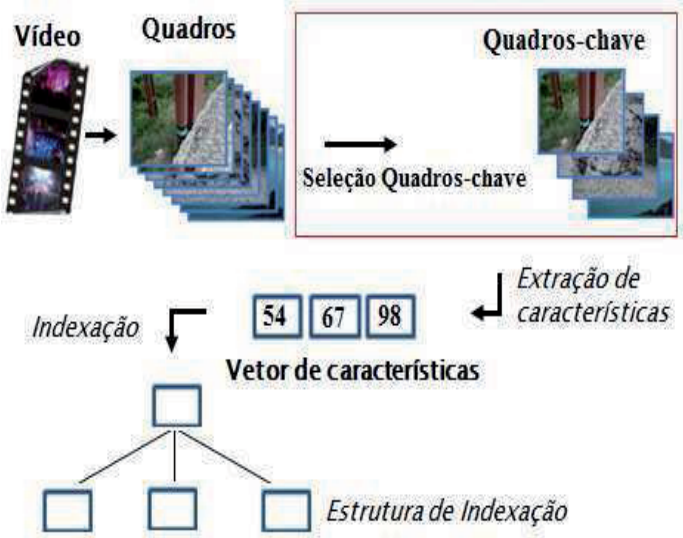


Figura 2 –Indexação das características do vídeo, Fonte: Adaptação de Silva [11]



FIGURA 3 – QUADROS PARA CONSULTA

A Figura 4 apresenta o processo de recuperação das imagens que possuem os objetos procurados no vídeo. Primeiramente, é feito um processo de extração das características da imagem do objeto que se deseja buscar no conteúdo do vídeo. Logo após, as características extraídas são armazenadas em um vetor de características que possui as informações sobre cor e textura [11]. É feita a busca na estrutura de indexação (gerada no processo de indexação das características do vídeo). Por fim, são retornadas todas as imagens que possuem características em comum com o objeto

procurado.

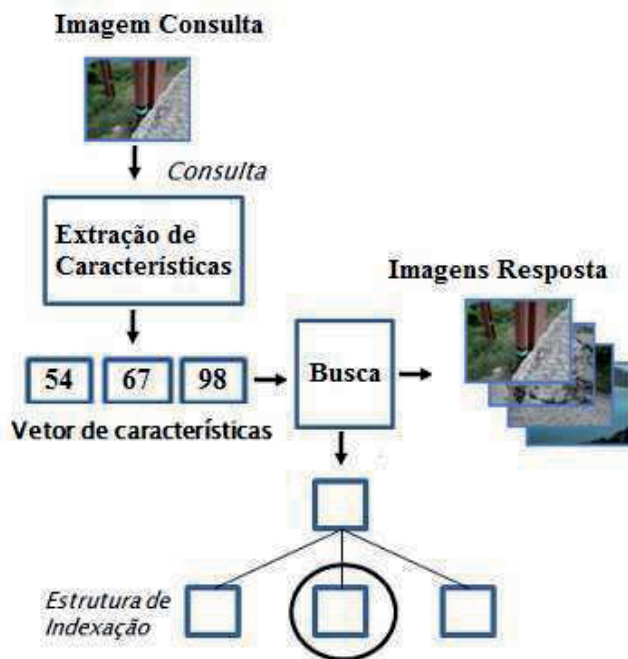


Figura 3–Recuperação de Informação por Conteúdo Visual ,
Fonte: Adaptação de Silva [11]

5. ANÁLISE E RESULTADO

Na análise dos resultados das consultas foi feito uma “rotulação dos objetos verdadeiros” para cada técnica de seleção de quadros-chave. Portanto, manualmente foi contabilizada a frequência com que os objetos selecionados para recuperação da imagem aparecem nos quadros-chave do vídeo. De acordo com o mesmo foi realizada a comparação dos resultados.

A Tabela 4 apresenta a análise quantitativa dos resultados da recuperação de informação visual da técnica de seleção de quadros-chave pixel-a-pixel.

TABELA 4
SELEÇÃO DE QUADRO-CHAVE TÉCNICA PIXEL-A-PIXEL

Consulta	Rotulação dos objetos verdadeiros	Acertos	% Revocação
Objeto 01	11	9	81,82%
Objeto 02	1	1	100%
Objeto 03	13	3	23,08%
Objeto 04	33	25	75,76%
Objeto 05	61	29	47,54%
Objeto 06	28	1	3,571%
Objeto 07	20	13	65%
Objeto 08	4	3	75%

Objeto 09	4	3	75%
Objeto 10	5	2	40%
Objeto 11	5	0	0%
Objeto 12	13	8	61,54%
Média Revocação			54,02%

A Tabela 5 apresenta a análise dos resultados da recuperação de informação visual da técnica de seleção de quadros-chave 3 em 3 segundos.

TABELA 5
SELEÇÃO DE QUADRO-CHAVE TÉCNICA 3 EM 3 SEGUNDOS

Consulta	Rotulação dos objetos verdadeiros	Acertos	% Revocação
Objeto 01	9	0	0%
Objeto 02	2	0	0%
Objeto 03	63	7	11,11%
Objeto 04	41	9	21,95%
Objeto 05	59	27	45,76%
Objeto 06	55	0	0%
Objeto 07	74	28	37,84%
Objeto 08	8	1	12,5%
Objeto 09	8	1	12,5%
Objeto 10	8	7	87,5%
Objeto 11	7	0	0%
Objeto 12	26	18	69,23%
Média Revocação			24,86%

Por fim, a Tabela 6 apresenta os resultados da recuperação de conteúdo visual da técnica seleção 1 em 1 segundo.

TABELA 6
SELEÇÃO DE QUADRO-CHAVE TÉCNICA 1 EM 1 SEGUNDO

Consulta	Rotulação dos objetos verdadeiros	Acertos	% Revocação
Objeto 01	35	0	0%
Objeto 02	07	0	0%
Objeto 03	184	4	2,17%
Objeto 04	131	2	1,53%
Objeto 05	171	5	2,92%
Objeto 06	157	23	14,6%
Objeto 07	219	29	13,2%
Objeto 08	21	2	9,52%
Objeto 09	27	0	0%
Objeto 10	19	0	0%
Objeto 11	18	3	16,7%
Objeto 12	80	12	15%
Média Revocação			6,30%

Foram utilizadas métricas de revocação na comparação das três técnicas de seleção de quadros-chave, que avaliou a capacidade de se recuperar as imagens mais relevantes para o usuário. Para o resultado calculou a quantidade de quadros que possui o objeto procurado, isto é, os acertos divididos pela quantidade de quadros da “rotulação dos objetos verdadeiros”.

De acordo com as técnicas implementadas pode-se observar que o método pixel-a-pixel é o método que apresentou o melhor resultado de forma geral, obteve 54,02% da média de resultados relevantes retornados no processo de recuperação de conteúdo visual. Obteve também 100% de acerto na consulta do objeto 02. O método de seleção de 3 em 3 segundos obteve resultados de algumas consultas de forma satisfatória. Um exemplo é a consulta 12 onde se teve 18 acertos, tendo o melhor resultado nesta consulta entre as três técnicas. Entretanto de forma geral alcançou apenas a média de 24,86% de resultados relevantes.

Por fim, o método de 1 em 1 segundo que teve uma revocação média de apenas 6,30%, também apresentou vantagens, sendo uma delas a consulta do Objeto 06, onde a mesma acertou 23 quadros relevantes. Todavia, deve ser levado em consideração que a mesma técnica possuía uma base 157 quadros para consulta. Em comparação com as outras técnicas que possuem uma maior quantidade de quadros na “rotulação dos objetos verdadeiros”.

Observar-se que apesar da base de quadros selecionados pelas técnicas 1 em 1 segundo é maior que a técnica pixel-a-pixel, esta última obteve os resultados mais relevantes da consulta na maioria dos casos. Pode-se dizer que o resultado final da recuperação de informação por conteúdo visual possui uma dependência das entradas e bases. Mesmo quando a base de quadros é maior, como na técnica de seleção de 1 em 1 segundo não houve influência de maneira geral no melhor resultado. Além disso, devido ao fato dos dois primeiros métodos (i.e., 1 em 1 segundo e 3 em 3 segundos) não realizarem nenhum processamento sobre as imagens (apenas selecionou os quadros em relação ao tempo do

vídeo) os resultados obtidos não foram tão relevante quanto à técnica pixel-a-pixel, que realiza a seleção dos quadros-chaves de maior relevância de acordo com critério de pré-processamento.

6. CONCLUSÕES

Com o compartilhamento de conteúdo visual na Internet, um grande volume de dados multimídias é publicado. Entretanto, há risco de usuários cometerem crimes cibernéticos através dos conteúdos multimídia publicado [12]. Devido ao volume destes dados, faz-se necessário ter formas automáticas tanto para realizar o gerenciamento dos mesmos, quanto para a realização de buscas. Esta automação pode ser realizada através de técnicas de recuperação por conteúdo visual.

De acordo com Sivic e Zisserman [10] este processo já é utilizado para gerenciamento de dados multimídia, sendo extremamente importante ter gerenciamento e controle do conteúdo visual publicado [4]. Entretanto, não aplicando técnicas de seleção de quadro-chave no processo inicial da recuperação por conteúdo de um vídeo, tem como consequência o excesso de informações armazenadas.

Portanto, o presente trabalho teve como objetivo avaliar o comportamento de diferentes técnicas de seleção de quadro-chave aplicado na recuperação de informação por conteúdo visual. Para isso, três técnicas de seleção de quadros-chave foram implementadas e avaliadas, são elas: pixel-a-pixel, 3 em 3 segundos e 1 em 1 segundo. Estas técnicas foram avaliadas utilizando o trabalho do Silva [11] conforme apresentado no experimento realizado. De acordo com os resultados dos experimentos, verifica-se que a utilização de técnicas de seleção de quadros-chave na recuperação de informação por conteúdo visual de fato influenciam nos resultados obtidos. A técnica de seleção de quadros-chave pixel-a-pixel foi a técnica que obteve os melhores resultados, obtendo a media de 54,02% de resultados relevantes.

Este trabalho contribui para a área de processamento digital de imagens apresentando uma avaliação de técnicas de seleção de quadros-chave no processo de recuperação de

informação por conteúdo visual. Os resultados obtidos indicam que a quantidade de quadros selecionados não garante os melhores resultados, e, além disso, a relevância dos quadros selecionados pode ser determinante na obtenção de melhores resultados na recuperação de informação por conteúdo visual. Ou seja, algoritmos que realizam uma análise para selecionar os quadros relevantes, como a técnica pixel-a-pixel, tende a gerar melhores resultados.

REFERÊNCIAS

- [1] LOPES, Ana Paula B.; AVILA, Sandra Eliza Fontes; PEIXOTO, Anderson N. A.; OLIVEIRA, Rodrigo Silva; COELHO, Marcelo; ARAÚJO, Arnaldo: “Nude Detection in Video Using Bag-of-Visual-Features” XXII Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI, IEEE Computer Society Press, Rio de Janeiro, RJ, Brasil, 2009.
- [2] PAIVA, Claudio Cardoso; SERRANO, Paulo Henrique Souto Maior. “Critérios de Categorização para os vídeos do Youtube”. X Congresso de Ciências da Comunicação na Região Nordeste, São Luis, MA, Junho 2008.
- [3] DESELAERS, Thomas; Lexi Hermann; PIMENIDIS, Ney. “Bag-of-Visual-Words models for Adult Image Classification and Filtering” in International Conference on Pattern Recognition (ICPR), Florida, USA, December 2008.
- [4] AVILA, Sandra Eliza Fontes de. “Uma Abordagem Baseada em Características de Cor para a Elaboração Automática e Avaliação Subjetiva de Resumos Estáticos de Vídeos”. Dissertação (Mestrado) - UFMG, CNPq, 2008.
- [5] GOMES, Marcio Luiz Rossato. “Recuperação de Vídeos por Conteúdo com Base em Informações Estáticas e Dinâmicas”. Dissertação (Mestrado em Informática) - Pontifícia Universidade Católica do Paraná, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, 2006.
- [6] MORIMOTO, Carlos H; SANTOS, Thiago Teixeira. “Estruturação e indexação de vídeo digital”. in WebMídia, XVI Brazilian Symposium on Computer Graphics and Image Processing, Outubro 2003.
- [7] PIMENTEL FILHO, C. A. F.; SAIBEL A. Celso; BUCK, T. A. “Integração de Métodos Baseados em Diferença de Quadros para Sumarização do Conteúdo de Vídeos”. In: WebMedia - XIV Simpósio Brasileiro de Sistemas Multimídia e Web, 2008, Vila Velha, v. 2. p. 1-4, 2008.
- [8] BUENO Márcio, SANTOS Andréa, FERREIRA André, KELNER Judith, SIQUEIRA Iony. “Refinamento de Reconstrução 3D através de Seleção Apropriada de Keyframes”. VIII Workshop de Realidade Virtual e Aumentada, Uberaba, MG, 2011
- [9] NASCIMENTO, Gilson A. do Nascimento; MANZATO, Marcelo G; GOULARTE, Rudinei. “Extração de quadros-chave como subsídio para personalização em vídeos digitais”. Instituto de Ciências Matemáticas e de Computação, USP, São Carlos, 2010.
- [10] SIVIC,Josef; ZISSERMAN, Andrew. “Efficient Visual Search of Videos Cast as Text Retrieval.”.IEEE Trans Pattern Anal Mach Intell 2009.
- [11] SILVA, Henrique Batista. “Análise de Estruturas Métricas para Recuperação de Vídeo Utilizando Vocabulário Visual”. 88f. Dissertação (Mestrado) - Pontifícia Universidade Católica de Minas Gerais, Programa de Pós-Graduação em Informática, Belo Horizonte, 2011.
- [12] Ministério Público Federal Procuradoria Da República No Estado De Sp Grupo De Combate Aos Crimes Cibernéticos.“ Crimes Cibernéticos Manual Prático De Investigação”. Abril, 2006.

Reconstrução de Trajetória de Voo para Simulações 3D em Sinistros Aeronáuticos

Bruno Gomes de Andrade

Instituto Nacional de Criminalística, Polícia Federal, Brasil
SPO – Lote 7 – Setores Complementares – 70.610-902, Brasília/DF - Brasil
andrade.bga@dpf.gov.br

RESUMO — A reconstrução da trajetória de voo é um importante passo na investigação das causas determinantes de sinistros aeronáuticos. A proposta deste artigo é expor e discutir a metodologia utilizada pelo NTSB (*National Transport Safety Board*) para reconstruir o movimento de uma aeronave e as forças que a produzem a partir de dados do FDR (*Flight Data Recorder*). O objetivo principal da metodologia apresentada é definir a posição e orientação da aeronave ao longo do voo, determinando respostas a comandos, falhas no sistema, perturbações externas ou outros fatores que possam afetar sua trajetória. A abordagem aqui utilizada permite também o cálculo das cargas produzidas ao longo do voo em qualquer ponto da aeronave.

1. INTRODUÇÃO

Em sinistros aeronáuticos, um dos aspectos importantes no estudo do desempenho da aeronave consiste no cálculo de sua posição e orientação ao longo do voo, baseados, principalmente, nas informações dos radares primários e secundários do controle de tráfego aéreo, dados do *Flight Data Recorder* (FDR) e gravação do *Cockpit Voice Recorder* (CVR).

Juntamente com a investigação e análise das causas determinantes do acidente, é comum, existindo dados disponíveis suficientes, a reconstrução dos últimos instantes de voos sinistrados através da reprodução por meio de simulação digital 3D.

Os parâmetros gravados diretamente pelo FDR, áudio do CVR, dados de radares, condições meteorológicas e demais informações derivadas do FDR, como a trajetória do voo, possibilitam, através da simulação 3D, a integração de todos os dados temporais e espaciais, proporcionando uma perspectiva visual dos acontecimentos.

A flexibilidade de combinar os dados, ângulos de câmeras e outros elementos visuais é uma grande vantagem na forma de apresentar a sequência de acontecimentos e, ao mesmo tempo, valorizar e facilitar ao destinatário da investigação o entendimento das conclusões do

laudo pericial.

Entretanto, antes da construção de uma simulação 3D que represente os últimos instantes de voos sinistrados é necessário reconstruir a trajetória da aeronave. O presente artigo visa a expor e discutir a metodologia aplicada pelo *National Transport Safety Board* (NTSB) [1], [2] em diversos casos.

2. CÁLCULOS DE DESEMPENHO DE AERONAVES BASEADOS EM DADOS DO FDR

O FDR, enquanto instrumento de gravação de dados, tem a função de registrar os parâmetros necessários para a reconstrução precisa da trajetória, velocidade e orientação da aeronave, bem como gravar dados relativos à potência, configuração e operação dos motores [3], [4].

Apesar de gravar uma série de parâmetros relacionados ao voo, o FDR não registra diretamente todas as variáveis que podem ser de interesse para a análise do desempenho da aeronave. Entretanto, diversos parâmetros podem ser derivados dos dados gravados diretamente pelo FDR, entre eles a trajetória.

Uma estimativa precisa da trajetória de voo da aeronave pode ser obtida através da dupla integração dos dados do sensor triaxial de aceleração (acelerômetros) instalado nas aeronaves. Entretanto, há alguns fatores intrínsecos relacionados ao sensor que devem ser entendidos e levados em consideração para a correta interpretação dos dados gravados.

Uma delas é o fato de que, geralmente, os acelerômetros não estão localizados coincidentemente com o centro de gravidade (CG) da aeronave. Dessa forma, antes de proceder à integração, os fatores de carga (quantidades efetivamente medidas pelos acelerômetros) registrados pelos sensores

devem ser corrigidos para valores relativos à posição do CG. Os ajustes são calculados através dos valores efetivamente medidos pelos acelerômetros e pelas velocidades e acelerações angulares da aeronave, descritas pelas derivadas dos ângulos de guinada, rolamento e arfagem, parâmetros estes gravados diretamente pelo FDR.

Ademais, sensores do tipo acelerômetro geralmente contém intrinsecamente pequenos erros de *offset*, ou erros de *bias*, que podem produzir grandes divergências entre a velocidade e posição reais e as calculadas, se não forem removidos antes da integração.

Os valores dos *offsets* presentes nos sensores podem ser determinados através de algoritmos de minimização, de tal forma que o resultado da posição da aeronave quando integrada pelos dados dos acelerômetros, esteja de acordo com a posição da aeronave vinda de outras fontes, como dados de radar ou GPS (geralmente são dados com frequência de amostragem bem mais baixa, mas que possuem maior precisão).

2.1 . CORREÇÕES NOS ACELERÔMETROS E INTEGRAÇÃO

Os fatores de carga registrados pelo FDR são parâmetros extremamente úteis pois quando integrados corretamente fornecem informações sobre a velocidade e posição inercial da aeronave. Ademais, são parâmetros que possuem frequência de amostragem bem superior a outros dados que também registram a trajetória da aeronave, como informações de radares e posição registrada por equipamentos GPS. Dados dos acelerômetros costumam ter frequência de registro pelo FDR da ordem de 120-250Hz enquanto que dados de GPS podem ter frequência de registro da ordem de 5s. Dessa forma, a trajetória obtida através das medidas dos acelerômetros apresenta uma imensa vantagem por apresentarem um maior detalhamento temporal.

Em relação aos problemas intrínsecos presentes nos acelerômetros, os dois erros levados em consideração na metodologia do NTSB são o erro devido à diferença entre a localização

dos acelerômetros e a posição do CG, e o erro devido à presença do *offset* na medição dos fatores de carga.

Geralmente, os acelerômetros são localizados bem próximos ao CG, de forma que o erro devido à localização do sensor tende a ser bem pequeno. Já o erro de *bias* tende a ser bastante significativo pois mesmo um pequeno *offset* pode gerar grandes erros quando integrados no tempo.

Ademais o *offset* presente nos sensores não é constante, mas tende a variar lentamente ao longo do voo. Com a finalidade de simplificar os cálculos, os erros de *offset* são considerados constantes durante pequenos intervalos de tempo (algo em torno de 30 a 60 segundos). Importante notar que essa simplificação não afeta o resultado final de forma significativa.

2.2 . ERRO DE POSICIONAMENTO DOS ACELERÔMETROS

As mesmas equações utilizadas para corrigir o erro relativo à diferença entre a posição dos acelerômetros e a posição do CG também fornecem os meios para calcular os fatores de carga em qualquer ponto da aeronave. A aceleração de um ponto P qualquer na aeronave é dada por:

$$\vec{a}_P = \dot{\vec{V}}_{CG} \Big|_{xyz} + \vec{\omega} \times \vec{V}_{CG} + \ddot{\vec{\omega}} \Big|_{xyz} \times \vec{r}_P + \vec{\omega} \times (\vec{\omega} \times \vec{r}_P) \quad \text{Eq. 1}$$

Onde $\Big|_{xyz}$ é o vetor distância entre o CG e o ponto P, e $\vec{\omega}$ é o vetor velocidade angular em torno do CG. A notação $\Big|_{xyz}$ na equação 1 significa que a derivada é relativa ao sistema de coordenadas do corpo, em oposição ao sistema inercial de coordenadas. A relação entre os dois sistemas de coordenadas é mostrada na fig. 1.

Logo, com $\vec{V}_{CG} \Big|_{xyz} = \{u \ v \ w\}$ temos:

$$\dot{\vec{V}}_{CG} \Big|_{xyz} = \begin{Bmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \end{Bmatrix} \quad \text{Eq. 2}$$

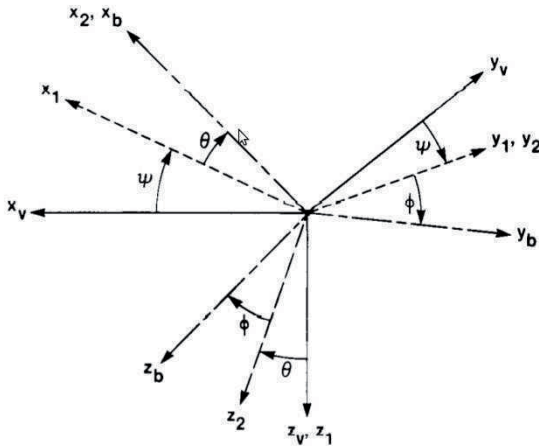


Figura 1: Relação entre os sistemas de coordenadas inercial projetada (x_v, y_v, z_v) e do corpo (x_b, y_b, z_b), ambos com origem no CG. [4]

Onde $\{u, v, w\}$ são as velocidades ao longo dos eixos X_b, Y_b e Z_b (coordenadas do corpo). Os componentes de $\vec{r}_p$ e $\vec{\omega}$ no eixo do corpo são:

$$\vec{r}_p = \hat{x}i + \hat{y}j + \hat{z}k = \begin{Bmatrix} x \\ y \\ z \end{Bmatrix} \quad \text{Eq. 3}$$

$$\vec{\omega} = \hat{P}i + \hat{Q}j + \hat{R}k = \begin{Bmatrix} P \\ Q \\ R \end{Bmatrix} \quad \text{Eq. 4}$$

Onde i, j e k são vetores unitários ao longo dos eixos X_b, Y_b e Z_b . Utilizando a seguinte notação mais simplificada $dx/dt = \dot{x}$ para denotar derivadas, as velocidades angulares P, Q e R são:

$$\begin{Bmatrix} P \\ Q \\ R \end{Bmatrix} = \begin{bmatrix} \sin \theta & 0 & 1 \\ \sin \phi \cos \theta & \cos \phi & 0 \\ \cos \phi \cos \theta & -\sin \phi & 0 \end{bmatrix} \begin{Bmatrix} \dot{\psi} \\ \dot{\theta} \\ \dot{\phi} \end{Bmatrix} \quad \text{Eq. 5}$$

Os ângulos de guinada (Ψ), arfagem, (Θ) e rolamento (Φ) são parâmetros gravados pelo FDR, sendo o ângulo de guinada considerado igual ao valor medido pela bússola magnética da aeronave, corrigido pelo valor da declinação magnética registrada no local do sinistro.

Da equação 4, temos que:

$$\vec{\omega} \Big|_{xyz} = \begin{Bmatrix} \dot{P} \\ \dot{Q} \\ \dot{R} \end{Bmatrix} \quad \text{Eq. 6}$$

Substituindo as equações 2, 3, 4 e 6 na equação 1 e realizando o produto vetorial, temos que:

$$\vec{a}_p = \begin{Bmatrix} \dot{u} + wQ - vR \\ \dot{v} + uR - wP \\ \dot{w} + vP - uQ \end{Bmatrix} + \begin{Bmatrix} Q(yP - xQ) + R(zP - xR) + (z\dot{Q} - y\dot{R}) \\ R(zQ - yR) + P(xQ - yP) + (x\dot{R} - z\dot{P}) \\ P(xR - zP) + Q(yR - zQ) + (y\dot{P} - x\dot{Q}) \end{Bmatrix} \quad \text{Eq. 7}$$

Por definição temos que a posição $\{x, y, z\}$ no CG é igual a $\{0, 0, 0\}$. Logo, o primeiro termo entre chaves na equação 7 deve ser a aceleração em torno do CG e o segundo termo entre chaves é o incremento na aceleração devido ao ponto P estar distante do CG. Reescrevendo a equação 7 de forma simplificada, temos que:

$$\vec{a}_p = \vec{a}_{CG} + \Delta \vec{a} \quad \text{Eq. 8}$$

Conforme mencionado anteriormente, um acelerômetro triaxial posicionado em um ponto P, mede o fator de carga segundo a equação:

$$\vec{n}_p = \frac{\vec{a}_p - \vec{g}}{g} = \frac{\vec{a}_{CG} + \Delta \vec{a} - \vec{g}}{g} = \vec{n}_{CG} + \frac{\Delta \vec{a}}{g} \quad \text{Eq. 9}$$

Onde a equação 8 foi usada para substituir $\vec{a}_p$. Os componentes de $\vec{n}$ são $\{n_x, n_y, n_z\}$.

Geralmente o fator de carga medido em relação ao eixo Z_b não segue a convenção tradicional de coordenadas do eixo do corpo, onde a medição positiva de uma grandeza tem direção para baixo (figura 1). A medição do fator de carga no eixo Z_b é invertida, sendo registrado pelo acelerômetro e gravado pelo FDR como:

$$n_{LF} = -n_z \quad \text{Eq. 10}$$

Geralmente o FDR registra os valores de $\{n_x, n_y, n_z\}$ como *Longitudinal Acceleration*, *Lateral Acceleration* e *Vertical/Normal Acceleration*. Os valores de $\{n_x, n_y, n_z\}$ em torno do CG podem ser encontrados utilizando os dados do FDR juntamente com as equações

9, 10 e 11 com $[x, y, z]$ na equação 7 sendo a distancia do sensor triaxial até o CG.

Uma vez calculado os fatores de carga no CG, as equações 7, 9 e 10 podem ser utilizadas para calcular fatores de carga em quaisquer outros pontos, como no centro de massa das superfícies de controle ou no *cockpit* da aeronave, caso seja de interesse da investigação.

2.3 . ERRO DE BIAS NOS ACELERÔMETROS

Os fatores de carga registrados pelos acelerômetros raramente correspondem aos valores reais, pois os sensores não são perfeitos. Enquanto os incrementos relativos registrados no fator de carga a partir de um ponto de referência possuem um erro muito pequeno, o valor absoluto do fator de carga de referência geralmente sofre de um erro de *offset* em relação ao valor real. Esse valor é facilmente visualizado na prática a partir da análise dos dados de acelerômetro enquanto a aeronave se encontra em posição de repouso. Os fatores de carga longitudinal, lateral e vertical nesse caso são dados pela equação 9 com $\vec{a}_p = 0$, ou seja, $\vec{n}_p = -\vec{g}/g$ onde:

$$\vec{g} = g \begin{pmatrix} -\sin \theta \\ \cos \theta \sin \phi \\ \cos \theta \cos \phi \end{pmatrix} \quad \text{Eq. 11}$$

e g é o valor da aceleração da gravidade.

Os fatores de carga registrados pelo FDR raramente refletem esses valores exatos, e isso é facilmente verificado em casos reais analisando-se tais variáveis, por exemplo, logo antes da corrida de decolagem, quando a aeronave se encontra em repouso.

Define-se o erro de *offset* pela equação:

$$\Delta \vec{n} = \vec{n}_{\text{atual}} - \vec{n}_{\text{Medido}} \quad \text{Eq. 12}$$

Os erros de *offset* que existem logo antes da decolagem não são necessariamente os mesmos que existem no instante do acidente pois, conforme já mencionado, o erro de *offset* pode divergir ligeiramente ao longo do voo. Dessa forma, o erro de *offset* que se aplica ao momento do acidente deve ser determinado em um instante anterior, mas o mais próximo

possível do evento da queda.

Os erros de *offset* relativos aos acelerômetros podem ser calculados para uma aeronave em voo através de um processo iterativo que compara a posição resultante da integração dos valores registrados pelos acelerômetros com a posição calculada utilizando dados de navegação de baixa frequência registrados pelo FDR.

A velocidade em relação ao solo (V_{solo}) e o ângulo de deriva (Ψ_{deriva}) são parâmetros geralmente registrados pelo FDR, assim como a latitude e longitude registrados pelo GPS. Entretanto, estas grandezas são registradas com uma frequência de amostragem bem mais baixa. Geralmente a informação de posição oriunda do GPS é mais precisa.

3 . VISÃO GERAL DO PROCEDIMENTO DE CÁLCULO DO ERRO DE BIAS NOS ACELERÔMETROS.

Nesta seção será detalhado o processo de cálculo do erro de *offset* e as equações envolvidas segundo o procedimento adotado pelo NTSB [1] e [2] em investigações de sinistros aeronáuticos que possuem dados do FDR disponíveis.

Primeiramente, a posição atual da aeronave é definida utilizando os dados de velocidade relativa ao solo e o ângulo de deriva registrados pelo FDR (posição relativa ao eixos inerciais X e Y). Com relação ao eixo inercial Z, uma melhor estimativa da altitude é calculada resolvendo a equação hidrostática com estimativas da densidade do ar atual em cada ponto e pode ser vista em maiores detalhes em [1], [2].

Em seguida uma estimativa do erro de *offset* dos acelerômetros é realizada através do cálculo dos fatores de carga que resultam da informação de posição que se obteve anteriormente e comparando esses fatores de carga com os fatores de carga registrados pelo FDR.

Nesse ponto as iterações tem início. Os fatores de carga são atualizados com a estimativa de erro de *offset* e depois integradas para obter a informação de posição. Esses valores de posição integradas são comparadas

com a posição inercial calculada previamente e o erro é calculado.

A sensibilidade na variação da posição devido à integração dos fatores de carga somada ao incremento da diferença no *offset* para cada sensor, é avaliada modificando os valores de *offset* ligeiramente e recalculando os erros. Com a avaliação do resultado do incremento de cada *offset* sobre a posição da aeronave, os valores de erro são atualizados e os valores de posição, erros e valor do *offset* são recalculados. Quando mudanças no valor do *offset* não mais resultam em uma significativa redução na diferença, então se chegou a melhor estimativa do valor do *offset* e os valores dos fatores de carga finais são calculados.

Os valores integrados finais de velocidade e posição são também calculados e representam a melhor estimativa da velocidade inercial.

Na próxima seção todo esse processo será visto em detalhes.

3.1 . DETALHES DO CÁLCULO DO OFFSET NOS ACELERÔMETROS.

O algoritmo para o cálculo do *offset* dos acelerômetros pode ser assim resumido:

Passo 1 - O valor atual da posição da aeronave é calculada através dos parâmetros de velocidade relativos ao solo (V_{solo}) e ângulo de deriva (Ψ_{deriva}). A posição é medida no sistema de coordenadas inercial, onde a origem do sistema é a posição da aeronave no início do cálculo, e onde o eixo X_{novo} está alinhado com o ângulo de trajetória relativo ao solo (figura 2, [2]). O ângulo de trajetória, ou seja, o *heading* do vetor velocidade, nesse sistema de coordenadas é dado por:

$$\Psi_{trajetória} = \Psi + \Psi_{deriva} - [\Psi + \Psi_{deriva}]_{t=0} \quad Eq. 13$$

O *heading* da aeronave nesse sistema de coordenadas é dado por:

$$\Psi_{novo} = \Psi - [\Psi + \Psi_{deriva}]_{t=0} \quad Eq. 14$$

Os componentes da velocidade inercial nos eixos X (Norte) e Y (Leste) são:

$$V_{X_{novo}} = V_{solo} \cos(\Psi_{trajetória}) \quad Eq. 15$$

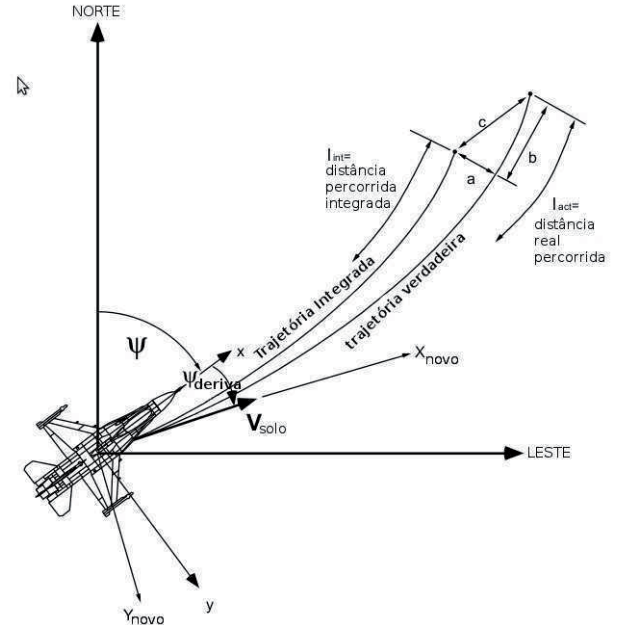


Figura 2: Convenções utilizadas no cálculo dos *offsets* dos sensores,[2] (figura adaptada)

$$V_{Y_{novo}} = V_{solo} \sin(\Psi_{trajetória}) \quad Eq. 16$$

A posição da aeronave então será:

$$X_{novo} = \int V_{X_{novo}} dt \quad Eq. 17$$

$$Y_{novo} = \int V_{Y_{novo}} dt \quad Eq. 18$$

Passo 2 - Os fatores de carga resultantes da utilização dos valores de posição inercial calculados previamente são computadas para fornecer uma estimativa inicial do *offset* dos acelerômetros.

Combinando as equações 7, 9 e 11 os fatores de carga no CG são dados por:

$$\vec{n}_{CG} = \frac{1}{g} \begin{pmatrix} \dot{u} + wQ - vR \\ \dot{v} + uR - wP \\ \dot{w} + vP - uQ \end{pmatrix} + \begin{pmatrix} -\sin\theta \\ \cos\theta \sin\phi \\ \cos\theta \cos\phi \end{pmatrix} \quad Eq. 19$$

Os ângulos de Euler (Ψ, θ, ϕ) são registrados diretamente pelo FDR, sendo o valor do ângulo de guinada Ψ considerado como o valor da medição da bússola magnética corrigida pelo valor da declinação magnética) e as velocidades angulares são dadas pela equação 6. As componentes de velocidade no sistema de coordenadas do corpo podem ser calculada pelos componentes das velocidades inerciais através da utilização da matriz de

cossenos diretores $T_{Inercial \rightarrow Corpo}$ (*Direction Cosine Matrix*, ou DCM):

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = [T_{Inercial \rightarrow Corpo}] \begin{bmatrix} V_{X_{novo}} \\ V_{Y_{novo}} \\ dh/dt \end{bmatrix} \quad Eq. 20$$

A DCM é mostrada na equação 24. Note que, como $T_{Inercial \rightarrow Corpo}$ é ortogonal, sua inversa é simplesmente sua transposta:

$$[T_{Corpo \rightarrow Inercial}] = [T_{Inercial \rightarrow Corpo}]^{-1} = [T_{Inercial \rightarrow Corpo}]^T \quad Eq. 21$$

Renomeando $\vec{n}_{CG}$ calculado utilizando as equações 19, 20 e 24 de $\vec{n}_{Inercial}$ e os valores dos fatores de carga corrigidos para a posição do CG de $\vec{n}_{Corrigido}$, uma estimativa inicial do *offset* vem por meio da equação 12 como:

$$\Delta \vec{n}_{Estimado} = \vec{n}_{Inercial} - \vec{n}_{Corrigido} \quad Eq. 22$$

Como na prática o *offset* dos acelerômetros ($\Delta \vec{n}$) são aproximadamente constantes durante o período de interesse, a média de $\Delta \vec{n}_{Estimado}$ do resultado da equação 22 durante o período de cálculo é usada como a estimativa de *offset* inicial.

Passo 3 - Agora se inicia o processo iterativo. O $\vec{n}_{Corrigido}$ é atualizado com a média $\Delta \vec{n}_{Estimado}$ do passo 2 para obter os novos valores de fatores de carga:

$$\vec{n}_{novo} = \vec{n}_{Corrigido} + \Delta \vec{n}_{Estimado} \Big|_{m\u00e9dia} \quad Eq. 23$$

Passo 4 - Os valores de $\vec{n}_{novo}$ e a equação 19 fornecem três equações diferenciais não lineares para as variáveis $\{u, v, w\}$. Essas equações são integradas numericamente utilizando os valores de $\{u, v, w\}$ dados pela equação 20 como valores iniciais para a integração.

Passo 5 - Os valores $\{u_f, v_f, w_f\}$ obtidos através do passo 4 são transformados para o

sistema de coordenadas inercial utilizando a inversa da matriz de cossenos diretores:

$$\begin{bmatrix} V_{X_f} \\ V_{Y_f} \\ \dot{h}_f \end{bmatrix} = [T_{Corpo \rightarrow Inercial}] \begin{bmatrix} u_f \\ v_f \\ w_f \end{bmatrix} \quad Eq. 25$$

Passo 6 - As posições X_f e Y_f integradas são calculadas utilizando 17 e 18 e as velocidades integradas do passo 5. A derivada da altitude dh_f/dt é integrada para obter o valor de altitude h_f .

As diferenças entre as posições integradas $\{X_f, Y_f, h_f\}$ e as posições inerciais $\{X_{novo}, Y_{novo}, h_{novo}\}$ são uma medida do erro ainda existente nos acelerômetros. Para calcular como os fatores de carga devem mudar para reduzir esse erro, o efeito da variação de cada incremento no *offset* na diferença entre a posição calculada e real deve ser entendido. Idealmente, o *offset* de um dos três acelerômetros que compõe o sensor triaxial deve ser associado com uma medida de erro única que é relativamente independente dos *offsets* dos outros acelerômetros e que pode ser monitorada para determinar os efeitos da variação do incremento no *offset*. Tal erro de medida para os *offsets* $\{n_x, n_y, n_z\}$ estão ilustrados na figura 2.

Para aeronaves de transporte comerciais que mantêm um pequeno ângulo arfagem (θ) e rolamento (Φ), um erro em n_x resulta em uma diferença entre a distância horizontal percorrida integrada e a verdadeira conforme ilustrado pelas curvas I_{it} e I_{at} na figura 2. Um erro em n_y resulta em uma diferença entre o ângulo de trajetória atual e integrados, no final do processo de integração, esse erro aparecerá como um erro de posição final (distância “a” na figura 2). Um erro em n_z vai resultar em uma diferença entre a altitude atual e a integrada.

I_{erro} , a diferença média no tempo entre I_{int} e I_{act} , é dada por:

$$Eq. 26$$

$$T_{Inercial \rightarrow Corpo} = \begin{bmatrix} \cos \theta \cos \psi_{novo} & \cos \theta \sin \psi_{novo} & -\sin \theta \\ \sin \phi \sin \theta \cos \psi_{novo} - \cos \phi \sin \psi_{novo} & \sin \phi \sin \theta \sin \psi_{novo} + \cos \phi \cos \psi_{novo} & \sin \phi \cos \theta \\ \cos \phi \sin \theta \cos \psi_{novo} + \sin \phi \sin \psi_{novo} & \cos \phi \sin \theta \sin \psi_{novo} - \sin \phi \cos \psi_{novo} & \cos \phi \cos \theta \end{bmatrix} \quad Eq. 24$$

$$I_{erro} = \frac{\int_0^t \sqrt{(I_{int} - I_{act})^2} dt}{\int_0^t dt}$$

onde

$$I_i(t) = \int_0^t \left(\frac{dI_i}{dt} \right) dt = \int_0^t v_i(t) dt \quad Eq. 27$$

$$v_i(t) = \sqrt{v_{x_i}^2 + v_{y_i}^2} \quad Eq. 28$$

e o subscrito “i” refere-se aos dados integrados em um caso, e o dado real no outro.

Uma equação aproximada para a distancia “a” na figura 2 que fornece a medida do erro em n_y é dada por:

$$a \approx \sqrt{c^2 - b^2} \quad Eq. 29$$

onde

$$b = I_f - I_{atual} \quad Eq. 30$$

e “c” é a distância entre a posição final integrada e a posição atual:

$$c = \sqrt{(X_f - X_{novo})^2 + (Y_f - Y_{novo})^2} \quad Eq. 31$$

Para que se possa utilizar “a” no cálculo do erro, é feita uma media no tempo do valor “a”:

$$a_{media} = \frac{\int_0^t a dt}{\int_0^t dt} \quad Eq. 32$$

O erro médio no tempo da altitude é

$$h_{erro} = \frac{\int_0^t \sqrt{(h_f - h)^2} dt}{\int_0^t dt} \quad Eq. 33$$

Para usar essas medidas de erro para determinar como o *offset* deve variar, a sensibilidade de cada erro para com seu *offset* deve ser calculado. Isso é conseguido através do incremento de cada *offset* independentemente por um pequeno valor e recalculando em seguida a posição e erros integrados. Se ξ_i é o erro associado com n_i , $\Delta\xi_i$ é o incremento no erro devido ao incremento

no *offset*, e ε é o incremento, então o novo n_i é:

$$n_i|_{novo} = n_i + \frac{\xi_i}{\Delta\xi_i} \varepsilon \quad Eq. 34$$

na prática, o incremento de n_i durante cada iteração não deve ser maior que um determinado valor (0.002g, por exemplo) sob pena de o processo de minimização não convergir.

Se após diversas iterações o incremento em ξ_i ficar abaixo de um determinado valor, então o *offset* mais adequado foi encontrado e o processo esta completo.

4 . RESULTADOS PRÁTICOS

Os casos que envolvem investigações de acidentes aéreos correm, geralmente, em segredo de justiça. Essa condição inviabiliza a apresentação dos resultados em casos reais de atuação da perícia criminal federal.

Entretanto, para mostrar a eficiência da metodologia, serão utilizados dados reais do acidente com a aeronave Airbus A320 que realizaria o voo US Airways 1549, ocorrido em 15 de janeiro de 2009, em Nova Jersey, Estados Unidos¹. Logo após a decolagem a aeronave teve um incidente de “bird strike” causando perda de potência em ambos os motores, culminando com o pouso de emergência no rio Hudson.

As figuras 3 e 4 mostram a coerência entre as posições inerciais da aeronave oriundos dos dados de baixa frequência e do resultado da aplicação da metodologia apresentada.

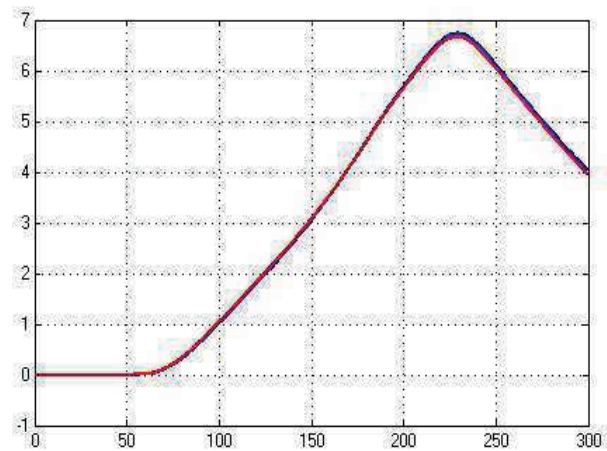


Figura 3: Gráfico mostrando a convergência entre as curvas de baixa frequência (vermelha) e integração pelos acelerômetros (eixo norte)

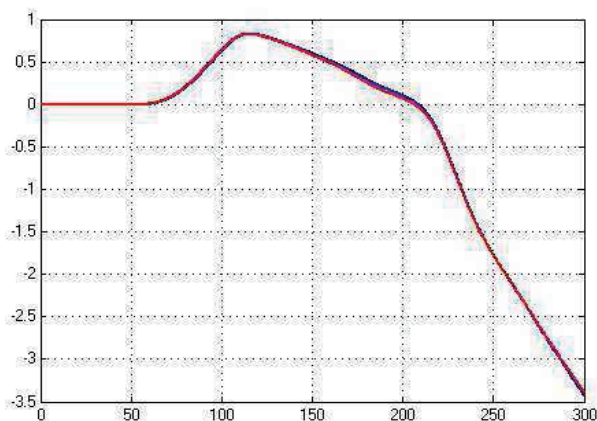


Figura 4: Gráfico mostrando a convergência entre as curvas de baixa frequência (vermelha) e integração pelos acelerômetros (eixo leste)

5 . ALTERNATIVA AO PROCESSO DE MINIMIZAÇÃO UTILIZADO PELO NTSB

A solução apresentada pelo NTSB para a correção dos erros de *offset* dos acelerômetros, apesar de efetiva, não é muito elegante, pois envolve um processo iterativo de minimização do tipo tentativa e erro.

Ao consultar a literatura científica sobre reconstrução de trajetória de voo a partir de parâmetros de desempenho, é fato que esse é um problema considerado resolvido pela literatura [6], [7], [8].

É possível encontrar diversas técnicas mais sofisticadas que utilizam filtragem adaptativa não linear, através do uso de filtros de Kalman do tipo estendidos, do tipo dessensibilizados, do tipo *sigma-point*, dentre outros.

Apesar de ser uma solução mais efetiva para o problema de erros de *offset* nos acelerômetros a filtragem adaptativa não linear é extremamente mais complexa e na literatura analisada não foram encontrados estudos dessa técnica aplicada a casos reais, onde a quantidade de parâmetros disponíveis oriundos do FDR é bem mais limitada, assim como a frequência de amostragem dos parâmetros é mais baixa.

6 . ALTERNATIVA AO PROCESSO DE MINIMIZAÇÃO UTILIZADO PELO NTSB

A solução apresentada pelo NTSB para a correção dos erros de *offset* dos acelerômetros, apesar de efetiva, não é muito elegante, pois envolve um processo iterativo de minimização do tipo tentativa e erro.

Ao consultar a literatura científica sobre reconstrução de trajetória de voo a partir de parâmetros de desempenho, é fato que esse é um problema considerado resolvido [6], [7], [8].

É possível encontrar diversas técnicas mais sofisticadas que utilizam filtragem adaptativa não linear, através do uso de filtros de Kalman do tipo estendidos, do tipo dessensibilizados, do tipo *sigma-point*, dentre outros.

Apesar de ser uma solução mais efetiva para o problema de erros de *offset* nos acelerômetros a filtragem adaptativa não linear é extremamente mais complexa e na literatura analisada não foram encontrados estudos dessa técnica aplicada a casos reais, onde a quantidade de parâmetros disponíveis oriundos do FDR é bem mais limitada, assim como a frequência de amostragem dos parâmetros é mais baixa.

7 . CONCLUSÃO

Este artigo procurou expor e discutir alguns aspectos da metodologia utilizada pelo NTSB para reconstrução da trajetória de voo a partir de parâmetros de desempenho gravados pelo FDR de aeronaves sinistradas.

A reconstrução da trajetória é o primeiro passo na construção de uma simulação 3D que vise a apresentar a sequência de acontecimentos e, ao mesmo tempo, valorizar e facilitar ao destinatário da investigação o entendimento das conclusões do laudo pericial sobre as causas do sinistro aeronáutico.

Apesar de comprovadamente efetiva, a metodologia utilizada pelo NTSB pode ser melhorada com a utilização de soluções mais modernas, como a utilização de filtragem adaptativa não linear através de filtros de Kalman.

Entretanto, é necessário levar em consideração as limitações que casos reais oferecem, como a quantidade de parâmetros de desempenho e suas frequências de amostragem.

REFERÊNCIAS

- [1] National Transportation Safety Board, Office of Research and Engineering, *Group Chairman's Aircraft Performance Study, US Airways Flight 1549, Airbus A320-214, Hudson River, New Jersey January 15, 2009*, NTSB Accident Number DCA09MA026, (Washington, DC: NTSB, December 10, 2009).
- [2] National Transportation Safety Board, Office of Research and Engineering, *Group Chairman's Aircraft Performance Study, American Airlines Flight 587, Airbus A300B4-605R, Belle Harbor, New York, November 12, 2001*, NTSB Accident Number DCA02MA001 (Washington, DC: NTSB, October 10, 2002).
- [3] Convention on International Civil Aviation, Annex 6, Operation of Aircraft, Part I - International and Commercial Air Transport (Aeroplanes), ICAO Standard, Ninth Edition, July 2010.
- [4] Convention on International Civil Aviation, Annex 6, Operation of Aircraft, Part II - International General Aviation (Aeroplanes), ICAO Standard, Seventh Edition, July 2008.
- [5] E. L. Duke, Antoniewicz R. F., Krambeer, K. D. "Derivation and Definition of a Linear Aircraft Model", *NASA RP-1207*, August 1988.
- [6] Mulder, J. A., Chu, Q. P., Sridhar, J. K., Breeman, J., and Laban, M. (1999). Non-linear aircraft flight path reconstruction review and new advances. *Progress in Aerospace Sciences*, 35:673–726.
- [7] Góes, L.C.S., Hemerly, E.M., Maciel, B.C.O., Neto, W.R., Mendonça, C. and Hoff, J. *PARAMETER ESTIMATION AND FLIGHT PATH RECONSTRUCTION USING OUTPUT-ERROR METHOD*. Inverse Problems, Design and Optimization Symposium, Rio de Janeiro, Brazil, 2004.
- [8] Teixeira B. O. S., Torres L. A. B., Iscold P., Aguirre L. A., "Flight path reconstruction – A comparison of nonlinear Kalman filter and smoother algorithms", *Aerospace Science and Technology*, 2011.

FI2 - UMA PROPOSTA DE FERRAMENTA PARA INVESTIGAÇÃO FORENSE EM IMAGENS

André P. Henriques, Tauan de O. Naves, Gustavo C. Bicalho, Flávio de B. Vidal, Dúbio L. Borges
Departamento de Ciência da Computação
Universidade de Brasília

RESUMO - Este trabalho trata-se do desenvolvimento de um software de processamento de imagens para investigação forense, destinado a peritos criminais denominado FI2 (Ferramenta de Investigação de Imagens). A ferramenta foi desenvolvida utilizando como base as bibliotecas OpenCV, para o processamento de imagens, e QT, para a interface da aplicação, com a linguagem C++. Por se tratar de uma aplicação voltada a investigações criminais, a ferramenta possui recursos durante a análise forense de uma imagem, de modo a robustecer sua validade em processos judiciais.

PALAVRAS CHAVE - Processamento de Imagens, Análise Forense de Imagens, FI2

1. Introdução

Fotos e vídeos são atualmente uma das principais fontes de informações para investigações. No entanto, nem sempre é possível conseguir imagens ou vídeos de boa qualidade, geralmente devido a problemas de equipamento e do ambiente. Com a tecnologia atual esses fatores não são impedimentos para se conseguir boas imagens e vídeos, para se ter provas de identificação válidas, e assim também informações úteis para as investigações. Para a utilização dessa tecnologia é necessário que imagens e vídeos sejam digitalizados a fim de serem processados por um software dedicado a melhorar as características úteis para a análise do material.

Características que podem afetar a qualidade da imagem: Baixa resolução de imagens,

necessária para o aumento do tamanho dos detalhes úteis; Falta de contraste; Diferentes tipos de ruídos e perturbações; - Tremido causado por movimento e/ou falta de foco; Distorções geométricas limitando, consequentemente, a reconstrução das dimensões dos objetos dentro da imagem;

Com o intuito de superar essas adversidades o *framework* proposto pelo software FI2 (Ferramenta de Investigação de Imagens) disponibiliza diversas ferramentas a fim de se obter uma imagem com significado relevante para o usuário. Cada operação realizada é registrada, certificada, além fornecer um conhecimento completo das ferramentas junto ao algoritmo para obter plena objetividade e garantir que o resultado possa ser reproduzido por outra pessoa.

Para a apresentação desta proposta de *framework* para análise forense de imagens, este artigo está dividido nas seguintes seções: A Seção 2 apresenta uma breve introdução ao tema de análise forense de imagens. Na Seção 3 são apresentadas as principais características de proposta do FI2. E finalmente na Seção 4 são apresentadas as conclusões e propostas futuras para esta proposta de ferramenta para análise forense de imagens.

2. Análise de Imagens para fins forenses

Apesar da demanda de ferramentas voltadas para aplicações, em que tenham o foco na análise forense de imagens digitais, atualmente

poucas aplicações atendem de forma satisfatória as necessidades dos agentes de investigação brasileiros. Atualmente, existem vários produtos que visam o processamento de imagens e vídeos, mas poucos são destinados à área forense, dos quais apenas alguns conseguem realizar os objetivos a que se propõem. Entretanto, a maioria desses programas tem uma solução completa de hardware e software tornando o seu custo elevado e ainda impossibilitando em muitas vezes sua utilização fora do ambiente laboratorial de perícia.

Todas essas aplicações possuem ferramentas como ajuste de contraste e brilho, equalização de histograma e edição, zoom, espelhamento, rotação das imagens que podem ser encontradas na literatura básica de processamento de imagens [1]. Alguns deles oferecem ferramentas onde se pode personalizar filtros. Quase todos armazenam um *log* de auditoria para cada caso realizado a fim de torná-los apropriados para apresentações em audiências. Apesar de oferecerem a capacidade de registrar as operações realizadas junto com seus parâmetros, esses softwares não dão acesso ao algoritmo aplicado, dificultando serem auditados.

Dentro desta linha, existe a necessidade de que seja criado um *framework* que realize todos estes procedimentos que as principais ferramentas realizam, permitem sua auditoria caso necessário, evitem a dependência entre hardware e software, sendo ainda acessível a todos os órgãos periciais nas esferas investigativas.

Em relação ao aspecto de auditoria, alternativamente, este trabalho proposto visa apenas a área de software, destinando-se ao processamento de imagem com algoritmos e técnicas que são reconhecidas na literatura especializada ou descritas e detalhadas pela própria aplicação, diferente de alguns daqueles, provendo acesso ao código ou ao algoritmo implementado.

3. Ferramenta de Investigação Forense - FI2

Nesta seção são apresentadas as principais características da ferramenta proposta, bem como suas principais funcionalidades.

A FI2 é uma ferramenta de processamento de imagens especializada em investigação forense. A ferramenta foi desenvolvida utilizando como base a biblioteca *OpenCV* [2] e a biblioteca e ambiente de desenvolvimento *Qt* [3] com a linguagem C++, mantendo a portabilidade entre os sistemas operacionais Linux, Windows e MacOS.

Por se tratar de uma aplicação focada nas necessidades para autoridades investigativas periciais, o sistema proposto possui registro das técnicas e manipulações realizadas nas imagens, incluindo datas, algoritmos e parâmetros, de modo a conferir sua validade em processos legais e judiciais. Esta estratégia permite a reprodutibilidade dos resultados, por outros profissionais permitindo maior confiabilidade nos resultados alcançados durante a análise.

O *framework* de aplicação do FI2 é descrito no diagrama UML (do inglês *Unified Modeling Language*) descrito na Figura 1.

Para se alcançar este *framework* foram utilizados como referência três principais norteadores: *desempenho*, *portabilidade* e *escalabilidade*, que são descritos a seguir.

A. Desempenho

O desempenho é um fator crítico pelo fato de que técnicas de processamento de imagens terem um custo computacional elevado. Associado a isto, deseja-se resposta (latência) rápida, com o intuito de o operador obter um resultado o mais célere possível para cada alteração de parâmetro de uma determinada manipulação, facilitando, assim, a comparação entre as diversas configurações possíveis de uma determinada técnica. Para satisfazer esse requisito, decidiu-se pelo projeto de uma aplicação nativa, utilizando-se a linguagem C++, que é servida de compiladores nas mais diversas plataformas.

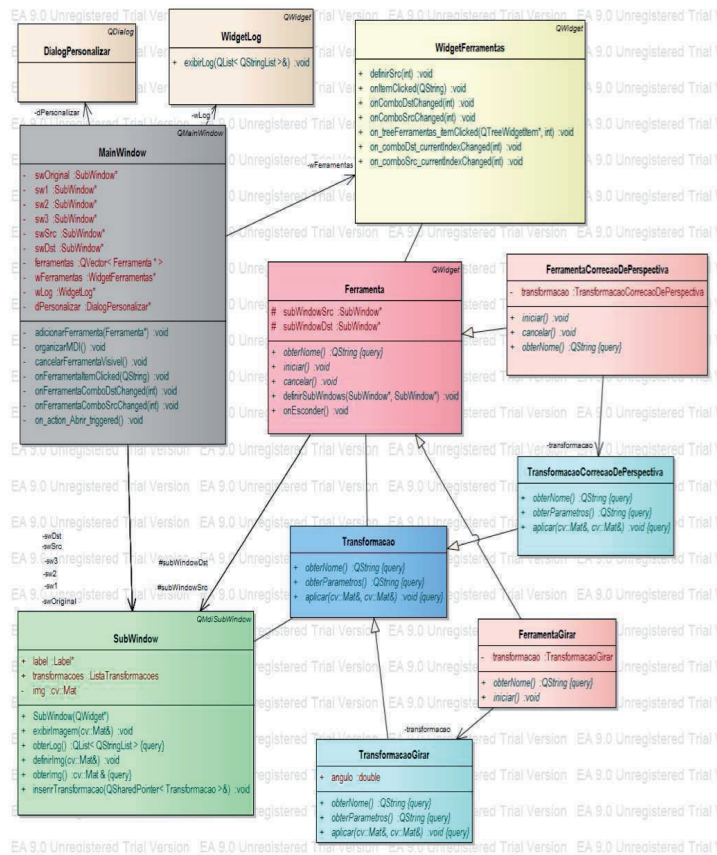


Figura 1. Diagrama UML do framework do FI2.

B. Portabilidade

A portabilidade é um aspecto importante tendo em vista o reaproveitamento de códigos e a alcançabilidade da ferramenta. Para obter a almejada portabilidade, estudou-se, dentre as possibilidades, ferramentas que pudessem ser interoperadas indistintamente entre os três principais sistemas operacionais: Windows, MacOS e Linux. Com isso, optou-se pelas bibliotecas OpenCV, para o auxílio no processamento de imagens, e QT para o desenvolvimento da interface.

C. Escalabilidade

A escalabilidade é importante pelo fato de se desejar uma ferramenta que englobe as mais diferentes e recentes ferramentas voltadas à investigação forense. Tais ferramentas

têm os mais diversos requerimentos, tanto a interface quanto as características internas do programa.

Para isso, desenvolveu-se uma interface única para o código de representação gráfica de todas as ferramentas e outra interface (descritos na Figura 2), também única, para encapsular o código de processamento de imagens das ferramentas, flexíveis o suficiente para atenderem as demandas que as ferramentas que se deseje implementar requerem. Enfim, procurou-se desenvolver uma aplicação que seja uma plataforma para que novas técnicas de processamento de imagens, em particular na área forense, fossem facilmente implementadas. As Figuras 3, 4 e 5 descrevem exemplos de algumas técnicas de processamento de imagens implementadas.

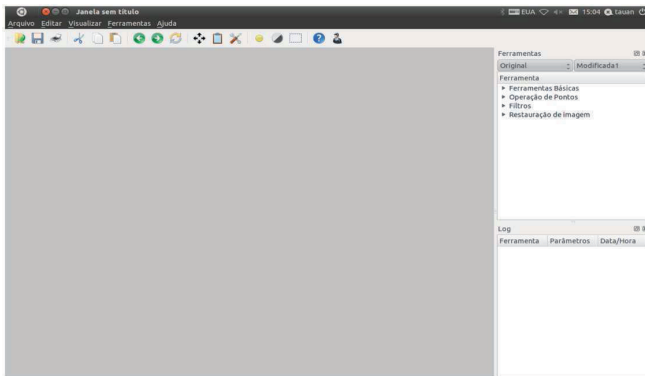


Figura 2. Interface com o usuário da ferramenta.

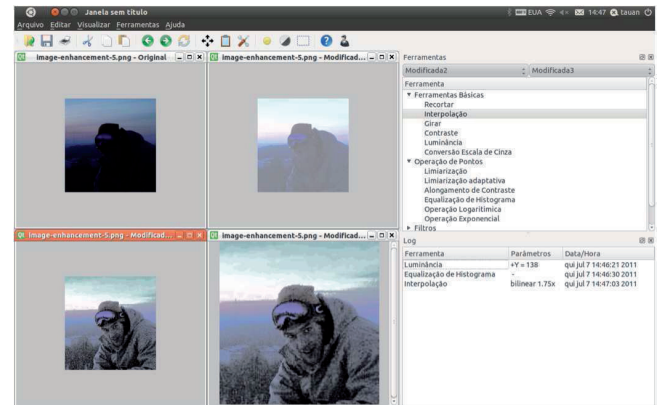


Figura 5. Ferramenta de correção de histograma.

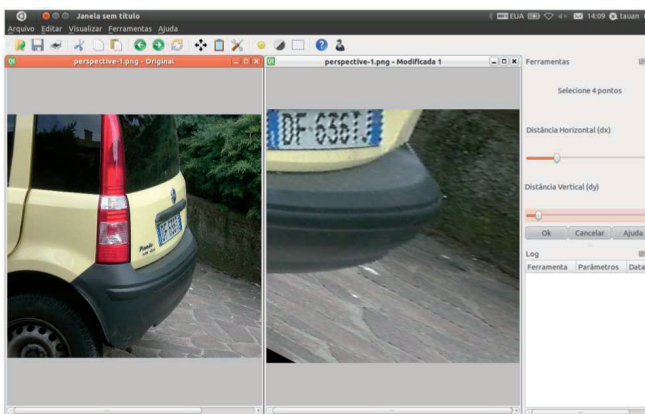


Figura 3. Ferramenta de correção de perspectiva.

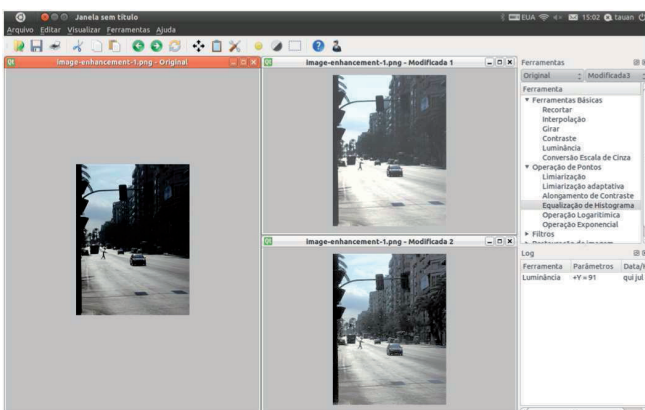


Figura 4. Ferramenta de realce.

D. Técnicas de Processamento de Imagens implementadas

A tabela 1 a seguir explicitam os recursos que foram implementados na primeira versão do FI2.

E. Interface

A aplicação foi desenvolvida de modo que novas ferramentas fossem futuramente incluídas no software, possibilitando que se preserve um modelo de interface entre todas elas. Foi utilizado o *QtCreator*[3] para a criação da interface. A interface foi feita com o objetivo de facilitar ao máximo a interação com o usuário.

Toda a interface pode ser personalizada de acordo com a necessidade/comodidade do usuário. As figuras 7 e 6 descrevem a interface para escolha de qual ferramenta será utilizada e o painel de seleção de funções, respectivamente.

F. Registro das Atividades Realizadas - Log

Uma funcionalidade utilizada nesta ferramenta é o Registro das atividades realizadas nas imagens processadas. É de fundamental importância que todo o trabalho de avaliação e análise das imagens seja documentado para posterior relato dos resultados alcança-

Tabela 1. Funcionalidades implementadas no FI2.

Ferramentas	Funções
Básicas	Cortar, Luminância, Contraste, Conversão em escala cinza e Girar/Rotacionar
Interpolação	Vizinhança mais próxima, Bilinear, Bicúbica, Interpolação Lanczos [4]
Operação de Pontos	Equalização de Histograma
Restauração da Imagem	Correção de Perspectiva

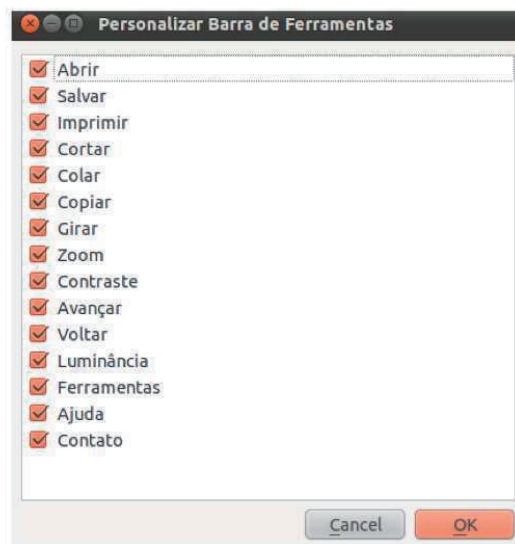


Figura 6. Seleção de ferramentas - Toolbox.

dos, bem como permitir sua reprodutibilidade em caso de necessidade de auditoria e/ou revisão dos resultados apresentados. Além dessas atribuições, o processo de *Log* também deverá ser criptografado, de maneira tal que este seja protegido contra alterações de terceiros, podendo invalidar todo o processo de análise das imagens. A Figura 8 mostra a interface de registro de *Log* da ferramenta FI2.

4. Conclusões e Trabalhos futuros

Neste artigo é apresentado um *framework* para processamento de imagens voltado para inves-

tigação forense e suas ferramentas, em que a principal motivação é ausência de aplicações no mercado brasileiro desse padrão de utilização. A criação de uma interface com a tentativa de facilitar ao máximo a interação aplicação/usuário permite fazer comparações entre várias imagens, aplicar várias ferramentas, localizar facilmente as ferramentas e a geração de um *log* que permite a reprodução dos procedimentos aplicados, bem como a auditoria das atividades realizadas pela ferramenta. Apesar da falta de algumas ferramentas, foi observada a grande margem de aplicações que podem ser feitas com o software, principal-

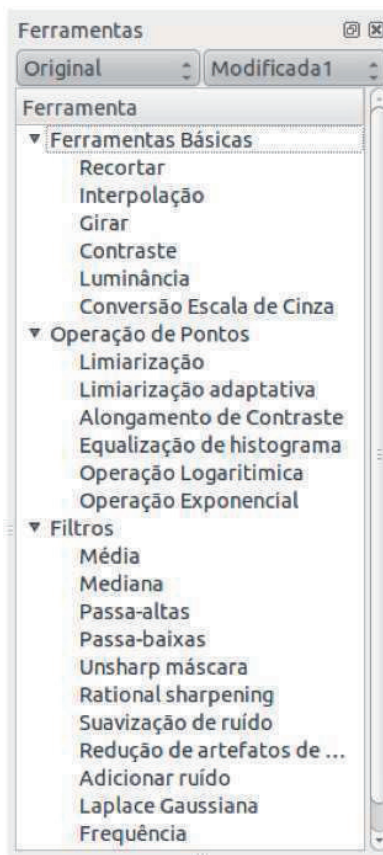


Figura 7. Painel de Seleção de Ferramentas.

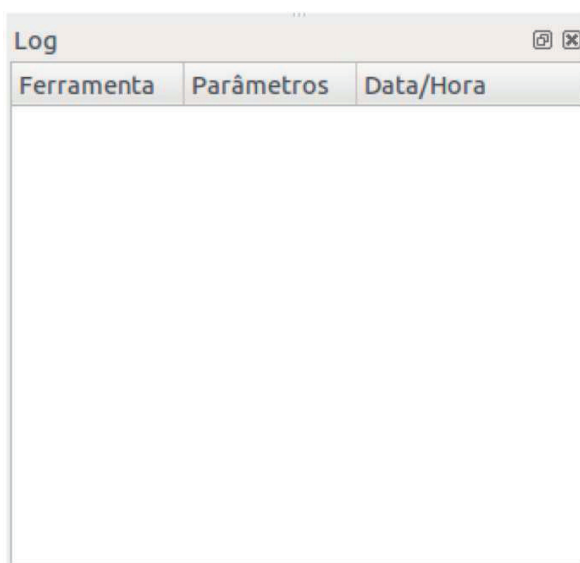


Figura 8. Interface de Registro de atividades - Log.

mente na área de investigação forense, onde foi possível verificar algumas aplicações com resultados relevantes.

Nosso objetivo principal com esta aplicação é a melhoria das investigações forenses, tanto em relação ao tempo quanto em qualidade, permitindo ao perito reduzir o tempo despendido e ao aperfeiçoamento da qualidade do trabalho com imagens. Um software que integre todas as ferramentas de análise forense de imagens com logs que permitam a reprodução fidedigna de todas as técnicas empregadas durante o exame pericial, bem como apresente uma interface prática e ágil na aplicação de suas ferramentas, possibilitará ao perito criminal realizar suas análises com a rapidez e reprodutibilidade necessárias a uma investigação técnica criminal.

Dentre as propostas para trabalhos futuros com este *framework* estão a incorporação de suporte a arquivos de vídeos para processamento com as ferramentas existentes, extração de *frames* específicos para processamento, reconhecimento facial em arquivo de imagens (comparação com retrato falado, imagens diversas) e extração automática de faces em vídeos (ex.: vídeos de segurança). Novos algoritmos que facilitem o trabalho pericial devem ser estudados e incorporados a esta plataforma com facilidade de uso que otimizem e documentem todas as análises realizadas.

Referências

- [1] R. Gonzales, R; Woods, *Digital Image Processing*, 2nd ed., 2002.
- [2] W. Garage. (2012, Maio). [Online]. Available: <http://opencv.willowgarage.com/wiki/Welcome>
- [3] Q. Nokia. (2012, maio). [Online]. Available: <http://qt.nokia.com/>
- [4] C. E. Duchon, "Lanczos filtering in one and two dimensions," *Journal of Applied Meteorology*, 1979.

Exames de Reconhecimento Facial na Polícia Federal

Gustavo H. M. de Arruda¹ e André L. da C. Morisson²

^{1,2}*Departamento de Polícia Federal*

¹*Superintendência Regional de Pernambuco
Av. Cais do Apolo, 321, 50030-230, Recife, PE, Brasil*

²*Instituto Nacional de Criminalística
SPO lote 7, Ed. INC, 70.675-607, Brasília, DF, Brasil
{gustavo.ghma, morisson.alcm}@dpf.gov.br*

RESUMO – A metodologia usada pelos peritos criminais federais nos exames de reconhecimento facial é definida pelo Instituto Nacional de Criminalística, responsável também por ações de capacitação, que contam com a presença inclusive de peritos criminais estaduais e do Distrito Federal. Neste artigo, procura-se verificar até que ponto a doutrina do Instituto Nacional de Criminalística segue as recomendações e orientações do FISWG (*Facial Identification Scientific Working Group*), cuja missão é apresentar metodologias, recomendações e manuais de boas práticas, que sejam consenso na área de identificação humana por meio de imagens..

1. INTRODUÇÃO

O exame de reconhecimento facial, dentro do escopo da perícia criminal, trata da comparação de duas faces presentes em registros visuais, ou de uma face em registro visual e um indivíduo conhecido, tendo por objetivo determinar se se trata da mesma pessoa. À face conhecida dá-se o nome de “face padrão”, e a face não conhecida é denominada “face questionada”.

Apesar da ampla bibliografia sobre sistemas biométricos de reconhecimento facial, são raras as publicações que abordam aspectos específicos do exame realizado por peritos criminais [1] [2] [3]. Tais exames são realizados de forma não automatizada, por meio de comparações individuais (um para um), e uma conclusão acerca da identidade da face questionada é emitida ao final.

Mesmo no estado da arte dos sistemas biométricos, ainda não há um método automatizado de reconhecimento facial que possa ser usado como prova pericial em um processo criminal [1]. Diferentemente de exames de DNA, a capacidade discriminatória dos métodos automatizados ainda não é suficiente para evitar erros [4], que em tais

situações, podem implicar na condenação de um inocente.

Por sua vez, a ausência de método científico na comparação facial é inaceitável, do ponto de vista pericial. Os métodos holísticos, assim definidos, trazem consigo um elevado grau de subjetividade, que não é compatível com a neutralidade que o exame pericial requer [5] [6].

Diante desse cenário, um grupo científico, organizado pelo *Federal Bureau of Investigation* (FBI), tem trabalhado no estabelecimento de padrões, manuais e requisitos, dentre outros, voltados para o exame forense de reconhecimento facial. O grupo, denominado *Facial Identification Scientific Working Group* (FISWG), conta com a participação de diversas agências americanas federais e estaduais, além de universidades e outras associações (Universidade John Hopkins, Laboratório Lincoln do MIT, Polícia Federal Australiana, Polícia Nacional Israelense, Instituto Forense da Holanda, dentre vários outros).

Neste artigo, procura-se estabelecer até que ponto a doutrina do Instituto Nacional de Criminalística (INC) da Polícia Federal segue as recomendações e orientações do FISWG, no que diz respeito aos exames periciais de reconhecimento facial. Para tanto, um breve resumo dos principais documentos publicados pelo FISWG até o momento são apresentados na seção 2. Em seguida, os principais aspectos da doutrina empregada na Polícia Federal são indicados, correlacionando-os com as recomendações do FISWG. Por fim, uma

análise qualitativa de laudos periciais de reconhecimento facial, recentemente elaborados, é discutida, e as considerações finais são apresentadas.

2. RECOMENDAÇÕES DO FISWG

O conjunto de documentos elaborados pelo FISWG trata de diversos aspectos relativos à área de reconhecimento facial. Para este artigo, os seguintes documentos serão considerados:

- 1) *Facial Comparison Overview* [7];
- 2) *Guidelines and Recommendations for Facial Comparison Training to Competency* [8];
- 3) *Guidelines for Facial Comparison Methods* [9];
- 4) *Recommendations for a Training Program in Facial Comparison* [10];
- 5) *Capture and Equipment Assessment for Facial Recognition Systems* [11].

2.1. CONSIDERAÇÕES INICIAIS

Os documentos do FISWG diferenciam o exame de reconhecimento facial (*facial examination*) do processo de análise facial (*facial review*). Ambos são procedimentos de comparação de faces (*facial comparison*), porém o exame de reconhecimento facial é realizado por profissionais especializados, usando métodos e procedimentos rigorosos.

Neste artigo, será adotada a nomenclatura “exame de reconhecimento facial” para indicar o exame realizado por peritos criminais, que tem por objetivo comparar duas faces, no intuito de determinar se pertencem ou não a um mesmo indivíduo, mesmo considerando não ser esta a nomenclatura mais adequada.

2.2. MÉTODOS

Segundo [9], os métodos de reconhecimento facial foram classificados em quatro tipos distintos:

- a) **Holísticos:** são métodos baseados na habilidade natural humana de classificar e reconhecer rostos, levando-se em conta todas as características visualizadas de maneira simultânea;

- b) **Morfológicos:** as características anatômicas das faces analisadas são identificadas e classificadas em relação à sua capacidade discriminatória, e uma conclusão é emitida com base em observações subjetivas;

- c) **Antropométricos:** exames realizados a partir de medidas sobre pontos cefalométricos nas faces comparadas, com a conclusão emitida a partir de limiares subjetivos para as diferenças observadas;

- d) **Por superposição:** por meio da criação de composições entre as faces questionada e padrão, devidamente alinhadas, no intuito de auxiliar a visualização das características observadas.

Conforme disposto em [9], em relação à utilização dos métodos para fins forenses, recomenda-se que:

- 1) Métodos holísticos sejam usados somente quando houver restrição de tempo, não sendo recomendado para realização de exames periciais;
- 2) Métodos morfológicos são considerados os métodos principais para realização de exames de comparação facial;
- 3) Métodos antropométricos não devem ser empregados em exames forenses, tanto de forma individual quanto conjunta a outros métodos, devido à incerteza provocada por condições não controladas, como ângulo de captura, iluminação, etc.;
- 4) Métodos por superposição podem ser utilizados apenas em conjunto com os métodos morfológicos.

2.3. CAPACITAÇÃO

A capacitação de pessoal também é abordada nos documentos emitidos pelo FISWG, que estabelece critérios e competências necessárias para o treinamento na área de reconhecimento facial. Em [10], são especificadas as diversas categorias de treinamento (nível introdutório, conhecimento de processos e relações, preparação e apresentação em juízo, etc.), bem como os níveis de avaliação (conhecimento, competência, proficiência e certificação). Tais especificações são então estabelecidas para a função desempenhada (gerência, supervisão, examinador, etc.). No presente caso, identifica-

se o cargo de perito criminal como tendo a função *facial image examiner*, o qual exige proficiência em todas as categorias, à exceção de instrutor.

Como disposto em [8], de maneira resumida, as seguintes competências são desejadas para peritos criminais:

- a) Conhecer a evolução histórica do exame pericial de reconhecimento facial;
- b) Compreender os princípios que regem a individualidade e a permanência, as diferenças entre classe e indivíduo, bem como as diferenças entre características estáveis e transitórias;
- c) Conhecimento amplo de princípios básicos de comparação facial, tais como avaliação da qualidade da face em uma imagem, processo ACE-V1, métodos de comparação, dentre outros;
- d) Conhecimento profundo acerca dos sistemas biométricos automatizados, sua operação e limitações, e aspectos gerais de algoritmos biométricos;
- e) Compreensão dos princípios que regem a representação por imagens, em processos fotográficos;

Por sua vez, os cursos de capacitação na área devem incluir os seguintes assuntos, quando voltados para a perícia criminal:

- a) Metodologia para obter conclusões a partir de comparações faciais;
- b) Limitações na representação da face por imagens, como distorções de perspectiva;
- c) Questões legais aplicáveis ao caso, como preparação e apresentação do material, e conhecimento de casos judiciais relevantes;

Além disso, deve-se demonstrar proficiência por meio de avaliação, e o curso deve ser especificamente voltado para a realização de exames de comparação facial.

2.4. EQUIPAMENTOS E PROCEDIMENTOS

Embora haja um conjunto de recomendações voltadas especificamente para equipamentos e

procedimentos de captura de imagens em sistemas biométricos [11], as especificações também trazem implicações na realização dos exames periciais.

Primeiramente, porque estabelecem especificações e parâmetros para a coleta de imagens que podem ser usadas em documentos de identificação. Assim, é possível estabelecer, eventualmente, se um determinado material questionado, oriundo de uma foto de documento, possui os requisitos mínimos necessários à consecução do exame.

Além disso, as especificações também são aplicáveis nas situações em que é necessário registrar as imagens da face padrão, interagindo diretamente com o indivíduo que terá sua face comparada à face questionada.

Esse último aspecto constitui um importante passo em exames periciais, usualmente denotado “coleta de material padrão”, aplicável em várias áreas de perícia, como fonética forense, balística, documentoscopia, dentre outras. Em geral, o êxito de um exame pericial, em que a coleta de padrão é necessária, está fortemente ligado à qualidade do material padrão obtido.

Dessa forma, fica clara a importância de se ter especificações de equipamentos e procedimentos relacionados à captura de imagens para comparação facial.

3. EXAMES NA POLÍCIA FEDERAL

Na Polícia Federal, a competência para realização dos exames de reconhecimento facial é dos peritos criminais federais, por se tratar de perícia oficial de natureza criminal [12].

Por questões históricas, além da afinidade com a área de Processamento Digital de Imagens e da realização frequente de exames em material audiovisual, os exames de reconhecimento facial têm sido majoritariamente realizados por peritos com formação em Engenharia Elétrica/Eletrônica. No entanto, não há restrições quanto à formação acadêmica para realização de tais exames.

¹ Do inglês, *Analysis, Comparison, Evaluation and Verification*

Na Polícia Federal do Brasil, o Serviço de Perícias em Audiovisual e Eletrônicos (SEPAEL) do INC tem a atribuição de propor normatizações e ações de capacitação na área. Até a data de elaboração deste artigo, foram ministrados três cursos de Reconhecimento Facial (2009, 2010 e 2012), com público alvo composto por peritos criminais federais, estaduais e do Distrito Federal.

3.1. ESTRUTURA DO CURSO

O curso de capacitação em Reconhecimento Facial está voltado para profissionais que realizam exames periciais de natureza criminal. Tais exames envolvem a comparação de imagens de faces ou indivíduos desconhecidos com um indivíduo conhecido, ou com imagens de sua face.

O curso tem carga horária de 80h e, em geral, é realizado em turmas com 20 participantes.

A estrutura do curso está dividida em três módulos distintos: Processamento Digital de Imagens (PDI); Anatomia Facial Forense (AFF); e Exame de Reconhecimento Facial (ERF). O detalhamento dos módulos é apresentado a seguir.

3.1.1. PROCESSAMENTO DIGITAL DE IMAGENS

O módulo de PDI tem como objetivo principal desenvolver habilidades com o tratamento de imagens estáticas digitais. Espera-se que, ao final do módulo, o perito criminal possua as seguintes competências e habilidades:

- Conhecer os fundamentos matemáticos da representação de imagens numéricas;
- Compreender os conceitos relativos à Transformada de Fourier Espacial aplicada em imagens digitais;
- Aplicar filtros no domínio espacial e frequencial em imagens digitais, compreendendo os efeitos de filtros passa-baixas, passa-altas, passa-faixa, notch, e outros;
- Compreender e aplicar técnicas de ampliação de contraste, equalização de histograma, filtragem homomórfica, dentre outras; e

- Compreender o processo de degradação de uma imagem por ruído e por meio da PSF (*Point Spread Function*), bem como técnicas de melhoria e tratamento;

3.1.2. ANATOMIA FACIAL FORENSE

O módulo de AFF está voltado para a análise detalhada da anatomia da face, as principais nomenclaturas e a metodologia de descrição das estruturas da face. Ao final deste módulo, as seguintes competências e habilidades são esperadas:

- Conhecer e identificar as diversas estruturas envolvidas na descrição e análise da face: estrutura básica da cabeça; classificações antropológicas da face; osteologia básica associada ao estudo da cabeça; e
- Compreender os aspectos e consequências do envelhecimento na face e suas estruturas (boca, orelha, sulcos, linhas de expressão, etc.);

3.1.3. EXAME DE RECONHECIMENTO FACIAL

O módulo de ERF é responsável por agrupar as ferramentas e métodos apresentados nos módulos de PDI e AFF, de maneira sistemática, inserindo-os na metodologia do exame a ser realizado pelo perito criminal. É particularmente neste módulo que a doutrina é estabelecida, a partir da definição da metodologia do exame, sendo discutidos ainda outros aspectos importantes relacionados.

Atualmente, a doutrina é baseada em método morfológico, sugerindo também, sempre que o material assim o permitir, uma análise complementar por superposição. Embora a análise por cefalometria seja apresentada no módulo de AFF, a utilização dos índices cefalométricos como ferramenta auxiliar foi descartada no caderno didático mais recente [13].

No método morfológico recomendado, a face é dividida em oito regiões, conforme mostrado na Figura 1. Em cada região, são apontadas as características principais a serem observadas, além de outras julgadas relevantes pelo perito criminal.

Em seguida, as características anotadas são classificadas como atributos de semelhança ou

divergência, sendo atribuído um dos seguintes graus: semelhante em detalhes, semelhante, sem comparação, diferente, diferente em detalhes. As características são ainda classificadas em relação à sua capacidade discriminatória, podendo ser: forte, moderada ou fraca, conforme análise subjetiva do perito criminal.

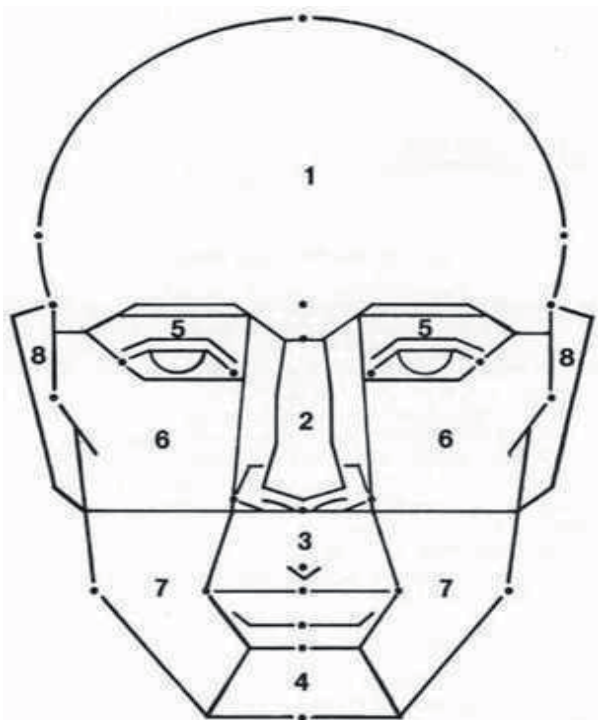


Figura 1 – Divisão da face em regiões: 1. frontal; 2. nasal; 3. labial; 4. mentoniana; 5. orbital; 6. zigomaxilar; 7. bucomandibular; e 8. auricular.

A metodologia baseada em superposição de imagens também é apresentada, porém de forma complementar ao exame. Sua utilização é recomendada apenas para ilustrar eventuais convergências ou divergências importantes, sendo a conclusão emitida com base nos resultados do método morfológico.

Além da definição da metodologia, outros aspectos também são abordados neste módulo, tais como:

- Verificação da viabilidade do exame, por meio das características do material questionado;
- Equipamentos e metodologia para coleta de imagens para comparação, incluindo sugestão de ambiente adequado e ajustes no equipamento; e

- Realização de procedimentos preparatórios para o exame, como digitalização de registros analógicos, correção de alinhamento e escala e ajustes de contraste/brilho.

Em relação à análise de viabilidade do exame, é importante destacar que havia uma recomendação de não se realizar o exame quando uma das faces (padrão ou questionada) apresentasse uma distância entre os centros dos olhos inferior a 40 pixels. Tal recomendação teria se originado nos valores mínimos possíveis para algoritmos biométricos, que se encontram na faixa de 40 a 60 pixels de distância entre os centros dos olhos [11].

No entanto, esta recomendação não se mostrou apropriada, uma vez que mesmo imagens com melhor resolução apresentavam outros fatores de degradação que impediam a realização do exame, tais como foco, distorção de lente, artefatos JPEG, dentre outros. Além disso, grande parte da casuística é voltada para análise de fotografias impressas em documentos e fichas criminais, de modo que a recomendação não seria aplicável.

No caderno didático mais recente, a análise de viabilidade é realizada de forma mais abrangente, abordando aspectos como definição da face questionada, oclusão ou obstrução parcial da face e impacto dos fatores de degradação na qualidade do exame.

4. ADEQUAÇÃO ÀS RECOMENDAÇÕES DO FISWG

Diante do que foi observado nas seções 2 e 3, passa-se a discutir as adequações da doutrina estabelecida na Perícia Criminal da Polícia Federal, face às recomendações do FISWG.

Inicialmente, verifica-se que o principal aspecto relativo aos exames de reconhecimento facial é o método a ser empregado. Desde as primeiras edições, o procedimento envolvia a análise da morfologia da face, o cálculo de índices cefalométricos e a sobreposição de imagens, quando as condições de alinhamento permitiam. O resultado seria conclusivo se fossem identificados aspectos morfológicos suficientemente discriminatórios nas faces analisadas.

A sobreposição de imagens e o cálculo de índices cefalométricos eram considerados exames complementares, que também embasavam a conclusão do laudo pericial. Essas eram recomendações contidas nas duas primeiras edições do curso, e logo surgiram indagações dos próprios peritos criminais acerca da confiabilidade do método ou de sua capacidade discriminatória.

Estudos recentes acabaram por mostrar que a análise por índices cefalométricos em imagens (ou foto-antropometria) não é confiável, mesmo em situações ideais [2] [3]. Assim, o cálculo de índices cefalométricos não é mais recomendado em [13], e tal orientação está em processo de divulgação entre os peritos criminais das primeiras turmas.

Em relação à capacitação e treinamento dos peritos criminais federais, ainda não há uma previsão de inclusão da matéria no Curso de Formação Profissional, haja vista se tratar de perícia geralmente associada à área de exames em material audiovisual. No entanto, a Academia Nacional de Polícia (ANP), o INC e o SEPAEL têm envidado esforços para garantir uma periodicidade anual do Curso de Reconhecimento Facial, sempre destinando vagas para peritos criminais dos estados e Distrito Federal.

A estrutura do curso atende, de um modo geral, às recomendações elencadas nos documentos emitidos pelo FISWG (vide seção 2).

5. LAUDOS PERICIAIS

A fim de verificar aspectos pontuais da metodologia empregada pelos peritos criminais federais, foram coletados dados de laudos periciais elaborados entre 2011 e 2012, cadastrados como não-sigilosos no Sistema Nacional de Criminalística² (SisCrim).

²O Sistema Nacional de Criminalística é um sistema de informação utilizado pela Perícia Criminal Federal para a gestão da Criminalística, e possui em sua base de dados todos os laudos periciais produzidos no DPF desde 2006, sendo acessíveis aos peritos criminais federais, desde que classificados como não sigilosos.

Inicialmente, foram anotadas as quantidades de laudos classificados como relativos aos exames de reconhecimento facial, que no SisCrim podem ser:

- a) Exame de Reconhecimento Facial (44);
- b) Exame de Reconhecimento de Indivíduo por Imagens (102);
- c) Exame Biométrico: Reconhecimento Facial (1);

Constatou-se que, entre janeiro/2011 e junho/2012 foram emitidos 147 laudos periciais relativos ao exame de reconhecimento facial.

Em seguida, em uma análise mais aprofundada, uma amostra desses laudos periciais foi analisada, com vistas a responder às seguintes questões:

- a) A metodologia é baseada em um método morfológico?
- b) Há a utilização de método baseado no cálculo de índices cefalométricos?
- c) Qual o tipo de material questionado analisado?
- d) Houve coleta de material padrão?
- e) O resultado é conclusivo³ ou parcialmente conclusivo⁴?
- f) Foi realizada análise de adequabilidade do material questionado?
- g) Foram empregados métodos de comparação por superposição?
- h) O exame tem resultado conclusivo?

Foi analisada uma amostra de 38 laudos periciais, sendo preenchida uma planilha contendo as respostas às questões elencadas. A partir do cruzamento dos dados obtidos, os seguintes resultados foram observados:

³ Entende-se por resultado conclusivo quando é possível apontar com elevado grau de certeza ou excluir a identidade da face questionada, quando comparada à face padrão.

⁴ O resultado é parcialmente conclusivo quando é possível excluir a face questionada comparada com vários padrões, exceto para pelo menos uma face padrão, onde o resultado é inconclusivo.

- a) Exames realizados em imagens de CFTV foram conclusivos em apenas 3 de 26 laudos examinados (11,5%);
- b) Quando o exame é realizado em fotografias de documentos (padrão 3 x 4cm), dos 6 laudos analisados, todos foram conclusivos (100%);
- c) Nos demais casos, que envolvem fotografias e filmagens diversas, de 6 laudos avaliados, 3 apresentaram resultados conclusivos (50%);
- d) Em apenas 4 dos 38 laudos examinados foi realizada a coleta de padrão de suspeitos (10,5%); Destes, apenas 1 apresentou resultado conclusivo;
- e) O uso de índices cefalométricos em conjunto com os exames morfológicos foi observado em 7 laudos, dos quais 6 apresentaram resultados conclusivos (85,7%), quando combinados ao exame morfológico;
- f) A maioria dos laudos em fotografia 3x4cm também levou em conta a análise por índices cefalométricos (4 laudos, representando 66,7%);

Uma importante constatação é a clara necessidade de se orientar os diversos órgãos públicos sobre a importância do projeto de sistemas de CFTV adequados. Na maioria das vezes, as imagens dos indivíduos investigados apresentam péssima qualidade, mesmo para uso de sistemas biométricos, que normalmente requerem de 40 a 60 pixels de distância entre os centros dos olhos. Outros fatores também contribuem para a baixa qualidade das imagens, tais como posicionamento da câmera, iluminação, distorção de lente, dentre outros. Exemplos típicos são mostrados na Figura 2.

Quando a análise é realizada em fotografias de documentos, aumentam significativamente as chances de resultados conclusivos, uma vez que se tem melhor acesso às estruturas anatômicas da face. Porém ainda há alguns fatores internos e externos que inviabilizam a realização do exame. Os principais fatores internos estão relacionados ao encaminhamento de material padrão ou questionado na forma de fotocópia preto-e-

branca ou impressão de baixa qualidade, ou ainda à coleta inadequada de padrão por policiais não habilitados. Tais problemas são normalmente sanados por meio de novas solicitações, para que o material adequado seja encaminhado, ou para que seja realizada nova coleta. Por sua vez, os fatores externos estão associados a aspectos da qualidade da fotografia para documentos ou fichas criminais, tais como posicionamento da face, oclusão das orelhas e expressão facial, que fogem ao controle da perícia criminal.

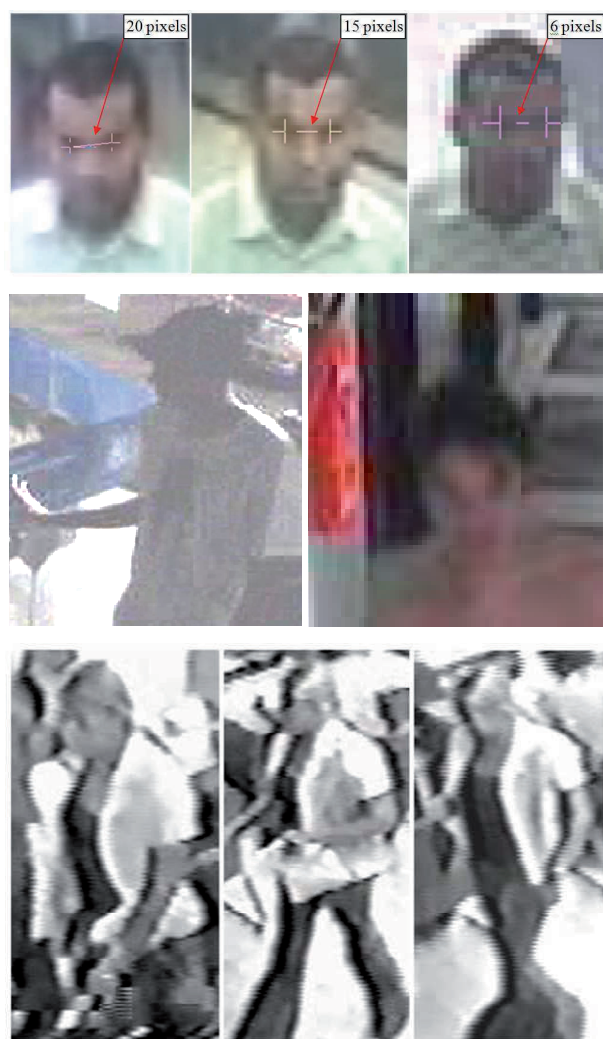


Figura 2 – Exemplos típicos de imagens de suspeitos capturadas por sistemas de CFTV, e encaminhadas para realização de exame de reconhecimento facial.

Mesmo assim, é importante divulgar entre as unidades que fazem o registro fotográfico da identificação criminal, institutos de identificação e demais órgãos, material didático informando da necessidade de pessoal treinado para realizar o procedimento, além de

estabelecer critérios, condições e métodos adequados para o registro fotográfico da face. Em um cenário ideal, a fotografia seria realizada por equipamento dedicado, com controle de iluminação, foco e enquadramento, dotado ainda de algoritmo de análise qualitativa da fotografia obtida e formato adequado, como o JPEG 2000.

Também se mostra necessário realizar ações de divulgação entre os peritos criminais federais acerca dos riscos em se incluir a análise por índices cefalométricos.

6. CONCLUSÃO

Os exames de comparação de faces realizados na Polícia Federal estão de acordo com os padrões recomendados pelo *Facial Identification Scientific Working Group*, que estabelece, dentre outros, recomendações e boas práticas internacionais para realização de exames periciais na área.

A análise de laudos periciais de reconhecimento facial realizados por peritos criminais federais mostrou que praticamente todos os casos com imagens oriundas de CFTV não resultaram em exames conclusivos. Por outro lado, os exames realizados em fotografias de documentos ou fichas criminais apresentam elevada probabilidade de resultado conclusivo. No entanto, ainda se constata a necessidade de padronizar procedimentos e realizar treinamentos para os profissionais responsáveis pelo registro fotográfico de indivíduos, ou empregar equipamentos dedicados a esse fim, com algoritmo de controle de qualidade das fotografias obtidas.

REFERÊNCIAS

- [1] Tauseef Ali, Raymond Veldhuis, and Luuk Spreeuwens, "Forensic Face Recognition: A Survey," Centre for Telematics and Information Technology (CTIT), University of Twente, Enschede, The Netherlands, Internal Report TR-CTIT-10-40 2010.
- [2] Krista F. Kleinberg, Peter Vanezis, and A. Mike Burton, "Failure of Anthropometry as a Facial Identification Technique Using High-Quality Photographs," *Journal of Forensic Science*, vol. 52, no. 4, pp. 779-783, 2007.
- [3] R. Moreton and J. Morley, "Investigation into the use of photoanthropometry in facial image comparison.," *Forensic Science International*, vol. 212, no. 1, pp. 231-237, 2011.
- [4] Anil K. Jain, Arun Ross, and Salil Prabhakar, "An Introduction to Biometric Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4-20, Jan 2004.
- [5] BRASIL, Recurso Extraordinário 363.889/DF. Ministro Dias Toffoli, voto do Ministro Luiz Fux, junho 02, 2011.
- [6] USA, *Daubert v. Merrell Dow Pharmaceuticals*. United States Supreme Court., 1993.
- [7] FISWG. (2010, June) Facial Identification Scientific Working Group. [Online].
http://www.fiswg.org/FISWG_Facial_Comparison_Overview_v1.0_2010.04.29.pdf
- [8] FISWG. (2010, Nov.) Facial Identification Scientific Working Group. [Online].
http://www.fiswg.org/FISWG_Training_Guidelines_Recommendations_v1.1_2010_11_18.pdf
- [9] FISWG. (2012, Feb.) Facial Identification Scientific Working Group. [Online].
http://www.fiswg.org/FISWG_GuidelinesforFacialComparisonMethods_v1.0_2012_02_02.pdf
- [10] FISWG. (2012, Feb.) Facial Identification Scientific Working Group. [Online].
http://www.fiswg.org/FISWG_RecommendationsForTrainingProgram_v1.0_2012_02_02.pdf
- [11] FISWG. (2011, Apr.) Facial Identification Scientific Working Group. [Online].
http://www.fiswg.org/FISWG_CaptureAndEquipmentAssessmentForFRSystems_v1.0_2011_05_05.pdf
- [12] BRASIL, Lei 12.030, de 17 de setembro de 2009.
- [13] Carlos E. P. Machado, Eduardo M. Lacerda Filho, Gustavo H. M. de Arruda, and Paulo M. G. I. Reis, *Reconhecimento Facial*, 2012.
- [14] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399-458, 2003.
- [15] A. S. Tolba, A.H. El-Baz, and A.A. El-Harby, "Face Recognition: A Literature Review," vol. 2, no. 2, pp. 88-103, 2006.
- [16] M. Evison et al., "Key parameters of face shape variation in 3D in a large sample," *Journal of Forensic Science*, vol. 55, no. 1, pp. 159-162, 2010.

Estimativa de Altura de Indivíduos em Imagens de Perspectiva Cônica com Distorções Esféricas (Caso Prático)

Luiz Vinícius Gontijo Laborda Larrain

Gerência de Perícias em Áudio e Vídeo, Perícia Oficial e Identificação Técnica do estado de Mato Grosso (POLITEC)
Av. Gonçalves Antunes de Barros, 3245, Bairro Novo Mato Grosso, Cuiabá, Mato Grosso, Brasil

luizlarrain@seguranca.mt.gov.br

RESUMO – O emprego de técnicas de fotogrametria possui aplicação em diversos casos de interesse forense. Apresenta-se neste artigo um caso prático em que é utilizado um algoritmo iterativo para a correção das distorções esféricas e é estimada na sequência a altura de um indivíduo conhecendo-se referências verticais na imagem. Utiliza-se para isto a conservação da razão cruzada entre os objetos na perspectiva cônica, empregando se necessário como ponto de referência auxiliar o ponto de fuga vertical. Esta técnica demonstra ser uma alternativa possível diante de impossibilidade de superposições de gabaritos métricos.

1. INTRODUÇÃO

Em diversas situações de interesse forense a imagem é analisada não só de forma qualitativa, mas também quantitativa, empregando-se técnicas fotogramétricas para extrair informações de medidas e distâncias relativas entre diferentes objetos que compõem a cena.

Para os registros que não oferecem condições para identificação dos indivíduos através dos exames de reconhecimento facial, seja por limitações qualitativas destes ou mesmo pela oclusão da face, a estimativa da altura dos indivíduos pode ter relevância significativa reduzindo possivelmente o universo dos suspeitos.

Em imagens sob a perspectiva cônica, nos casos em que o sistema de captura encontra-se indisponível para a equipe forense, impossibilitando assim a utilização de gabaritos métricos, pode-se utilizar a conservação da razão cruzada entre as medidas de dois outros objetos conhecidos na imagem para estimar uma medida questionada.

Dentre as dificuldades para este tipo de medição encontra-se a presença de distorções geométricas, principalmente do tipo “barril” ou

“almofada”, provocando a curvatura das retas presentes na imagem. Na prática a identificação de duas referências verticais conhecidas na imagem, a fim de ser explorada a conservação da razão cruzada sob a perspectiva cônica, também se torna um desafio.

No caso de só se ter disponível apenas uma referência vertical uma alternativa é utilizar como ponto de referência auxiliar o ponto de fuga vertical do objeto em análise e da referência utilizada.

Este artigo apresentará um caso prático da criminalística do estado de Mato Grosso em que fora estimada a altura de um indivíduo em imagens contendo distorções esféricas através da conservação da razão cruzada entre os objetos na perspectiva cônica.

2. CORREÇÃO DE DISTORÇÕES ESFÉRICAS

As distorções esféricas são aberrações ópticas frequentemente encontradas nas imagens oriundas de sistemas de CFTV (Circuito Fechado de Televisão). Explica-se esta ocorrência devido à utilização de lentes com ângulos acentuados, comumente empregadas nos conjuntos ópticos destes equipamentos de captura. A utilização dos conjuntos ópticos desta natureza visa proporcionar maior área de cobertura das imagens, característica geralmente desejada nestes sistemas de monitoramento e segurança.

Para viabilizar a realização de medições através do emprego das técnicas fotogramétricas, as distorções esféricas presentes nas imagens devem ser amenizadas.

Pode-se constatar na imagem questionada, exibida na figura 1, a presença da distorção do tipo “barril”, evidenciada através da curvatura aparente presente na coluna da edificação (indicada pela seta vermelha) localizada na região esquerda da imagem.



FIGURA 1 – DISTORÇÃO DO TIPO “BARRIL” NA IMAGEM QUESTIONADA.

2.1. MODELO DE CONRADY-BROWN

As distorções esféricas podem ser reproduzidas segundo o modelo clássico de distorções de lentes, desenvolvido por Conrady [1] e Brown [2], [3]. A expressão do modelo de Brown [4] é:

$$x_u = x_d + (x_d - x_c)(K_1 r^2 + K_2 r^4 + \dots) + (P_1(r^2 + 2(x_d - x_c)^2) + 2P_2(x_d - x_c)(y_d - y_c))(1 + P_3 r^2 + \dots)$$

$$y_u = y_d + (y_d - y_c)(K_1 r^2 + K_2 r^4 + \dots) + (P_2(r^2 + 2(y_d - y_c)^2) + 2P_1(x_d - x_c)(y_d - y_c))(1 + P_3 r^2 + \dots)$$

Os parâmetros K_n referem-se às distorções radiais enquanto os parâmetros P_n às distorções tangenciais [4].

Na prática a representação das distorções do tipo “barril” ou “almofada” pode ser satisfatoriamente modelada considerando apenas as distorções radiais (K_1 e K_2).

2.2. APLICAÇÃO DO ALGORITMO PARA A CORREÇÃO DAS DISTORÇÕES ESFÉRICAS

Para a correção da distorção do tipo “barril” presente na imagem questionada foi utilizado o plugin Lens_unDistortionSMT [5], exibido na figura 2, que implementa a versão simplificada do modelo de Brown (considerando apenas as distorções radiais). Este plugin foi desenvolvido para execução no software de análise e processamento de imagens ImageJ [6].

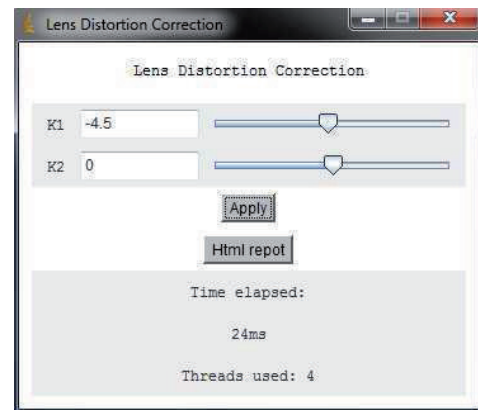


FIGURA 2 – INTERFACE DO PLUGIN LENS_UNDISTORTIONSMT

Dentre as vantagens para a utilização deste plugin, destaca-se além da sua disponibilidade gratuita, tal como o software ImageJ, a sua interatividade. Após escolher um referencial para o ajuste na imagem questionada, pode-se facilmente, através de uma análise perceptiva, amenizar a distorção esférica presente na imagem ajustando os parâmetros K_1 e K_2 . O plugin também permite a visualização instantânea do resultado obtido após aplicação dos parâmetros.

Segue na figura 3 o resultado da imagem corrigida. Evidencia-se através da reta na cor vermelha a correção tomando como referência a coluna da edificação localizada a esquerda da imagem.



FIGURA 3 – IMAGEM COM AS DISTORÇÕES ESFÉRICAS AMENIZADAS.

3. FOTOGRAMETRIA

A seguir serão descritos passo-a-passo a aplicação de técnicas fotogramétricas com o intuito de estimar a altura do indivíduo questionado nas imagens em análise.

O conteúdo contido no item 3.1 fora extraído e adaptado da apostila de Fundamentos de Processamento Digital de Imagens, do Instituto Nacional de Criminalística [7].

3.1. PERSPECTIVA CÔNICA E A RAZÃO CRUZADA

Ao visualizar uma fotografia, imagem dotada de apenas duas dimensões, o cérebro humano tem a capacidade de interpretar de forma consistente as relações de distâncias dos objetos em três dimensões, tal como a cena real capturada.

As linhas retas paralelas na cena real são convertidas, no plano imagem, em linhas retas que convergem em pontos denominados “pontos de fuga”. Essa é uma característica da perspectiva linear cônica, que é de fato a perspectiva de projeção do plano real no plano imagem.

Na perspectiva cônica, dois ou mais pontos de fuga determinam uma “linha de fuga”, ou seja, uma linha reta que corresponde ao encontro de infinitos planos paralelos.

É fato conhecido da geometria projetiva que, em uma imagem sob perspectiva cônica, **preserva-se** o que denominamos de **razão cruzada**.

A razão cruzada pode então ser utilizada nessas situações como uma forma se medir a dimensão de objetos em uma imagem a partir da informação conhecida da dimensão de outros objetos também contidos na mesma imagem.

3.1.1. ESTIMATIVA DE ALTURA QUESTIONADA A PARTIR DE DUAS REFERÊNCIAS VERTICAIS

Apresenta-se na figura 4 um diagrama traçado em perspectiva cônica. Nessa ilustração, observa-se a existência dos pontos de fuga “P1” e “P2”, formados pelo encontro do padrão quadriculado do plano horizontal. Os pontos “P1” e “P2” definem, portanto, a linha de fuga horizontal.

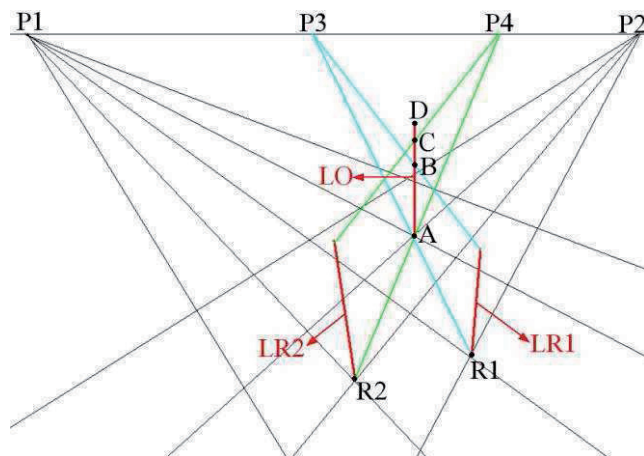


FIGURA 4 – DIAGRAMA EM PERSPECTIVA CÔNICA

Observa-se na figura 4, a presença de um objeto de dimensão desconhecida (LO) e dois elementos de referência, denominados “R1” e “R2”, todos localizados na posição vertical sobre o solo. Supondo que as dimensões das referências sejam conhecidas, LR1 e LR2, é possível obter a dimensão do objeto (LO) utilizando-se a conservação da razão cruzada. A partir da base de cada uma das referências, “R1” e “R2”, é traçada uma reta que passa também pela base do objeto, identificado pelo ponto “A”, determinando assim sobre a linha de fuga os pontos “P3” e “P4” respectivamente.

A partir de “P3” e “P4” são traçadas retas que passam pelo topo de cada uma das referências, determinando sobre o objeto os pontos “B” e “C” respectivamente. Como as retas que convergem para o mesmo ponto na linha de fuga horizontal são paralelas ao plano horizontal, os segmentos de reta AB e AC correspondem à segmentos obtidos pela projeção de pontos com dimensões iguais às de “R1” e “R2”, se marcados sobre o objeto a partir do solo.

Dessa forma pode-se determinar o comprimento real do objeto, a partir da propriedade da preservação da razão cruzada na perspectiva cônica:

$$\frac{AC}{AD} \bigg/ \frac{BC}{BD} = \frac{LR2}{LO} \bigg/ \frac{(LR2 - LR1)}{(LO - LR1)} = K$$

O valor da constante “K” pode ser obtido diretamente da imagem a partir dos segmentos, em pixels, AC, AD, BC e BD. Nesse caso a

dimensão do objeto poderá ser obtida diretamente por:

$$LO = \frac{LR2 \cdot LR1}{LR2 + K (LR1 - LR2)}$$

3.1.2. ESTIMATIVA DE ALTURA QUESTIONADA UTILIZANDO-SE APENAS UMA REFERÊNCIA VERTICAL

Nas situações em que há apenas um objeto vertical de referência com dimensões conhecidas na imagem, pode-se utilizar como ponto de referência auxiliar o ponto de fuga vertical “I” formado pelo objeto e a referência.

Apresenta-se na figura 5 um diagrama traçado em perspectiva cônica. É possível ver os pontos de fuga “P1” e “P2” e o ponto de fuga vertical “I”, formado pelo encontro das retas verticais traçadas a partir da base do objeto e da referência.

Os pontos “P4”, “A” e “C” são definidos da mesma maneira que descrito no subitem 3.1.1.

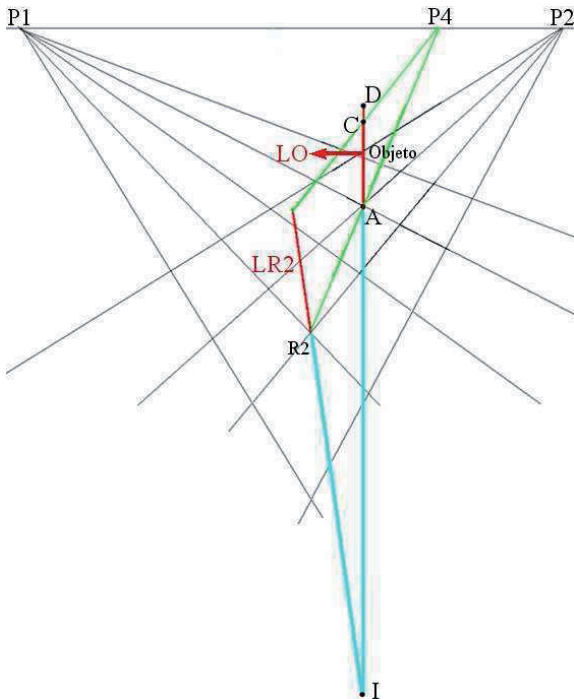


FIGURA 5 – LINHA DE FUGA E O PONTO DE FUGA VERTICAL

Nesse caso, também através da conservação da razão cruzada, com perdão a falta de rigor matemático na notação do limite ao infinito, encontra-se a seguinte relação entre os segmentos de reta:

$$\frac{AC}{AD} \Big/ \frac{IC}{ID} = \frac{LR2}{LO} \Big/ \frac{(\infty)}{(\infty)} = K$$

Assim pode-se calcular novamente a constante K diretamente da imagem, e a dimensão do objeto é estimada em:

$$LO = \frac{LR2}{K}$$

3.2. ESTIMATIVA DO DESVIO DA ALTURA DO OBJETO QUESTIONADO

Para se estimar adequadamente a altura do objeto questionado, de modo a minimizar o erro em relação à medida real, recomenda-se a repetição completa deste procedimento por no mínimo dez vezes, calculando o valor médio para a constante K e seu desvio padrão de acordo com as expressões [8]:

$$\hat{K} = \frac{1}{N} (K_1 + K_2 + \dots + K_i + \dots + K_n)$$

$$\hat{\sigma}_k = \sqrt{\frac{(K_1 - K_{est})^2 + \dots + (K_n - K_{est})^2}{N - 1}}$$

Assim, pode-se estabelecer o desvio da altura do objeto questionado (LO) através da propagação de erros, exibida na tabela 1 [8]:

TABELA 1
PROPAGAÇÃO DE ERROS

Relação funcional	Erro propagado
$\bar{R} = R(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n)$	$\sigma_{\bar{R}} = \sqrt{\left(\frac{\partial R}{\partial a_1}\right)^2 \sigma_{a_1}^2 + \left(\frac{\partial R}{\partial a_2}\right)^2 \sigma_{a_2}^2 + \dots + \left(\frac{\partial R}{\partial a_n}\right)^2 \sigma_{a_n}^2}$
$\bar{R} = \bar{a} \pm \bar{b}$	$(\sigma_{\bar{R}})^2 = (\sigma_{\bar{a}})^2 + (\sigma_{\bar{b}})^2$
$\bar{R} = \bar{a} \bar{b}$ ou $\bar{R} = \frac{\bar{a}}{\bar{b}}$	$\left(\frac{\sigma_{\bar{R}}}{\bar{R}}\right)^2 = \left(\frac{\sigma_{\bar{a}}}{\bar{a}}\right)^2 + \left(\frac{\sigma_{\bar{b}}}{\bar{b}}\right)^2$
$\bar{R} = \bar{a}^r$	$\frac{\sigma_{\bar{R}}}{\bar{R}} = r \frac{\sigma_{\bar{a}}}{\bar{a}}$
$\bar{R} = \ln \bar{a}$	$\sigma_{\bar{R}} = \frac{\sigma_{\bar{a}}}{\bar{a}}$
$\bar{R} = e^{\bar{a}}$	$\frac{\sigma_{\bar{R}}}{\bar{R}} = \sigma_{\bar{a}}$

3.3. APLICAÇÃO EM CASO PRÁTICO

Inicialmente foi selecionado o quadro mais favorável para a estimativa da altura do indivíduo questionado. A postura e a posição do indivíduo na imagem foram os parâmetros avaliados para a escolha deste quadro, uma vez que a região localizada ao centro da imagem é a área menos afetada pelas distorções esféricas.

Em uma primeira análise fora realizada a estimativa da altura do indivíduo questionado (LO) utilizando-se apenas uma referência vertical na imagem (LR2), sendo apresentados os resultados obtidos em 3.3.1.

Posteriormente observou-se a possibilidade de utilizar uma segunda referência vertical (LR1). De modo a verificar a metodologia já utilizada foi inicialmente estimada a altura desta segunda referência vertical (LR1) utilizando a mesma metodologia utilizada no item 3.3.1, sendo apresentados os resultados obtidos em 3.3.2.

E finalmente apresenta-se no item 3.3.3 o cálculo da altura do indivíduo questionado utilizando-se as duas referências verticais encontradas na imagem (LR1 e LR2).

3.3.1. ESTIMATIVA DA ALTURA QUESTIONADA (LO) UTILIZANDO-SE APENAS UMA REFERÊNCIA VERTICAL (LR2)

A medida utilizada inicialmente como referência neste caso foi a altura da faixa branca presente no poste localizado na região direita do vídeo, conforme exibido na figura 6. Foi constatado pelo perito no local, utilizando uma trena de alta precisão, que a altura desta faixa branca em relação ao solo é de 210,50cm.



FIGURA 6 – COMPRIMENTO DA REFERÊNCIA UTILIZADA

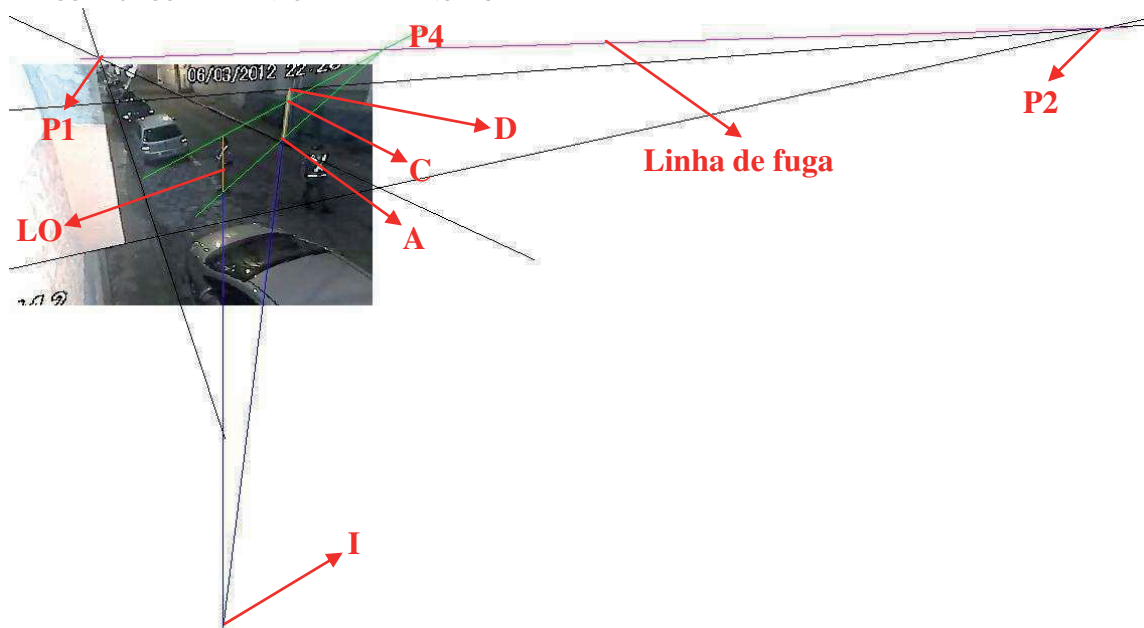


FIGURA 8 – DIAGRAMA CONTENDO A LINHA DE FUGA E O PONTO DE FUGA VERTICAL

O ponto de fuga "P1" foi definido através do encontro das retas traçadas a partir das extremidades da via enquanto que o ponto de fuga "P2" foi definido através do encontro das retas traçadas a partir da base da coluna da edificação e da calçada localizada a direita da imagem, conforme exibido nas figuras 7 e 8.

Realizou-se este procedimento dez vezes, utilizando-se em cinco destes a imagem tratada pelo filtro "Find Edges", com intuito de realçar as bordas dos objetos, conforme demonstrado na figura 7.

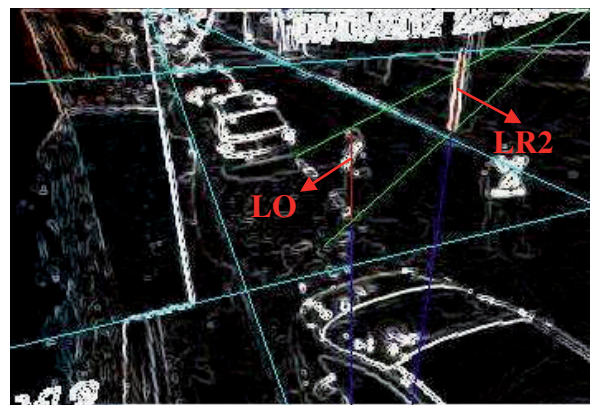


FIGURA 7 – IMAGEM COM AS BORDAS REALÇADAS

Segue na figura 8 a aplicação da técnica fotogramétrica descrita em 3.1.2, para o caso em análise. A linha na cor rosa é a linha de fuga e o ponto formado pelo encontro das duas retas verticais na cor azul é o ponto de fuga vertical "I". Também encontra-se explicitados na figura 8 os pontos "A", "C" e "D".

Seguem na tabela 2 os resultados obtidos nas dez medições realizadas.

TABELA 2
RESULTADOS DAS MEDIÇÕES REALIZADAS

Medições	AC (px)	AD (px)	IC (px)	ID (px)	K
1ª	37,056	48,125	495,571	506,794	0,787433
2ª	39,485	50,653	456,155	467,327	0,798611
3ª	36,878	49,645	453,198	466,127	0,764026
4ª	38,468	50,629	452,656	464,713	0,78004
5ª	37,381	49,77	504,744	517,041	0,769373
6ª	37,877	50,323	462,213	474,54	0,772751
7ª	38,823	49,387	512,73	523,562	0,802705
8ª	37,957	49,896	393,188	405,175	0,783914
9ª	37,816	50,364	462,077	474,734	0,771421
10ª	36,528	48,87	524,503	536,627	0,76473

Dessa forma, foram obtidos os resultados exibidos na tabela 3.

TABELA 3
RESULTADOS OBTIDOS

$\hat{K}$	0,779500388
$\hat{\sigma}_k$	0,013570273
$\hat{\sigma}_{LO}$	2,856808391

Dessa forma, a altura estimada do indivíduo questionado, na ocasião em análise, utilizando apenas uma referência vertical (LR2), é aproximadamente:

$$LO = 164,08 \pm 2,86 \text{ (cm)}$$

3.3.2. ESTIMATIVA DA ALTURA DA SEGUNDA REFERÊNCIA VERTICAL (LR1) UTILIZANDO-SE APENAS UMA REFERÊNCIA VERTICAL (LR2)

Em um segundo momento vislumbrou-se a possibilidade de identificar uma segunda referência vertical na imagem (LR1), exibida na figura 9. Foi constatado pelo perito no local, utilizando uma trena de alta precisão, que a altura desta construção em relação a calçada é de 107 cm. Vale ressaltar que devido ao aclive desta via esta medida não permanece constante ao longo da imagem.

Inicialmente a fim de verificar a viabilidade de utilização desta nova referência vertical além da própria confiabilidade do método

apresentado em 3.3.1, estimou-se a altura de LR1 utilizando apenas uma referência vertical (LR2).



FIGURA 9 – ALTURA LR1 NA IMAGEM QUESTIONADA

Segue na figura 9, a aplicação do diagrama utilizado para estimativa de LR1 a partir da altura de LR2 utilizando como ponto de referência auxiliar o ponto de fuga vertical.

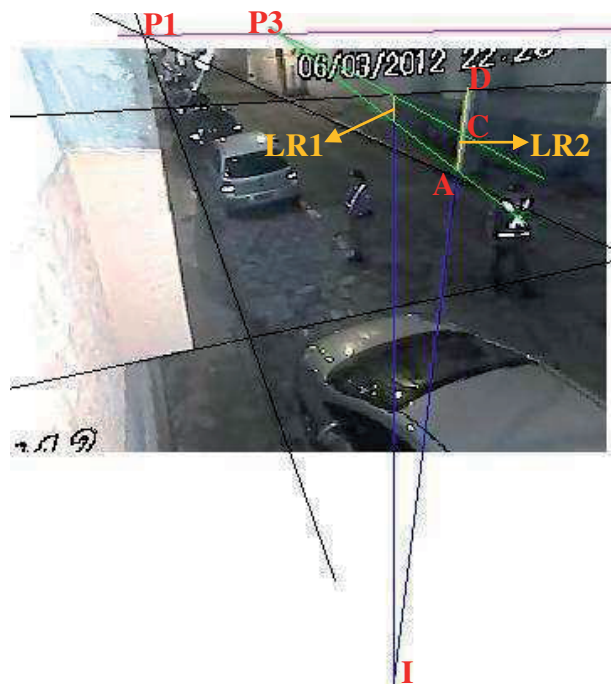


FIGURA 10 – DIAGRAMA CONTENDO O PONTO DE FUGA VERTICAL PARA ESTIMATIVA DE LR1

Seguem na tabela 4 as dez medições realizadas e na tabela 5 os resultados obtidos a partir destas medições.

TABELA 4
RESULTADOS DAS MEDIÇÕES REALIZADAS

Medições	AC (px)	AD (px)	IC (px)	ID (px)	K
1ª	23,084	49,345	345,812	372,076	0,5033378
2ª	23,36	50,57	283,797	311,043	0,506282

3ª	23,16	50,567	320,975	348,321	0,4970268
4ª	23,12	50,14	280,983	308,051	0,505529
5ª	22,374	48,555	354,173	380,521	0,4950771
6ª	23,691	50,953	287,793	315,199	0,509235
7ª	23,179	50,304	338,616	365,727	0,4976703
8ª	23,402	50,794	287,859	315,123	0,5043602
9ª	23,622	50,66	266,281	293,368	0,5137171
10ª	23,02	49,89	328,65	355,398	0,4989685

TABELA 5
RESULTADOS OBTIDOS

$\hat{K}$	0,503120392
$\hat{\sigma}_k$	0,005923853
$\hat{\sigma}_{LR1}$	1,24722475

Dessa forma, a altura estimada de LR1, utilizando apenas uma referência vertical (LR2) na imagem é:

$$LR1 = 105,91 \pm 1,25 \text{ (cm)}$$

3.3.3. ESTIMATIVA DA ALTURA QUESTIONADA UTILIZANDO DUAS REFERÊNCIAS VERTICAIS.

A partir da identificação das duas referências verticais conhecidas na imagem (LR1 e LR2), pode-se calcular a altura do indivíduo questionado (LO) a partir do método apresentado no item 3.1.1.

Apresenta-se na figura 11 a aplicação do diagrama para o cálculo da altura questionada (LO) utilizando a conservação da razão e as duas referências conhecidas: LR1 e LR2, medindo 107 e 210,50 cm respectivamente.

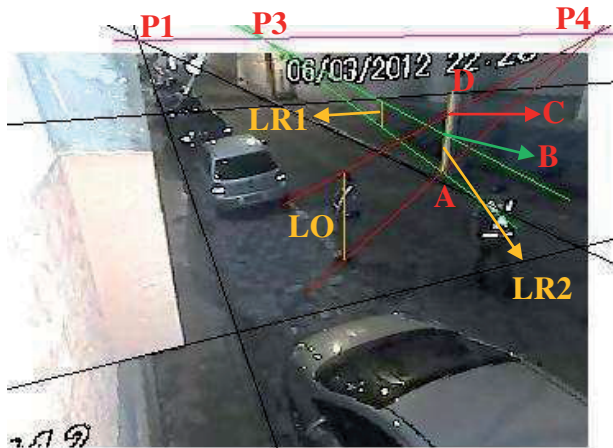


FIGURA 11 – DIAGRAMA CONTENDO A RAZÃO CRUZADA PARA O CÁLCULO DA ALTURA QUESTIONADA (LO)

Seguem na tabela 6 as dez medições realizadas e na tabela 7 os resultados obtidos a partir destas medições.

TABELA 6
RESULTADOS DAS MEDIÇÕES REALIZADAS

Medições	AC (px)	AD (px)	BC (px)	BD (px)	K
1ª	37,403	50,22	12,25	25,135	1,528173
2ª	37,623	50,934	12,536	25,801	1,5202786
3ª	37,342	50,505	13,27	26,436	1,4729501
4ª	36,478	49,754	14,191	27,589	1,4253646
5ª	36,401	50,685	13,196	27,47	1,4950311
6ª	37,595	50,54	14,204	27,277	1,4285018
7ª	35,403	48,529	13,052	26,255	1,467485
8ª	37,408	50,57	14,237	27,46	1,4267687
9ª	36,422	50,477	12,998	27,339	1,5176665
10ª	36,351	50,495	13,127	27,326	1,4985753

TABELA 7
RESULTADOS OBTIDOS

$\hat{K}$	1,478079475
$\hat{\sigma}_k$	0,040239349
$\hat{\sigma}_{LO}$	2,176211672

Dessa forma, a altura estimada do indivíduo questionado, na ocasião em análise, utilizando as duas referências verticais conhecidas (LR1 e LR2), é aproximadamente:

$$LO = 160,34 \pm 2,18 \text{ (cm)}$$

4. CONCLUSÃO

O emprego das técnicas fotogramétricas apresentadas neste artigo demonstra ser uma alternativa possível para as diversas situações em que o sistema de captura encontra-se indisponível a equipe forense, impossibilitando assim a utilização de gabaritos métricos.

Verificou-se no caso em análise que o método para estimativa de uma altura questionada utilizando apenas uma referência vertical e o ponto de fuga vertical, conforme resultados obtidos exibidos no item 3.3.2, apresentou pequena variação em relação a medida real conhecida (LR1). Cabe ressaltar

que a medida real encontra-se dentro da margem definida pelo desvio obtido.

Em relação a estimativa da altura do indivíduo questionado (LO), o método utilizando apenas uma referência e o ponto de fuga vertical, cujo os resultados obtidos foram exibidos no item 3.3.1, apresentou divergências em relação ao resultados obtidos através do método utilizando as duas referências verticais conhecidas, cujo os resultados obtidos foram exibidos no item 3.3.3.

Foi constatada uma pequena faixa comum, 1,61m a 1,62m, para os resultados obtidos através dos dois métodos (Item 3.3.1 e 3.3.3). Entretanto nota-se que o resultado obtido em 3.3.1, utilizando apenas uma referência e o ponto de fuga vertical, contemplou faixas maiores de altura (1,62 a 1,66) enquanto que resultado obtido em 3.3.3, utilizando as duas referências verticais, contemplou faixas menores de altura (1,58 a 1,62).

Apesar de o autor considerar provável que a altura questionada encontre-se contida na faixa comum aos resultados dos dois métodos, recomenda-se que a estimativa de alturas de indivíduos, sempre que possível, utilize duas referências verticais presentes na imagem.

AGRADECIMENTOS

A diretoria da Perícia Oficial e Identificação Técnica do Estado de Mato Grosso (POLITEC) por viabilizar a capacitação dos peritos lotados na Gerência de Perícias em Áudio e Vídeo.

Aos profissionais do SEPAEL (Serviço de Perícias em Audiovisual e Eletrônicos) do Instituto Nacional de Criminalística da Polícia Federal, pela difusão de conhecimento e cooperação técnica quando necessário.

REFERÊNCIAS

- [1] CONRADY, A., "Decentered lens systems," Monthly Notices of the Royal Astronomical Society 79, 384–390, 1919.
- [2] BROWN, D.. "Decentering distortion of lenses," Photogrammetric Engineering 7, 444–462, 1966.
- [3] BROWN, D.. "Close range camera calibration," Photogrammetric Engineering 8, 855–855, 1971.
- [4] DE VILLIERS, J. P.; LEUSCHNER, F. W.; GELDENHUYS, R.. "Centi-pixel accurate real-time inverse distortion correction," in Proceedings of the 2008 International Symposium on Optomechatronic Technologies, ser. ISOT2008, vol. 7266, pp. 1-8, 2008.
- [5] Plugin Lens_unDistortionSMT. Disponível em: <http://svg.dmi.unict.it/iplab/imagej/Plugins/Forensics/LensUnDistortion/Lens_unDistortionSMT.htm>. Acesso em 15/07/2012.
- [6] ImageJ Features. Disponível em: <<http://imagej.nih.gov/ij>>. Acesso em 15/07/2012.
- [7] REIS, P. M. G. I., Fundamentos de Processamento Digital de Imagens, Academia Nacional de Polícia, Instituto Nacional de Criminalística, Brasília, 2012.
- [8] VUOLO, J. H.. Fundamentos da Teoria de Erros". Ed. Edgard Blücher Ltda, São Paulo, 1992.

On the practical aspects of applying the PRNU approach to device identification tasks

Hermeson B. Costa¹, Ronaldo F. Zampolo¹, Diego M. Carmo¹, Adalbery R. Castro^{1,2}, Eurípedes P. Santos¹

¹ *Computer and Telecommunications Department, Federal University of Pará
Augusto Correa, 1, Belém, Pará, 66075-110, Brazil*

² *Centro de Perícias Científicas Renato Chaves
Rodovia Transmangueirão, s/n, Belém, Pará, Brazil*

hermeson.costa@itec.ufpa.br, zampolo@ufpa.br,
diego.carmo@itec.ufpa.br, adalbery@ufpa.br, epsantos@ufpa.br

ABSTRACT – This paper addresses some practical issues regarding the device identification problem when using the sensor photo-response nonuniformity (PRNU). The PRNU is unique to each digital imaging sensor due to imperfections in the manufacturing process. The motivation for this work comes from some problems found in identifying devices by using threshold strategies presented so far. Experiments with five digital cameras, from two different manufacturers, were performed and their corresponding results are discussed. Although such results are preliminary, they may point out correct usage and help to find out limitations of the forensic techniques based on PRNU.

1. INTRODUCTION

This work presents a brief discussion, supported by experimental results, about the use of the photo-response nonuniformity (PRNU) of camera sensors as a strategy to identify acquisition devices from digital photographs or videos.

In forensics, device identification means verifying a possible connection between a signal under analysis and an acquisition equipment. We are particularly interested in digital camera identification. In this context, both the digital photograph and the suspicious camera must be available. Thus, the question to be answered is “Did this camera take the photo?” In order to find out, the following approaches can be adopted alone or combined: file header evaluation, watermark checking, and sensor noise pattern analysis [1].

The simplest strategy is the file header evaluation. Such an approach consists in extracting the information contained in the image or video file header which is written automatically by the acquisition equipment itself. For camera identification, the information of interest is the model and serial

number of the camera. Its simplicity however is accompanied by the characteristic of being a weak evidence provider, once the file header can be easily modified or erased.

On the other hand, analysis of watermark consistency requires a camera capable to insert a digital watermark in files it produces. In addition, such a watermark must be robust to re-quantisation, resizing, and filtering operations, in order to serve as a reliable feature for identification purposes. Then, this watermark would be recovered later from a digital picture/video to determine its origin. The drawback here is exactly the necessity of having an acquisition equipment with watermarking capabilities, which is not common if one considers the typical universe of Brazilian forensic institutes.

The third approach aforementioned is the analysis of the sensor noise pattern, also known as PRNU. The PRNU is due to differences among photo-sensor elements in converting light to electrical signal [2], [3]. Such differences lead to a sort of inherent watermark, which characterises uniquely each photo-sensor (sensor fingerprint). The PRNU can be estimated from a camera and then compared with the PRNU found in a photograph under analysis.

The PRNU has some properties, which make it useful for forensic purposes: a) the noise pattern is unique to each sensor; b) every sensor exhibits it and every picture as well (with the exception of completely dark images); c) the PRNU is robust to a wide range of image processing operations, such as lossy compression, filtering, and gamma adjustment;

and d) the PRNU is stable in time and under a wide range of physical conditions.

Images under test are filtered in order to separate as much as possible the image signal from its PRNU component, which is then compared with the estimated camera PRNU. Such a comparison consists in assessing the statistical similarity between camera and image PRNUs. This image PRNU extraction is not an easy task, because the noise sensor is multiplicative and consequently correlated with image pixels.

The remaining of this text is organised as follows. Section 2 introduces a model of the image sensor output as well as the PRNU estimation process. In Section 3, the fingerprint detection procedure in test images is addressed. Section 4 shows and discusses some experimental results. In Section 5, concluding remarks are drawn.

2. SENSOR OUTPUT MODEL AND FINGERPRINT ESTIMATION

In the formation process of a non computer generated digital image, real scene information passes through several stages, as wavelength filtering (usually by using a colour filter array), light-to-electrical current conversion, signal amplitude quantisation, colour channel interpolation (demosaicing), and lossy coding. Such stages are common to the most of video cameras, cellphones and camcorders spread worldwide. Each of those stages inserts distortions into the original image signal, which must be taken into account in modelling the image formation process. A general model for the image sensor output [2] is given by

$$I = g^\gamma \cdot [(1 + K)Y + \Omega]^\gamma + Q, \quad (1)$$

where g is the gain factor (different for each colour channel); γ represents the gamma correction factor; Y denotes the original image in the absence of any noise; Ω is a combination of several types of noise; Q models the combined distortion due to quantisation and/or lossy compression (usually from JPEG coding); and K is the device PRNU (camera fingerprint).

Based on (1), a maximum likelihood (ML) estimator for the PRNU can be derived as shown in Expression 2 [2].

$$\hat{K} = \frac{\sum_{k=1}^d W_k I_k}{I_k^2}, \quad (2)$$

where, d is the number of camera images used; I_k represents the k -th camera image; and W_k corresponds to the residual noise of I_k , which in turn is given as follows.

$$W_k = I_k - \hat{I}_k^{(0)}, \quad (3)$$

where the $\hat{I}_k^{(0)}$ is a denoised version of I_k .

The larger the number of camera images (d), the better the PRNU estimation. Typically, for acceptable quality estimation, d ranges from 30 to 50.

3. FINGERPRINT DETECTION

By fingerprint detection, we mean extracting the PRNU from the image under investigation and its comparison with the estimated camera PRNU. Actually, the best we can do is to infer the image PRNU by filtering. Details of the procedure used in this paper can be found in [2], [4]. From this section on, we consider the scenario in which that image has the same dimensions of the estimated camera PRNU and does not have undergone any kind of geometrical transformation, such as scaling or rotation.

3.1. HYPOTHESIS TESTING

For detection purposes, the following two-channel hypothesis testing problem is assumed [3]:

$$\begin{aligned} H_0 : K_1 &\neq K_2 \\ H_1 : K_1 &= K_2 \end{aligned}, \quad (4)$$

where K_1 and K_2 denotes the camera and the image under investigation fingerprints, respectively.

The null hypothesis H_0 means the image was not taken by the given camera (i.e. their fingerprints are different), while the alternative hypothesis H_1 says it was.

Two figures are associated to the hypothesis testing problem in (4): the probability of false rejection, and the probability of false acceptance. False rejection or false negative occurs when the hypothesis H_I is rejected, given the analysed image had indeed been acquired by the suspicious camera. In turn, false acceptance or false positive happens when H_I is accepted for an image that had not been taken by the suspicious camera. Such figures are actually the ultimate quality indicators of any technique conceived to solve the camera identification problem. For forensic applications, in particular, the most important figure seems to be the false acceptance probability, since in general this means that a compromising photograph/video is linked to a camera owned by an innocent person. Obviously, the smaller those probabilities, the better the identification system. However, there is a trade-off between such probabilities, meaning that the decrease of one of them results in the increase of the other.

3.2. STATISTICAL SIMILARITY METRICS

In order to test the hypothesis in (4), some metric to quantify the statistical similarity between estimated PRNUs is required. Two possibilities are given next.

The first metric presented here is the well-known Pearson's correlation coefficient, which is defined as

$$\rho = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [X(i, j) - \bar{X}][Y(i, j) - \bar{Y}]}{\|X - \bar{X}\| \|Y - \bar{Y}\|}, \quad (5)$$

where X and Y denotes two random samples with dimensions $m \times n$; $\bar{X}$ and $\bar{Y}$ are the arithmetic means of X and Y , respectively; and $\|\cdot\|$ represents the L_2 norm.

Another metric used in this paper is the *peak-to-correlation energy* (*PCE*) [5]:

$$PCE = \frac{NCC(0,0)^2}{\frac{1}{MN} \sum_i \sum_j NCC(i, j)^2}, \quad (6)$$

where MN denotes the dimension of the NCC matrix, and NCC is the normalized cross

correlation, whose expression is given as follows.

$$NCC(i, j) = \frac{\sum_{k=0}^{m-1} \sum_{l=0}^{n-1} [X(k+i, l+j) - \bar{X}][Y(k+i, l+j) - \bar{Y}]}{\|X - \bar{X}\| \|Y - \bar{Y}\|}, \quad (7)$$

Those metrics are related, since $NCC(0,0)$ equals ρ by definition. In spite of that, *PCE* attains better results (see Section 4) due to the inherent comparison between peak and average energy of NCC .

3.3. DECISION MAKING

For simplicity reasons, in the following discussion we refer only to ρ but the same points also apply to *PCE*. Thus, after the calculation of ρ , its value is then compared with a decision threshold ρ_{th} . If ρ exceeds ρ_{th} , one decides for hypothesis H_I (the image had been taken by the camera), otherwise, one decides for H_0 . Rigorously, we would need two thresholds, one corresponding to the probability of false positive and other associated to the false negative probability. The determination of those thresholds is an issue, asking for training procedures or the availability of some previous statistical model.

Undoubtedly, having general statistical models that describe the behaviour of ρ , one for each hypothesis H_0 and H_I , would be very useful and time-saving because of the avoidance of training sessions. However, such models are not easily obtained, as their statistical properties seem to be camera-dependent. Thus, training comes up as a necessary step in practical situations.

When the suspicious camera is available, developing specific models with regard to similarity metrics, considering test images had been taken by the suspicious camera, is perhaps cumbersome but not technically difficult. As a result, decision thresholds for given probabilities of false negative can be easily derived. However, practical models (even specific ones) for the behaviour of similarity metrics, when images had not been taken by the suspicious camera, have not been achieved so far to the best of our knowledge.

Existing attempts fail when applied to real situations [2], [3].

4. EXPERIMENTS AND RESULTS

This section begins by describing the experimental procedure adopted. Then, the corresponding results are presented and discussed.

4.1. EXPERIMENTAL PROCEDURE

The experiments conducted in this paper were conceived to evaluate the performance of the two metrics addressed in Section 3 as features intended for device identification. We have used five different cameras, three of the model Fuji-FinePix JZ300 with serial numbers OAQ38568, OAQ38657, and OBQ39278 (for convenience, identified as Fuji_1, Fuji_2 and Fuji_3, respectively); and other two of the model Sony-DSC-W210, serial numbers 6507300 and 6507323 (named Sony_1 and Sony_2, respectively).

For each camera, 30 images of a cloudy sky were taken to estimate the camera fingerprint, and other 30 images of random content serve as test signals.

4.2. EXPERIMENTAL RESULTS

The first set of tests regards Pearson's correlation coefficient. Results are shown in Figures 1-5 for cameras Fuji_1, Fuji_2, Fuji_3, Sony_1 and Sony_2, respectively. In Figures 1 to 3, one can see the correlation coefficients from those images taken by the suspicious camera present higher values and the existence of a threshold ρ_{th} can be clearly perceived. Figures 4 and 5, however, present what we can call a class mixture, where one must tolerate some misclassification. In addition, note that the variation between camera models of the underlying ρ_{th} cannot be overlooked.

The results concerning *PCE* are shown in Figures 6-10 for suspicious cameras Fuji_1, Fuji_2, Fuji_3, Sony_1 and Sony_2, respectively. By taking the results separately, one realizes *PCE* outperforms the correlation coefficient, since no class confusion is verified. Again, underlying decision threshold values vary from model to model. Figure 11 emphasises this last aspect, where all *PCE* data

is displayed together. In addition, such a figure points out the difficulty of obtaining general models as, even for the *PCE*, a lot of class confusion can be noticed.

4.3. DISCUSSION

As stated before, a camera identification system is evaluated from its false positive and negative probabilities. Tuning thresholds for each of such quality figures relies on statistical models of PRNU similarity metrics. False negative probability depends on models derived from images taken by the suspicious camera (H_1 is true). This can be easily done, but false negative is the less important figure for a forensic application. In turn, false positive probability depends on models derived from images taken by cameras that are different from the suspicious camera (H_0 is true). Note that, in this case, our results suggest that as new cameras are included the model parameters change and system overall performance decays.

5. CONCLUSIONS

This paper has presented a discussion about the practical use of PRNU approaches for digital camera identification. Experimental results have demonstrated that statistical models for similarity PRNU metrics, which allow threshold tuning for a given false positive rate, are not easy to obtain. Those results suggest that as new cameras are included the model parameters change and system overall performance decays.

Future research efforts include a deeper investigation of statistical modelling under true H_0 condition.

ACKNOWLEDGEMENTS

The authors would like to thank UFPA and CPCRC for the financial support.

REFERENCES

- [1] J. Redi, W. Taktak, J. -L. Dugelay, "Digital image forensics: a booklet for beginners", *Multimedia Tools and Applications*, vol. 51 (1), pp. 133-162, 2011.
- [2] J. Fridrich, "Digital image forensics", *IEEE Signal Processing Magazine*, vol. 26(2), pp 16-37, 2009.
- [3] M. Goljan and J. Fridrich, "Camera identification from cropped and scaled images", in *Proc. of the SPIE Electronic Imaging*,

- [4] M. K. Mihcak, I. Kozintsev, and K. Ramchandran, "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising," Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Mar. 1999, vol. 6, pp. 3253–3256.
- [5] B. V. K. V. Kumar and L. Hassebrook, "Performance measures for correlation filters", *Applied Optics*, Vol. 29(20) (1990), pp. 2997-3006.

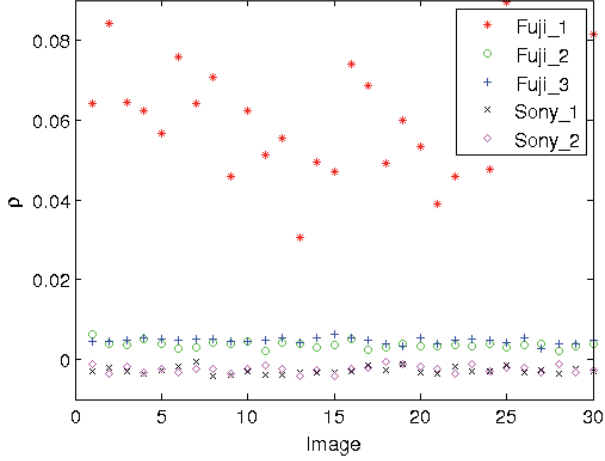


Figure 1: Pearson's correlation coefficients ρ for 30 images of each camera with the camera fingerprint estimated from Fuji_1.

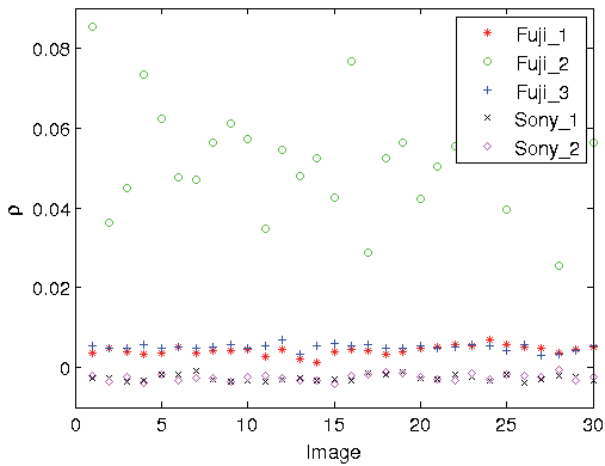


Figure 2: Pearson's correlation coefficients ρ for 30 images of each camera with the camera fingerprint estimated from Fuji_2.

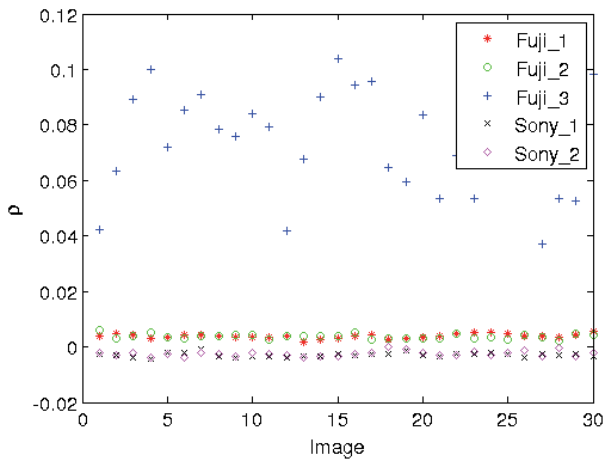


Figure 3: Pearson's correlation coefficients ρ for 30 images of each camera with the camera fingerprint estimated from Fuji_3.

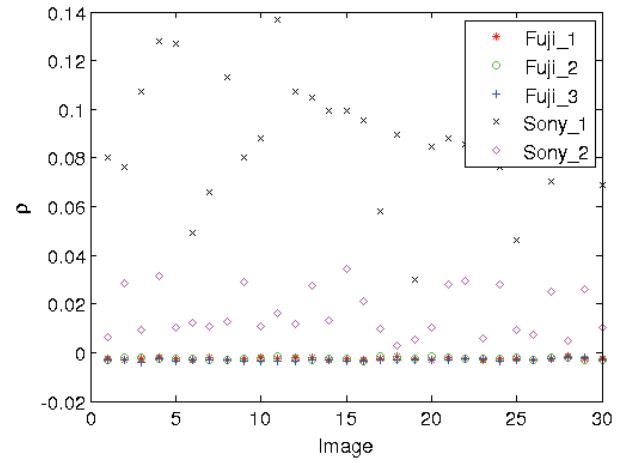


Figure 4: Pearson's correlation coefficients ρ for 30 images of each camera with the camera fingerprint estimated from Sony_1.

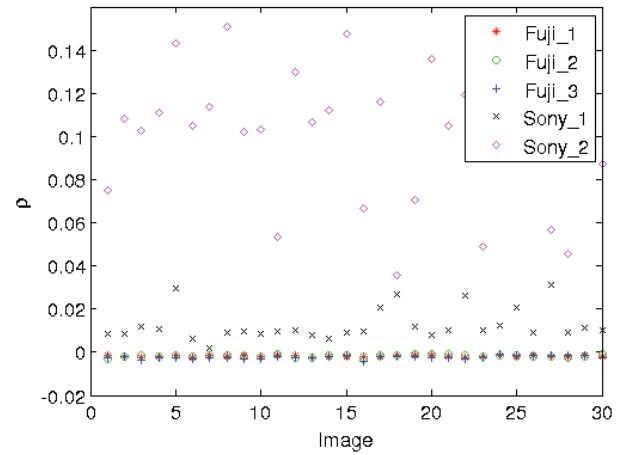


Figure 5: Pearson's correlation coefficients ρ for 30 images of each camera with the camera fingerprint estimated from Sony_2.

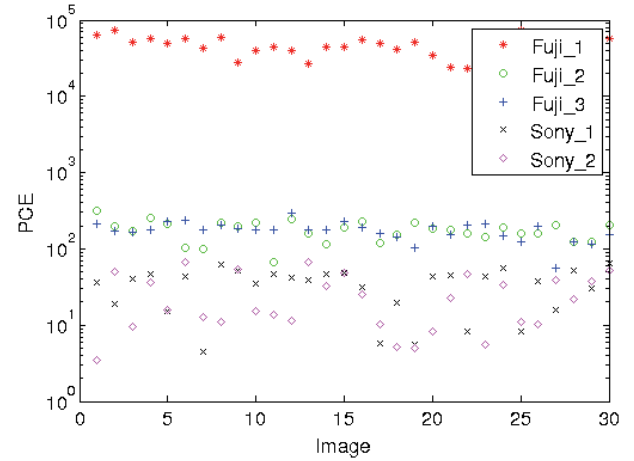


Figure 6: Peak-to-correlation energy (PCE) for 30 images of each camera with the camera fingerprint estimated from Fuji_1.

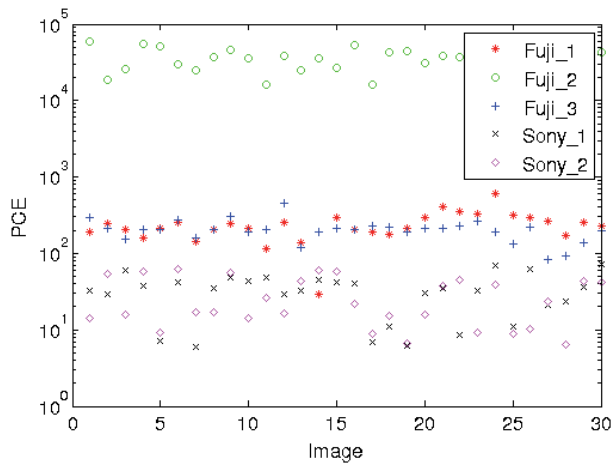


Figure 7: Peak-to-correlation energy (PCE) for 30 images of each camera with the camera fingerprint estimated from Fuji_2.

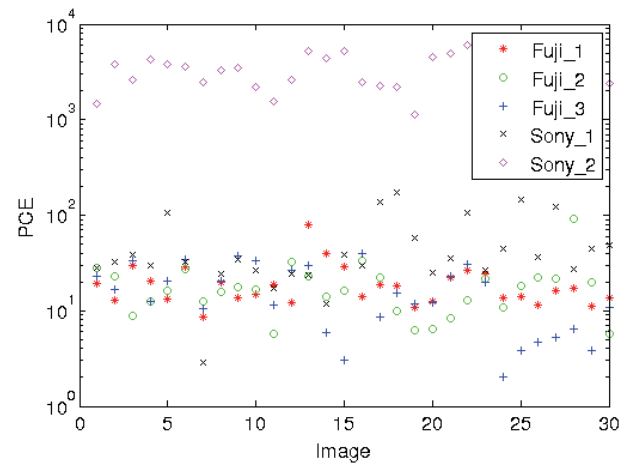


Figure 10: Peak-to-correlation energy (PCE) for 30 images of each camera with the camera fingerprint estimated from Sony_2.

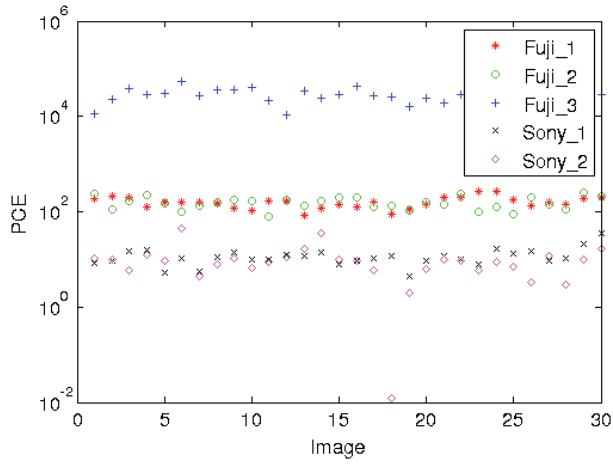


Figure 8: Peak-to-correlation energy (PCE) for 30 images of each camera with the camera fingerprint estimated from Fuji_3.

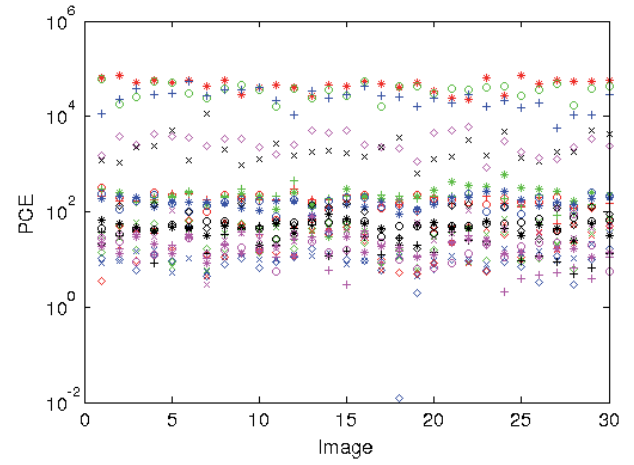


Figure 11: Peak-to-correlation energy (PCE) showed on Figure 6, Figure 7, Figure 8, Figure 9 and Figure 10 together.

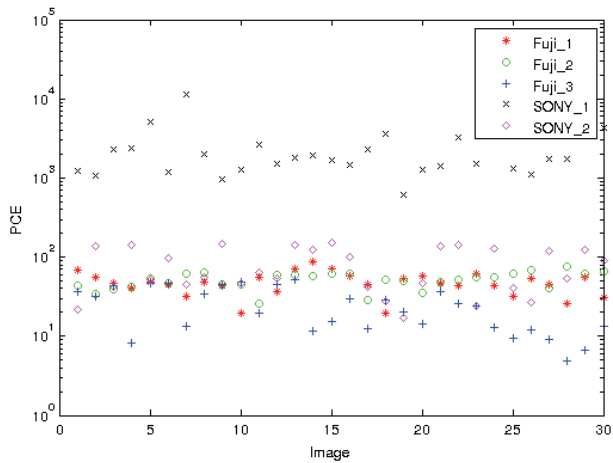


Figure 9: Peak-to-correlation energy (PCE) for 30 images of each camera with the camera fingerprint estimated from Sony_1.

Avaliação de Classificadores Haar Projetados para Detecção Facial

R. Padilla¹, C. F. F. Costa Filho², M. G. F. Costa²

¹ Instituto Nokia de Tecnologia

Av. Torquato Tapajós, 7200 – Col. Terra Nova, Manaus-AM. CEP: 69048-660 Brasil

² Universidade Federal do Amazonas

Av. Gal. Rodrigo Otávio Jordão Ramos, 3000, Manaus-AM. CEP: 69010-110 Brasil

rafael.padilla@indt.org.br, cffcfilho@gmail.com, marly.costa@uol.com.br

RESUMO – Nos últimos anos vários estudos foram realizados na área de detecção de faces. A face humana contém características importantes que podem ser utilizadas em sistemas biométricos para a identificação de indivíduos. A localização facial, a primeira etapa de sistemas biométricos pela face, localiza faces na imagem de entrada. O *framework* Viola-Jones tem sido amplamente utilizado por pesquisadores para detecção e localização de objetos e faces. Classificadores de detecção facial são compartilhados pela comunidade científica e acadêmica. Este trabalho tem como objetivo avaliar classificadores localizadores de face através de pontos faciais, auxiliando pesquisadores na escolha do melhor classificador de acordo com a necessidade de cada um.

1. INTRODUÇÃO

Embora o reconhecimento de um indivíduo através da face seja uma tarefa trivial para seres humanos, é um desafio para sistemas de automação baseados em imagens. Este tema tem sido estudado por pesquisadores de diferentes áreas envolvendo processamento digital de imagens, redes neurais, estatística, reconhecimento de padrões, antropometria e visão computacional. Sistemas de automação baseados em imagens fazem o reconhecimento e identificação facial em diversas aplicações comerciais, tais como autenticação por biometria, interação homem-máquina, vigilância, jogos e entretenimento.

Ao contrário de outras biometrias, o reconhecimento de face é não-invasivo e não necessita de contato físico do indivíduo com o sistema, tornando-se uma tecnologia muito aceitável no mercado. Sistemas automáticos baseados em visão para o reconhecimento de indivíduos pela face pode ser dividido em 4 etapas: detecção da face, pré-processamento da imagem, extração de características e reconhecimento [1].

A imagem capturada da face pode ser pré-processada para eliminar as variações de iluminação [2]. A extração de características é o processo onde um modelo geométrico ou vetorial é obtido reunindo características importantes presentes na face. A extração de características pode ser dividida em 3 abordagens: holística, baseada em características e híbrida. São exemplos de abordagens holísticas: Análise de Componentes Principais [3] [4], Análise Discriminante de Fisher [5] [6] e Máquina de Vetores de Suporte [7]. Extrair características de uma determinada face é uma tarefa difícil, uma vez que suas características, tais como olhos, boca, nariz e queixo, possuem, em geral, a mesma configuração geométrica. Abordagens baseadas em características utilizam as relações geométricas das características faciais. O Modelo de Forma Ativa pode ser aplicado para obter informações importantes presentes em alguns pontos faciais [8]. Classificadores estatísticos, tais como Distância Euclidiana [9], Classificador de Bayes [10], Distância de Mahalanobis [11] e Classificadores Neurais [12] podem ser usados na comparação do vetor de características com outras classes (indivíduos) durante a etapa de reconhecimento.

A detecção de faces teve sua acurácia melhorada com a aplicação de características tipo Haar, com a contribuição do *framework* de detecção de objetos Viola-Jones. Implementações deste *framework* tornaram-se populares e proporcionam a criação de classificadores por autores que utilizaram diferentes conjuntos de imagens para sua

geração. O desempenho e a confiabilidade desses classificadores varia muito [13].

Este artigo tem como objetivo avaliar classificadores faciais considerando características faciais contidas nas faces encontradas. Propomos um método utilizando diferentes pontuações dadas para cada característica facial contida na face localizada. Dois bancos de dados de faces (FEI e Yale) foram utilizados para avaliar 10 classificadores faciais.

2. MATERIAIS

2.1. BASE DE FACES YALE

A base de faces Yale [14] contém imagens dos rostos de 15 indivíduos, com 11 fotos cada um, tiradas em diferentes condições de iluminação. Os indivíduos apresentam expressões faciais variadas (com óculos, triste, sonolento, piscando, surpresos). O tamanho de cada imagem é 320x243 pixels.

A Figura 1 mostra exemplos de imagens da base de faces Yale.

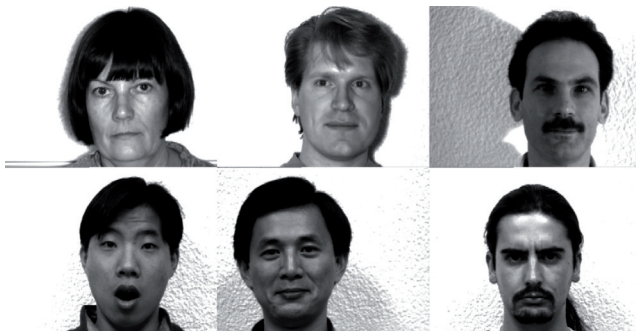


Figura 1 – Base de faces Yale

2.2. BASE DE FACES FEI

A base de faces FEI [15] é um banco de dados brasileiro contendo 14 imagens para cada um dos 200 indivíduos, com um total de 2800 imagens. As imagens são coloridas e apresentam faces em diferentes rotações, com expressões neutras e sorrindo. Utilizamos 2 imagens frontais por indivíduo, considerando as imagens com sorriso e neutras, totalizando 400 imagens. O tamanho original de cada imagem é 640x480 pixels. Exemplos de imagens da base de faces FEI podem ser vistos na Figura 2.



Figura 2 – Base de faces FEI

2.3. DETECTORES DE FACE VIOLA-JONES

Motivados pelo desafio de detecção de faces, [16] propuseram um *framework* detector objetos utilizando características tipo Haar. Este *framework* tem sido amplamente utilizado por outros pesquisadores e trabalhos, não só para a detecção de faces, mas também para localização de objetos em imagens e vídeos.

Graças à implementação da biblioteca OpenCV [17], o *framework* Viola-Jones tornou-se popular e motivou a comunidade a gerar seus próprio classificadores de objetos. Esses classificadores usam características tipo Haar que são aplicadas sobre a imagem.

O conceito de características tipo Haar provém de funções Haar de Base [18]. Existem basicamente 4 tipos de características, que são demonstradas na Figura 3.

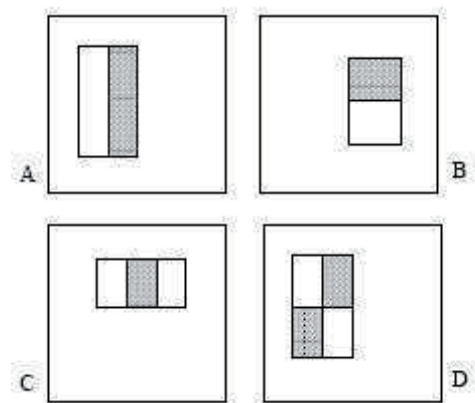


Figura 3 – Características tipo Haar posicionadas sobre uma janela de amostragem. FONTE: P. Viola and M. Jones [16]

Cada característica posicionada sobre uma parte da imagem, chamada de sub-janela, expressa um valor obtido através da soma dos pixels da região branca subtraídos da soma dos pixels da região escura.

Uma vez obtido o conjunto de características e seus limiares, o processo de teste é feito utilizando uma técnica de verificação em cascata, conforme a Figura 4. Nesta figura temos um classificador formado por N camadas. Cada camada é composta por várias características. Para o classificador comprovar que uma sub-janela possua uma face, esta sub-janela deve passar por todas as N camadas do classificador, de modo que não seja rejeitada por nenhuma delas.

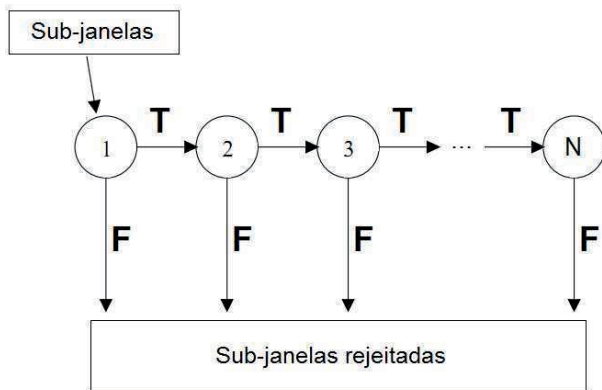


Figura 4 – Detecção em cascata

Alguns classificadores de localização de faces são distribuídos na implementação do OpenCV [17]. Estes classificadores são arquivos XML (Extensible Markup Language) estruturados em estágios como mostra a Figura 5.

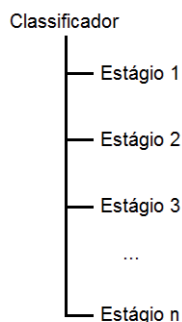


Figura 5 – Estrutura de um classificador Viola-Jones em XML

Cada classificador é formado por vários estágios. O classificador determina o tamanho mínimo da sub-janela que será analisada pelos estágios até que seja reprovada por um deles. Caso a sub-janela seja reprovada, a mesma é deslocada pela imagem, gerando uma nova sub-janela a ser analisada pelo classificador. Ao passar por toda imagem, a sub-janela terá suas dimensões ampliadas e o processo

recomeça através de toda a imagem, conforme mostra a Figura 6.

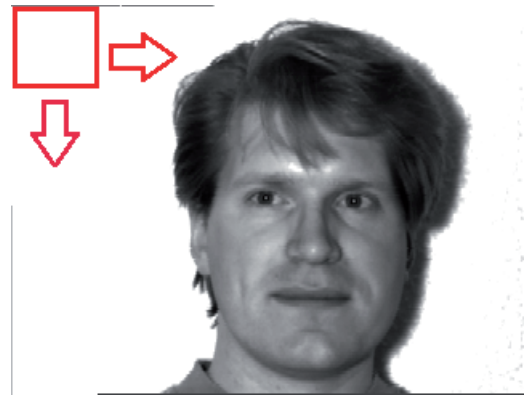


Figura 6 – Sub-janela deslocada através de uma imagem

Cada estágio é formado por características e por um limiar de estágio. Cada característica é formada por retângulos, por um limiar de característica, valor esquerdo e valor direito. A Figura 7 mostra a estrutura de um único estágio.

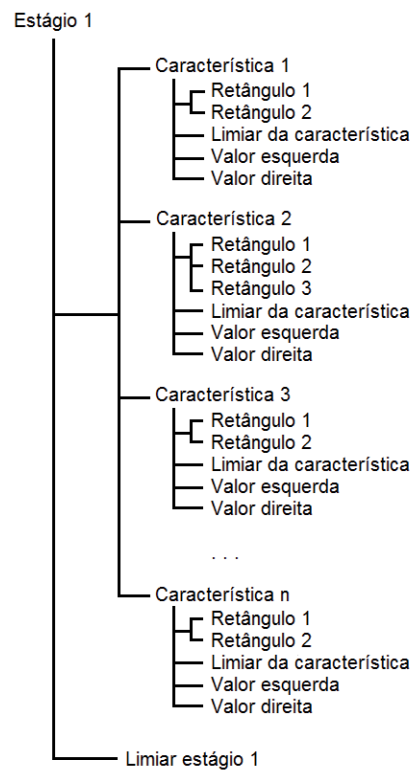


Figura 7 – Estrutura de um estágio utilizado no classificador de objetos

Cada característica possui um valor esquerda e um valor direita. Estes valores são pesos encontrados no processo de treinamento e atribuídos a cada característica, que são somados e comparados com o limiar do estágio.

O valor de cada característica é calculado pela aplicação da mesma na sub-janela. Se o valor da característica for menor do que seu limiar, acumula-se o valor da esquerda, caso contrário, o valor acumulado é o valor da direita. Para uma sub-janela ser aprovada por um estágio, o valor acumulado por todas as características deve ser menor do que o limiar do estágio.

Neste trabalho, foram utilizados 10 classificadores, apresentados na Tabela I. Mantivemos a mesma descrição utilizada por [13] para facilitar comparações.

TABELA I
CLASSIFICADORES

Classificador	Tamanho	Estágios	Referências	Posição da face
FD	24x24	25	[19,20]	Frontal
FA1	20x20	22	[19,20]	Frontal
FAT	20x20	46	[19,20]	Frontal
FA2	20x20	20	[19,20]	Frontal
FW	30x30	19	[21]	Frontal
FWQ	30x30	20	[21]	Rotacionada $\frac{1}{4}$
FWH	25x30	20	[21]	Rotacionada $\frac{1}{2}$
PR	20x20	26	[22]	Perfil
HS1	22x18	30	[23]	Parte superior do corpo
HS2	22x20	19	[24]	Cabeça e ombros

Selecionamos o maior número de classificadores disponíveis pela comunidade científica. A escolha desses classificadores resultou de uma pesquisa detalhada realizada na literatura, de onde extraímos as referências listadas na Tabela 1 e seus correspondentes classificadores. Os classificadores FW, FWQ, FWH, HS1 e HS2 foram gentilmente cedidos pelos autores através de contato realizado. Os classificadores PR, FD, FAT, FA2, FA1 estão disponíveis na versão 2.4 da biblioteca OpenCV.

3. PONTOS FACIAIS

A detecção de pontos faciais é importante não somente para gerar um modelo geométrico do rosto, mas também para ser utilizado na detecção de faces [24]. [25] compararam diferentes algoritmos para a localização de

pontos faciais e propuseram um conjunto de ferramentas que facilitam a integração com outras bases de dados. [26] propuseram uma técnica para segmentação facial usando Modelo de Forma Ativa com base em pontos de fronteira do rosto. [27] utilizou um modelo geométrico facial com base na distância dos olhos para estipular a posição dos pontos da face. Este modelo geométrico é mostrado na Figura 8.

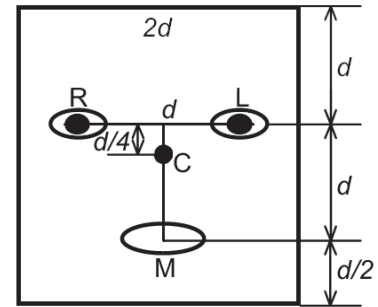


Figura 8 – Detecção em cascata

O projeto FGnet publicou a localização de 22 traços faciais de cada face do banco de faces AR [28]. Nós também marcamos manualmente estes 22 pontos faciais nas bases de Yale e FEI utilizadas neste trabalho. A Figura 9 mostra uma imagem com os pontos faciais marcados.

No total, 565 imagens foram utilizados e para cada um dos 22 pontos de referência, uma pontuação foi dada (ver Tabela 2). A pontuação dada para cada ponto foi 1 ou 2. Para pontos localizados no contorno da face a maior pontuação foi dada. A aplicação das pontuações será explicado na sessão seguinte.



Figura 9 – Exemplo de pontos faciais marcados manualmente

TABELA 2
PONTOS FACIAIS E SUAS PONTUAÇÕES

Ponto facial	Descrição	Pontuação
0	Centro do olho direito	1
1	Centro do olho esquerdo	1
2	Canto direito da boca	1
3	Canto esquerdo da boca	1
4	Canto direito da sobrancelha direita	2
5	Canto esquerdo da sobrancelha direita	2
6	Canto direito da sobrancelha esquerda	2
7	Canto esquerdo da sobrancelha esquerda	2
8	Limite superior direito da face	2
9	Canto direito do olho direito	1
10	Canto esquerdo do olho direito	1
11	Canto direito do olho esquerdo	1
12	Canto esquerdo do olho esquerdo	1
13	Limite superior esquerdo da face	2
14	Ponta do nariz	1
15	Narina direita	1
16	Narina esquerda	1
17	Limite superior labial	1
18	Limite inferior labial	1
19	Queixo	2
20	Limite direito da face	2
21	Limite esquerdo da face	2

4. EXPERIMENTOS

Primeiramente coletamos 10 classificadores de localização facial. Eles diferenciam-se pelo número de estágios e pelo tamanho mínimo de faces que podem ser detectadas. Eles foram projetados para detectar rostos em diferentes posições (frontal e de perfil) e somente a cabeça ou cabeça e ombros e juntos.

Dois critérios foram utilizados para determinar a precisão de cada classificador, e são descritos a seguir.

4.1. CRITÉRIO I

A precisão de um classificador foi mensurada pela pontuação obtida pelas faces localizadas. A região do rosto detectado deve conter o máximo de pontos faciais possível. A pontuação de cada característica facial é somada se esta encontra-se dentro do retângulo que representa a face.

A Figura 10 mostra o exemplo de uma face localizada por um dado classificador. Note que existem 2 pontos faciais importantes que não

estão dentro da face localizada - pontos 13 e 21.

Ao analisarmos diferentes faces, estimamos que um rosto é bem localizado, se sua pontuação total for maior ou igual a 27. Faces com pontuações mais baixas deixam importantes pontos faciais fora da imagem do rosto.

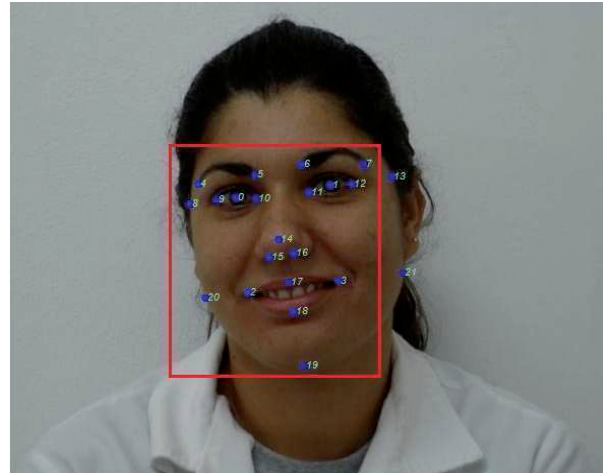


Figura 10 – Face localizada com pontuação igual à 27

4.2. CRITÉRIO II

Nós também consideramos o mesmo critério utilizado por [13] para avaliar a precisão do tamanho da face detectada. Uma face é corretamente detectada se sua altura e largura são menores que quatro vezes a distância entre os olhos. Diferentemente de [13], não avaliamos o tempo de processamento de cada classificador. Nosso objetivo é avaliar a acurácia dos classificadores, considerando pontos faciais (critério I) e o tamanho das faces localizadas (critério II).

4.3. RESULTADOS

Podemos ver os resultados da aplicação dos dez classificadores nas duas bases nas Figuras 11 e 12.

A Figura 11 mostra os resultados obtidos quando o critério I foi utilizado e a Figura 12 com a utilização do critério II.

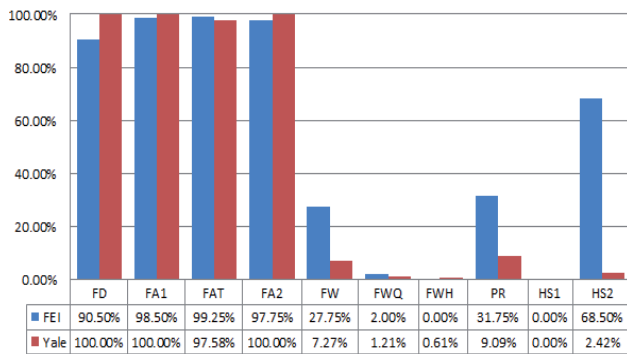


Figura 11 – Resultados utilizando o critério I

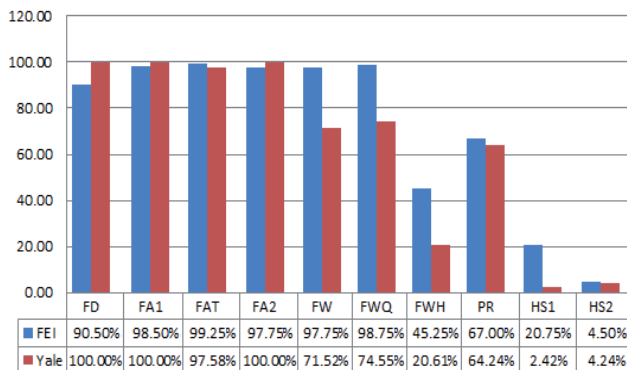


Figura 12 – Resultados utilizando o critério II

A partir dos dados obtidos, podemos concluir que:

a) Os melhores índices de aprovação foram alcançados pelos classificadores FA1, FA2 e FAT;

b) Nota-se que o índice de 100% de aprovação só foi conseguido em alguns classificadores, utilizando-se a base de imagens Yale;

c) Os classificadores FD, FA1, FAT, FA2 tiveram índices de aprovação semelhantes ao se utilizar os Critérios I e II, em ambas as bases;

d) Os classificadores FW, FWQ, FWH, PR e HS1 tiveram índices de aprovação bem menores, em ambas as bases, ao se utilizar o Critério I do que o Critério II;

e) O classificador HS2 teve um índice de aprovação maior na Base FEI ao se utilizar o Critério I, enquanto que o mesmo teve um índice de aprovação maior na Base Yale ao se utilizar o Critério II.

Assim sendo, a partir das conclusões anteriores podemos afirmar que os melhores classificadores avaliados foram FA1, FA2 e FAT. Em termos dos percentuais de

aprovação, o critério I, proposto neste trabalho, foi, na maioria dos casos, mais exigente que o critério II. A única exceção a essa regra foi verificada com o classificador HS2, quando avaliado na base FEI.

REFERÊNCIAS

- [1] J. Fagertun, 2005. Face Recognition. Dissertação de Mestrado, Universidade Técnica da Dinamarca (DTU).
- [2] Y. Gang, L. Jiawei, L. Jiayu, M. Qingli and Y. Ming, "Illumination Variation in Face Recognition: A Review", *IEEE Second International Conference on Intelligent Networks and Intelligent Systems (ICINIS 2009)*, pp. 309-311.
- [3] M.A. Turk, A.P. Pentland, "Face Recognition Using Eigenfaces", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'91)*, 3-6 Junho 1991, Maui, Hawaii, USA, pp. 586-591.
- [4] H. Moon, P.J. Phillips, "Computational and Performance Aspects of PCA-based Face Recognition Algorithms", *Perception*, Vol. 30, 2001, pp. 303-321.
- [5] K. Etemad, R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images", *Journal of the Optical Society of America A*, Vol. 14, No. 8, Agosto 1997, pp. 1724-1733.
- [6] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, "Face Recognition Using LDA-Based Algorithms", *IEEE Transaction on Neural Networks*, Vol. 14, No. 1, Janeiro 2003, pp. 195-200.
- [7] B. Heisele, P. Ho, T. Poggio, "Face Recognition with Support Vector Machines: Global versus Component-based Approach", *Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV'01)*, Vol. 2, 09-12 Julho 2001, Vancouver, Canadá, pp. 688-694.
- [8] A. Lanitis, C.J. Taylor, T.F. Cootes, "Automatic Interpretation and Coding of Face Images Using Flexible Models", *IEEE Transaction Pattern Analysis and Machine Intelligence* (1997), pp. 743-756.
- [9] E. Gomathi, K. Baskaran, "Recognition of Faces Using Improved Principal Component Analysis", (ICMLC'10), pp.198-201.
- [10] C. Liu and H. Wechsler, "Probabilistic Reasoning Models for Face Recognition", in: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'98)*, pp.827-832.
- [11] Y. Ji, T. Lin, and H. Zha, "Mahalanobis Distance Based Non-negative Sparse Representation for Face Recognition", em *The Eighth International Conference on Machine Learning and Applications (ICMLA'09)*, pp.41-46.
- [12] V. Kabeer & N. K. Narayanan, "Face recognition using state space parameter and Artificial Neural Network Classifier", *IEEE International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'07)*, Sivakasi, Índia Vol.3, Dezembro, 2007, pp 250-254.
- [13] M. Castrillón, O. Déniz, D. Hernández, and J. Lorenzo. "A Comparison of Face and Facial Feature Detectors based on the Viola-Jones General Object Detection Framework". *Machine Vision and Applications*, vol. 22 issue 3, 2011.
- [14] P. N. Bellhumer, J. Hespanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Issue on Face Recognition*, 1997, pp. 711-720.
- [15] L. L. de Oliveira Junior and C. E. Thomaz. "Captura e Alinhamento de Imagens: Um Banco de Faces Brasileiro". *Relatório Final de Graduação*. Departamento de Engenharia Elétrica, FEI, São Bernardo do Campo, São Paulo, Brasil, Junho 2006.
- [16] P. Viola and M. Jones, "Robust real-time object detection," *International Journal of Computer Vision*, 2002 vol. 57, no. 2, pp. 137-154.
- [17] Intel, Intel Open Source Computer Vision Library, v1.10.0, <http://sourceforge.net/projects/opencvlibrary> (Outubro 2011).
- [18] PAPANAGIOTOU, C.; OREN, M.; POGGIO, T. "A general framework for object detection" *International Conference on Computer Vision*, 1998.

- [19] Lienhart, R., Kuranov, A., Pisarevsky, V., "Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection", em *25th Pattern Recognition Symposium (DAGM'03)*, pp. 297-304.
- [20] Lienhart, R., Liang, L., Kuranov, A., "A Detector Tree of Boosted Classifiers for Real-Time Object Detection and Tracking", em *IEEE International Conference on Multimedia and Expo (ICME'03)*, pp. 277-280.
- [21] Wilson, P. I., Fernandez, J., "Facial feature detection using haar classifiers", em *Journal of Computing Sciences in Colleges* (2006), pp. 127-133.
- [22] Bradley, D. "Profile face detection" <http://www.cs.cmu.edu/~dbradley/publications/bradley-iurac-03.swf>, (2003) ultimo acesso em 10/12/2011.
- [23] Kruppa, H., Castrillón Santana, M., Schiele, B., "Fast and Robust Face Finding via Local Context", em *IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS'03)*, pp. 157-164.
- [24] G. M. Beumer, Q. Tao, A. M. Bazen and R. N. J. Veldhuis, "A Landmark Paper in Face Recognition". *IEEE International Conference on Automatic Face and Gesture Recognition (FGE'02)*, pp. 73-78.
- [25] M. Koestinger, P. Wohlhart, P. M. Roth and H. Bischof, "Annotated Facial Landmarks in the Wild: A Large-scale, Real-world Database for Facial Landmark Localization", *IEEE International Workshop on Benchmarking Facial Image Analysis Technologies (BeFIT'11)*.
- [26] M. Jian-Wei and F. Yu-Hua, "Face segmentation algorithm based on ASM", *IEEE Conference on Intelligent Computing and Intelligent Systems (ICIS'09)*, 2009, pp 495-499.
- [27] Z. Liu, W. Li, X. Zhang and J. Yang, "Efficient Face Segmentation Based on Face Attention Model and Seeded Region Merging", *9th International Conference on Signal Processing (ICSP'08)*, pp. 1116-1119.
- [28] A.M. Martínez and R. Benavente, "The AR Face Database", CVC Technical Report #24, Junho - 1998.

Identificação de Câmeras de Celulares Usando Sensor Pattern Noise

Leonardo Torres¹, Valter Ramos¹, Alisson S. Nascimento¹, José Alencar-Neto¹, Alejandro C. Frery²

¹*Departamento de Pesquisa & Desenvolvimento – Investigação e Perícia Digital, RASTRU
57057-780, Maceió, AL – Brasil*

²*Instituto de Computação – Universidade Federal de Alagoas, UFAL
57072-970, Maceió, AL – Brasil*

{ljmtorres, valter, alisson.sa.nascimento, alencar}@rastru.com.br, acfrery@gmail.com

Resumo—Neste trabalho, propomos uma metodologia para identificação de câmeras usando *Sensor Pattern Noise*. Para cada câmera investigada são obtidas características de n -imagens, pela média e pela mediana do ruído residual por técnicas de filtragem, denotadas *fingerprints*. Então o ruído residual da imagem investigada é comparado com as *fingerprints* da câmera para mensurar sua correlação. Finalmente, no processo de decisão, o algoritmo utiliza métodos estatísticos e um limiar adaptativo nos resultados provenientes das correlações. No melhor caso obtivemos 100% de acurácia e como pior resultado uma taxa de falsa aceitação FAR = 0% com falsa rejeição FRR = 13.33% para ambos os padrões estimados usando 10-imagens.

1. INTRODUÇÃO

A análise forense de vídeos e imagens digitais é uma área recente, cujo interesse da comunidade científica vem aumentando no decorrer dos anos. Uma das técnicas bastante utilizada na análise forense são as *blind techniques*, que usam atributos extraídos da própria imagem em análise. Assim, não há necessidade de inserir características em qualquer etapa da formação da imagem para posterior verificação ou validação, características que levam vantagens sobre outras técnicas, como exemplo, as *watermarks*.

Devido aos avanços tecnológicos e redução dos custos de dispositivos e softwares para aquisição e manipulação de imagens digitais, mais pessoas se tornam usuárias destas tecnologias usando livremente seus recursos, legal ou ilegalmente. Assim, estes usuários trazem novos desafios para investigação na ciência forense em imagens digitais.

Uma imagem digital contém bem mais que somente a cena a qual se captou, cada imagem leva consigo características inerentes a sua formação, como por exemplo: (i) o coeficientes gerados por

uma tabela de quantização utilizada como parâmetro no processo de compressão JPEG [1] e [2]; (ii) a correlação entre os pixels referente a interpolação das cores utilizando *Color Filter Array* (CFA) [1]; (iii) o ruído padrão do fotossensor deixado na imagem, denotado *Sensor Pattern Noise* (SPN) [3] e [4]. Estas e outras informações obtidas de uma imagem digital podem provar sua autenticidade e integridade, identificar o dispositivo de origem ou a correlação entre imagens como provenientes de um mesmo dispositivo, auxiliando a análise forense em crimes digitais.

Este trabalho apresenta uma metodologia para identificação de câmeras digitais baseado no ruído padrão do sensor. Apresenta as principais condições para captura das imagens do dispositivo investigado e um modelo entidade-relacional para armazenamento destas imagens. As imagens investigadas passam por um processo de extração do ruído residual, denotado atributos. Estes atributos servem como entrada para o processo de formação do padrão do modelo do dispositivo. Neste processo são identificados dados estatísticos comuns e inerentes às imagens. Tais características extraídas servem de entrada para o processo de decisão. O objetivo principal é constatar de qual dispositivo as imagens investigadas se originam.

O artigo está organizado da seguinte forma: Na Seção II apresenta o estado da arte, onde são expostas as principais técnicas utilizadas. Na Seção III descreve a metodologia abordada, ressaltando as entradas e saídas de cada processo. Na Seção IV são expostos os resultados alcançados e por fim, na Seção V nós discutimos as contribuições e trabalhos futuros.

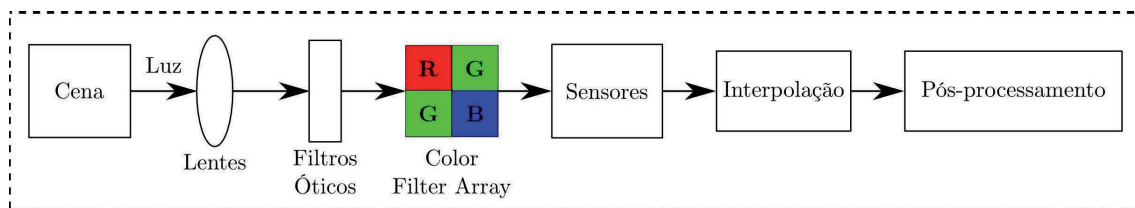


Figura 1. Modelo de aquisição de imagem digital.

2. TRABALHOS RELACIONADOS

A análise forense de imagens digitais, mais especificamente em identificação de dispositivos, é comumente utilizada para responder às questões abaixo dando suporte à comunidade forense no combate aos crimes eletrônicos:

- Dada uma imagem pode-se provar se ela foi ou não produzida por uma certa câmera ou dispositivo digital (*Camera Identification*)?
- Duas ou mais imagens podem ser correlacionadas como originadas de um mesmo dispositivo desconhecido (*Device Linking*)?

A. Aquisição de Imagem Digital

A Figura 1 mostra um modelo de aquisição de imagem digital: a luz refletida da cena e dos objetos nela contidos passa através de lentes e de filtros óticos, para que seja capturada pelos sensores de cores. A maioria das câmeras digitais usam sensores *Charge-coupled Device* (CCD) ou *Complementary Metal-Oxide-Semiconductor* (CMOS) e capturam as imagens usando um CFA. Os CFAs são matrizes, em sua maioria, com três filtros de cores em padrão RGB (*red, green, blue*), acoplados a cada sensor. Cada pixel de uma imagem digital tem três valores, um de cada cor ou banda, que compõem a cor do pixel, porém cada ponto do CFA captura um único valor de uma banda específica, os valores das outras bandas são inexistentes, até então. O processo de interpolação usa os valores da vizinhança do pixel nas outras bandas para estimar os valores inexistentes, afim, de torna possível a exibição da imagem e, finalmente, pode ser aplicado um pós-processamento onde, por exemplo, pode ser feito uma compressão JPEG.

B. Ruído Padrão do Sensor

Muitos autores vêm se dedicando no uso do SPN como principal característica para identificar um dis-

positivo, como em [3], [4], [5] e [6]. Diferentemente de outras técnicas, a análise por SPN não necessita de conhecimento *a priori* das características físicas do dispositivo, evitando-se a necessidade de abordagem intrusiva sobre os dispositivos na investigação de *Camera Identification*, ou mesmo a necessidade da posse do dispositivo original nas perícias de *Device Linking*.

Uma outra objeção em se utilizar técnicas baseadas em características físicas é a facilidade de adulteração das informações que as imagens levam no cabeçalho (e.g. EXIF), por exemplo, diminuindo sua credibilidade.

A motivação deste trabalho é a utilização de técnicas não intrusivas e uma estrutura de difícil adulteração. Essas premissas levam à utilização do ruído padrão do sensor, o qual se torna eficaz visto que cada dispositivo deixa na imagem, um tipo de impressão digital (*fingerprint*), uma marca única e inerente a cada sensor, mesmo em marcas e modelos iguais. Com esta *fingerprint* é possível identificar o dispositivo da qual uma imagem foi capturada, ou pelo menos determinar quais dispositivos a imagem se origina [7].

Lukáš et. al [4] descrevem que o ruído produzido no processo de aquisição pode ser dividido em duas categorias: (i) *Fixed Pattern Noise* (FPN) e (ii) *Photoresponse Nonuniformity* (PRNU). O FPN é formado por pontos escuros, pontos do sensor (CCD ou CMOS) não expostos à luz. Em uma imagem natural o que predomina é a PRNU, que é formada pelo ruído devido a não uniformidade dos pixels (*Pixel Nonuniformity* – PNU) e pelos componentes de baixa frequência. Todos os sensores não recebem a mesma intensidade de luz da cena, tornando-se mais uma característica do processo de fabricação do dispositivo. Porém, diferentemente do FPN, a temperatura não é considerada para contribuição na PNU.

C. Ruído Residual

Os componentes de baixa frequência de uma imagem digital advêm dos elementos da cena capturada. A luz refratada por partículas de poeira, superfícies ópticas e configurações de *zoom* também geram componentes de baixa frequência.

A supressão do ruído é um passo importante no processo de investigação e identificação de dispositivos através do SPN. Dos trabalhos como de [3], [4], [6], [8] e [9], obtém-se o ruído residual, denotado w_k , utilizando técnicas de filtragem e dado por

$$w_k = I_k - F(I_k), \quad (1)$$

em que I_k representa uma imagem natural e F um função de filtragem. Com isso, o SPN pode ser estimado pela média do ruído residual de n -imagens, denotado η e dado por

$$\eta = \frac{1}{n} \sum_{k=1}^n w_k, \quad (2)$$

em que n representa o número total de imagens usadas para estimar a impressão digital do sensor.

A Figura 2 apresenta um exemplo do processo de pré-processamento em imagem digital para obter o ruído residual utilizando a Equação (1).



Figura 2. Pré-processamento em imagem para obter o ruído residual utilizando a Equação (1).

D. Identificação e correlação usando SPN

Para identificar um dispositivo através do SPN, utiliza-se características únicas que podem ser provenientes de imperfeições dos componentes, de efeitos do ambiente e condições de operação [4]. Estas características dão unicidade ao ruído produzido no processo de captação da imagem, fazendo com que seja viável a identificação do dispositivo somente através de imagens originadas deste dispositivo.

Obtido o ruído residual, como subsídio para o processo de classificação, alguns autores extraem características do ruído (*features*) [5], [10], outros estimam a PNRU [11], [8] e outros ainda usam o próprio ruído residual como entrada [3], [6]. Alinhado a primeira abordagem, este trabalho utiliza características estatísticas extraídas do ruído residual como entrada para o processo de classificação.

E. Processo de Classificação

O processo de decisão é responsável pela autenticação do dispositivo investigado. Esta proposta utiliza a autenticação do tipo verificação “Um para Um” e o tipo de identificação “Um para Muitos”. O primeiro tipo verifica se uma imagem investigada pertence a um certo dispositivo conhecido, ou não. Enquanto, o segundo tipo identifica para um dado conjunto de câmeras se uma imagem investigada pertence a alguma das câmeras ou a nenhuma delas.

Muitos autores utilizam o SVM (*Support Vector Machine*) como processo de decisão supervisionado. Os resultados alcançados são satisfatórios, como em [10], [12], [5], [3].

A Curva ROC (*Receiver Operating Characteristic*) é um método estatístico, comumente utilizado para classificar indivíduos em grupos. Este método é fortemente apoiado no grupo e amostras de treinamento. A quantificação da medida do erro ou da chance do algoritmo validar corretamente é o objetivo deste método, como visto em [10], [3].

As taxas de falsa rejeição (*False Rejection Rate* – FRR) e as taxas de falsa aceitação (*False Acceptance Rate* – FAR) são um dos métodos estatísticos utilizados para traçar a Curva ROC. Porém, trabalhos como de [4], [8] tratam diretamente com estes dados para classificar os grupos com limiares adaptativos. Mancomunado com esta última abordagem, este trabalho propõe um limiar adaptativo que é mensurado pelas taxas de FAR e FRR.

3. METODOLOGIA ADOTADA

Esta seção apresenta as principais etapas para alcançar os objetivos elencados.

A. Banco de Dados de Imagens

A primeira etapa da metodologia é a aquisição e armazenamento das imagens capturadas para formação do modelo do dispositivo e das imagens investigadas. Nesta etapa definimos alguns aspectos, como os dispositivos a serem investigados, quantidade de imagens capturadas, ambiente (urbano, em laboratório), dentre outros. Podemos definir também, dependendo da necessidade, se serão usadas imagens já existentes de uma base de dados disponível.

A coleta e armazenamento de imagens de várias câmeras não é uma tarefa fácil e nem muito econômica. Gloe e Böhme [13] criaram o *Dresden Image Database for Benchmarking Digital Image Forensics*, que contém mais de 14 mil imagens em JPEG e TIFF de cenas variadas, capturadas de 73 modelos de câmeras diferentes. Este banco de dados é uma ferramenta de alto valor para a comunidade forense, pois apresenta uma visão clara do estado da arte em identificação de câmeras.

Inspirados pelo trabalho de Gloe e Böhme [13], nós utilizamos um modelo entidade-relacional para o armazenamento de imagens e informações do processo de aquisição. Um Banco de Dados (BD) é uma coleção de dados inter-relacionados, representando informações sobre um domínio específico e um Sistema Gerenciador de Banco de Dados (SGBD) provê ferramentas para acesso a estes dados.

A Figura 3 mostra um modelo simplificado do BD proposto. As entidades são representadas por retângulos, os relacionamentos são representados por losangos e os atributos são as elipses. O relacionamento entre um Fabricante e um Modelo caracteriza um Dispositivo. Podem existir vários Modelos para cada Fabricante, assim como vários Dispositivos com um mesmo Modelo e Fabricante. O Ambiente tem tipo (urbano, céu azul, natural) e uma *flag* indicando se é indoor ou outdoor. As condições em que uma imagem foi capturada (câmera fixa, distância focal, flash) são registradas no relacionamento Condições, caracterizado pelo relacionamento entre o Ambiente e um Dispositivo. Cada Imagem recebe um identificador (Id) criptografado. Se a imagem for

de um banco de dados externo ou online, o atributo Base Externa deve ser preenchido com uma descrição desta base.

B. Extração de Atributos

Uma vez populado o BD, colhemos n -imagens de cada câmera registrada para estimar a *fingerprint*. Foram utilizados dois operadores: (i) a média, seguindo a Equação (2) e (ii) a mediana. Cada operador é aplicado em cada ponto correspondente das n -imagens, que resulta em uma imagem de intensidade, denotada *fingerprint*.

C. Processo de Decisão

Após estimarmos a *fingerprint* e definimos um conjunto de m -imagens para treinamento (obtidas do mesmo dispositivo que originou a *fingerprint*). Calculamos as correlações entre a *fingerprint* e as m -imagens de treinamento usando a correlação de Pearson em blocos 8×8 . O coeficiente de correlação de Pearson é dado por

$$\rho_v(X, Y) = \frac{\sum_{r=1}^u (x_r - \bar{x})(y_r - \bar{y})}{\sqrt{\sum_{r=1}^u (x_r - \bar{x})^2 (y_r - \bar{y})^2}} \quad (3)$$

em que

$$\bar{x} = \frac{1}{u} \sum_{r=1}^u x_r, \quad \bar{y} = \frac{1}{u} \sum_{r=1}^u y_r$$

são as médias dos blocos referente as imagens do ruído padrão da câmera X e a imagem de treinamento Y , respectivamente, no r -ésimo pixel e $v = 1, \dots, b$, onde b é a quantidade de blocos 8×8 das imagens. Em posse das matrizes de correlações Z_j , $j = 1, \dots, m$ uma para cada imagem de treinamento, calculamos o coeficiente de variação ς_{Z_j} destas matrizes, que é uma medida de dispersão relativa definida por

$$\varsigma_{Z_j} = \frac{s_{Z_j}}{\bar{Z}_j} \quad (4)$$

em que s_{Z_j} é o desvio padrão e $\bar{Z}_j$ a média das matrizes de correlação Z_j .

Assim, podemos definir um limiar $\mathcal{L}_{C_i}$ seguro para identificar o dispositivo de origem, dado por

$$\mathcal{L}_{C_i} = \bar{\varsigma}_Z \pm (3 s_{\varsigma_Z}) \quad (5)$$

em que s_{ς_Z} é o desvio padrão e $\bar{\varsigma}_Z$ a média dos coeficientes de variação das Z_j , $j = 1, \dots, m$ de m -imagens de treinamento.

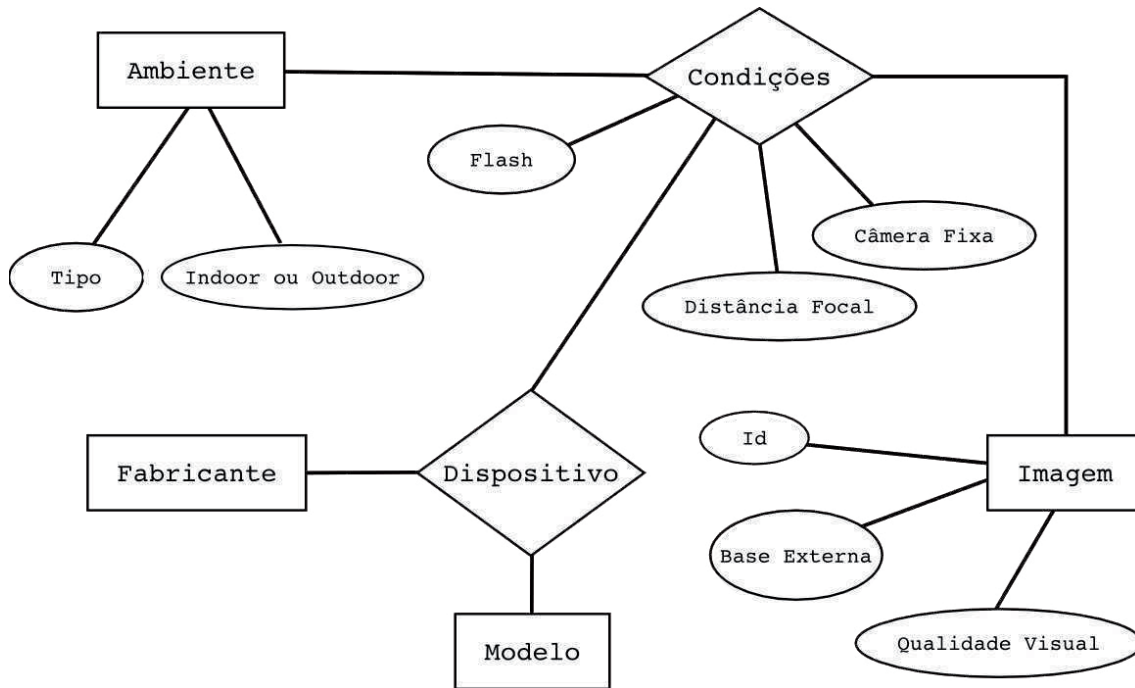


Figura 3. Modelo Entidade-Relacional simplificado do BD usado para registrar dados e imagens de entrada dos experimentos.

A Figura 4 apresenta um diagrama em blocos da metodologia abordada para identificação de dispositivos e o processo de validação.

4. EXPERIMENTOS E RESULTADOS

Nesta seção apresentamos os resultados obtidos com a metodologia proposta para identificação de dispositivos utilizando o ruído residual, que foi extraído com o processo de filtragem baseado em *wavelet* 2-D.

Para validar a proposta, nós registramos no BD 20 imagens com resolução nativa provenientes das 3 câmeras listadas na Tabela I, num total de 60 imagens. As imagens tiveram uma variação entre *indoor* e *outdoor*, luz natural e iluminação controlada, sem *flash*, com variação de *zoom* em dias claros e nublados. Além disso, nenhuma imagem passou por conversão de formatos ou recompressão.

Dentre as 20 imagens de cada câmera, 10 imagens foram utilizadas para extrair o ruído padrão do sensor, 5 imagens para determinar o limiar $\mathcal{L}_{C_i}$ e as 5 imagens remanescentes foram utilizadas nos testes.

Realizamos um recorte central quadrado em todas as imagens disponíveis no BD. A ordem do recorte é igual a menor dimensão dentre as resoluções nativas

Tabela I
FABRICANTES E MODELOS DAS CÂMERAS UTILIZADAS NOS EXPERIMENTOS.

ID	Modelo da Câmera	Resolução Nativa
C_1	Samsung Galaxy Ace	2560×1920
C_2	Samsung Galaxy X	2592×1944
C_3	Motorola EX115	2046×1536

das câmeras (Tabela I), neste experimento foi 1536 da C_3 .

Em seguida, obtermos as *fingerprints* de cada câmera, denotada f_1 a obtida pela média e f_2 a obtida pela mediana. No experimento, cada *fingerprint* é correlacionada com as amostras destinadas para treinamento utilizando a Equação (3).

Ao final deste processo, utilizando a Equação (4) obtermos duas tabelas: (i) Na Tabela II apresentamos os coeficientes de variação entre a *fingerprint* f_1 e todas as amostras para treinamento; (ii) Na Tabela III são expostos os resultados obtidos utilizando a *fingerprint* f_2 .

Com o suporte nestas tabelas (Tabelas II e III) calculamos os limiares para cada dispositivo utilizando a Equação (5). A Tabela IV apresenta os limiares para as câmeras envolvidas neste experimento.

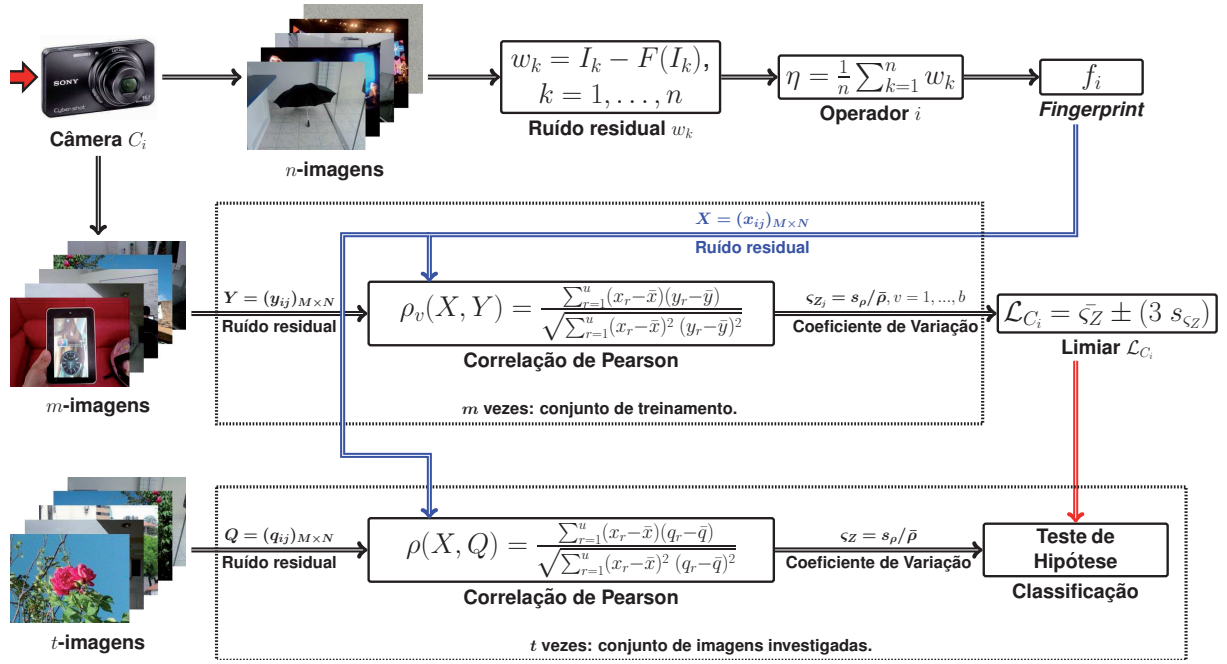


Figura 4. Diagrama em blocos da metodologia adotada e o processo de validação.

Tabela II

COEFICIENTES DE VARIAÇÕES ς_{Z_j} OBTIDOS COM A CORRELAÇÃO ENTRE O PADRÃO f_1 E AS IMAGENS PARA TREINAMENTO.

Amostras	ς_{Z_j} da C_1	ς_{Z_j} da C_2	ς_{Z_j} da C_3
#1	73.6045246	43.3224565	8.89200948
#2	37.1437447	50.7092309	16.4527572
#3	52.8831584	37.4681718	9.88767764
#4	84.8770245	35.5886513	10.1710338
#5	79.6299762	35.8433435	9.54454742

Tabela III

COEFICIENTES DE VARIAÇÕES ς_{Z_j} OBTIDOS COM A CORRELAÇÃO ENTRE O PADRÃO f_2 E AS IMAGENS PARA TREINAMENTO.

Amostras	ς_{Z_j} da C_1	ς_{Z_j} da C_2	ς_{Z_j} da C_3
#1	38.0277898	25.4793670	10.3750708
#2	21.9204324	18.0073213	19.4063936
#3	26.2725724	18.3630436	12.1477842
#4	42.8895011	17.9749418	10.7467727
#5	31.9896522	19.2894891	10.7732969

Submetemos as 15 imagens de teste para validar a metodologia proposta. Correlacionamos cada uma delas com as *fingerprint* f_1 e f_2 das três câmeras utilizando a Equação (3). Em seguida, extraímos de cada resultado o coeficiente de variação ς_{Z_j} . Por fim, verificamos se cada coeficiente está dentro do

Tabela IV

LIMIARES DAS CÂMERAS ENVOLVIDAS NO EXPERIMENTO.

Câmeras	$\mathcal{L}_{C_i}$ da f_1	$\mathcal{L}_{C_i}$ da f_2
C_1	65.62769 ± 60.07475	32.21999 ± 25.50330
C_2	40.58637 ± 19.39327	19.82283 ± 9.619036
C_3	10.98961 ± 9.544547	12.68986 ± 11.44459

limiar $\mathcal{L}_{C_i}$ (veja Tabela IV) correspondente de cada *fingerprint* para cada modelo de câmera.

Ambos os testes realizados com as *fingerprint* f_1 e f_2 foram satisfatórios e obtivemos no melhor caso uma acurácia de 100% ($FAR = 0\%$ e $FRR = 0\%$) para as câmeras C_1 e C_3 . Houveram duas falsas rejeições associadas ao padrão da câmera C_2 proporcionando uma $FRR = 13.33\%$, contudo, não houveram falsas aceitações ($FAR = 0\%$).

As 2 imagens, que resultaram na falsa rejeição em ambos os testes no padrão da câmera C_2 foram capturadas com um *zoom* acentuado e, provavelmente, sua rejeição está relacionada a esta característica. Estudos mais profundos devem ser direcionados ao entendimento deste fenômeno.

As Tabelas V e VI apresentam uma sumarização dos dados provenientes dos testes com as *fingerprint* f_1 e f_2 , respectivamente.

Tabela V
COEFICIENTES DE VARIAÇÕES ς_{Z_j} OBTIDOS COM A CORRELAÇÃO
ENTRE O PADRÃO f_1 E TODAS AS IMAGENS DE TESTE.

Amostras	ς_{Z_j} da C_1	ς_{Z_j} da C_2	ς_{Z_j} da C_3
#1 da C_1	40.014724	360.978214	134.552898
#2 da C_1	40.865016	208.343016	153.370379
#3 da C_1	33.251045	987.985970	265.912326
#4 da C_1	32.707328	228.554239	114.971440
#5 da C_1	32.370200	226.186325	5938.93515
#1 da C_2	214.283909	43.605546	292.500198
#2 da C_2	167.854468	4523.2382	112.740948
#3 da C_2	215.626974	34.337964	80.5481029
#4 da C_2	276.034535	37.349068	331.834290
#5 da C_2	259.536631	334.91339	120.501912
#1 da C_3	207.305097	84.3339083	8.4803013
#2 da C_3	365.319795	318.032940	14.351982
#3 da C_3	190.351127	93.5151662	9.1345649
#4 da C_3	76.3076703	68.6430355	10.360035
#5 da C_3	589.618371	132.979152	9.3097546

Tabela VI
COEFICIENTES DE VARIAÇÕES ς_{Z_j} OBTIDOS COM A CORRELAÇÃO
ENTRE O PADRÃO f_2 E TODAS AS IMAGENS DE TESTE.

Amostras	ς_{Z_j} da C_1	ς_{Z_j} da C_2	ς_{Z_j} da C_3
#1 da C_1	20.645812	723.010023	327.962533
#2 da C_1	18.739010	1386.01104	160.740719
#3 da C_1	16.212133	783.042479	3815.42664
#4 da C_1	19.063436	160.647400	128.627354
#5 da C_1	16.167803	124.934004	254.222252
#1 da C_2	3485.16331	24.227291	143.456115
#2 da C_2	237.097587	316.44011	203.522081
#3 da C_2	332.106203	19.942202	77.8998215
#4 da C_2	285.208588	16.812953	1245.56721
#5 da C_2	4790.17089	330.59364	272.922369
#1 da C_3	196.266328	161.075700	9.5609000
#2 da C_3	2616.05192	176.474426	18.632235
#3 da C_3	12520.0039	82.9365908	10.823300
#4 da C_3	58.5053499	72.2391886	11.861742
#5 da C_3	1688.85154	139.116071	10.215513

Fazendo uma análise com os trabalhos de Lukáš et. al [4] e Chen et. al [8], enquanto eles utilizam 300 e 30 imagens de cada câmera para obter o padrão do sensor, respectivamente, nós extraímos o padrão com 10 imagens; um número alto de imagens, neste caso, dificulta a replicação dos experimentos.

Lukáš et. al [4] avaliam sua proposta fixando a taxa de falsa aceitação em $FAR = 10^{-3}$ e para cada câmera com um limiar determinado obteve diferentes taxas de falsa rejeição e em um único caso

um $FRR = 0\%$. Chen et. al [8] utilizam o critério de correlação Neyman–Pearson e definem as taxas de erro por uma função de probabilidade com limiar determinado. Os resultados obtidos foram uma taxa de falsa aceitação em $FAR = 10^{-5}$ para todos os teste e uma taxa de falsa rejeição variada por dispositivo investigado.

5. CONCLUSÃO

Este trabalho propôs uma metodologia para análise forense em imagens digitais que responda à questão: Dada uma imagem, pode-se provar se ela foi ou não produzida por uma certa câmera ou dispositivo digital?

A metodologia proposta foi validada utilizando o *Sensor Pattern Noise* e duas *fingerprints* foram estimadas para cada dispositivo. Como método de organização e armazenamento apresentamos um modelo entidade-relacional simplificado, que abrange as informações das imagens capturadas e seus respectivos dispositivos.

Um grande diferencial foi adotar um limiar adaptativo baseado no coeficiente de variação ς_{Z_j} , onde para cada *fingerprint* existe um intervalo que identifica o dispositivo. Nos melhores casos obtivemos 100% de acurácia (nas câmeras C_1 e C_3 nas Tabelas V e VI) e como pior resultado uma taxa de falsa aceitação $FAR = 0\%$ com falsa rejeição de $FRR = 13.33\%$ (câmera C_2) para ambos os padrões estimados usando 10-imagens. Frente a literatura obtemos resultados satisfatórios utilizando poucas imagens para obter o ruído padrão do sensor.

Com a evolução desta metodologia espera-se utilizar um número ainda menor de imagens para estimar o padrão do sensor. Tendo em vista, que a aquisição destas imagens na área forense, em certas condições, pode ser onerosa ou não ser possível. Outros métodos para classificação devem ser abordados para futuras comparações e junções de métodos.

Com a análise do erro relacionado às imagens com *zoom* acentuado, iremos direcionar testes mais focados para avaliar o comportamento do ruído em diversas condições.

Outros passos estão relacionados a ampliação do Banco de Dados para que os dados dos experimentos, da etapa de extração de atributos e do processo de decisão possam ser registrados e replicados facilmente por outros peritos.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro concedido pelo CNPq através Programa RHAE – Pesquisador na Empresa (550670/2011-4), assim como, o suporte financeiro da FAPEAL pelo Programa PDT-TI (2010-02). Tais apoios vêm viabilizando os resultados até então alcançados nos projetos de P&D da empresa RASTRU em parceria com a Universidade Federal de Alagoas.

REFERÊNCIAS

- [1] H. Farid, “Image forgery detection,” *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 16–25, Mar 2009.
- [2] T. Bianchi and A. Piva, “Detection of nonaligned double jpeg compression based on integer periodicity maps,” *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 842–848, Apr 2012.
- [3] X. Kang, Y. Li, Z. Qu, and J. Huang, “Enhancing source camera identification performance with a camera reference phase sensor pattern noise,” *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 393–402, Apr 2012.
- [4] J. Lukáš, J. Fridrich, and M. Goljan, “Digital camera identification from sensor pattern noise,” *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, Jun 2006.
- [5] O. Celiktutan, B. Sankur, and I. Avcibas, “Blind identification of source cell-phone model,” *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 553–566, Sep 2008.
- [6] C.-T. Li, “Source camera identification using enhanced sensor pattern noise,” *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 280–287, Jun 2010.
- [7] J. A. Redi, W. Taktak, and J.-L. Dugelay, “Digital image forensics: a booklet for beginners,” *Multimedia Tools and Applications*, vol. 51, no. 1, pp. 133–162, Jan. 2011.
- [8] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, “Determining image origin and integrity using sensor noise,” *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 74–90, Mar 2008.
- [9] M. Goljan, J. Fridrich, and M. Chen, “Defending against fingerprint-copy attack in sensor-based camera identification,” *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 1, pp. 227–236, march 2011.
- [10] H. Gou, A. Swaminathan, and M. Wu, “Noise features for image tampering detection and steganalysis,” in *IEEE International Conference on Image Processing*, vol. 6, Oct. 2007, pp. VI–97–100.
- [11] C.-T. Li and Y. Li, “Color-decoupled photo response non-uniformity for digital image forensics,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 2, pp. 260–271, Feb. 2012.
- [12] H. Gou, A. Swaminathan, and M. Wu, “Robust scanner identification based on noise features,” in *Security, Steganography, and Watermarking of Multimedia Contents*, San Jose, CA, Jan. 2007.
- [13] T. Gloe and R. Böhme, “The dresden image database,” Online, 2010, acessado em 17 de julho de 2012. [Online]. Available: <http://forensics.inf.tu-dresden.de/ddimgdb/licenses>

PATROCÍNIO



APOIO INSTITUCIONAL



REALIZAÇÃO E ORGANIZAÇÃO



ASSOCIAÇÃO NACIONAL
DOS PERITOS CRIMINAIS FEDERAIS

COORGANIZAÇÃO



ABRID
Associação Brasileira das Empresas
de Tecnologia em Identificação Digital