Cybersecurity

Multimedia Forensics: a brief introduction

Mauro Barni
University of Siena

Summary

- Motivations
  - Examples
- Introduction to Multimedia forensics
- General principle underlying MF
- Some simple examples

The problem
(focus on visual data)

Seeing is believing ?

Photographic images have lost their innocence (if they ever had one) a long time ago ...

Seeing is believing ?

Seeing is believing ?

Seeing is believing ?

With the diffusion of digital images, the validity of photos as witnesses of real events is definitely lost

You only need to listen to everyday news

Was it for Gossip only !!!

Frontal light

Side light

Frightening enemies (dictatorship)

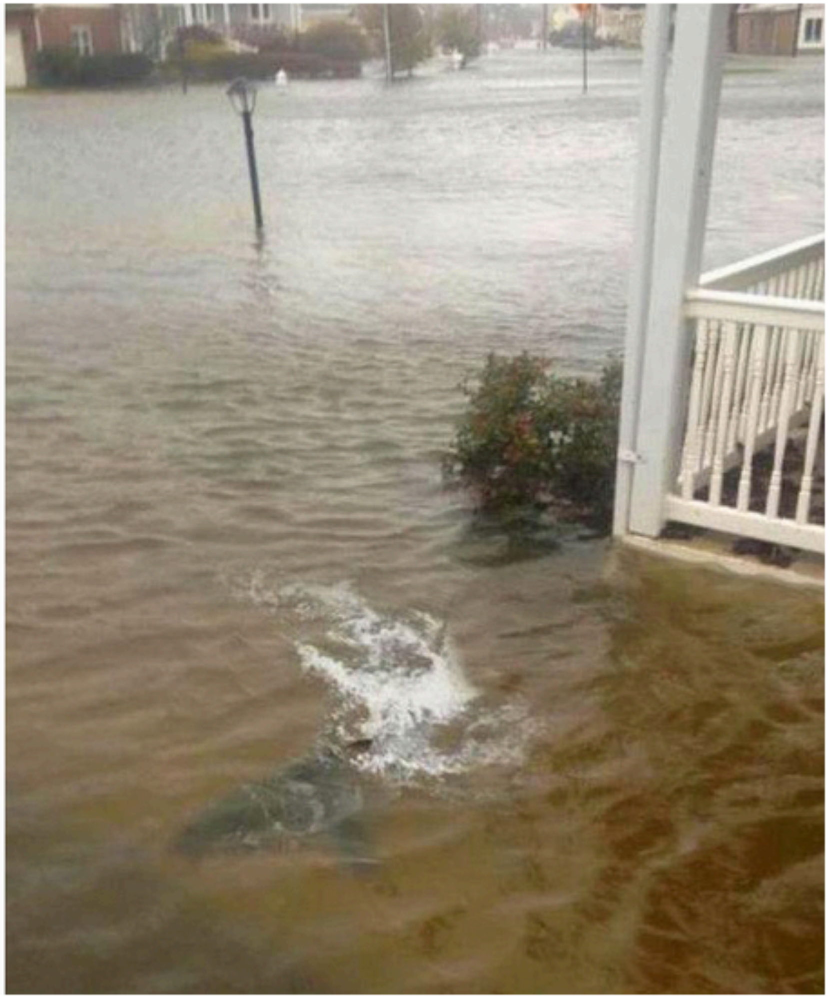
Artificially-augmented support

Conveying a message the picture does not tell

=

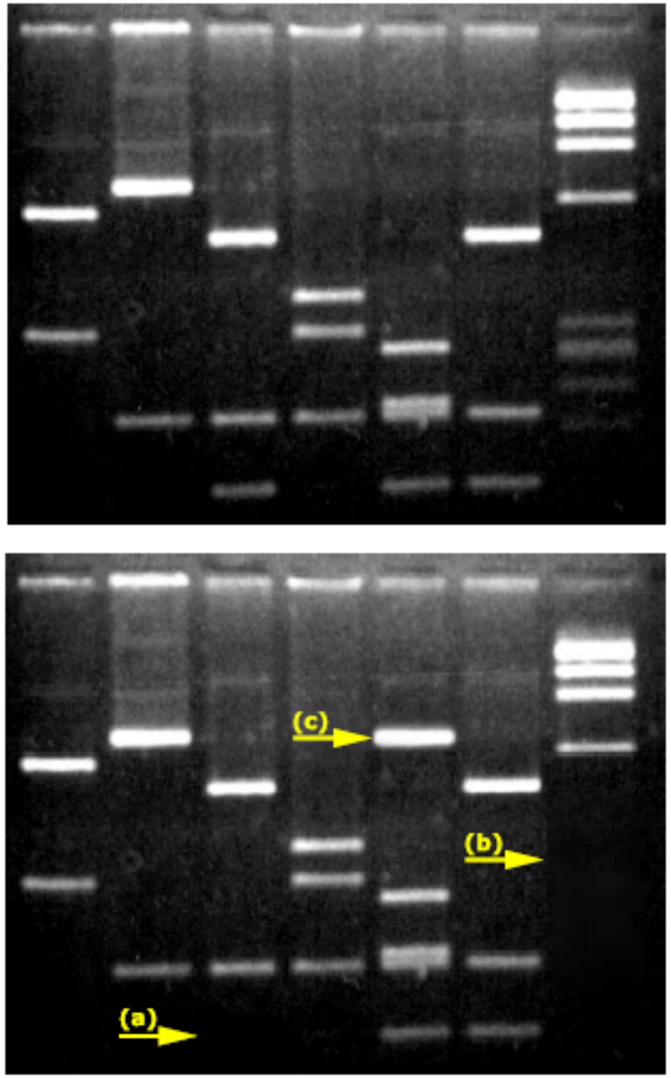
+

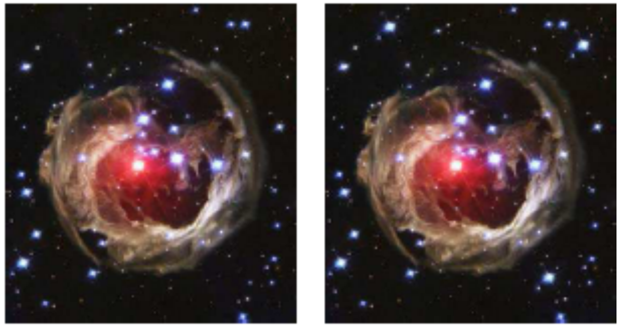
Let alone the web !!!

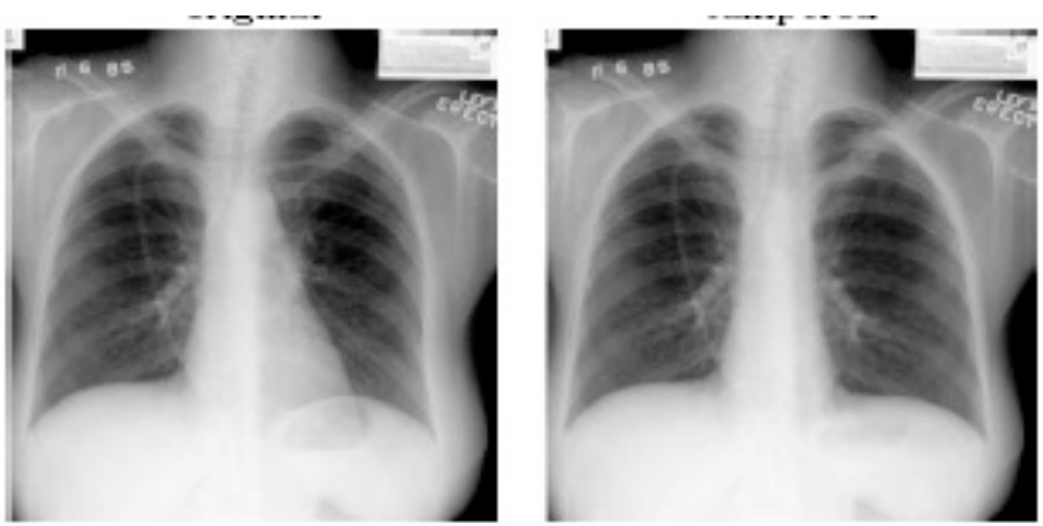


Impressions from Hurrican Sandy

Even scientists







Not only photomontages

CG

CG

Real

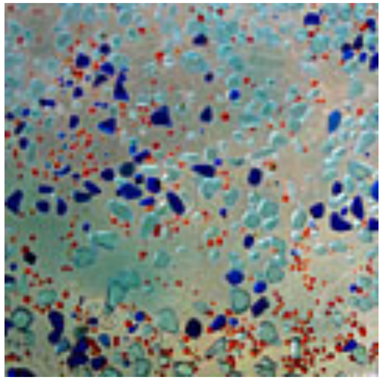
CG

Real

CG

Real

CG



Real

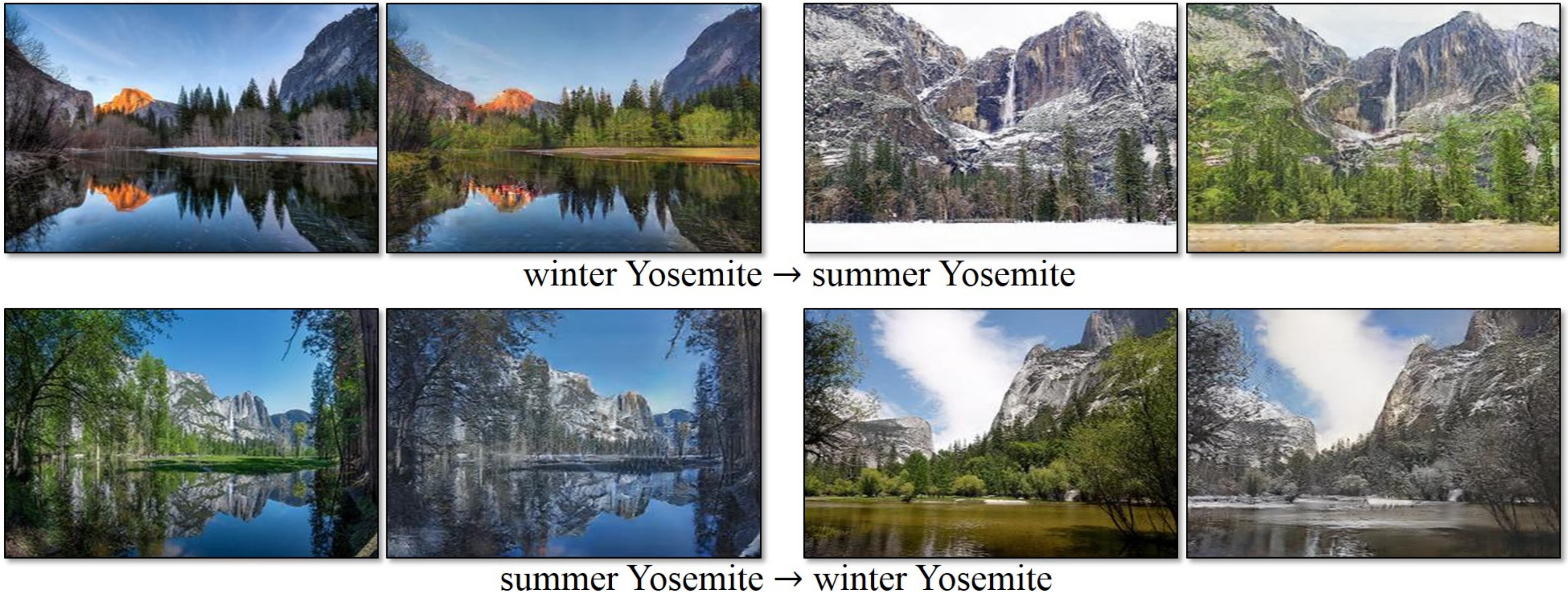
Real

Not only images

A dark side-effect of the AI revolution

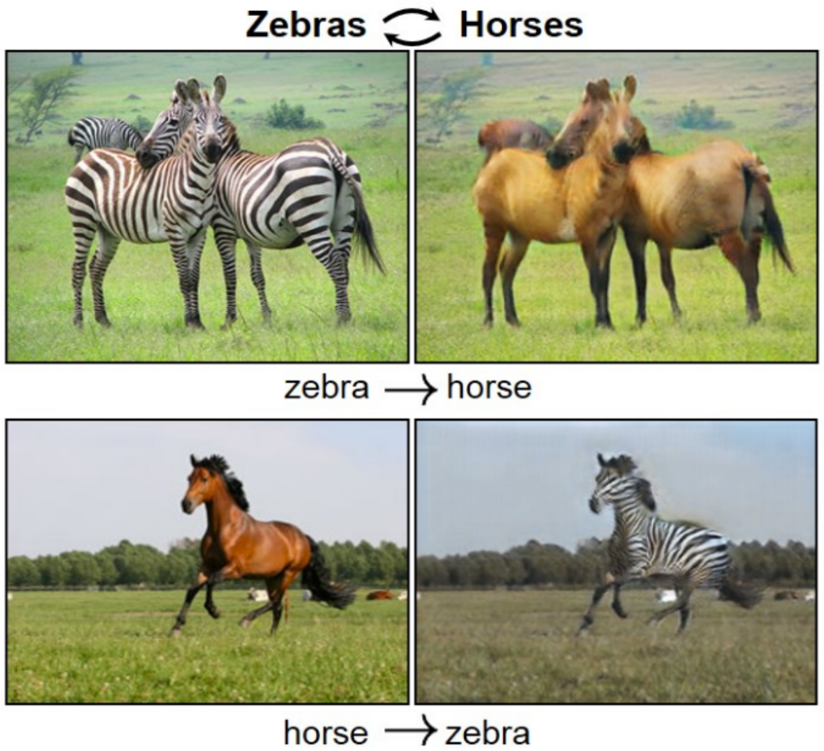
Fake HUMANS - more here

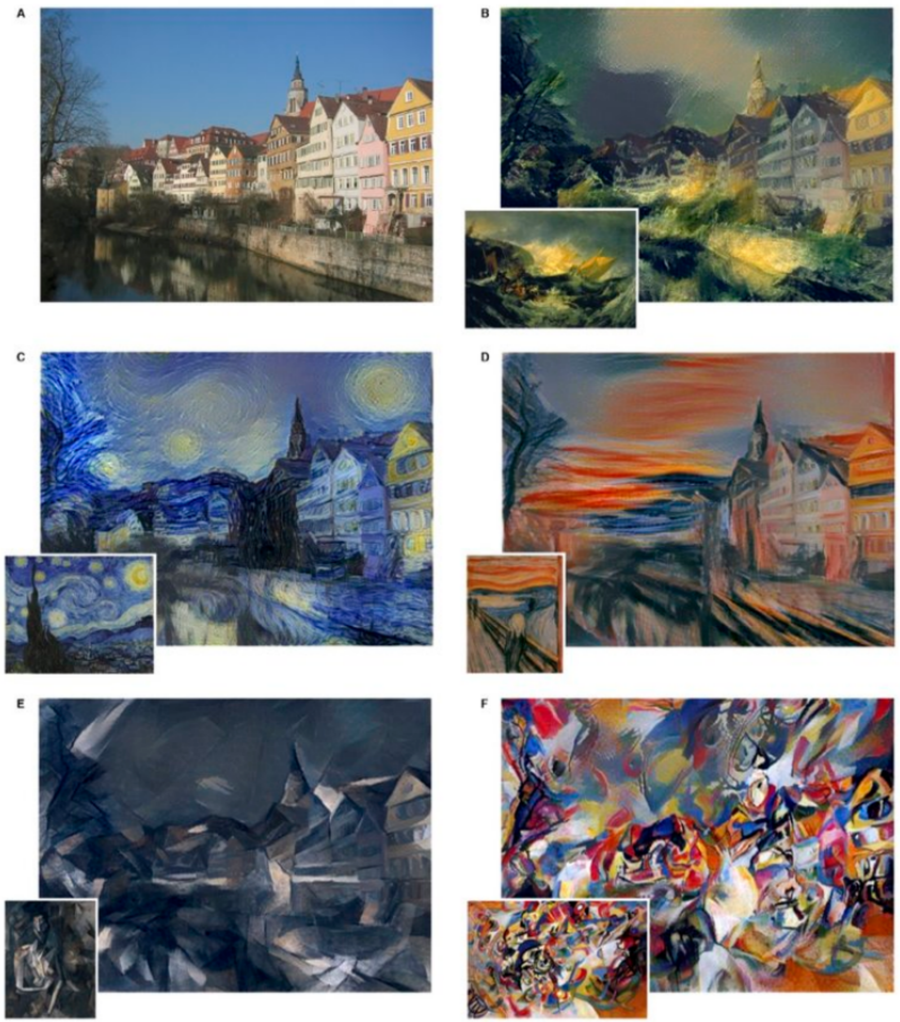
A dark side-effect of the AI revolution



Style (season) transfer

A dark side-effect of the AI revolution





And movies …

Why should we care ?

- Opinion manipulation
- Social impact: undermines one of our primary source of information
- Probatory value of digital images, videos, audios

- Scientific question: ultimate reliability of digital media as trustful representation of reality

The solution(s)

Two approaches to MM authentication

- **Active approach:**
  - Cryptographic Signature:
    - Extracting features for generating authentication signature at the source side and verifying the image integrity by signature comparison at the receiver side.
    - Possibly coupled with blockchain technology
    - It requires a complete cryptographic infrastructure
  - (same) Main problem: does not survive D/A and A/D conversion

Two approaches to MM authentication

- **Active approach:**
  - Fragile/Semi Fragile Digital Watermarking
    - Inserting digital watermark at the source side and using the watermark to verify integrity at the detection side.
    - Two approaches based on
      - Fragile watermarking
      - Robust watermarking

Two approaches to MM authentication

- **Passive and blind approach: multimedia forensics**
  - Without any prior information, verifying whether an image is authentic or not
  - Advantages:
    - No need for watermark embedding or signature generation at the source side.
    - No need for a standard
    - No need for a priori knowledge about the acquisition device

Forensic Science: forensics

forensics

use of scientific methods for gaining probative facts (from physical or digital evidences)

digital forensics

analog forensics

computer forensics

multimedia forensics

Multimedia Forensics

- Given a digital data (i.e. image), multimedia forensic techniques try to answer a number of forensic questions related to:
  - source identification
    What is the origin of the data ?
  - integrity verification / tampering detection
    Has the data undergone some processing ?

Source identification

- How was the image captured?
- Which CLASS of device was used?

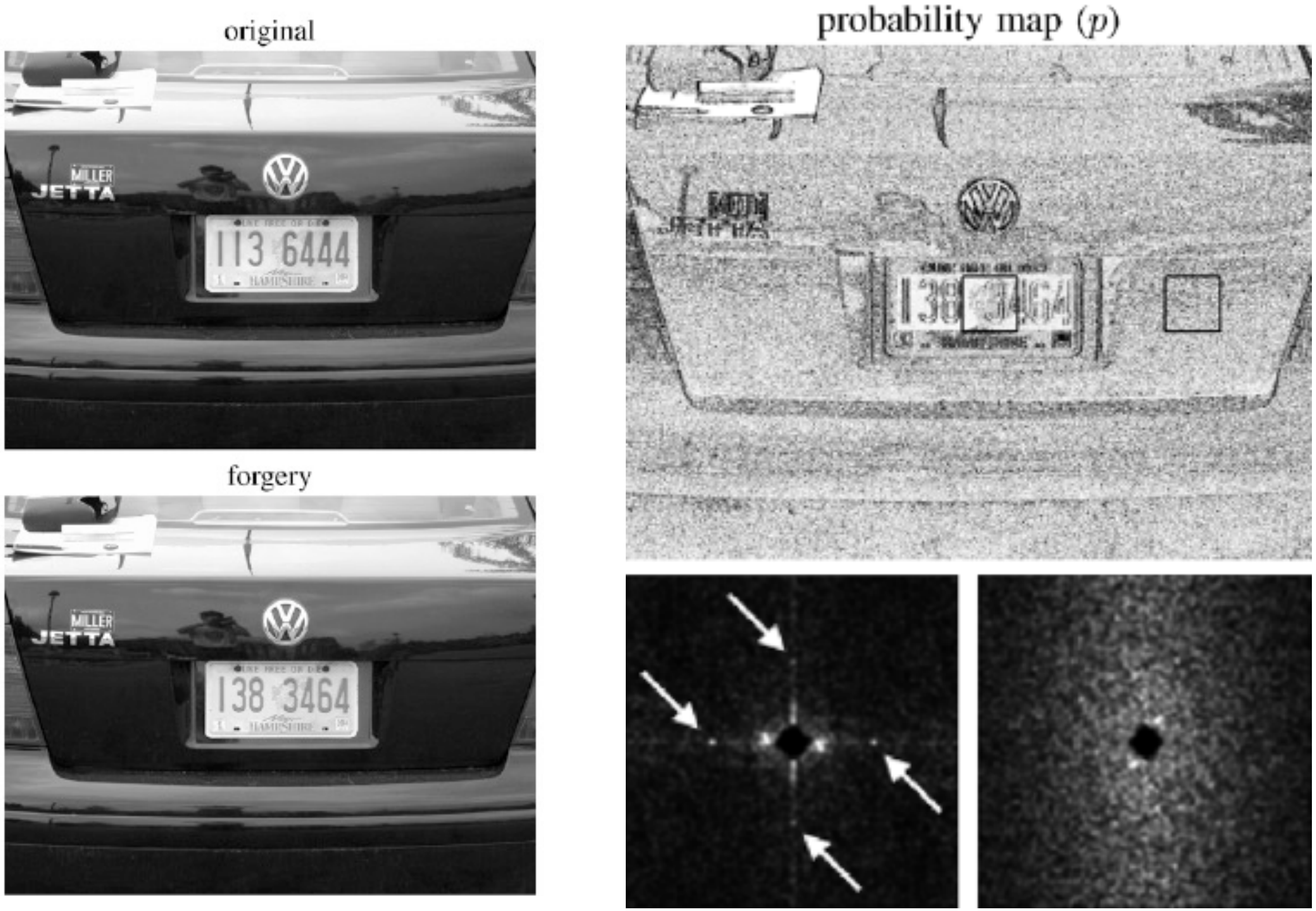



- Which BRAND / MODEL / SPECIFIC DEVICE?

Manipulation detection

- Is the image authentic ?
- How was it tampered with ?

Manipulation detection



The basic idea

- Multimedia forensics is based on the idea that inherent traces (like digital fingerprints) are left behind in a digital media during both the creation phase and any other subsequent processing.

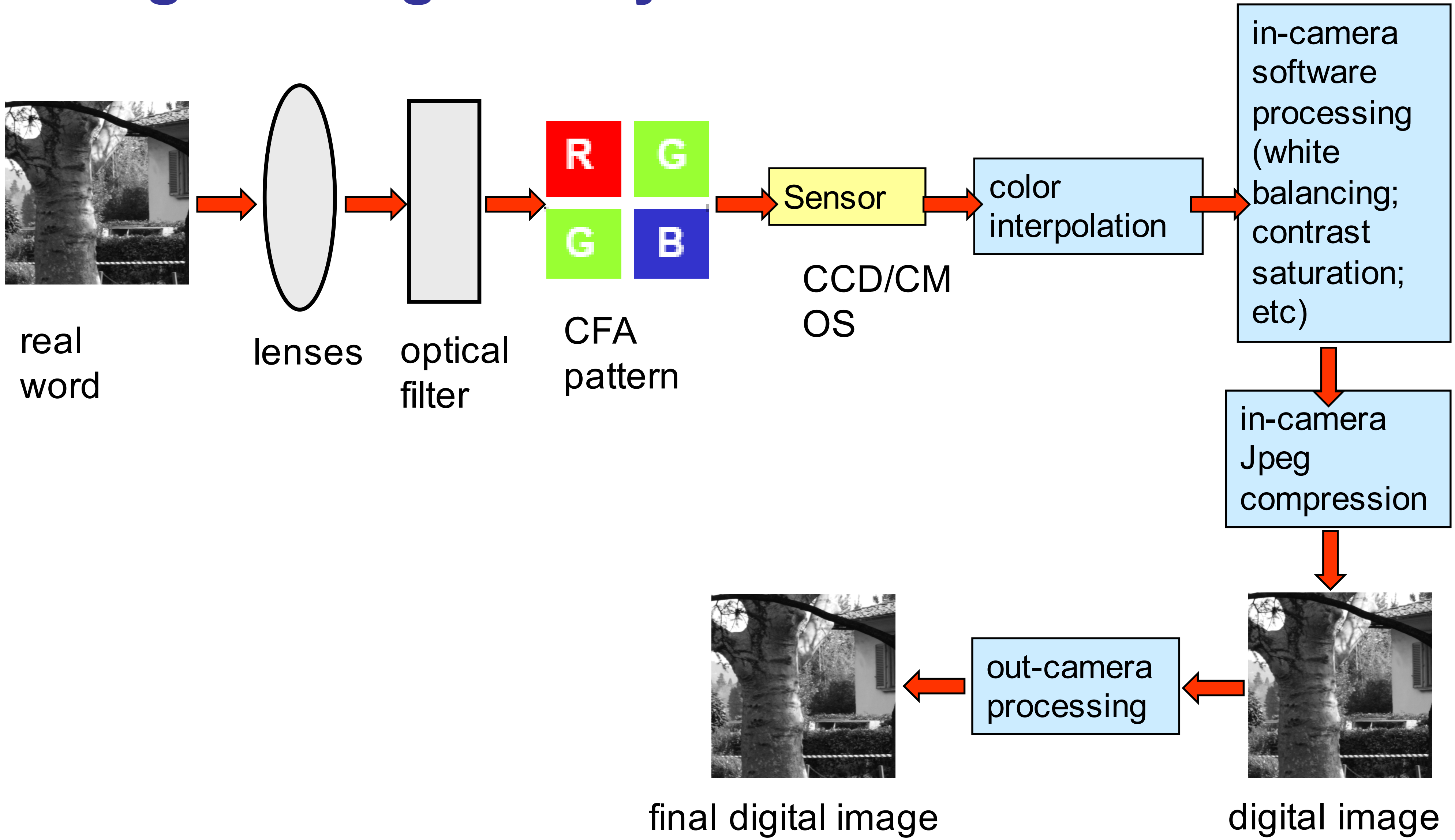
Digital fingerprints

- **In-camera fingerprints**: each component in the acquisition device leaves intrinsic fingerprints in the final output, due to the specific optical system, color sensor and camera software.
- **Out-camera fingerprints**: each processing applied to digital media modifies their properties (e.g. statistical, geometrical, etc.) leaving peculiar traces.
- **Scene (geometric) fingerprints**: the real world has specific properties depending on the content, like lighting properties, which characterize the reproduced scene (illuminant direction, specular highlights in the eye)

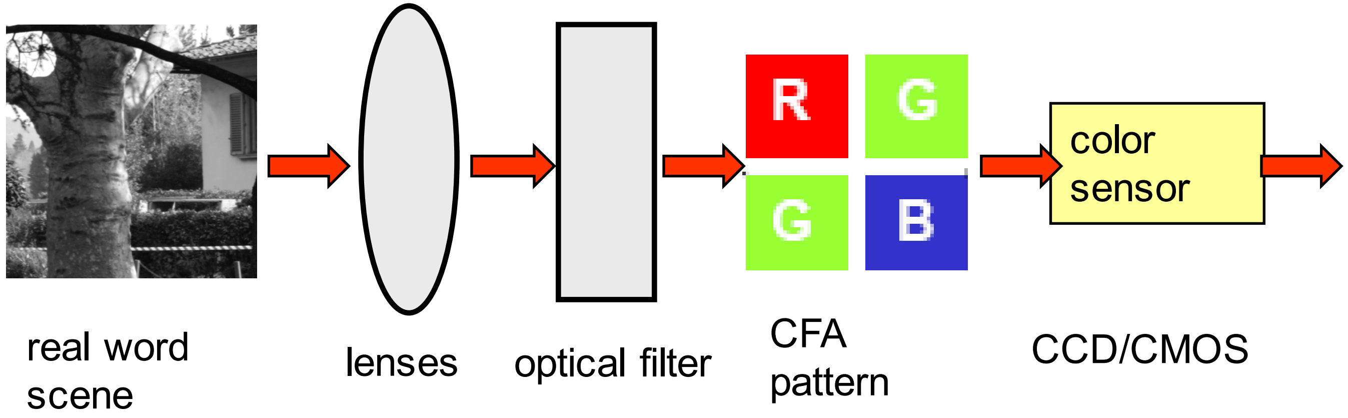
Use of digital fingerprints

- For **source identification**:
  - Fingerprints are usually extracted and then compared with a dataset of possible fingerprints specific for each class/brand/model of acquisition (creation) devices
- For **forgery detection**:
  - detect non-uniformity or absence of fingerprints within the analyzed data
  - detect the presence of fingerprints pointing to a specific post-processing

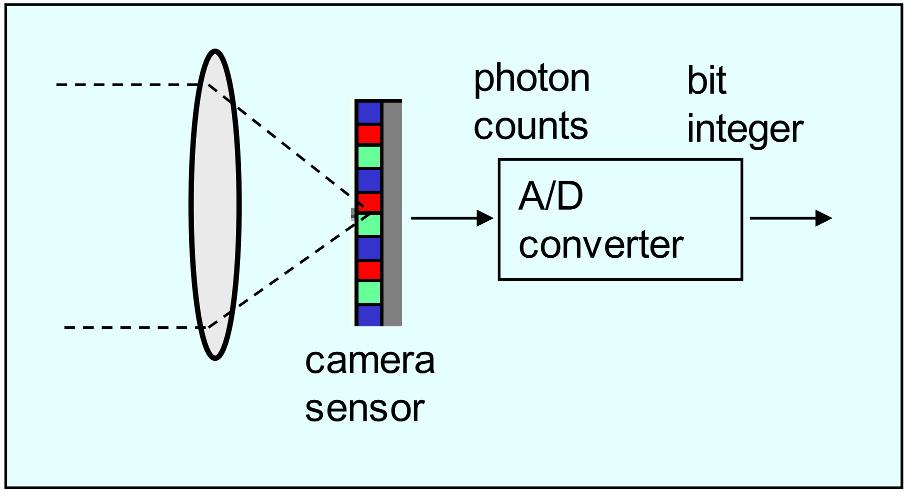
Digital image life cycle



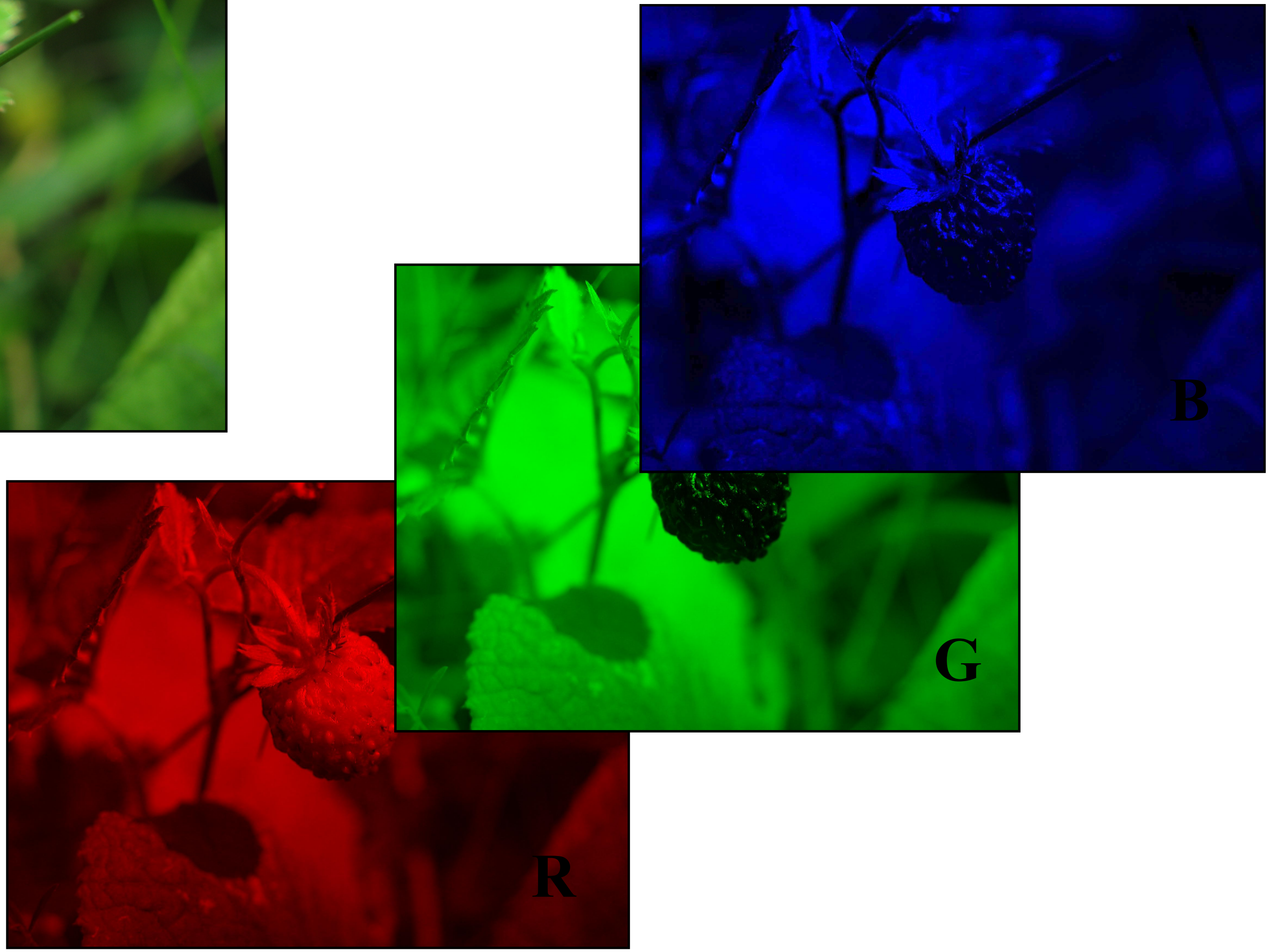
Digital Camera Model



- Light is focused by the lenses on a 2D array of CCD/CMOS (pixels).
- Such elements are hit by the photons and convert them into voltage signals which are then sampled by an A/D converter.
- Before reaching the sensor, the rays from the scene are filtered by the CFA (Colour Filter Array)

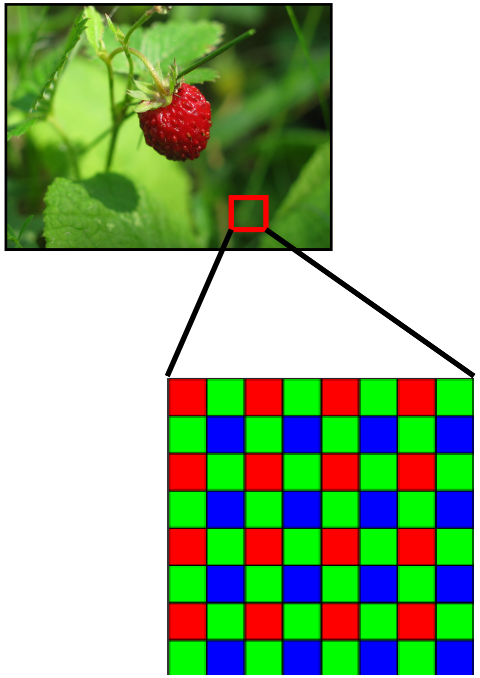


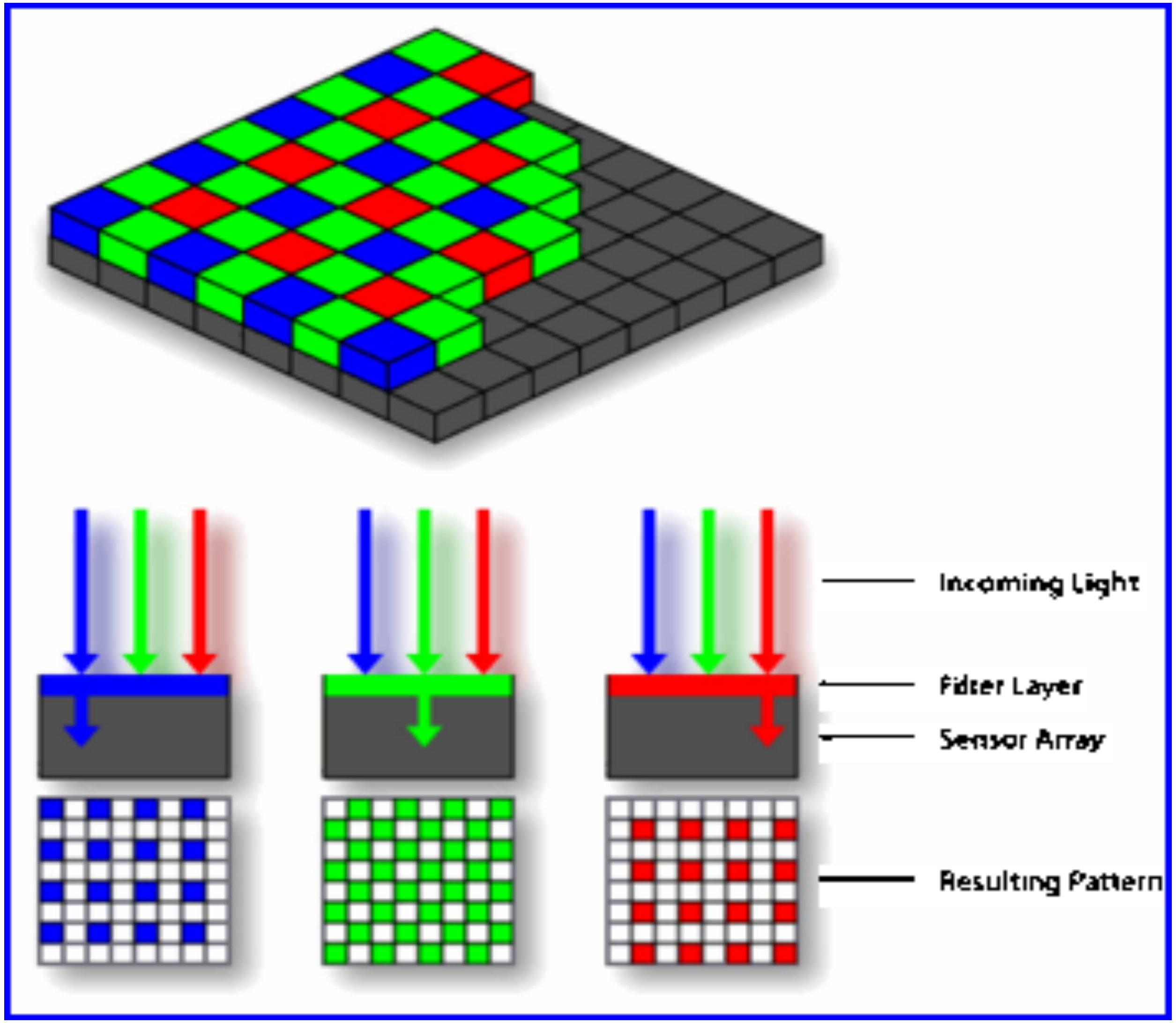
Bayer color array



Bayer color array

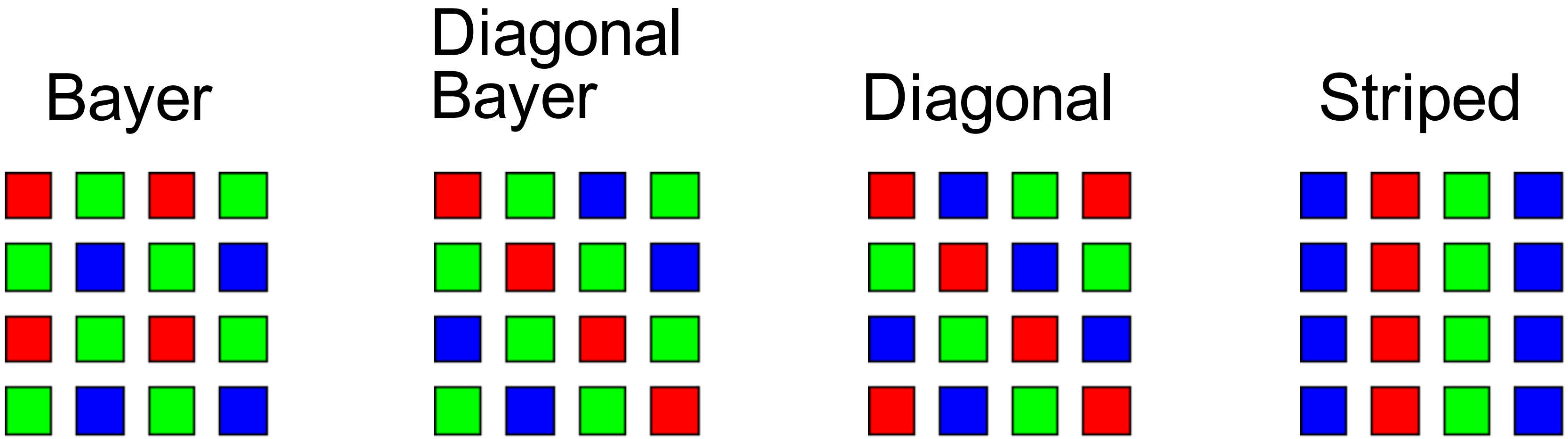
- Half pixels are Green, a quarter Red and a quarter Blue



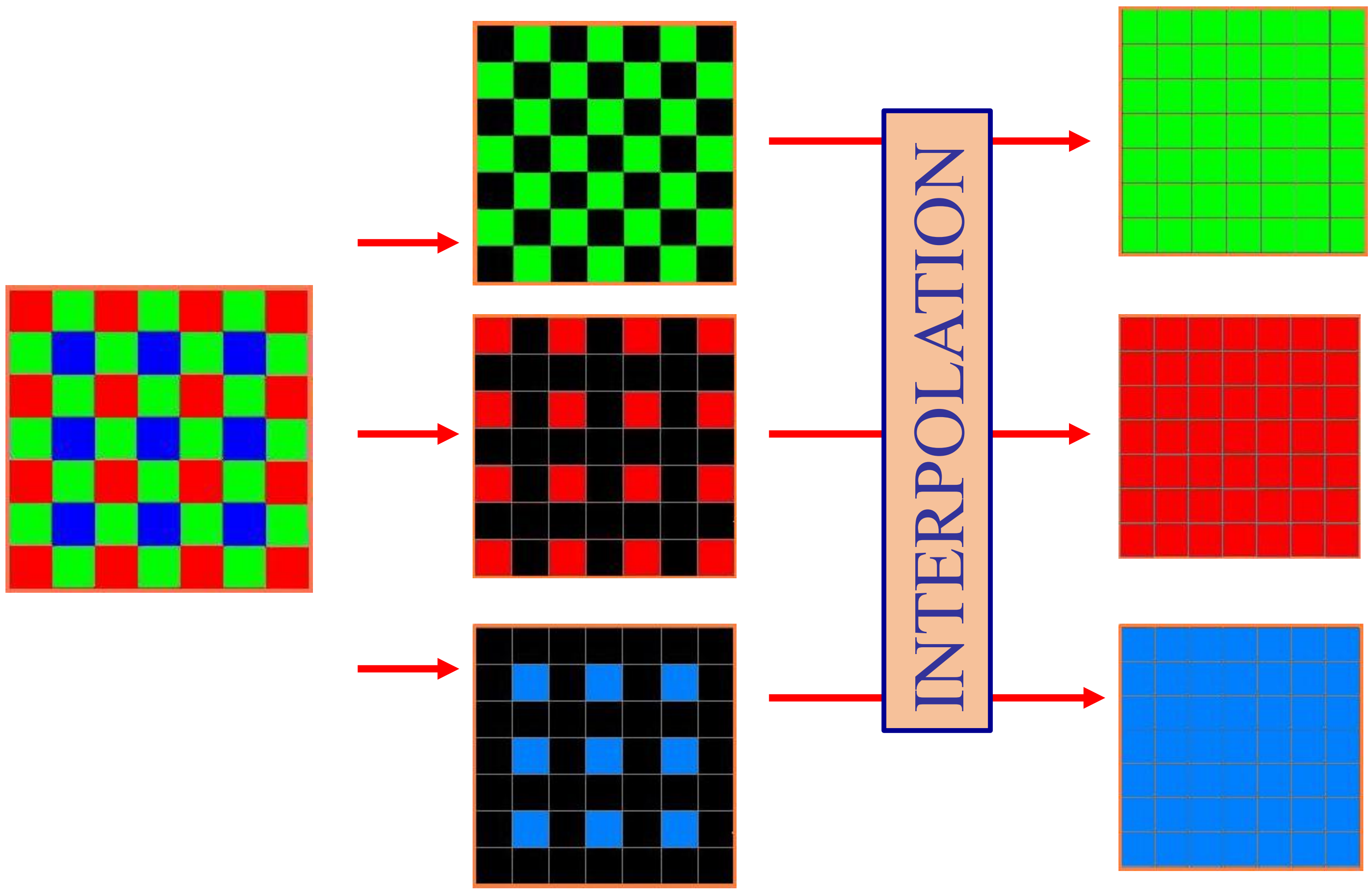


Bayer color array

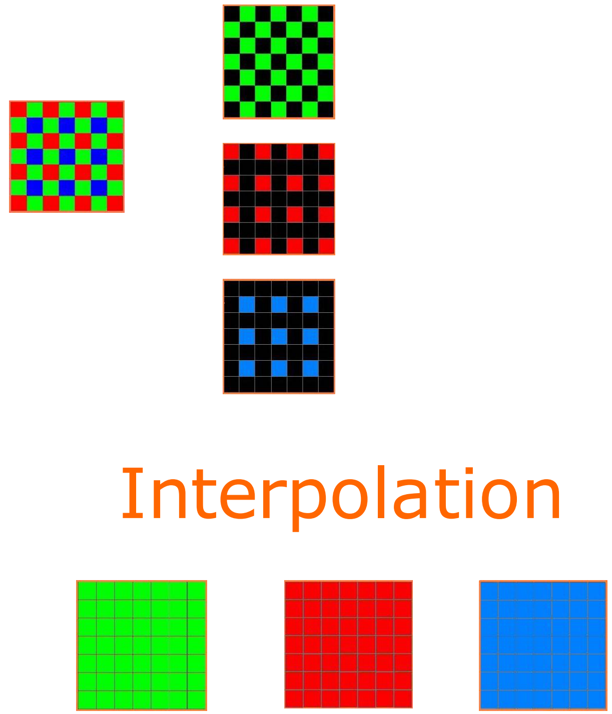
- Several possible patterns



Digital cameras - forming color



Source identification

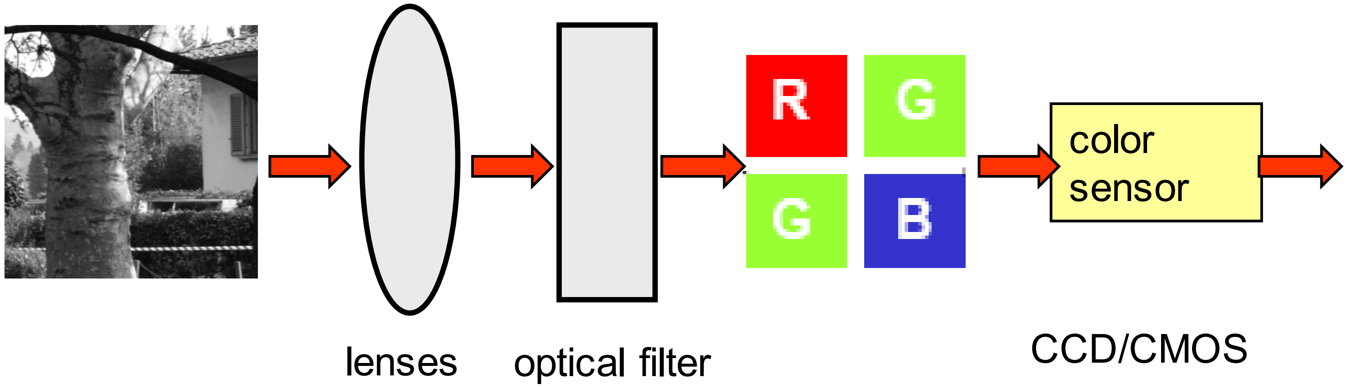


- Bayer Array for almost all digital cameras
- Color Interpolation different for each make of Digital Camera
- In the same way we can distinguish between different devices: scanners, CG images

Tampering detection

Incongruencies in CFA fingerprint can be used to detect tampering

Sensors imperfections: noise



- Sensor noise has 2 main components
- Fixed Pattern Noise (FPN)
  - pixel to pixel difference in dark conditions
  - additive noise
- Photo-Response Non-Uniformity (PRNU)
  - dominant part of the pattern noise: multiplicative noise

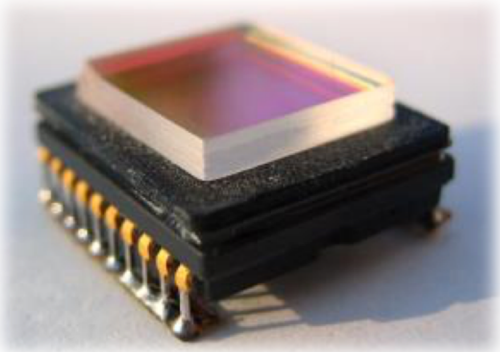
Sensors imperfections: FPN

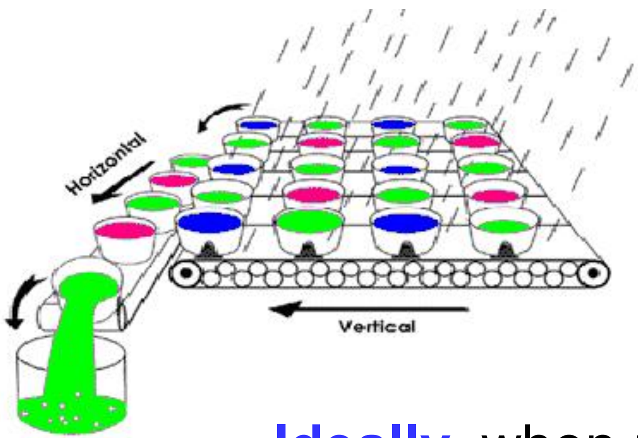
- The FPN is the pixel to pixel difference when the sensor is not exposed to light
- In most digital cameras this difference is equalized by subtracting a dark frame (mask) from the picture.

$$\begin{bmatrix} -2 & 0 & 3 & -5 \end{bmatrix} - \begin{bmatrix} -2 & 0 & 3 & -5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}$$

Sensors imperfections: PRNU

Typically, a digital camera has a 2D array of **several million CCDs**, each of which is responsible of the acquisition of a single pixel





A CDD is often exemplified as a **bucket collecting rain** (photons) until a certain level (the pixel value) is reached

Ideally, when uniform light falls on a camera sensor, each pixel should output exactly the same value……

Practically, small variations in cell size and substrate material result in slightly different output values

Photo Response Non Uniformity

- Given an image I of size MxN, image imperfections can be modeled as (*simplified model*):

$$I(x,y) = I_0(x,y) + I_0(x,y)\,K(x,y) + N(x,y)$$

- I_0(.) is the noise-free image, K(.) multiplicative noise term (PRNU), N(.) is an additive noise term (other disturbs).
- Goal: extract signal of interest K() from observed data I() -> use of denoising tools

Photo Response Non Uniformity

- Estimate of PRNU of camera C obtained by averaging noise residuals of a number of training images taken from C

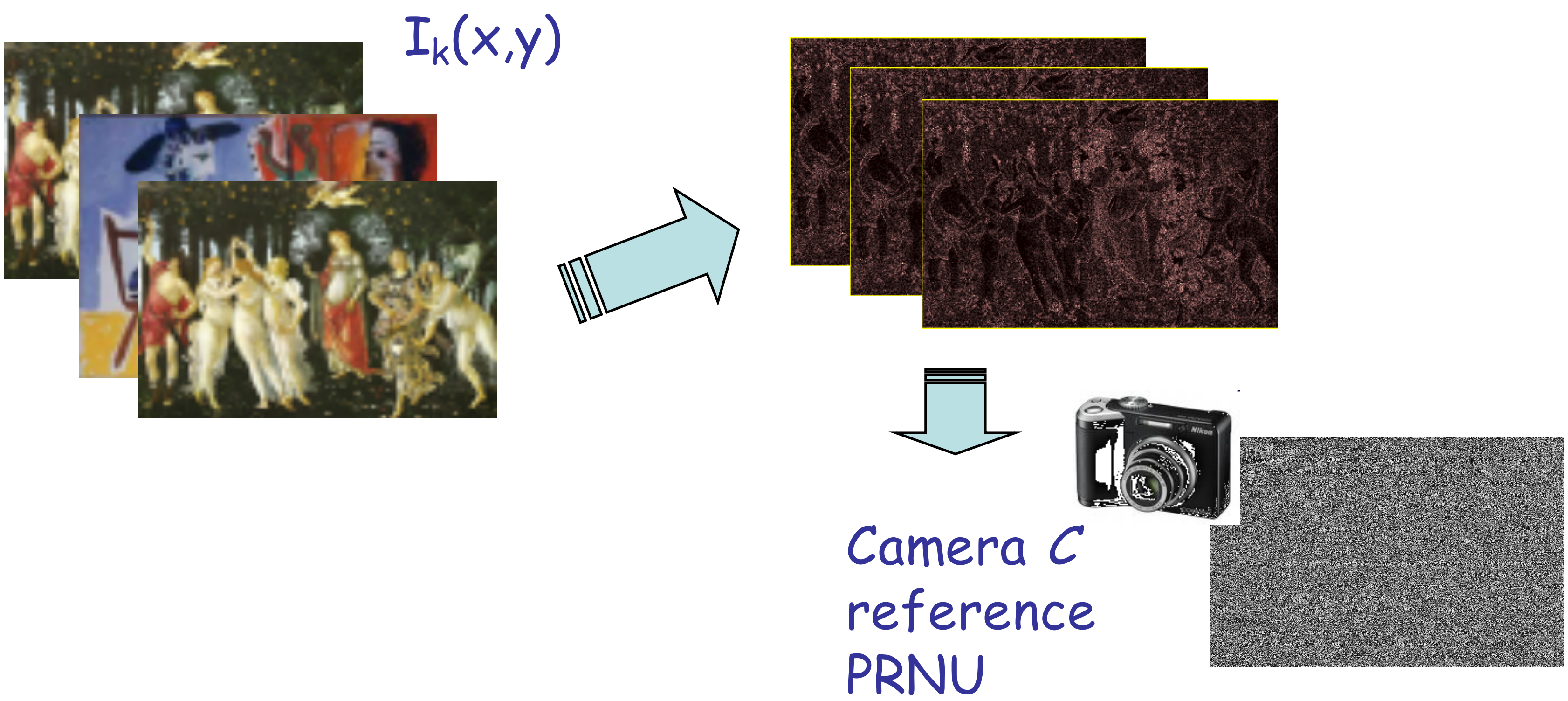


Image residual: standard filters

$$W_I = I - F(I)$$

- **Gaussian smoothing, 2D-Wiener …**
- **Advantages**
  - Simple implementation
  - Very fast
- **Disadvantages**
  - Image content left behind in the pattern **alters the correlation** between reference PRNU and the residual of the image under analysis

Image residual: best filter

$$W_I = I - F(I)$$

- **Wavelet based denoising**
- **Advantages**
  - Significantly more accurate
  - Better PRNU fingerprint estimation
- **Disadvantages**
  - Slower
  - Higher complexity

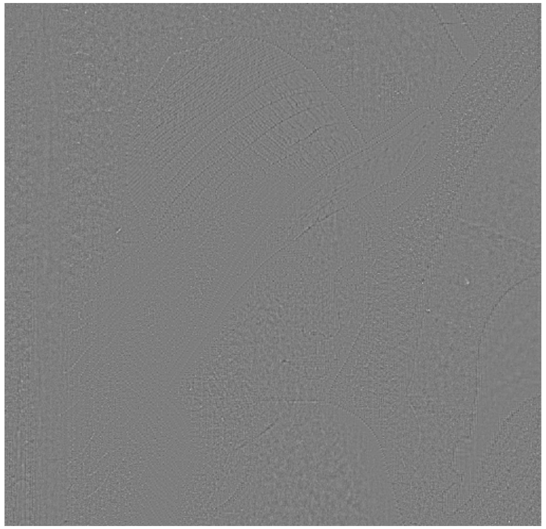
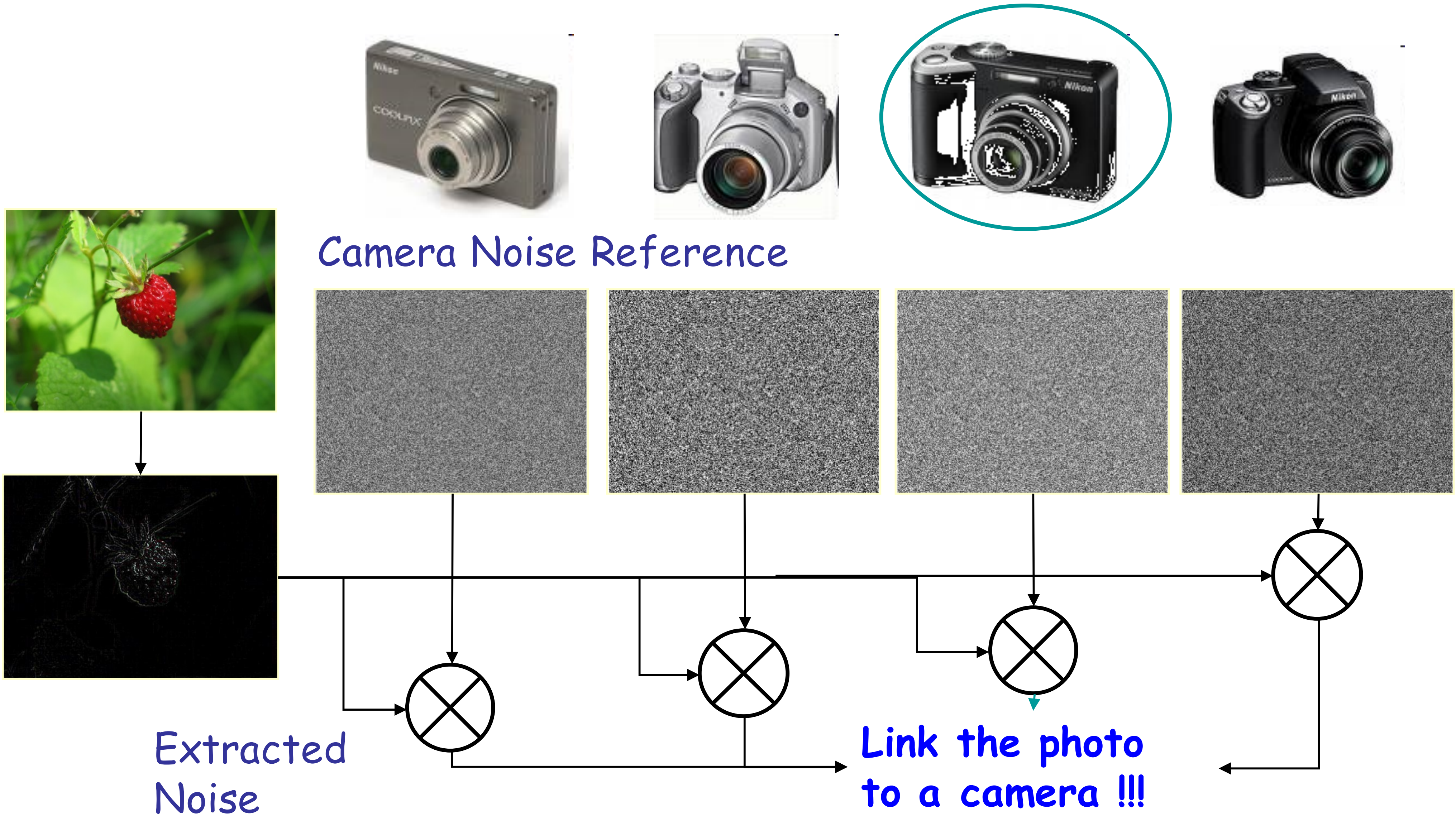
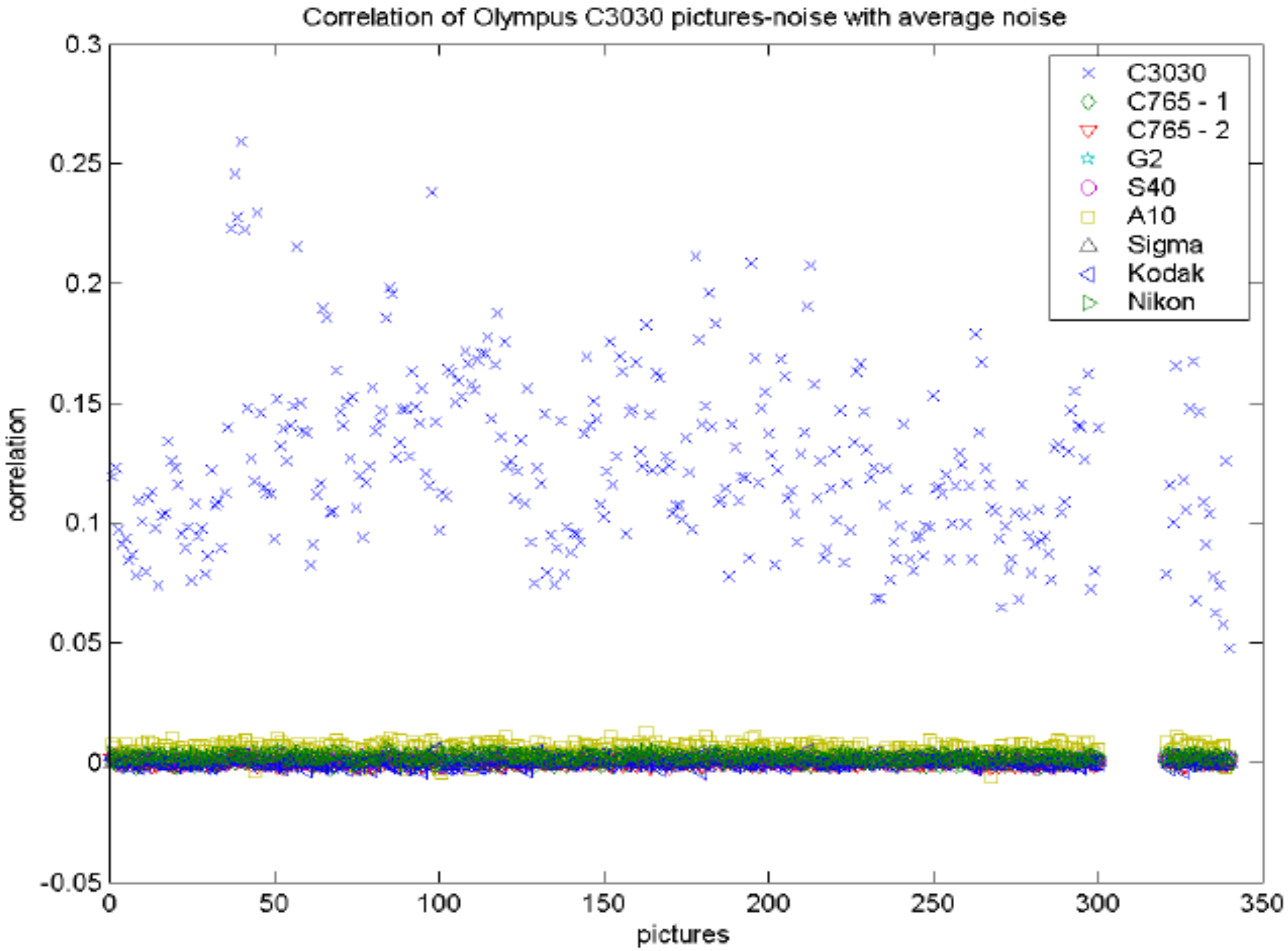


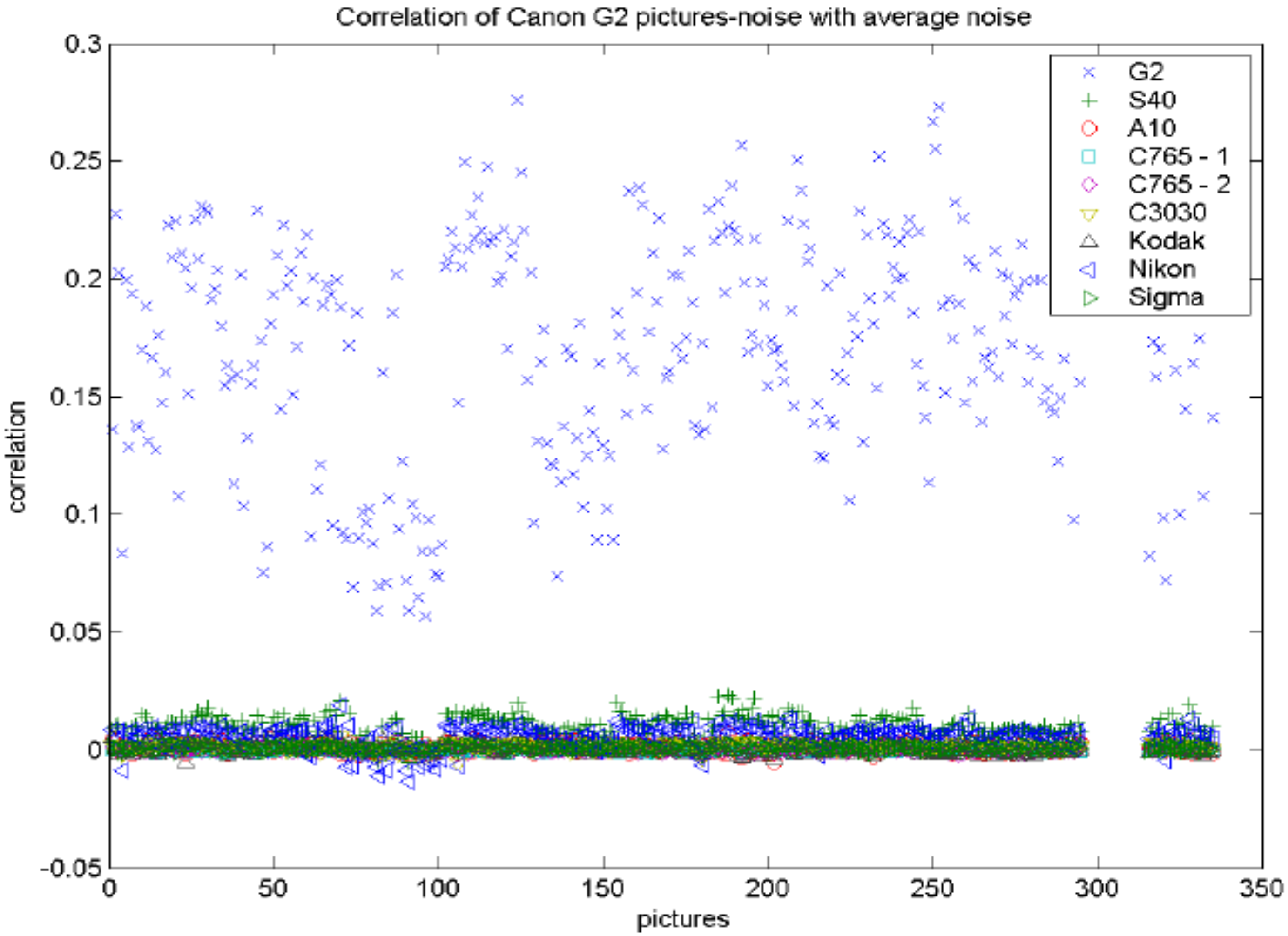
Photo Response Non Uniformity



Olympus 3030 (all JPEGs)



Canon G2 (raw)



Correlation of Canon G2 pictures-noise with average noise
correlation
pictures
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
0
50
100
150
200
250
300
350
G2
S40
A10
C765 - 1
C765 - 2
C3030
Kodak
Nikon
Sigma

Digital Camera Model

- The signal undergoes additional processing such as: white balancing, color processing, image sharpening, contrast enhancement, gamma correction.
- It is stored in the camera memory in a customized format, (for commercial devices JPEG format is usually preferred).
- All these steps introduce traces that can be exploited for MF analysis

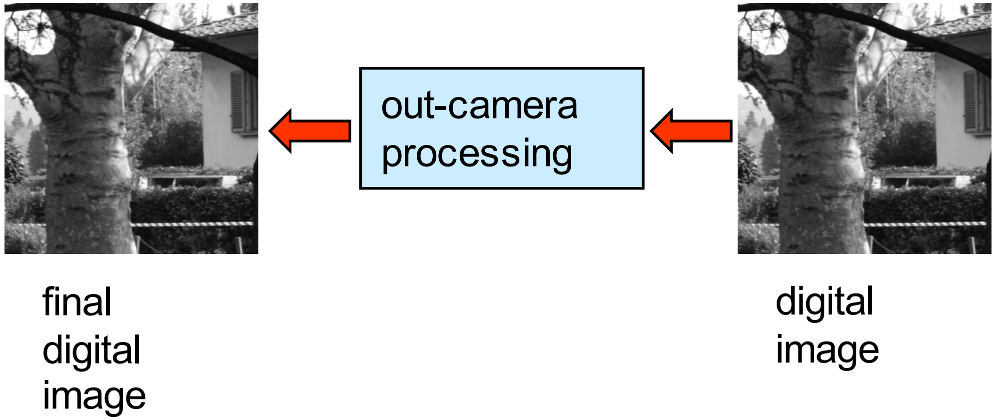
in-camera software processing (white balancing; contrast saturation; etc)

in-camera Jpeg compression

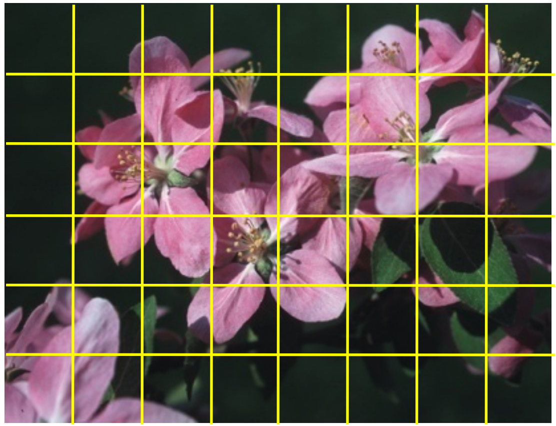
digital image

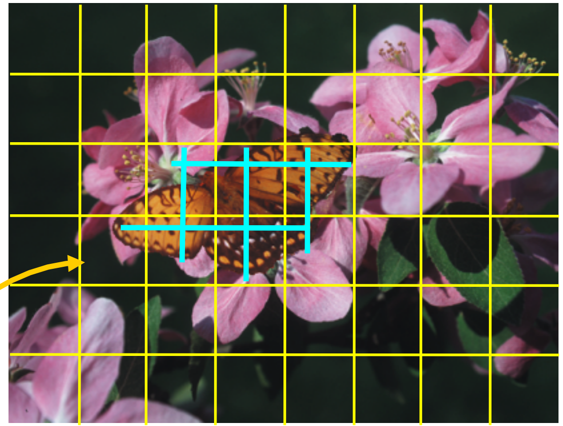
Out-camera processing

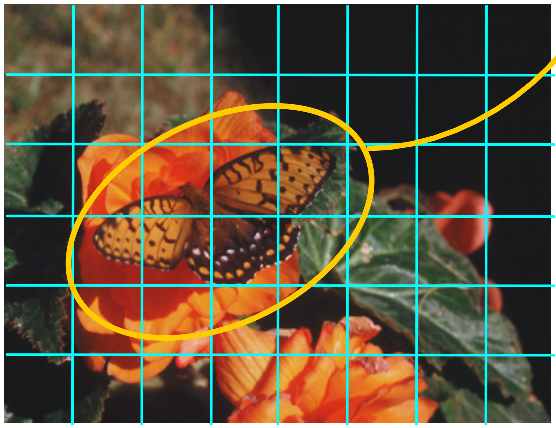
- Several kinds of processing
  can be applied to an image during its life:
  - compression
  - geometric transformation (rotation, scaling, …)
  - blurring and sharpening
  - contrast adjustment
  - …



Double JPEG artifacts







- JPEG compression leaves artifacts at the border of 8x8 blocks
- In case of double compression, the traces of old and new compression stages are likely to be deynchronized thus opening the door to MF analysis

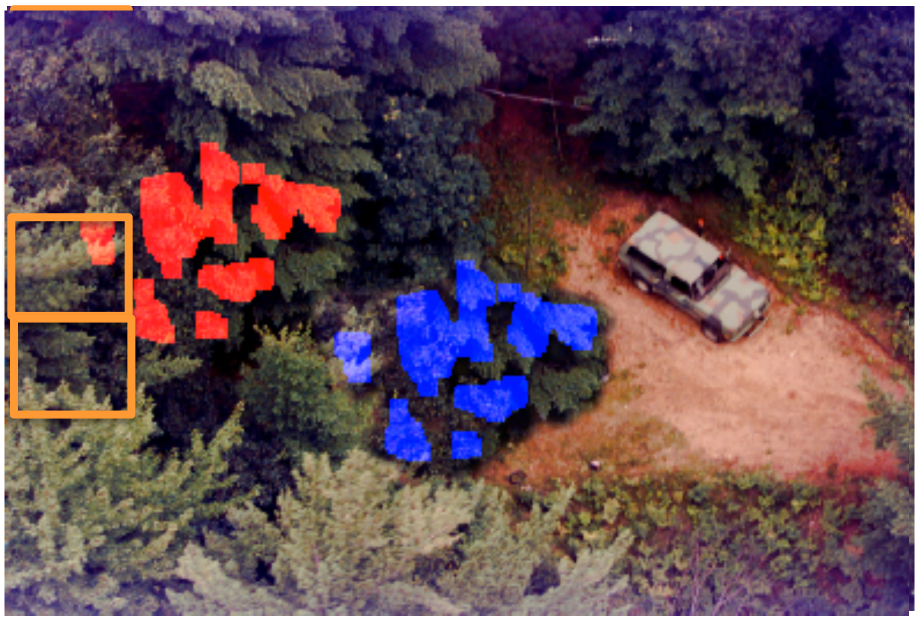
Double JPEG artifacts

Geometric fingerprint: copy move

- It's a particular kind of tampering: part of an image is duplicated to cover some undesired details

Geometric fingerprint: copy move

- Possible solution: we analyze the image with a sliding window looking for improbable duplicates



- It is not possible to understand which between the duplicated parts is the original

Geometric fingerprint: highlights

- We model eyes as spheres and infer the direction of light source from highlights

- Very likely this picture was taken in three different time instants

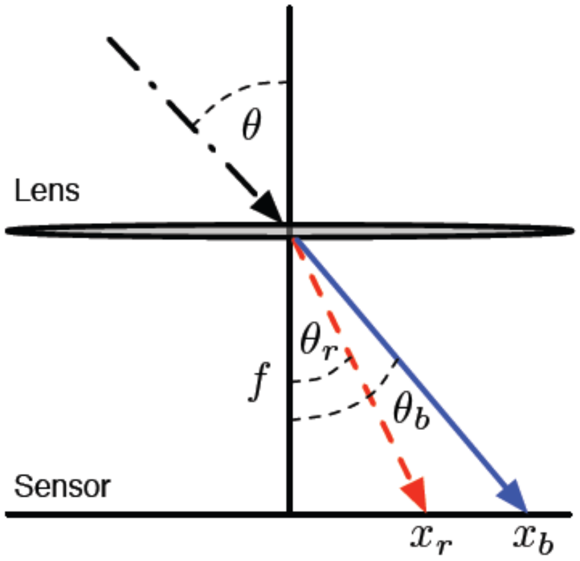
Geometric fingerprint: shadows

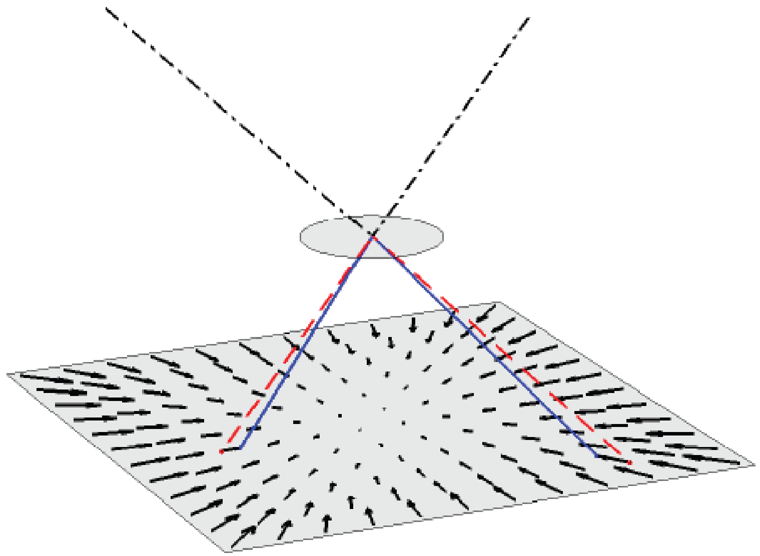
- Under certain assumptions it is possible to derive the direction of light from shadows

Geometric fingerprint: shadows

- Creating a photomontage by preserving the coherence of light and shadows is not an easy task

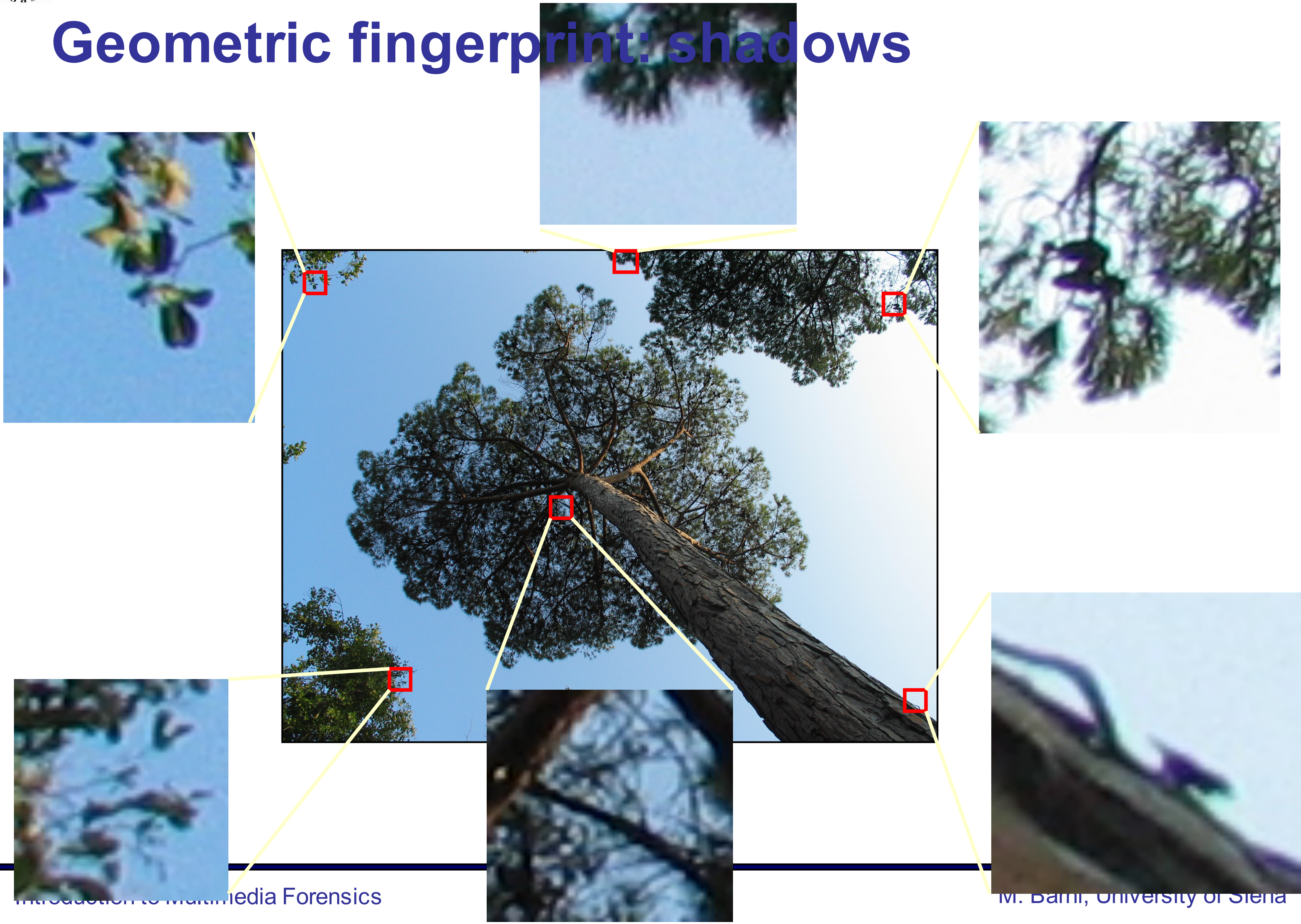
Geometric fingerprint: light aberration





Non-planar lenses create a rainbow effect due to light aberration, which can be used to detect cut & paste tampering

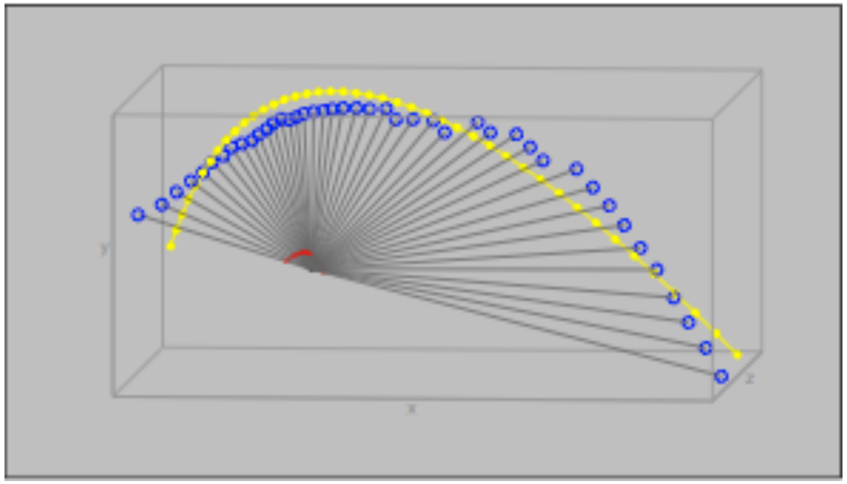
Geometric fingerprint: shadows

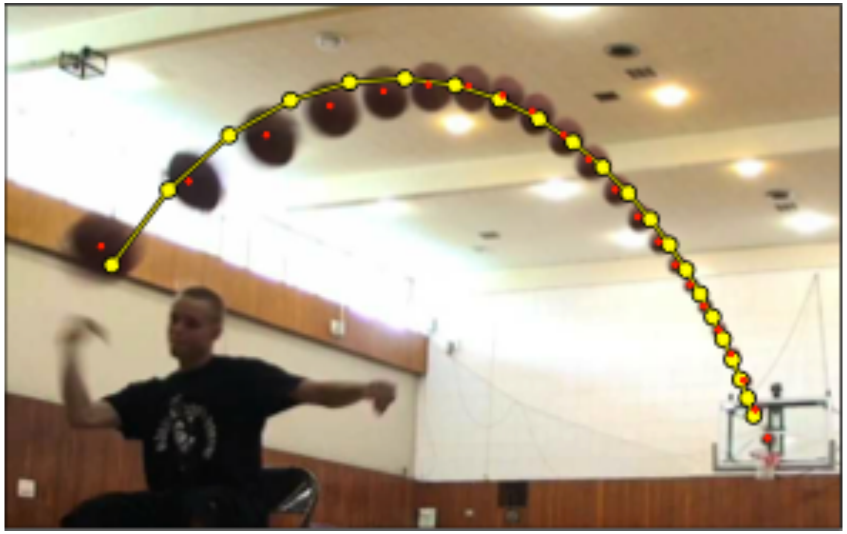


Geometric fingerprint: video

Geometric fingerprint: video

- After perspective compensation we can reconstruct the 3D trajectory of the ball and compare it against the expected trajectory according to physics
- Comparing the expected and apparent trajectories we can deduce that the video is **FAKE !!**





AI from threat to defense

- AI (deep learning) capabilities can be exploited to identify image source, detect processing operations, tampering detection
- Significant advances made in last 5-6 years by applying CNN architectures for forensics
- Race of arms between AI and AI
- Vulnerability to intentional informed attacks
- … work in progress