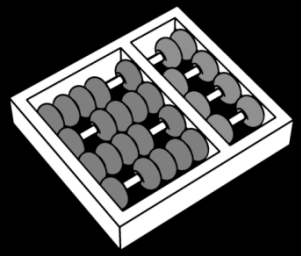


Digital Image Forensics for Device Attribution

Anselmo Ferreira and Anderson Rocha

Reasoning for Complex Data Lab.
Institute of Computing
State University of Campinas

Summary

- 1. The problem
- 2. Moving towards the solution
  - 2.1 How devices work
  - 2.2 Understand devices fingerprinting
  - 2.3 Devices attribution
    - 2.3.1 State of the art
- 3. Anti-Forensics approaches
- 4. Conclusion

The problem

- High availability of electronic devices:
  - They are cheap!
- Our daily documents are now created with these devices:
  - Images.
  - Images created by scanning processes.
  - Videos.
  - Printed Documents.

The problem

- Problem: not only legal documents are created:
  - Terrorist plans.
  - Fake currency.
  - Child Pornography and animal abuse photos.
- How to prove the ownership of these criminal documents?

The problem

Israeli soldier posts Instagram image of Palestinian child in crosshairs of rifle

Military investigates Mor Ostrovski, 20, as row grows over spate of offensive images posted online by Israeli soldiers

Phoebe Greenwood
theguardian.com, Monday 18 February 2013 10.06 GMT

Israeli soldier Mor Ostrovski, 20, has sparked controversy after posting this image on his Instagram account. Photograph: electronicintifada.net

Figure 1: Controversial Image posted in Instagram highlights the problem of device source attribution. Extracted from [1].

The problem

GIZMODO

It's Surprisingly Easy to Print Fake Money on an Inkjet Printer

Sarah Zhang
Filed to: MONEY 5/07/14 6:00pm

62,544 2 ★

If you're running an international counterfeiting ring, then yes, you're gonna need some expensive equipment. But for the small-time counterfeiter about town, it's all too easy. Just grab your everyday inkjet printer.

As Bloomberg News reports, 34-year-old hairstylist and janitor Tarshema Brice faked up to $20,000 in counterfeit bills.

Figure 2: one example on how devices can be used for criminal purposes. Extracted from[2].

The problem

Cocoa man charged with child porn

J.D. Gallop, FLORIDA TODAY 6:05 p.m. EDT May 23, 2014

(Photo: Cocoa Police Department)

f 42 CONNECT | TWEET | in LINKEDIN | 1 COMMENT | EMAIL | MORE

Cocoa police arrested a man suspected of keeping dozens of illicit images of underage children on his computer, a month after authorities said he used a phone app to search for runaway youth in a case that involved a lewd act.

Barry Gill Vest, 53, was charged with 175 counts of possession of child pornography today and ordered held on a $175,000 at the Brevard County Detention Center in Sharpes. Many of the illicit images involved multiple children, police report.

Figure 3:child porn photographies arise the importance of devices attribution. Extracted from [3].

How to move towards the solution

- What is Device Attribution?

  "A set of computer vision techniques applyied in a digital version of a document, aimed at pointing out which device is the source of the document."

How to move towards the solution

Figure 4: which printer printed which photo? Extracted from [4].

How to move towards the solution

Figure 5: which device created which document? Extracted from [4].

How to move towards the solution

- Devices Attribution can be done by searching for two kinds of signatures:
  - Intrinsic: inserted by the device in the document.
  - Extrinsic (Blind): given by the analysis of the resulting document.

How to move towards the solution

- Device Attribution involves answering two questions:

1. Which device model and brand produced a given document?
2. What specific device produced a given document?

How to move towards the solution

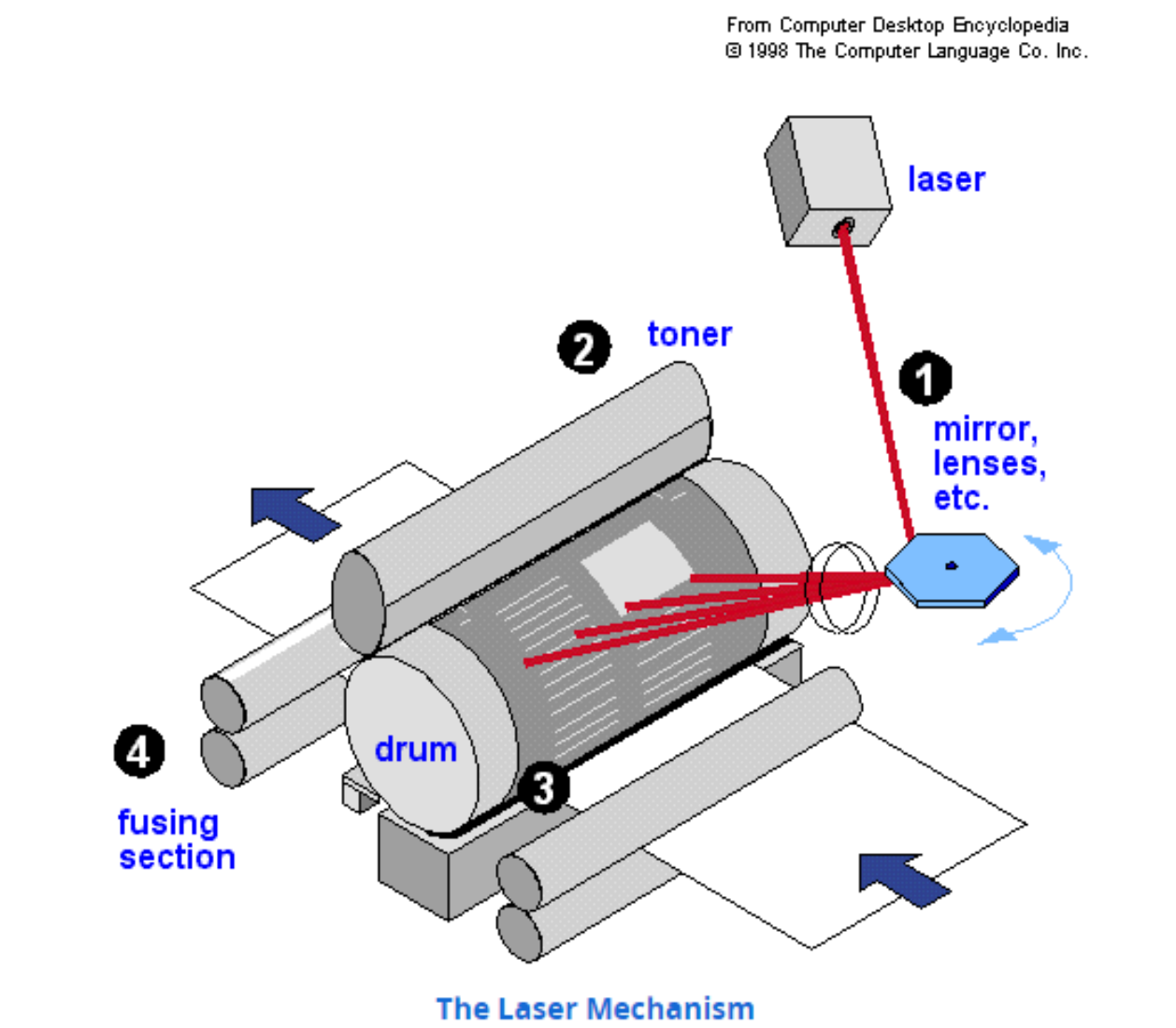
- Steps for device attribution:

1. Understand how these devices work.
2. Find unique behavior of each device in the document (e.g., Noise, Texture and Distortions).
3. Describe these behaviors for device attribution.

STEP ONE
HOW THESE DEVICES WORK

HOW LASER PRINTERS WORK

- Data written by a laser beam, which discharges certain places in a drum where ink must be put.
- Positively charged ink is then sticked in the discharged places of the drum.
- The data with ink is spread on the paper by the fuser.



STEP TWO
INVESTIGATING DEVICES FINGERPRINTING

INVESTIGATING DEVICES FINGERPRINTING

- Common fingerprints:
  - ✓ Texture
  - ✓ Distortion
  - ✓ Noise
  - ✓ Imperfections such as dust, scratches, etc.
  - ✓ Among others
- They are commonly yielded by devices manufacturing process.

INVESTIGATING DEVICES FINGERPRINTING

- LASER PRINTERS
  - ✓ Have moving parts that behave differently.
  - ✓ Differences seen on halftones of printed material.
  - ✓ **Banding** [5]**:** nonuniform light and dark lines printed horizontaly.

INVESTIGATING DEVICES FINGERPRINTING

- LASER PRINTERS ATTRIBUTION
  - ✓ Uses a digitalized (scanned) version of a document.
    - Approaches:
      - ✓ Frequency Analysis of Banding (Fourier Spectrum Analysis) in halftones for color documents
      - ✓ Texture among printed material for text.
      - ✓ Etc.

INVESTIGATING DEVICES FINGERPRINTING

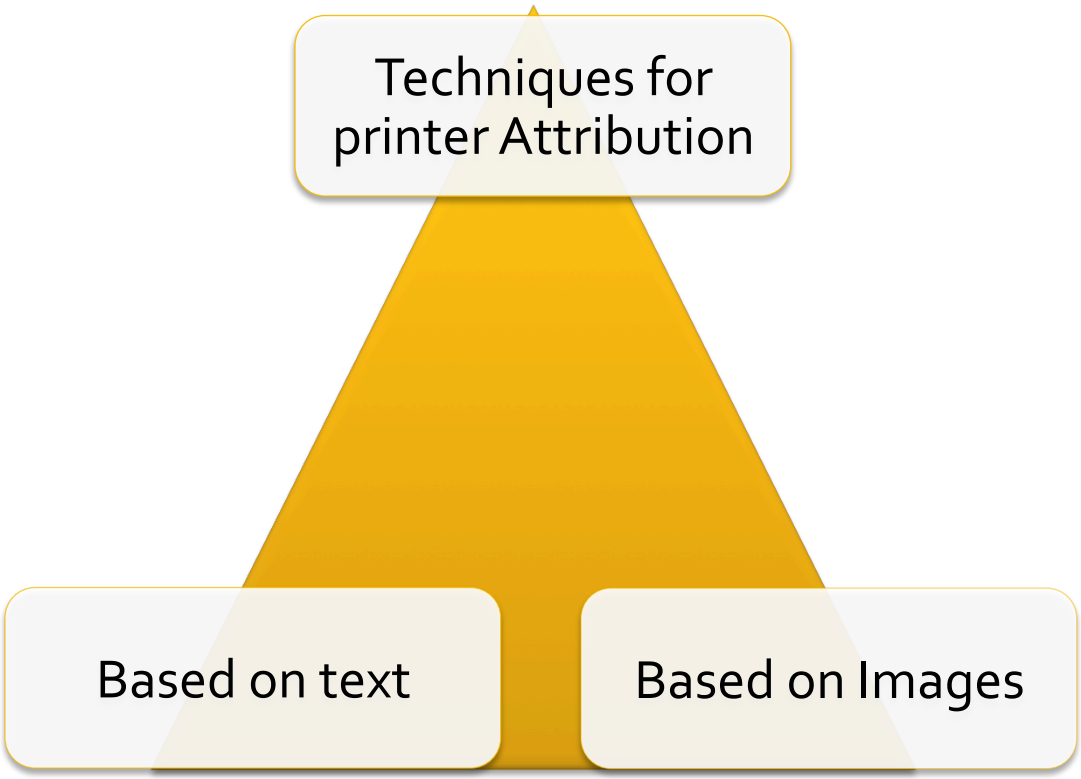


Figure 5: Approaches for Printer Attribution differ when applied to text or color documents.

INVESTIGATING DEVICES FINGERPRINTING

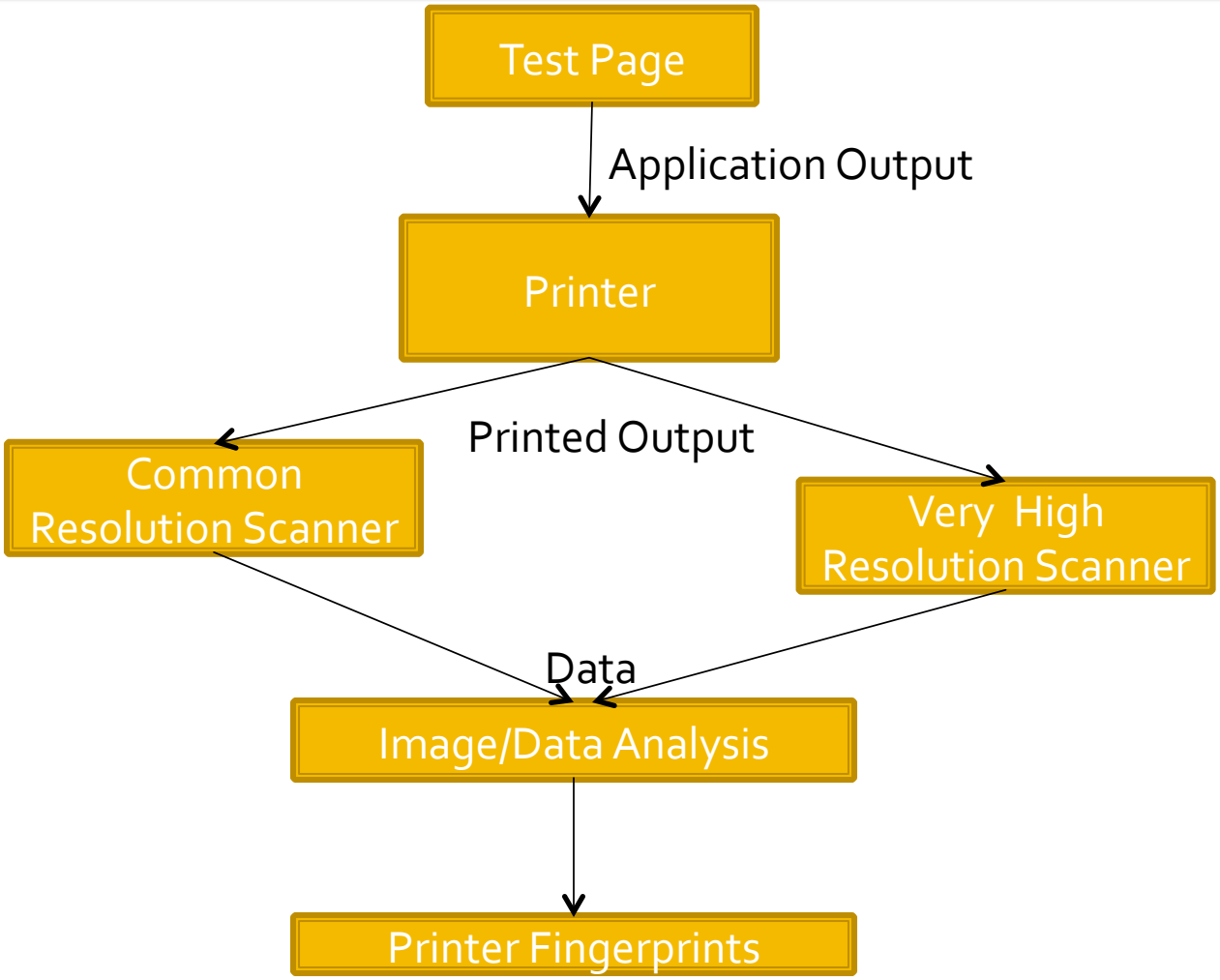


Figure 6: steps for printer attribution

INVESTIGATING DEVICES FINGERPRINTING

- SCANNERS
  - Attribution based on source camera identification (i.e., sensor noise).
  - Two-dimensional (2-D) noise pattern of the image is used to identify source scanner

INVESTIGATING DEVICES FINGERPRINTING

- SCANNERS ATTRIBUTION
  - Approaches:
    - ✓ Statistical features.
    - ✓ high-frequency wavelet coefficients.
    - ✓ neighborhood prediction errors.
    - ✓ Dust or scratches.

STEP THREE- Describing devices fingerprinting for devices attribution

Describing devices fingerprinting for devices attribution

- LASER PRINTERS






Describing devices fingerprinting for devices attribution

- Approaches used for laser printer attribution.
  - halftone-based [6, 7]: applied only in color documents.
  - texture-based [8-13]: applied on text documents.
  - noise-based [14, 15, 16]: applied in both.

- Machine learning plays an important role.

Describing devices fingerprinting for devices attribution

- The following areas of a document can be used for analysis in laser printer attribution:
  - Letters -> applied only in text
  - Frames -> applied in text and images
  - Whole Document-> applied in text and images

Describing devices fingerprinting for devices attribution

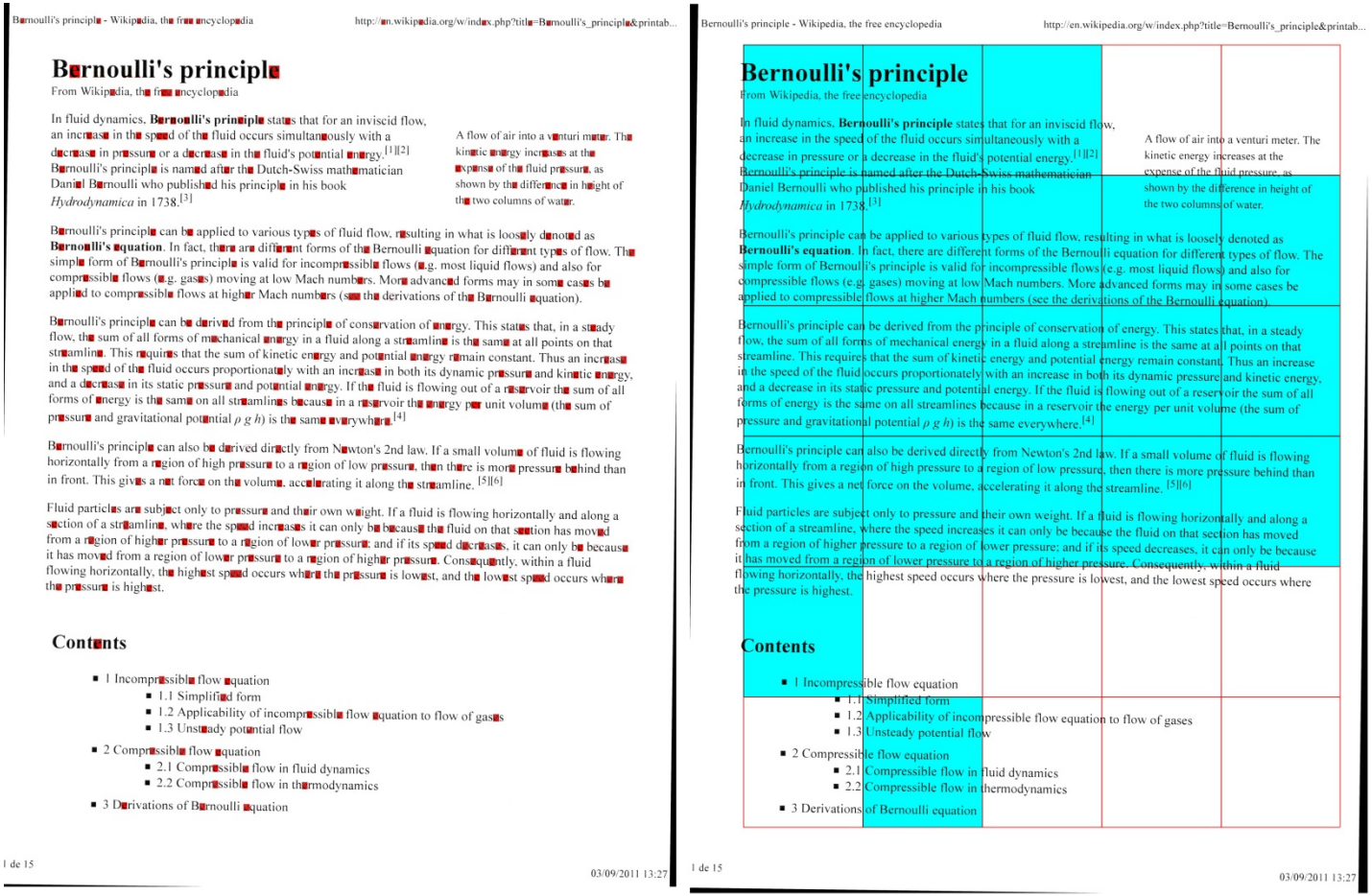


Figure 7: Letter (left) and Frame approach (right).

Describing devices fingerprinting for devices attribution

- Laser Printer attribution by Ali et al [8]:
  - Applied in letters of text (letter 'I')
  - The projection (pixel values) are used as fingerprints.
  - Gaussian mixture model classifier is used to recognize these texture patterns for each printer.

Describing devices fingerprinting for devices attribution

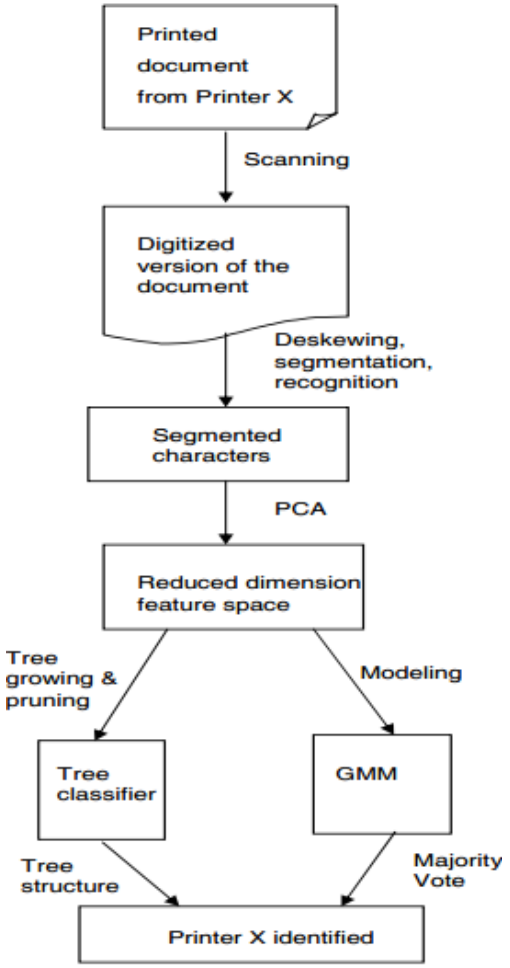


Figure 8: Ali et al's approach for laser printer attribution. Extracted from [8].

Describing devices fingerprinting for devices attribution

- Laser Printer attribution by Lee et al [14]:
  - Applied in color documents (images)
  - Documents are scanned and converted to CMY color space
  - noise of CMY image is isolated.

Describing devices fingerprinting for devices attribution

- Laser Printer attribution by Lee et al [14]:
  - Texture information is calculated by statistics of five GLCMs.
  - A machine learning classifer is used to recognize these texture patterns

Describing devices fingerprinting for devices attribution

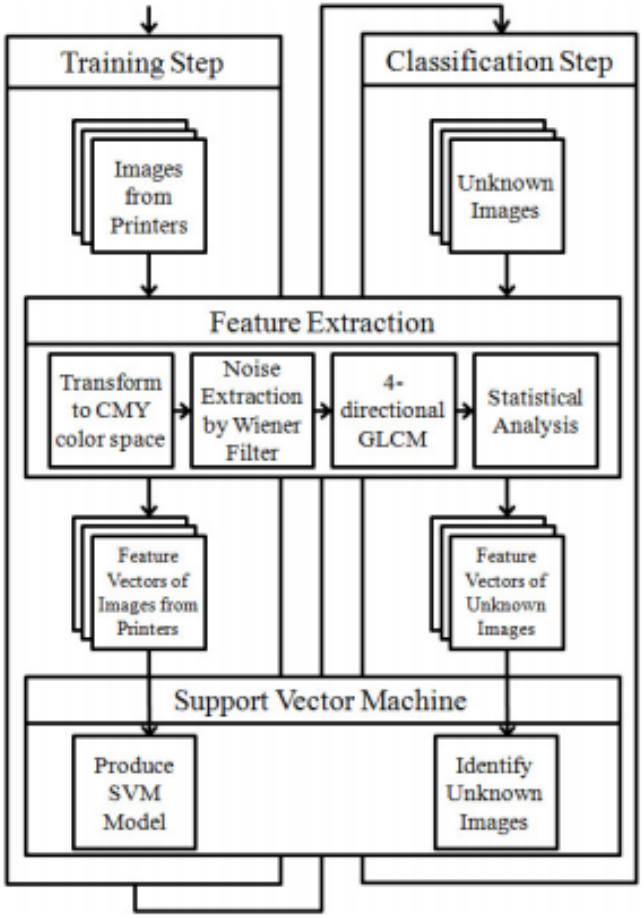


Figure 9: Lee et al's approach for laser printer attribution. Extracted from [14].

Describing devices fingerprinting for devices attribution

- Laser Printer Attribution by Mikkilineni et al [9]
  - Applied in text documents.
  - Letters "e" are extracted.
  - Statistics over one GLCM is used with machine learning.

Describing devices fingerprinting for devices attribution

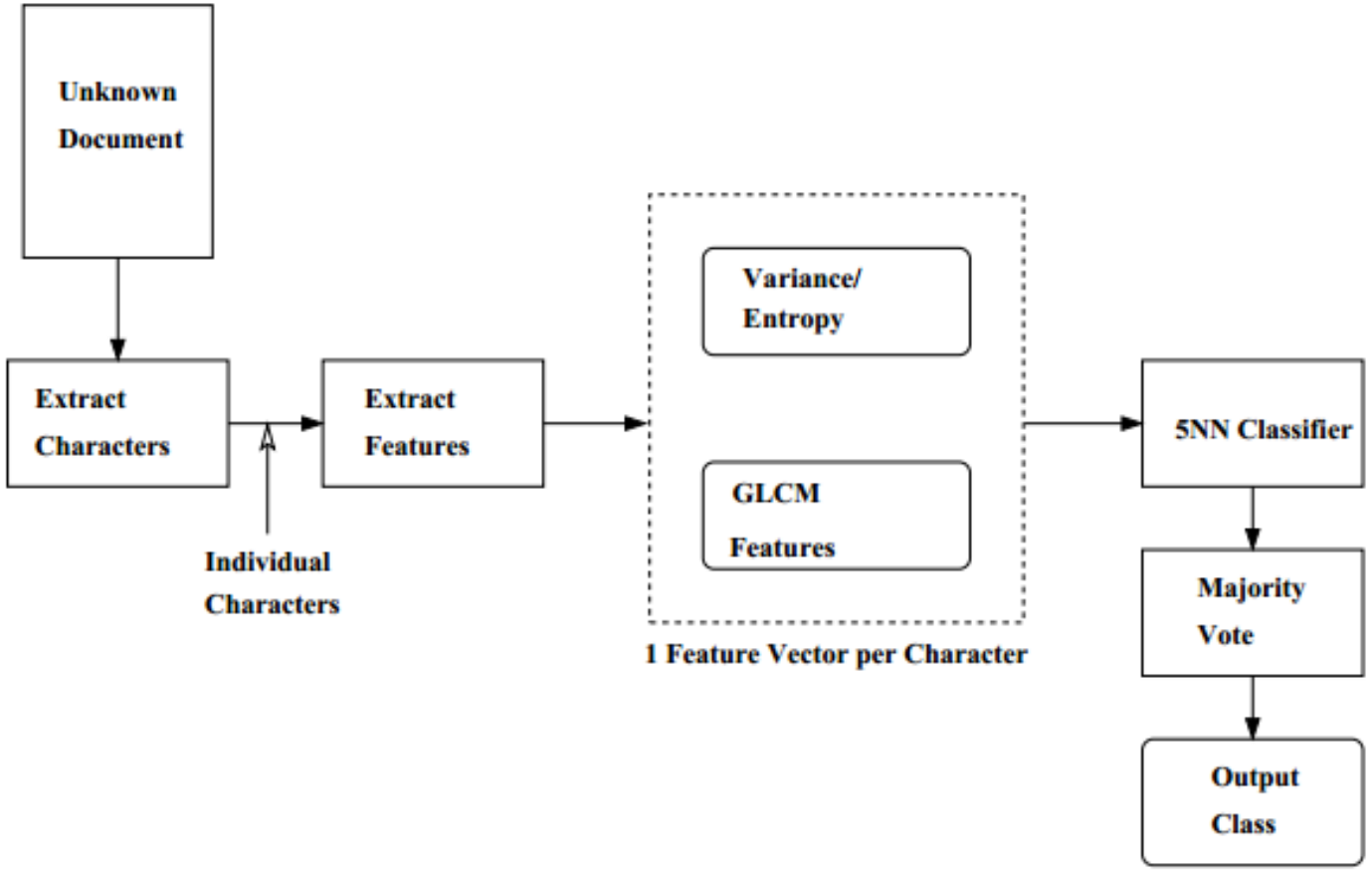


Figure 10: Mikkilineni et al's approach for laser printer attribution. Extracted from [9].

Describing devices fingerprinting for devices attribution

- Laser Printer Attribution by Kee and Farid [10]
  - Applied in text documents
  - Letters "e" are extracted.
  - Technique is divided in three steps
    - Pre-processing
    - Printer Profile
    - Ballistics

Describing devices fingerprinting for devices attribution

- Laser Printer Attribution by Kee and Farid [10]
  - Pre-processing.
  1. A reference letter is chosen.
  2. Similar letters are searched, preprocessed by histogram normalization and registered with the reference letter.

Describing devices fingerprinting for devices attribution

- Laser Printer Attribution by Kee and Farid [10]
  - Printer Profile.
    1. Aligned letters are used as columns of a Matrix *D*.
    2. PCA is performed in *D.*
    3. Printer profile: both the mean letter $\vec{u}$ and the top *p* eigenvalue eigenvectors yielded by PCA: $e_{ij}, i \in [1, p]$

Describing devices fingerprinting for devices attribution

- Laser Printer Attribution by Kee and Farid [10]
  - Ballistic.
    1. Test letters in vector form c_j are first aligned to the reference character.
    2. Each letter is then projected onto each printer space:

$$\alpha_{ji} = (c_j - \vec{u})e_i \ (1)$$

Describing devices fingerprinting for devices attribution

- Laser Printer Attribution by Kee and Farid [10]
  - Ballistic.
    3. the letter in printer space is then reconstructed, using the printer´s space parameters.

$$r_j = \vec{u} + \sum_{i=1}^{p} \alpha_{ji} \vec{e}_i \quad (2)$$

Describing devices fingerprinting for devices attribution

- Laser Printer Attribution by Kee and Farid [10]
  - Ballistic.
    4. The reconstruction error is then calculated. It must be minimum for the source printer.

$$E_j = \sqrt{(\vec{c}_j - \vec{r}_j)^T (\vec{c}_j - \vec{r}_j)} \quad (3)$$

    5. Again, this is done per printer space (i.e., per printer)

Describing devices fingerprinting for devices attribution

- Laser Printer Attribution by Choi et al [16]
  - Applied in color documents (images)
  - Based on statistics of DWT from CMYK color bands.
  - 39 statistical features are extracted from HH sub-band per image.

Describing devices fingerprinting for devices attribution

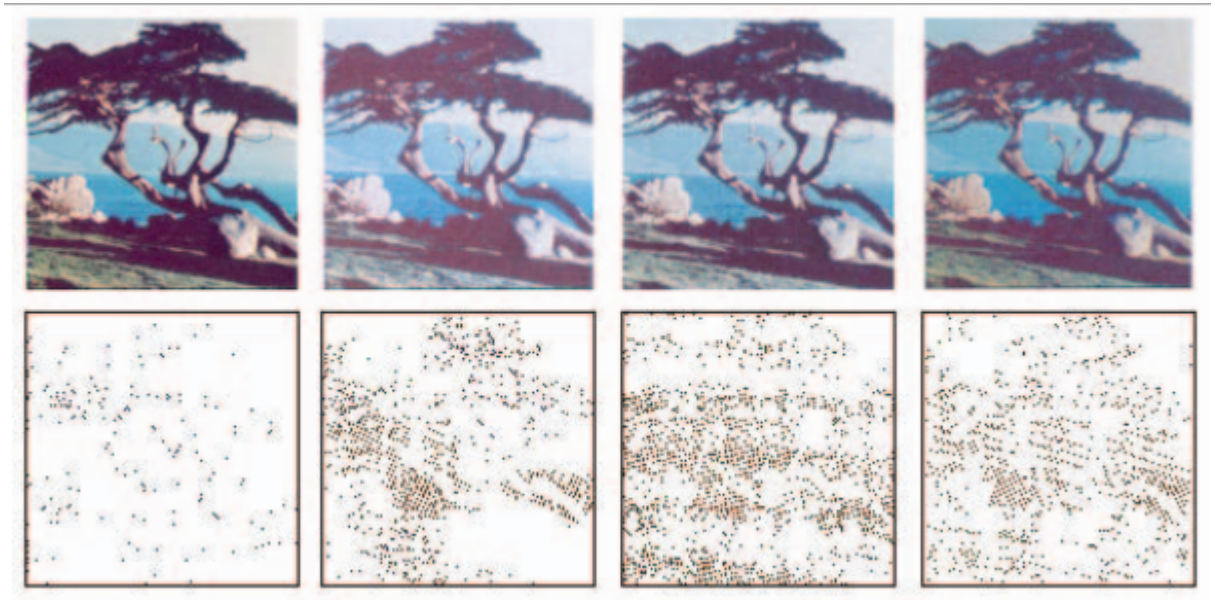


Figure 11: Printer noise signatures seen from HH sub-band of DWT from the same image printed with 4 different printers. Extracted from [16].

Describing devices fingerprinting for devices attribution

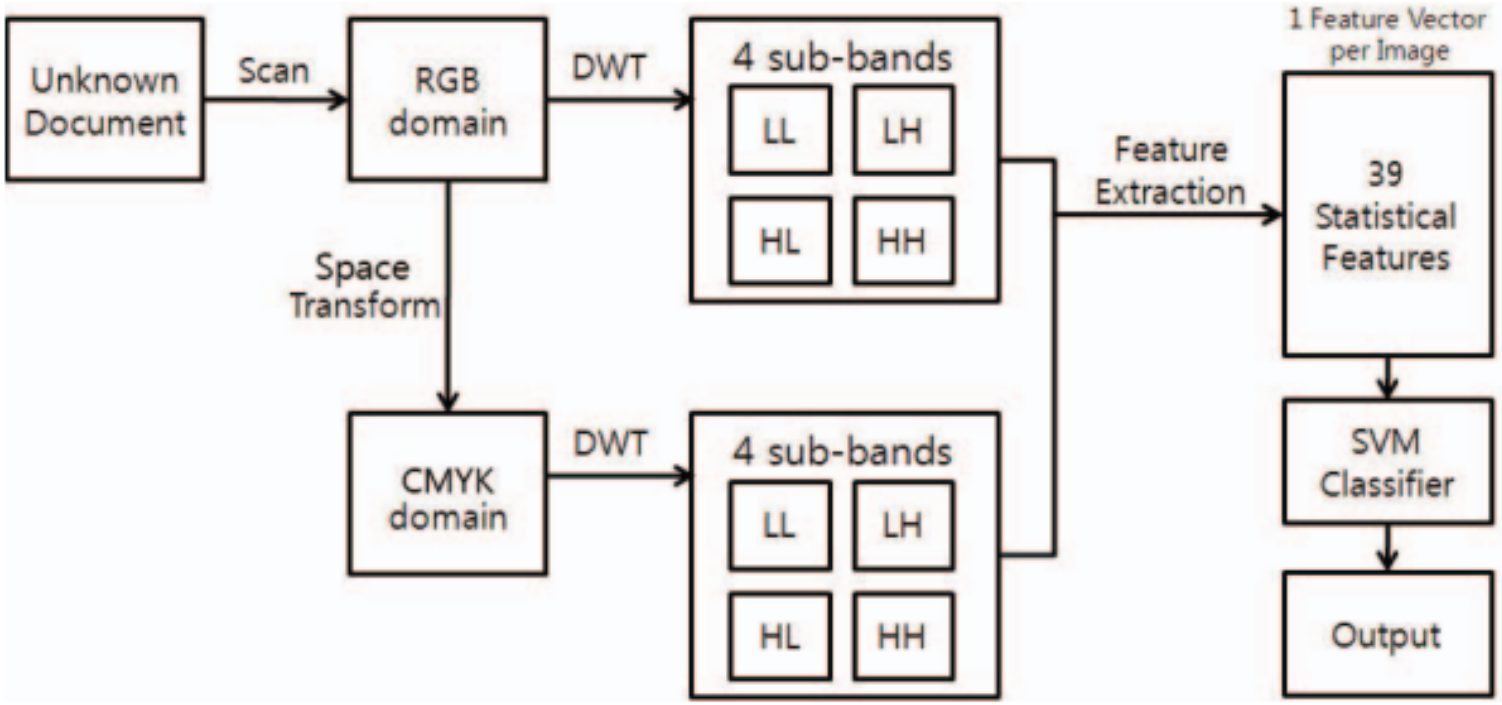


Figure 12: Choi et al's approach for laser printer attribution. Extracted from [16].

Describing devices fingerprinting for devices attribution

OUR SOLUTION!

Describing devices fingerprinting for devices attribution

- Our solution:
  - Multiscale and multidirectional texture analysis inside printed material.
  - Works on images, texts or both.
  - Can be applied in whole document, in letters or frames.

Describing devices fingerprinting for devices attribution

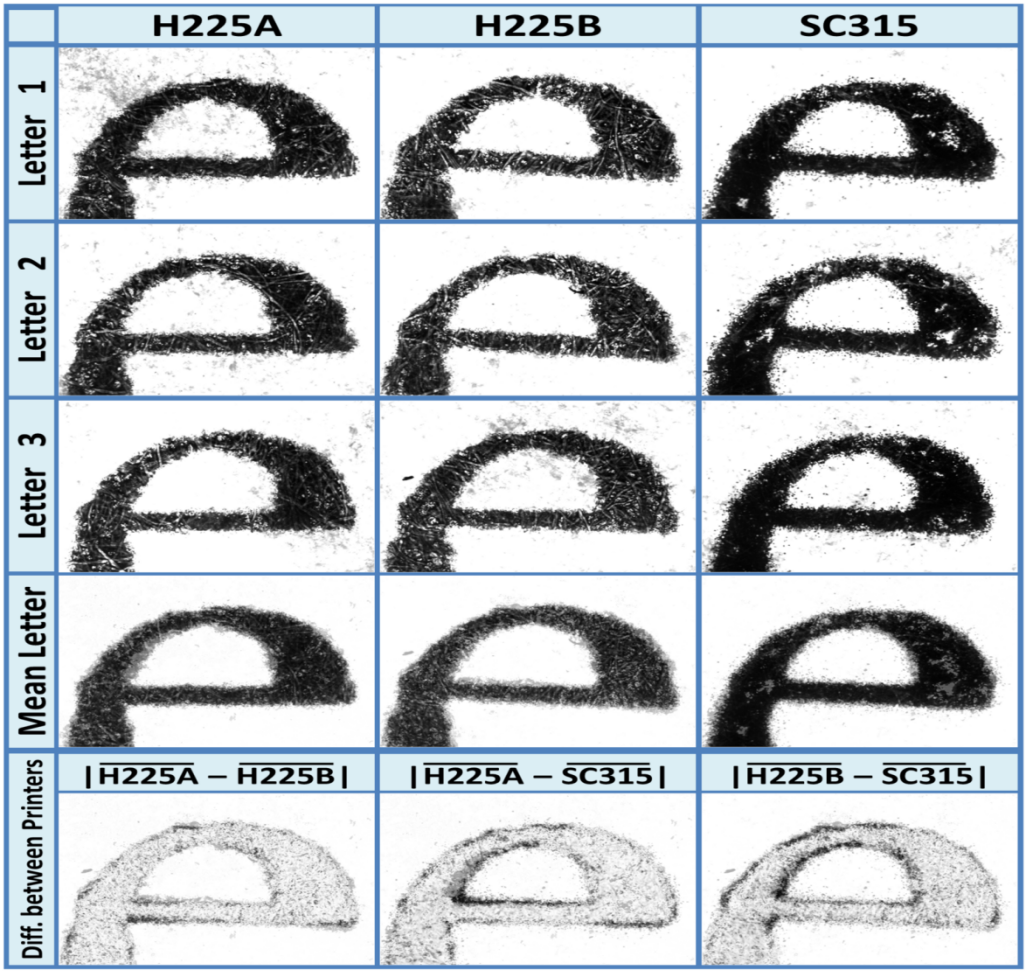


Figure 14: Microscope analysis of printed text shows that inside printed material there are multidirectional and multiscale texture patterns.

Describing devices fingerprinting for devices attribution

- Our Contributions
  1. Multidirectional GLCM approach
  2. Multidirectional/multiscale GLCM approach
  3. Convolutional gradient multidirectional and multiscale texture filter

Describing devices fingerprinting for devices attribution

- Our Contributions
  4. More realistic dataset
  5. Investigation on chunks of documents (frames)
  6. Dimensionality reduction approach

Describing devices fingerprinting for devices attribution

1. Multidirectional GLCM approach

   - GLCMs: 2d histograms that describe the pixel neighborhood in a given direction and distance (offset).

   - A series of statistics are calculated from these matrices and are used for image description.

   - We use more directions (eight) in the GLCM approach.

Describing devices fingerprinting for devices attribution

1. Multidirectional GLCM approach

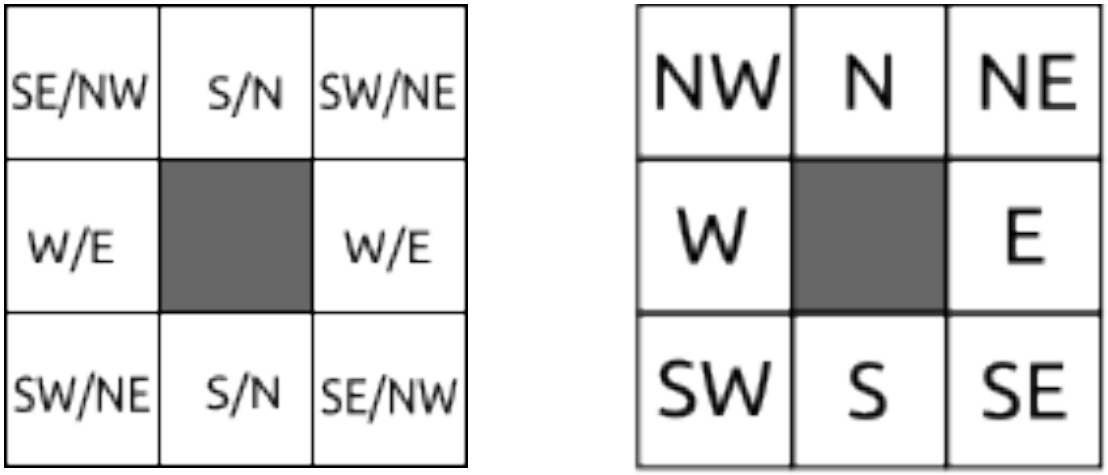


Figure 15: (Left) original approach yields 4 GLCMs (Right) our proposed approach yields 8 GLCMs.

Describing devices fingerprinting for devices attribution

1. Multidirectional GLCM approach

   - At each direction (GLCM), 22 statistics are calculated.

   - With 8 matrices, a 22 x 8=176 dimensional feature vector is used to classification.

Describing devices fingerprinting for devices attribution

2. Multidirectional/Multiscale GLCM approach

- We used the gaussian pyramidal image decomposition here.
- Four scales: the original, two downscales and one up-scale.
- At each, 176 statistical features are extracted as in the previous approach.

Describing devices fingerprinting for devices attribution

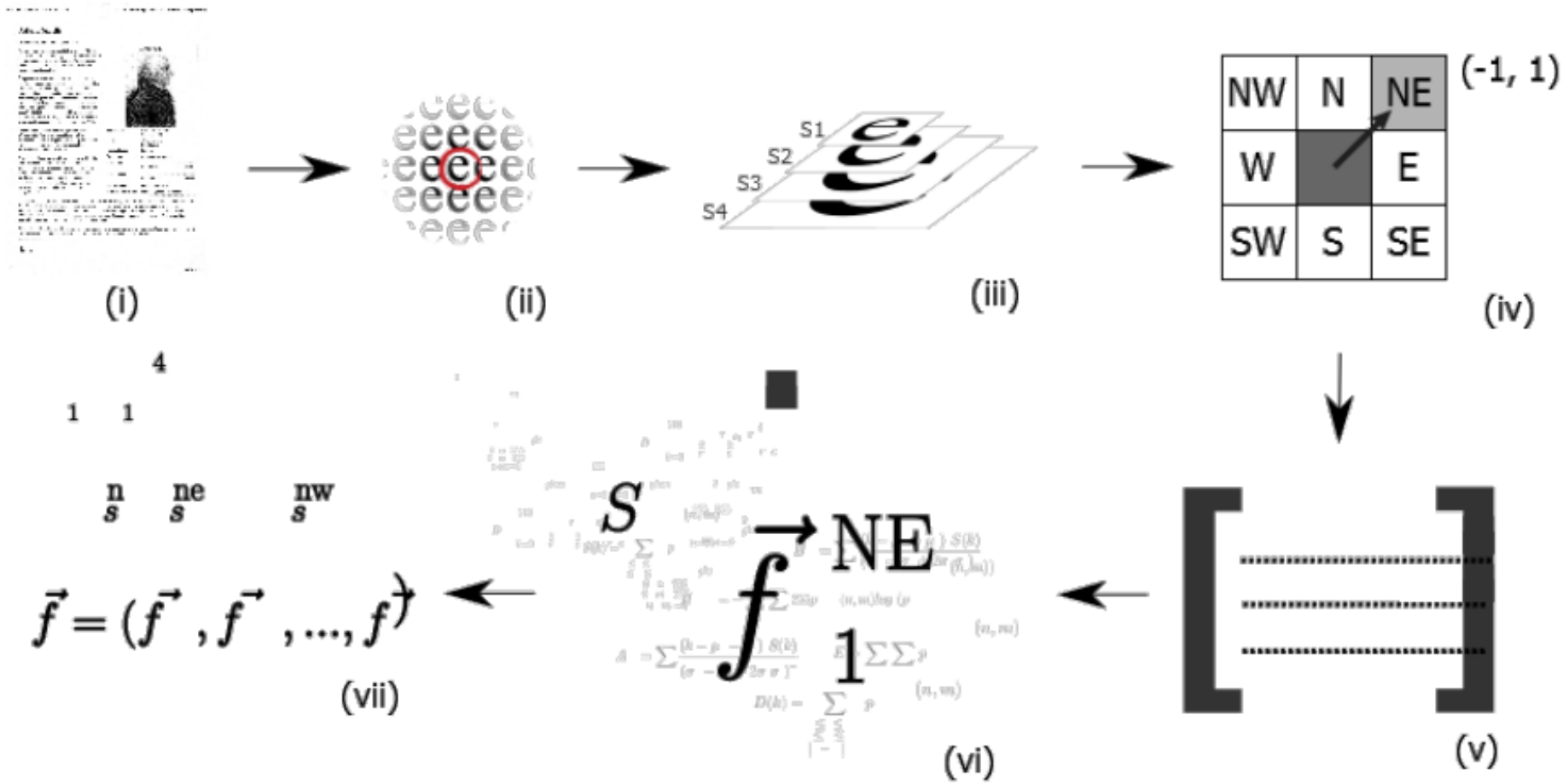


Figure 16: Our multidirectional and multiscale GLCM approach.

Describing devices fingerprinting for devices attribution

3. Convolutional gradient multidirectional and multiscale texture filter

   - Textures on areas with small gradient value are generated differently by different printer firmware.

   - We propose a filter to analyze texture in these areas: the Convolution Texture Gradient Filter (CTGF).

Describing devices fingerprinting for devices attribution

3. Convolutional gradient multidirectional and multiscale texture filter

   - CTGF filter the low-gradient textures in a set of n x n pixel neighborhood.

   - Seven transformations are applied to find the printer signature in a scanned Document, which we call *S*.

Describing devices fingerprinting for devices attribution

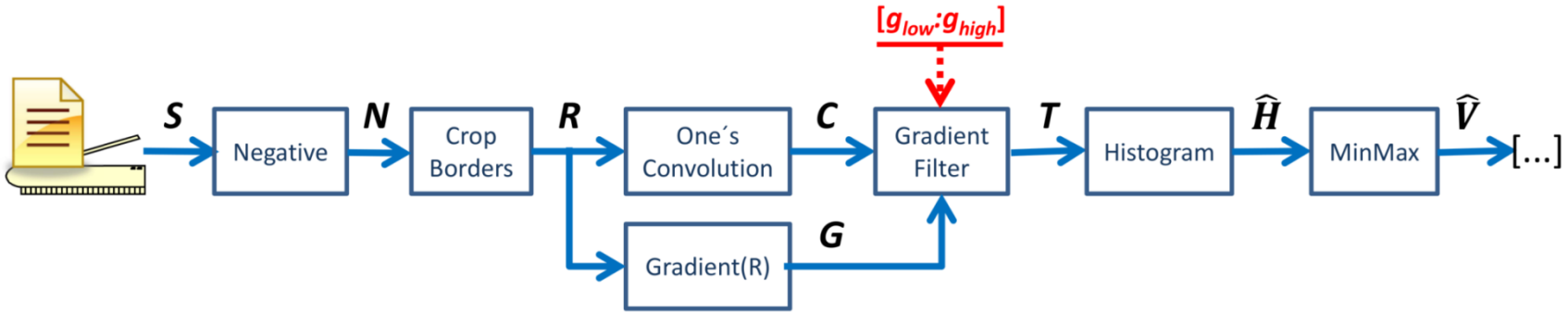


Figure 17: Our Convolutional Gradient Filter Approach

Describing devices fingerprinting for devices attribution

- **Step 1: Negative**
  - Pre-processing step: image pixels in S are inverted.
  - Values close to zero will mean white pixels and 255, black pixels.
  - This is made for convenience, it yields a negative image *N* .

Describing devices fingerprinting for devices attribution

- **Step 2: Crop borders**
  - Eliminate scanning noise at image borders generated by external light.
  - N is cropped, eliminating 6% of pixels in each border.
  - New matrix: *R.*

Describing devices fingerprinting for devices attribution

- **Step 3.1: Convolution with ones.**
  - The texture is calculated by replacing the central pixel with the sum pixels in an *n x n* area centered at that pixel.
  - This is done by a convolution with a mask of ones.
  - This results in the matrix of textures sums *C* .

Describing devices fingerprinting for devices attribution

- **Step 3.1: Convolution with ones**

$$C = R * \begin{vmatrix} 1 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & 1 \end{vmatrix} \quad (4)$$

Describing devices fingerprinting for devices attribution

- **Step 3.2: Gradient(R)**
  - In parallel, a gradient between a pixel and its n^2-1 neighbors is calculated.
  - The maximum gradient is used.
  - A new matrix, *G* is created.

Describing devices fingerprinting for devices attribution

- **Step 3.2: Gradient Filter**.
  - (g_{low}) and (g_{high}) define the range of gradient values that are valuable for printer signature.
  - This will filter textures with gradients of interest.
  - (g_{low}) and (g_{high}) are selected from a previous validation.

Describing devices fingerprinting for devices attribution

- **Step 3.2: Gradient Filter**
  - The matrix *T* of texture codes (sums) is then created by filtering textures that are in the defined range.

$$T = \begin{bmatrix} t_{1,1} & \dots & t_{1,c-2} \\ \dots & t_{i,j} & \dots \\ t_{r-2,1} & \dots & t_{r-2,c-2} \end{bmatrix}, \text{where} \begin{cases} t_{i,j} = c_{i,j} & \text{if } g_{low} \leq g_{i,j} \leq g_{max} \\ t_{i,j} = 0 & \text{otherwise} \end{cases}. \quad (5)$$

Describing devices fingerprinting for devices attribution

- **Step 4: Histogram**
  - A histogram of low-gradient textures is then built to identify that printer.

$$H = hist(T, 1 : 255 \times n^2) \ (6)$$

Describing devices fingerprinting for devices attribution

- **Step 5: MinMax.**
  - Final feature vector V is generated by applying a Min-Max normalization on H.

$$V_{ij} = \frac{H_{ij} - u}{v - u} \quad u = min(H), v = max(H) \quad (7)$$

  - The final feature vector has $(255 \times n^2) - 1$ dimensions.

Figure 18: Filtered textures using the proposed CTGF in printed text (top) and images (bottom) are different in different printers.

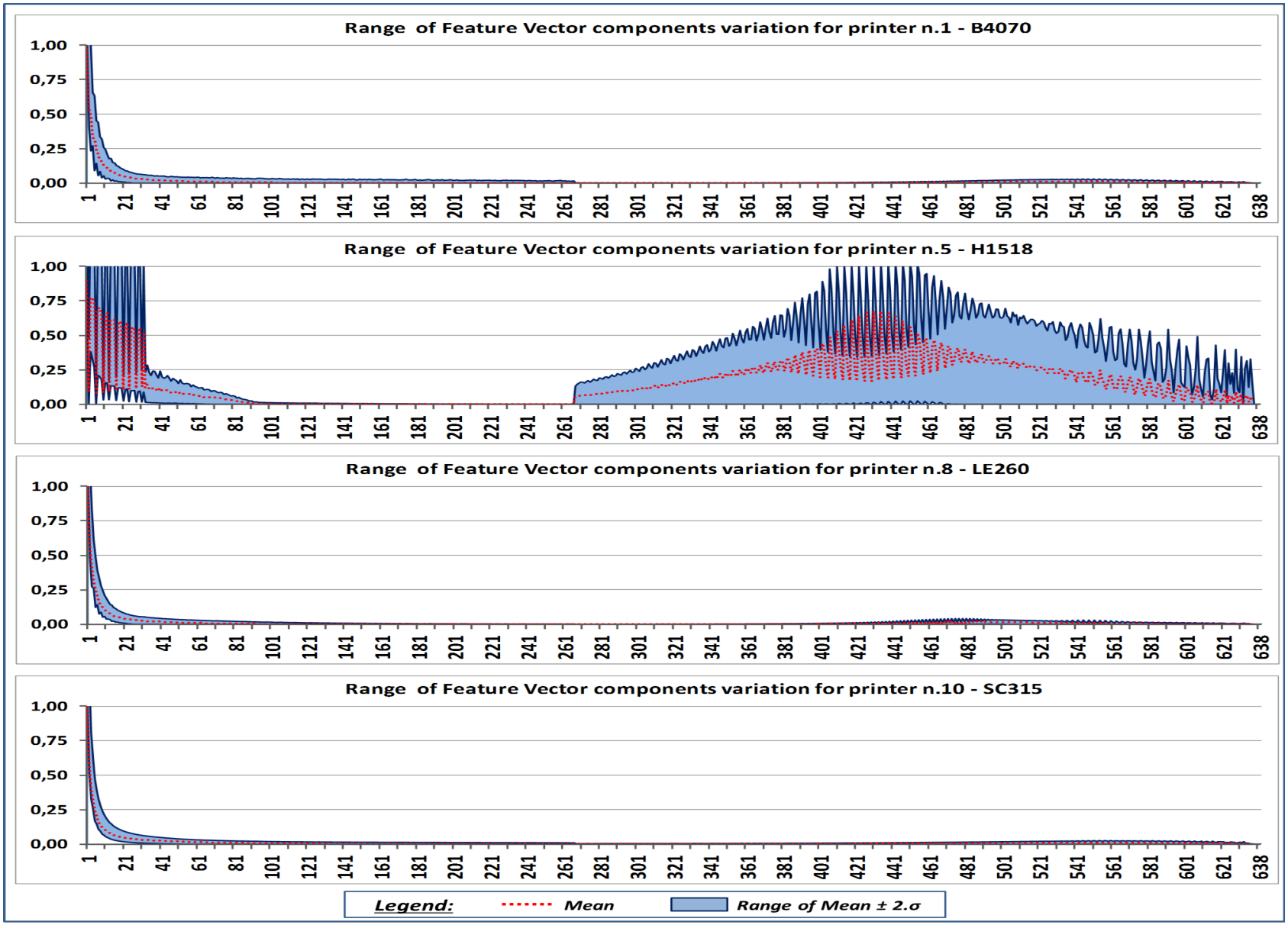


Figure 19: Printer signatures from different printers using the proposed CTGF.

Describing devices fingerprinting for devices attribution

4. More realistic dataset
   - Databases used in prior works :
     - always consider fonts of same size and style.
     - composed by text or figures
     - they expect high resolutions in scanning process

Describing devices fingerprinting for devices attribution

4. More realistic dataset
   - Wikipedia documents with one, two or three pages
   - Contains different letter sizes, fonts and figures.
   - Total of 1,184 documents.

Describing devices fingerprinting for devices attribution

4. More realistic dataset
   - Separated by two factors: Language (English or Portuguese) and Figures (With or Without).
   - Scanner: Canon 8800 Scanner at a 600 dpi resolution
   - Images saved in TIFF format.

Describing devices fingerprinting for devices attribution

#	Printer ID	Manufacturer	Laser Printer Model	Number of Printed Documents
1	B4070	Brother	HL-4070CDW	120
2	C1150	Canon	D1150	116
3	C3240	Canon	MF3240	120
4	C4370	Canon	MF4370DN	120
5	H1518	Hewlett Packard	CP1518	120
6	H225A	Hewlett Packard	CP2025A	119
7	H225B	Hewlett Packard	CP2025B	110
8	LE260	Lexmark	E260DN	119
9	OC330	OKI Data	C330DN	120
10	SC315	Samsung	CLP-315	120
			Total	**1,184**

Table 1: number of documents per printer used in our proposed dataset.

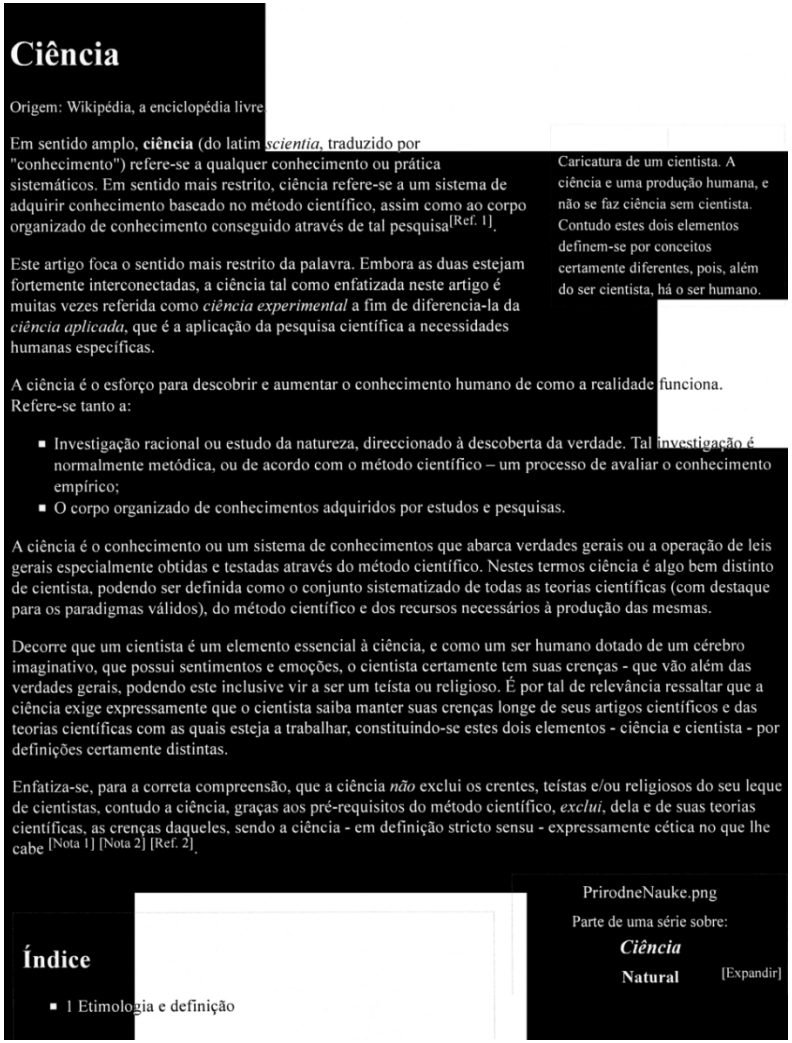
Describing devices fingerprinting for devices attribution

5. Investigation on chunks of documents (frames)
   - We proposed analysis in segmented areas.
   - Useful when only parts of documents are available.
   - They are rectangular areas with sufficient printed material.

Describing devices fingerprinting for devices attribution

5. Investigation on chunks of documents (frames)
   - Document (A4 paper) divided in a matrix of frames with five columns by six rows.
   - Minimum accepted ratio between dark pixels and blank ones in a frame should be 0.02.

Describing devices fingerprinting for devices attribution



Ciência

Origem: Wikipédia, a enciclopédia livre

Em sentido amplo, **ciência** (do latim *scientia*, traduzido por "conhecimento") refere-se a qualquer conhecimento ou prática sistemáticos. Em sentido mais restrito, ciência refere-se a um sistema de adquirir conhecimento baseado no método científico, assim como ao corpo organizado de conhecimento conseguido através de tal pesquisa[Ref. 1].

Caricatura de um cientista. A ciência e uma produção humana, e não se faz ciência sem cientista. Contudo estes dois elementos definem-se por conceitos certamente diferentes, pois, além do ser cientista, há o ser humano.

Este artigo foca o sentido mais restrito da palavra. Embora as duas estejam fortemente interconectadas, a ciência tal como enfatizada neste artigo é muitas vezes referida como *ciência experimental* a fim de diferencia-la da *ciência aplicada*, que é a aplicação da pesquisa científica a necessidades humanas específicas.

A ciência é o esforço para descobrir e aumentar o conhecimento humano de como a realidade funciona. Refere-se tanto a:

- Investigação racional ou estudo da natureza, direccionado à descoberta da verdade. Tal investigação é normalmente metódica, ou de acordo com o método científico – um processo de avaliar o conhecimento empírico;
- O corpo organizado de conhecimentos adquiridos por estudos e pesquisas.

A ciência é o conhecimento ou um sistema de conhecimentos que abarca verdades gerais ou a operação de leis gerais especialmente obtidas e testadas através do método científico. Nestes termos ciência é algo bem distinto de cientista, podendo ser definida como o conjunto sistematizado de todas as teorias científicas (com destaque para os paradigmas válidos), do método científico e dos recursos necessários à produção das mesmas.

Decorre que um cientista é um elemento essencial à ciência, e como um ser humano dotado de um cérebro imaginativo, que possui sentimentos e emoções, o cientista certamente tem suas crenças - que vão além das verdades gerais, podendo este inclusive vir a ser um teísta ou religioso. É por tal de relevância ressaltar que a ciência exige expressamente que o cientista saiba manter suas crenças longe de seus artigos científicos e das teorias científicas com as quais esteja a trabalhar, constituindo-se estes dois elementos - ciência e cientista - por definições certamente distintas.

Enfatiza-se, para a correta compreensão, que a ciência *não* exclui os crentes, teístas e/ou religiosos do seu leque de cientistas, contudo a ciência, graças aos pré-requisitos do método científico, *exclui*, dela e de suas teorias científicas, as crenças daqueles, sendo a ciência - em definição stricto sensu - expressamente cética no que lhe cabe [Nota 1] [Nota 2] [Ref. 2].

PrirodneNauke.png

Parte de uma série sobre:

Ciência

Natural [Expandir]

Índice

- 1 Etimologia e definição

Figure 20: Frames of interest(in black) extracted from a scanned document.

Describing devices fingerprinting for devices attribution

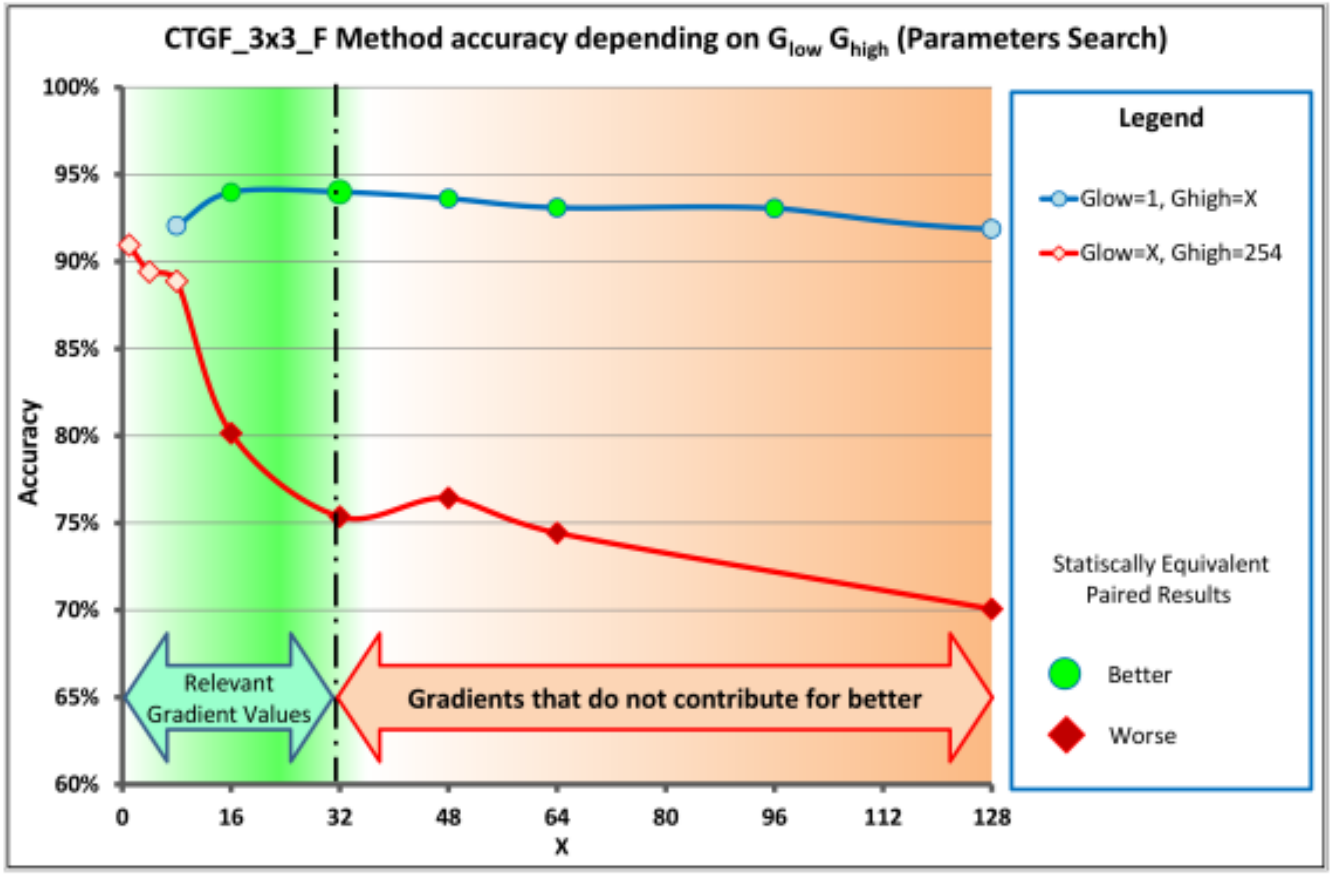


Figure 21: Experiments show that the best gradient values for the proposed CTGF filter is (1,32).

Describing devices fingerprinting for devices attribution

Method	Accuracy Statistics on Crossfolding 5x2 Experiments					
	Min	Mean	Max	σ	Mean-2σ	Mean+2σ
CTGF_MDMS_F	**98.97**	**99.23**	**99.49**	**0.18**	**98.86**	**99.60**
GLCM_MDMS_F	**97.10**	**98.38**	**99.32**	**0.72**	**96.95**	**99.81**
GLCM_MDMS_C	**96.63**	**97.60**	**98.99**	0.72	**96.15**	**99.05**
GLCM_MD_F	95.40	97.15	98.30	0.84	95.47	98.84
GLCM_MD_C	95.78	96.99	98.82	0.94	95.12	98.87
HOG_C [36]	94.42	95.79	96.79	0.83	94.14	97.45
LBP_F [33]	94.22	95.21	96.25	**0.59**	94.03	96.39
CTGF_3x3_F	93.17	94.44	96.59	1.03	92.38	96.50
GLCM_C [3,4]	92.24	94.19	96.45	1.36	91.47	96.91
CTGF_MDMS_D	88.70	93.70	96.63	2.07	89.56	97.84
GLCM_F [3,4]	91.82	93.62	95.23	1.13	91.36	95.89
LBP_C [33]	88.16	90.20	91.71	1.22	87.77	92.64
GLCM_MD_D	84.32	89.31	92.72	2.28	84.75	93.87
GLCM_MDMS_D	86.34	88.58	90.69	1.58	85.43	91.74
LBP_D [33]	86.17	88.08	90.19	1.50	85.08	91.07
CTGF_5x5_F	85.32	87.78	90.46	1.67	84.44	91.11
RECONST_ERROR_C [7]	81.79	84.87	87.82	2.09	80.69	89.04
CTGF_7x7_F	80.89	83.80	86.50	2.09	79.62	87.98
CTGF_3x3_D	82.29	83.78	85.11	0.78	82.23	85.34
GLCM_D [3,4]	78.75	82.57	87.31	***2.68***	77.21	87.93
CTGF_5x5_D	77.74	80.29	82.23	1.45	77.38	83.19
HOG_D [36]	78.00	79.66	81.90	1.37	76.92	82.41
CTGF_7x7_D	73.52	76.91	81.62	***2.45***	72.02	81.80
HOG_F [36]	72.74	74.36	75.47	0.81	72.74	75.97
NOISE_STATS_C [18]	67.29	68.87	70.66	1.10	66.68	71.06
NOISE_STATS_F [18]	40.48	42.27	43.76	1.12	40.03	44.51
NOISE_STATS_D [18]	37.27	39.82	42.13	1.50	36.82	42.82
DWT_STATS_D [22]	***33.00***	***36.57***	***39.93***	1.94	***32.68***	***40.46***
DWT_STATS_F [22]	***32.08***	***34.34***	***37.65***	1.81	***30.73***	***37.95***
DWT_STATS_C [22]	***24.96***	***28.70***	***33.56***	***3.19***	***22.32***	***35.07***

Legend:

xx.xx = Three best methods in the column metric

xx.xx = Three worst methods in the column metric

Table 2: Laser printer attribution experiments comparing our proposed technique against the state the art in characters (c), frames (f) and documents (d).

Describing devices fingerprinting for devices attribution

Method	Mean f-measure by Printer on Crossfolding 5x2 Experiments									
	B4070	C1150	C3240	C4370	H1518	H225A	H225B	LE260	OC330	SC315
	1	2	3	4	5	6	7	8	9	10
CTGF_MDMS_F	99.44	**99.63**	99.14	99.32	**99.83**	98.81	***98.81***	99.13	***98.73***	99.59
GLCM_MDMS_F	99.24	99.04	99.16	99.24	98.73	***94.22***	***94.45***	99.41	**100.00**	**99.92**
GLCM_MDMS_C	99.59	95.45	99.01	99.03	94.67	***94.36***	***94.19***	**99.68**	**100.00**	99.59
GLCM_MD_F	98.51	98.08	97.44	97.90	98.89	***90.22***	***90.70***	99.25	**100.00**	**99.92**
GLCM_MD_C	99.60	95.02	97.95	97.59	94.14	***92.69***	***92.92***	**100.00**	**100.00**	99.51
HOG_C [36]	95.24	92.63	97.44	98.28	93.74	***91.58***	***91.33***	97.41	**100.00**	**99.51**
LBP_F [33]	**100.00**	97.25	98.15	99.41	97.22	***82.06***	***77.46***	99.21	**100.00**	99.67
CTGF_3x3_F	97.85	96.51	89.59	91.90	95.35	***87.89***	***89.31***	96.23	**99.65**	**99.50**
GLCM_C [3,4]	97.90	***89.55***	90.54	94.94	93.35	***88.27***	90.69	96.26	**100.00**	**99.58**
CTGF_MDMS_D	95.33	***90.61***	91.13	92.20	**99.03**	91.41	***89.55***	94.09	96.31	**96.80**
GLCM_F [3,4]	97.23	87.65	91.95	96.13	95.32	***84.12***	***86.44***	97.58	**99.17**	**99.92**
LBP_C [33]	98.86	92.22	94.50	95.94	93.61	***73.84***	***49.39***	94.71	**99.83**	**99.43**
GLCM_MD_D	**95.85**	88.78	91.32	88.69	94.17	***76.51***	***75.85***	90.71	94.16	**95.54**
GLCM_MDMS_D	92.79	83.88	88.29	91.02	**93.83**	***76.51***	***78.11***	93.54	90.25	**96.32**
LBP_D [33]	92.82	87.24	87.87	90.23	93.96	***72.68***	***71.18***	91.02	**94.79**	**97.40**
CTGF_5x5_F	87.47	83.42	84.30	81.59	93.20	***77.90***	***78.37***	94.26	**97.13**	**98.81**
RECONST_ERROR_C [7]	87.43	90.75	90.34	92.74	92.47	***43.72***	***48.11***	95.18	**98.01**	**97.96**
CTGF_7x7_F	85.46	78.89	***69.58***	83.64	93.48	***71.14***	74.18	88.04	**96.48**	**97.42**
CTGF_3x3_D	86.51	81.04	80.09	***77.55***	85.81	80.64	***76.77***	81.30	**93.03**	**94.16**
GLCM_D [3,4]	**93.90**	73.33	81.89	76.77	92.81	***71.86***	***69.03***	85.95	85.82	**93.61**
CTGF_5x5_D	79.73	***73.37***	77.13	77.74	85.68	75.23	***73.89***	80.23	**86.92**	**92.57**
HOG_D [36]	85.41	71.43	81.59	81.14	**92.31**	***54.01***	***53.90***	89.78	91.79	**93.37**
CTGF_7x7_D	73.03	***67.62***	71.18	75.61	**87.38**	70.72	***67.70***	81.29	86.04	**91.22**
HOG_F [36]	77.57	64.18	71.90	68.28	**94.05**	***51.20***	***46.60***	86.87	**92.70**	86.09
NOISE_STATS_C [18]	***45.21***	54.60	***32.56***	57.71	92.88	69.81	48.99	72.49	**93.07**	**96.42**
NOISE_STATS_F [18]	55.04	18.94	***1.87***	38.65	**77.15**	***11.40***	40.39	18.47	35.56	**59.59**
NOISE_STATS_D [18]	38.01	27.25	51.94	38.93	**67.51**	26.01	31.02	***20.75***	***18.44***	**75.09**
DWT_STATS_D [22]	15.01	***12.27***	21.78	21.40	**92.60**	39.74	***10.23***	28.08	37.64	**53.93**
DWT_STATS_F [22]	19.32	15.16	***0.65***	14.68	**94.24**	34.31	***0.00***	42.53	15.58	**43.94**
DWT_STATS_C [22]	29.19	2.00	***0.00***	0.30	**93.82**	3.96	***0.00***	8.16	**56.24**	25.60

Legend:
xx.xx = Two best *f-measure* results of the method
xx.xx = Two worst f-measure results of the method

Table 3: Laser printer attribution experiments comparing f-measure per printer of our proposed technique against the state the art.

Describing devices fingerprinting for devices attribution

Method	CTGF_MDMS_F	GLCM_MDMS_F	GLCM_MDMS_C	GLCM_MD_F	GLCM_MD_C	LBP_F [33]	HOG_C [36]	CTGF_3x3_F	GLCM_C [3,4]	CTGF_MDMS_D	GLCM_F [3,4]	RECONST_ERROR_C [7]	CTGF_3x3_D	NOISE_STATS_C [18]	DWT_STATS_D [22]
CTGF_MDMS_F	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
GLCM_MDMS_F	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
GLCM_MDMS_C	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
GLCM_MD_F	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
GLCM_MD_C	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
LBP_F [33]	*-1*	*-1*	0	0	0	0	0	1	1	1	1	1	1	1	1
HOG_C [36]	*-1*	*-1*	*-1*	*-1*	0	0	0	0	0	0	0	1	1	1	1
CTGF_3x3_F	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	0	0	0	0	0	1	1	1	1
GLCM_C [3,4]	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	0	0	0	0	0	1	1	1	1
CTGF_MDMS_D	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	0	0	0	0	0	1	1	1	1
GLCM_F [3,4]	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	0	0	0	0	0	1	1	1	1
RECONST_ERROR_C [7]	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	0	0	0	1
CTGF_3x3_D	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	0	0	0	0
NOISE_STATS_C [18]	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	0	0	0	0
DWT_STATS_D [22]	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	*-1*	0	0	0

Legend:
1 = Row method is statistically better than column method
0 = Row method is statistically equivalent to column method
-1 = Row method is statistically worse than column method

Table 4: Laser printer attribution experiments comparing f-measures.

Describing devices fingerprinting for devices attribution

- Scanners






Describing devices fingerprinting for devices attribution

- Scanners
  - Scanner have almost similar function than cameras.
  - Their identification are based on sensor pattern noise [17-18] or sensor imperfections [19].

Describing devices fingerprinting for devices attribution

- Scanner attribution by Khanna et al [18]
  - Based on texture of noise pattern in row and column direction.
  - Two steps:
    - Reference Pattern construction for profiling
    - Noise Correlation for ballistics

Describing devices fingerprinting for devices attribution

- Scanner attribution by Khanna et al [18]
  - Reference Pattern Construction
    - Every scanned image I^k per printer has its noise extracted by subtracting it from the denoised image.
    - Wiener Filter is used.

Describing devices fingerprinting for devices attribution

- Scanner attribution by Khanna et al [18]
  - Reference Pattern Construction

$$I^k_{noise} = I^k - I^k_{denoised} \quad (6)$$

Describing devices fingerprinting for devices attribution

- Scanner attribution by Khanna et al [18]
  - Reference Pattern Construction
    - The 2D reference Pattern is the mean of each noise from K images.

$$\tilde{I}_{noise}^{array}(i,j) = \frac{1}{K}\sum_{k=1}^{k=K} I_{noise}^{k}(i,j) \quad (7)$$

Describing devices fingerprinting for devices attribution

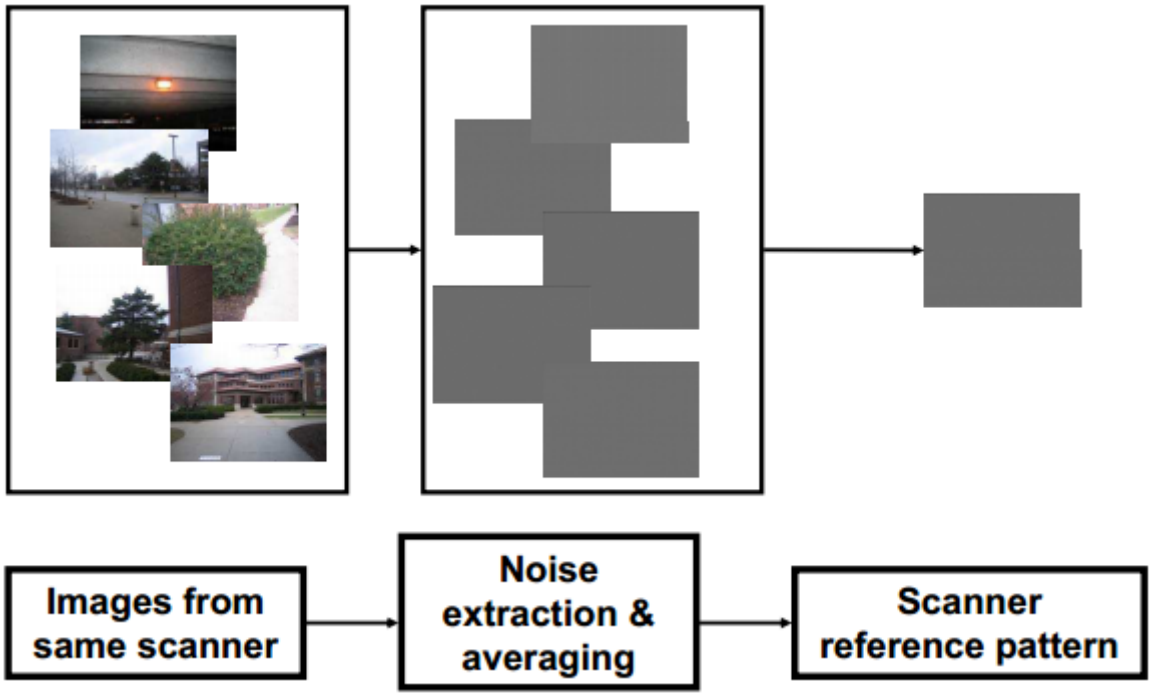


Figure 22: Reference pattern extracted from scanned images proposed by Khanna et al [18].

Describing devices fingerprinting for devices attribution

- Scanner attribution by Khanna et al [18]
  - Noise Correlation
    - The noise of a suspected document is compared to the reference pattern of x-th printer by correlation.
    - Higher correlation wins!

$$C(I_{noise}^{suspect}, \tilde{I_x}_{noise}^{array}) = \frac{(I_{noise}^{suspect} - \bar{I}_{noise}^{suspect}) \cdot (\tilde{I_x}_{noise}^{array} - \bar{\tilde{I}_x}^{array}{}_{noise})}{||I_{noise}^{suspect}|| - \bar{I}_{noise}^{suspect}|| \cdot ||\tilde{I_x}_{noise}^{array} - \bar{\tilde{I}_x}^{array}{}_{noise}||} \quad (8)$$

Describing devices fingerprinting for devices attribution

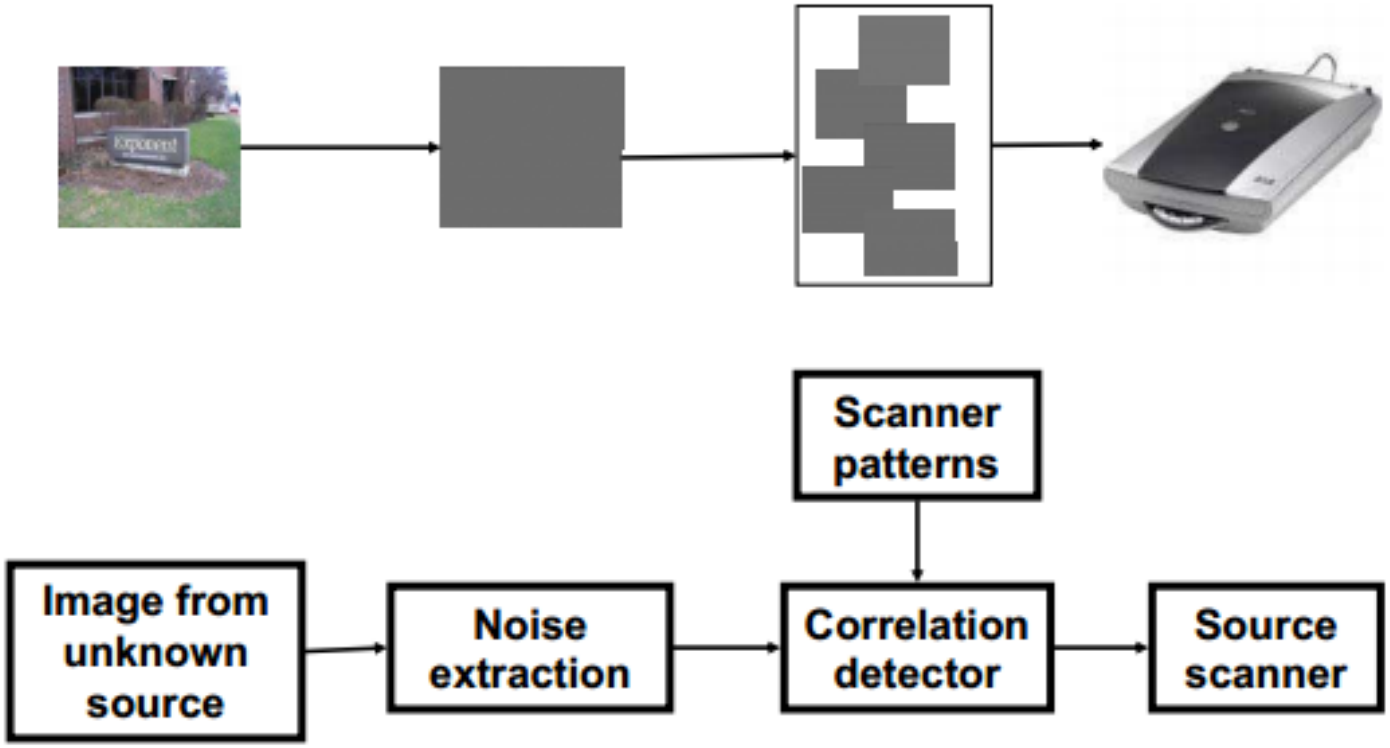


Figure 23: Correlation based scanner attribution by Khanna et al [18].

Describing devices fingerprinting for devices attribution

- Scanner attribution by Khanna et al [18]
  - Extension: assuming mean of lines and columns of noise as two vectors.

$$\tilde{I}^{l}_{noise}(1,i) = \frac{1}{N}\sum_{j=1}^{N} \tilde{I}^{array}_{noise}(i,j) \quad (9)$$

$$\tilde{I}^{c}_{noise}(1,j) = \frac{1}{M}\sum_{i=1}^{M} \tilde{I}^{array}_{noise}(i,j) \quad (10)$$

Describing devices fingerprinting for devices attribution

- Scanner attribution by Khanna et al [18]
  - A set of features are extracted from these vectors:
    1. Mean of $\tilde{I}^{l}_{noise}$ and $\tilde{I}^{c}_{noise}$.
    2. Correlation between each line of $\tilde{I}^{suspected}_{noise}$
    3. and $\tilde{I}^{l}_{noise}$
    4. Correlation between each column of $\tilde{I}^{suspected}_{noise}$ and $\tilde{I}^{c}_{noise}$
    5. Statistics over each vector.

Describing devices fingerprinting for devices attribution

- Scanner attribution by Gou et al [19]
  - Scanning noise taken from multiple perspectives:
    - image denoising
    - wavelet analysis,
    - neighborhood prediction
  - Statistical noise features is taken from each of them.

Describing devices fingerprinting for devices attribution

- Scanner attribution by Gou et al [19]
  1. Image denoising:
     - Noise extracted as in previous approach.
     - The noise suffers $\log_2$ image transform.
     - Two statistics are taken from it:
       - mean
       - standard deviation.

Describing devices fingerprinting for devices attribution

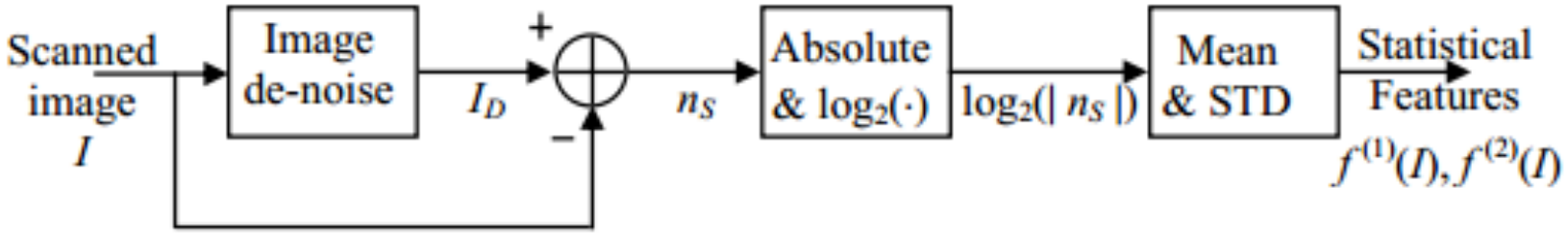


Figure 24: Image denoise approach for scanner attribution by Gou et al [19].

Describing devices fingerprinting for devices attribution

2. Wavelet Analysis
   - Done in a normalized version of the image.
   - Subbands HH, HL, LH are used for analysis.
   - Statistical measures are calculated:
     - the mean of coefficients.
     - standard deviation of coefficients.
     - goodness of Gaussian fitting.

Describing devices fingerprinting for devices attribution

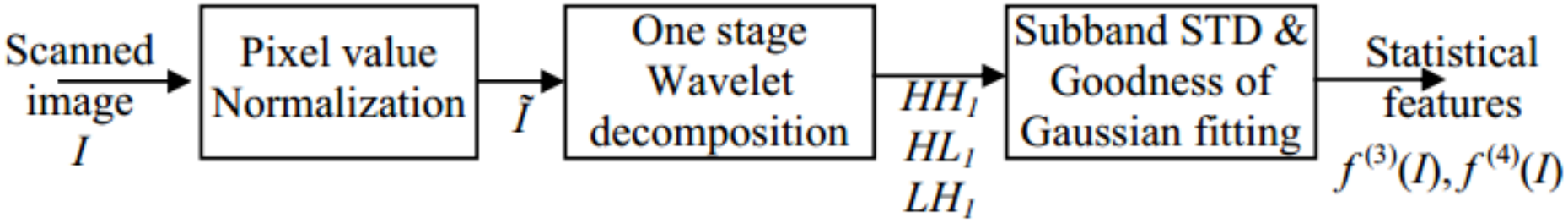


Figure 25: Wavelet statistical analysis approach for scanner attribution by Gou et al [19].

Describing devices fingerprinting for devices attribution

3. Neighborhood prediction
   - Scanned image must be first normalized
   - smooth areas are found by gradient and intensity thresholds.
   - At each region, its center pixel value is predicted using a linear model on its eight neighbors.

Describing devices fingerprinting for devices attribution

3. Neighborhood prediction
   - Absolute prediction errors are calculated
   - Mean and standard deviation as features
   - Done in each color channel
   - 30 + 18 + 12 = 60 statistical is extracted by the whole technique.

Describing devices fingerprinting for devices attribution

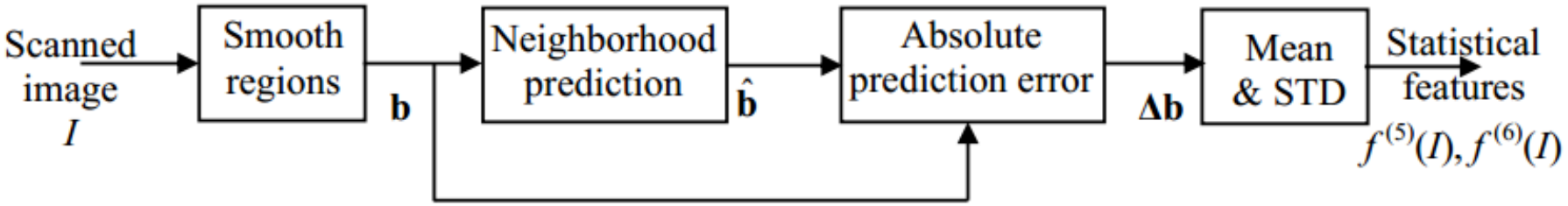


Figure 26: Neighborhood prediction approach for scanner attribution by Gou et al [19].

Describing devices fingerprinting for devices attribution

- Scanner Attribution by Dirik et al [20]
  - Uses traces of dust, dirt, and scratches over scanner platen on scanned images.
  - Two steps:
    - Dust and scratch reference construction per scanner
    - Source scanner identification

Describing devices fingerprinting for devices attribution

1. Dust and scratch reference construction
   - With ‘black scans’, just dust and scratch positions are detected. Two are enough.
   - Dust and scratches in two images are not aligned due to the vertical and horizontal scanner head position shifts.

Describing devices fingerprinting for devices attribution

1. Dust and scratch reference construction
   - Matching of images through cross correlation.
   - Scanner dust and scratch reference:
     - Hadamard product of the correctly aligned images.

Describing devices fingerprinting for devices attribution

2. Source scanner identification
   - Given a suspected document, its scratches and dust are detected.
   - They are then correlated with scanners templates as shown before.
   - High correlation wins!
   - But, how to find dust and scratches in a scanned image?

Describing devices fingerprinting for devices attribution

- How to find dust and scratches
  - High frequencies components of the image are found.
  - A model is searched through normalized cross correlation (NCC).
  - High NCC means dust and scratch locations.

Describing devices fingerprinting for devices attribution

Figure 27: Dust/scratch model for high pass filtered scanned image. Extracted from [20].

Describing devices fingerprinting for devices attribution

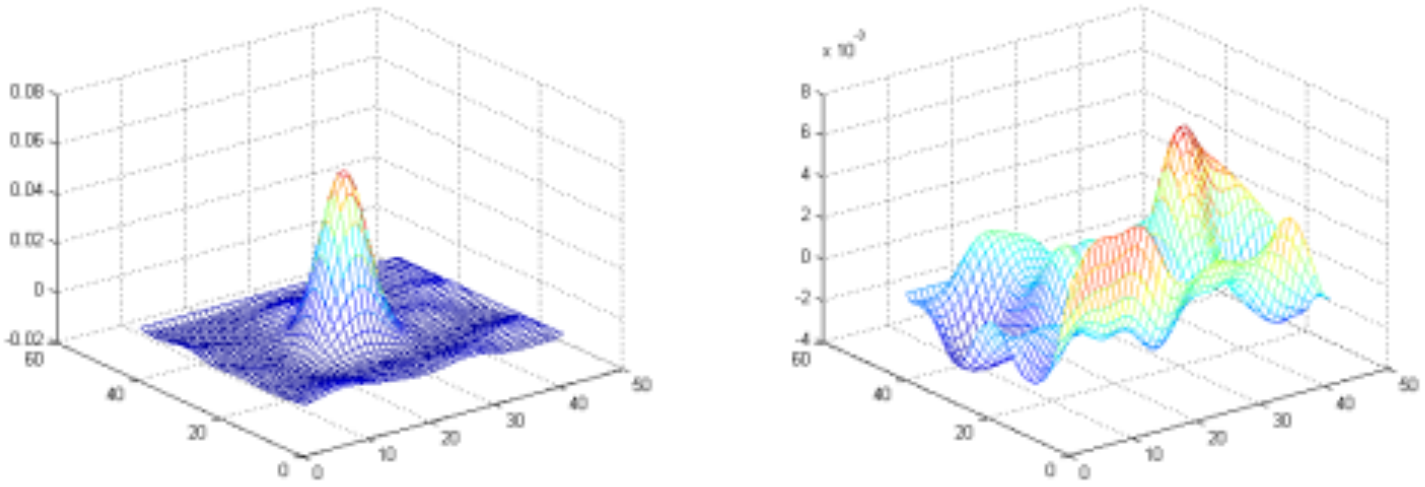


Figure 28: Cross-correlation results for scanner identification. (Left) he image dust/scratch positions are matched with the scannerdust/scratch template. (Right) there is no matching. Extracted from [20].

ANTI-FORENSICS

- Proposed techiques so far are vulnerable to attacks [20]:
  - Signature removal .
  - Signature replacement (spoofing).

ANTI-FORENSICS

- Printers
  - OCR + Another Printer: remove any extrinsic or intrinsic signature of laser printer in a text document.
  - Printed halftone images could be scanned, converted to continuoustone, and then reprinted after using standard watermark attacks to remove any embedded watermark.

ANTI-FORENSICS

- Scanners
  - A process called flatfielding can replace the sensor noise.
  - This operation includes the subtraction of light and dark patterns from the image.
  - Then, these same patterns can be replaced by others from another sensor.

Conclusion

- Device Attribution is a **hot** research field in forensics
  - Criminal investigations.
  - Documents Authentication.
- Machine Learning with Computer Vision play na important role.
- New device technologies are coming! how to deal with this? How to deal with anti-forensics?

Conclusion

- New devices (and challenges) are coming!

The terrifying reality of 3D-printed guns: Devices that ANYONE can make are quickly evolving into deadly weapons

- It is one year since the first 3D-printed gun was unveiled to the world
- Over the last 12 months the designs have become better and better
- Many enthusiasts across the globe have been showing off their designs
- MailOnline spoke to several users of 3D-printing gun site Fosscad
- Some suggested the guns could be on a par with real guns in a year or two

By JONATHAN O'CALLAGHAN

PUBLISHED: 13:31 GMT, 16 May 2014 | UPDATED: 16:42 GMT, 16 May 2014

Share 108 View comments

Almost a year ago in May 2013 the world's first gun made with a 3D printer was unveiled.

At the time it sparked major controversy – some derided it as nothing more than a toy, others warned it was a serious security risk that was undetectable by metal detectors.

Now, one year on, the chilling reality of 3D-printed guns has been revealed as enthusiasts across the world show off their 'toys'.

In May 2013 Cody Wilson showed off the world's first fully 3D-printed gun, known as the Liberator (pictured). Now, one year one, users of the online community Fosscad reveal to MailOnline how far the technology has progressed, and how dangerous it could become in the future

Figure 29: Which 3D printer printed this gun? extracted from [21].

Conclusion

- New devices (and challenges) are coming!

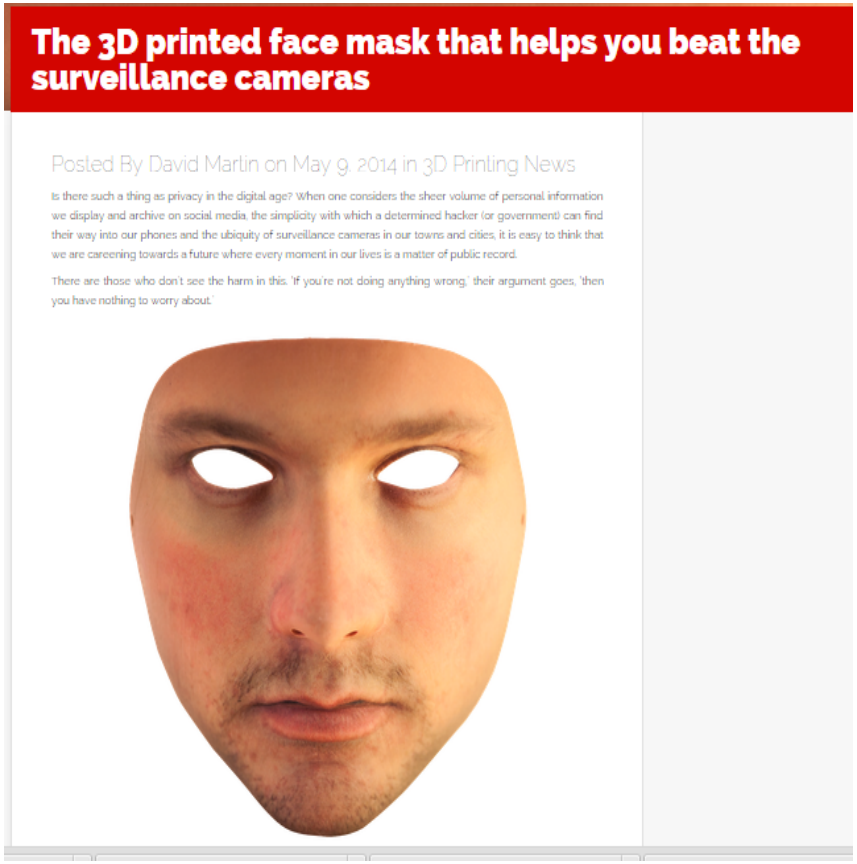


Figure 30: Which 3D printer printed this mask? extracted from [22].

Conclusion

- New devices (and challenges) are coming!

Secretly Record Your Friends and Enemies With Handy Pocket Camera Drone

Because camera phones weren't enough.

BY MOLLY MULSHINE | 1/21 11:30AM

Like 1k Tweet 227 Share 22 Compartilhar 22 Email More

That? Oh, that's just a mosquito, now tell me more about your top secret project.

Do you ever feel frustrated that you can't keep tabs on your significant other at all times? Nervous that everyone is hanging out without you? Curious about what the heck your neighbors are doing over there?

Thankfully, there's a Pocket Drone currently being funded on Kickstarter that will solve these problems and more. It only takes 20 seconds to unpack and launch. Then, you

Figure 31: Which photos this drone took? Extracted from[23].

References

[1] Phoebe Greenwood. **Israeli soldier posts Instagram image of Palestinian child in crosshairs of rifle.** Available at http://www.theguardian.com/world/ 2013/feb/18/israeli-soldier-posts-instagram -palestinian. Access in September, 13, 2014.

[2] Sarah Zhang. **It's Surprisingly Easy to Print Fake Money on an Inkjet Printer.** Available at http://gizmodo.com/its-surprisingly-easy-to-print-fake-money-on-an-inkjet-1573134734. Access in September, 13, 2014.

[3] J.D. Gallop. **Cocoa man charged with child porn. Available at http://www.floridatoday.com/story/news/crime/2014/05/23/cocoa-man-char ged-with-child-porn/9509521/.** Access in September, 13, 2014.

[4] Anderson Rocha. **Aula 1**. Available at http://www.ic.unicamp.br/~rocha/ teaching/2012s2/mo447/aulas/aula-01-analise-forense-documentos.pdf. Access in September, 13, 2014

References

[5] G. N. Ali, P.-J. Chiang, A. K. Mikkilineni, J. P. Allebach, G. T.-C. Chiu, and E. J. Delp, **Intrinsic and extrinsic signatures for information hiding and secure printing with electropho tographic devices**, in Proc. IS&T's NIP19: Int. Conf. Digital Printing Technologies, New Orleans, LA, Sept. 2003, vol. 19, pp. 511–515.

[6] O. Bulan, J. Mao, G. Sharma, **Geometric distortion signatures for printer identification**, in: Intl. Conference on Acoustics, Speech and Signal Processing (ICASSP), 2009, pp. 1401-1404.

[7] Y. Wu, X. Kong, X. You, Y. Guo, **Printer forensics based on page document's geometric distortion**, in: Intl. Conference on Image Processing (ICIP), 2009,pp.2909-2912.

References

[8] G. N. Ali, P. ju Chiang, A. K. Mikkilineni, G. T. Chiu, E. J. Delp, J. P. Allebach, **Application of principal components analysis and gaussian mixture models to printer identification**, in: Intl. Conference on Digital Printing Technologies, 2004, pp. 301-305.

[9] A. K. Mikkilineni, P. ju Chiang, G. N. Ali, G. T.-C. Chiu, J. P. Allebach, E. J. Delp, **Printer identication based on graylevel co-occurrence features for security and forensic applications**, in: Intl. Conference on Security, Steganography, and Watermarking of Multimedia Contents, 2005, pp. 430-440.

[10] Eric Kee and Hany Farid. 2008. **Printer profiling for forensics and ballistics**. In *Proceedings of the 10th ACM workshop on Multimedia and security* (MM\&\#38;Sec '08). ACM, New York, NY, USA, 3-10.

[11] A. K. Mikkilineni, P. ju Chiang, G. N. Ali, G. T. c. Chiu, J. P. Allebach, E. J.Delp, **Printer identification based on textural feature**s, in: Intl. Conference on Digital Printing Technologies, 2004, pp. 306-311.

References

[12] A. K. Mikkilineni, O. Arslan, P. ju Chiang, R. M. Kumontoy, J. P. Allebach, G. T. c, **Printer forensics using svm techniques**, in: Intl. Conference on Digital Printing Technologies, 2005, pp. 223-226.

[13] M.-J. Tsai, J.-S. Yin, I. Yuadi, J. Liu, **Digital forensics of printed source identification for chinese characters**, Multimedia Tools and Applications (2013) pp.1-27.

[14] A. K. Mikkilineni, N. Khanna, E. J. Delp**, Forensic printer detection using intrinsic signatures**, in: Intl. Society for Optics and Photonics (SPIE), Vol. 7880, 2011, pp. 78800R7-8800R11

[15] H.-Y. Lee, J.-H. Choi, **Identifying color laser printer using noisy feature and support vector machine**, in: Intl. Conference on Ubiquitous Information Technologies and Applications, 2010, pp. 1-6.

References

[16] J.-H. Choi, H.-K. Lee, H.-Y. Lee, Y.-H. Suh, **Color laser printer forensics with noise texture analysis**, in: ACM Workshop on Multimedia and Security, 2010, pp.19-24.

[17] S. Elkasrawi, F. Shafait, **Printer identication using supervised learning for document forgery detection**, in: Intl. Workshop on Document Analysis Systems, 2014, pp. 146-150.

[18] Khanna, N.; Mikkilineni, AK.; Delp, E.J., **Scanner Identification Using Feature-Based Processing and Analysis**, *IEEE Transactions on Information Forensics and Security*,, vol.4, no.1, pp.123-139, March 2009.

[19] Gou, H., Swaminathan, A. and Wu, M. **Robust Scanner Identification Based on Noise Features**. Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents. San Jose, USA.

References

[20] Pei-Ju Chiang; Khanna, N.; Mikkilineni, AK.; Segovia, M.V.O.; Sungjoo Suh; Allebach, J.P.; Chiu, G.T.-C.; Delp, E.J., **Printer and scanner forensics**, *Signal Processing Magazine, IEEE* , vol.26, no.2, pp.72,83, March 2009

[21] CALLAGHAN, J. **The terrifying reality of 3D-printed guns: Devices that ANYONE can make are quickly evolving into deadly weapons**. Available at http://www.dailymail.co.uk/sciencetech/article-2630473/The-terrifying-reality-3D-printed-guns-Devices-ANYONE-make-quickly-evolving-deadly-weapons.html. Access in September, 17, 2014

[22] Martin, D. **The 3D printed face mask that helps you beat the surveillance cameras.** Available at http://www.top43dprinting.com/the-3d-printed-face-mask-that-helps-you-beat-the-surveillance-cameras/. Access in September, 17, 2014.

[23] Mulshine, M. Secretly Record Your Friends and Enemies With Handy Pocket Camera Drone. Available at http://betabeat.com/2014/01/secretly-record-your-friends-and-enemies-with-handy-pocket-camera-drone/. Access in September, 17, 2014.