



# Digital Forensics

MO447 / MC919

***Prof. Dr. Anderson Rocha***

*Microsoft Research Faculty Fellow  
Affiliate Member, Brazilian Academy of Sciences  
Reasoning for Complex Data (Recod) Lab.*

[anderson.rocha@ic.unicamp.br](mailto:anderson.rocha@ic.unicamp.br)  
<http://www.ic.unicamp.br/~rocha>

**Reasoning for Complex Data (RECOD) Lab.**

Institute of Computing,  
University of Campinas (Unicamp)

Av. Albert Einstein, 1251 – Cidade Universitária  
CEP 13083-970 • Campinas/SP – Brasil

# Organização

# Organização

- ▶ Atribuição de Fontes (Modelo e Tipo Específico)
  - Câmera
  - *Scanner*
- ▶ Identificação de Criações Sintéticas
- ▶ Técnicas Contra-Forenses

# **Atribuição de Fontes**

# Atribuição de Fontes

- ▶ Câmera
- ▶ *Camcorder*
- ▶ *Scanner*
- ▶ Impressora

# Atribuição de Fontes

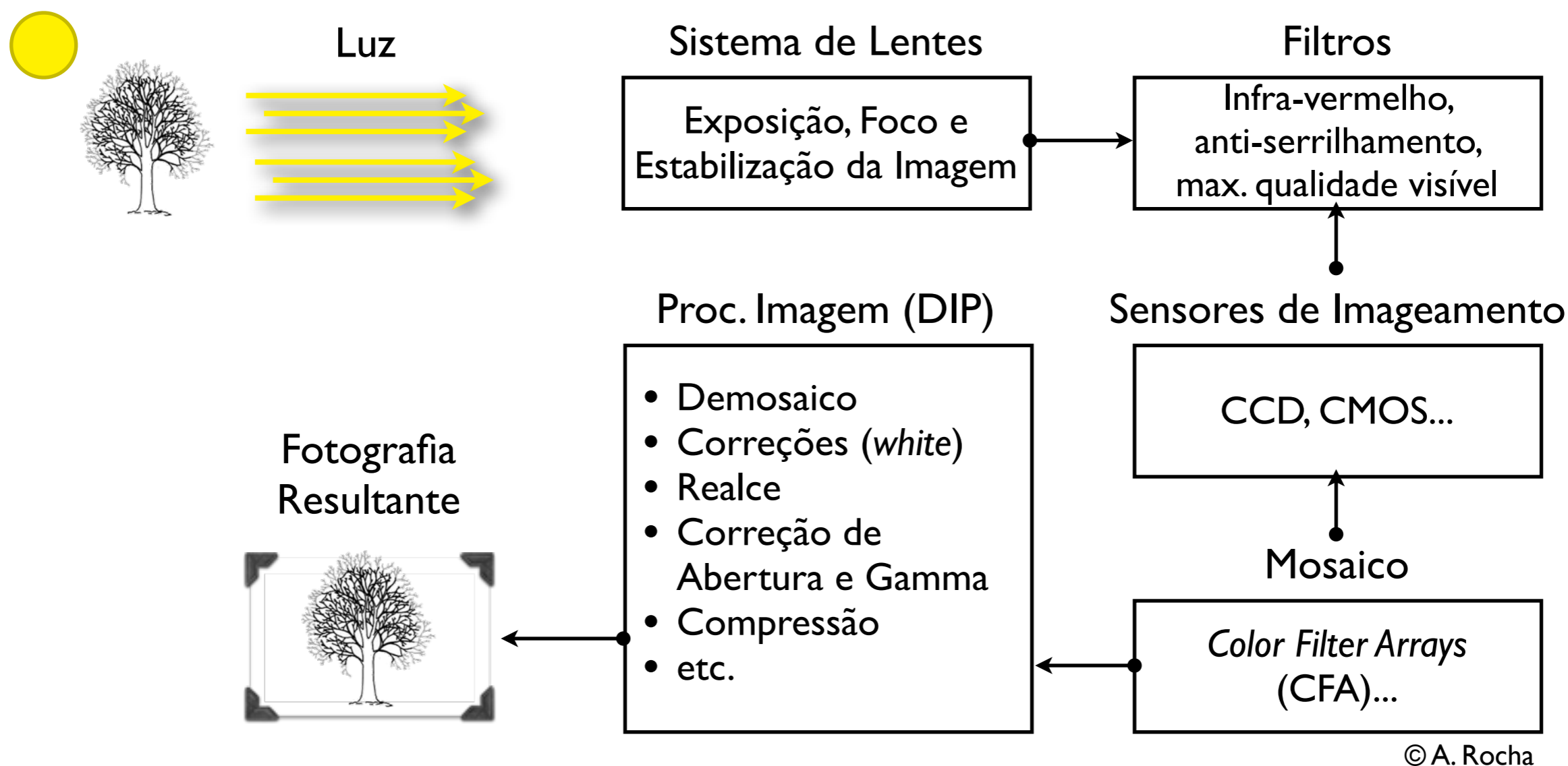
- ▶ Técnicas para apontar/atribuir
  - O modelo do dispositivo de aquisição
  - O dispositivo exato

# **Atribuição do Modelo de Câmera**

# Atribuição do Modelo

- ▶ Informações relacionadas ao processo de aquisição
  - Características das lentes
  - Tipo e tamanho dos sensores de aquisição
  - Tipo de filtro de mosaico/demosaico
  - Informações dos algoritmos da DIP

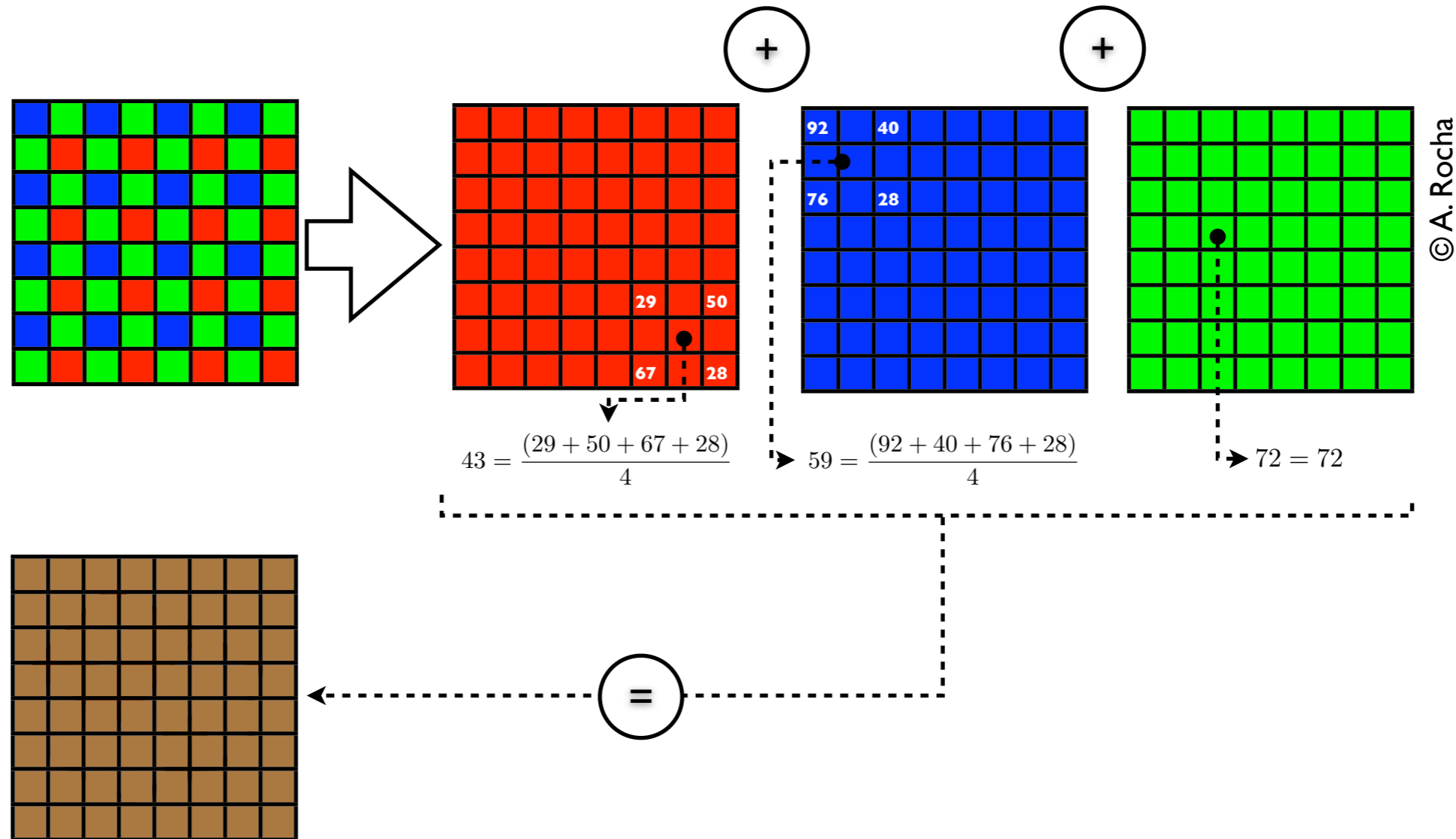
# Pipeline Câmera



# Estado da Arte

- ▶ **Análise de pós-processamento**  
[Kharrazi et al. 2004]
- ▶ **Análise de artefatos de mosaico/demosaico**  
[Popescu & Farid 2004; Bayram et al. 2005b]
- ▶ **Artefatos em tabelas de quantização**  
**JPEG**[Popescu 2004]
- ▶ **Distorções de lentes**  
[Choi et al. 2006]

# Mosaico/Demosaico



# Análise de Artefatos CFA

## ▶ Hipótese:

- linhas e coluna de imagens interpoladas provavelmente possuem correlações com seus vizinhos
- relação de vizinhança pode ser fornecida por *kernels* de tamanhos específicos (e.g.,  $3 \times 3$ ,  $5 \times 5$ , etc.)

# Análise de Artefatos CFA

- ▶ [Popescu 2004] apresentam um algoritmo de maximização da esperança (EM)
- ▶ Dois estágios
  - Esperança
  - Maximização

# Análise de Artefatos CFA

- ▶ **No estágio da esperança (E)**, estima-se a probabilidade de cada amostra pertencer a um modelo em particular
- ▶ **No estágio da maximização (M)**, estima-se a forma específica das correlações entre as amostras

# Análise de Artefatos CFA

- ▶ Mais especificamente, podemos assumir que cada amostra pertence a um de dois modelos possíveis
- ▶ Se a amostra é linearmente correlacionada com seus vizinhos ela pertence ao Modelo 1. Caso contrário, ao Modelo 2

# Análise de Artefatos CFA

- ▶ Função de correlação linear

$$f(x, y) = \sum_{u, v=-k}^k \alpha_{u, v} f(x + u, y + v) + \mathcal{N}(x, y),$$

- ▶  $f$  é um canal de cor (RGB) após o demosaico,  $k$  é um inteiro e  $N$  distribuição normal iid
- ▶  $u, v$  são *offsets* dos *pixels*
- ▶  $\alpha$  é um vetor de coeficientes lineares que expressa as correlações

# Análise de Artefatos CFA

- ▶ Esperança estima a prob. de cada amostra pertencer ao Modelo 1

$$\Pr\{f(x, y) \in \mathcal{M}_1 | f(x, y)\} = \frac{\Pr\{f(x, y) | f(x, y) \in \mathcal{M}_1\} \Pr\{f(x, y) \in \mathcal{M}_1\}}{\sum_{i=1}^2 \Pr\{f(x, y) | f(x, y) \in \mathcal{M}_i\} \Pr\{f(x, y) \in \mathcal{M}_i\}},$$

- ▶  $\Pr\{f(x, y) \in \mathcal{M}_1\}$  e  $\Pr\{f(x, y) \in \mathcal{M}_2\}$  são prob. *a priori* (1/2)

# Análise de Artefatos CFA

- ▶ Se assumirmos que uma amostra é gerada pelo Modelo 1, a probabilidade de isso ocorrer é

$$\Pr\{f(x, y) | f(x, y) \in \mathcal{M}_1\} = \frac{1}{\sigma\sqrt{2\pi}} \left[ -\frac{1}{2\sigma^2} \left( f(x, y) - \sum_{u, v=-k}^k \alpha_{u, v} f(x + u, y + v) \right)^2 \right].$$

- ▶ Assumimos que o Modelo 2 tem uma distribuição uniforme
- ▶ Estimamos a variância no estágio M

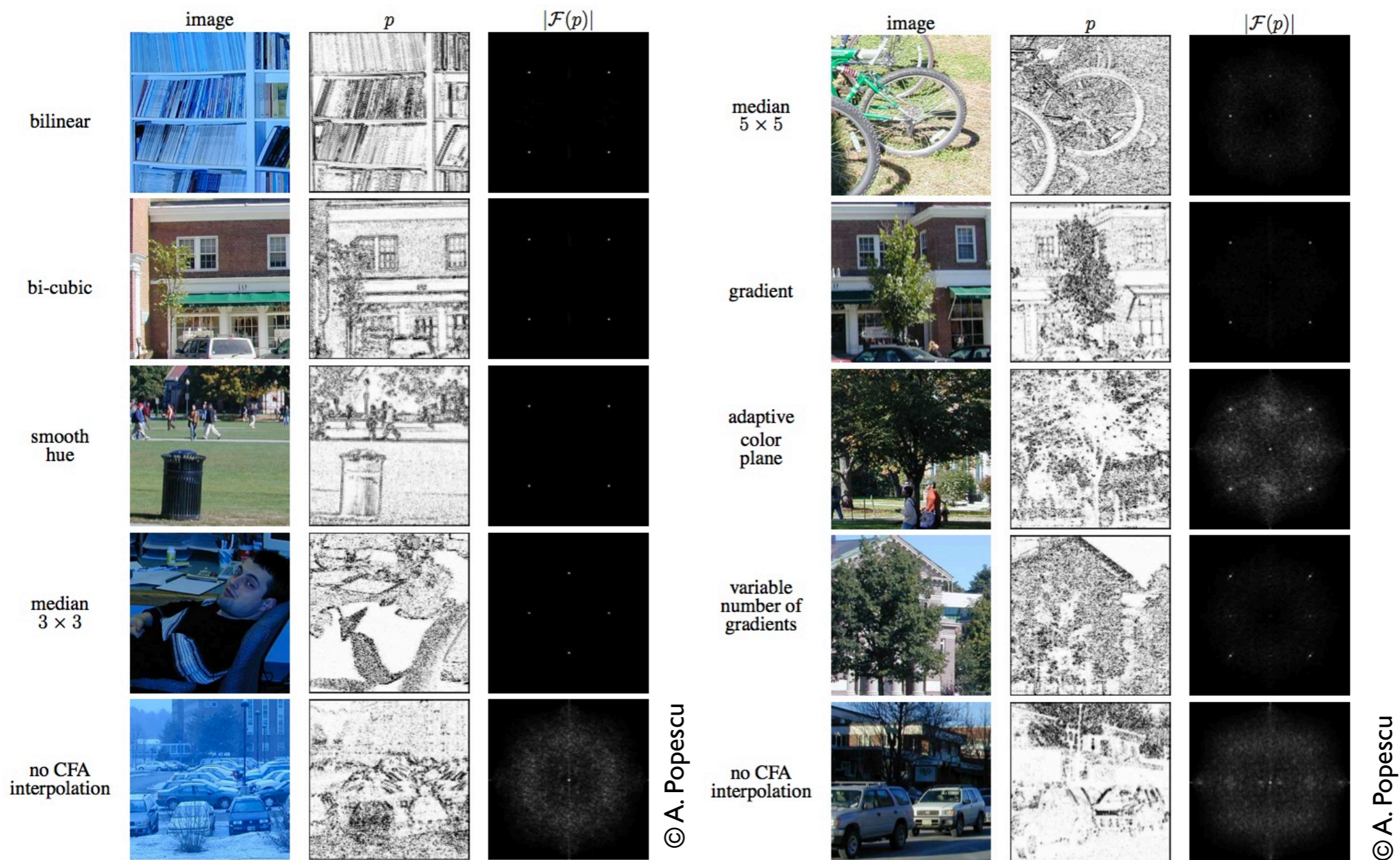
# Análise de Artefatos CFA

- ▶ No estágio M precisamos estimar os coeficientes da correlação
- ▶ Usamos método dos mínimos quadrados

$$E(\vec{\alpha}) = \sum_{x,y} w(x,y) \left( f(x,y) - \sum_{u,v=-k}^k \alpha_{u,v} f(x+u, y+v) \right)^2.$$

- ▶ Os pesos  $w$  são equivalentes a  $\Pr\{f(x,y) \in \mathcal{M}_1 | f(x,y)\}$

# Análise de Artefatos CFA

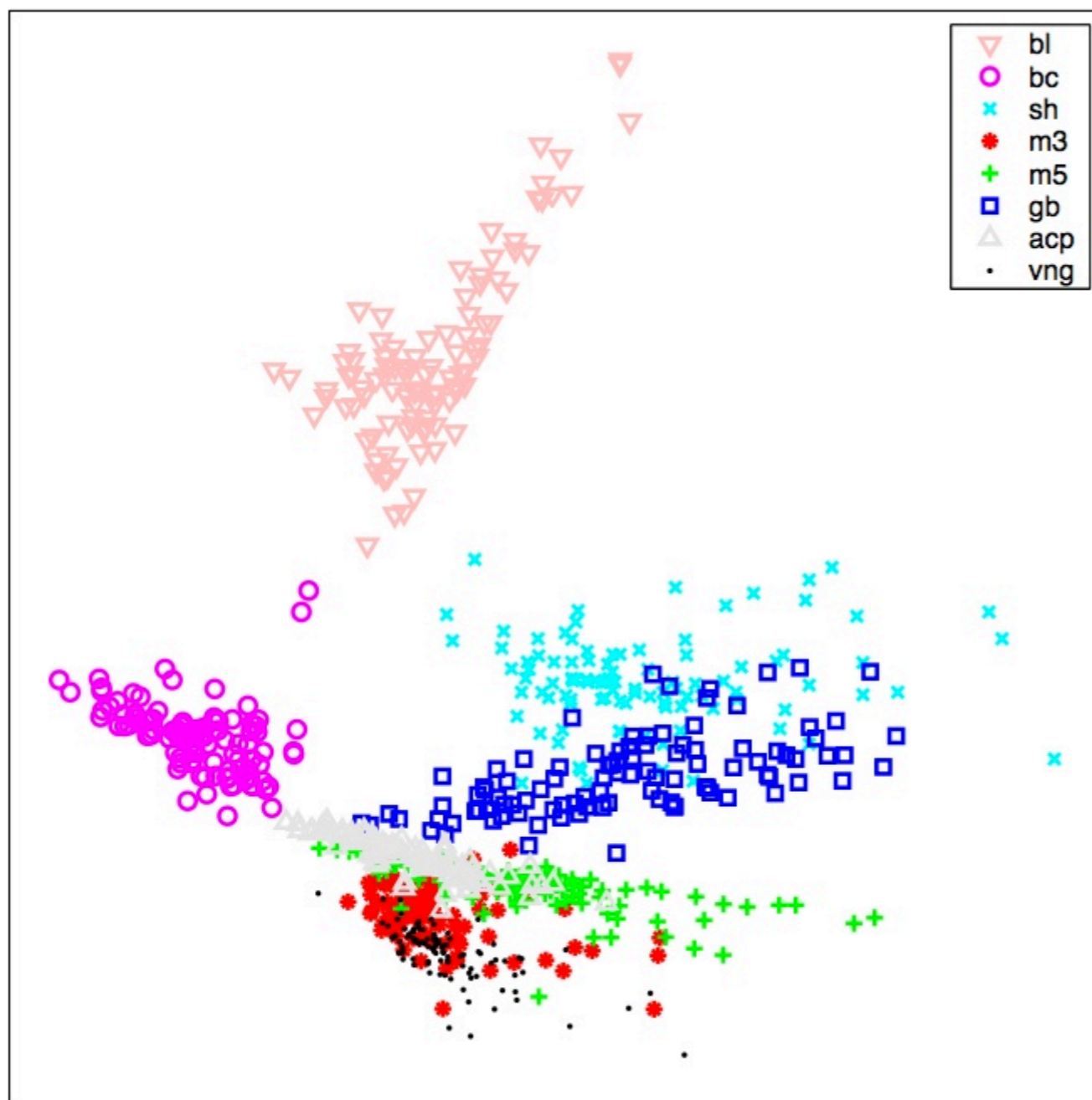


© A. Popescu

© A. Popescu

# Análise de Artefatos CFA

© A. Popescu



1. Cem imagens com interpolação CFA

2. Algoritmos

1. Bilinear
2. Bicubic
3. Smooth hue
4. Median (3 & 5)
5. Gradient
6. Adaptive Color Plane
7. Variable Gradients

# Análise de Artefatos CFA

- ▶ Problemas?
- ▶ Quais são os possíveis ataques a essa abordagem?
- ▶ Podemos empregar aprendizado de máquina?
- ▶ Podemos usar isso no cenário de investigação de adulterações de imagens?

# **Atribuição da Câmera em Específico**

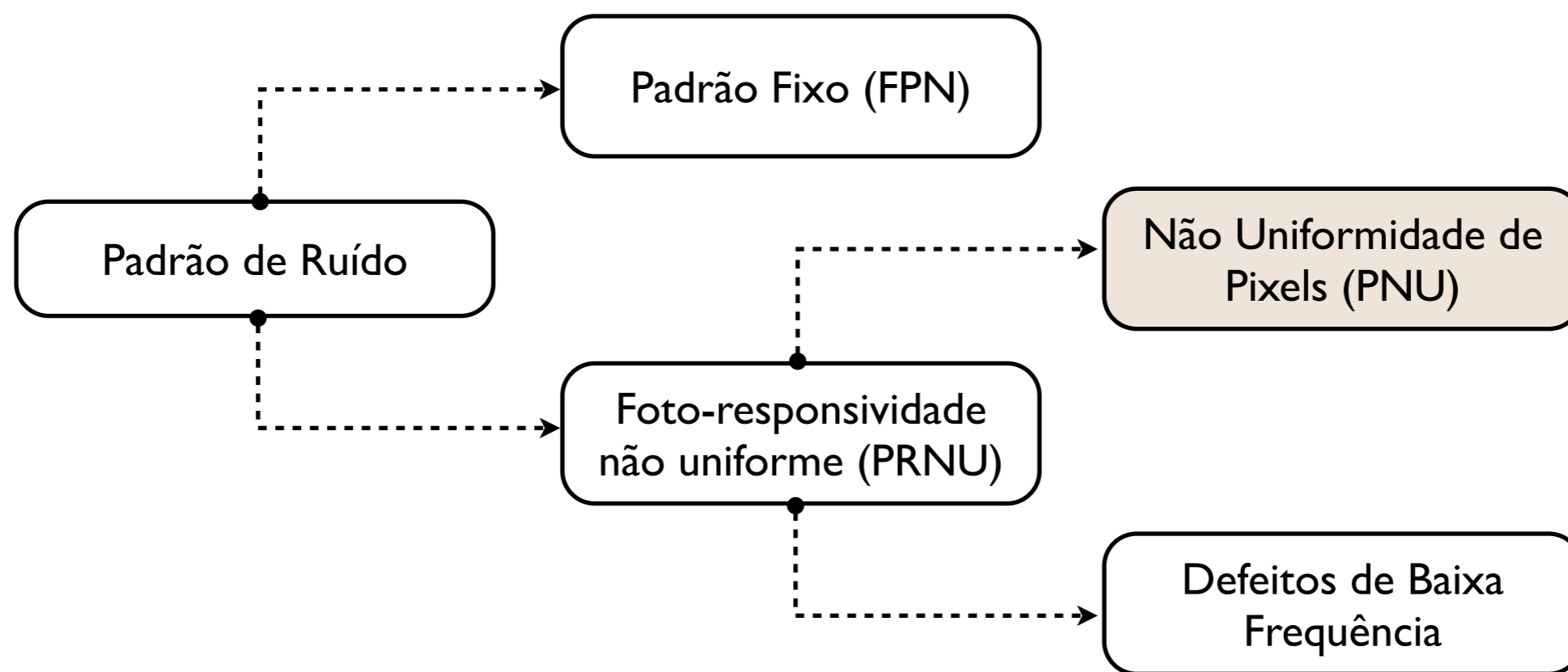
# Atribuição direta

- ▶ Características únicas da câmera em análise
  - Imperfeições dos componentes
  - Defeitos e falhas decorrentes do ambiente e condições de operação
  - *Dead e cold pixels, pixel traps etc.*

# Estado da Arte

- ▶ **Análise dos efeitos do ruído inserido no processo de captura**  
[Lukas et al. 2006]
- ▶ **Artefatos originados pela presença de poeira nos sensores no momento de aquisição**  
[Dirik et al. 2007]
- ▶ **Efemeridade**

# Análise de Ruído



© A. Rocha

# Análise de Ruído

## ▶ FPN

- Aditivo e decorrente de *dark-currents*
- dependente do tempo de exposição e temperatura
- Pode ser removido “*in-camera*” extraindo-se um quadro preto (*dark frame*)

# Análise de Ruído

## ▶ PRNU

- Multiplicativo
- Causado pelo ruído não uniforme dos *pixels* (PNU) e defeitos de baixa frequência
- Definido como a sensibilidade que diferentes *pixels* têm em relação à luz
- Causado por inconsistências na fabricação do sensor de captura

# Análise de Ruído

## ▶ PRNU

- Os defeitos de baixa frequência são causados pelos efeitos da refração da luz nas partículas devido a configurações de *zoom* e superfície ótica
- Não utilizado no trabalho de [Lukas et al. 2006]

# Análise de Ruído

- ▶ Para utilizar o ruído PNU no cenário forense, este precisa ser isolado
- ▶ Em um cenário forense certamente não teremos um padrão de referência
- ▶ Temos que estabelecer um padrão de referência  $P_c$ , uma aproximação do ruído PNU

# Análise de Ruído

- ▶ No processo de aproximação, fazemos a média de  $K$  imagens de uma cena uniforme (*lit scene*) como o céu, por exemplo

$$\bar{I}^{(k)} = \frac{1}{K} \sum_{k=1}^K I^k.$$

- ▶ Esta aproximação pode ser otimizada de modo a suprimir o conteúdo da cena

# Análise de Ruído

- ▶ Filtro de supressão/remoção de ruído (*denoising*)

$$\bar{\xi}^{(k)} = (I^{(k)} - \lambda(I^{(k)})) / K$$

- ▶ *Wavelet denoising* é uma boa escolha
- ▶ Como determinar se uma dada imagem  $I_{\text{teste}}$  pertence a uma determinada câmera?

# Análise de Ruído

- ▶ Como determinar se uma dada imagem  $I_{\text{teste}}$  pertence a uma determinada câmera?
- ▶ Calcula-se a correlação entre o ruído residual da imagem em questão e o padrão de referência

$$\rho_c(I) = \frac{(\xi - \bar{\xi}) \cdot (P_c - \bar{P}_c)}{\|\xi - \bar{\xi}\| \|P_c - \bar{P}_c\|}.$$

# Análise de Ruído

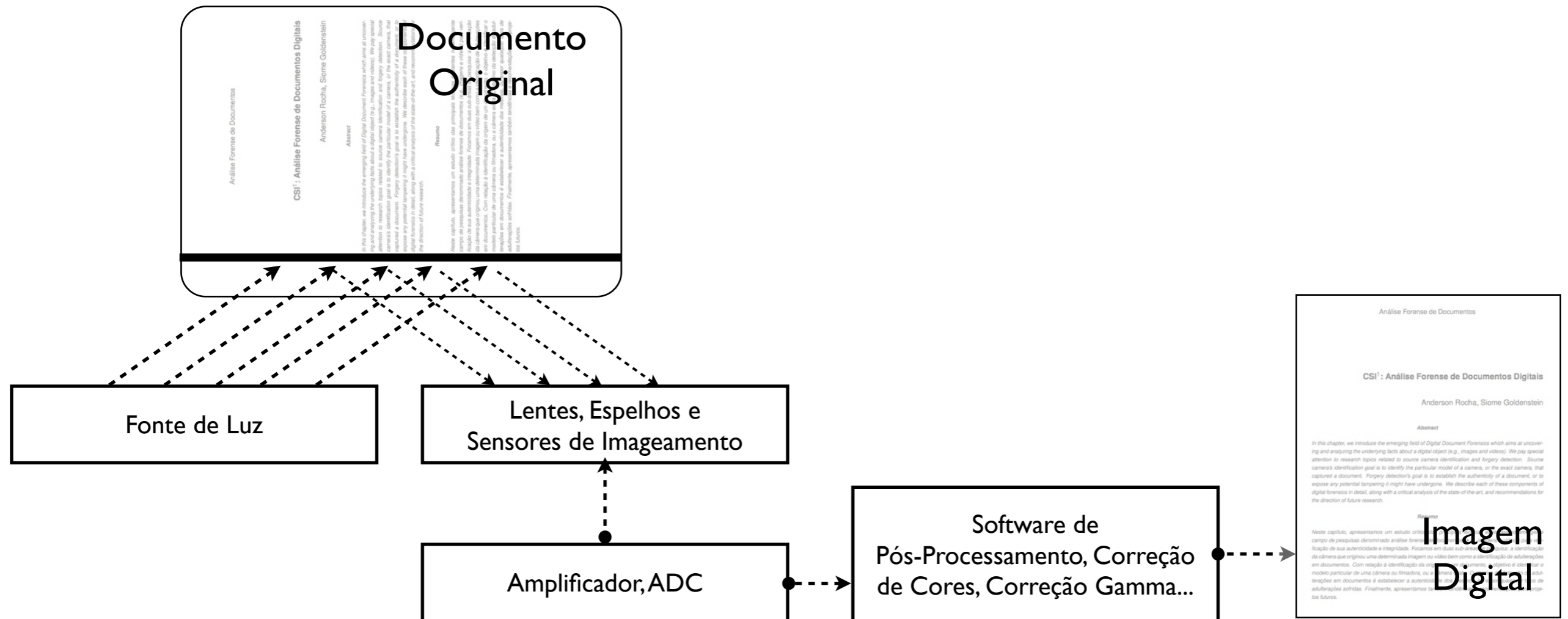
- ▶ Método com ótimos resultados
- ▶ Problemas
  - Sincronização (modificações de escala e recorte)
  - Força bruta para achar as transformações
- ▶ Podemos empregar aprendizado de máquina?

# **Atribuição de Scanner**

# Atribuição de Scanner

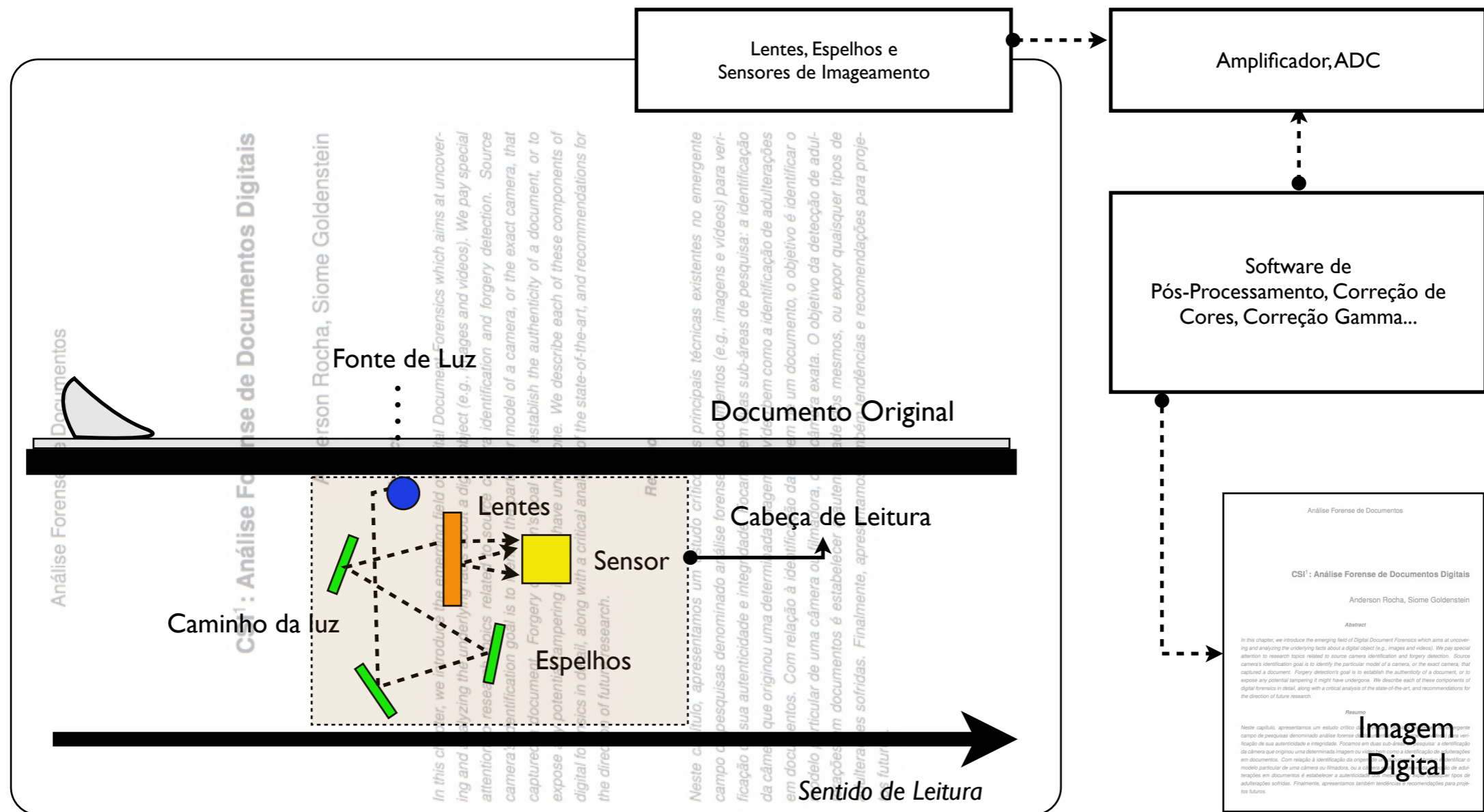
- ▶ O processo de fabricação de qualquer dispositivo eletrônico de captura introduz defeitos nos sensores de imageamento
- ▶ Como funciona o processo de captura de um scanner?

# Aquisição via Scanner (recap)



© A. Rocha

# Aquisição via Scanner (recap)



© A. Rocha

# Atribuição de Scanner

- ▶ Scanners possuem
  - Resolução horizontal (sensor) e vertical (motor de passo)
  - Motor de passo
  - Estabilizador

# Atribuição de Scanner

- ▶ Como os *scanners* fazem escaneamentos em resolução não-nativa?
  - Sub-amostragem
  - Amostragem normal seguida de ajustes *in-scanner* (maioria)

# Estado da Arte

- ▶ Extensões da análise de ruído FPN  
[Gloe et al. 2007a] e [Gou et al. 2007]
- ▶ **Análise de ruído considerando as propriedades de captura unidimensionais e periódicas do scanner**  
[Khanna et al. 2009]

# Ruído e Periodicidade

- ▶ Construção do padrão de referência
- ▶ Média das linhas (devido ao motor de passo)
- ▶ Média de múltiplas imagens
- ▶ Análise de correlação

# Ruído e Periodicidade

- ▶ Filtro de supressão/redução de ruído

$$I_{noise}^k = I^k - I_{denoised}^k$$

- ▶ Com  $K$  imagens podemos construir o padrão de referência (2D) de um dado scanner

$$\tilde{I}_{noise}^{array}(i, j) = \frac{1}{K} \sum_{k=1}^K I_{noise}^k(i, j);$$

$$1 \leq i \leq M \text{ e } 1 \leq j \leq N$$

# Ruído e Periodicidade

- ▶ Padrão de referência ID

$$\tilde{I}_{noise}^{linear}(1, j) = \frac{1}{M} \sum_{i=1}^M \tilde{I}_{noise}^{array}(i, j); \quad 1 \leq j \leq N.$$

- ▶ Como atribuir uma imagem em específico a um *scanner*?

# Ruído e Periodicidade

- ▶ Correlação entre a assinatura de um *scanner* e um padrão de referência

$$C(X, Y) = \frac{(X - \bar{X}) \cdot (Y - \bar{Y})}{\|X - \bar{X}\| \cdot \|Y - \bar{Y}\|}.$$

- ▶ Técnica com bons resultados na prática
- ▶ Palavra final?

# Ruído e Periodicidade

- ▶ Palavra final?
- ▶ Problemas
  - *Scanners* usam parcialmente a superfície de captura (vidro)
  - Esta abordagem requer condições similares de captura para funcionar

# Ruído e Periodicidade

- ▶ [Khanna et al. 2009] propõe uma análise estatística sobre os vetores de assinatura
- ▶ Média das linhas e colunas em  $\tilde{I}_{noise}^l$  e  $\tilde{I}_{noise}^c$
- ▶ Cálculo da correlação entre cada linha de  $I_{noise}$  e a média das linhas de  $\tilde{I}_{noise}^l$
- ▶ Similarmente para as colunas

# Ruído e Periodicidade

- ▶ Estatísticas de ordem sobre as novas características (média, variância, moda, curtose, etc.)
- ▶ Classificador de padrões
- ▶ Bons resultados

# **Técnicas Contra- Forenses em Atribuição**

# Abordagens Contra-Forenses

- ▶ Técnicas de atribuição de dispositivos de captura vs. Técnicas contra-forenses
- ▶ [Gloe et al. 2007b] observam que um filtro de supressão/redução de ruído baseado em *wavelets* não elimina o ruído totalmente
- ▶ Pode aplicar um método de *flatfielding* para capturar o ruído restante

# Abordagens Contra-Forenses

- ▶ Para estimar o ruído fixo (FPN) pode-se utilizar imagens de um quadro preto

$$I_{dark\_estimate} = \frac{1}{J} \sum_J I_{dark}.$$

- ▶ Para estimar o ruído PRNU, usa-se K imagens de uma cena homogênea (e.g., céu azul) subtraindo-se o a estimativa FPN

# Abordagens Contra-Forenses

- ▶ Estimativa de *flatfielding*

$$I_{flatfield} = \frac{1}{K} \sum_K (I_{light} - I_{dark\_estimate}).$$

- ▶ Com as estimativas em mãos, pode-se suprimir as características de ruído em uma imagem de uma câmera em específico
- ▶ Como?

# Abordagens Contra-Forenses

- ▶ Retirando a assinatura de ruído

$$\hat{I} = \frac{I - I_{dark\_estimate}}{I_{flatfield}}.$$

- ▶ *Flatfielding* é difícil devido ao grande número de parâmetros que precisam ser levados em conta
  - tempo de exposição
  - velocidade de captura
  - ISO etc.

# Abordagens Contra-Forenses

- ▶ Após a extração de padrão de ruído, como substituí-lo como outro padrão?
- ▶ O padrão de ruído de uma câmera pode ser substituído com a operação de *flatfielding inverso*

$$\hat{I}_{forge} = \hat{I} \cdot I_{flatfield\_forge} + I_{dark\_forge}.$$

# **Identificação de Criações Sintéticas**

# Identificação de Criações Sintéticas

## ▶ CPPA

- Possessão de imagens de menores de idade é considerado crime
- Imagens geradas em computador não são crime
- E se alguém fotografa um menor e modifica propriedades no computador?

# Estado da Arte

- ▶ **Decomposição multi-escala e análise estatística**  
[Lyu 2005]
- ▶ **Diferenças em modelos de superfície (imagens naturais vs. imagens sintéticas)**  
[Ng et al. 2005]
- ▶ **Propriedades de aquisição – ruído**  
[Dehnie et al. 2006]

# Estado da Arte

- ▶ **Análise do comportamento de imagens naturais e geradas em computador mediante sucessivas perturbações**  
[Rocha & Goldenstein 2007,  
Rocha & Goldenstein 2010]
- ▶ **Propriedades de aquisição – artefatos de mosaico/demosaico**  
[Dirik et al. 2007]

# Decomposição multi-escala e análise estatística

- ▶ [Lyu 2005], apresentaram uma técnica para identificação de imagens sintéticas baseado na decomposição *wavelet* de imagens
- ▶ Extensão de um trabalho anterior para o cenário de detecção de mensagens escondidas em imagens digitais

# Decomposição multi-escala e análise estatística

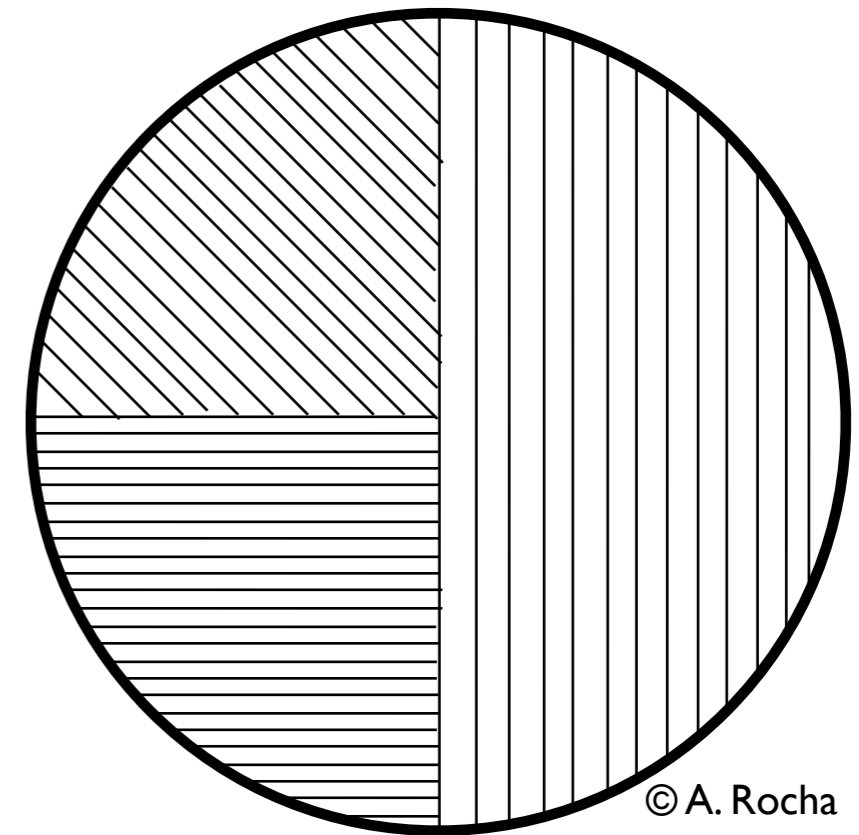
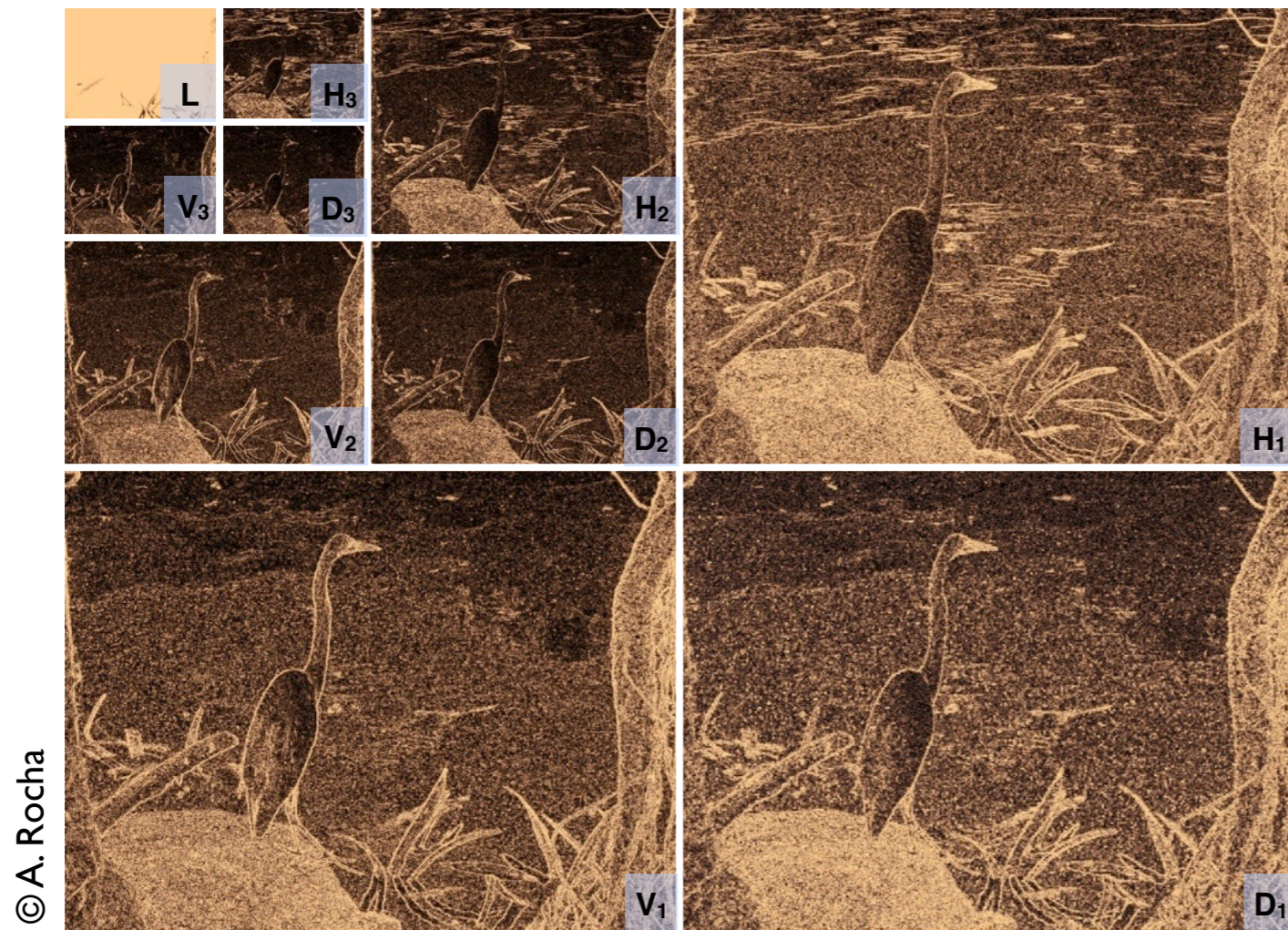
## ▶ Hipótese

- imagens naturais possuem regularidades detectáveis com estatísticas de alta ordem
- o processo de criação de uma imagem em computador insere artefatos estatísticos nas imagens produzidas

# Decomposição multi-escala e análise estatística

- ▶ Como funciona o método?
  - Decomposição QMF da imagem em múltiplas escalas e orientações
  - análise estatística em duas fases
    - ▶ análise direta
    - ▶ predição linear dos erros de magnitude

# Decomposição multi-escala e análise estatística



# Decomposição multi-escala e análise estatística

- ▶ A análise direta das múltiplas escalas e orientações consiste em calcular
  - média
  - variância
  - moda
  - curtose

# Decomposição multi-escala e análise estatística

- ▶ A predição linear dos erros de magnitude consiste em estimar as possíveis correlações entre um *pixel* e seus vizinhos multi-escala

# Decomposição multi-escala e análise estatística

- ▶ Preditor linear para a sub-banda vertical

$$\begin{aligned} V_i(x, y) = & w_1 V_i(x - 1, y) + w_2 V_i(x + 1, y) + w_3 V_i(x, y - 1) \\ & + w_4 V_i(x, y + 1) + w_5 V_{i+1}\left(\frac{x}{2}, \frac{y}{2}\right) + w_6 D_i(x, y) \\ & + w_7 D_{i+1}\left(\frac{x}{2}, \frac{y}{2}\right), \end{aligned}$$

- ▶  $w_k$  denota os valores escalares de peso dos coeficientes

# Decomposição multi-escala e análise estatística

- ▶ Como calcular os coeficientes e erro?

$$E(w) = [V - Qw]^2,$$

$$w = (w_1, \dots, w_7)^T$$

- ▶  $V$  contém os coeficientes de magnitude de  $V_i(x, y)$
- ▶  $F = 2 \times 3 \times 4 \times s = 72$  características para três escalas

# Diferenças físicas nos processos de captura e geração

- ▶ [Ng et al. 2005] desenvolvem dois níveis para separação de imagens naturais e sintéticas
  - Autenticidade em nível de processamento (imagens capturadas por algum sensor)
  - Autenticidade em nível de cena (*snapshot* de um campo de luz)
- ▶ Definição de diversas características

# Diferenças físicas nos processos de captura e geração

- ▶ **Dimensão fractal** local para capturar a complexidade de texturas
- ▶ **Vetores de patches** locais para capturar propriedades de arestas
- ▶ **Superfície gradiente** para capturar a forma de resposta de uma câmera
- ▶ **Geometria quadrática local** para capturar propriedades poligonais dos objetos computadorizados

# **Técnicas Contra- Forenses em CGI vs. Naturais**

# Abordagens Contra-forenses

- ▶ Ataques de recaptura
- ▶ Como ficam abordagens baseadas em análise de artefatos de captura?
- ▶ Como ficam abordagens baseadas em propriedades de textura e regularidade?
- ▶ A abordagem de [Ng et al. 2005]

# Referências

# Referências

1. [Bayram et al. 2005b] **Bayaram, S., Sencar, H., Memon, N. e Avcibas, I.** (2005b). *Source camera identification based on CFA interpolation*. In Intl. Conf. on Image Processing (ICIP), Genova, Italy.
- 2.[Choi et al. 2006] **Choi, K. S., Lam, E. e Wong, K.** (2006). *Automatic source camera identification using the intrinsic lens radial distortion*. Optics Express, 14(24):11551-11565.
- 3.[Dehnie et al. 2006] **Dehnie, S., Sencar, T. e Memon, N.** (2006). *Identification of computer generated and digital camera images for digital image forensics*. In Intl. Conf. on Image Processing (ICIP), Atlanta, USA.
- 4.[Dirik et al. 2007] **Dirik, E., Sencar, H. e Memon, N.** (2007). *Source camera identification based on sensor dust characteristics*. In IEEE Intl. Workshop on Signal Processing Applications for Public Security and Forensics (SAFE), pp. 1-6, Washington DC, USA.
- 5.[Gloe et al. 2007a] **Gloe, T., Franz, E. e Winkler, A.** (2007a). *Forensics for flat-bed scanners*. In SPIE Intl. Conf. on Security, Steganography, Watermarking of Multimedia Contents, pp. 65051-1.
- 6.[Gloe et al. 2007b] **Gloe, T., Kirchner, M., Winkler, A. e Bohme, R.** (2007b). *Can we trust digital image forensics?*
- 7.[Gou et al. 2007] **Gou, H., Swaminathan, A. e Wu, M.** (2007). *Robust scanner identification based on noise features*. In SPIE Security, Steganography, and Watermarking of Multimedia Contents (SSWMC), San Jose, USA.
- 8.[Khanna et al. 2009] **Khanna, N., Mikkilineni, A. K. e Delp, E. J.** (2009). *Scanner identification using feature-based processing and analysis*. IEEE Transactions on Information Forensics and Security (TIFS), 4(1):123-139.

# Referências

- 10.[Kharrazi et al. 2004] **Kharrazi, M., Sencar, H. e Memon, N. (2004)**. *Blind source camera identification*. In Intl. Conf. on Image Processing (ICIP), Singapore.
- 11.[Lukas et al. 2006] **Lukas, J., Fridrich, J. e Goljan, M. (2006)**. *Digital camera identification from sensor noise sensor*. IEEE Transactions on Information Forensics and Security (TIFS), 1(2):205-214.
- 12.[Lukas et al. 2006] **Lukas, J., Fridrich, J. e Goljan, M. (2006)**. *Digital camera identification from sensor noise sensor*. IEEE Transactions on Information Forensics and Security (TIFS), 1(2):205-214.
- 13.[Lyu 2005] **Lyu, S. (2005)**. *Natural Image Statistics for Digital Image Forensics*. Phd thesis, Dep. of Computer Science - Dartmouth College, Hanover, USA.
- 14.[Ng et al. 2005] **Ng, T.-T., Chang, S.-F. e Tsui, M.-P. (2005)**. *Physics-motivated features for distinguishing photographic images and computer graphics*. In ACM Multimedia (ACMMM), pp. 239-248, Singapore.
- 15.[Popescu 2004] **Popescu, A. C. (2004)**. *Statistical Tools for Digital Image Forensics*. Phd thesis, Dep. of Computer Science - Dartmouth College, Hanover, USA.
- 16.[Rocha & Goldenstein 2007] **Rocha, A. e Goldenstein, S. (2007)**. *PR: More than meets the eye*. In Intl. Conf. on Computer Vision (ICCV), pp. 1-8.

---

***Obrigado!***

---