DIGITAL FORENSICS

Manoranjan Mohanty

August 19, 2019

COMPSCI 725

SYSTEMS SECURITY

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

THE PLAN

- Quick overview
- Image forensics
  - A branch of digital forensics
  - Discuss some image forensics techniques

WHAT IS FORENSICS?

Police

Murder scene

Mr. X

- Police believes Mr. X committed the crime. Court agrees?
- Court needs evidence
- Forensics: Providing evidence in a way acceptable to the court – scientific evidence

"Application of science to solve a legal problem"
– John Sammons

WHAT THE COURT REQUIRES

- Evidence Authenticity and Integrity
  - Chain of custody: Detailed record of who has handled evidence and in what way
  - Witness
  - Able to prove that the original evidence from the crime scene has not been modified (either intentionally or non-intentionally)
  - Handled by professionals
  - Crime scene reconstruction can be required

FORENSICS LIFE CYCLE

- Identification of evidence source
- Secure collection of evidence
  - Without disturbing the crime scene
- Storage of evidence without modification
- Analysis of evidence for reaching conclusion
- Presentation of evidentiary conclusion in the court (in a layman way)

WHAT IS DIGITAL FORENSICS?

- Dealing with evidence form crime involving digital source (device and/or data)
  - Crime involving digital device/data:
    - Sending threatening email
    - Accessing child pornography
    - Hacking
    - Stealing business data
    - etc.

DIGITAL FORENSICS DEFINATION

"The use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitating or furthering the reconstruction of events found to be criminal, **or helping to anticipate unauthorized actions shown to be disruptive to planned operations**."

- Digital Forensic Research Workshop (DFRWS)

DIGITAL FORENSICS GOALS

- To prove a crime
- To prevent a crime – by anticipating crime through (live) analysis

Community	Primary Objective	Secondary Objective	Environment
Law Enforcement	Prosecution		After the fact
Military	Continuity of Operations	Prosecution	Real Time
Business and Industry	Availability of Services	Prosecution	Real Time

DIGITAL FORENSICS TYPES

Computer forensics

Network forensics

Mobile forensics

Cloud forensics

Memory forensics

Multimedia forensics

IDENTIFICATION OF THE EVIDENCE SOURCE

- “Who” can identify (and collect) depends on “what” can be taken to the forensics lab
  - Common practice: Take all the digital device(s) and related accessories from crime scene to the lab
    - Can be done by a police officer having basic knowledge about the devices (e.g., this is a computer, this is a flash drive, this is a smartphone)
  - For live analysis / much bigger crime scene (e.g., a company), a professional is required for identifying the source

DIGITAL EVIDENCE COLLECTION

- "Old school" method in evidence collection
  - *Pull the plug*

    - Remove power to the computer and shut down
    - Disconnect the device from the network
    - Seal the device so that it is not disturbed by electricity and network signal
    - Document everything and send the device to the lab
  - Clone the device: Create a "true copy" of the device
    - Use court-certified hardware and software
    - Bit by bit copy: Use cryptographic hash to validate
    - Document everything
    - Store the original device. From now on, use the clone

EVIDENCE COLLECTION CONT.

- Collect as much evidence as possible from the clone using court-certified tools
  - Recover deleted files
    - We will discuss how a deleted JPEG image can be recovered
  - Look for evidence in non-conventional places where normal user may not be aware of
    - Windows hibernation file
    - Internet history
    - Image header etc.
  - Document every bit of operation
    - The court will need to know why a key was pressed

EVIDENCE COLLECTION DILEMMA

- The “big” question: Pull the plug (shut down) or not to pull the plug?
  - Pulling the plug can shut down the device with minimal user (forensics expert) interaction
  - However, by pulling the plug, some devices can be corrupted due to improper shut down
    - Proper shutting down requires interaction with the system – a problem if done by a non-professional (court has to be convinced)
  - By shutting down the system, some evidence (e.g., in the RAM) will be lost
  - Whole disk encryption can be enforced when a shut down system is restarted
  - etc.

EVIDENCE COLLECTION CONT.

- So pull the plug (shut down) or not to pull the plug?
  - No clear policy
  - Forensic expert can take a call, but the action should be documented with the minute details
  - It is up to the court to decide if a well-knowledgeable person took the right decision for collecting evidence in the best possible way with modifying the device as little as possible

- Evidence identification and collection becoming harder
  - Cloud, Edge, IoT, Blockchain
  - Real time
  - Some of the papers in the reading list deals with this

ANALYSIS OF EVIDENCE

- Must be done using court-certified hardware and software
- Different types of analysis required for different objectives
  - For finding a virus: Anti virus software
  - For identifying a crime image: Crime image detection tool
  - For finding source camera of a crime image: Source camera attribution tool
- Everything must be documented

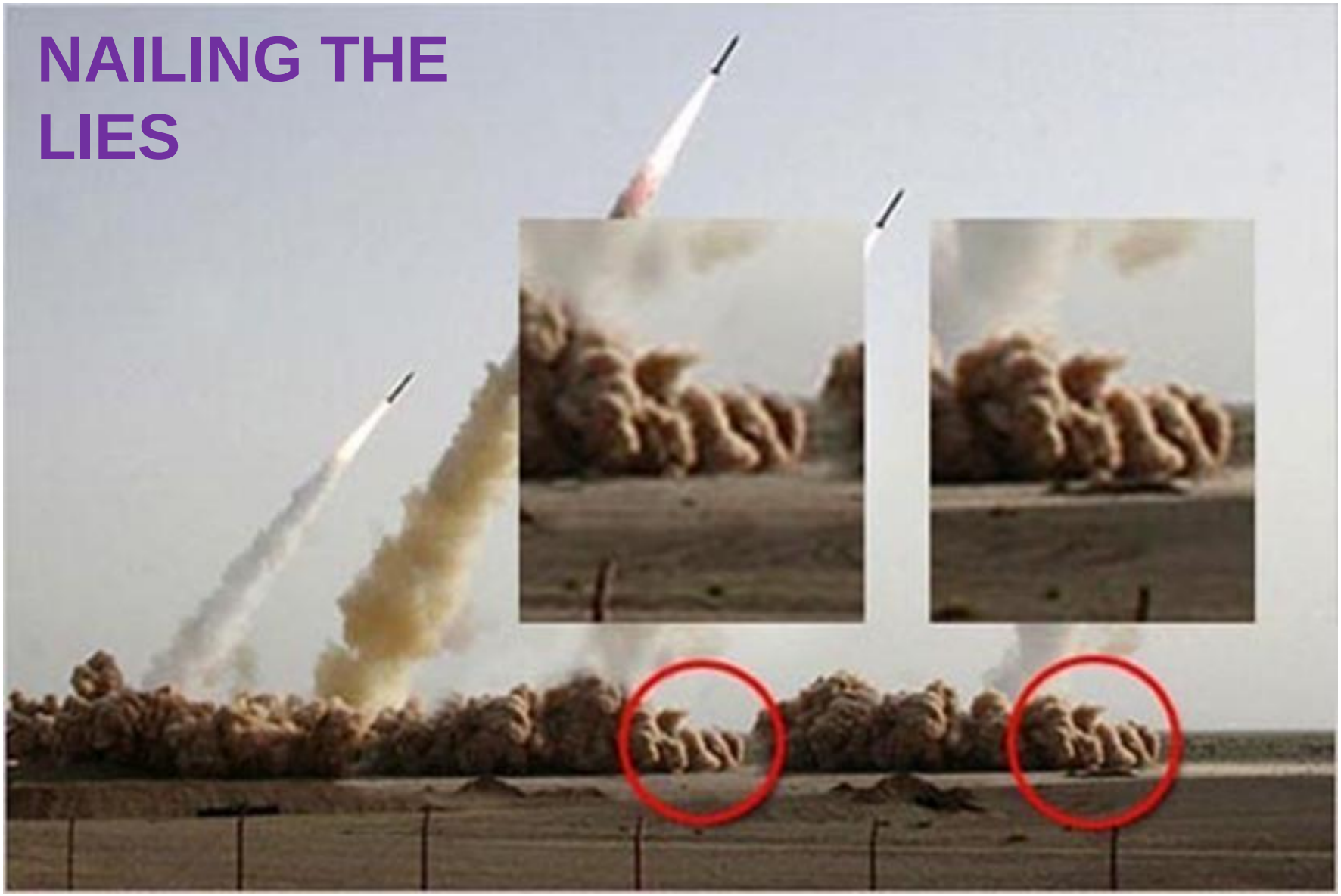
IMAGE FORENSICS

- Computer forensics
- Mobile forensics
- Network forensics
- Cloud forensics
- Memory forensics
- Multimedia forensics
  - Image forensics
  - Video forensics
  - Audio Forensics

IMAGE

IMAGE

IMAGE FORENSICS

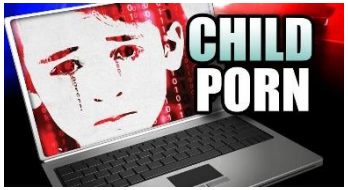


DEEP FAKE

IMAGE FORENSICS – THE PLAN

- Discuss some software techniques
  - Evidence collection
    - Recovering/carving deleted JPEG images
  - Evidence analysis
    - Detecting a crime (child porn) image
    - Finding source camera of an image

A SCENARIO



- Police believes that suspect X has photographed and stored child porn images

Suspect X

- Only a laptop was seized from the suspect
  - There is no image in the laptop – Images might have been deleted
    - How to recover/carve deleted JPEG images?
  - There are thousands of images
    - How to automatically detect child porn images?
  - No camera was found
    - How to verify if suspect's camera has captured the crime image without accessing the camera?

JPEG CARVING

Finding deleted JPEG images from a secondary drive

HOW WE DELETE AN IMAGE ?

- Recycle bin

  DEL

  – OS allows recovery

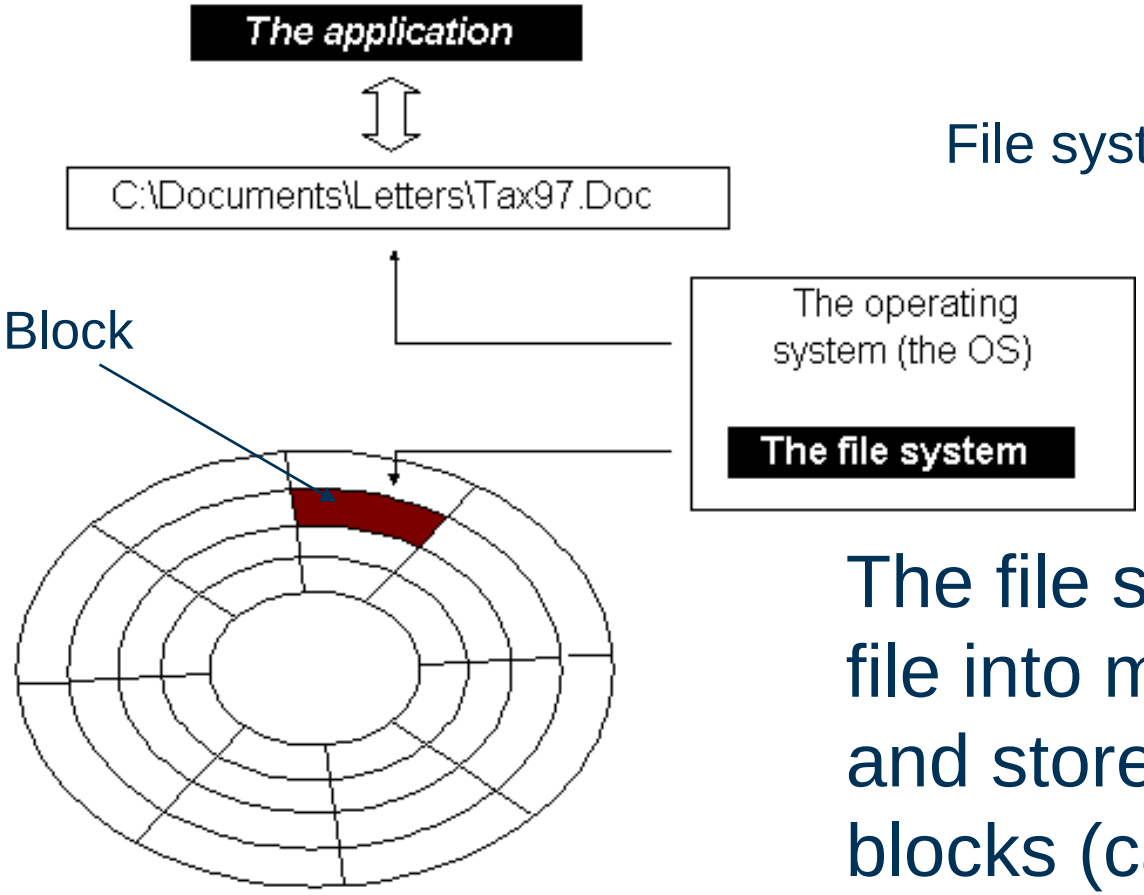
- “Permanent” deletion

  SHIFT + DEL

  – Emptying recycle bin
  – OS does not allow recovery
  – By deletion, we will mean “Permanent” deletion

HOW OS DELETES A FILE (IMAGE) ?

- How files are stored and handled?

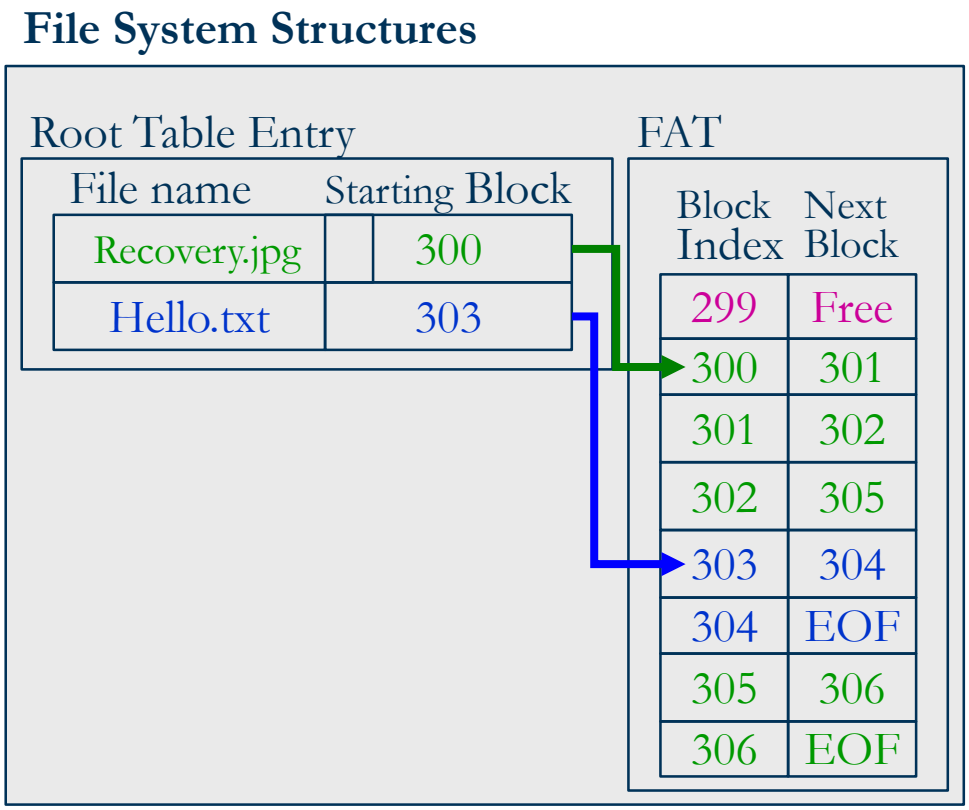


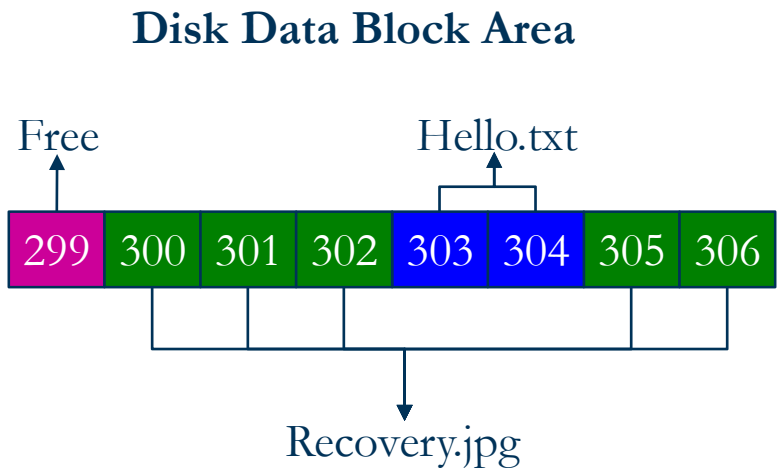
http://www.karbosguide.com/hardware/module6a1.htm

File system: FAT, NTFS, etc.

The file system divides a file into multiple parts, and stores the parts in file blocks (can be non-contiguous)

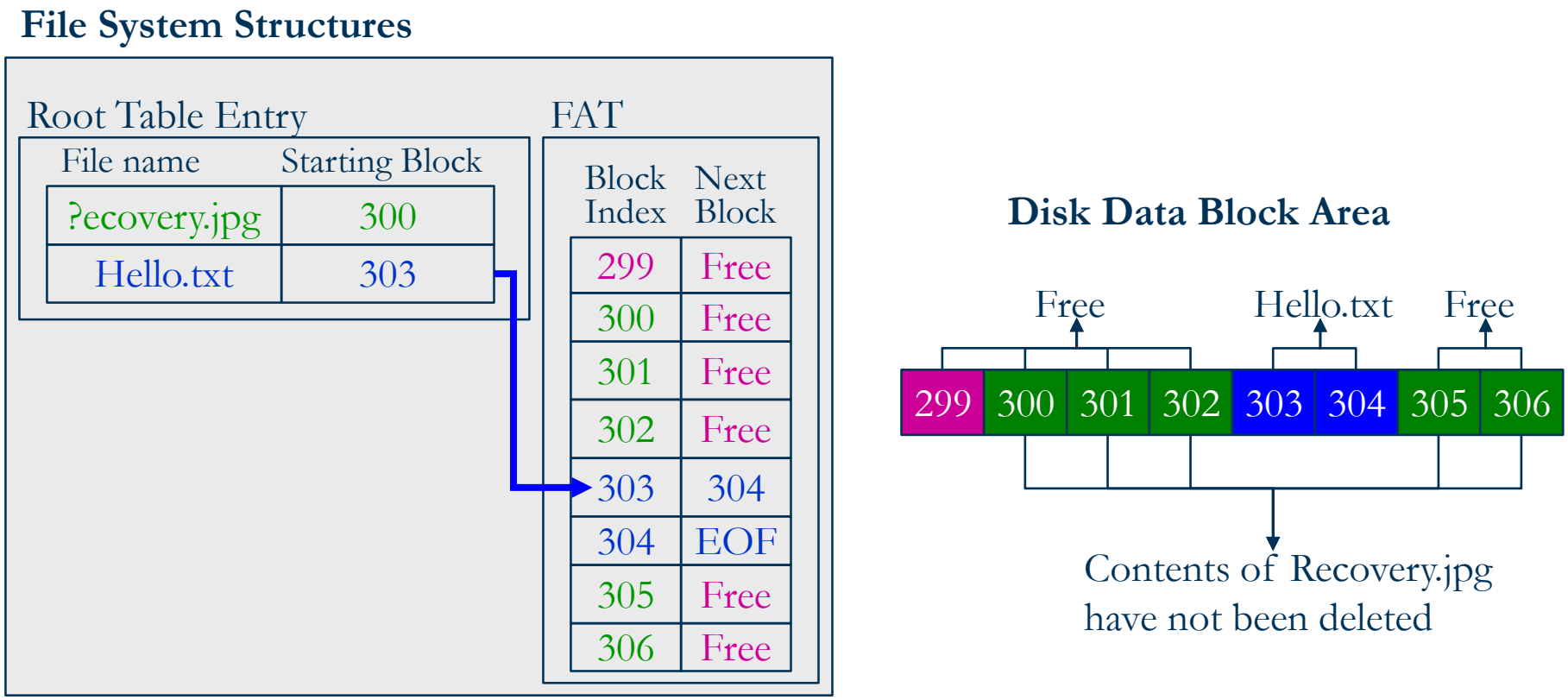
FILE ALLOCATION IN FAT EXAMPLE





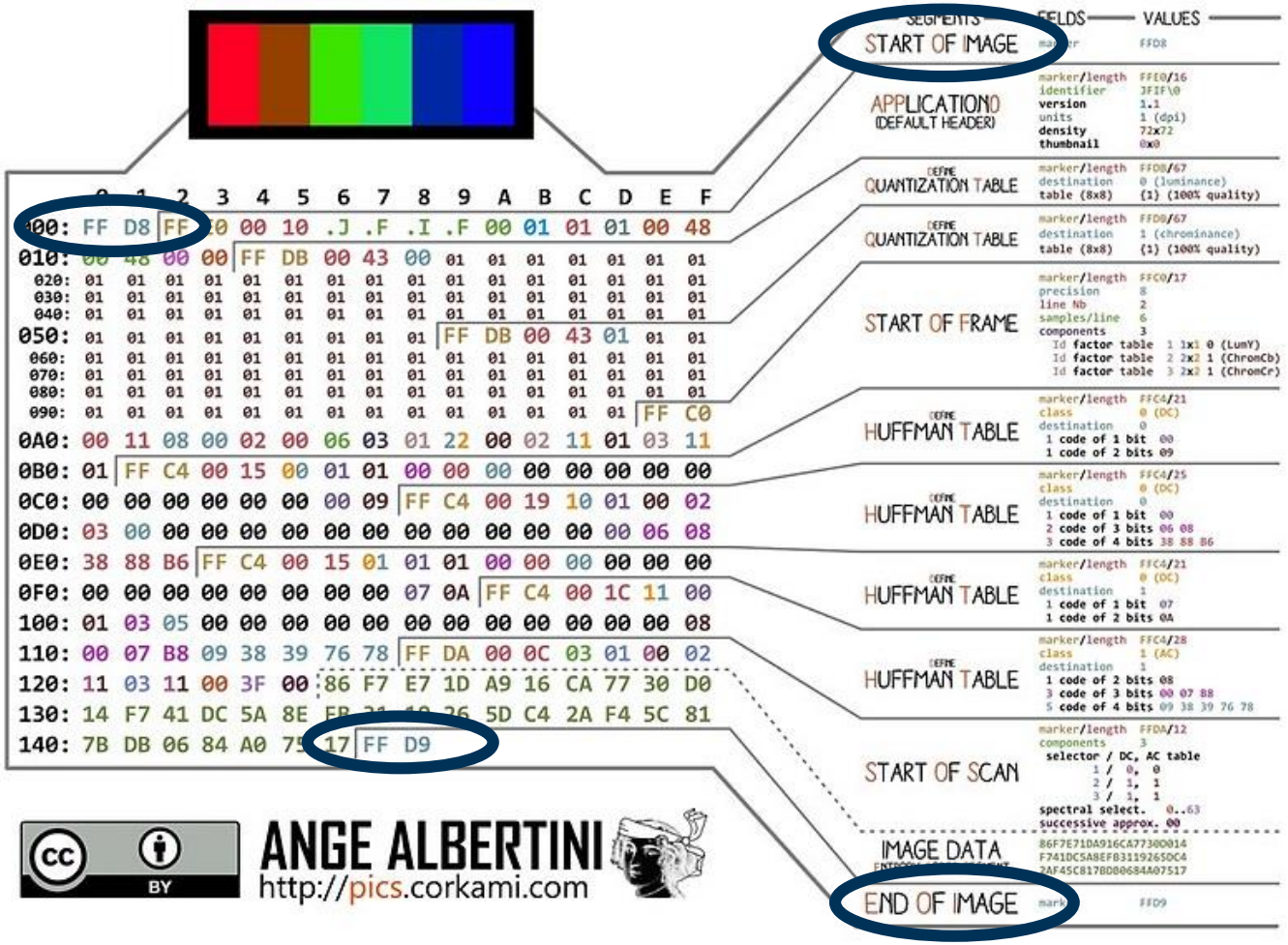
The file system keeps a record of which file-part goes to what disk-block.

FILE DELETION IN FAT EXAMPLE



- File deletion only marks the blocks as free (data still sitting on the disk)
- Carving: Collecting deleted (by file system) and non-overwritten blocks, and construct a file

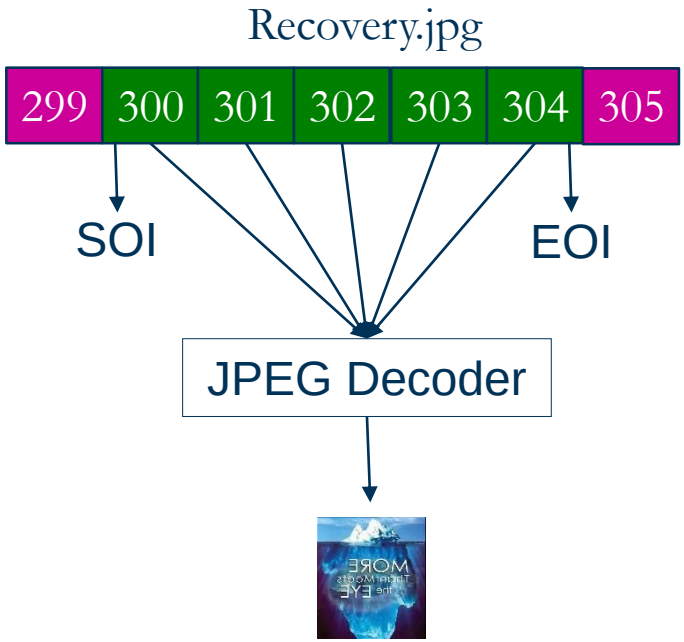
JPEG FILE FORMAT



HEADER

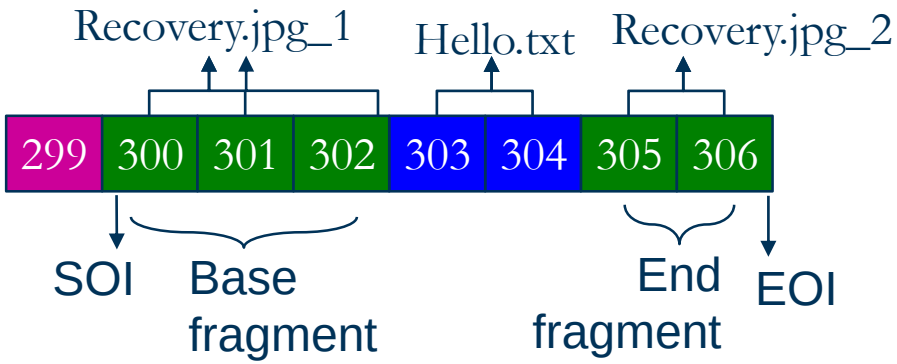
JPEG CARVING

- File is stored in contiguous disk blocks



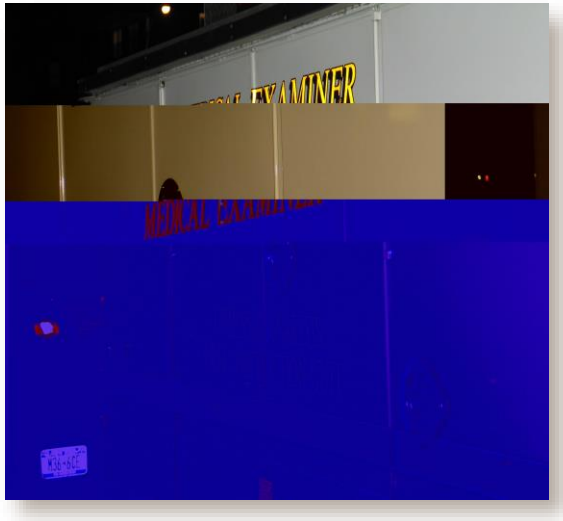
JPEG CARVING CONT.

- Things may not be that simple: Files stored in multiple fragments

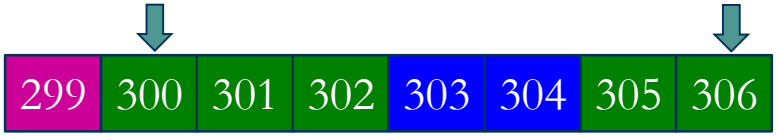


Two fragments

- Solution 1: Carve n blocks starting with the SOI block
  - Problem: Partial image



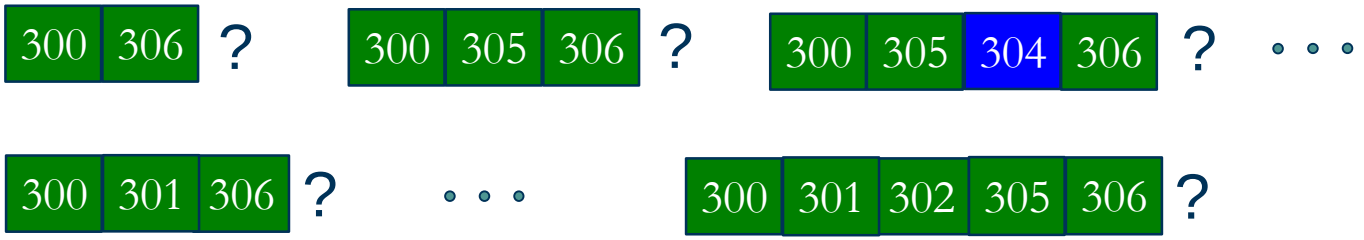
JPEG CARVING CONT.



- Identify SOI and EOI blocks (FFD8, FFD9)
- Try carving all possibilities between them
- Bi-fragment gap carving
  - Garfinkel S. *Carving contiguous and fragmented files with fast object validation*. In Proceedings of the 2007 digital forensics research workshop, DFRWS, Pittsburgh, PA; August 2007.

BI-FRAGMENT GAP CARVING

- Works for two fragments
  - Start of first fragment (SOI) and end of second fragment (EOI) known
  - Trying all possibilities for combining blocks between header and footer blocks
  - Validating a combination using validator (e.g., JPEG decoder)



What is the problem with this approach?

CARVING CAN BE DIFFICULT

- More than one fragments from one JPEG file
- The “gap” between the fragments can be large
- Identify and decode the fragments
  - Use restart markers: **FF D0 – FF D7**
  - What else?
  - Decode and validate each fragment using header information
  - How to reassemble the partially decoded image fragments efficiently ?
  - Reassembly can be difficult with more images and random order fragments

REASSEMBLY

Top Fragment

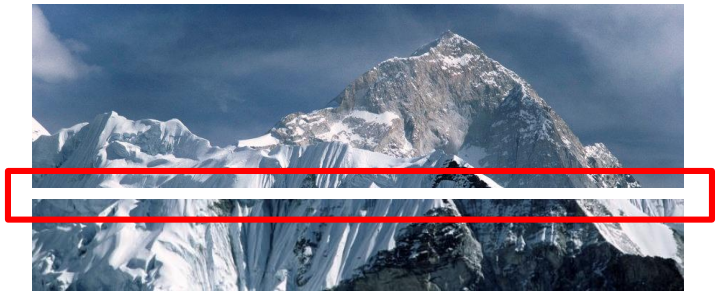
Bottom Fragment 1

Bottom Fragment 2

- Which bottom fragment goes with Top Fragment?
- How do you know?
- How your computer will know?
  - A popular method: Sum of squared difference (SSD)

SSD

I



J

- Sum of squared differences between boundary pixels
- For m, n: $(I_{m,n} - J_{m,n})$^2

- The result should be less than a threshold for two fragments to be put into one image

REASSEMBLY

Fragment 1

Fragment 2

Fragment 3

- The problem can be difficult
  - Some papers in the *reading list*

- Other methods are also possible
  - For example, deeplearning method can also be useful

CARVING CAN BE MORE DIFFICULT

- More than one image to be carved
- Missing header
- Missing data (image body) blocks
- Many fragments stored randomly on the disk
  - SSD drive
  - Flash drive
  - This is an active area of research

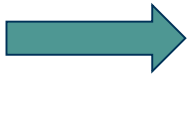
- Some papers given as resources towards the end of this deck of slides

PRNU-BASED CAMERA ATTRIBUTION

Your camera has also a fingerprint!

SOURCE CAMERA ATTRIBUTION

- Has this camera captured that image?


?

  - Can be any camera: DSLR, compact, smartphone
  - Without accessing the camera physically
- A number of approaches
  - Header information
  - Camera model identification
  - Photo Response Non-Uniformity Noise (We will discuss this)

IMAGE NOISE

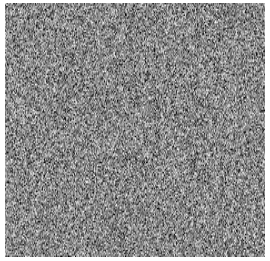
Left

Right

Which image is better?

- The left image is better because it has less “noise”
- Noise: Unwanted signal (information)
- Each image (we see) has “some” noise

IMAGE NOISE CONT


What is noise and what is signal?

- “noise” is subjective
- A “filter” to separate signal from the noise
- Noise can be of different types => Different types of filters can be required

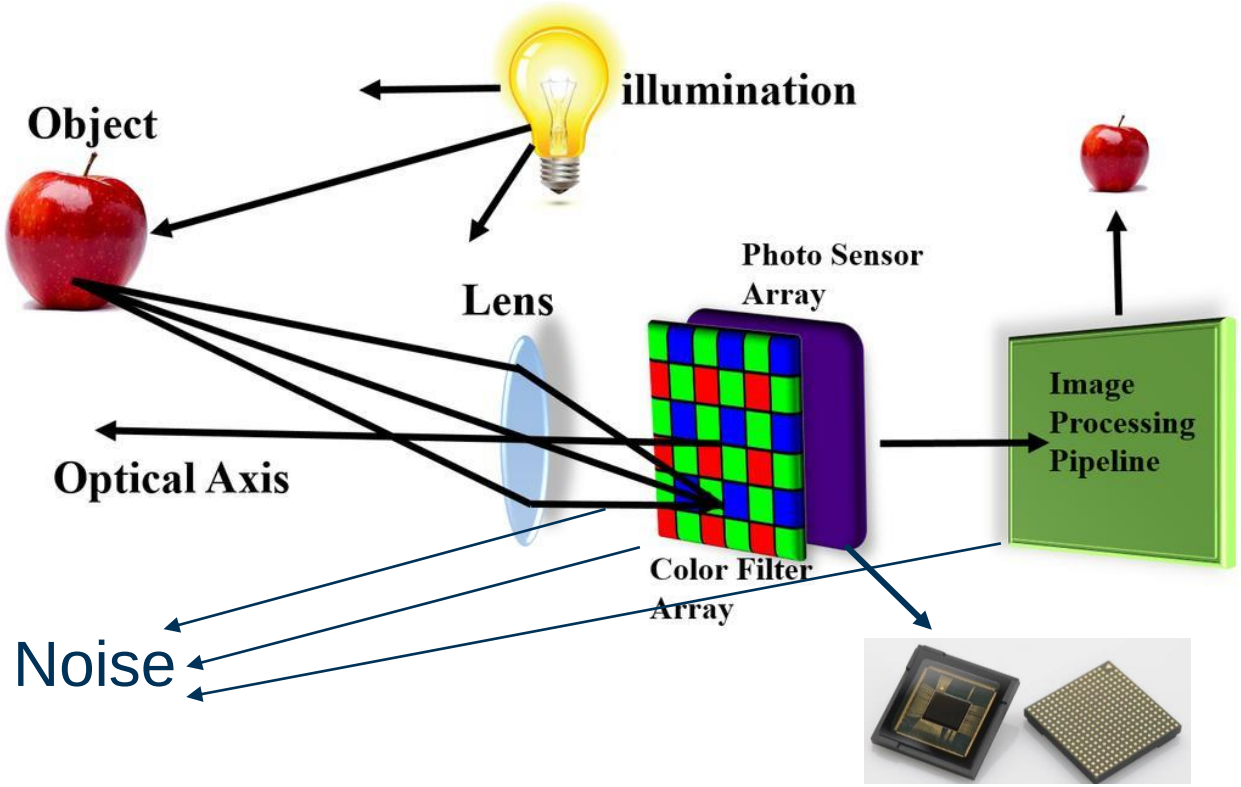

Sea water


Salt

The “filter” to remove water is different than the “filter” to remove other minerals

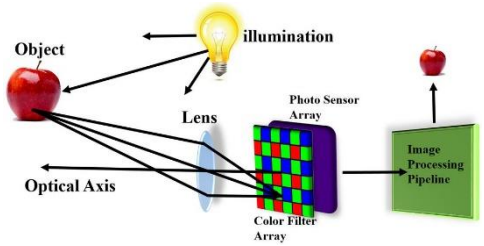
IMAGE NOISE CONT

- Filter depends on type of noise (source of noise)
- Types of noise can depend on the specific stage at the image capturing pipeline
- Typical image capturing pipeline

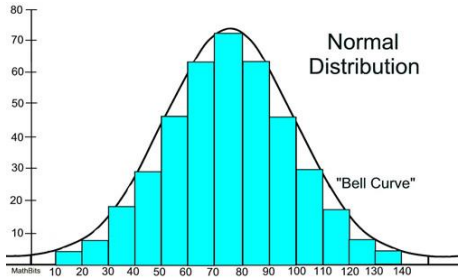


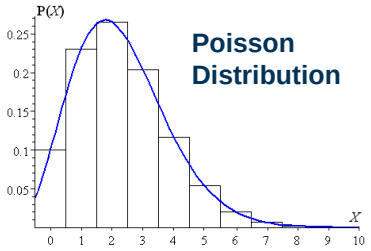
http://hobbydocbox.com/Photography/67994933-Three-dimensional-moving-pictures-with-a-single-imager-and-microfluidic-lens.html

IMAGE NOISE CONT



- Understanding noise
  - Some noise are added while some are multiplied with the signal
  - Can be modelled as probabilistic distribution (Gaussian, Poisson, etc)

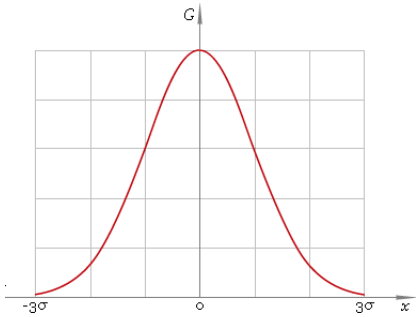




- Designing a filter for additive Gussian noise

$$S = S^0 + \phi$$

Obtained signal — Noise free signal — Additive noise

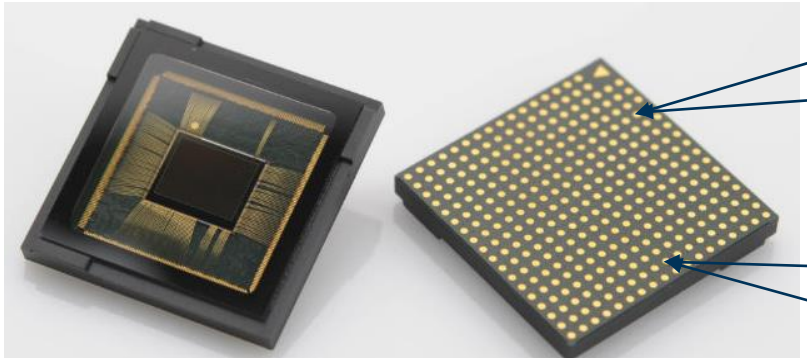


$$\dots S^0 + \phi_1, S^0 - \phi_2, S^0 - \phi_2, S^0 + \phi_2 \dots$$

Average filter?

PRNU NOISE

- Due to the imperfection with which the sensor pixels (semiconductor) are manufactured



- Each camera has this manufacturing error
- The noise (due to the error) is unique to the camera sensor
  - Unique to a camera irrespective of the camera model
- Also known as “Camera Fingerprint”

PRNU WITHOUT ACCESSING CAMERA

- Each image has a “degraded” version of the noise
- A multiplicative Gussian noise

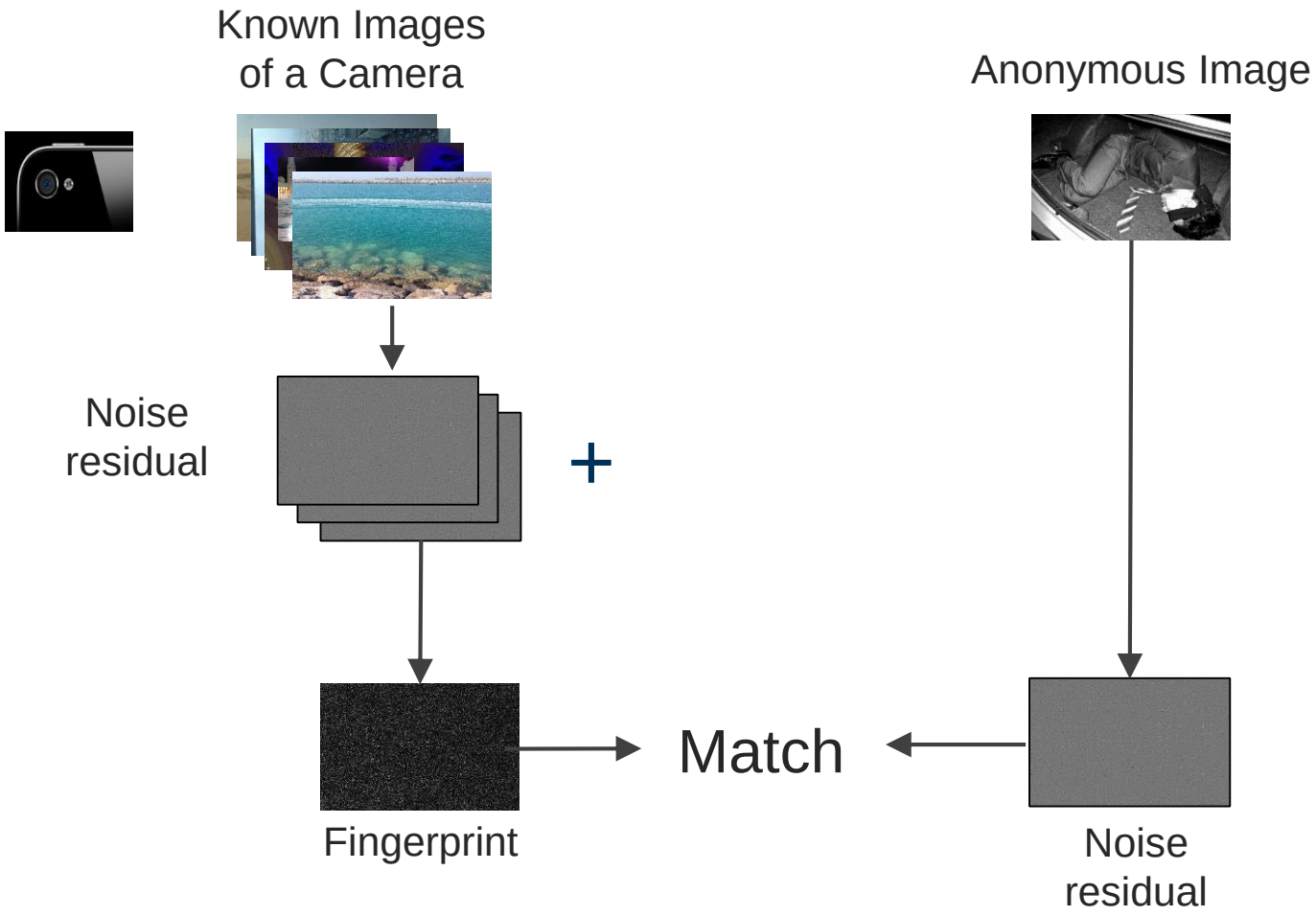
$$I = I^0 + I^0K + \psi$$

Obtained image — Noise free image — PRNU — Other noise

Can be filtered using Wavelet Filter, BM3D Filter, etc.

- Great for source camera attribution
  - Check if two images have the same PRNU noise
  - The PRNU noise from an image may not be of good quality (due to presence of other noise)
  - Better to consider a number of images

CAMERA ATTRUBUTION USING PRNU



CAMERA ATTRUBUTION USING PRNU CONT

- How to find if K' matches with W' ?

- $K' \in R^{M \times N}$ $W' \in R^{M \times N}$
- $K' =? W'$ feasible?
- Pearson Correlation Coefficient

$$r_{K'W'} = \frac{\sum_{i=1}^{M \times N}(K_i' - \overline{K'})(W_i' - \overline{W'})}{\sqrt{\sum_{i=1}^{M \times N}\left(K_i' - \overline{K'}\right)^2 \sum_{i=1}^{M \times N}\left(W_i' - \overline{W'}\right)^2}}$$

- If $r_{K'W'} \geq \tau$, where τ is threshold, then K' matches with W'
- Limitation: τ is camera dependent

CAMERA ATTRUBUTION USING PRNU CONT

- Other matching methods
  - Peak-to-Correlation Energy (PCE) - can provide a universal threshold for all the cameras
  - Normalized Cross Correlation (NCC) – useful when the image is cropped

PROPERTIES OF PRNU

- Dimensionality: The matrix K appears random and unique to each sensor.
- Universality: All imaging sensors exhibit PRNU.
- Generality: Present in every picture independently of settings, or scene content,
- Stability: The factor K is stable in time and under wide range of environmental conditions (temperature, humidity).
- Robustness: PRNU survives lossy compression, filtering, gamma correction etc

USE OF PRNU IN FORENSICS

- Has been rigorously tested and shown to uniquely identify cameras
- Withstands all kinds of image processing including compression, blurring, scaling, cropping etc.
- Given the camera accuracy high (high TPR, low FPR)
- Without camera can still get good accuracy.
- Just a handful of pictures can be enough to get reasonable accuracy

USE OF PRNU IN FORENSICS CONT

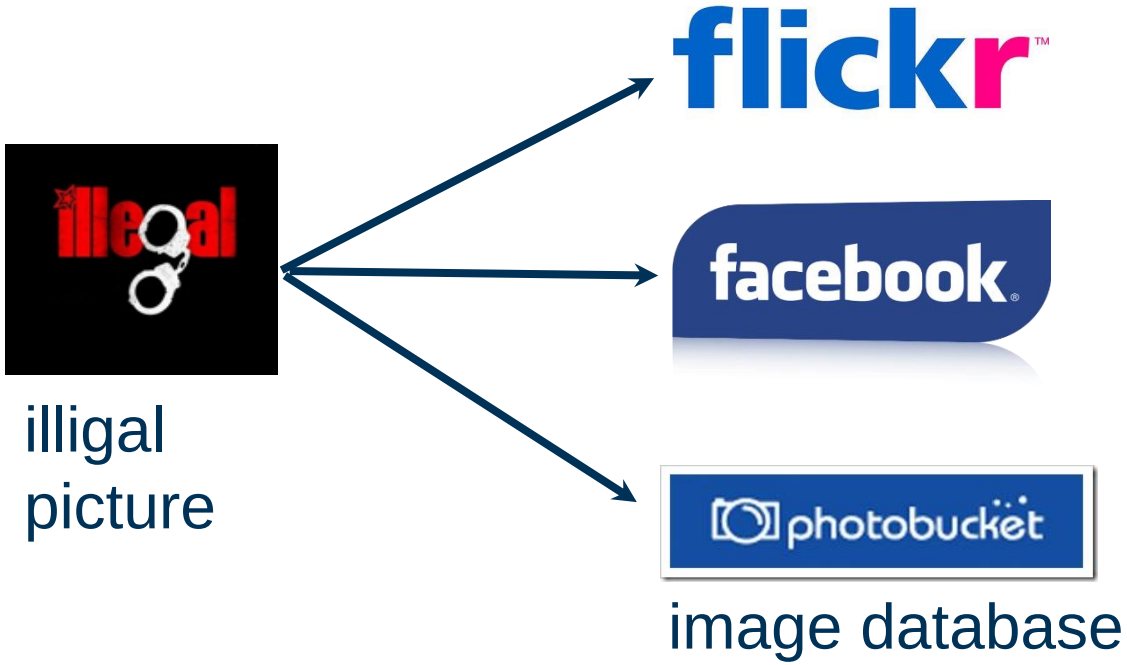
- Verification

illegal picture

Suspect device

USE OF PRNU IN FORENSICS CONT

- Identification



illigal picture

image database

- Efficiency is a major issue here as have to deal with a large database of fingerprints.

USE OF PRNU IN FORENSICS CONT

- Tamper (mosaic) detection

USE OF PRNU IN FORENSICS CONT

- Authentication (weak)

Authentication System

Alice's phone

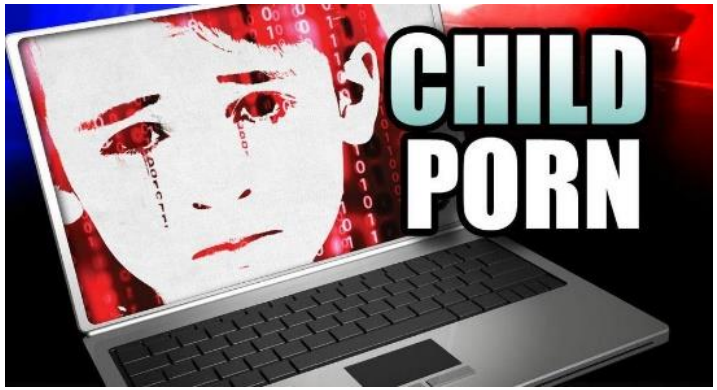
- Some papers in reading list regarding PRNU-based authentication

- Image clustering (based on camera)

PORNOGRAPHIC IMAGE DETECTION

The world still does not have a good tool for it!

PORNOGRAPHIC IMAGE DETECTION

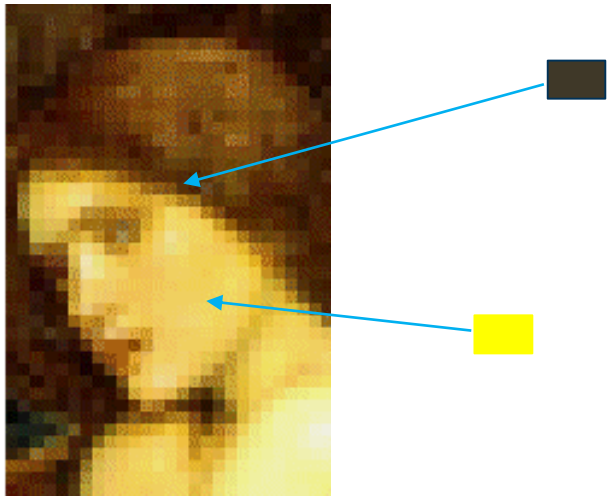


- Manually
  - Problem: Scalability (e.g., 10's of thousands of images)
- Automatically (using a software tool)
  - Performance

By detection, we will mean automatic detection

SKIN-TONE BASED METHOD

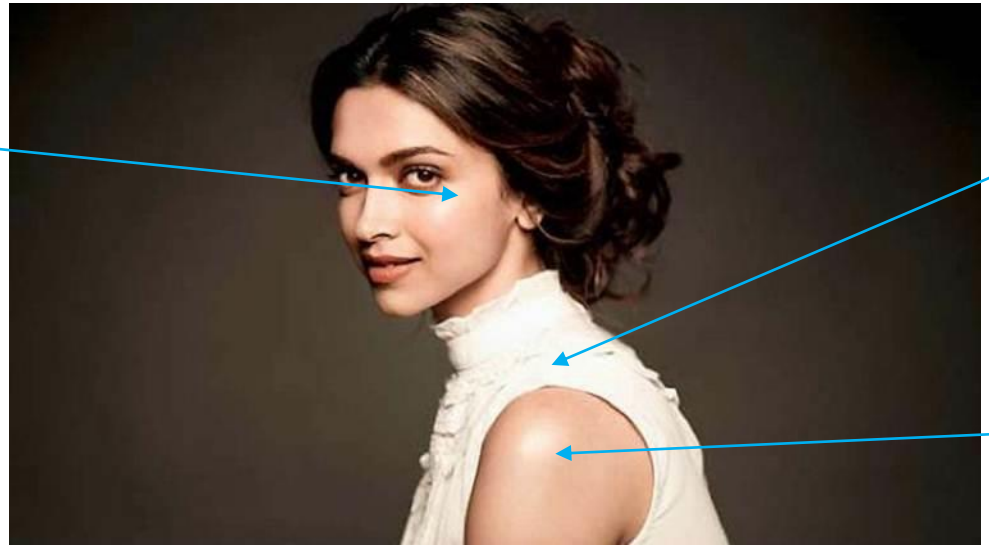
- Find exposed skin in the image and determine if the exposure implies nudity
- What is an (uncompressed) image?



- A matrix of coloured dots, aka pixels. Different pixels can have different colors. When they are arranged in the right way, a meaningful image is formed
- Binary image, gray image, color image

SKIN-TONE BASED METHOD: TWO MAIN STEPS

- Determining which images contain large areas of skin (skin pixels, smoothness via texture)
- Within colored regions, finding regions those resemble human body part



SKIN-TONE BASED METHOD CONT.

- Detecting skin pixels – explicit definition

(R, G, B) is classified as skin if:

$R > 95$ and $G > 40$ and $B > 20$ and

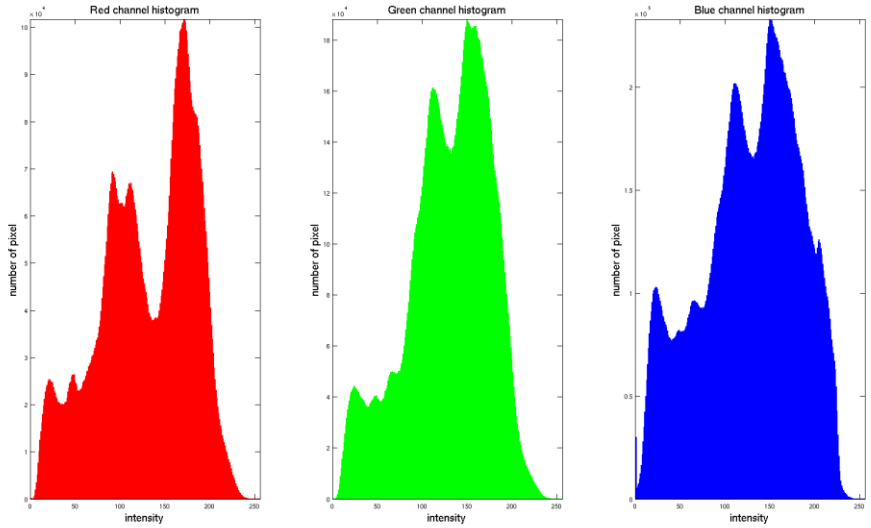
$\max(R, G, B) - \min(R, G, B) > 15$ and

$|R - G| > 15$ and $R > G$ and $R > B$

- How to get rule for a color space

SKIN-TONE BASED METHOD CONT.

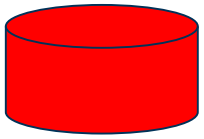
- Detecting skin pixels – histogram-based method
  - Given a large number of training images (skin and non-skin labelled)
  - Convert the images to a color space (RGB, YUV, HSB, etc.)
  - Find histogram for skin and non-skin color



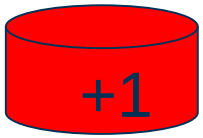
Example of RGB histogram taken from http://www.sci.utah.edu/~acoste/uou/Image/project1/Arthur_COSTE_Project_1_report.html

HOW HISTOGRAM WORKS

- A number of bins (max 256) for each color component

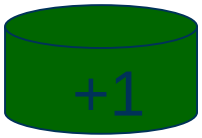


0-127

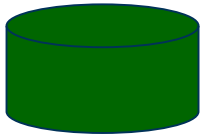


128-255

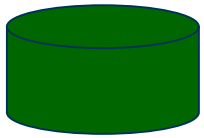
Labelled images



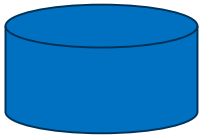
0-85



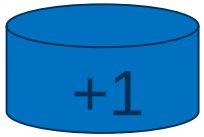
86-70



171-255



0-127



128-255

A “skin” pixel of an image has value (R=200, G=20, B= 190)

FORENSICS AND PRIVACY

- Forensics person: “What privacy? Do not you use Facebook?”
- Privacy person: “On the name of (national) security, they are stealing your data”.

Questions?

Thanks for your attention!

RESOURCES

- The Best Damn Cybercrime and Digital Forensics Book Period by Jack Wiles and Anthony Reyes. 2007. Syngress Publishing.

- The Basics of Digital Forensics: The Primer for Getting Started in Digital Forensics (1st ed.) by John Sammons. 2012. Syngress Publishing.

RESOURCES

- A. Pal and N. Memon, "The evolution of file carving," in IEEE Signal Processing Magazine, vol. 26, no. 2, pp. 59-71, March 2009.
- Garfinkel S. Carving contiguous and fragmented files with fast object validation. In: Proceedings of the 2007 digital forensics research workshop, DFRWS, Pittsburgh, PA; August 2007.
- Pal A, Memon N. Automated reassembly of file fragmented images using greedy algorithms. IEEE Transactions on Image processing February 2006:385–93.
- A. Pal, H. T. Sencar, N. Memon, Detecting file fragmentation point using sequential hypothesis testing, Proc. Dig. Invest. 5, p. 2-13, 2008.

RESOURCES

- E. Durmus, M. Mohanty, S. Taspinar, E. Uzun and N. Memon, "Image carving with missing headers and missing fragments," 2017 IEEE Workshop on Information Forensics and Security (WIFS), Rennes, 2017, pp. 1-6.
- Ho, A. and Li, S., Eds. Handbook of Digital Forensics of Multimedia Data and Devices, 1st ed., Wiley-IEEE Press, 2015. ISBN: 978-1118640500. DOI: 10.1002/9781118705773.2

RESOURCES

- Statistical Color Models with Application to Skin Detection; Michael J. Jones James M. Rehg; Compaq; http://www.hpl.hp.com/techreports/Compaq-DEC/CRL-98-11.pdf

- A survey on pixel-based skin color detection techniques; Vezhnevets et al.; Graphicon 2003; https://pdfs.semanticscholar.org/bc1b/5ff4fdb70c10a9aa0e9b8f6b260b2e1f4fed.pdf

- Automatic Detection of Human Nudes. D. A. Forsyth and M. M. Fleck; International Journal of Computer Vision; http://luthuli.cs.uiuc.edu/~daf/papers/ko5.pdf

RESOURCES

- Human skin colour clustering for face detection; Jure Kovac, Peter Peer, and Franc Solina; http://eprints.fri.uni-lj.si/2113/1/Human_Skin_Colour_Clustering_for_Face_Detection.pdf

- Towards automatic detection of child pornography; N. Sae-Bae, X. Sun, H. T. Sencar and N. D. Memon; ICIP 2014; https://ieeexplore.ieee.org/document/7026079/

RESOURCES

- J. Lukas, J. Fridrich and M. Goljan, "Digital camera identification from sensor pattern noise," in IEEE Transactions on Information Forensics and Security, vol. 1, no. 2, pp. 205-214, June 2006
- Miroslav Goljan, Jessica Fridrich, Tomáš Filler, "Large scale test of sensor fingerprint camera identification", Proc. SPIE 7254, Media Forensics and Security, 72540I (4 February 2009).
- Jessica Fridrich, “Digital Image Forensics Using Sensor Noise”, Tutorial.
- S. Taspinar, M. Mohanty and N. Memon, "PRNU-Based Camera Attribution From Multiple Seam-Carved Images," in IEEE Transactions on Information Forensics and Security, vol. 12, no. 12, pp. 3065-3080, Dec. 2017.