Hermeson Barbosa da Costa

UMA ESTRATÉGIA DE ESTIMAÇÃO DA PRNU DE DISPOSITIVOS DE AQUISIÇÃO EQUIPADOS COM UM ARRANJO DE FILTROS DE COR

Dissertação submetida ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina para a obtenção do grau de Mestre em Engenharia Elétrica.
Orientador: Prof. Eduardo Luiz Ortiz Batista, Dr.
Coorientador: Prof. Rui Seara, Dr.

Florianópolis
2016

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Costa, Hermeson Barbosa da
Uma estratégia de estimação da PRNU de dispositivos de aquisição equipados com um arranjo de filtros de cor / Hermeson Barbosa da Costa; orientador, Eduardo Luiz Ortiz Batista; coorientador, Rui Seara. - Florianópolis, SC, 2016.
90 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Elétrica.

Inclui referências

1. Engenharia Elétrica. 2. Análise forense. 3. Identificação de dispositivo de aquisição. 4. PRNU. 5. Ruído de interpolação. I. Batista, Eduardo Luiz Ortiz. II. Seara, Rui. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Elétrica. IV. Título.

Hermeson Barbosa da Costa

UMA ESTRATÉGIA DE ESTIMAÇÃO DA PRNU DE DISPOSITIVOS DE AQUISIÇÃO EQUIPADOS COM UM ARRANJO DE FILTROS DE COR

Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia Elétrica, Área de Concentração Comunicações e Processamento de Sinais, e aprovada em sua forma final pelo Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.

Florianópolis, 05 de julho de 2016.

Prof. Marcelo Lobo Heldwein, Dr. sc. ETH
Coordenador do Programa de Pós-graduação em Engenharia Elétrica

Prof. Eduardo Luiz Ortiz Batista, Dr. – UFSC
Orientador

Prof. Rui Seara, Dr. – UFSC
Coorientador

Banca Examinadora:

Prof. Walter Pereira Carpes Júnior, Dr. – UFSC

Prof. Danilo Silva, Ph.D. – UFSC

Prof. Ciro André Pitz, Dr. – IFSC

Dedico este trabalho à minha família e ao meu tio Moraes.

Agradecimentos

Agradeço a Deus, por me confortar nos momentos mais difíceis.

Aos meus pais e aos meus irmãos, por todo o apoio, carinho e educação, por serem minhas fontes de inspirações, e por terem me apoiado a sair de casa em busca de minha formação profissional.

Ao meu tio, Moraes, por ter me disponibilizado aquela casa de madeira para eu morar durante o período do curso pré-vestibular e durante a graduação. Além disso, pelos conselhos e conversas que ajudaram um menino do interior a ter uma visão mais ampla do mundo.

À minha namorada, pelo apoio e pelas várias revisões realizadas durante a redação deste trabalho.

Ao meu orientador, Prof. Eduardo Luiz Ortiz Batista, Dr., e ao meu coorientador, Prof. Rui Seara, Dr., pelas orientações que contribuíram para o enriquecimento científico deste trabalho.

Ao meu orientador na graduação, Prof. Ronaldo Zampolo, Dr., por ter me incentivado a cursar o mestrado e por ter me indicado o LINSE.

A todos os meus amigos do LINSE, pelos momentos de descontração. Em especial, aos alunos da pós-graduação pelo bolo da feirinha de quarta. Além disso, vale ressaltar as contribuições do Marcos Matsuo e do Eduardo Kuhn através das discussões realizadas durante o desenvolvimento deste trabalho. Também agradeço ao Elton Fontão, pelas contribuições a realização deste trabalho e por ter me ajudado a descobrir que sou um goleador nato (apenas) no futebol de terça.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro e ao LINSE pela estrutura fornecida para o desenvolvimento deste trabalho.

Uma imagem vale mais do que mil palavras
(Autor desconhecido)

Resumo

Em análise forense, as técnicas de identificação de dispositivos de aquisição são usadas para verificar se uma imagem investigada foi adquirida por um determinado dispositivo suspeito (câmera digital, por exemplo). Para tentar solucionar esse problema, uma característica dos equipamentos de aquisição de imagens que vem sendo amplamente explorada é a não uniformidade da fotorresposta (*photo-response nonuniformity* – PRNU). A PRNU é um padrão de ruído, originário das imperfeições do processo de manufatura, que caracteriza cada fotossensor de forma única. Atualmente, a taxa de sucesso (identificação correta da fonte da imagem) das técnicas baseadas em PRNU, utilizando imagens de baixa resolução, tende a ser limitada. Tal limitação se deve ao fato de que imagens de baixa resolução possuem poucas amostras (*pixels*) para estimar a PRNU, diferente de imagens de alta resolução. Este trabalho propõe uma nova abordagem baseada em PRNU para identificação de dispositivos de aquisição de imagens. Tal abordagem baseia-se na decomposição da imagem em subimagens, visando reduzir o efeito da interpolação da imagem no processo de estimação da PRNU. Além disso, essa abordagem utiliza todos os canais da imagem, sem distinção entre eles, para obter uma estimativa da PRNU. Assim, resultados melhores são obtidos para identificação da fonte, tanto de imagens de alta resolução quanto de baixa resolução, conforme observado em experimentos realizados com imagens adquiridas por câmeras digitais e câmeras digitais de telefones celulares.

Palavras-chave: Análise forense. Identificação de dispositivo de aquisição. Imagem digital. PRNU. Ruído de interpolação.

Abstract

In forensic analysis, techniques of acquisition device identification are used for checking whether an image under investigation has been acquired by a particular suspect device (digital camera, for example). For trying to solve this problem, a characteristic of image acquisition device that has been widely explored is the photo-response nonuniformity (PRNU). The PRNU is a pattern noise originating from the imperfections of the manufacturing process, which characterizes each photo sensor. Currently, the success rate (correct identification of the image source) of techniques based on PRNU, using low-resolution images, tends to be limited. Such a limitation is due to the fact that low-resolution images have few samples (pixels) to estimate the PRNU, unlike high-resolution images. This paper proposes a new approach based on PRNU to perform image acquisition device identification. This approach is based on the decomposition of the image into sub-images, aiming to reduce the image interpolation effect in the PRNU estimation process. Moreover, this approach uses all image channels, without distinguishing among them, to estimate the PRNU. Thereby, improved results are obtained for identifying the source for both high-resolution and low-resolution images, as illustrated in experiments carried out by using images from digital cameras and cellphone digital cameras.

Keywords: Forensic analysis. Acquisition device identification. Digital image. PRNU. Noise interpolation.

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

2D Duas dimensões

CC *Pearson's correlation coefficient* (coeficiente de correlação de Pearson)

CCN *Correlation over circular cross-correlation norm* (razão entre a correlação e a norma da correlação cruzada circular)

CDMLE *Color-decoupled maximum likelihood estimator*

CFA *Color filter array* (arranjo de filtros de cor)

CRLB *Cramer-Rao lower bound* (limite inferior de Cramer-Rao)

i.i.d Independentes e identicamente distribuídas

JPEG *Joint Photographic Experts Group*

LDA *Linear discriminant analysis* (análise de discriminantes lineares)

MACE *Minimum average correlation energy*

ME *Mean estimator* (estimador da média)

MLE *Maximum likelihood estimator* (estimador de máxima verossimilhança)

NCC *Normalized cross-correlation* (correlação cruzada normalizada)

PCA *Principal component analysis* (análise de componentes principais)

PCAI8 *Predictor based on eight-neighbor context-adaptive interpolation algorithm* (preditor baseado no algoritmo de interpolação de contexto adaptativo com vizinhança de 8 *pixels*)

PCE *Peak-to-correlation energy* (pico de energia de correlação)

PDF *Probability density function* (função densidade de probabilidade)

PME *Phase mean estimator*

PRNU *Photo-response nonuniformity* (não uniformidade da fotorresposta)

RGB *Red, green and blue* (vermelho, verde e azul)

ROC *Receiver operating characteristic*

SNR *sinal-to-noise ratio* (relação sinal ruído)

SVM *Support vector machine* (máquina de vetores de suporte)

TFP Taxa de falso positivo

TVP Taxa de verdadeiro positivo

Lista de Símbolos

- $\oslash$ Divisão de Hadamard (divisão elemento-a-elemento)
- $\succeq$ Inequação (maior ou igual) elemento-a-elemento
- $^{\circ}$ Potência de Hadamard (potência elemento-a-elemento)
- $\circ$ Produto de Hadamard (produto elemento-a-elemento)
- $\otimes$ Produto de Kronecker

Sumário

Capítulo 1

Introdução

Este trabalho aborda o problema de identificação de dispositivo de aquisição de imagem em análise forense. Tal problema consiste em verificar se uma imagem sob investigação foi adquirida por um determinado dispositivo (como câmera digital, por exemplo) sob suspeição [1] em casos que envolvam, por exemplo, pirataria, pornografia infantil ou pedofilia. Algumas técnicas de identificação de dispositivo exploram informações inseridas no arquivo de imagem pelo dispositivo de aquisição, como o número de série do aparelho no cabeçalho do arquivo ou uma marca d'água digital [1]. Outras técnicas exploram artefatos deixados na imagem durante o processo de aquisição, como, por exemplo, distorções das lentes [2], padrão de interpolação [3], poeiras no sensor [4], defeitos dos elementos fotossensores [5] e ruído de padrão fixo [6]. Um dos problemas dessas abordagens é que elas podem ser facilmente fraudadas utilizando operações de processamento de imagens, tais como, compressão com perdas, rotação e redimensionamento. Como alternativa, uma outra peculiaridade do processo de aquisição que vem sendo explorada é a não uniformidade da fotorresposta (*photo-response nonuniformity* – PRNU) [7]-[35]. A PRNU é um padrão de ruído causado pelas imperfeições do processo de manufatura do fotossensor e apresenta as seguintes características: a) é única para cada sensor; b) está presente em todas as imagens adquiridas pelo sensor; c) é robusta às várias operações de processamento de imagem; d) é estável no tempo e sob diferentes condições ambientais. Assim, a PRNU é usualmente considerada a *impressão digital* ou *marca d'água natural* do dispositivo, o que motiva o seu uso em aplicações forenses de identificação de dispositivo de aquisição.

O processo de identificação de dispositivo de aquisição baseado em

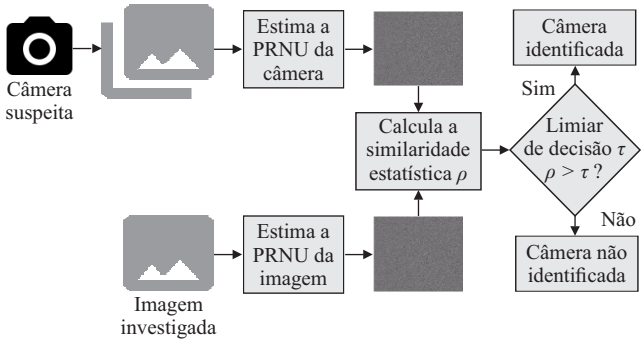


Figura 1.1: Processo de identificação de dispositivo de aquisição de imagem baseado em PRNU.

PRNU envolve duas etapas: *estimação* e *detecção*. Na primeira etapa, estima-se a impressão digital (PRNU) da câmera suspeita a partir de um conjunto de imagens comprovadamente adquiridas por ela. Essa estimativa pode ser realizada através da média de várias imagens de cenas aleatórias. Para reduzir o número de imagens necessárias, é utilizado um filtro de extração de ruído para suprimir o conteúdo da imagem. Na segunda etapa do processo de identificação de dispositivo baseada em PRNU (a detecção), utiliza-se uma métrica de similaridade estatística para detectar a impressão digital do dispositivo na imagem investigada. Caso o valor obtido por essa métrica seja maior do que um limiar de decisão pré-estabelecido, diz-se que a imagem foi adquirida pela câmera suspeita [7]. A Figura 1.1 ilustra esse processo de identificação de dispositivo baseado em PRNU. Diversos trabalhos encontrados na literatura abordam diferentes particularidades desse problema de identificação de dispositivos. A seguir, são descritos alguns desses trabalhos.

1.1 Estado da arte

Diferentes abordagens de estimação da PRNU são encontradas na literatura. Por exemplo, Lukáš *et al.* [8] propuseram um estimador da PRNU

baseado na média do ruído residual (obtido com o filtro de extração de ruído) de várias imagens. Chen *et al.* [9] desenvolveram um estimador de máxima verossimilhança da PRNU. Além disso, esses autores observaram que a estimativa da PRNU contém alguns artefatos causados por operações de aquisição de imagens como interpolação dos canais de cor e compressão com perdas. Para suprimir tais artefatos, a média de cada linha e coluna da PRNU é removida e, em seguida, um filtro de Wiener no domínio da frequência é aplicado [9]. Li [10] apresentou vários métodos para atenuar artefatos presentes no ruído residual da imagem devido a detalhes (bordas) da imagem. Em [11], foi elaborada uma técnica para reduzir o efeito do ruído de interpolação no desempenho do filtro de extração de ruído. Tal técnica decompõe cada canal da imagem em quatro subimagens e aplica, em cada subimagem, o filtro de extração de ruído. Então, os ruídos extraídos das subimagens são reagrupados para formar o ruído residual que é utilizado pelo estimador da PRNU. Visando eliminar os artefatos produzidos pelo processo de aquisição de imagem mencionados em [9], o estimador desenvolvido em [12] gera uma estimativa da PRNU com espectro de frequência plano. A técnica proposta em [13] estima a PRNU através da filtragem homomórfica da média de um conjunto de imagens extraídas (tiradas) de ruídos aleatórios, exibidos em um monitor de alta definição. Kang *et al.* [14] elaboraram um estimador da média da PRNU utilizando um filtro de extração de ruído baseado no algoritmo de interpolação de contexto adaptativo. A abordagem discutida em [15] seleciona blocos de imagens com pouca textura (os quais contém menos influência dos detalhes da imagem), ponderados por diferentes valores, para compor a impressão digital do dispositivo.

Alguns trabalhos tratam do problema de definição da métrica de similaridade estatística a ser utilizada na etapa de detecção do processo de identificação de dispositivo. O objetivo é utilizar a métrica que melhor separe a classe de imagens adquiridas pela câmera suspeita da classe de imagens não adquiridas. Algumas métricas utilizadas nesse contexto são: coeficiente de correlação de Pearson (*Pearson's correlation coefficient* – CC) [7],

CC janelado [9], [16], razão entre o maior e o segundo maior pico (*ratio between the primary peak to the secondary peak* – PSR) da correlação cruzada circular [17], pico de energia de correlação (*peak-to-correlation energy* – PCE) [18], e razão entre a correlação e a norma da correlação cruzada circular (*correlation over circular cross-correlation norm* – CCN) [12].

Após definir a métrica de similaridade estatística a ser utilizada, é preciso definir o limiar de decisão. Para tal, duas soluções geralmente aplicadas são: 1) modelar a função densidade de probabilidade (*probability density function* – PDF) da métrica utilizada; 2) empregar alguma técnica de aprendizado de máquina, como máquina de vetores de suporte (*support vector machine* – SVM), por exemplo. Lukáš *et al.* [7] representaram a PDF do CC como um modelo gaussiano generalizado. Os parâmetros desse modelo são estimados a partir dos CCs de um conjunto de imagens de testes. Chen *et al.* [9] utilizaram o CC janelado e o CC predito – obtido a partir de características da imagem, tais como, intensidade e textura – para calcular os parâmetros do modelo gaussiano generalizado. Goljan [19] considerou a PDF do PCE como uma distribuição tipo chi-quadrado com um grau de liberdade para testes com imagens de mesma resolução que a PRNU. Em [20], Goljan *et al.* calcularam o valor de PCE para mais de um milhão de imagens – obtidas da *web* – adquiridas por quase 6900 câmeras. Esses autores concluíram que o valor de PCE (limiar de decisão) 60 equivale a uma taxa de falso positivo (dizer que a imagem foi adquirida pela câmera suspeita quando ela não foi) de 10^{-5}. Tsai *et al.* [21] utilizaram os três primeiros momentos centrais de cada canal da PRNU e 34 valores de medidas de qualidade de imagens para treinar um classificador SVM. Em [16], o classificador SVM foi treinado utilizando como características a média do CC janelado e a média do CC predito.

Também há trabalhos que avaliam a robustez da PRNU em relação a operações de rotação e redimensionamento de imagem, por exemplo. Nas técnicas desenvolvidas em [22], o fator de escalamento da imagem é encontrado experimentalmente, testando os valores que maximizam o PCE, e os parâmetros de corte da imagem são determinados a partir do máximo da cor-

relação cruzada normalizada (*normalized cross-correlation* – NCC). Essas técnicas são utilizadas em [23] para realizar identificação de dispositivo de aquisição a partir de imagens impressas. A versão digital da imagem impressa é obtida por um *scanner*. Goljan *et al.* [24] desenvolveram estratégias para revelar atividade contra-forense, na qual a impressão digital de uma câmera é inserida em uma imagem adquirida por outra câmera. Em [25], os autores demonstraram que imagens adquiridas com distâncias focais (efeito *zoom*) diferentes apresentam impressões digitais dessincronizadas. Essa dessincronização é causada pelas operações de correção das distorções causadas pelo efeito *zoom*. PRNUs dessincronizadas resultam em baixo valor da métrica de similaridade estatística utilizada no processo de identificação de um dispositivo. Para reduzir esse problema, os autores de [25] adotaram um modelo de distorção e buscaram os parâmetros desse modelo que maximizam o PCE.

Um outro grupo de trabalhos aborda o problema de armazenamento da impressão digital e de buscas por uma impressão em um grande banco de impressões digitais. A PRNU é um ruído aleatório com as mesmas dimensões do fotossensor. Por isso, armazenar uma grande quantidade de impressões digitais requer bastante espaço em memória. Esse problema ocorre em aplicações, por exemplo, de buscas em uma base de imagens na *web* [26] e em agrupamento (clusterização) de imagens [27]. Os trabalhos que tratam de tal problema objetivam comprimir a PRNU, sem aumentar a taxa de erro do sistema de identificação. A abordagem proposta em [28] utiliza um conjunto fixo dos maiores valores (em magnitude) da impressão digital e os índices (posições) desses valores. Bayram *et al.* [29] elaboraram uma abordagem de quantização binária ($+1$ ou -1) da impressão digital. Valsesia *et al.* [27] propuseram a redução da dimensão da PRNU utilizando projeções aleatórias para gerar uma matriz de projeção. Posteriormente, os mesmos autores desenvolveram um método de busca mais rápido [26] e aplicaram esse processo de compressão da PRNU a imagens escaladas [30]. Nessa abordagem, a impressão digital da imagem escalada é mapeada para o espaço comprimido

da PRNU, visando tornar a compressão robusta ao escalamento da imagem. Em [31], os autores utilizaram análise de componentes principais (*principal component analysis* – PCA) e análise de discriminantes lineares (*linear discriminant analysis* – LDA) após a estimação da PRNU. Um método de agrupamento de imagens utilizando árvore de busca binária foi apresentado em [32].

A PRNU também é utilizada para identificar a câmera que adquiriu um determinado vídeo. A abordagem apresentada em [18] estima a PRNU de um vídeo através do estimador de máxima verossimilhança e utiliza o PCE na etapa de detecção. Pande *et al.* [33] desenvolveram uma arquitetura em *hardware* para realizar identificação de dispositivo de aquisição de vídeos em tempo real em uma rede de câmeras. Hyun *et al.* [34] utilizaram um filtro MACE-MRH (*minimum average correlation energy mellin radial harmonic*) – o qual é menos sensível a operações de escalamento –, após a estimação da PRNU, para realizar identificação de dispositivo de aquisição de vídeos escalados.

1.2 Objetivos

O foco deste trabalho de pesquisa reside na etapa de estimação da impressão digital do dispositivo. Nesse contexto, este trabalho apresenta uma abordagem de estimação da PRNU para dispositivos de aquisição equipados com um arranjo de filtros de cor (*color filter array* – CFA). A maioria dos equipamentos de aquisição utiliza um CFA, para separar as informações de cor da luz, e um processo de interpolação para capturar uma imagem digital colorida [36]. Usualmente, o processo de estimação da PRNU a partir de um conjunto de imagens coloridas é realizado aplicando uma das seguintes estratégias: 1) as estimativas da PRNU de cada canal de cor são obtidas e, em seguida, essas estimativas são combinadas (tipicamente, por uma conversão para escala de cinza) [7], [35]; 2) a PRNU é estimada apenas para o canal de cor verde [12]; 3) a PRNU é obtida apenas do canal de luminância (que é

uma combinação linear dos canais vermelho, verde e azul) [14]. Em contraste com essas abordagens, o estimador da PRNU aqui proposto considera cada canal da imagem como uma amostra da PRNU, sem fazer distinção entre eles. Além disso, cada canal da imagem é decomposto em subimagens para atenuar o impacto do ruído de interpolação, assim como em [11]. Porém, diferentemente da abordagem de [11], todo o processo de estimação da PRNU é realizado com as subimagens. Assim, sub-PRNUs são obtidas a partir das subimagens, sendo tais sub-PRNUs agrupadas para formar a PRNU que é utilizada na etapa de detecção do processo de identificação de dispositivo. Resultados de experimentos realizados com imagens obtidas por câmeras digitais e câmeras digitais de telefones celulares mostram que a abordagem aqui proposta apresenta melhor desempenho, em termos de taxa de sucesso na identificação do dispositivo, comparada às metodologias apresentadas em [9], [11], [12] e [14], particularmente, para imagens de baixa resolução.

1.3 Organização do trabalho

Este trabalho está organizado da seguinte forma. O Capítulo 2 apresenta um modelo simplificado do processo de aquisição de imagens digitais e revisita o problema de identificação de dispositivo de aquisição baseada em PRNU. O Capítulo 3 introduz a abordagem proposta neste trabalho. O Capítulo 4 descreve os experimentos e discute os resultados obtidos. Por fim, o Capítulo 5 apresenta as conclusões e as considerações finais deste trabalho.

Capítulo 2

Identificação de dispositivo de aquisição baseada em PRNU

Este capítulo descreve o modelo típico de aquisição de imagem digital de duas dimensões (2D) e revisita as metodologias de estimação da PRNU apresentadas em [9], [11], [12] e [14], as quais foram derivadas a partir de tal modelo. Além disso, este capítulo descreve o processo de detecção da impressão digital da câmera suspeita na imagem investigada.

2.1 Modelo de aquisição de imagem

A Figura 2.1 ilustra o processo de aquisição de imagem digital comum à maioria dos dispositivos de aquisição. Primeiro, a lente foca os raios de luz capturados a partir de uma cena real sobre o CFA. O CFA tem por finalidade separar as informações de cor da luz, uma vez que os fotossensores não possuem essa capacidade. Geralmente, o CFA empregado é o filtro de *Bayer*, que decompõe a luz em seus componentes vermelho, verde e azul, – padrão RGB (*red, green and blue*). Após ser filtrada pelo CFA, a luz incide sobre o fotossensor, que é uma matriz de elementos semicondutores fotossensíveis que capturam a luz incidente, convertendo os fótons em elétrons (efeito foto-elétrico). O sinal que sai do fotossensor é amplificado e convertido em sinal digital por um conversor analógico-digital. Então, o sinal digitalizado segue até a unidade de processamento de sinais que realiza as seguintes operações: escala cada canal de cor por um fator de ganho para atingir o balanço de branco (*white balance*) adequado; interpola as amostras para formar a imagem colorida; corrige a cor e o brilho (fator gama) da imagem para fins de

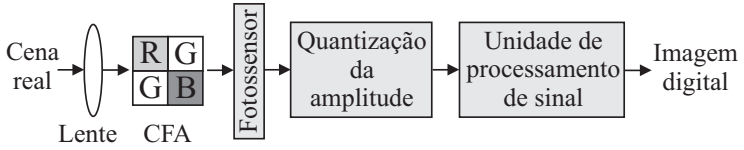


Figura 2.1: Processo de aquisição de imagem digital de duas dimensões (2D).

exibição em um monitor; e por fim, codifica a imagem para ser armazenada, usualmente no formato JPEG (*Joint Photographic Experts Group*) [9], [37].

Seja $\mathbf{Y}$ um canal de cor de dimensão $m \times n$ de uma imagem digital colorida no padrão RGB e $\mathbf{X}$, uma matriz com as mesmas dimensões de $\mathbf{Y}$, a intensidade luminosa que incide sobre o fotossensor. Dessa forma, o processo de aquisição de imagem previamente descrito pode ser modelado para um canal de cor como [38]

$$\mathbf{Y} = g^{\gamma} \left(\mathbf{X} + \mathbf{X} \circ \mathbf{K} + \mathbf{\Omega} \right)^{\circ \gamma} + \mathbf{Q} \tag{2.1}$$

onde g denota o fator de ganho (diferente para cada canal cromático), γ representa o fator de correção gama (geralmente, $\gamma \cong 0,45$), $\circ$ denota o produto de Hadamard (produto elemento-a-elemento), $\mathbf{K}$ corresponde à PRNU, $\mathbf{\Omega}$ representa a combinação de várias outras fontes de ruído (como corrente negra e ruído de leitura [39]), $^{\circ}$ denota a potência de Hadamard (potência elemento-a-elemento) e $\mathbf{Q}$ caracteriza o ruído de quantização e/ou codificação introduzido pelo sistema de aquisição. A faixa de valores dos *pixels* de $\mathbf{Y} = \{\mathbf{Y}(i,j) | 1 \leq i \leq m, 1 \leq j \leq n\}$ depende da profundidade de *bits*. A profundidade de *bits* descreve a quantidade de *bits* utilizados para representar a cor de um *pixel* em uma imagem digital. A maioria das imagens digitais são adquiridas por câmeras digitais que utilizam 8 *bits* por canal de cor. Assim, o canal de cor vermelho, por exemplo, tem uma variação de cor de 256 níveis (de 0 até 255) [37]. Neste trabalho, as operações matemáticas são realizadas em ponto flutuante com precisão de 64 *bits*.

Para simplificar o modelo de aquisição de imagem descrito em (2.1),

aplica-se o seguinte procedimento: primeiro, o termo dominante $\mathbf{X}$ é colocado em evidência, isto é,

$$\mathbf{Y} = (g\mathbf{X})^{\circ\gamma} \circ (\mathbf{1} + \mathbf{K} + \mathbf{\Omega} \oslash \mathbf{X})^{\circ\gamma} + \mathbf{Q} \tag{2.2}$$

onde $\mathbf{1}$ representa uma matriz com todos os elementos iguais a um, e $\oslash$ denota a divisão de Hadamard (divisão elemento-a-elemento); em seguida, aplica-se uma expansão em série de Taylor à parcela $(\mathbf{1} + \mathbf{K} + \mathbf{\Omega} \oslash \mathbf{X})^{\circ\gamma}$, para $\mathbf{K} + \mathbf{\Omega} \oslash \mathbf{X} = \mathbf{0}$, sendo $\mathbf{0}$ uma matriz com todos os elementos iguais a zero, e mantém-se apenas os termos de primeira ordem dessa expansão. Assim, o modelo de aquisição de imagem é simplificado para

$$\begin{aligned} \mathbf{Y} &= (g\mathbf{X})^{\circ\gamma} \circ (\mathbf{1} + \gamma\mathbf{K} + \gamma\mathbf{\Omega} \oslash \mathbf{X}) + \mathbf{Q} \\ &= \mathbf{X}^{(0)} + \mathbf{X}^{(0)} \circ \mathbf{K} + \mathbf{\Lambda} \end{aligned} \tag{2.3}$$

onde $\mathbf{X}^{(0)} = (g\mathbf{X})^{\circ\gamma}$ representa a saída do sensor na ausência de fontes de ruído, $\mathbf{K} = \gamma\mathbf{K}$ para facilitar a notação, $\mathbf{X}^{(0)} \circ \mathbf{K}$ denota o termo da PRNU e $\mathbf{\Lambda} = \mathbf{X}^{(0)} \circ \gamma\mathbf{\Omega} \oslash \mathbf{X} + \mathbf{Q}$ modela a combinação das demais fontes de ruídos [38].

2.2 Estimação da PRNU

As técnicas de estimação da PRNU objetivam obter uma estimativa de $\mathbf{K}$ a partir de uma imagem ou de um conjunto de imagens adquiridas pela câmera suspeita. Dado que $\mathbf{K}$ é um sinal com nível de potência muito baixa [38], geralmente, utiliza-se um filtro de extração de ruído para melhorar a razão sinal-ruído (*sinal-to-noise ratio* – SNR) entre o termo da PRNU ($\mathbf{X}^{(0)} \circ \mathbf{K}$) e a imagem observada $\mathbf{Y}$ [veja (2.3)]. Dessa maneira, as técnicas de estimação são derivadas a partir do ruído residual $\mathbf{W}$ da imagem, o qual é dado por

$$\mathbf{W} = \mathbf{Y} - F(\mathbf{Y}) = \mathbf{X}^{(0)} \circ \mathbf{K} + \mathbf{\Xi} \tag{2.4}$$

onde $F(\cdot)$ é o filtro de extração de ruído, e $\boldsymbol{\Xi}$ representa $\boldsymbol{\Lambda}$ [os ruídos indesejados do modelo de imagem descrito em (2.3)] e as distorções causadas pelo filtro de extração de ruído. Neste trabalho, são considerados dois filtros de extração de ruído: um filtro de Wiener no domínio da transformada *wavelet* [38], [40]; e um filtro baseado no algoritmo de interpolação de contexto adaptativo seguido de um filtro de Wiener no domínio da sequência [14].

Em aplicações práticas, não se tem acesso a $\mathbf{X}^{(0)}$ (já que esse sinal é a saída ideal do sensor). A alternativa empregada é fazer $\mathbf{X}^{(0)} = F(\mathbf{Y})$ ou $\mathbf{X}^{(0)} = \mathbf{Y}$ [35], sendo esta última a solução mais usual [9]. Assim, o ruído residual da imagem é tratado aqui como sendo

$$\mathbf{W} = \mathbf{Y} \circ \mathbf{K} + \boldsymbol{\Xi}. \tag{2.5}$$

A seguir, são descritos quatro estimadores da PRNU, desenvolvidos a partir de (2.5), os quais são considerados neste trabalho.

2.2.1 Abordagem MLE

Esta seção apresenta o estimador de máxima verossimilhança (*maximum likelihood estimator* – MLE) da PRNU proposto em [9]. Tal abordagem considera um conjunto de d imagens adquiridas pela câmera suspeita e, a partir de (2.5), escreve-se para $k = 1, ..., d$,

$$\mathbf{W}_k \oslash \mathbf{Y}_k = \mathbf{K} + \boldsymbol{\Xi}_k \oslash \mathbf{Y}_k. \tag{2.6}$$

Então, assume-se que a sequência de d amostras de cada elemento $\boldsymbol{\Xi}_k(i,j)$ é um ruído gaussiano branco com variância σ^2. Desse modo, a função log-verossimilhança de $\mathbf{W}_k \oslash \mathbf{Y}_k$ dado $\mathbf{K}$ é

$$\begin{aligned} L(\mathbf{K}) = &- \frac{d}{2} \sum_{k=1}^{d} \log\left(2\pi\sigma^2 \mathbf{1} \oslash \mathbf{Y}_k^{\circ 2}\right) \\ &- \sum_{k=1}^{d} \left(\mathbf{W}_k \oslash \mathbf{Y}_k - \mathbf{K}\right)^{\circ 2} \oslash \left(2\sigma^2 \mathbf{1} \oslash \mathbf{Y}_k^{\circ 2}\right). \end{aligned} \tag{2.7}$$

Derivando $L(\mathbf{K})$ em relação aos elementos de $\mathbf{K}$ e igualando a expressão resultante a zero, a seguinte estimativa de máxima verossimilhança da PRNU é obtida:

$$\hat{\mathbf{K}} = \left(\sum_{k=1}^{d} \mathbf{W}_k \circ \mathbf{Y}_k \right) \oslash \sum_{k=1}^{d} \mathbf{Y}_k^{\circ 2} \tag{2.8}$$

onde $\hat{\mathbf{K}}$ é uma estimativa de $\mathbf{K}$. A partir da derivada de segunda ordem de $L(\mathbf{K})$, tem-se o limite inferior de Cramer-Rao (*Cramer-Rao lower bound* – CRLB) da variância de $\hat{\mathbf{K}}$,

$$\mathrm{var}(\hat{\mathbf{K}}) \succeq \mathbf{1} \oslash -\mathrm{E}\left[\partial^2 L(\mathbf{K}) \oslash \partial^2 \mathbf{K}^{\circ 2}\right] = \sigma^2 \mathbf{1} \oslash \sum_{k=1}^{d} \mathbf{Y}_k^{\circ 2} \tag{2.9}$$

onde $\succeq$ representa a inequação maior ou igual aplicada elemento-a-elemento e $\mathrm{E}[\cdot]$ representa o operador valor esperado.

Como o modelo utilizado para obter $\hat{\mathbf{K}}$ é linear, o MLE da PRNU é o estimador não-polarizado de variância mínima e o CRLB é a variância desse estimador [9]. Portanto, conclui-se que as imagens que minimizam $\mathrm{var}(\hat{\mathbf{K}})$ são aquelas de alta luminosidade e pequena σ^2 (o que significa imagens com textura suave). Assim, recomenda-se o uso, por exemplo, de imagens de céu nublado. A quantidade de imagens d a ser utilizada depende da câmera em questão. Em geral, estimativas adequadas são obtidas com $d = 20$ conforme descrito em [35].

2.2.2 Abordagem CDMLE

Como descrito no Capítulo 1, a maioria dos equipamentos de aquisição de imagens utiliza um processo de interpolação para gerar a imagem digital colorida. Tal processo correlaciona as amostras da imagem, como também as amostras do ruído presente na imagem. Portanto, a interpolação prejudica o desempenho do filtro de extração de ruído [11] – o qual, geralmente, assume que as amostras do ruído são independentes e identicamente distribuídas [38]. A abordagem apresentada em [11], denominada CDMLE

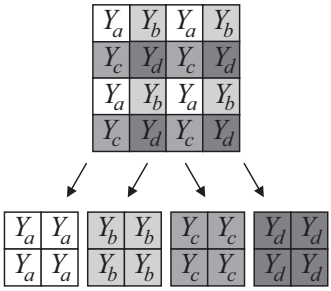


Figura 2.2: Separação de uma imagem (matriz) de resolução $m \times n$ em quatro subimagens (submatrizes) de tamanho $m/2 \times n/2$.

(*color-decoupled maximum likelihood estimator*), tem por objetivo reduzir essa influência da interpolação no desempenho do filtro de extração de ruído. Tal metodologia decompõe cada canal da imagem em quatro subimagens, como ilustrado na Figura 2.2, e aplica, em cada subimagem, o filtro de extração de ruído. Assim, em algumas das subimagens estarão presentes apenas as amostras capturadas pelo fotossensor, o que resulta nas melhores estimativas do ruído residual da imagem. Em seguida, fazendo-se o processo inverso ao ilustrado na Figura 2.2, os ruídos extraídos das subimagens são reagrupados para formar o ruído residual que é utilizado pelo MLE da PRNU (2.8).

2.2.3 Abordagem PME

Segundo [12], o ruído residual da imagem, descrito por (2.5), contém alguns artefatos indesejados causados pelos detalhes da cena capturada e por operações do processo de aquisição, tais como, interpolação dos canais de cor e projeto do fotossensor. Esses artefatos aparecem como picos no domínio da frequência. Para suavizar tais artefatos, os autores de [12] transformaram $\mathbf{W}_k$ em um sinal com espectro de frequência plano, isto é,

$$\mathbf{\Phi}_k = \mathcal{F}(\mathbf{W}_k) \oslash |\mathcal{F}(\mathbf{W}_k)| \tag{2.10}$$

onde $\mathcal{F}(\cdot)$ é o operador de transformada de Fourier e $|\mathcal{F}(\mathbf{W}_k)|$ é a magnitude de $\mathcal{F}(\mathbf{W}_k)$. A estimativa da PRNU é então obtida por

$$\hat{\mathbf{K}} = \mathrm{Re}\left[\mathcal{F}^{-1}\left(\frac{1}{d}\sum_{k=1}^{d}\mathbf{\Phi}_k\right)\right] \tag{2.11}$$

onde $\mathrm{Re}[z]$ retorna apenas a parte real de z. Esse estimador é chamado de PME (*phase mean estimator*) [12].

2.2.4 Abordagem PCAI8 ME

As metodologias de estimação da PRNU descritas nas Seções 2.2.1 a 2.2.3 utilizam, como filtro de extração de ruído, o filtro de Wiener no domínio da transformada *wavelet* [38], [40]. Por outro lado, o estimador da média (*mean estimator* – ME) proposto em [14] utiliza o preditor baseado no algoritmo de interpolação de contexto adaptativo com vizinhança de 8 *pixels* (*predictor based on eight-neighbor context-adaptive interpolation algorithm* – PCAI8) e um filtro de Wiener no domínio da sequência. Primeiro, os pixels da imagem são estimados com o PCAI8 e, em seguida, o erro entre a imagem original e a imagem predita é filtrado usando um filtro de Wiener no domínio da sequência. A estimativa da PRNU é então dada por

$$\hat{\mathbf{K}} = \frac{1}{d}\sum_{k=1}^{d}\mathbf{W}_k. \tag{2.12}$$

De acordo com [14], esse estimador requer no mínimo $d = 100$ imagens para estimar a impressão digital do dispositivo. No restante deste trabalho, tal abordagem é chamada de PCAI8 ME.

2.2.5 Pós-processamento da PRNU

Como discutido em [9], a estimativa $\hat{\mathbf{K}}$ da PRNU contém artefatos introduzidos por operações do processo de aquisição de imagem, como, por

exemplo, interpolação dos canais de cor e compressão com perdas. Esses artefatos são comuns a impressões digitais de dispositivos de aquisição de mesmo modelo (ou mesmo projeto) do fotossensor e, portanto, eles podem aumentar a taxa de erro do sistema de identificação. Os artefatos causados pelos algoritmos de interpolação são, em geral, periódicos e dependem do padrão do CFA implementado. Para atenuar esses ruídos, $\hat{\mathbf{K}}$ é decomposta em quatro submatrizes (como ilustrado na Figura 2.2) e, de cada submatriz, as médias de cada linha e coluna são removidas. Em seguida, as submatrizes são reagrupadas (fazendo o processo inverso ao descrito na Figura 2.2). Então, os demais artefatos presentes na PRNU são mitigados por um filtro de Wiener no domínio da frequência [9], isto é,

$$\mathbf{F} = \mathcal{F}(\hat{\mathbf{K}}), \quad \hat{\mathbf{K}} \leftarrow \mathrm{Re}\left(\mathcal{F}^{-1}\left\{\mathbf{F} \circ \left[|\mathbf{F}| - W(|\mathbf{F}|, \sigma^2)\right] \oslash |\mathbf{F}|\right\}\right) \tag{2.13}$$

onde $W(\cdot,\cdot)$ é o filtro de Wiener, e σ^2 é a variância de $\hat{\mathbf{K}}$ [35]. A diferença entre as PRNUs sem e com pós-processamento é chamado de padrão linear [9]. Esse padrão linear pode ser utilizado em aplicações de identificação do modelo do dispositivo de aquisição [41].

2.2.6 Estimação da PRNU a partir de imagens coloridas

Usualmente, a estimação da impressão digital utilizando imagens coloridas envolve uma dentre as seguintes abordagens:

1. As estimativas da PRNU de cada canal da imagem são obtidas, sendo tais estimativas posteriormente combinadas utilizando uma conversão para escala de cinza, como, por exemplo,

$$\hat{\mathbf{K}} = 0,3\hat{\mathbf{K}}_{\mathrm{R}} + 0,6\hat{\mathbf{K}}_{\mathrm{G}} + 0,1\hat{\mathbf{K}}_{\mathrm{B}} \tag{2.14}$$

   onde $\hat{\mathbf{K}}_{\mathrm{R}}$, $\hat{\mathbf{K}}_{\mathrm{G}}$ e $\hat{\mathbf{K}}_{\mathrm{B}}$ são as estimavas da PRNU dos canais vermelho, verde e azul, respectivamente [7], [35].

2. Apenas a estimativa da PRNU do canal de cor verde é utilizada [12].

Figura 2.3: Diferentes estruturas de CFA no padrão RGB.

Geralmente, o CFA implementado pelas câmeras digitais é um daqueles no padrão RGB mostrados na Figura 2.3. Nessas configurações, metade das amostras capturadas pelo fotossensor são do canal de cor verde. Portanto, esse canal contém mais informações da PRNU do que os canais vermelho e azul.

3. A estimativa da PRNU é obtida a partir do canal de luminância [14]. Antes do processo de estimação, os canais de cor da imagem no padrão RGB são combinados para formar o canal de luminância, isto é,

$$\mathbf{Y}_{\mathrm{L}} = 0,3\mathbf{Y}_{\mathrm{R}} + 0,6\mathbf{Y}_{\mathrm{G}} + 0,1\mathbf{Y}_{\mathrm{B}} \tag{2.15}$$

onde $\mathbf{Y}_{\mathrm{L}}$ é o canal de luminância, e $\mathbf{Y}_{\mathrm{R}}$, $\mathbf{Y}_{\mathrm{G}}$ e $\mathbf{Y}_{\mathrm{B}}$ representam os canais de cor vermelho, verde e azul, respectivamente.

Como será mostrado no Capítulo 3, a abordagem proposta neste trabalho trata os canais de cor de forma igual, utilizando todos os canais para obter apenas uma estimativa da PRNU, diferindo assim das abordagens de estimação descritas neste capítulo.

2.3 Detecção da PRNU

Na etapa de detecção do processo de identificação de dispositivo baseada em PRNU, o objetivo é utilizar uma métrica de similaridade estatística para avaliar a presença da impressão digital $\hat{\mathbf{K}}$ da câmera suspeita na imagem investigada $\mathbf{Y}_{\mathrm{t}}$. Esse processo pode ser representado por um problema de

teste de hipótese binária como

$$\begin{aligned} \mathrm{H}_0 &: \hat{\mathbf{K}} \neq \hat{\mathbf{K}}_\mathrm{t} \\ \mathrm{H}_1 &: \hat{\mathbf{K}} = \hat{\mathbf{K}}_\mathrm{t} \end{aligned} \tag{2.16}$$

onde $\hat{\mathbf{K}}_\mathrm{t}$ é a PRNU da imagem avaliada. A hipótese nula H_0 significa que a imagem investigada não foi adquirida pela câmera suspeita (isto é, as impressões digitais são diferentes), enquanto a hipótese alternativa H_1 afirma o contrário. Para avaliar essas hipóteses, este trabalho utiliza a CCN – uma métrica de similaridade estatística que apresenta menor taxa de erro em sistemas de identificação de dispositivo de aquisição baseado em PRNU [12]. A CCN é definida como [12]

$$\mathrm{CCN}(\mathbf{A}, \mathbf{B}) = \frac{R_{\mathbf{AB}}(0,0)}{\sqrt{\frac{1}{mn-|\mathbf{S}|}\sum_{i,j\notin\mathbf{S}} R^2_{\mathbf{AB}}(i,j)}} \tag{2.17}$$

onde $R_{\mathbf{AB}}(i,j)$ representa a correlação cruzada circular entre os sinais $\mathbf{A}$ e $\mathbf{B}$, $\mathbf{S}$ denota uma área em torno de $R_{\mathbf{AB}}(0,0)$, e $|\mathbf{S}|$ caracteriza o tamanho da área $\mathbf{S}$. Em [20], $\mathbf{A}$ e $\mathbf{B}$ são definidos como $\mathbf{A} = \hat{\mathbf{K}} \circ \mathbf{Y}_\mathrm{t}$ e $\mathbf{B} = \mathbf{W}_\mathrm{t}$.

A decisão por H_0 ou H_1 é tomada comparando o valor de CCN com um limiar de decisão τ. Então, caso $\mathrm{CCN}(\mathbf{A}, \mathbf{B}) > \tau$, H_1 é aceita, isto é, a imagem $\mathbf{Y}_\mathrm{t}$ é dita adquirida pelo dispositivo suspeito. Caso contrário, H_1 é rejeitada.

2.4 Conclusões

Neste capítulo, o modelo de aquisição de imagens digitais e um processo utilizado para simplificar tal modelo foram revisitados. Também foram descritas as abordagens de estimação da PRNU apresentadas em [9], [11], [12] e [14], as quais serão utilizadas posteriormente para fins de comparação com a abordagem proposta neste trabalho. Ainda neste capítulo, foram discutidas as técnicas de pós-processamento da PRNU e as estratégias utilizadas na estimação da PRNU a partir de um conjunto de imagens coloridas. Por fim,

foi descrito o processo de detecção da impressão digital da câmera suspeita na imagem investigada.

Capítulo 3

Abordagem Proposta de Estimação da PRNU

No melhor dos cenários para se estimar a PRNU, tem-se acesso a um conjunto de imagens no formato RAW garantidamente adquiridas pela câmera suspeita. Imagens nesse formato contêm os *pixels* capturados pelo fotossensor sem que eles tenham sido submetidos a operações de processamento de imagem, como interpolação e compressão com perdas, por exemplo. No entanto, na maioria dos casos práticos, as imagens utilizadas para estimar a PRNU foram previamente interpoladas e comprimidas no formato JPEG. Visando amenizar o impacto do processo de interpolação e compressão na estimação da PRNU e com isso obter melhores resultados no processo de identificação de dispositivo, a abordagem proposta neste trabalho baseia-se inicialmente na decomposição da imagem em subimagens de forma similar à abordagem CDMLE [11] (discutida na Seção 2.2.2). Em seguida, contrastando com a abordagem CDMLE [11], a abordagem aqui proposta aplica os processos de extração de ruído e estimação da PRNU em cada uma das subimagens. Além disso, aqui não é feita qualquer distinção entre os canais de cor de imagens. Antes de apresentar a abordagem proposta, este capítulo discute a influência do ruído de interpolação no processo de extração de ruído da imagem e, consequentemente, na estimação da PRNU.

3.1 Influência do ruído de interpolação na estimação da PRNU

Na Seção 2.1, foi mencionado que a maioria dos equipamentos de aquisição de imagem digital implementa um CFA para decompor a luz em seus componentes cromáticos. Dessa forma, considerando um padrão RGB genérico, como aqueles apresentados na Figura 2.3, nota-se que, em cada

posição (i, j) da imagem digital, apenas um componente cromático é efetivamente adquirido pelo fotossensor. Os demais componentes são resultantes do processo de interpolação. A interpolação combina as amostras dentro de uma janela de observação centrada na posição (i, j) para obter o *pixel* desejado. As técnicas de interpolação podem ser classificadas como não adaptativas e adaptativas [42]. Nas técnicas não adaptativas, os coeficientes, que ponderam a contribuição de cada amostra dentro da janela de observação, são fixos para todos os *pixels* de um mesmo canal, como, por exemplo, nas interpolações bilinear e bicúbica. Por outro lado, as técnicas adaptativas buscam o conjunto de coeficientes que apresente o menor erro de interpolação, tais como a interpolação baseada no gradiente e a de plano de cor adaptativo [42]. Independentemente do algoritmo utilizado, a interpolação insere um tipo específico de correlação entre as amostras da imagem [36], o que afeta a estimação da PRNU e prejudica o processo de identificação do dispositivo.

Para avaliar o impacto da interpolação na estimação da PRNU, considera-se que $\mathbf{Y}(i, j)$ representa o *pixel* na posição (i, j) de um canal de cor de uma imagem adquirida por um dispositivo com uma configuração CFA no padrão RGB (veja exemplos na Figura 2.3), e que o algoritmo de interpolação seja linear e utilize apenas amostras do mesmo canal de cor. Dessa forma, pode-se dizer que $\mathbf{Y}(i, j)$ pertence a um dos seguintes grupos: M_1 se a amostra foi efetivamente capturada pelo fotossensor ou M_2 caso ela seja resultante da operação de interpolação. Assim, o modelo de imagem descrito em (2.1) pode ser reescrito como

$$\mathbf{Y} = \begin{cases} g^{\gamma}\left[\mathbf{X} + \mathbf{X} \circ \mathbf{K} + \mathbf{\Omega}\right]^{\circ\gamma} + \mathbf{Q}, & \mathbf{Y}(i,j) \in M_1 \\ g^{\gamma}\left[\sum\limits_{l,m} \alpha_{l,m}\left(\mathbf{X} + \mathbf{X} \circ \mathbf{K} + \mathbf{\Omega}\right)\right]^{\circ\gamma} + \mathbf{Q}, & \mathbf{Y}(i,j) \in M_2 \end{cases} \tag{3.1}$$

onde α representa os coeficientes do algoritmo de interpolação utilizado, $\alpha = \{\alpha_{l,m} | -N \leq l, m \leq N\}$, para N um número inteiro. Para facilitar o entendimento, os índices (i, j) não são representados em (3.1). Anali-

sando (3.1), é possível verificar que as amostras do grupo M_2 não apresentam informações novas sobre a PRNU, que é o sinal de interesse. Portanto, o melhor cenário para estimar $\mathbf{K}$ a partir de imagens interpoladas é aquele em que a configuração CFA da câmera sob suspeição é conhecida. Nesse caso, apenas as amostras capturadas pelo fotossensor (grupo M_1) podem ser utilizadas. No entanto, a maioria dos equipamentos de aquisição não informa a estrutura CFA utilizada [11], o que impede o uso de tal informação para seleção dos *pixels* e estimação da PRNU.

Com o objetivo de utilizar apenas as amostras capturadas pelo fotossensor no processo de identificação de dispositivo, pode ser utilizada alguma técnica de identificação de configuração CFA, como aquela apresentada em [42]. A abordagem proposta em [42] considera todas as estruturas de CFAs compostas pelo padrão RGB que se repetem a cada conjunto de 2×2 elementos (veja exemplos na Figura 2.3) e busca entre os 36 padrões de CFA aquele que apresenta menor erro entre a imagem avaliada e a imagem obtida utilizando um determinado algoritmo de interpolação. O padrão RGB arranjado a cada conjunto de 2×2 elementos é o padrão de CFA mais popular e implementado pela maioria das câmeras digitais. O problema da abordagem apresentada em [42] é que algumas configurações CFAs, como o super CCD EXR (mostrado na Figura 3.1) – utilizado por alguns modelos de câmeras da Fujifilm –, implementam um padrão CFA diferente dos 36 padrões avaliados em [42]. As câmeras equipadas com o super CDD EXR adquirem duas imagens da mesma cena: uma obtida com alta sensibilidade[1] e a outra com baixa. Após adquirir as duas imagens, as câmeras combinam essas imagens para produzir uma única imagem de melhor qualidade [43]. O algoritmo de detecção de estrutura CFA proposto em [42] só identifica um padrão de CFA se ele fizer parte do espaço de busca do algoritmo. Dado que não se conhecem todos os padrões de CFAs implementados, tal abordagem de identificação de CFA não é utilizada neste trabalho.

[1] A sensibilidade do sensor refere-se à quantidade de variação de luz que ele é capaz de detectar. Sensores com alta sensibilidade detectam pequenas variações de luz.

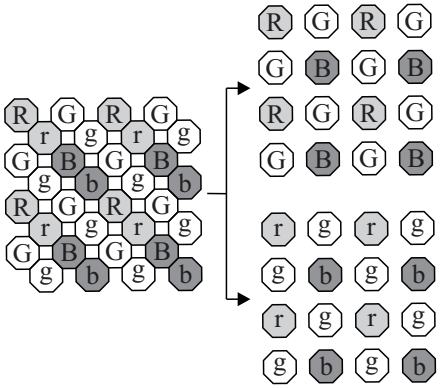


Figura 3.1: Super CCD EXR produzido pela Fujifilm [43]. As células do CFA com letras maiúsculas (R, G, B) representam os *pixels* da imagem adquirida com alta sensibilidade e as de letras minúsculas (r, g, b) os *pixels* de baixa sensibilidade.

As abordagens de estimação da PRNU descritas na Seção 2.2 estimam a PRNU para cada canal da imagem sem distinguir os *pixels* do grupo M_1 do M_2. Desse modo, baseado em (3.1), as amostras da PRNU $\hat{\mathbf{K}}$ podem ser classificadas como

$$\hat{\mathbf{K}} = \begin{cases} \mathbf{K} + \mathbf{\Theta}, & \hat{\mathbf{K}}(i,j) \in M_1 \\ \sum_{l,m} \beta_{l,m}\mathbf{K} + \mathbf{\Theta}, & \hat{\mathbf{K}}(i,j) \in M_2 \end{cases} \tag{3.2}$$

onde $\mathbf{\Theta}$ modela a combinação dos artefatos causados por operações de compressão com perdas, interpolação e erro de estimação [9], e β representa os coeficientes que ponderam a combinação das amostras de M_1, $\beta = \{\beta_{l,m} | -M \leq l,m \leq M\}$, para M um número inteiro. Então, $\hat{\mathbf{K}}$ é submetido a operações de pós-processamento (discutidas na Seção 2.2.5) que incluem o filtro de Wiener descrito em (2.13) para suprimir $\mathbf{\Theta}$. Esse filtro considera que as amostras da PRNU são independentes e identicamente distribuídas

(i.i.d) [40], o que, de acordo com (3.2), não é o caso para $\hat{\mathbf{K}}$. Assim, o filtro de Wiener tende a atenuar também informações relevantes da PRNU. A fim de amenizar tal problema, uma solução, elaborada durante este trabalho de pesquisa, consiste em decompor $\hat{\mathbf{K}}$ em quatro submatrizes e aplicar, para cada submatriz, o filtro de Wiener. Resultados de experimentos utilizando essa abordagem foram apresentados em [44].

3.2 Abordagem MLE proposta

Esta seção apresenta o MLE da PRNU proposto neste trabalho. Para um conjunto de d imagens coloridas, compostas por c canais de cor, adquiridas pela câmera suspeita, cada canal $\mathbf{Y}_{j,k}$ de cada imagem $\mathbf{Y}_k$ é decomposto em quatro subimagens, similarmente à abordagem de [11]. Esse processo de decomposição (ilustrado na Figura 2.2) é representado aqui por

$$\begin{cases} \mathbf{Y}_{1,j,k} = \left(\mathbf{I}_{m/2} \otimes \mathbf{u}_0^{\mathrm{T}}\right) \mathbf{Y}_{j,k} \left(\mathbf{I}_{n/2} \otimes \mathbf{u}_0\right) \\ \mathbf{Y}_{2,j,k} = \left(\mathbf{I}_{m/2} \otimes \mathbf{u}_0^{\mathrm{T}}\right) \mathbf{Y}_{j,k} \left(\mathbf{I}_{n/2} \otimes \mathbf{u}_1\right) \\ \mathbf{Y}_{3,j,k} = \left(\mathbf{I}_{m/2} \otimes \mathbf{u}_1^{\mathrm{T}}\right) \mathbf{Y}_{j,k} \left(\mathbf{I}_{n/2} \otimes \mathbf{u}_0\right) \\ \mathbf{Y}_{4,j,k} = \left(\mathbf{I}_{m/2} \otimes \mathbf{u}_1^{\mathrm{T}}\right) \mathbf{Y}_{j,k} \left(\mathbf{I}_{n/2} \otimes \mathbf{u}_1\right) \end{cases} \tag{3.3}$$

onde $\mathbf{Y}_{s,j,k}$, para $s = 1, 2, 3, 4$, denota a s-ésima subimagem do canal j, para $j = 1, 2, ..., c$ (para imagens no padrão RGB, $c = 3$), da k-ésima imagem, m representa o número de linhas da imagem, n é o número de colunas, $\mathbf{I}_{m/2}$ representa uma matriz identidade de dimensão $m/2$, $\mathbf{u}_0 = [1\ 0]^{\mathrm{T}}$, $\mathbf{u}_1 = [0\ 1]^{\mathrm{T}}$, e $\otimes$ representa o produto de Kronecker [45]. Após a decomposição, cada subimagem $\mathbf{Y}_{s,j,k}$ passa por um filtro de extração de ruído $F(\cdot)$ para obter o ruído residual $\mathbf{W}_{s,j,k}$, isto é,

$$\begin{aligned} \mathbf{W}_{s,j,k} &= \mathbf{Y}_{s,j,k} - F(\mathbf{Y}_{s,j,k}) \\ &= \mathbf{Y}_{s,j,k} \circ \mathbf{K}_s + \mathbf{\Xi}_{s,j,k}. \end{aligned} \tag{3.4}$$

Aqui, assim como em [9], assume-se que a sequência de amostras de cada elemento $\boldsymbol{\Xi}_{s,j,k}(i,j)$ é um ruído gaussiano branco com variância σ^2. Porém, ao invés de ser considerada apenas a sequência de d amostras de cada canal, aqui são incluídas as d amostras de todos os c canais da imagem. Como consequência, o número de amostras de $\boldsymbol{\Xi}_{s,j,k}(i,j)$ é resultante do produto cd. Conforme discutido na Seção 3.1, em cada posição (i,j) de uma imagem digital de c canais, adquirida por uma câmera que implementa um CFA, há uma amostra realmente capturada pelo fotossensor. Portanto, há apenas uma amostra da PRNU na posição (i,j). Desse modo, em um conjunto de cd amostras, há d amostras realmente capturadas pelo fotossensor. Assim, aplicando o procedimento utilizado por [9], o qual foi discutido na Seção 2.2.1, obtém-se o estimador de máxima verossimilhança da PRNU para o conjunto de subimagens s dado por

$$\hat{\mathbf{K}}_s = \left(\sum_{j=1}^{c} \sum_{k=1}^{d} \mathbf{W}_{s,j,k} \circ \mathbf{Y}_{s,j,k} \right) \oslash \sum_{j=1}^{c} \sum_{k=1}^{d} \mathbf{Y}_{s,j,k}^{\circ 2}, \qquad s = 1,2,3,4 \tag{3.5}$$

onde $\hat{\mathbf{K}}_s$ é a subestimativa da PRNU para o conjunto de subimagens s. O CRLB de $\hat{\mathbf{K}}_s$ é

$$\operatorname{var}(\hat{\mathbf{K}}_s) \succeq \sigma^2 \mathbf{1} \oslash \sum_{j=1}^{c} \sum_{k=1}^{d} \mathbf{Y}_{s,j,k}^{\circ 2}. \tag{3.6}$$

Assim como em [9], o CRLB (3.6) é a variância de $\hat{\mathbf{K}}_s$.

Após a estimação das quatro subestimativas da PRNU, são aplicadas as operações de pós-processamento (descritas na Seção 2.2.5): remoção da média de cada linha e coluna de $\hat{\mathbf{K}}_s$, e aplicação do filtro de Wiener no domínio da frequência. Então, as sub-PRNUs são agrupadas (fazendo o processo inverso ao descrito na Figura 2.2) para formar a impressão digital da câmera, isto é,

$$\hat{\mathbf{K}} = \hat{\mathbf{K}}_1 \otimes \left(\mathbf{u}_0 \mathbf{u}_0^{\mathrm{T}}\right) + \hat{\mathbf{K}}_2 \otimes \left(\mathbf{u}_0 \mathbf{u}_1^{\mathrm{T}}\right) + \hat{\mathbf{K}}_3 \otimes \left(\mathbf{u}_1 \mathbf{u}_0^{\mathrm{T}}\right) + \hat{\mathbf{K}}_4 \otimes \left(\mathbf{u}_1 \mathbf{u}_1^{\mathrm{T}}\right). \tag{3.7}$$

Note que a diferença entre a abordagem aqui proposta e aquelas apre-

sentadas na Seção 2.2 está no fato de não haver distinção entre os canais de cor das imagens. Com isso, melhores resultados podem ser obtidos levando em conta que: $i)$ a correlação entre as amostras da subimagem é menor, proporcionando um melhor desempenho do filtro de extração de ruído [38], [40]; e $ii)$ apenas as amostras capturadas pelo fotossensor estarão presentes em algumas das subimagens, resultando em melhores estimativas da PRNU.

Neste trabalho, os canais das imagens são decompostos utilizando apenas a configuração descrita em (3.3). Outra configuração pode ser usada, como, por exemplo, separar cada canal da imagem de 3 em 3 *pixels*. Dessa forma, a tendência é que as amostras do ruído presente na imagem fiquem menos correlacionadas, permitindo um melhor desempenho do filtro de extração de ruído. No entanto, tal abordagem descorrelaciona também as amostras da imagem, o que, por outro lado, pode afetar o desempenho do filtro de extração de ruído.

3.2.1 Análise da variância do MLE proposto

A Seção 2.2.6 descreveu três formas utilizadas na literatura para obter a impressão digital do dispositivo a partir de um conjunto de imagens coloridas. Uma das soluções estima a PRNU apenas do canal de cor verde das imagens. As outras duas estratégias envolvem uma combinação linear das estimativas da PRNU de cada canal de cor ou dos canais de cor da imagem antes do processo de estimação da PRNU. O objetivo desta seção é mostrar que a variância do MLE da PRNU aqui proposto (3.4) é menor ou igual à variância dos estimadores da PRNU que envolvem a combinação linear e o MLE descrito na Seção 2.2.1. Quanto menor a variância do estimador, menor será o erro entre a estimativa e o real valor do sinal analisado [46]. Como a estratégia de estimação da PRNU que envolve apenas o canal de cor verde da imagem utiliza menos amostras do que a estratégia proposta neste trabalho, ela não será avaliada nesta seção.

Para facilitar a notação nesta seção, são consideradas as seguintes de-

finições: $\mathbf{R}$, $\mathbf{G}$ e $\mathbf{B}$ representam os canais de cor de uma imagem no padrão RGB; $\hat{\mathbf{K}}_{\mathrm{p}}$ é a estimativa da PRNU obtida com o estimador MLE proposto neste trabalho; e $\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{p}})$ é dada por

$$\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{p}}) = \sigma^2 \mathbf{1} \oslash \left(\sum_{k=1}^{d} \mathbf{R}_k^{\circ 2} + \sum_{k=1}^{d} \mathbf{G}_k^{\circ 2} + \sum_{k=1}^{d} \mathbf{B}_k^{\circ 2} \right). \tag{3.8}$$

Vale ressaltar que $\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{p}})$ é uma matriz com as mesmas dimensões da PRNU, na qual cada elemento corresponde à variância de cada amostra de $\hat{\mathbf{K}}_{\mathrm{p}}$. Por isso, não é utilizada nesta seção a notação considerada anteriormente para representar subimagens.

Na primeira abordagem descrita na Seção 2.2.6, a estimativa final da PRNU é obtida através da conversão para escala de cinza (2.14) das estimativas da PRNU de cada canal do padrão RGB. Essa combinação é representada aqui por

$$\hat{\mathbf{K}}_{\mathrm{c}} = \alpha_1 \hat{\mathbf{K}}_{\mathrm{R}} + \alpha_2 \hat{\mathbf{K}}_{\mathrm{G}} + \alpha_3 \hat{\mathbf{K}}_{\mathrm{B}} \tag{3.9}$$

onde $\hat{\mathbf{K}}_{\mathrm{c}}$ representa a versão da impressão digital do dispositivo em escala de cinza, e α_1, α_2 e α_3 ponderam, respectivamente, a combinação das estimativas $\hat{\mathbf{K}}_{\mathrm{R}}$, $\hat{\mathbf{K}}_{\mathrm{G}}$ e $\hat{\mathbf{K}}_{\mathrm{B}}$ obtidas por (2.8) a partir dos canais de cor do padrão RGB. Assumindo que as estimativas $\hat{\mathbf{K}}_{\mathrm{R}}$, $\hat{\mathbf{K}}_{\mathrm{G}}$ e $\hat{\mathbf{K}}_{\mathrm{B}}$ são independentes, a variância de $\hat{\mathbf{K}}_{\mathrm{c}}$ é dada por

$$\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{c}}) = \alpha_1^2 \,\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{R}}) + \alpha_2^2 \,\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{G}}) + \alpha_3^2 \,\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{B}}). \tag{3.10}$$

Como $\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{R}})$, $\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{G}})$ e $\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{B}})$ são dadas por (2.9), então

$$\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{c}}) = \sigma^2 \left(\alpha_1^2 \mathbf{1} \oslash \sum_{k=1}^{d} \mathbf{R}_k^{\circ 2} + \alpha_2^2 \mathbf{1} \oslash \sum_{k=1}^{d} \mathbf{G}_k^{\circ 2} + \alpha_3^2 \mathbf{1} \oslash \sum_{k=1}^{d} \mathbf{B}_k^{\circ 2} \right). \tag{3.11}$$

Todas as operações em $\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{c}})$ e $\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{p}})$ são do tipo elemento-a-elemento, portanto, mostrar que $\mathrm{var}(\hat{\mathbf{K}}_{\mathrm{c}}) \succeq \mathrm{var}(\hat{\mathbf{K}}_{\mathrm{p}})$ para o elemento na posição (i, j) é suficiente para provar que a variância do estimador MLE proposto neste trabalho é menor ou igual a do MLE desenvolvido em [9]. Além

disso, utilizar os elementos de $\mathrm{var}(\hat{\mathbf{K}}_\mathrm{c})$ e $\mathrm{var}(\hat{\mathbf{K}}_\mathrm{p})$ permite eliminar as operações de Hadamard (divisão e multiplicação) nos procedimentos que serão empregados a seguir – facilitando o entendimento de tais procedimentos. Então, considerando $\hat{K}_\mathrm{p} = \hat{\mathbf{K}}_\mathrm{p}(i,j)$ e $\hat{K}_\mathrm{c} = \hat{\mathbf{K}}_\mathrm{c}(i,j)$, necessita-se mostrar que

$$\mathrm{var}(\hat{K}_\mathrm{c}) \geq \mathrm{var}(\hat{K}_\mathrm{p}), \tag{3.12}$$

o que equivale a provar que

$$\mathrm{var}(\hat{K}_\mathrm{c}) - \mathrm{var}(\hat{K}_\mathrm{p}) \geq 0. \tag{3.13}$$

Substituindo (3.8) e (3.11) em (3.13), e considerando $\bar{R} = \sum_{k=1}^{d} \mathbf{R}_k^{\circ 2}(i,j)$, $\bar{G} = \sum_{k=1}^{d} \mathbf{G}_k^{\circ 2}(i,j)$ e $\bar{B} = \sum_{k=1}^{d} \mathbf{B}_k^{\circ 2}(i,j)$, tem-se

$$\sigma^2 \left(\frac{\alpha_1^2}{\bar{R}} + \frac{\alpha_2^2}{\bar{G}} + \frac{\alpha_3^2}{\bar{B}} \right) - \frac{\sigma^2}{\bar{R}+\bar{B}+\bar{G}} \geq 0 \tag{3.14}$$

e portanto

$$\left(\alpha_1^2 \bar{G}\bar{B} + \alpha_2^2 \bar{R}\bar{B} + \alpha_3^2 \bar{R}\bar{G}\right)\left(\bar{R}+\bar{G}+\bar{B}\right) - \bar{R}\bar{G}\bar{B} \geq 0. \tag{3.15}$$

Sabe-se que o quadrado da soma de três escalares a, b e c é

$$(a+b+c)^2 = a^2 + b^2 + c^2 + 2ab + 2ac + 2bc. \tag{3.16}$$

Expandindo (3.15) e utilizando (3.16), tem-se

$$\begin{aligned} &\left(\alpha_1+\alpha_2+\alpha_3\right)^2 \bar{R}\bar{G}\bar{B} + \left(\alpha_2\bar{B} - \alpha_3\bar{G}\right)^2 \bar{R} \\ &\quad + \left(\alpha_1\bar{B} - \alpha_3\bar{R}\right)^2 \bar{G} + \left(\alpha_1\bar{G} - \alpha_2\bar{R}\right)^2 \bar{B} - \bar{R}\bar{G}\bar{B} \geq 0. \end{aligned} \tag{3.17}$$

Geralmente, em uma conversão para escala de cinza como a de (2.14), tem-se $\alpha_1+\alpha_2+\alpha_3 = 1$. Então, substituindo $\alpha_1+\alpha_2+\alpha_3 = 1$ em (3.17), obtém-se

$$\left(\alpha_2\bar{B} - \alpha_3\bar{G}\right)^2 \bar{R} + \left(\alpha_1\bar{B} - \alpha_3\bar{R}\right)^2 \bar{G} + \left(\alpha_1\bar{G} - \alpha_2\bar{R}\right)^2 \bar{B} \geq 0. \tag{3.18}$$

Como as parcelas da soma em (3.18) serão sempre positivas, conclui-se que

$\mathrm{var}(\hat{K}_\mathrm{c}) \geq \mathrm{var}(\hat{K}_\mathrm{p})$ e, portanto, $\mathrm{var}(\hat{\mathbf{K}}_\mathrm{c}) \succeq \mathrm{var}(\hat{\mathbf{K}}_\mathrm{p})$.

A outra estratégia de estimação da PRNU, discutida na Seção 2.2.6, combina os canais de cor das imagens para formar o canal de luminância e estima a PRNU a partir desse canal. Essa combinação é representada aqui como

$$\mathbf{Y} = \alpha_1\mathbf{R} + \alpha_2\mathbf{G} + \alpha_3\mathbf{B} \tag{3.19}$$

onde $\mathbf{Y}$ representa o canal de luminância da imagem. Como consequência dessa combinação linear, o ruído residual (2.5) torna-se

$$\mathbf{W} = \mathbf{Y} - F(\mathbf{Y}) = (\alpha_1\mathbf{R} + \alpha_2\mathbf{G} + \alpha_3\mathbf{B}) \circ \mathbf{K} + \boldsymbol{\Xi} \tag{3.20}$$

onde $\boldsymbol{\Xi}$ representa um ruído gaussiano branco com variância $\sigma^2_{\boldsymbol{\Xi}}$ dada por

$$\sigma^2_{\boldsymbol{\Xi}} = \sigma^2 \left(\alpha_1^2 + \alpha_2^2 + \alpha_3^2\right). \tag{3.21}$$

Desse modo, a estimativa da PRNU $\hat{\mathbf{K}}_\mathrm{l}$ obtida do canal de luminância das imagens é dada por (2.8) e

$$\mathrm{var}(\hat{\mathbf{K}}_\mathrm{l}) = \sigma^2 \left(\alpha_1 + \alpha_2 + \alpha_3\right) \mathbf{1} \oslash \sum_{k=1}^{d} \left(\alpha_1\mathbf{R}_k + \alpha_2\mathbf{G}_k + \alpha_3\mathbf{B}_k\right)^{\circ 2}. \tag{3.22}$$

Novamente, para provar que $\mathrm{var}(\hat{\mathbf{K}}_\mathrm{l}) \succeq \mathrm{var}(\hat{\mathbf{K}}_\mathrm{p})$ será feita a análise apenas para o elemento na posição (i, j). Fazendo $\hat{K}_\mathrm{p} = \hat{\mathbf{K}}_\mathrm{p}(i, j)$ e $\hat{K}_\mathrm{l} = \hat{\mathbf{K}}_\mathrm{l}(i, j)$, necessita-se mostrar que

$$\mathrm{var}(\hat{K}_\mathrm{l}) - \mathrm{var}(\hat{K}_\mathrm{p}) \geq 0. \tag{3.23}$$

Usando (3.8) e (3.22) em (3.23), e fazendo $R = \mathbf{R}(i, j)$, $G = \mathbf{G}(i, j)$ e $B = \mathbf{B}(i, j)$, tem-se

$$\frac{\sigma^2 \left(\alpha_1 + \alpha_2 + \alpha_3\right)}{\sum\limits_{k=1}^{d} \left(\alpha_1 R_k + \alpha_2 G_k + \alpha_3 B_k\right)^2} - \frac{\sigma^2}{\sum\limits_{k=1}^{d} \left(R_k^2 + G_k^2 + B_k^2\right)} \geq 0 \tag{3.24}$$

e portanto

$$\left(\alpha_1^2+\alpha_2^2+\alpha_3^2\right)\sum_{k=1}^{d}\left(R_k^2+G_k^2+B_k^2\right)-\sum_{k=1}^{d}\left(\alpha_1 R_k+\alpha_2 G_k+\alpha_3 B_k\right)^2\geq 0 \tag{3.25}$$

o que resulta em

$$\sum_{k=1}^{d}\left[\left(\alpha_1^2+\alpha_2^2+\alpha_3^2\right)\left(R_k^2+G_k^2+B_k^2\right)-\left(\alpha_1 R_k+\alpha_2 G_k+\alpha_3 B_k\right)^2\right]\geq 0. \tag{3.26}$$

Nota-se que (3.26) necessita ser provada apenas para $d=1$, pois as parcelas do somatório para um determinado valor de k precisam ser positivas para que (3.26) seja verdadeira para qualquer $d>1$. Assim, expandindo (3.26) para $d=1$ e utilizando novamente (3.16), obtém-se

$$\left(R^2+G^2+B^2\right)\left(\alpha_1^2+\alpha_2^2+\alpha_3^2\right)-\left(\alpha_1 R+\alpha_2 G+\alpha_3 B\right)^2\geq 0 \tag{3.27}$$

o que resulta em

$$\begin{aligned}\left(\alpha_1 R+\alpha_2 G+\alpha_3 B\right)^2+\left(\alpha_2 R-\alpha_1 G\right)^2+\left(\alpha_3 R-\alpha_1 B\right)^2\\ +\left(\alpha_3 G-\alpha_2 B\right)^2-\left(\alpha_1 R+\alpha_2 G+\alpha_3 B\right)^2\geq 0\end{aligned} \tag{3.28}$$

e finalmente

$$\left(\alpha_2 R-\alpha_1 G\right)^2+\left(\alpha_3 R-\alpha_1 B\right)^2+\left(\alpha_3 G-\alpha_2 B\right)^2\geq 0. \tag{3.29}$$

Portanto, dado que todas as parcelas de (3.29) são positivas, conclui-se que $\mathrm{var}(\hat{K}_\mathrm{l})\geq\mathrm{var}(\hat{K}_\mathrm{p})$, então $\mathrm{var}(\hat{\mathbf{K}}_\mathrm{l})\succeq\mathrm{var}(\hat{\mathbf{K}}_\mathrm{p})$.

3.3 Abordagem PCAI8 ME proposta

Este trabalho também propõe uma versão do estimador PCAI8 ME [14] descrito na Seção 2.2.4. Aqui, o ruído residual (3.4) da subimagem $\mathbf{Y}_{s,j,k}$ é

obtido com o filtro PCAI8 e a estimativa $\hat{\mathbf{K}}_s$ é dada por

$$\hat{\mathbf{K}}_s = \frac{1}{cd} \sum_{j=1}^{c} \sum_{k=1}^{d} \mathbf{W}_{s,j,k}, \quad s = 1, 2, 3, 4. \tag{3.30}$$

Em seguida, são aplicadas as operações de pós-processamento (discutidas na Seção 2.2.5) em cada subestimativa $\hat{\mathbf{K}}_s$ e, por fim, as sub-PRNUs são reagrupadas usando (3.7). Nota-se que a diferença entre (2.12) e (3.30) está no fato de esta última estimar a PRNU a partir de subimagens e utilizar todos os canais da imagem para obter uma subestimativa da PRNU.

3.4 Estimação da PRNU da imagem investigada

Neste trabalho, a estimativa da PRNU da imagem investigada é obtida por

$$\hat{\mathbf{K}}_s = \frac{1}{c} \sum_{j=1}^{c} \mathbf{W}_{s,j}, \quad s = 1, 2, 3, 4 \tag{3.31}$$

onde $\hat{\mathbf{K}}_s$ representa a subestimativa da PRNU da imagem investigada e $\mathbf{W}_{s,j}$ é o ruído residual (3.4). As operações de pós-processamento descritas na Seção 2.2.5 também são aplicadas em $\hat{\mathbf{K}}_s$. Então, as sub-PRNUs são agrupadas por (3.7) para formar a impressão digital $\hat{\mathbf{K}}_{\mathrm{t}}$ da imagem investigada. Resultados de experimentos mostraram que (3.31) apresenta uma menor taxa de erro do que (3.5) com $d = 1$, por isso que (3.31) é utilizada neste trabalho.

Como discutido na Seção 2.3, após a etapa de estimação das PRNUs da câmera suspeita e da imagem investigada, o próximo passo do processo de identificação de dispositivo é a detecção, no qual este trabalho utiliza a $\mathrm{CCN}(\mathbf{A}, \mathbf{B})$ (2.17) como métrica de similaridade estatística. Aqui, as variáveis $\mathbf{A}$ e $\mathbf{B}$ são definidas como $\mathbf{A} = \hat{\mathbf{K}}$ e $\mathbf{B} = \hat{\mathbf{K}}_{\mathrm{t}}$.

3.5 Conclusões

Neste capítulo, foi mostrado que o processo de interpolação, utilizado pela maioria dos equipamentos de aquisição de imagem digital, correlaciona as amostras do ruído presente em uma imagem, prejudicando o processo de estimação da PRNU. Para atenuar o efeito da interpolação, foi proposta uma abordagem de estimação da PRNU que consiste em decompor cada canal da imagem em subimagens e aplicar em cada subimagem, um filtro de extração de ruído. Além disso, nessa abordagem não é feita distinção entre os canais de cor de uma imagem. Utilizando tal abordagem, foram propostos dois estimadores da PRNU: um baseado no MLE [9] e outro no PCAI8 ME [14]. Também foi mostrado que o MLE da PRNU aqui proposto apresenta menor variância do que as abordagens de combinação das estimativas da PRNU de cada canal de cor e daquelas que combinam os canais de cor antes do processo de estimação da PRNU.

Capítulo 4

Experimentos e Resultados

Este capítulo apresenta resultados experimentais obtidos visando comparar o desempenho das abordagens de estimação da PRNU propostas neste trabalho (MLE proposto e PCAI8 ME proposto) com as metodologias descritas na Seção 2.2: MLE [9], CDMLE [11], PME [12] e PCAI8 ME [14]. As técnicas MLE proposto, MLE[1], CDMLE e PME utilizam como filtro de extração de ruído o filtro de Wiener no domínio da transformada *wavelet* [40], enquanto a PCAI8 ME e a PCAI8 ME proposto utilizam o filtro PCAI8 [14]. Em todos os experimentos, são utilizadas imagens coloridas, no padrão RGB, com valores de *pixel* entre 0 e 255. Para todos os estimadores, a variância do ruído é $\sigma^2 = 4$ [38]. Com exceção dos estimadores propostos neste trabalho, os demais estimadores utilizam a primeira estratégia descrita na Seção 2.2.6 para obter a estimativa da PRNU a partir de imagens coloridas.

O primeiro experimento, descrito na Seção 4.1, avalia o impacto da interpolação na estimativa da PRNU, comparando a estimativa da PRNU obtida com imagens no formato RAW com as estimativas obtidas após a interpolação. A Seção 4.2 avalia a taxa de sucesso (identificação correta da câmera que adquiriu a imagem) das técnicas de identificação de dispositivo utilizando imagens adquiridas por câmeras digitais de telefones celulares. A Seção 4.3 considera o mesmo procedimento do experimento da Seção 4.2, porém com imagens do banco de imagens Dresden [47]. Esse banco contém imagens, de ambientes *indoor* e *outdoor*, obtidas por câmeras digitais de várias marcas e modelos, utilizando diferentes configurações como distância focal (efeito *zoom*) e *flash*. De acordo com [25], as operações de correções

[1]Os códigos referentes ao MLE da PRNU, implementados em Matlab, são encontrados em: `http://dde.binghamton.edu/download/camera_fingerprint/`. (Acessado pela última vez em 10 de dezembro de 2015).

das distorções causadas pelo efeito *zoom* influenciam o valor da métrica de similaridade estatística utilizada na etapa de detecção do processo de identificação de dispositivo. Assim, apesar de o banco de imagens Dresden [47] não conter as melhores imagens para se estimar a PRNU, ele serve para avaliar o desempenho das técnicas de identificação de dispositivo baseada em PRNU para imagens adquiridas com diferentes configurações, tais como distância focal e *flash*.

4.1 Análise do ruído de interpolação na estimação da PRNU

Para avaliar o efeito da interpolação na estimação da PRNU, este trabalho calcula o CC, entre a estimativa da PRNU obtida a partir de um conjunto de imagens no formato RAW e as estimativas obtidas após a interpolação desse conjunto de imagens. O CC é dado por

$$\mathrm{CC}(\mathbf{A},\mathbf{B}) = \frac{\mathrm{E}[(\mathbf{A}-\mu_{\mathbf{A}})(\mathbf{B}-\mu_{\mathbf{B}})]}{\sigma_{\mathbf{A}}\sigma_{\mathbf{B}}} \tag{4.1}$$

onde $\mu_{\mathbf{A}}$ e $\mu_{\mathbf{B}}$ representam a média aritmética de $\mathbf{A}$ e $\mathbf{B}$, respectivamente, e $\sigma_{\mathbf{A}}$ e $\sigma_{\mathbf{B}}$ denotam o desvio padrão de $\mathbf{A}$ e $\mathbf{B}$, nessa ordem. Aqui, $\mathbf{A}$ e $\mathbf{B}$ representam, respectivamente, as PRNUs obtidas antes e após a interpolação das imagens no formato RAW.

As estimativas de máxima verossimilhança e da média da PRNU utilizando imagens no formato RAW são obtidas por (3.5) e (3.30), respectivamente, fazendo $c = 1$. A estimativa de máxima verossimilhança da PRNU obtida a partir de imagens RAW é utilizada para calcular o CC com as estimativas da PRNU obtidas após a interpolação pelas técnicas MLE, CDMLE, PME e MLE proposto, pois todas utilizam o filtro de extração de ruído proposto por [40]. Por outro lado, a estimativa da média da PRNU obtida de imagens no formato RAW é utilizada para calcular o CC com as estimativas da PRNU obtidas pelas técnicas PCAI8 ME e PCAI8 ME proposto.

As imagens RAW utilizadas nesta seção são as imagens adquiridas por

duas câmeras Nikon D200 equipadas com um filtro de Bayer no padrão RGB encontradas na base de imagens Dresden [47]. Para cada câmera, são selecionadas 100 imagens de resolução 3872×2592 *pixels* obtidas com a mesma distância focal. Devido ao tempo necessário para processar imagens com essa resolução, os testes são realizados considerando blocos de 1024×1024 *pixels* retirados da parte central das imagens. A partir desse conjunto de imagens no formato RAW recortadas, são selecionadas aleatoriamente d fotos para estimar a PRNU. Em seguida, as d imagens são interpoladas utilizando os seguintes algoritmos de interpolação: bilinear, bicúbica, transição suave de matiz, filtro da mediana (com uma janela de observação de 3×3 *pixels*), interpolação baseada no gradiente, plano de cor adaptativo e número de gradientes variável. Detalhes sobre esses algoritmos podem ser encontradas em [36]. Após a interpolação, cada um dos 7 conjuntos de d imagens interpoladas é utilizado para estimar a PRNU pelas técnicas citadas. Em todas as estimativas da PRNU, incluindo aquelas obtidas de imagens no formato RAW, são aplicadas as técnicas de pós-processamento discutidas na Seção 2.2.5. Por fim, calcula-se o CC entre a PRNU estimada com as imagens RAW e cada uma das estimativas obtidas após a interpolação. Esse processo de selecionar as imagens, estimar as PRNUs e calcular os CCs, é repetido 30 vezes para $d = 5, 10, 15, ..., 60$.

As médias dos CCs calculados entre as estimativas da PRNU obtidas antes e após a interpolação são mostradas nas Figuras 4.1 a 4.4. A partir desses resultados, nota-se que as interpolações bilinear e bicúbica, que são consideradas interpolações não adaptativas por possuírem coeficientes de interpolação fixos [42], influenciam mais no processo de estimação da PRNU do que as demais técnicas de interpolação avaliadas, as quais são consideradas adaptativas e apresentam um menor erro de interpolação [42]. Ainda nas Figuras 4.1 a 4.4, observa-se que as abordagens propostas neste trabalho apresentam maiores médias de CCs, o que leva a concluir que tais abordagens são mais robustas às operações de interpolações do que as demais técnicas avaliadas. Dessa forma, espera-se obter uma menor taxa de erro no sistema de

identificação de dispositivo de aquisição utilizando a metologia de estimação da PRNU aqui proposta.

4.2 Imagens adquiridas por câmeras digitais de telefones celulares

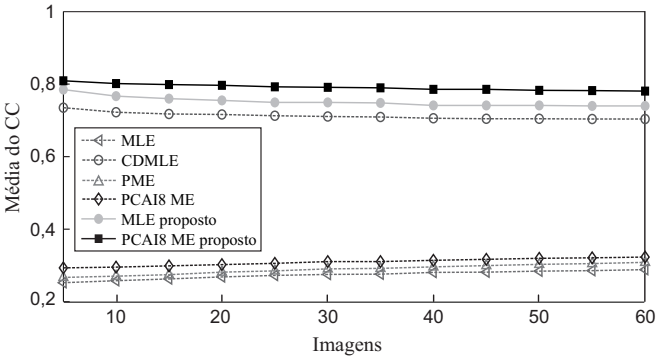
A Tabela 4.1 descreve as 12 câmeras digitais de telefones celulares utilizadas nos experimentos. Com cada uma dessas câmeras, são adquiridas 30 imagens de céu nublado (as quais resultam em melhores estimativas da PRNU [38]) para estimar a PRNU e 100 imagens de cenas aleatórias, em ambientes *indoor* e *outdoor*, para testes. Para cada um dos 11 modelos de celulares, são obtidas via *web*[2] 100 imagens do mesmo modelo e de mesma resolução. No total, 2300 imagens de teste estão disponíveis. Assim, nos experimentos com uma dada câmera, são utilizadas 100 imagens adquiridas por ela (classe positiva) e 2200 não adquiridas (classe negativa). Na estimação da impressão digital do dispositivo, são utilizados blocos de dimensão 1024×1024 *pixels* extraídos da parte central da imagem. Das imagens de testes, são obtidos blocos de dimensão 128×128, 256×256, 512×512 e 1024×1024 *pixels*, também da parte central.

Após o cálculo da CCN (2.17), são analisadas as curvas ROCs (*receiver operating characteristic*) das técnicas de identificação avaliadas. A curva ROC apresenta a taxa de verdadeiro positivo (TVP) *versus* a taxa de falso positivo (TFP) em função da variação do limiar de decisão τ. Caso a imagem seja adquirida pela câmera suspeita e o valor de $CCN(\mathbf{A}, \mathbf{B}) > \tau$, diz-se que ocorre um verdadeiro positivo. Caso a imagem não seja adquirida pela câmera suspeita e tem $CCN(\mathbf{A}, \mathbf{B}) > \tau$, então tem-se um falso positivo. Por fim, divide-se a quantidade de verdadeiros positivos e falsos positivos pelo número de imagens de cada classe, 1200 e 26400, respectivamente, para obter a TVP e TFP.

As Figuras 4.5 e 4.6 apresentam as curvas ROCs obtidas com imagens

[2] As imagens obtidas via *web* são encontradas na página do Flickr (`https://www.flickr.com`).

(a) Bilinear.



(b) Bicúbica.

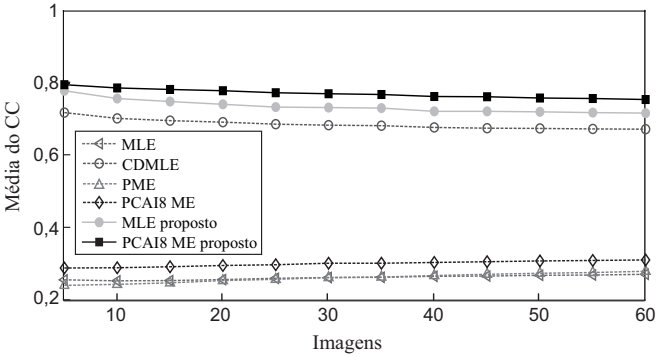
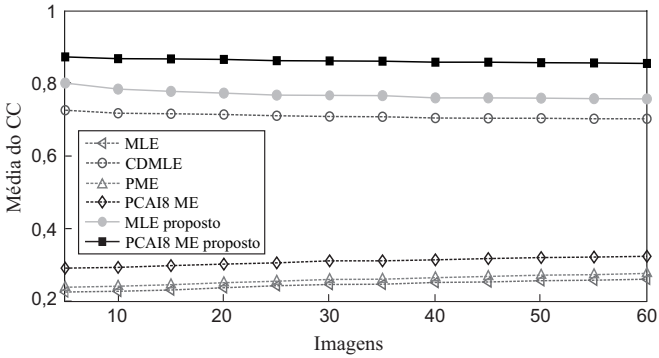


Figura 4.1: Média de 30 CCs calculado entre estimativas da PRNU obtidas de um conjunto de imagens de 1024×1024 *pixels* no formato RAW e após a interpolação desse conjunto de imagens utilizando diferentes técnicas de estimação da PRNU. (a) Bilinear. (b) Bicúbica.

(a) Transição suave de matiz.



(b) Filtro da mediana 3×3.

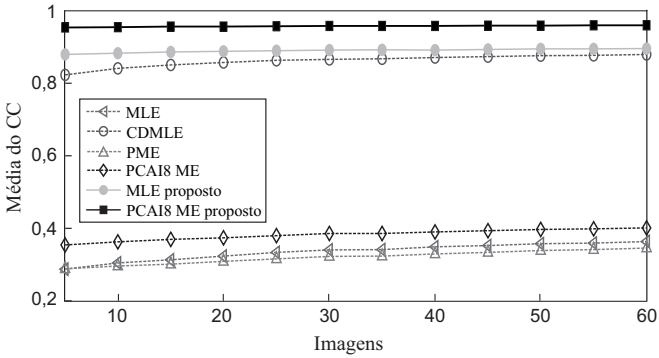
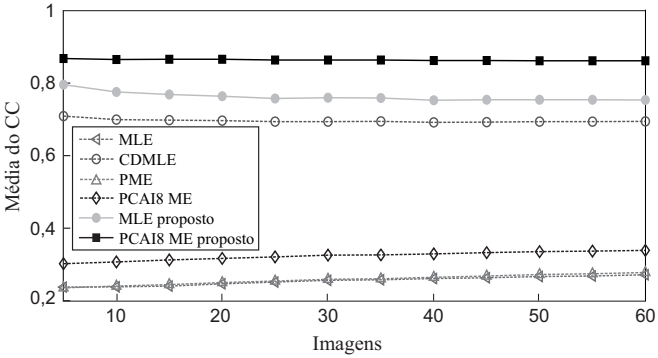


Figura 4.2: Média de 30 CCs calculado entre estimativas da PRNU obtidas de um conjunto de imagens de 1024×1024 *pixels* no formato RAW e após a interpolação desse conjunto de imagens utilizando diferentes técnicas de estimação da PRNU. (a) Transição suave de matiz. (b) Filtro da mediana 3×3.

(a) Gradiente.



(b) Plano de cor adaptativo.

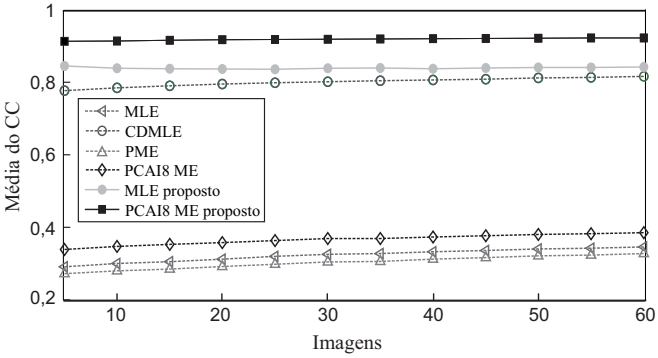


Figura 4.3: Média de 30 CCs calculado entre estimativas da PRNU obtidas de um conjunto de imagens de 1024×1024 *pixels* no formato RAW e após a interpolação desse conjunto de imagens utilizando diferentes técnicas de estimação da PRNU. (a) Gradiente. (b) Plano de cor adaptativo.

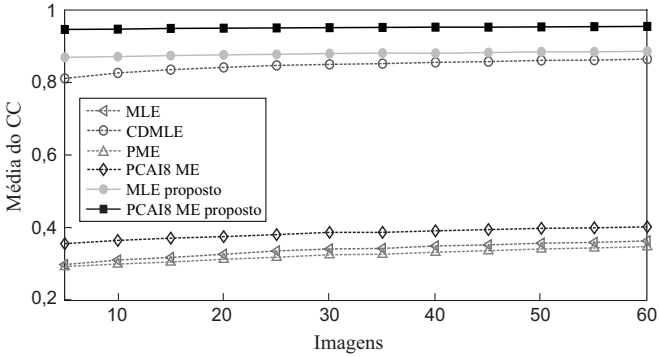


Figura 4.4: Média de 30 CCs calculado entre estimativas da PRNU obtidas de um conjunto de imagens de 1024×1024 *pixels* no formato RAW e após a interpolação, com o algoritmo número de gradientes variável, desse conjunto de imagens utilizando diferentes técnicas de estimação da PRNU.

Tabela 4.1: Detalhes das 12 câmeras digitais de telefones celulares utilizadas nos experimentos

Modelo	Número de celulares	Resolução em *pixels*
Apple iPhone 4	1	2592×1936
Apple iPhone 4S	1	3264×2448
LG E400	1	2048×1536
Nokia C2 06	1	1600×1200
Samsung GT I9070	1	2560×1920
Samsung GT I9300	2	3264×2448
Samsung GT S5301B	1	1600×1200
Samsung GT S5367	1	2048×1536
Samsung GT S6102B	1	2048×1536
Sony C1604	1	2048×1536
Sony ST25a	1	2592×1944

Tabela 4.2: Taxa de verdadeiro positivo para taxa de falso positivo igual a 10^{-3} de vários métodos de identificação de dispositivo de aquisição utilizando as operações de pós-processamento da PRNU e imagens adquiridas por câmeras digitais de telefones celulares

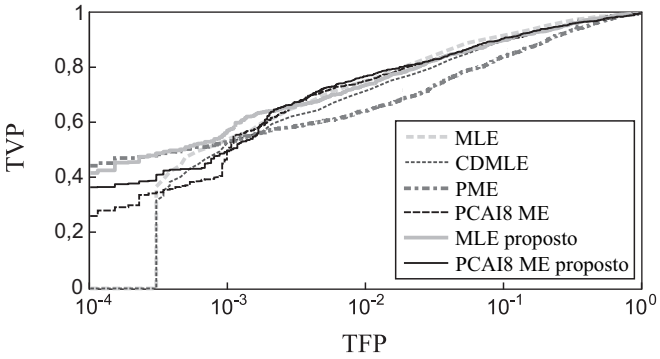
Método	Tamanho da imagem em *pixels*			
	128×128	256×256	512×512	1024×1024
MLE	0.5392	0.8092	0.9833	1.0000
CDMLE	0.5525	0.8275	0.9867	1.0000
PME	0.4800	0.7308	0.9758	1.0000
PCAI8 ME	0.5258	0.8100	0.9925	1.0000
MLE proposto	0.6283	0.8817	0.9983	1.0000
PCAI8 ME proposto	0.6383	0.9158	0.9992	1.0000

de celular, sem utilizar as operações de pós-processamento da PRNU descritas na Seção 2.2.5, e as Figuras 4.7 e 4.8 apresentam os resultados obtidos utilizando as operações de pós-processamento da PRNU. As abordagens propostas (MLE proposto e PCAI8 ME proposto) apresentam as maiores TVPs utilizando as operações de pós-processamento da PRNU. Para imagens de teste de 1024×1024 *pixels*, todas as abordagens avaliadas, com as operações de pós-processamento da PRNU, apresentam TVP $= 1$ para TFP $= 10^{-3}$. A Tabela 4.2 apresenta a TVP para a TFP $= 10^{-3}$ das curvas ROCs das Figuras 4.7 e 4.8. A partir desses resultados nota-se que o método aqui proposto possui melhor desempenho do que as outras técnicas avaliadas para imagens de baixa resolução adquiridas por câmeras digitais de telefones celulares.

4.3 Imagens adquiridas por câmeras digitais

Este experimento utiliza imagens, da base de imagens Dresden [47], adquiridas por 31 câmeras digitais. Detalhes dessas câmeras e das imagens adquiridas por elas são mostrados na Tabela 4.3. Para cada câmera, são se-

(a) 128×128.



(b) 256×256.

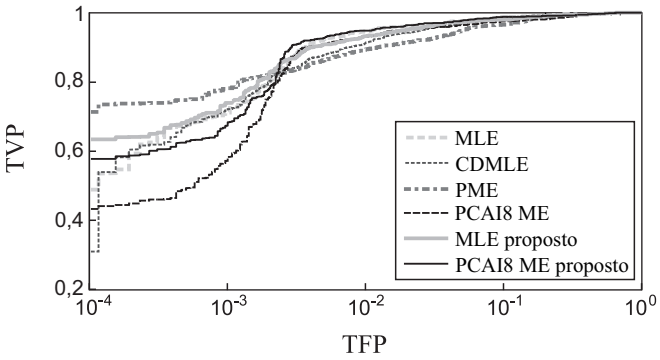
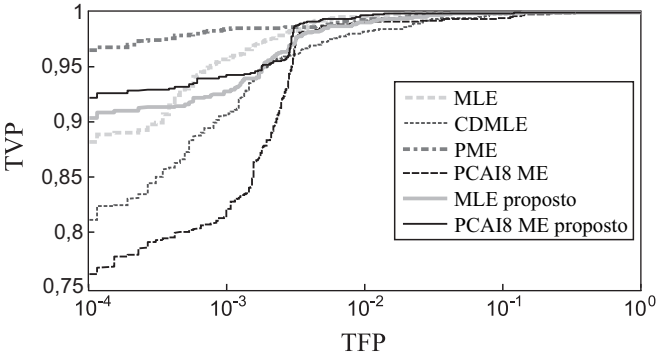


Figura 4.5: Curvas ROCs de diversos métodos de identificação de dispositivo de aquisição sem utilizar as operações de pós-processamento da PRNU e com imagens de teste adquiridas por câmeras digitais de telefones celulares. (a) 128×128 *pixels*. (b) 256×256 *pixels*.

(a) 512×512.



(b) 1024×1024.

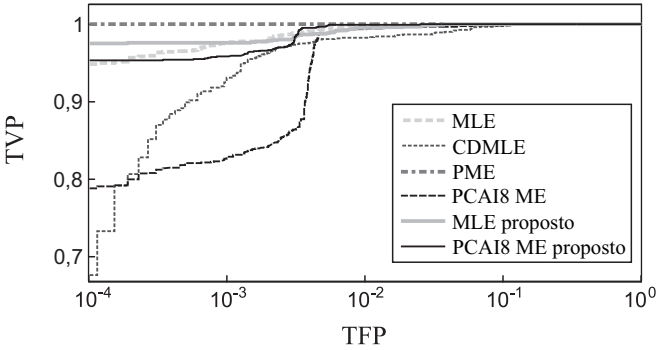
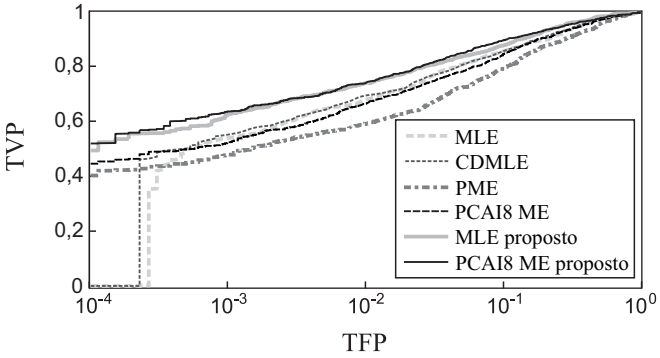


Figura 4.6: Curvas ROCs de diversos métodos de identificação de dispositivo de aquisição sem utilizar as operações de pós-processamento da PRNU e com imagens de teste adquiridas por câmeras digitais de telefones celulares. (a) 512×512 *pixels*. (b) 1024×1024 *pixels*.

(a) 128×128.



(b) 256×256.

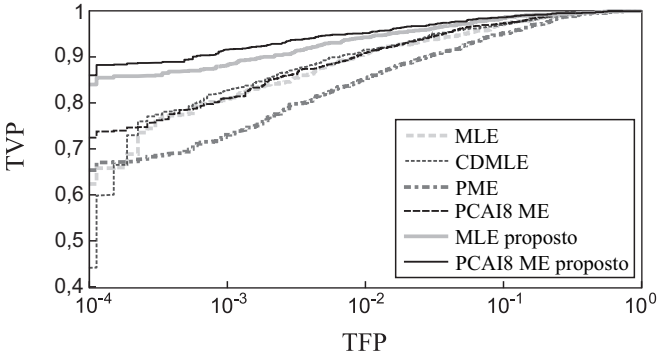
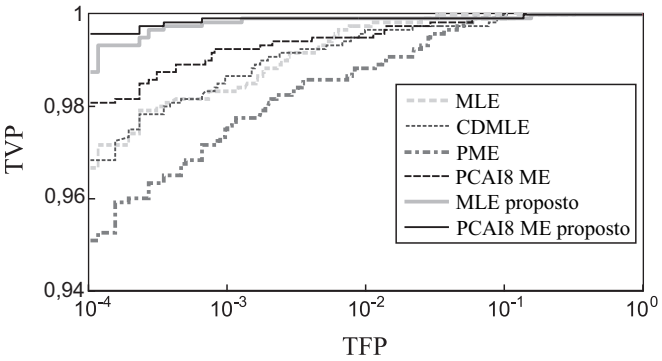


Figura 4.7: Curvas ROCs de diversos métodos de identificação de dispositivo de aquisição utilizando as operações de pós-processamento da PRNU e com imagens de teste adquiridas por câmeras digitais de telefones celulares. (a) 128×128 *pixels*. (b) 256×256 *pixels*.

(a) 512×512.



(b) 1024×1024.

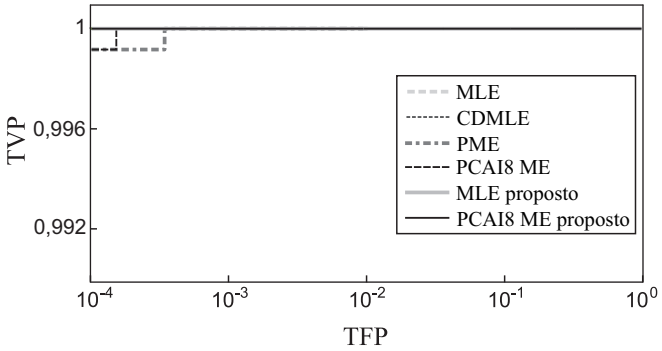


Figura 4.8: Curvas ROCs de diversos métodos de identificação de dispositivo de aquisição utilizando as operações de pós-processamento da PRNU e com imagens de teste adquiridas por câmeras digitais de telefones celulares. (a) 512×512 *pixels*. (b) 1024×1024 *pixels*.

lecionadas 100 imagens para estimar a PRNU e as restantes são utilizadas como imagens de teste. Nesse experimento, são utilizadas 100 imagens, em vez de 30 como na Seção 4.2, pois essas 100 imagens são de cenas aleatórias e foram adquiridas com diferentes configurações (como *flash* e efeito *zoom*). Com exceção da quantidade de imagens consideradas, os demais procedimentos de estimação da PRNU e teste das imagens são os mesmos utilizados no experimento da Seção 4.2.

Os resultados obtidos sem utilizar as operações de pós-processamento da PRNU são apresentados nas Figuras 4.9 e 4.10, e os resultados obtidos aplicando as técnicas de pós-processamento são apresentados nas Figuras 4.11 e 4.12. Assim como no experimento com imagens adquiridas por câmeras digitais de telefones celulares, as abordagens MLE proposto e PCAI8 ME proposto apresentam os melhores resultados, em termos de curvas ROCs, utilizando as operações de pós-processamento da PRNU. A Tabela 4.4 apresenta as TVPs para TFP $= 10^{-3}$. Comparando os resultados das Tabelas 4.2 e 4.4, percebe-se que as TVPs desta última são menores. Apesar de os resultados serem de experimentos realizados com imagens adquiridas por tipos de dispositivos diferentes (câmeras digitais de telefones celulares e câmeras digitais), a diminuição das TVPs é causada, principalmente, pela dessincronização das impressões digitais e não necessariamente pelo tipo de dispositivo. Como discutido na Seção 1.1, essa dessincronização das impressões digitais é resultante das operações de correções das distorções causadas nas imagens adquiridas com efeito *zoom* [25]. Portanto, mesmo para imagens adquiridas com efeito *zoom*, a abordagem de estimação da PRNU proposta neste trabalho apresenta melhores resultados.

4.4 Conclusões

Neste capítulo, o desempenho da abordagem de estimação da PRNU aqui proposta foi comparada com as técnicas MLE [9], CDMLE [11], PME [12] e PCAI8 ME [14], através de experimentos realizados com imagens adquiri-

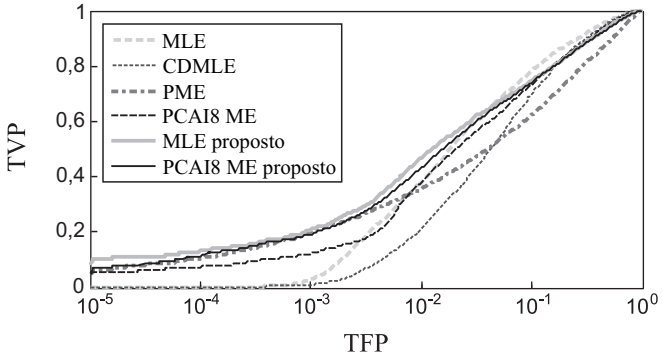
Tabela 4.3: Detalhes das 31 câmeras digitais da base de imagens Dresden [47] usadas nos experimentos

Modelo	Número de câmeras	Número de imagens por câmera	Resolução em *pixels*
AgfaPhoto DC-733s	1	250	3072×2304
AgfaPhoto DC-830i	1	250	3264×2448
AgfaPhoto Sensor 530s	1	250	4032×3024
Canon Ixus 55	1	224	2592×1944
Fujifilm FinePix J50	3	205, 210, 215	3264×2448
Kodak M1063	5	212, 219, 221, 250, 250	3664×2748
Nikon D200	1	250	3872×2592
Olympus u1050SW	5	202, 204, 207, 209, 218	3648×2736
Panasonic DMC-FZ50	3	209, 215, 226	3648×2736
Praktica DCZ 5.9	4	200, 205, 206, 209	2560×1920
Samsung L74wide	3	224, 231, 231	3072×2304
Samsung NV15	3	210, 214, 217	3648×2736

Tabela 4.4: Taxa de verdadeiro positivo para taxa de falso positivo igual a 10^{-3} de vários métodos de identificação de dispositivo de aquisição utilizando as operações de pós-processamento da PRNU e imagens, da base de imagens Dresden [47], tiradas por câmeras digitais

Método	Tamanho da imagem em *pixels*			
	128×128	256×256	512×512	1024×1024
MLE	0.2380	0.6230	0.8608	0.9442
CDMLE	0.2415	0.6059	0.8565	0.9394
PME	0.1563	0.5196	0.8277	0.9356
PCAI8 ME	0.2658	0.6297	0.8587	0.9420
MLE proposto	0.2688	0.6572	0.8990	0.9618
PCAI8 ME proposto	0.2752	0.6554	0.9052	0.9634

(a) 128×128.



(b) 256×256.

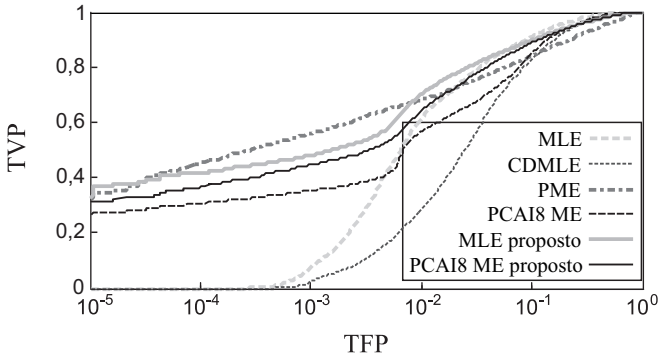
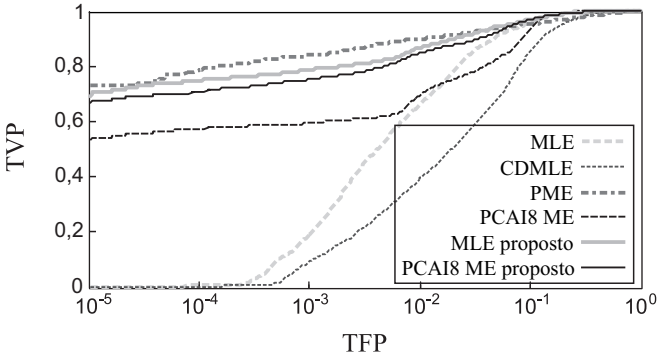


Figura 4.9: Curvas ROCs de diversos métodos de identificação de dispositivo de aquisição sem utilizar as operações de pós-processamento da PRNU e com imagens de teste obtidas da base de imagens Dresden [47], adquiridas por câmeras digitais. (a) 128×128 *pixels*. (b) 256×256 *pixels*.

(a) 512×512.



(b) 1024×1024.

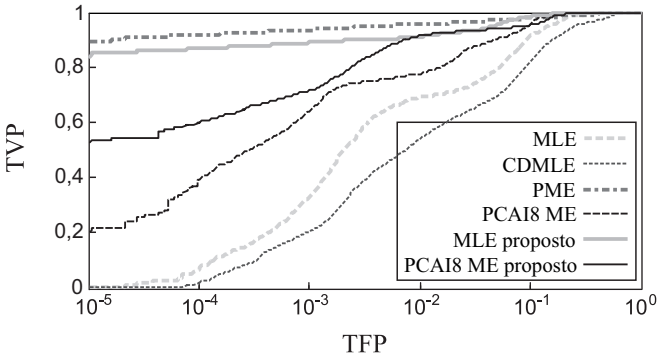
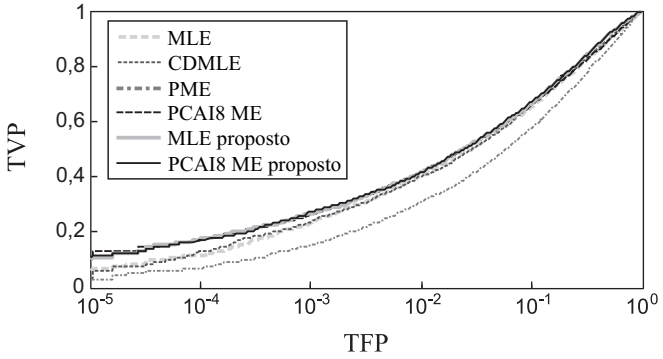


Figura 4.10: Curvas ROCs de diversos métodos de identificação de dispositivo de aquisição sem utilizar as operações de pós-processamento da PRNU e com imagens de teste obtidas da base de imagens Dresden [47], adquiridas por câmeras digitais. (a) 512×512 *pixels*. (b) 1024×1024 *pixels*.

(a) 128×128.



(b) 256×256.

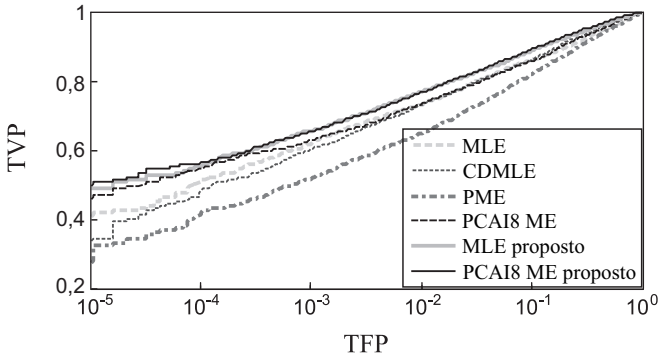
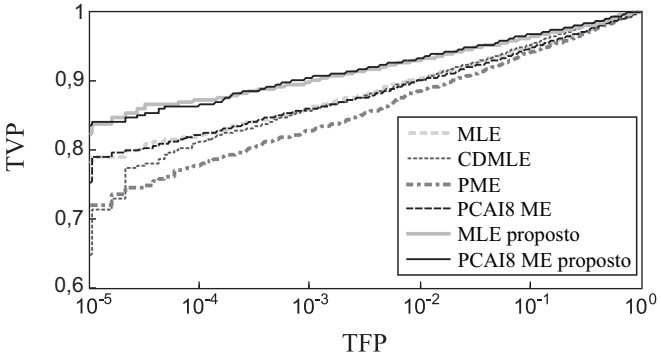


Figura 4.11: Curvas ROCs de diversos métodos de identificação de dispositivo de aquisição utilizando as operações de pós-processamento da PRNU e com imagens de teste obtidas da base de imagens Dresden [47], adquiridas por câmeras digitais. (a) 128×128 *pixels*. (b) 256×256 *pixels*.

(a) 512×512.



(b) 1024×1024.

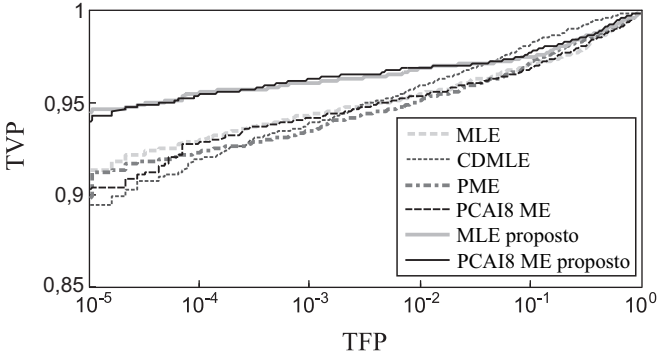


Figura 4.12: Curvas ROCs de diversos métodos de identificação de dispositivo de aquisição utilizando as operações de pós-processamento da PRNU e com imagens de teste obtidas da base de imagens Dresden [47], adquiridas por câmeras digitais. (a) 512×512 *pixels*. (b) 1024×1024 *pixels*.

das por câmeras digitais de telefones celulares e câmeras digitais. Nesses experimentos foram avaliados o efeito da interpolação no processo de estimação da PRNU e as curvas ROCs das técnicas citadas anteriormente em aplicações de identificação de dispositivo de aquisição com imagens de baixa resolução. Considerando o cenário no qual são utilizadas as técnicas de pós-processamento da PRNU, a abordagem aqui proposta obteve resultados superiores aos das técnicas avaliadas. Nos cenários sem o uso das técnicas de pós-processamento da PRNU, todas as abordagens avaliadas apresentam baixas TVPs em comparação com cenários que utilizam as técnicas de pós-processamento. Vale ressaltar que as metodologias de estimação da PRNU MLE proposto e PCAI8 ME proposto, apresentam resultados semelhantes em termos de curvas ROCs. No entanto, essas duas técnicas utilizam estimadores distintos (MLE e ME) como também filtros de extração de ruído diferentes (filtro de Wiener no domínio da transformada *wavelet* [40] e PCAI8 [14]).

Capítulo 5

Conclusões e Considerações Finais

Este trabalho apresentou uma abordagem de estimação da PRNU (impressão digital) para dispositivos equipados com um CFA. Conforme discutido no Capítulo 1, a PRNU é um padrão de ruído causado pelas imperfeições do processo de manufatura do fotossensor que pode ser usado para identificar cada fotossensor. Por essa razão, a PRNU é utilizada em análise forense para realizar perícias de identificação de dispositivo de aquisição e identificação de fraudes em imagens e vídeos digitais. Neste capítulo, são discutidos os principais resultados obtidos neste trabalho e algumas sugestões de trabalhos futuros.

5.1 Discussão dos resultados

Neste trabalho, foi mostrado que o processo de interpolação de imagem, utilizado pelos dispositivos de aquisição de imagem (como câmera digital e câmera digital de telefone celular, por exemplo) que implementam um CFA, correlaciona as amostras do ruído presente na imagem e prejudica o processo de estimação da PRNU. Para amenizar esse problema, a abordagem de estimação da PRNU aqui proposta consiste em separar a imagem em subimagens e utilizar as subimagens no processo de estimação da PRNU. Além disso, todos os canais de cor das imagens são utilizados, sem distinção, para obter uma subestimativa da PRNU. Em seguida, essas subestimativas da PRNU são agrupadas para formar a estimativa da impressão digital do dispositivo de aquisição. Baseado nesse procedimento, foram apresentadas duas estratégias de estimação da PRNU: uma baseada no MLE [9] e outra no PCAI8 ME [14].

A análise da variância do MLE da PRNU, proposto no Capítulo 3 deste trabalho, mostrou que ela é sempre menor ou igual à variância dos estimadores da PRNU que envolvem a combinação linear dos canais de cor da imagem antes do processo de estimação e das técnicas que aplicam a combinação linear às estimativas da PRNU de cada canal de cor. Em termos práticos, isso significa que a abordagem aqui proposta obtém uma estimativa da PRNU mais próxima da original, conforme mostrado no Capítulo 4 a partir dos resultados dos experimentos realizados.

Com o objetivo de comparar o desempenho da abordagem aqui proposta com aquelas apresentadas em [9], [11], [12] e [14], foram realizados três experimentos. O primeiro experimento avaliou a influência do processo de interpolação na estimativa da PRNU. Para tal, foram usadas imagens no formato RAW (as quais contém os *pixels* antes do processo de interpolação). A estimativa da PRNU obtida a partir das imagens no formato RAW foi comparada com as estimativas obtidas após a interpolação. No segundo experimento, foram comparadas as TVPs e TFPs (curvas ROCs) das técnicas de identificação de dispositivo com imagens de baixa resolução adquiridas por câmeras digitais de telefones celulares, sendo alguns telefones celulares do mesmo modelo. Também foi avaliado o efeito das técnicas de pós-processamento da PRNU, as quais são usadas para suprimir artefatos presentes nas estimativas da PRNU. No terceiro e último experimento, foram utilizadas imagens da base Dresden [47] adquiridas por câmeras digitais. A base Dresden [47] contém milhares de imagens adquiridas por várias câmeras digitais diferentes, com algumas sendo do mesmo modelo, utilizando diferentes configurações, tais como efeito *zoom* e *flash*. Os resultados dos três experimentos realizados mostraram que a abordagem aqui proposta apresenta melhor desempenho quando comparada com aquelas apresentadas em [9], [11], [12] e [14].

5.2 Trabalhos futuros

Os seguintes tópicos são sugestões de trabalhos futuros:

- Analisar o desempenho da técnica aqui proposta com imagens de teste que tenham sofrido algum tipo de transformação geométrica, tais como rotação e redimensionamento. Imagens que passaram por esse tipo de transformação tendem a apresentar um baixo valor da métrica de similaridade estatística utilizada no processo de identificação de dispositivo de aquisição.

- Aplicar a abordagem proposta neste trabalho em problemas de identificação de fraudes em imagens baseada em PRNU. Dado que a abordagem de estimação da PRNU proposta neste trabalho apresenta melhores resultados do que as técnicas ora apresentadas na literatura, espera-se que esta abordagem apresente uma maior taxa de acerto de identificação das regiões de uma imagem que sofreram algum tipo de fraude, como, por exemplo, a fraude de copiar e colar (que consiste em apagar parte da imagem utilizando cópias de outras partes da mesma imagem).

- Avaliar o desempenho da abordagem aqui proposta para realizar identificação de dispositivo de aquisição de vídeos digitais. Geralmente, o processo de aquisição de vídeo degrada mais o sinal capturado pelo fotossensor do que o processo de aquisição de imagem.

Referências Bibliográficas

[1] J. A. Redi, W. Taktak, and J.-L. Dugelay, "Digital image forensics: a booklet for beginners," *Multimedia Tools Appl.*, vol. 51, no. 1, pp. 133–162, Jan. 2011.

[2] K. S. Choi, E. Y. Lam, and K. K. Y. Wong, "Automatic source camera identification using the intrinsic lens radial distortion," *Opt. Express*, vol. 14, no. 24, pp. 11 551–11 565, Nov 2006.

[3] S. Bayram, H. T. Sencar, and N. Memon, "Source camera identification based on CFA interpolation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, vol. 3, Genoa, Italy, Sept. 2005, pp. 69–72.

[4] A. E. Dirik, H. T. Sencar, and N. D. Memon, "Digital single lens reflex camera identification from traces of sensor dust," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 3, pp. 539–552, Sept. 2008.

[5] Z. J. Geradts, J. Bijhold, M. Kieft, K. Kurosawa, K. Kuroki, and N. Saitoh, "Methods for identification of images acquired with digital cameras," in *Proc. SPIE, Enabling Technologies for Law Enforcement and Security*, vol. 4232, Boston, MA, Feb. 2001, pp. 505–512.

[6] K. Kurosawa, K. Kuroki, and N. Saitoh, "CCD fingerprint method-identification of a video camera from videotaped images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, vol. 3, Kobe, Japan, Oct. 1999, pp. 537–540.

[7] J. Lukáš, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise." *IEEE Trans. Inf. Forensics Security*, vol. 1, no. 2, pp. 205–214, Jun. 2006.

[8] ——, "Determining digital image origin using sensor imperfections," in *Proc. SPIE, Image and Video Communications and Processing*, vol. 5685, San Jose, CA, Apr. 2005, pp. 249–260.

[9] M. Chen, J. Fridrich, and M. Goljan, "Digital imaging sensor identification (further study)," in *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents*, vol. 6505, San Jose, CA, Feb. 2007, pp. 0P–0Q.

[10] C.-T. Li, "Source camera identification using enhanced sensor pattern noise," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 280–287, Jun. 2010.

[11] C.-T. Li and Y. Li, "Color-decoupled photo response non-uniformity for digital image forensics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 2, pp. 260–271, Feb. 2012.

[12] X. Kang, Y. Li, Z. Qu, and J. Huang, "Enhancing source camera identification performance with a camera reference phase sensor pattern noise," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 393–402, Apr. 2012.

[13] Z. Qu, X. Kang, J. Huang, and Y. Li, "Forensic sensor pattern noise extraction from large image data set," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, Vancouver, Canada, May 2013, pp. 3023–3027.

[14] X. Kang, J. Chen, K. Lin, and P. Anjie, "A context-adaptive SPN predictor for trustworthy source camera identification," *EURASIP J. Image and Video Processing*, vol. 2014, no. 1, p. 19, Apr. 2014.

[15] R. Li, C.-T. Li, and Y. Guan, "A reference estimator based on composite sensor pattern noise for source device identification," in *Proc. SPIE, Media Watermarking, Security and Forensics*, vol. 9028, San Francisco, CA, Feb. 2014, pp. 0O 1–7.

[16] L.-H. Chan, N.-F. Law, and W.-C. Siu, "A two dimensional camera identification method based on image sensor noise," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 1741–1744.

[17] M. Goljan, M. Chen, and J. Fridrich, "Identifying common source digital camera from image pairs," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, vol. 6, San Antonio, TX, Sep. 2007, pp. VI–125–VI–128.

[18] M. Chen, J. Fridrich, M. Goljan, and J. Lukáš, "Source digital camcorder identification using sensor photo-response nonuniformity," in *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents*, vol. 6505, San Jose, CA, Feb. 2007, pp. 1G–1H.

[19] M. Goljan, "Digital camera identification from images – estimating false acceptance probability," in *Digital Watermarking,* vol. 5450*, Lecture Notes in Computer Science*, H.-J. Kim, S. Katzenbeisser, and A. Ho, Eds., Busan, Korea: Springer, 2009, pp. 454–468.

[20] M. Goljan, J. Fridrich, and T. Filler, "Large scale test of sensor fingerprint camera identification," in *Proc. SPIE, Electronic Imaging, Media Forensics and Security*, vol. 7254, San Jose, CA, Feb. 2009, pp. 17–21.

[21] M.-J. Tsai, C.-S. Wang, and J. Liu, "A hybrid model for digital camera source identification," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Cairo, Egypt, Nov. 2009, pp. 2901–2904.

[22] M. Goljan and J. Fridrich, "Camera identification from cropped and scaled images," in *Proc. SPIE, Security, Forensics, Steganography, and Watermarking of Multimedia Contents*, vol. 6819, San Jose, CA, Feb. 2008, pp. 0E–1–13.

[23] M. Goljan, J. Fridrich, M. Goljan, and J. Lukáš, "Camera identification from printed images," in *Proc. SPIE, Security, Forensics, Steganо-*

graphy, and Watermarking of Multimedia Contents, vol. 6819, San Jose, CA, Feb. 2008, pp. 0I–1–12.

[24] M. Goljan, J. Fridrich, and M. Chen, "Defending against fingerprint-copy attack in sensor-based camera identification," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 227–236, Mar. 2011.

[25] M. Goljan and J. Fridrich, "Sensor-fingerprint based identification of images corrected for lens distortion," in *Proc. SPIE, Electronic Imaging, Media Watermarking, Security and Forensics*, vol. 8303, San Francisco, CA, Jan. 2012, pp. 0H 1–13.

[26] D. Valsesia, G. Coluccia, T. Bianchi, and E. Magli, "Large-scale image retrieval based on compressed camera identification," *IEEE Trans. Multimedia*, vol. 17, no. 9, pp. 1439–1449, Sep. 2015.

[27] ——, "Compressed fingerprint matching and camera identification via random projections," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 7, pp. 1472–1485, Jul. 2015.

[28] M. Goljan, J. Fridrich, and T. Filler, "Managing a large database of camera fingerprints," in *Proc. SPIE, Media Forensics and Security*, vol. 7541, San Jose, CA, Jan. 2010, pp. 08–01–08–12.

[29] S. Bayram, H. Sencar, and N. Memon, "Efficient sensor fingerprint matching through fingerprint binarization," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 4, pp. 1404–1413, Aug. 2012.

[30] D. Valsesia, G. Coluccia, T. Bianchi, and E. Magli, "Scale-robust compressive camera fingerprint matching with random projections," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 1697–1701.

[31] R. Li, C.-T. Li, and Y. Guan, "A compact representation of sensor fingerprint for camera identification and fingerprint matching," in *Proc.*

IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP), Brisbane, Australia, Apr. 2015, pp. 1777–1781.

[32] S. Bayram, H. Sencar, and N. Memon, "Sensor fingerprint identification through composite fingerprints and group testing," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 3, pp. 597–612, Mar. 2015.

[33] A. Pande, S. Chen, P. Mohapatra, and J. Zambreno, "Hardware architecture for video authentication using sensor pattern noise," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 1, pp. 157–167, Jan. 2014.

[34] D.-K. Hyun, S.-J. Ryu, M.-J. Lee, J.-H. Lee, H.-Y. Lee, and H.-K. Lee, "Source camcorder identification from cropped and scaled videos," in *Proc. SPIE, Media Watermarking, Security and Forensics*, vol. 8303, San Francisco, CA, Jan. 2012, pp. 0E 1–8.

[35] J. Fridrich, "Sensor defects in digital image forensic," in *Digital Image Forensics: There is More to a Picture Than Meets the Eye*, H. T. Sencar and N. Memon, Eds., New York: Springer, 2013, pp. 179–218.

[36] A. C. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3948–3959, Oct. 2005.

[37] R. C. Gonzalez and R. E. Woods, "Digital image fundamentals," in *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, Inc., 2007, pp. 57–125.

[38] J. Fridrich, "Digital image forensics," *IEEE Signal Process. Mag.*, vol. 26, no. 2, pp. 16–37, Mar. 2009.

[39] G. E. Healey and R. Kondepudy, "Radiometric CCD camera calibration and noise estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 3, pp. 267–276, Mar. 1994.

[40] M. K. Mihçak, I. Kozintsev, and K. Ramchandran, “Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising,” in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process. (ICASSP)*, vol. 6, Phoenix, AZ, Mar. 1999, pp. 3253–3256.

[41] T. Filler, J. Fridrich, and M. Goljan, “Using sensor pattern noise for camera model identification,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, San Diego, CA, Oct. 2008, pp. 1296 – 1299.

[42] A. Swaminathan, M. Wu, and K. J. R. Liu, “Nonintrusive component forensics of visual sensors using output images,” *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 1, pp. 91–106, Mar. 2007.

[43] F. Corporation. (2008, Sep.) Revolutionary sensor blazes the way to a new era of high image quality. Cologne, Germany. Acessado em 28 de outubro de 2015. [Online]. Disponível em: http://www.fujifilm.com/products/digital_cameras/topics/2008/0922_01.html.

[44] H. B. Costa, E. L. O. Batista, e R. Seara, “Uma estratégia de identificação de dispositivo de aquisição de imagens baseada em PRNU,” em *Anais do Simpósio Brasileiro de Telecomunicações (SBrT)*, Juiz de Fora, MG, Brasil, Set. 2015, pp. 79–83.

[45] J. W. Brewer, “Kronecker products and matrix calculus in system theory,” *IEEE Trans. Circuits Syst.*, vol. 25, no. 9, pp. 772–781, Sep. 1978.

[46] S. M. Kay, “Minimum variance unbiased estimation,” in *Fundamentals of Statistical Signal Processing: Estimation Theory*, Upper Saddle River, NJ: Prentice-Hall, Inc., 1993, pp. 15–26.

[47] T. Gloe and R. Böhme, “The Dresden image database for benchmarking digital image forensics,” *Journal of Digital Forensic Practice*, vol. 3, no. 2-4, pp. 150–159, 2010.