

Multimedia forensics

an Overview
of Recent Advances and Open Issues

Giulia Boato
University of Trento

MuVer 2017

CNN Today



Related Article: Trump releases some, but not all, JFK assassination records



Donald J. Trump ✓

@realDonaldTrump

Follow

JFK Files are being carefully released. In the end there will be great transparency. It is my hope to get just about everything to public!

4:38 AM - Oct 27, 2017

10,638 9,478 41,466



Seeing is not believing



EARLY PHOTOGRAPH

Seeing is not believing



=

+

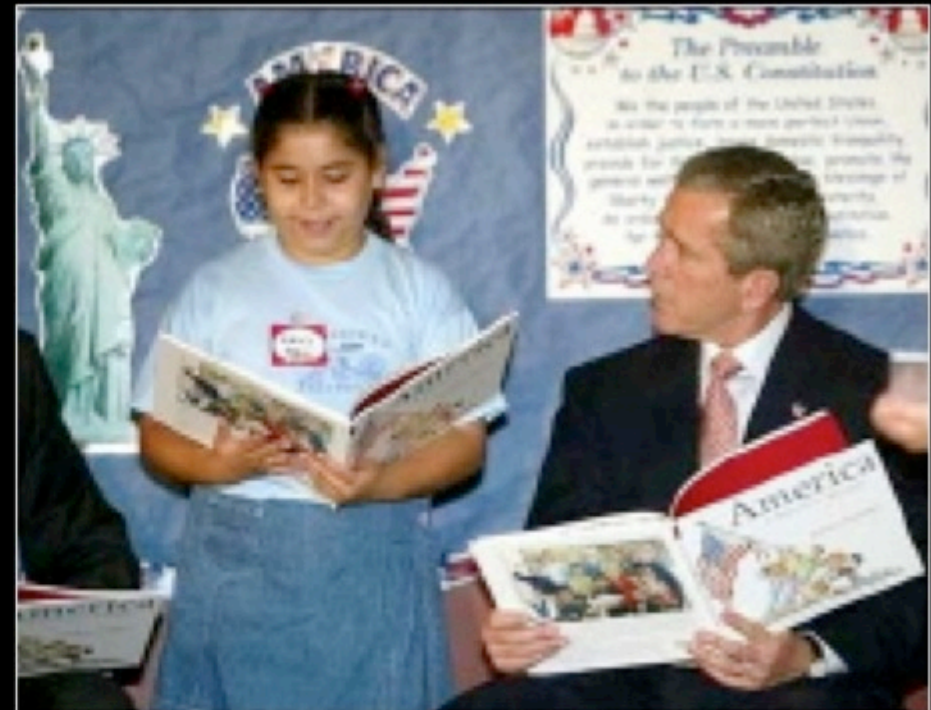
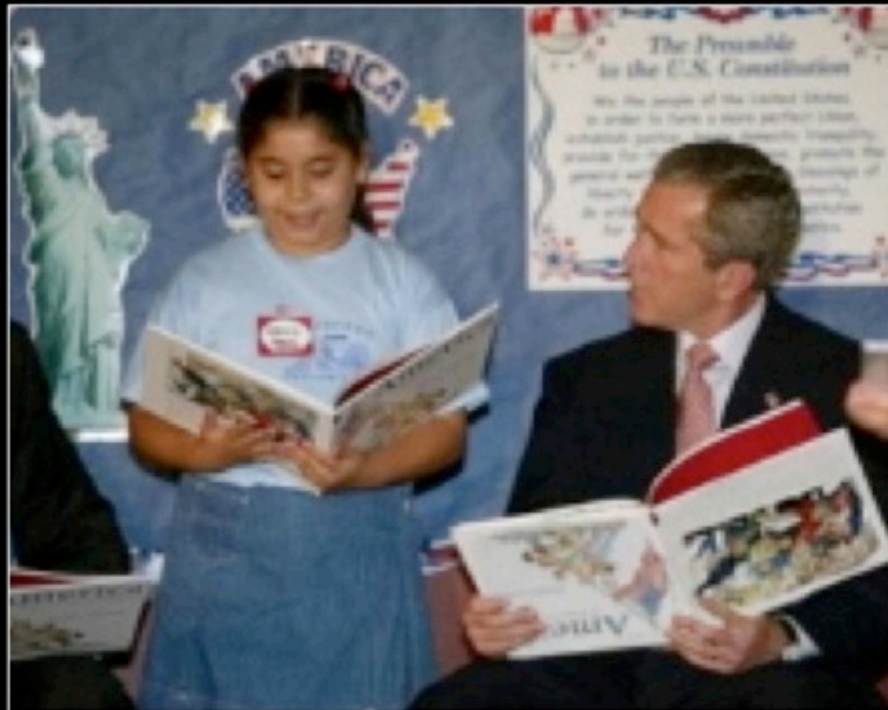


+



EARLY PHOTOGRAPH

Seeing is not believing



POLITICS

Seeing is not believing



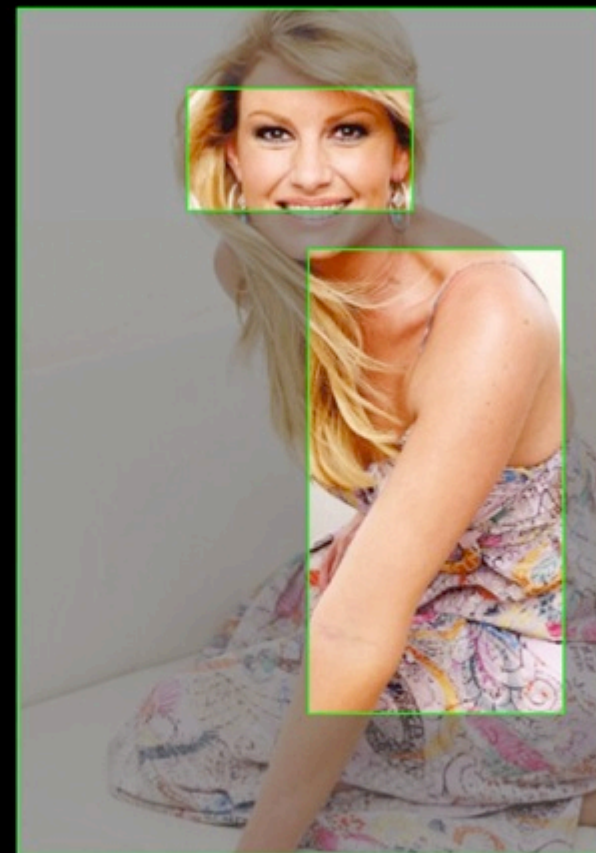
POLITICS

Seeing is not believing



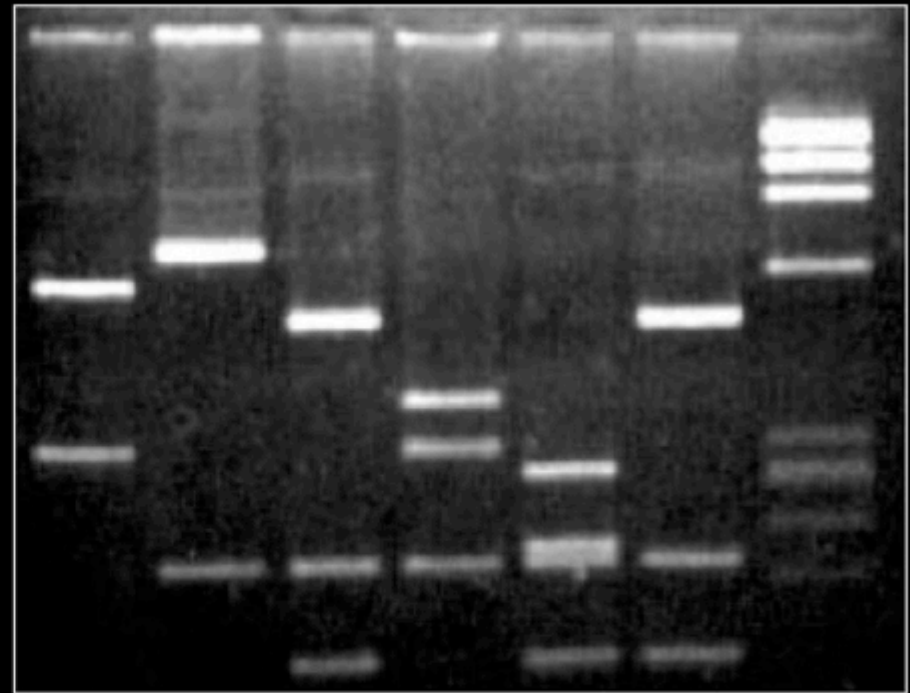
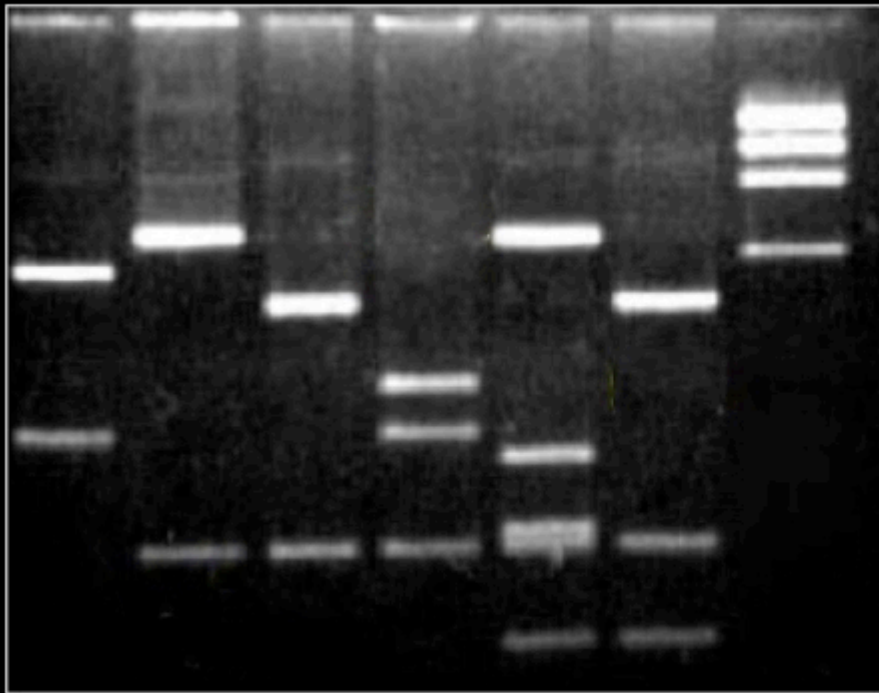
MAGAZINES

Seeing is not believing



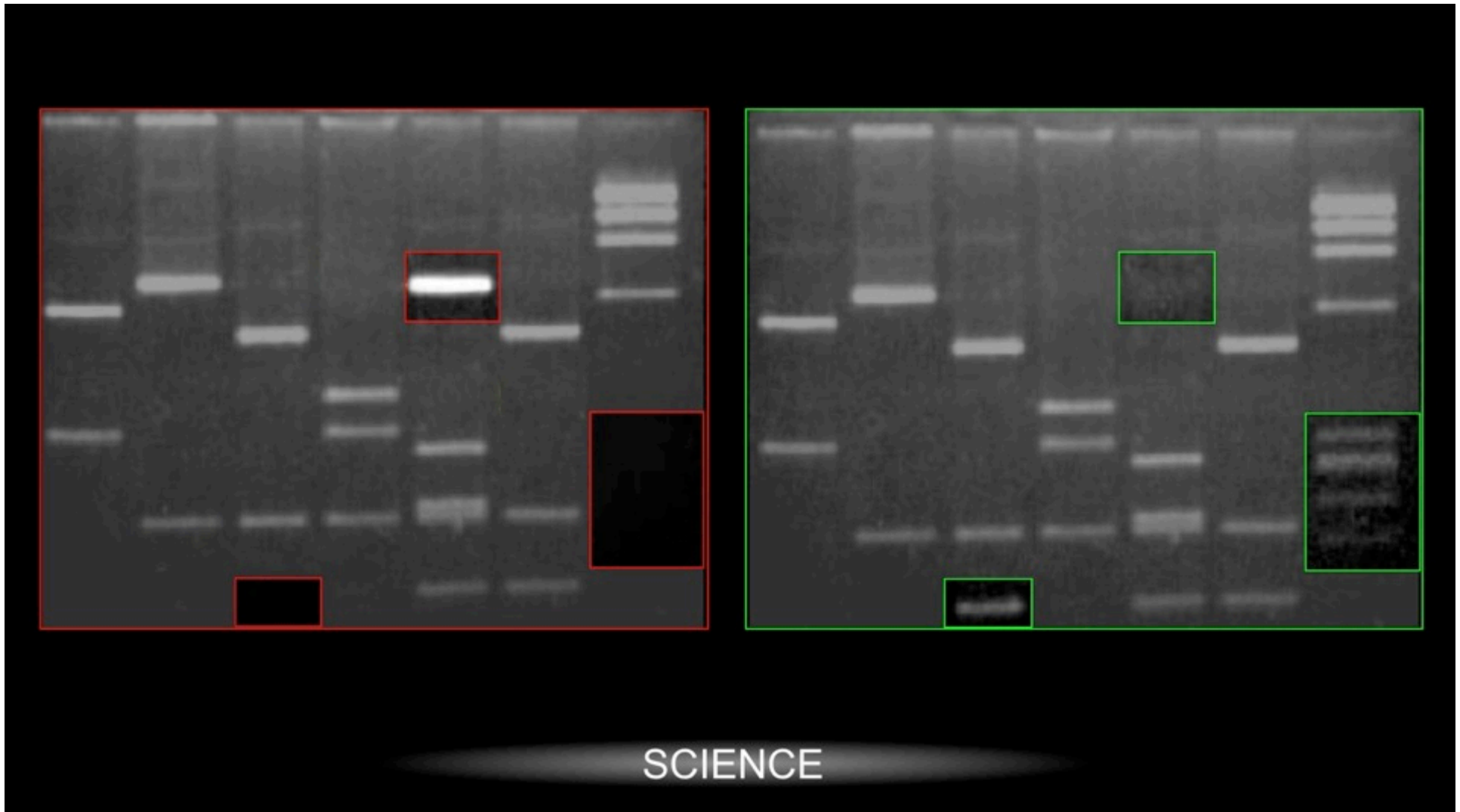
MAGAZINES

Seeing is not believing



SCIENCE

Seeing is not believing



Seeing is not believing



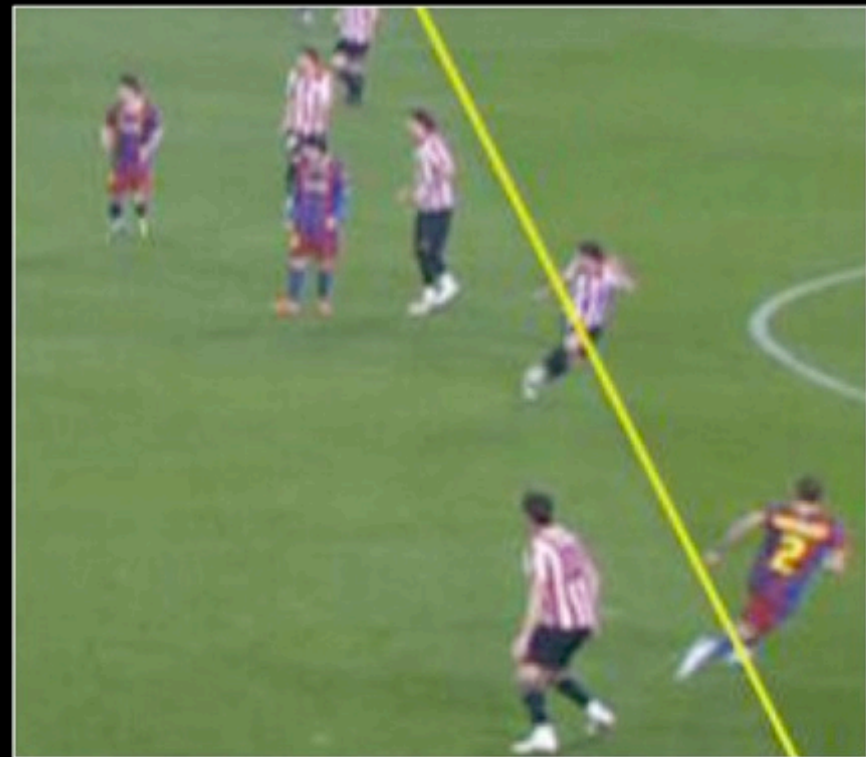
NATIONAL SECURITY

Seeing is not believing



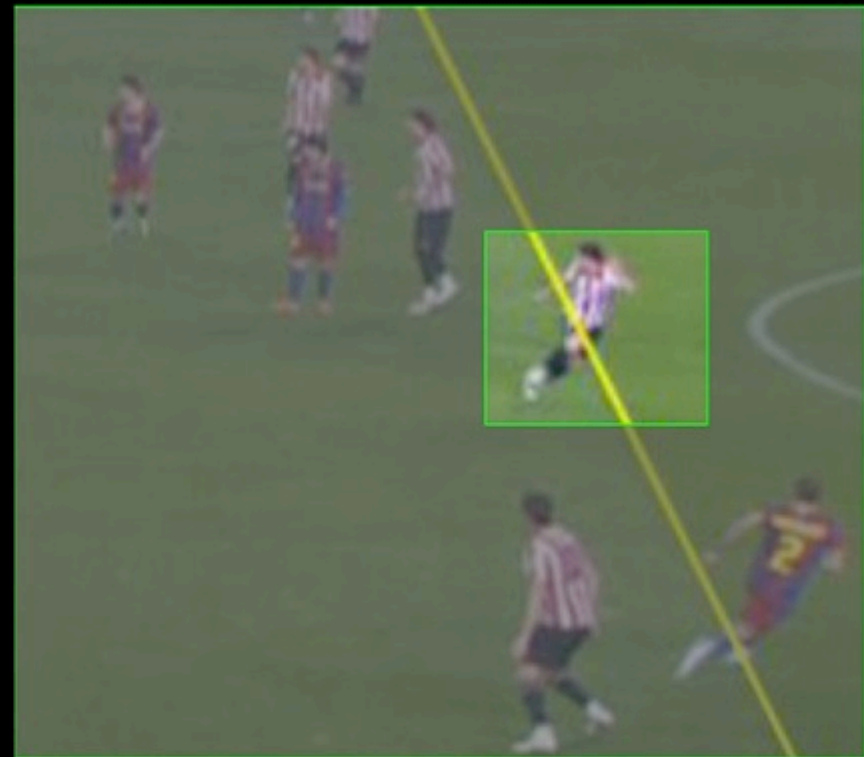
NATIONAL SECURITY

Seeing is not believing



SPORT

Seeing is not believing



SPORT

Seeing is not believing



VIDEO

Seeing is not believing



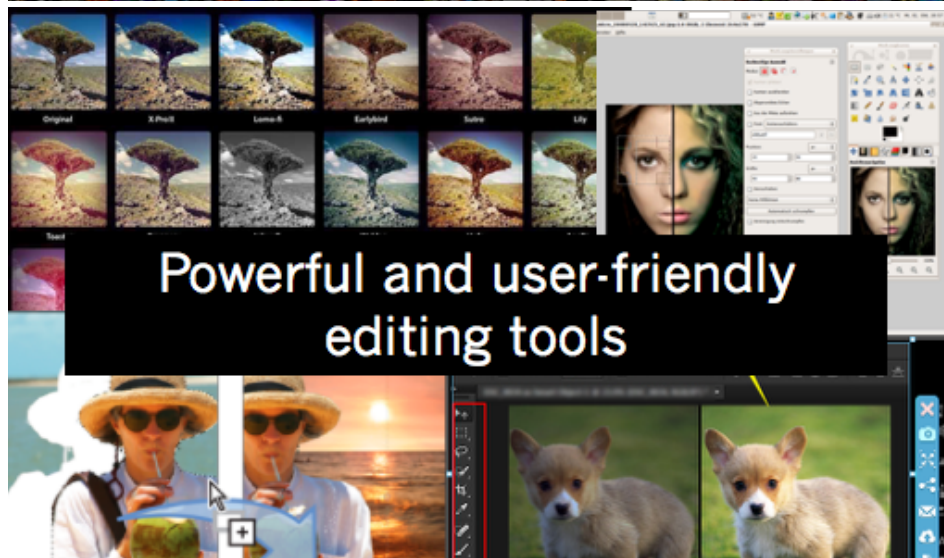
VIDEO

Multimedia forensics – why



Fast and large-scale propagation of modified multimedia content, potentially carrying a distorted semantic message

Multimedia forensics – why



Multimedia forensics – why



Forensic investigation
Journalism, photojournalism
Communication, social media



Multimedia forensics – what

Identify the device that took a particular multimedia data

Determine possible tampering

Discriminate multimedia objects generated by real cameras, scanners, or via CG softwares

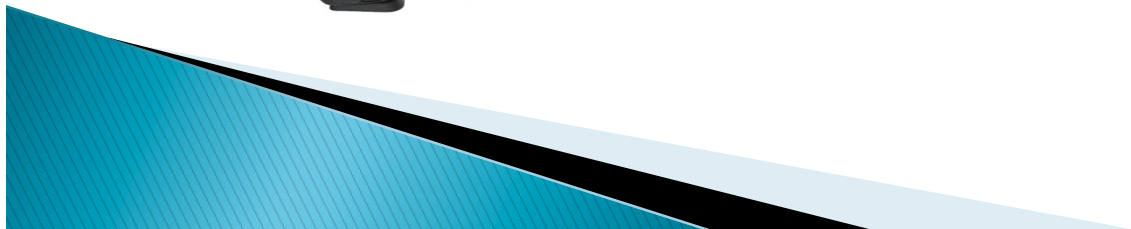
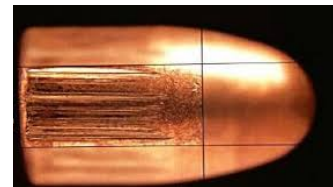


Multimedia forensics – what ... keeping in mind the CSI effect



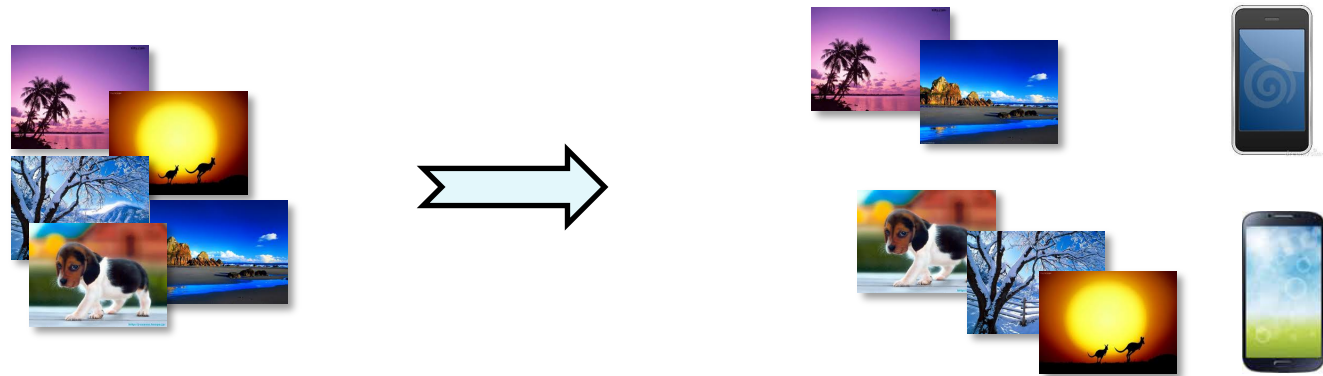
Source identification

Every time a device acquires a multimedia object it leaves imperceptible traces on it

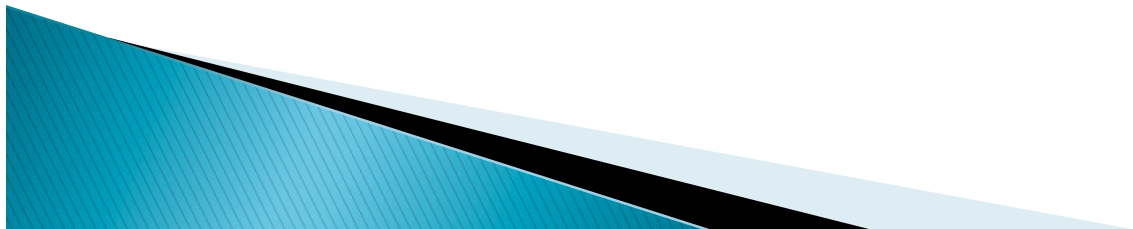


Source identification

- ▶ In general, **source identification** problem links multimedia content to a particular (class of) acquisition device(s).



- ▶ It can be solved at different levels ...



Source identification

▶ Level 1

- Which Device ?

Which Coolpix P60?



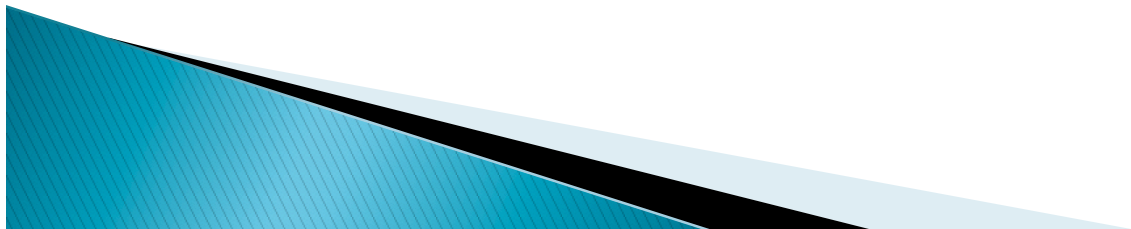
Serial Number
000111101



Serial Number
000111103



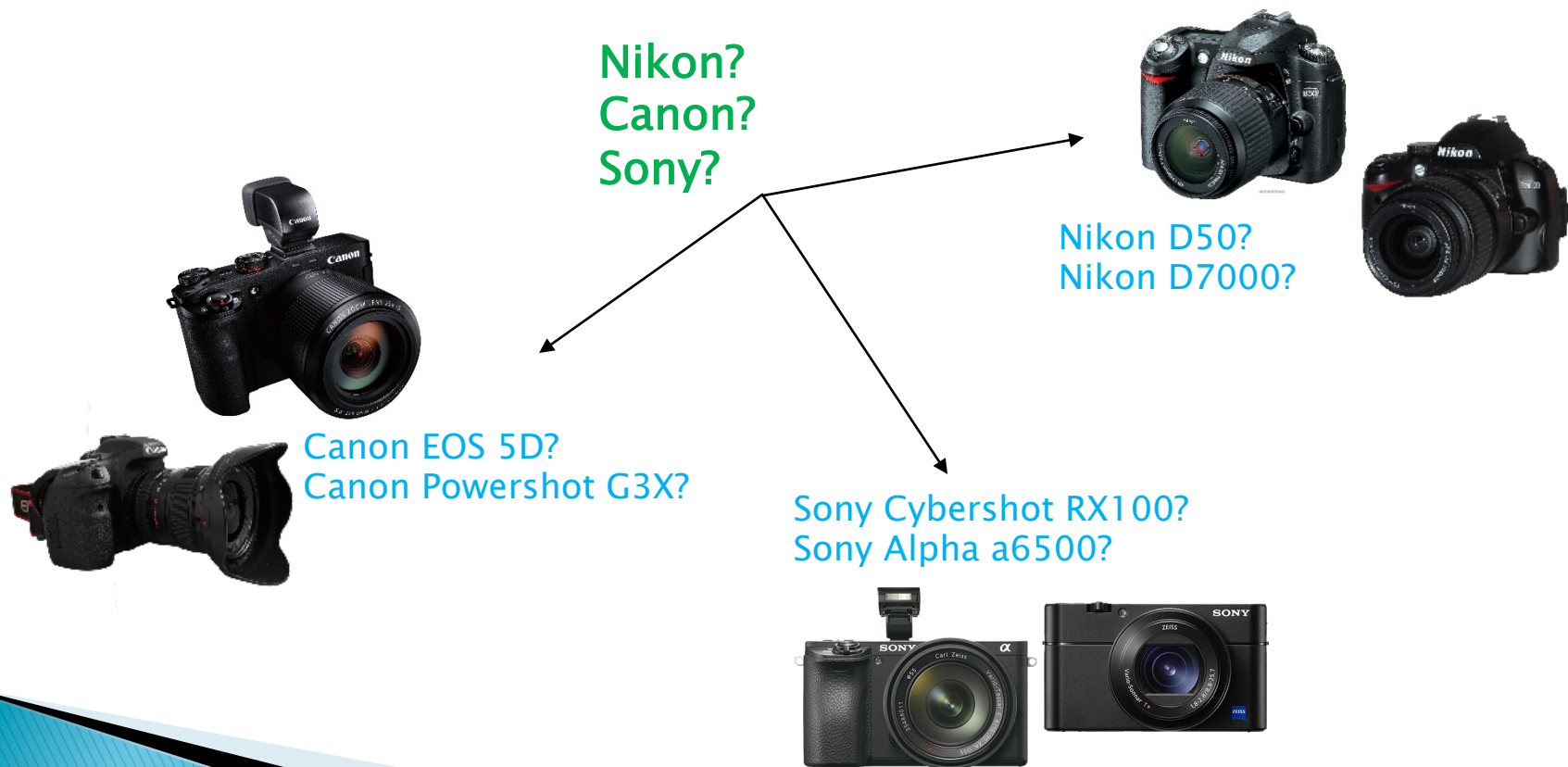
Serial Number
000111106



Source identification

▶ Level 2

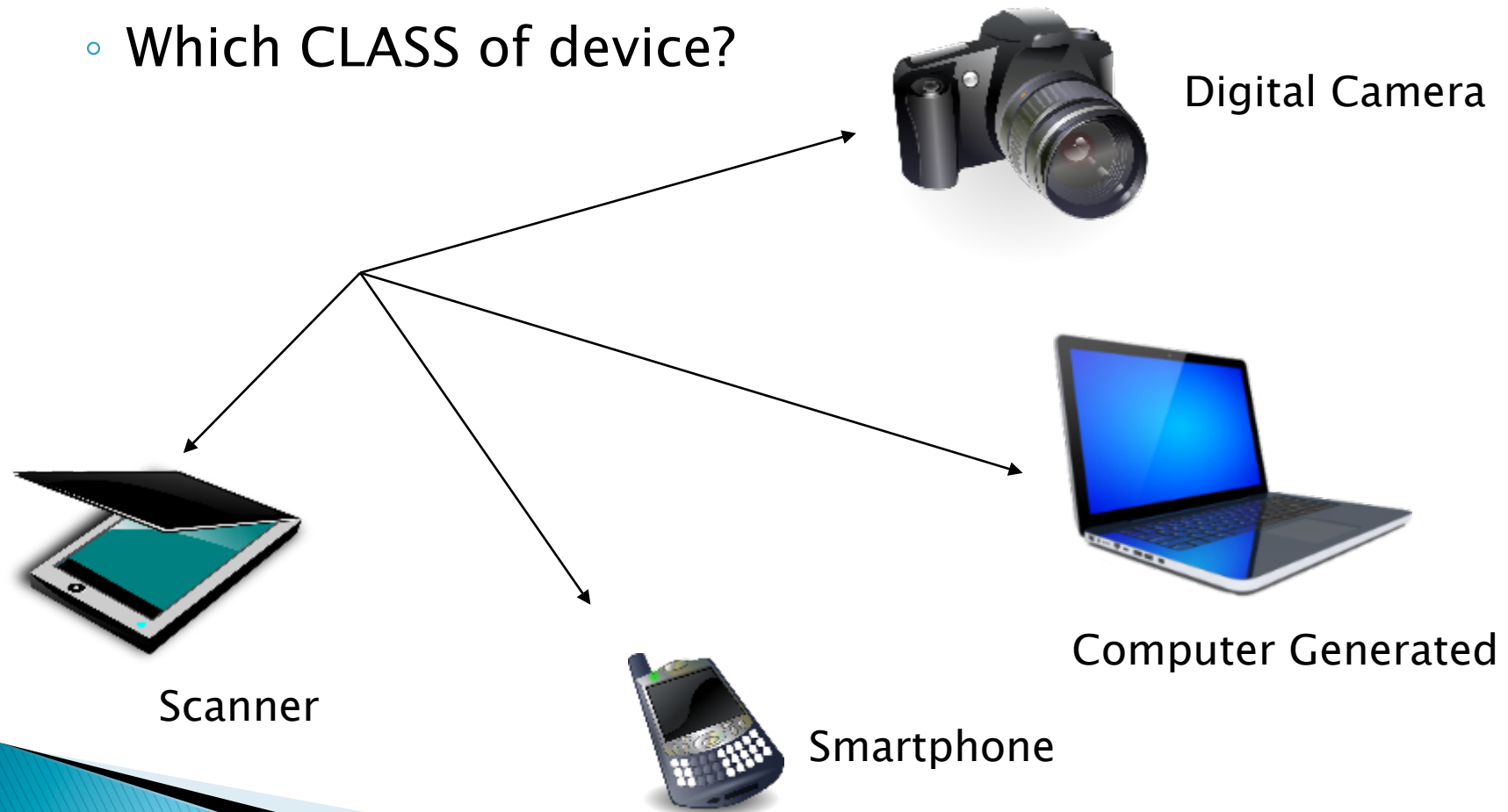
- Which BRAND/MODEL?



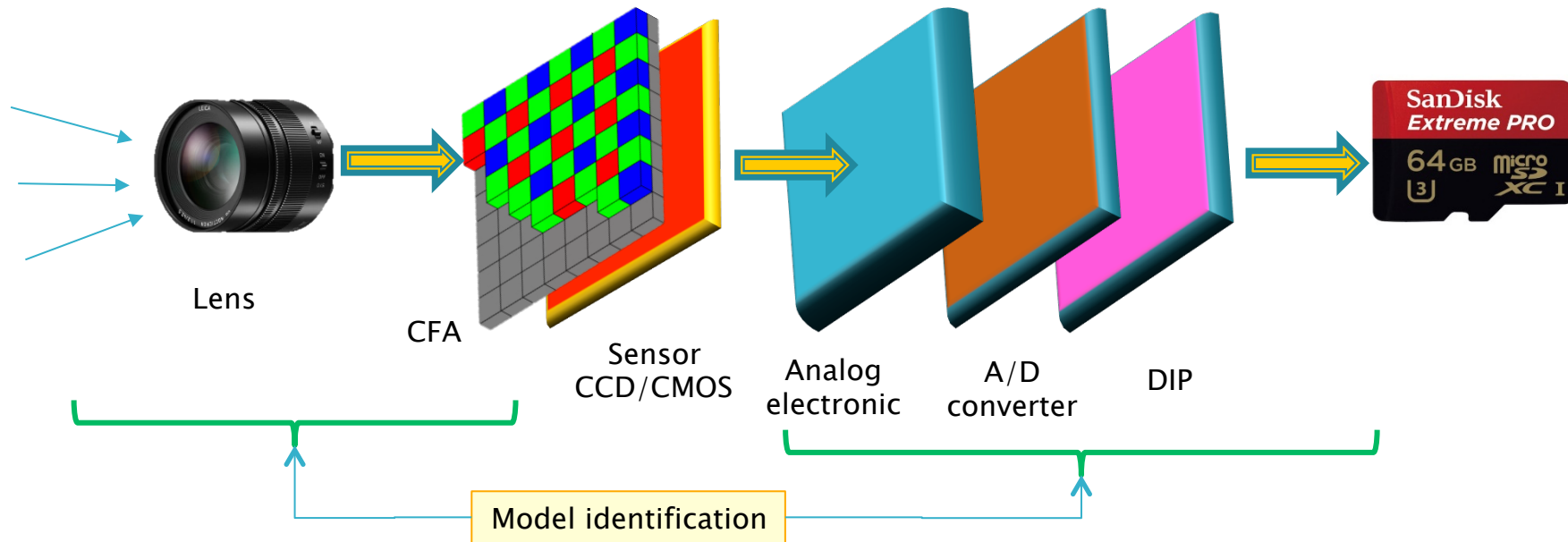
Source identification

▶ Level 3

- Which CLASS of device?



Source identification – device



- ▶ Sensor CCD/CMOS presents intrinsic imperfection
- ▶ It leaves a systematic noise on each acquired image (video)
- ▶ Such noise is imperceptible

Source identification – device

- ▶ Such noise has two components



Fixed Pattern
Noise (FPN)

Dark current (usually
compensated directly
in the device)

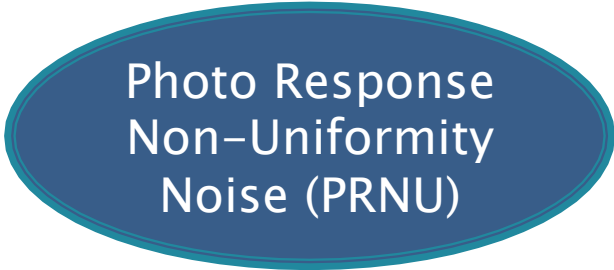
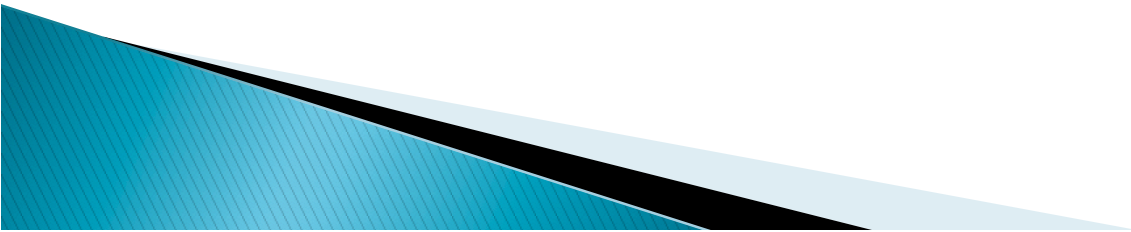


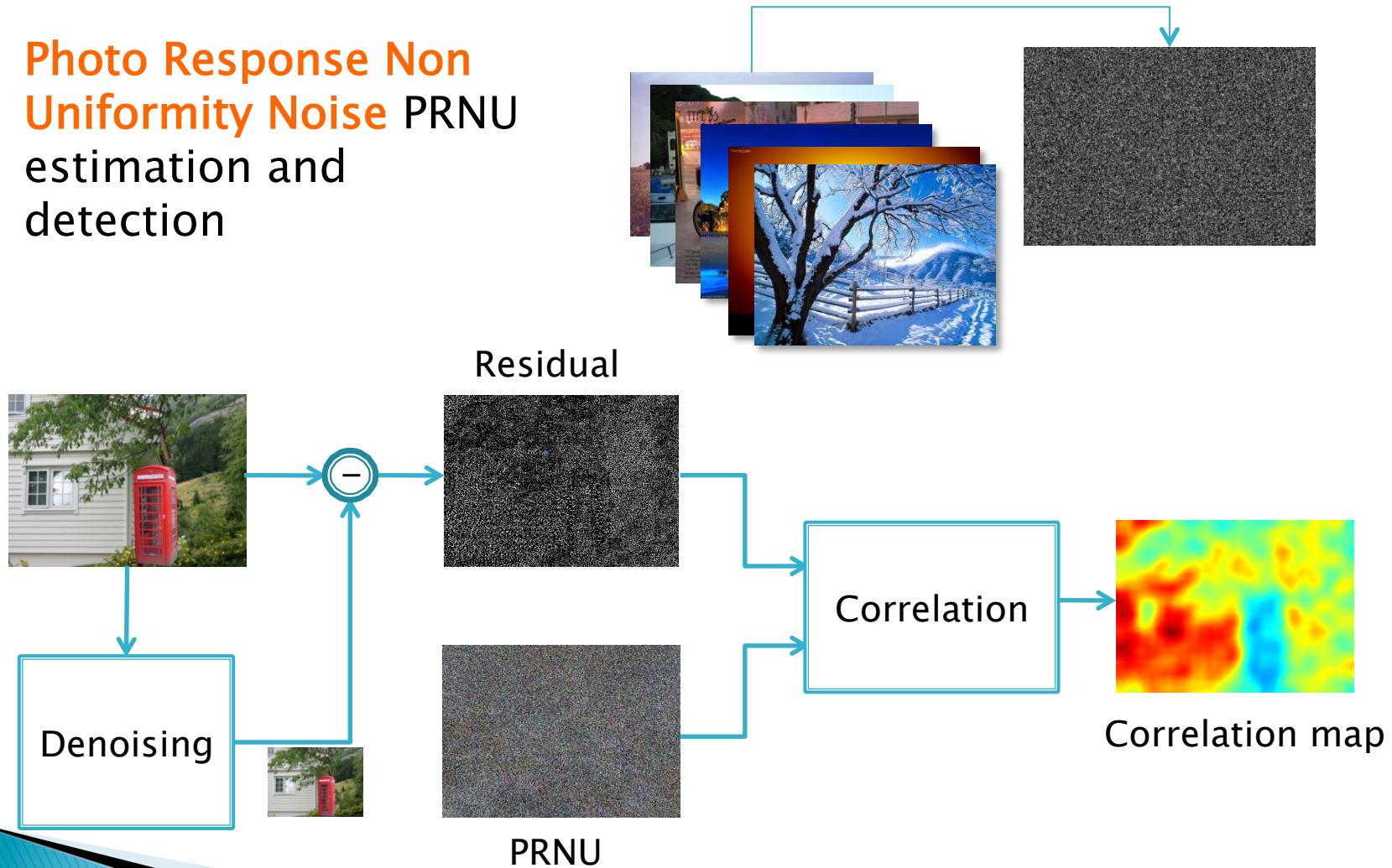
Photo Response
Non-Uniformity
Noise (PRNU)

Not uniform response to
the same light impulse



Source identification – device

- ▶ **Photo Response Non Uniformity Noise** PRNU estimation and detection



Example of application: photo on Facebook



BOB

BOB

BOB

This block represents the user 'BOB'. It includes a grey silhouette profile picture, a grid of various photos, and a black and white fingerprint scan. The name 'BOB' is written above the profile picture and to the right of the fingerprint.

ALICE



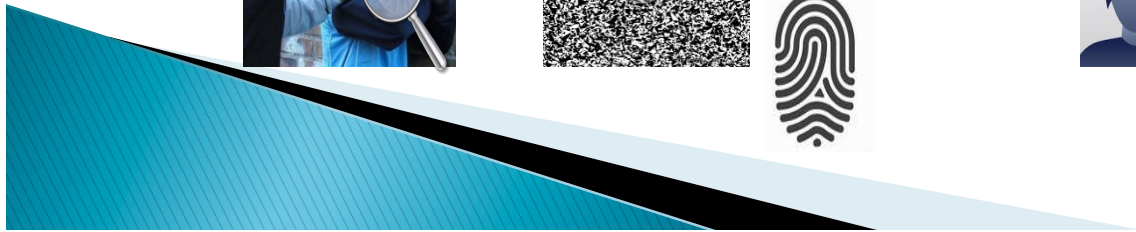
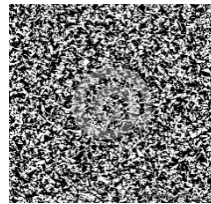
ALICE

JOHN



JOHN

Who is the author of this image?



Open issues

▶ **Robustness**

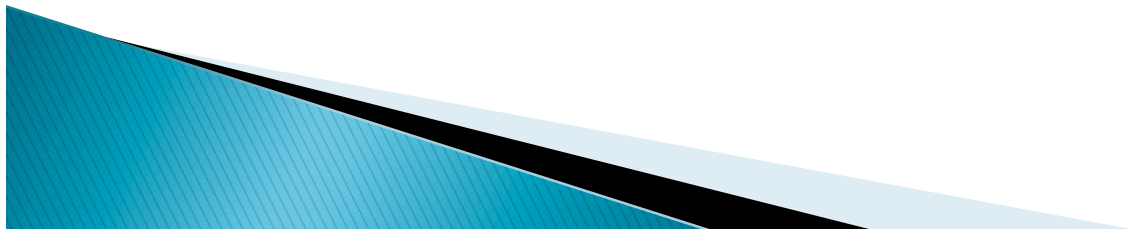
- JPEG compression
- Different resolutions
- Various processing

▶ **Security**

- Fingerprint cancellation or replacement
- Image quality preserved

▶ **Computational complexity**

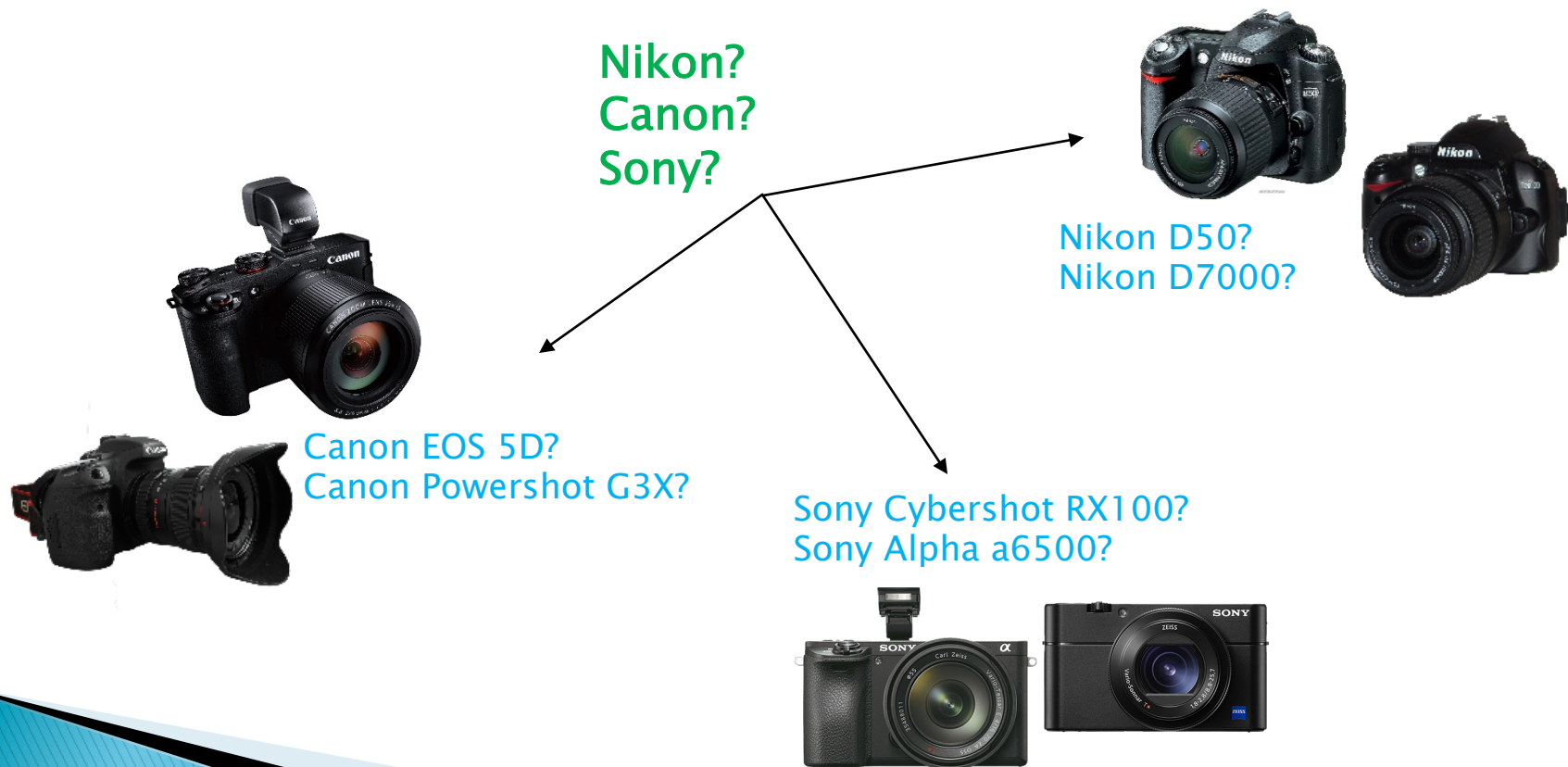
- Dataset of fingerprints
- Fingerprint digest



Source identification – model

▶ Level 2

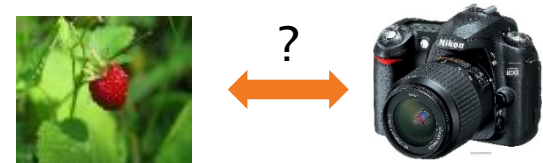
- Which BRAND/MODEL?



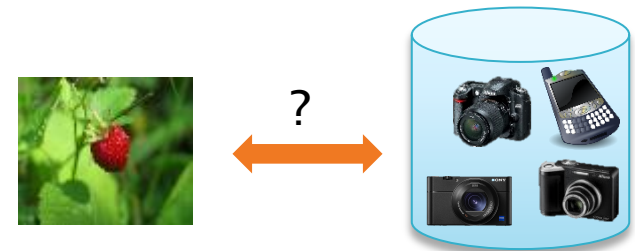
Source identification – model

▶ Several applications:

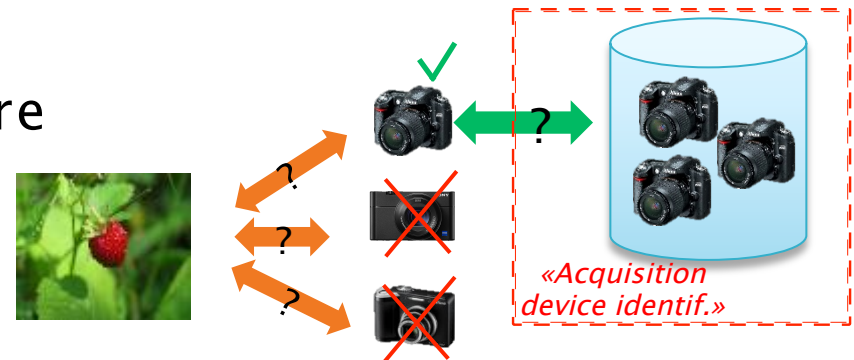
1. Check compatibility w.r.t an alleged provenance device (even if it's not available!)



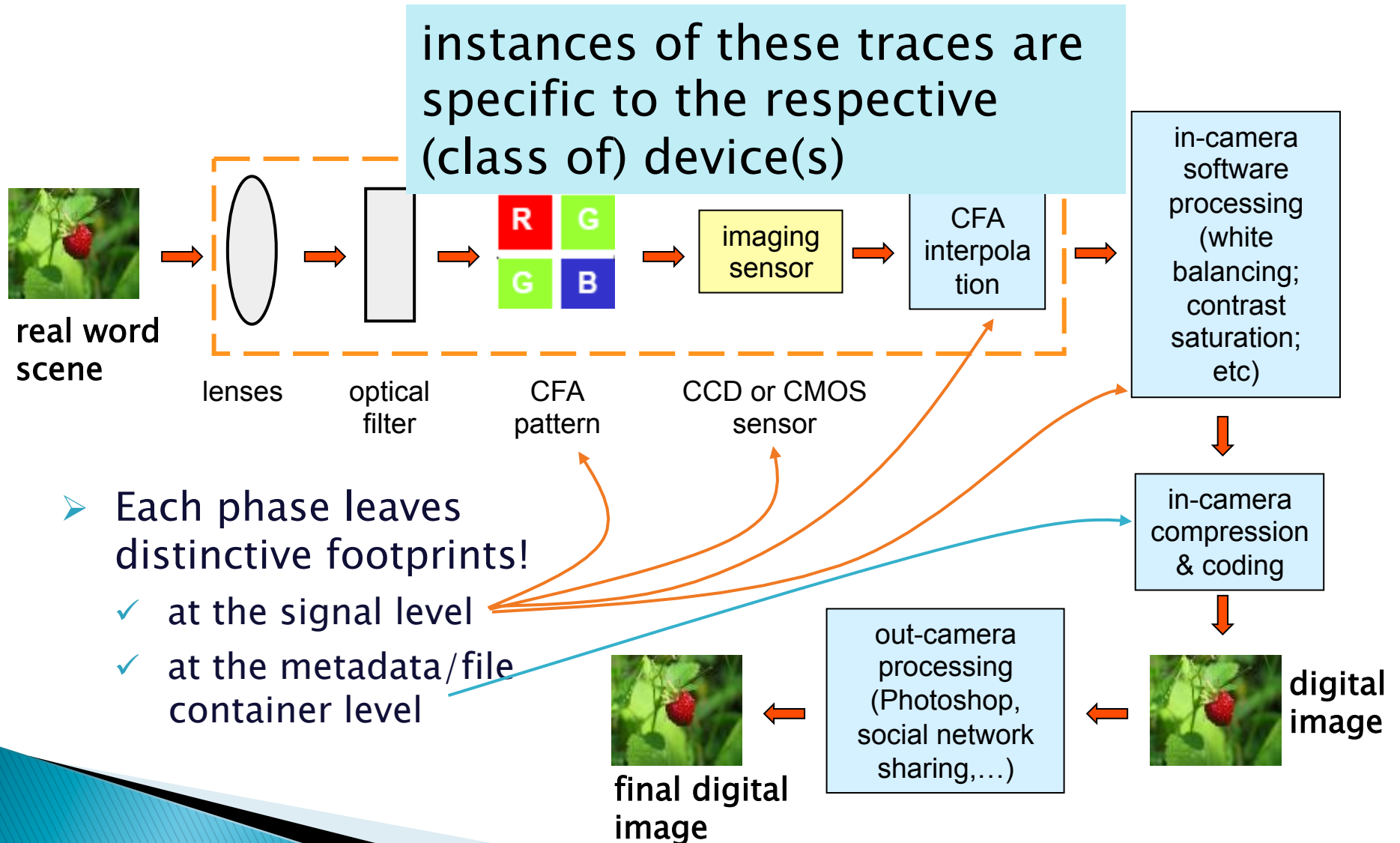
2. Trace back a picture to the kind of device that captured it



3. Pre-filtering step before more specific analyses

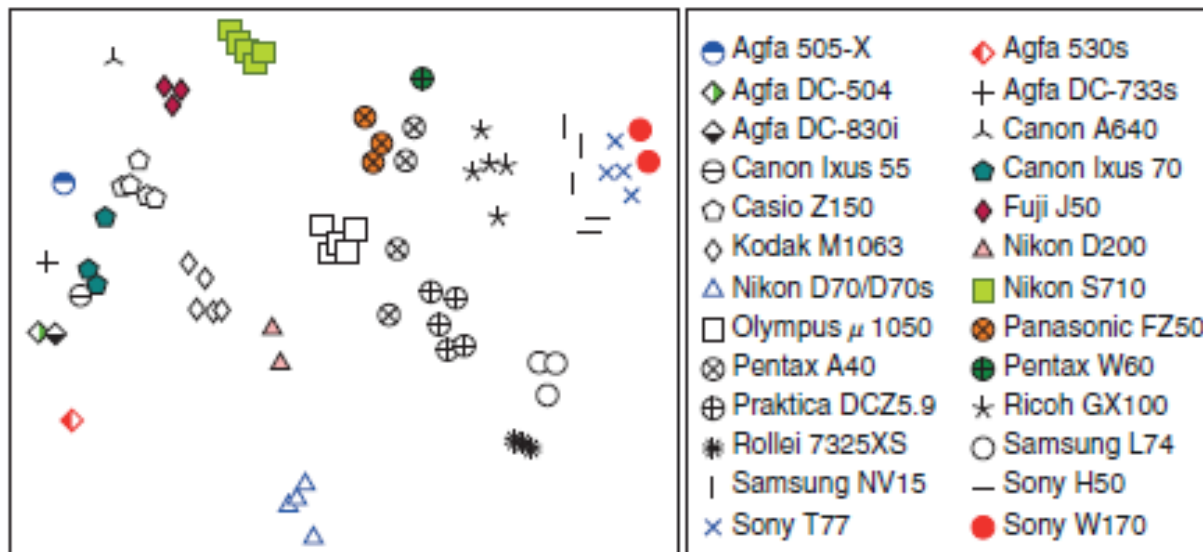


Source identification – model



Scenarios – Perfect knowledge

- ▶ Finite set of known camera models
 - for each model we have enough training images to carry out reliable estimates of features of interest
- ▶ We can use a multiclass classifier and will be able to compute a full confusion matrix



Scenarios – Limited knowledge

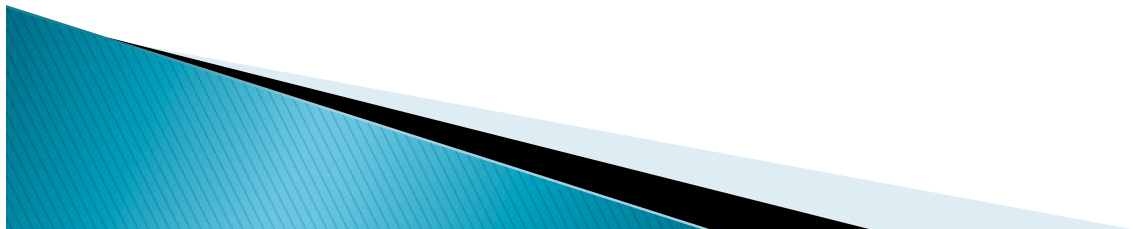
- ▶ Full knowledge on target model, but know nothing about the number of other models
- ▶ Practically, we have a large number of training images, classified as either belonging to the target model or another (unknown) model
 - Here, we can use only a one-class classifier



Other models

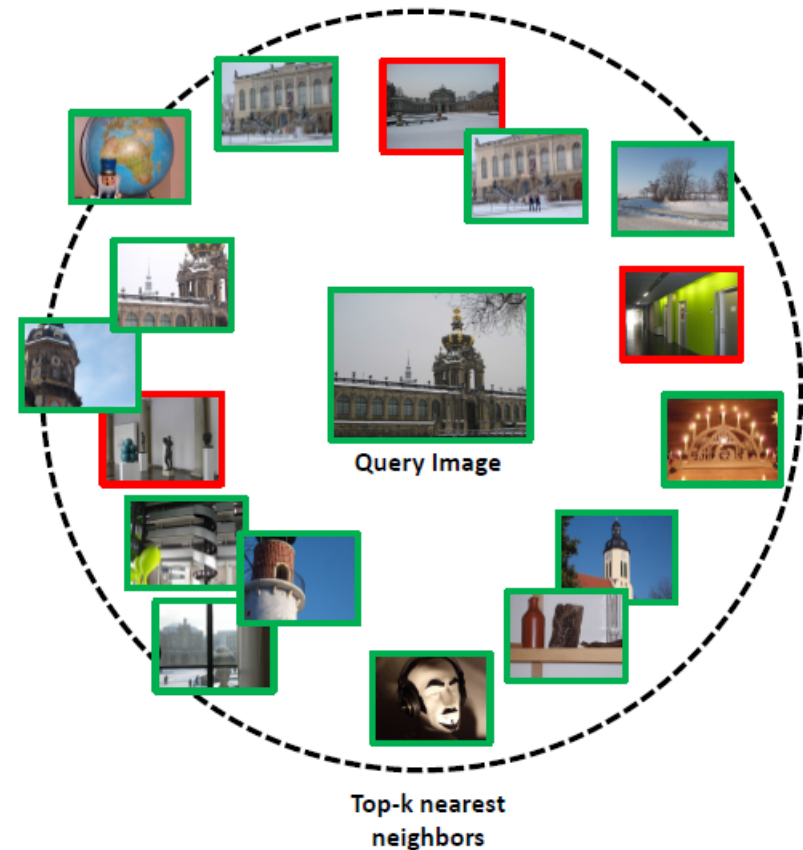


Query image model



Scenarios – Zero knowledge

- ▶ No prior information on the number of camera models
- ▶ There are only a large number of images
- ▶ We can only retrieve other images taken from the same model, which may help for subsequent investigations



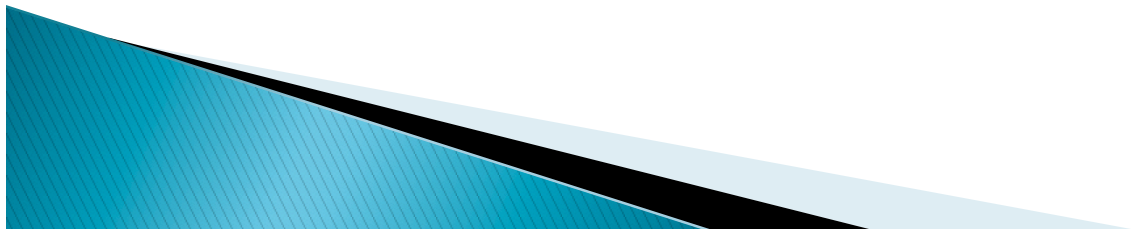
Source identification – model

Signal level

- Artifacts related to specific in-camera processing steps:
 - Lens, CFA, In-camera processing
- General intrinsic features, irrespective of their physical meaning (deep learning)

Format level

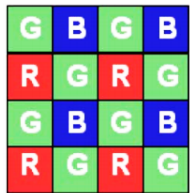
- Codec
- Metadata
- File container



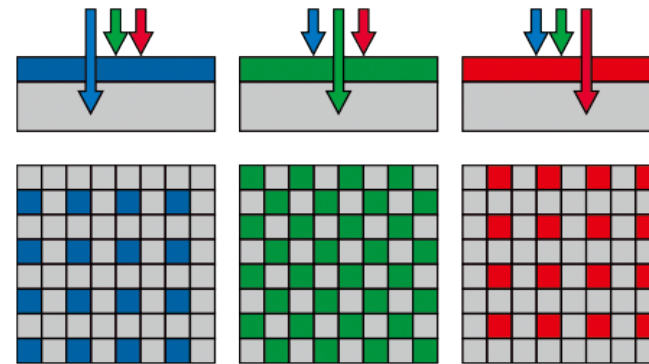
Signal-level Traces: CFA-based

▶ Colour Filter Array

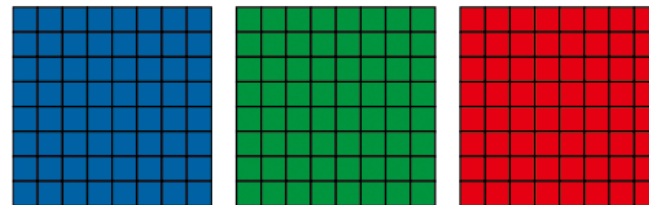
- A thin film that selectively permits a certain component of light to pass through it to the sensor.
- In practice, to each pixel only one particular colour is gathered.



- ▶ The sensor output is successively interpolated to obtain all the 3 colours for each pixel (**demosaicking**)

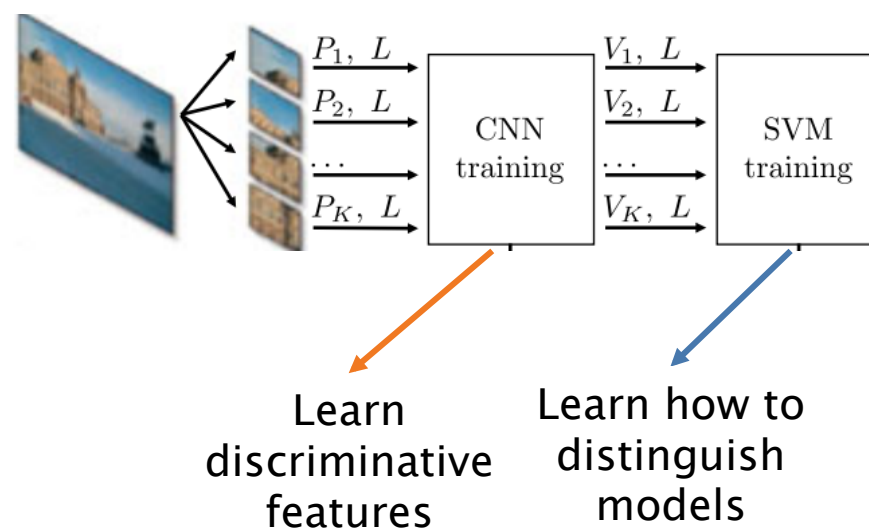
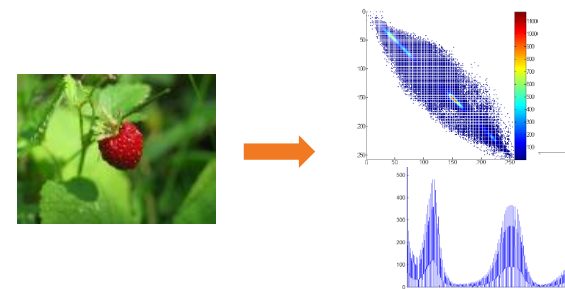


CFA interpolation
(demosaicking)



Signal-level Traces: deep learning

- ▶ Data Driven Approach: learn **intrinsic footprints** directly from images, rather than imposing any model or handcrafted feature recipes
- ▶ Convolutional Neural Networks (CNNs) proved to be an excellent tool to learn brand/model distinctive traces
 - Learn from pixels of image patches, no need to devise complex features
 - Learned CNNs generalize well to new brand/models



Format-level Traces: JPEG

- ▶ Basic structures of JPEG file formats are marker segments
 - occurrence and sequence of these segments differ

- ▶ A key parameter is the **quantization table**
 - different brands/models usually employ different tables

- ▶ **EXIF metadata** including camera properties and employed image acquisition settings

marker id	short value	JIF	JFIF	EXIF	description
SOI	0xFF D8	×	×	×	start of image
APP _n	0xFF E _n				application data
APP0	0xFF E0		×		(e.g., JFIF)
APP1	0xFF E1			×	(e.g., EXIF)
DQT	0xFF DB	×	×	×	quant. tables
DHT	0xFF C4	(×)	×	×	Huffman tables
SOF	0xFF C _n	×			start of frame
SOF	0xFF C0		×	×	(baseline DCT)
SOS	0xFF DA	×	×	×	start of scan
DRI	0xFF DD				restart interval
RST _n	0xFF D _n				<i>n</i> th restart
COM	0xFF FE				comment
EOI	0xFF D9	×	×	×	end of image

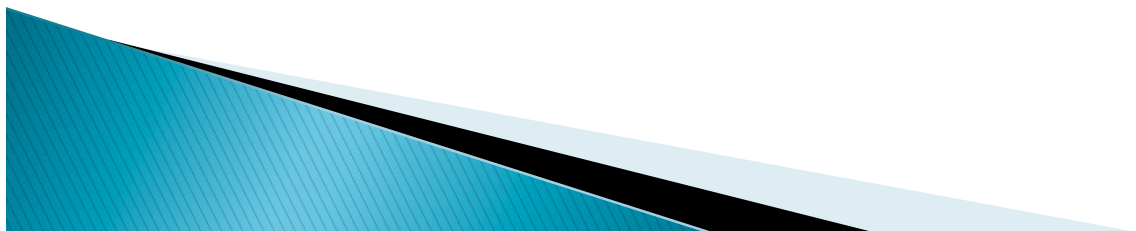
$Q =$

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	60	63	64
18	22	37	56	68	76	80	81
24	35	55	64	81	90	95	96
49	64	78	87	103	112	117	118
72	92	95	98	112	121	126	127

Entry	Meaning
Image	
• Make	Canon
• Model	Canon EOS 450D
• Orientation	top/left
• X Resolution	72
• Y Resolution	72
• Resolution Unit	inch
• Date Time	2011-02-27 16:14:23
• YCbCr Position...	co-sited
• Exif IFD Pointer	Offset: 196
Camera	
• Exposure Time	1/60"
• F Number	F5.6
• Exposure Prog...	Portrait mode
• ISO Speed Rati...	400
• Exif Version	Version 2.21
• Date Time Orig...	2011-02-27 16:14:23

Open issues

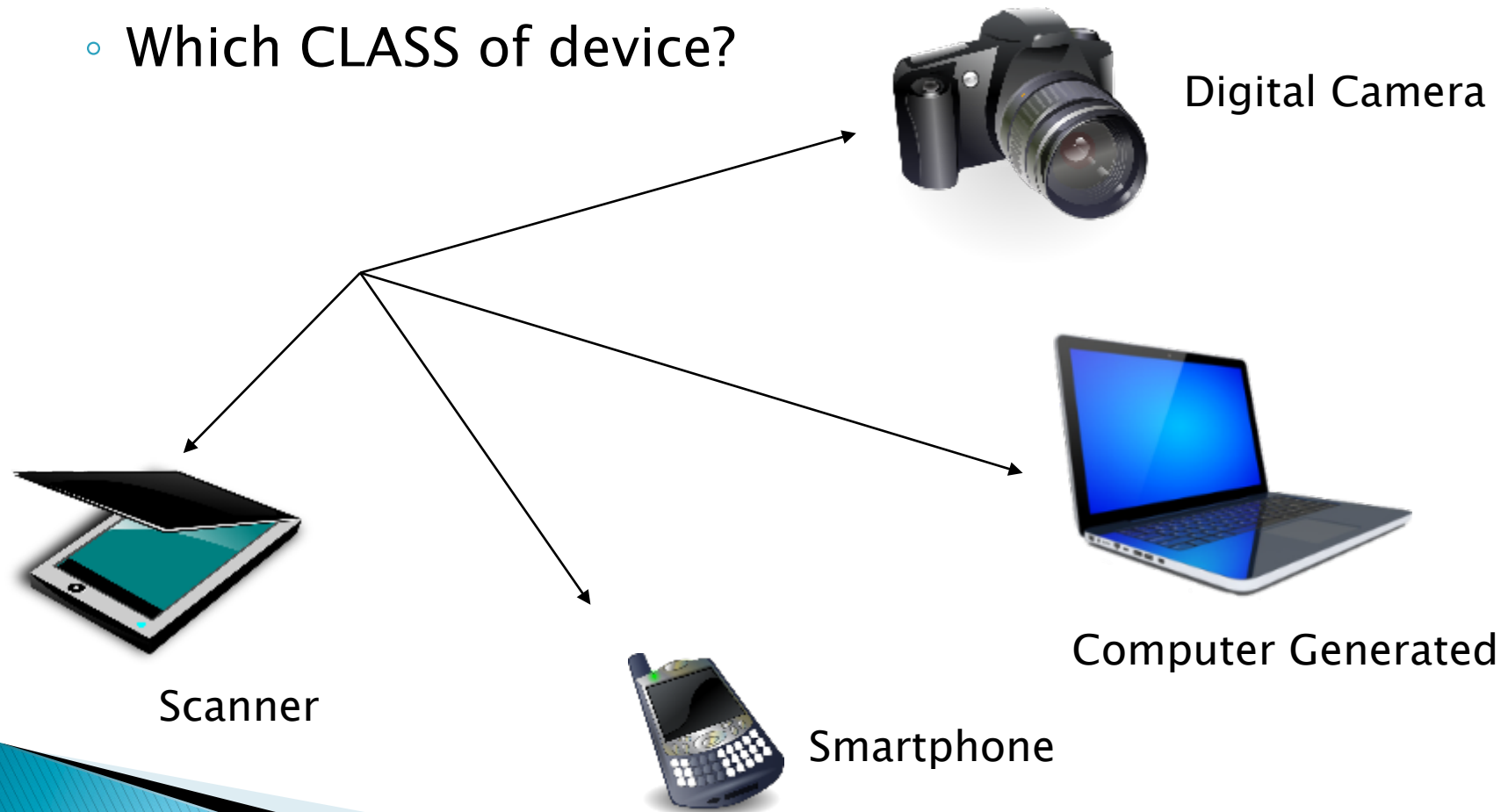
- ▶ Most of solutions are tested on Perfect knowledge scenarios only
 - What if the content comes from **unknown devices**?
- ▶ How to deal with a **new device**
 - Classifier needs re-training for adding new models
- ▶ Camera identification for **videos** is still unexplored
 - While they are becoming the most diffused information



Source identification – class

▶ Level 3

- Which CLASS of device?



Source identification – class

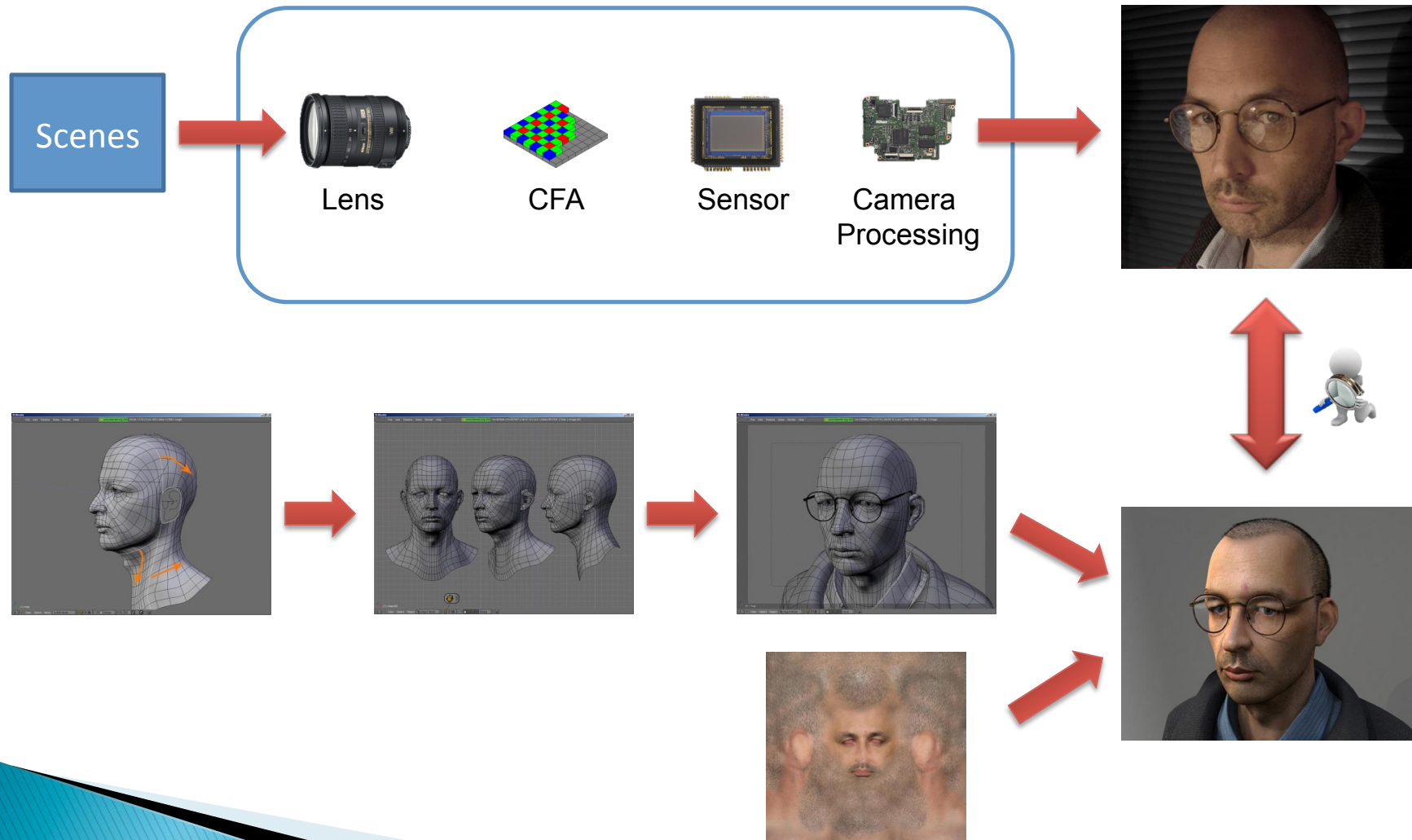
- ▶ Given that the **increasing photorealism of computer generated (CG) imagery** created by sophisticated 3D graphic tools, the task to differentiate between real and CG images is nowadays challenging at only visual inspection



CG faces discrimination



CG faces discrimination




CG faces discrimination

Faces in images

- Statistics on wavelets
- Face symmetry
- Repetitive expressions

Faces in videos

- Repetitive expressions
 - Face dynamics
 - Geometric deformations
 - Physiological signals
- 

INPUT

Human Faces
Extraction

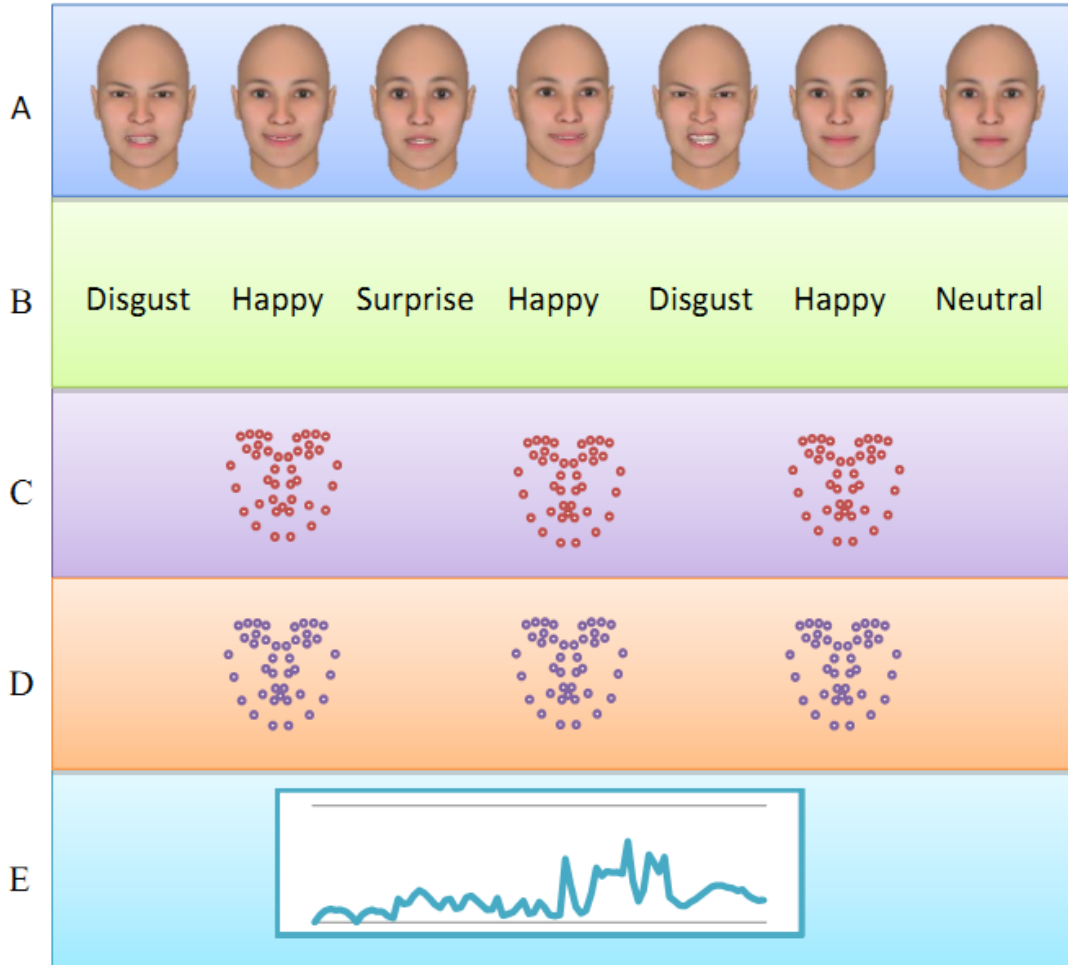
Facial Expression
Recognition

Active Shape Model
Extraction

Normalized
Face Computation

Variation Analysis

RESULT

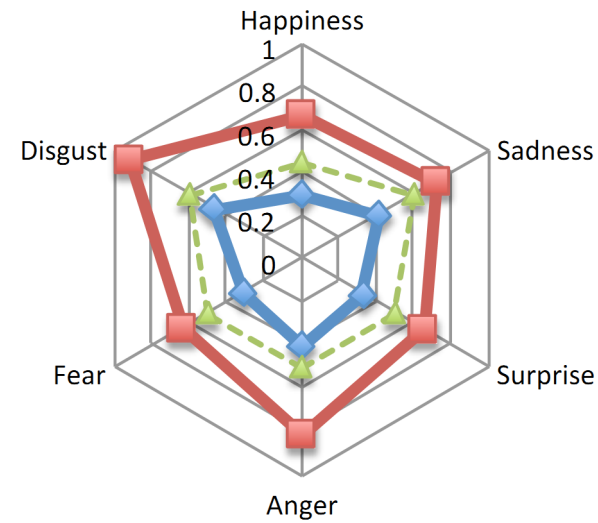
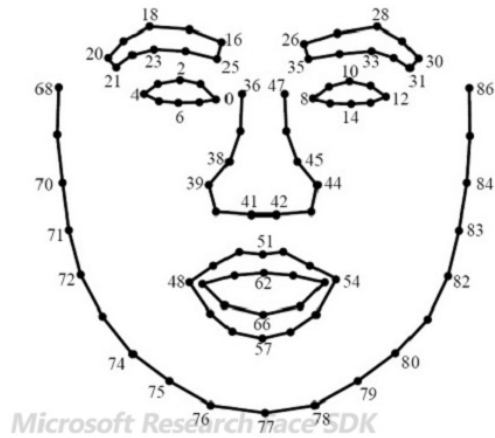


?

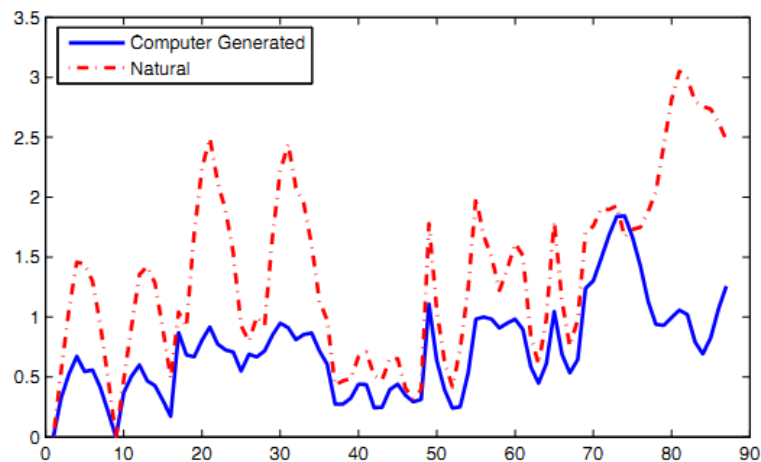


with Dang Nguyen et al. WIFS 2012

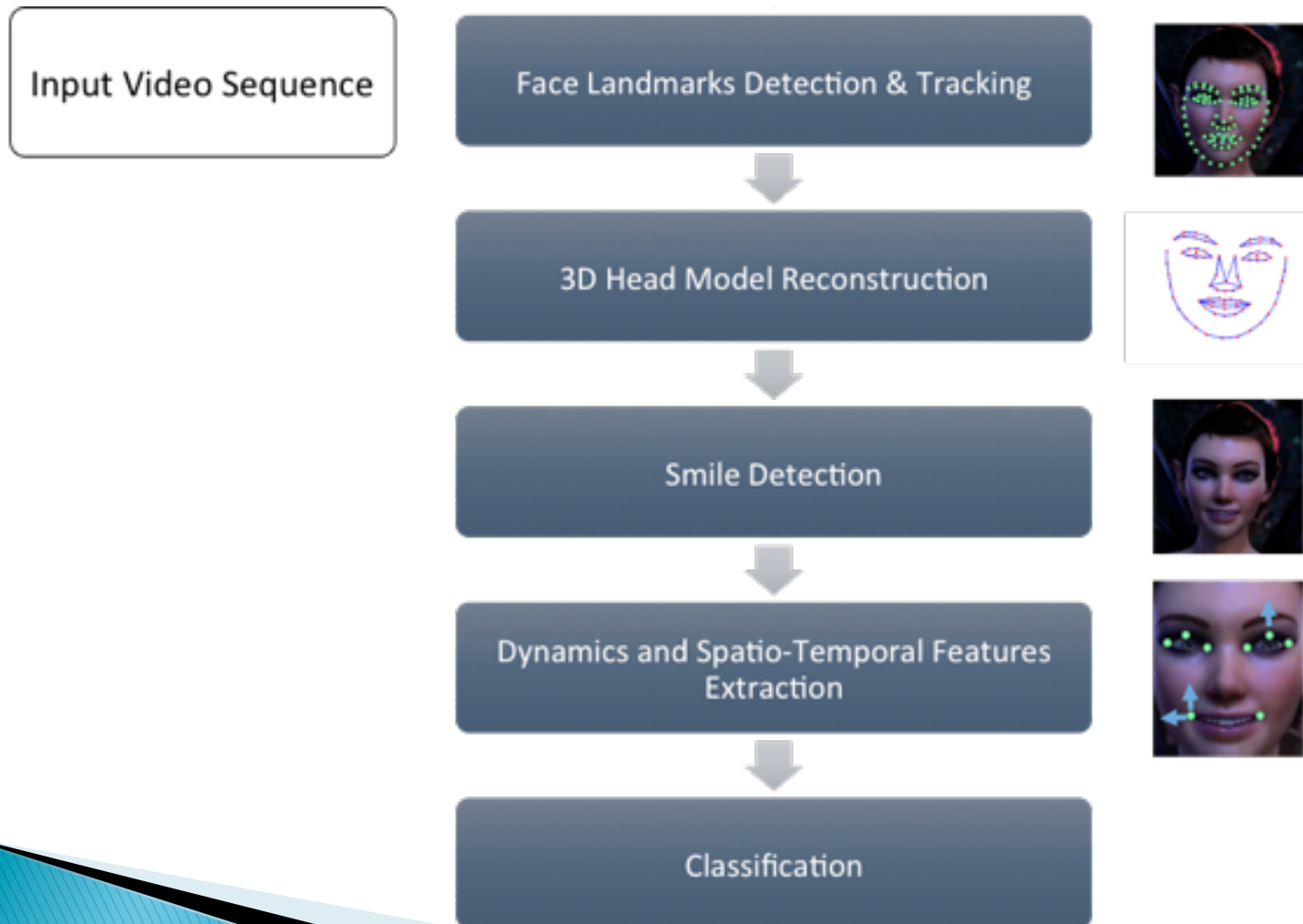
CG faces discrimination: repetitive expressions



—◆— Computer Generated —■— Natural - - -▲- - - Threshold

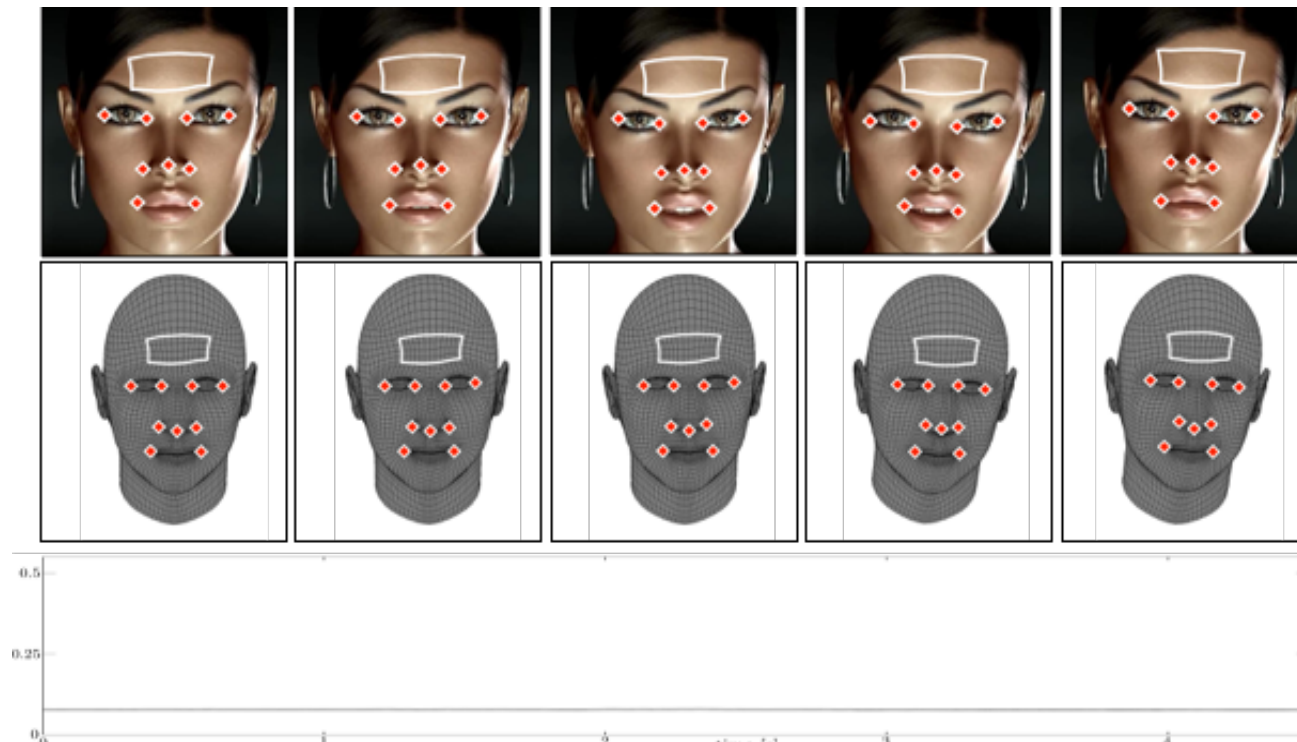


CG faces discrimination: face dynamics



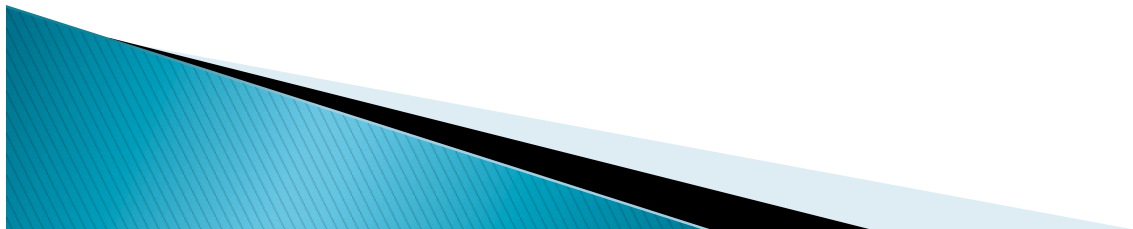
CG faces discrimination: physiological signals

- ▶ Discrimination based on physiological signals, e.g., **heart rate** (estimated from video sequences)



Multimedia forensics – what

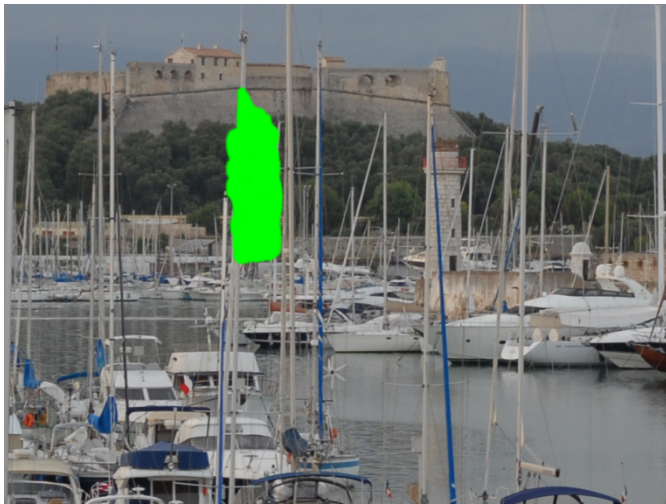
- ▶ Multimedia forensics aims at
 - Determine the source of the multimedia object
 - Device
 - Camera model/brand
 - Class of acquisition devices: camera, scanner, CG
 - **Detect possible tampering** in the multimedia object
 - Global processing
 - Splicing
 - Copy move
 - Local tampering



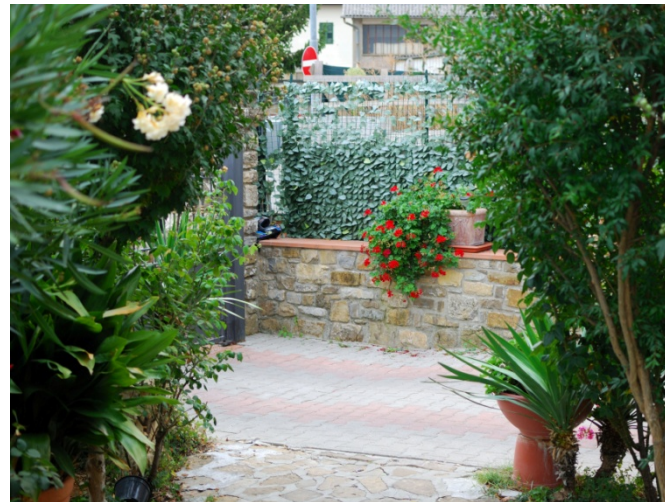
Tampering detection



Tampering detection



Tampering detection

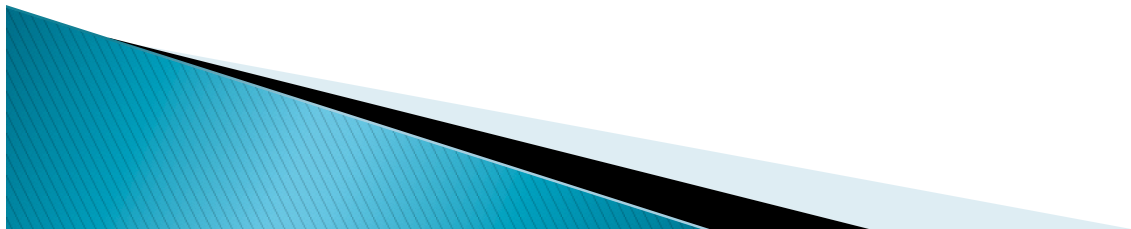


Tampering detection

- ▶ Manipulation **detection**



- ▶ Manipulation **localization**



Tampering detection



Forgery Detection

Extract features

Classify

Y

N

Training set:
Genuine+Forged

Forgery Localization

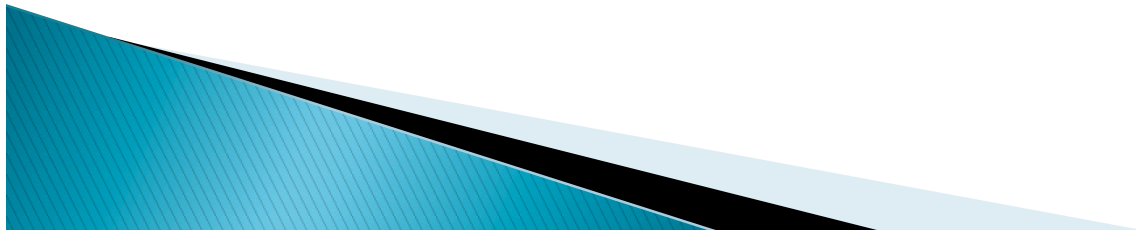
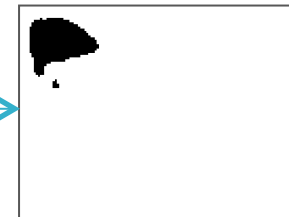


Sliding window

Extract features

Classify

Aggr. & Decide



Tampering detection blind



Forgery Detection



Forgery Localization



Tampering detection

Pixel level traces

- CFA
- PRNU
- Possible filtering, e.g., median

Physical and geometric traces

- Light direction
- Shadows
- Perspective geometry analysis

Format based traces

- Multiple JPEG compression
- Coding

Semantic level

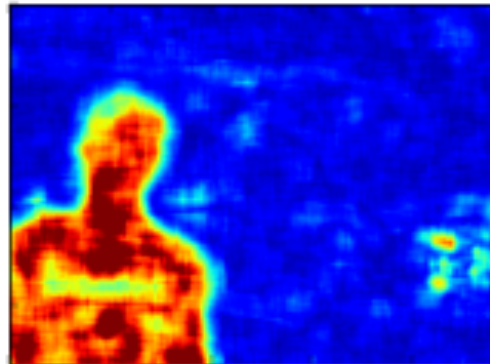
- Date and place are coherent?



Pixel level traces



Image



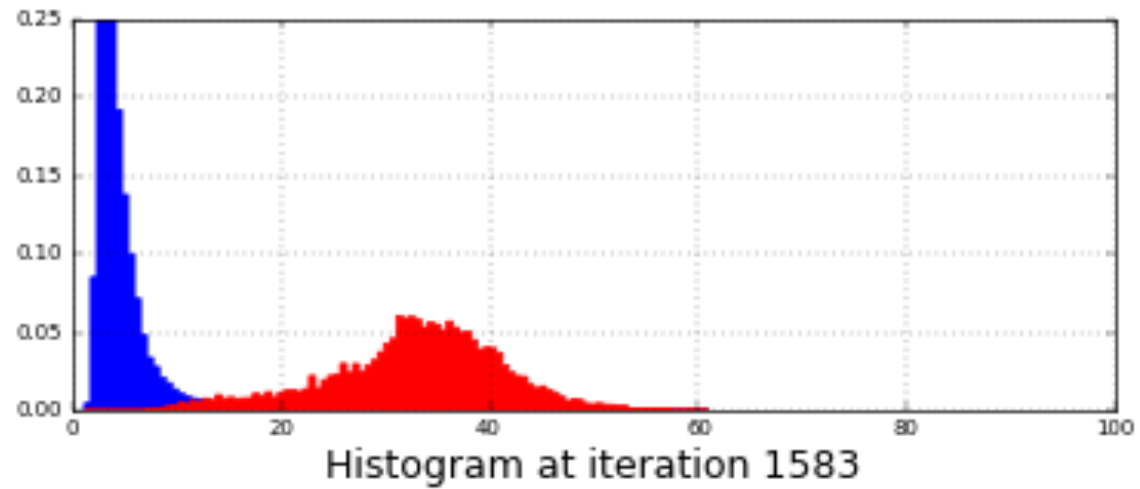
Heat map



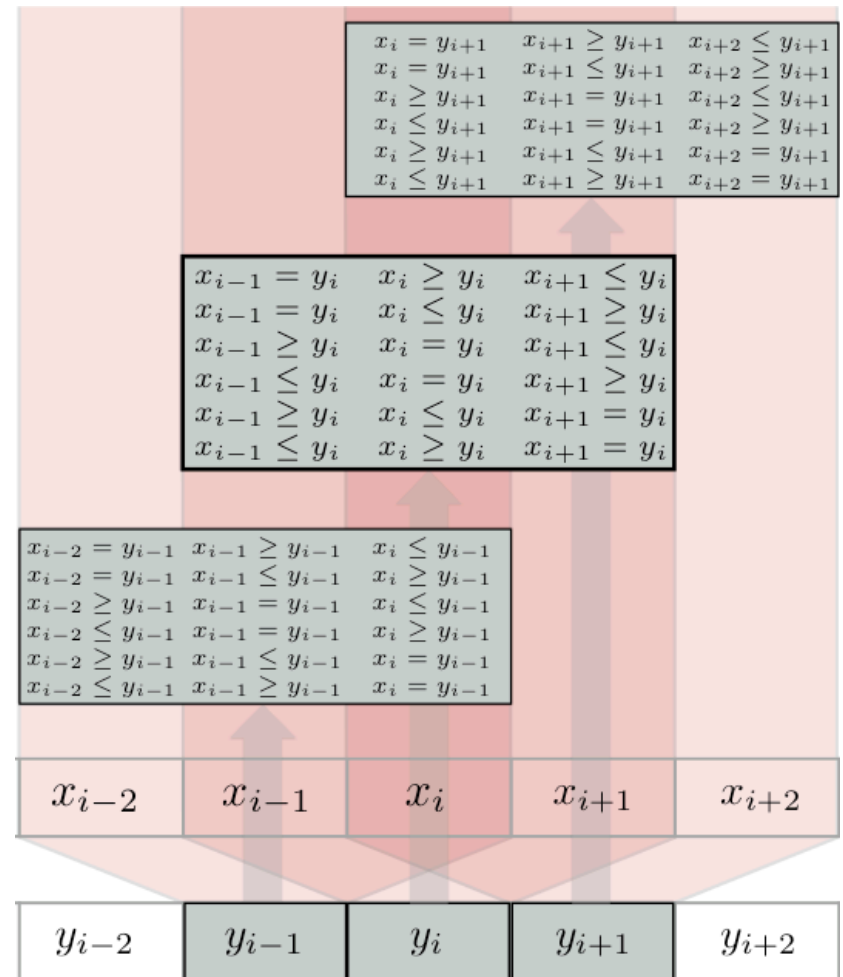
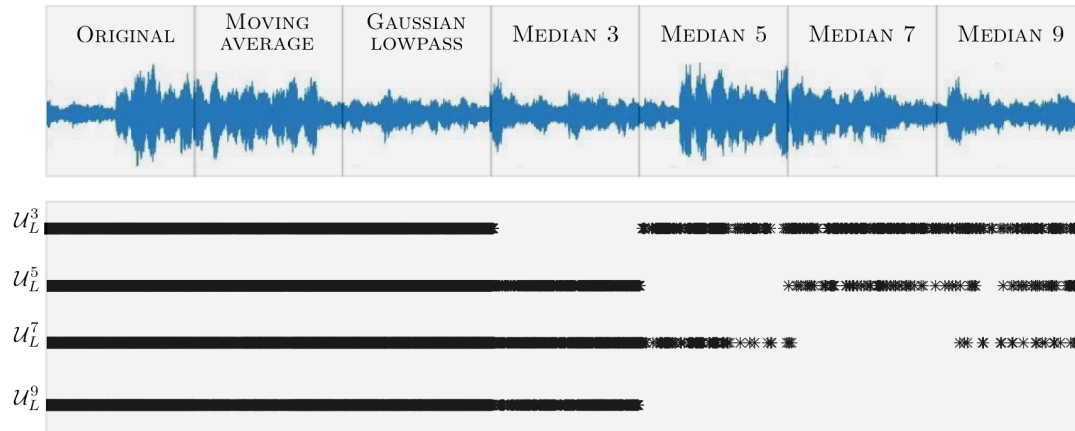
Binary map



Ground truth



Median filtering traces



Morphological filtering traces

« Respect, Monsieur Eddy ! »
N°2398 LE MITCHELL ACADEMIQUE

Dans votre numéro du 21 octobre, une critique de l'ultime opus d'Eddy Mitchell m'a fait ouvrir la médiation pour vérifier les écrits de Mme Sophie Delacour. Après une lecture attentive de l'album, je me suis précipité chez mon éditeur...

Notes bien que je n'échappe pas systématiquement Eddy Mitchell, mais j'ai trouvé dans ce disque un résumé habillé du parcours du chanteur : entre croquer et rocker, il nous balance son humour et en fait même sur des harmonies bien trouvées (Pierre Papsidiamants) et des rythmes d'enfer (En garde à vue). Je ne connais pas vos goûts médians S.D., mais je vous conseille de réécouter l'une ou l'autre un peu plus attentive ce « come-back » et comme le disaient Otis et Aretha : respect, Monsieur Eddy ! Serge Lelièvre (internet)

« Une façon claire, juste et objective »

N°2398 CONSERVER OU CONSTRUIRE
J'ai lu avec beaucoup d'attention l'article de M. Denis Olivennes et je trouve enfin quelqu'un qui ose traiter de façon claire, juste et objective le problème des retraités en France. Dans cet océan d'opinion généralisée qui nous comble, c'est vraiment rafraîchissant. Situations qu'enfin nos responsables politiques, syndicalistes et médiatiques renvoient le sens des responsabilités. Félicitations et sincères remerciements. Jacques Bailé, 33170 Gradignan

Inacceptable !

Quels directeurs des programmes (tous), proches de l'Église catholique, se sont permis du 1^{er} au 5 novembre sur le service public dans une France dite républicaine et laïque, d'imposer à des millions de téléspectateurs – dont un grand nombre ont pris leurs distances avec la religion – une semaine de « reportages » complaisants auprès d'une communauté de bonnes sœurs confites dans leur naïveté divine (...). Cette irresponsable propagande, cette catéchisation télévisuelle mériterait à tout le moins une réprimande du CSA. A quand une semaine consacrée au culte protestant, israélite, musulman, à celui des moines, des Téninois de Jéboval, des radicaux ou, tout bonnement et en toute logique, au quotidien des agnostiques et des athées ? Jean-Béard (internet)

LES PLUS DE L'ACTUALITÉ

Algérie : les blessures toujours ouvertes

Plus de quarante ans après l'indépendance, comme le dit Claudine Rimes, « la guerre d'Algérie et l'écrit restent un blessure intime pour tous ceux qui les ont vécus et aucune analyse rationnelle ne peut en rendre compte. Il y a, tout simplement, le souvenir de l'Algérie, quittée définitivement il y a quarante-huit ans (...). Je voudrais relire l'image partielle que vous donnez (solitairement ?) des "Français" d'Algérie. Il y avait aussi de pauvres gens, battus par l'histoire, qui n'ont pas connu grandes choses ou manipulations politiques des uns et des autres, mais qu'on a manipulés eux-mêmes puisque leur seule conviction résidait dans cet attachement à un pays où ils étaient nés. Leur tort ? N'avoir pas su porter un vrai regard sur la vraie réalité du pays, ne pas avoir fait l'effort de comprendre que leur mode de pensée et de vie n'était pas universels (...). D'avoir ainsi, par leur comportement irresponsable, participé à un des plus lamentables et effroyables maux de l'histoire contemporaine qui est retourné les a ébranlés. Est-ce cela être d'extrême droite ? Je ne pense pas que ma contribution soit publiée dans votre courrier de lecteurs, car elle ne correspond pas au bien-pensant attendu sur le sujet ». Et bien, si Madama, nous le publions car, quoi que vous pensiez, nous avons dans ce dossier laissé place à des témoignages similaires. Réviser la page 34.

Jean-Philippe Delcambre, lui, voudrait faire partager son sentiment de frustration en tant qu'ancien rapatrié pour qui aucun engagement n'a été tenu : « En 2007, nous avons eu droit à Nour à la suite de M. Estrosi, promet-tant formellement un désendettement via un moratoire. Le temps passe, nous sommes désormais à près d'un demi-siècle de l'indépendance de l'Algérie. La course contre la montre va s'imposer dans la double perspective de cette célébration et de la présidentielle de 2012. L'heure de régler les comptes aussi... »

Jean-Marcel Bouguereau
jmbouguereau@nouvelobs.com

Voir le blog de Jean-Marcel Bouguereau :
« Derrière les plus de l'actualité »

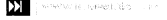


Denis Marchand précise que, se trouvant à Alger en 1982, il a pu constater l'absence totale d'assistance de la France lorsque il s'est agi d'aider un rapatriement des pieds-noirs. « C'est qui embourbaient à cette époque étaient les plus humbles. Ces gens devaient quitter leur logement sans quasiment rien emporter ». Ce lecteur nous reproche d'avoir parlé de « l'Etat tortionnaire ». Pour lui, « c'est honteux. Vous oubliez d'alléger les nombreuses exactions des "moussaires" ». « C'est encore plus qu'une guerre (...) ». Non, nous n'oublions rien mais, comme la rappelle Jean Daniel, « la torture est positive : le procédé le plus humiliant pour la victime, le plus dévalorisant pour le bourreau ».

Et dans cette guerre il y a une autre catégorie de victimes, celles des barbes, qui pour beaucoup payent encore même si, comme le souligne Jean Peyroudet, « c'est grâce à l'honneur de certains officiers qui ont desobéi aux ordres que quelques milliers de barbas (...) ont pu embarquer pour la métropole droite ». Et dans cette guerre il y a une autre catégorie de victimes, celles des barbis, qui pour beaucoup payent encore même si, comme le souligne Jean Peyroudet, « c'est grâce à l'honneur de certains officiers qui ont desobéi aux ordres que quelques milliers de barbis (...) ont pu embarquer pour la métropole et s'échapper aux massacres du FLN ». Cette guerre, au refus de l'appeler par son nom. On parlait des « événements », comme en atteste le témoignage longtemps refusé d'une élève infirmière, Lucie Marsol, témoin d'une crise entre Algériens et de la mort de l'un d'eux dans l'indifférence générale : « Ne vous occupez pas de ce barba, laissez ces gens-là se débrouiller ». D'autres phrases se précipitent, je suis sûr que je pourrais secourir un jeune homme, si c'était à refaire (...). Qu'est-il devenu ? Une mère, une sœur, une femme le pleure-t-elle ? »

Jean-Marcel Bouguereau
jmbouguereau@nouvelobs.com

Voir le blog de Jean-Marcel Bouguereau :
« Derrière les plus de l'actualité »



« Respect, Monsieur Eddy ! »
N°2398 LE MITCHELL ACADEMIQUE

Dans votre numéro du 21 octobre, une critique de l'ultime opus d'Eddy Mitchell m'a fait ouvrir la médiation pour vérifier les écrits de Mme Sophie Delacour. Après une lecture attentive de l'album, je me suis précipité chez mon éditeur...

Notes bien que je n'échappe pas systématiquement Eddy Mitchell, mais j'ai trouvé dans ce disque un résumé habillé du parcours du chanteur : entre croquer et rocker, il nous balance son humour et en fait même sur des harmonies bien trouvées (Pierre Papsidiamants) et des rythmes d'enfer (En garde à vue). Je ne connais pas vos goûts médians S.D., mais je vous conseille de réécouter l'une ou l'autre un peu plus attentive ce « come-back » et comme le disaient Otis et Aretha : respect, Monsieur Eddy ! Serge Lelièvre (internet)

« Une façon claire, juste et objective »

N°2398 CONSERVER OU CONSTRUIRE
C'est dans les prochains jours que l'ex ministre de Lionel Jospin devrait annoncer le lancement de sa propre fondation, l'École d'avenir libérée comme beaucoup d'autres par l'Institut de France, qui réunira quel Con les cinq grandes Académies, dont les immat. tels en habit... vert. Une « paraver fondationn. tiques, syndicalistes et médiatiques renouvant le sens des responsabilités. Félicitations et sincères remerciements. Jacques Bailé, 33170 Gradignan

Inacceptable !

Quels directeurs des programmes (tous), proches de l'Église catholique, se sont permis du 1^{er} au 5 novembre sur le service public, dans une France dite « républicaine et laïque », d'imposer à des millions de téléspectateurs – dont un grand nombre ont pris leurs distances avec la religion – une semaine de « reportages » complaisants auprès d'une communauté de bonnes sœurs confites dans leur naïveté divine (...). Cette irresponsable propagande, cette catéchisation télévisuelle mériterait à tout le moins une réprimande du CSA. A quand une semaine consacrée au culte protestant, israélite, musulman, à celui des moines, des Téninois de Jéboval, des radicaux ou, tout bonnement et en toute logique, au quotidien des agnostiques et des athées ? Jean-Béard (internet)

LES PLUS DE L'ACTUALITÉ

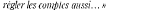
Algérie : les blessures toujours ouvertes

Plus de quarante ans après l'indépendance, comme le dit Claudine Rimes, « la guerre d'Algérie et l'écrit restent une blessure intime pour tous ceux qui les ont vécus et aucune analyse rationnelle ne peut en rendre compte. Il y a, tout simplement, le souvenir de l'Algérie, quittée définitivement il y a quarante-huit ans (...). Je voudrais relire l'image partielle que vous donnez (solitairement ?) des "Français" d'Algérie. Il y avait aussi de pauvres gens, battus par l'histoire, qui n'ont pas connu grandes choses ou manipulations politiques des uns et des autres, mais qu'on a manipulés eux-mêmes puisque leur seule conviction résidait dans cet attachement à un pays où ils étaient nés. Leur tort ? N'avoir pas su porter un vrai regard sur la vraie réalité du pays, ne pas avoir fait l'effort de comprendre que leur mode de pensée et de vie n'était pas universels (...). D'avoir ainsi, par leur comportement irresponsable, participé à un des plus lamentables et effroyables maux de l'histoire contemporaine qui est retourné les a ébranlés. Est-ce cela être d'extrême droite ? Je ne pense pas que ma contribution soit publiée dans votre courrier de lecteurs, car elle ne correspond pas au bien-pensant attendu sur le sujet ». Et bien, si Madama, nous le publions car, quoi que vous pensiez, nous avons dans ce dossier laissé place à des témoignages similaires. Réviser la page 34.

Jean-Philippe Delcambre, lui, voudrait faire partager son sentiment de frustration en tant qu'ancien rapatrié pour qui aucun engagement n'a été tenu : « En 2007, nous avons eu droit à Nour à la suite de M. Estrosi, promettant formellement un désendettement via un moratoire. Le temps passe, nous sommes désormais à près d'un demi-siècle de l'indépendance de l'Algérie. La course contre la montre va s'imposer dans la double perspective de cette célébration et de la présidentielle de 2012. L'heure de régler les comptes aussi... »

Jean-Marcel Bouguereau
jmbouguereau@nouvelobs.com

Voir le blog de Jean-Marcel Bouguereau :
« Derrière les plus de l'actualité »

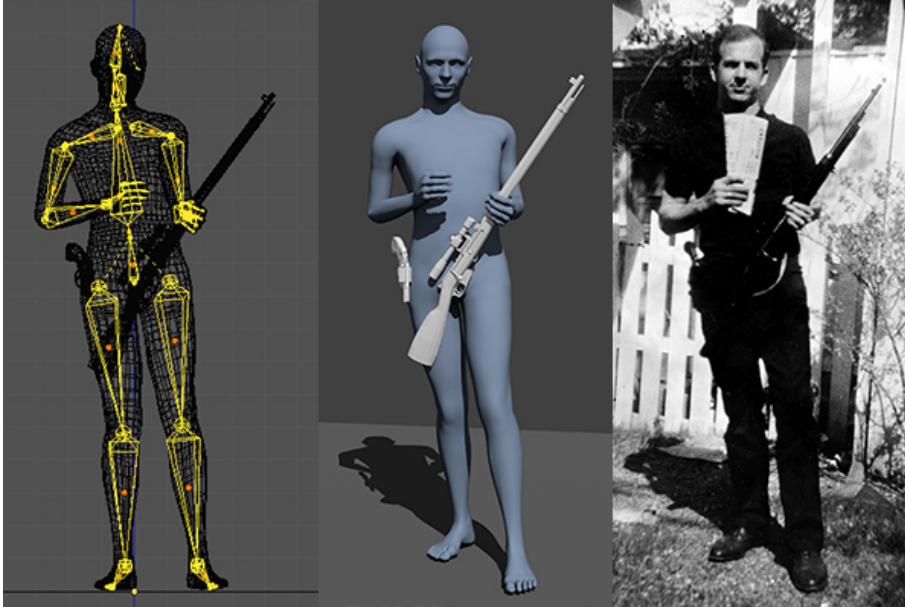


Light direction traces

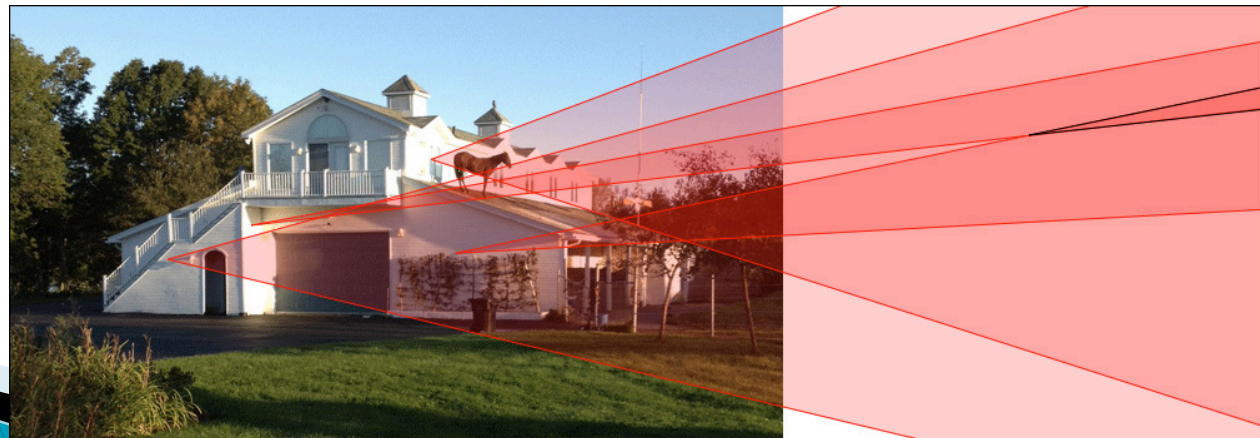
- ▶ In compositing of images, it is difficult to match lighting effects under which each image was taken
- ▶ Inconsistencies in lighting are used to detect splicing



Reflections, shadows traces



Farid et al. JDFSL 2015



Geometric traces



Geometric traces

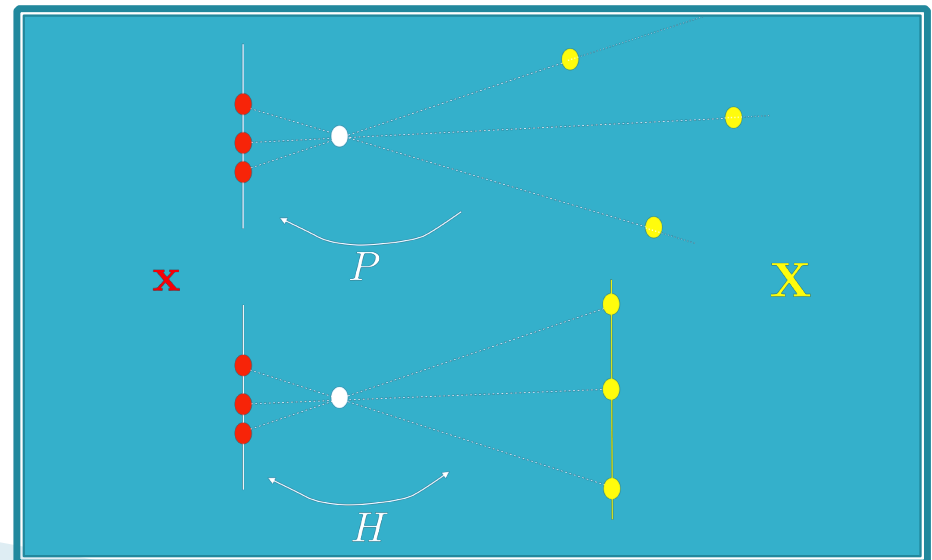


with Farid and Conotter ICIP 2010

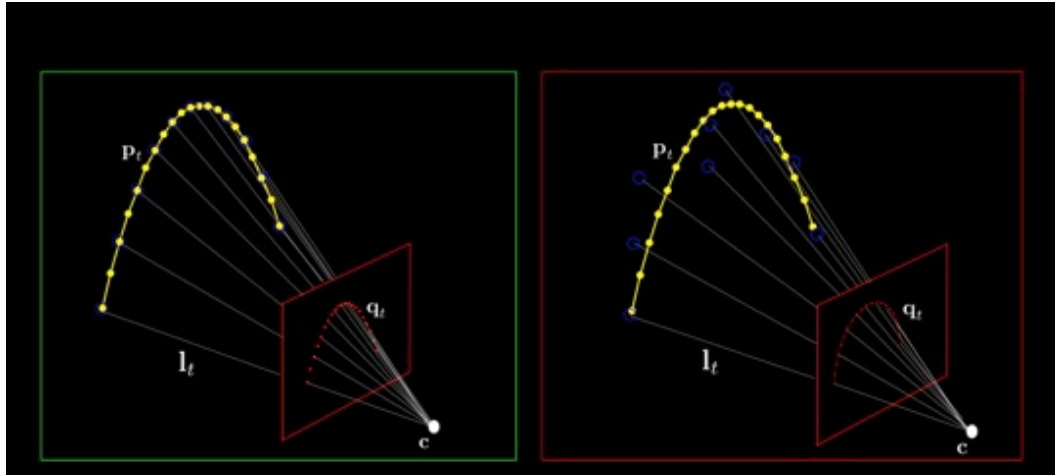
Geometric traces

LACASA = H LACASA

ALICE \neq H ALICE



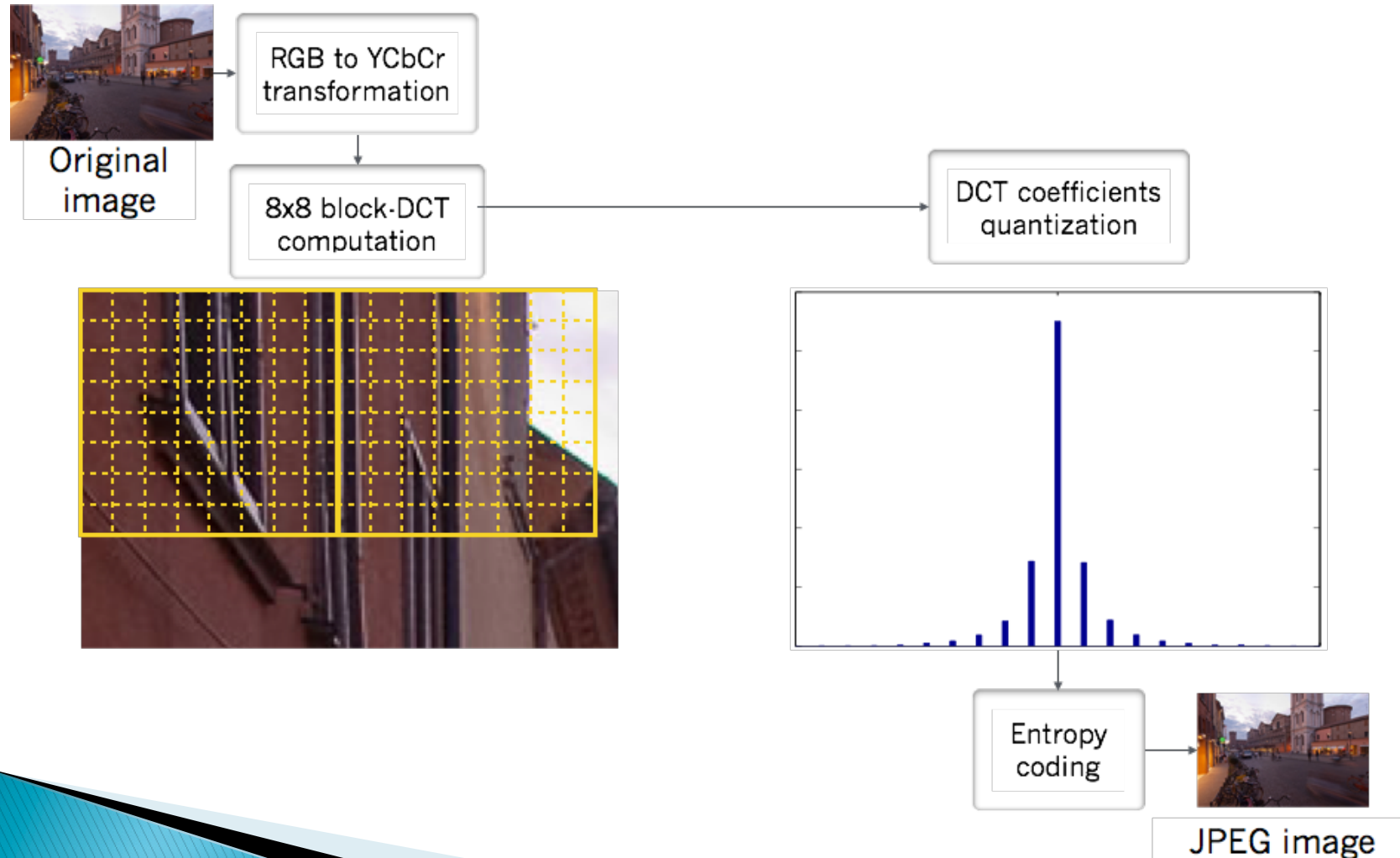
Geometric traces



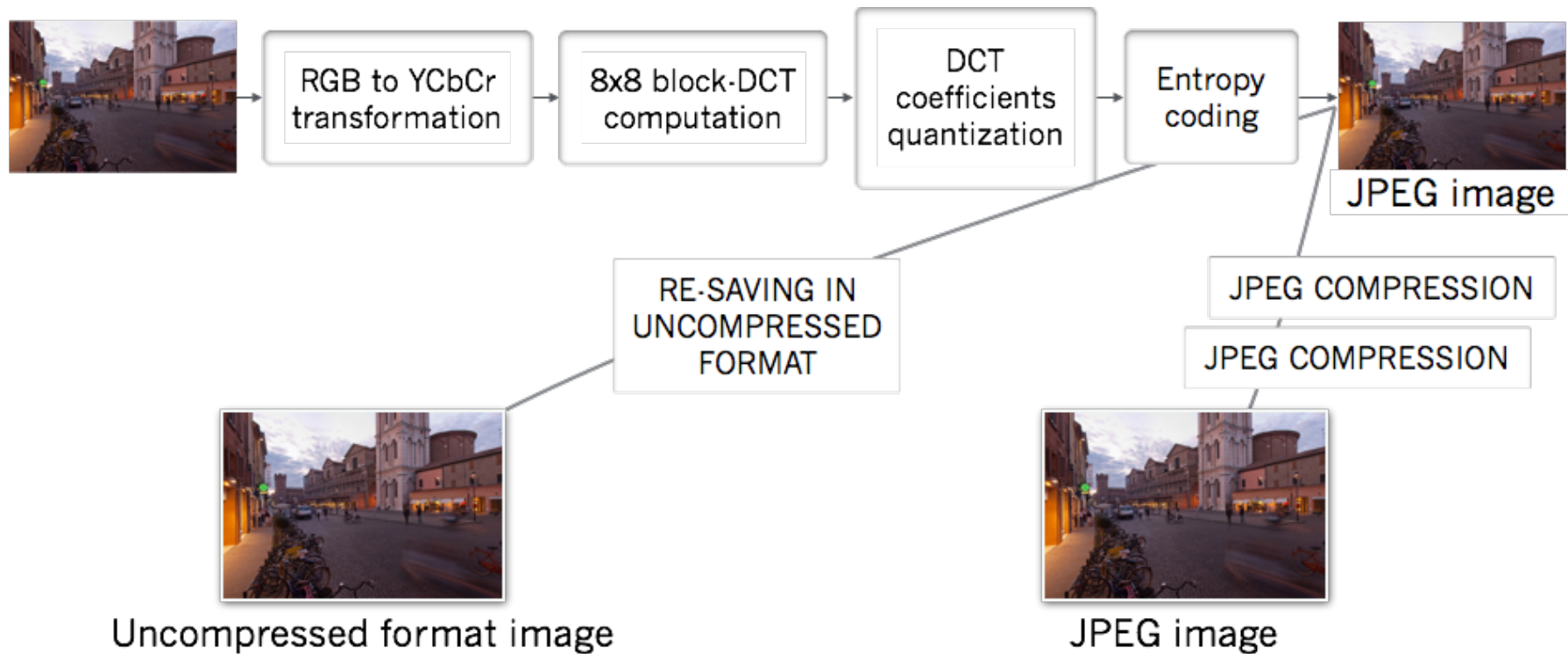
- ▶ A geometric technique to detect physically **implausible trajectories** of objects in video sequences



JPEG compression traces



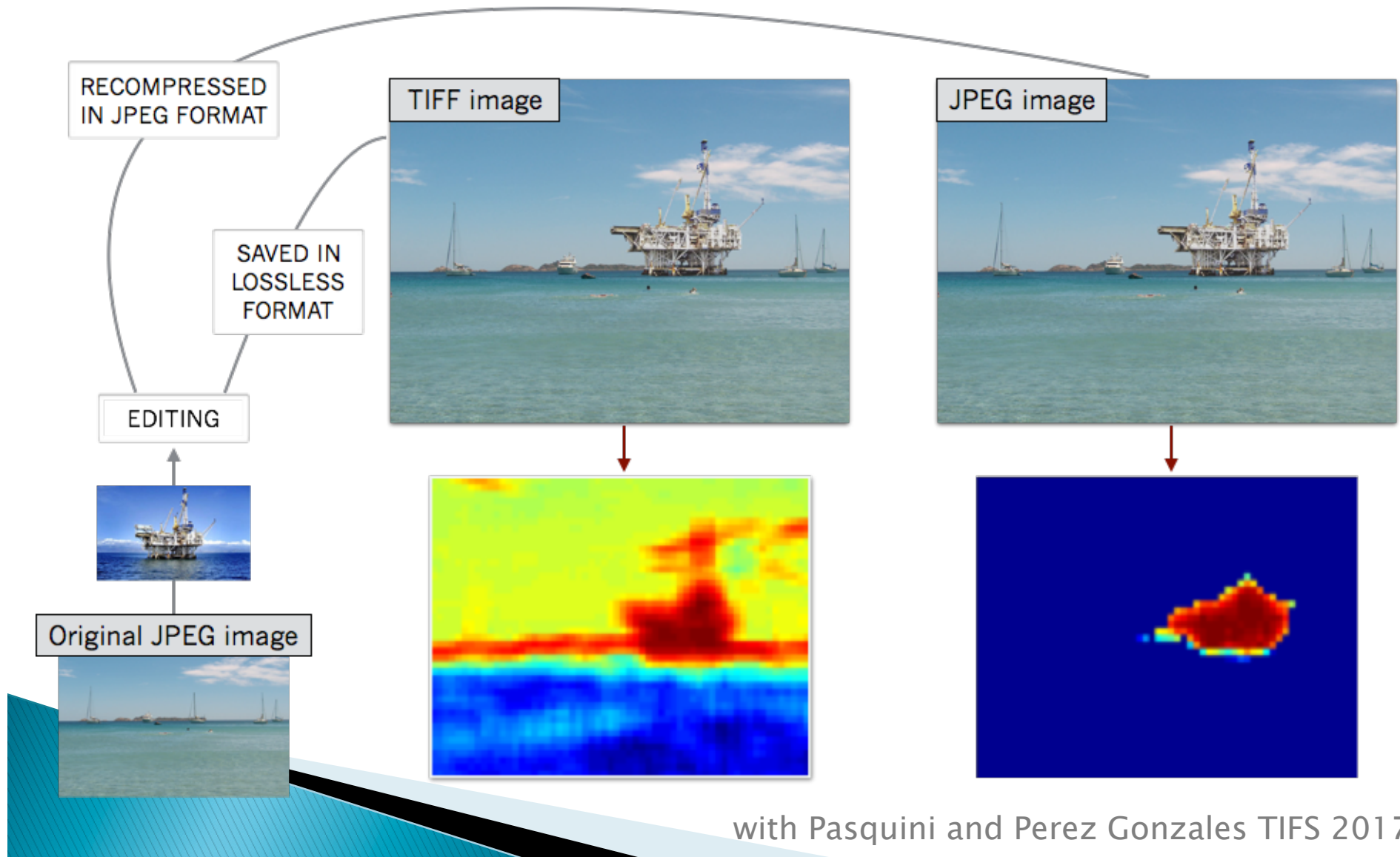
JPEG compression traces



SINGLE COMPRESSION DETECTION
detection of single JPEG compression traces in uncompressed format images

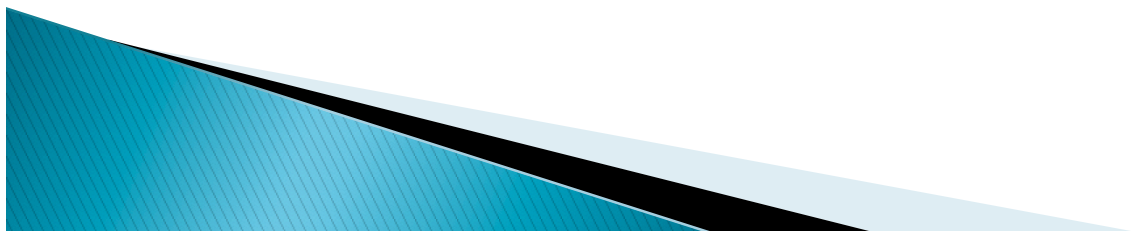
MULTIPLE COMPRESSION DETECTION
detection of multiple JPEG compression traces in JPEG images

JPEG compression traces



Open issues

- ▶ Lots of specific tools but no **universal** solution
- ▶ How to **combine** results from different tools?
- ▶ Missing large-scale **dataset** with different types of manipulation
- ▶ What about **robustness** to current information sharing?



New trends

- ▶ **In the wild** multimedia forensics
 - Real-world scenarios
 - Not only lab conditions
- ▶ Verifying Multimedia Use at MediaEval 2015 and 2016
 - Given a post (image+metadata), return a decision (fake, real, unknown) on whether the information presented by the post reflects the reality



Real photo
captured April 2011 by WSJ
but
heavily tweeted during
Hurricane Sandy
(29 Oct 2012)

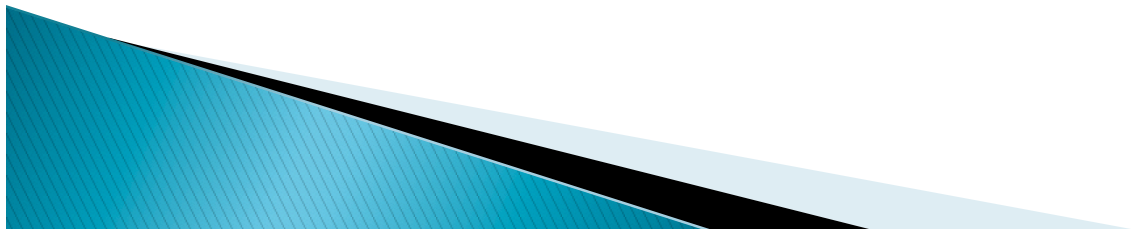
New trends

- ▶ Multimedia data **phylogeny**
 - History of the data
 - Social network and sharing platforms profiling



New trends

- ▶ **Adversary aware** multimedia forensics
 - Adversarial machine learning
 - Game theory
 - Current analysis only in very limited scenarios



New trends and open issues

In the wild forensics

Social network forensics

Adversary aware forensics

Deep learning for forensics

Video forensics

Cross-media forensics

Thanks to D.T. Dang Nguyen, C. Pasquini, V. Conotter,
F. De Natale, A. Piva, R. Caldelli, L. Verdoliva