PRIVACY PROTECTION IN VISUAL DATA: OVERVIEW AND CHALLENGES

Frédéric Dufaux

Laboratoire des Signaux et Systèmes (L2S)
CNRS - CentraleSupelec - Université Paris-Sud

frederic.dufaux@l2s.centralesupelec.fr

Outline

- **Context**
- **Privacy Protection in Visual Data**
  - Pixelization, Blur, Masking
  - Scrambling/encryption
  - Smart camera
- **Benchmarking of Privacy Protection Solutions**
- **New Trends in Imaging Technology**
- **Conclusions**

Context

Video Surveillance

- **4-6M cameras in U.K.**
- **500’000 cameras in Greater London**
  - Londoner recorded more than 300 times a day
- **History of abuse**

Video Surveillance

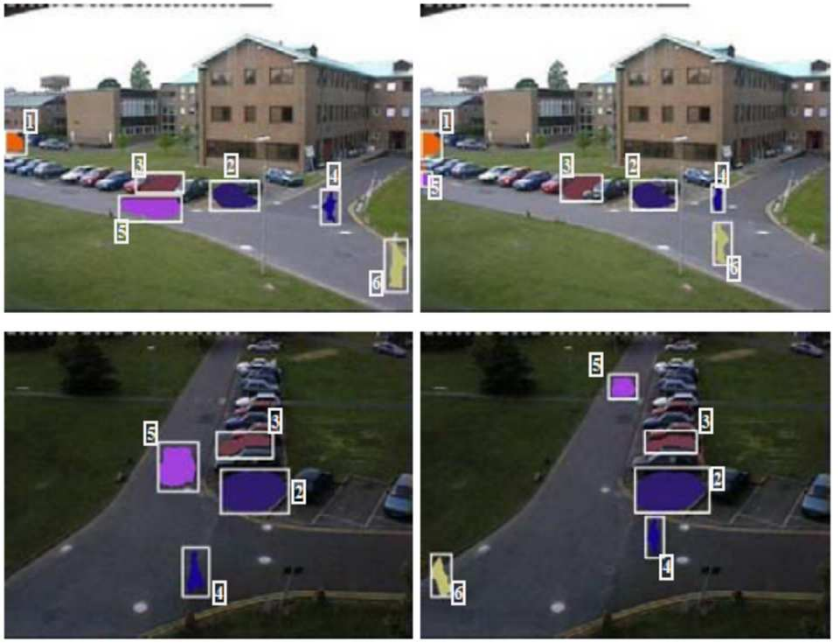
- **Intrusion detection**
  - Residential surveillance, retail surveillance, …
- **Traffic control**
  - Speed control
- **Access to places**
  - Car license plate recognition
- **Event detection**
  - Child/Elderly care
- **Marketing/statistics**
  - Customers habits
  - Number of visitors
- **…**

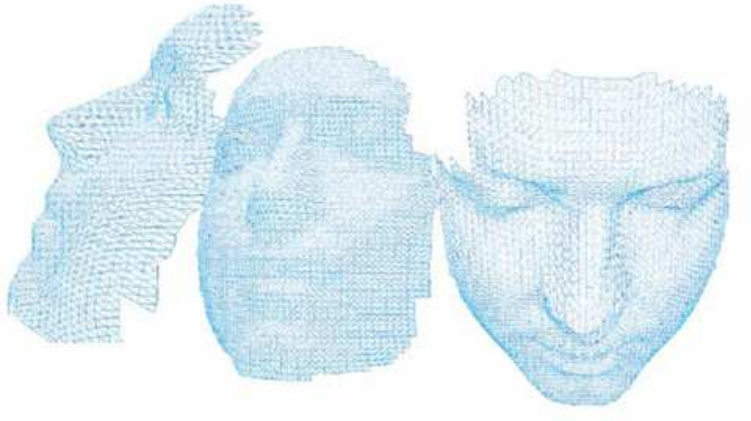
History of Abuse

- **Criminal abuse**
  - Criminal misuse by law enforcement officers
- **Institutional abuse**
  - Spy upon political demonstrations and political activists
- **Discrimination**
  - Racial discrimination
- **Voyeurism**
  - Bored male operators spying on women
  - Footage of public cameras made publicly available

Advanced Video Analytics

- **Advances video analytics**
  - Object detection and tracking
  - Face detection and recognition
  - People in the scene
  - Cars license plates





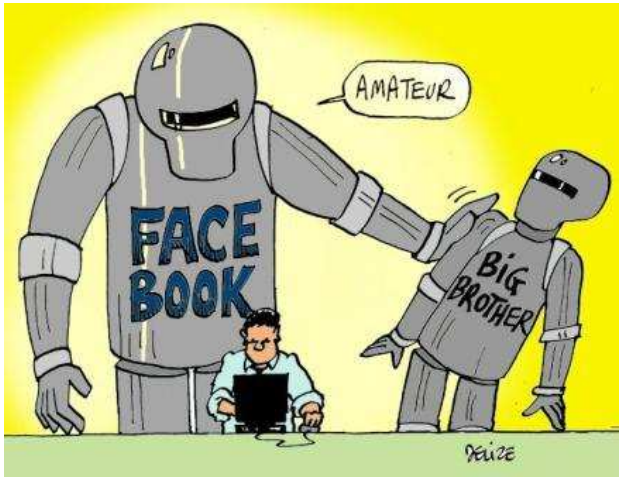
- **Big media data analysis**
- **Deep learning**

Context

Social Media

- **2-3 Terabytes of photos uploaded every day**

- **300 hours of video uploaded every minute**



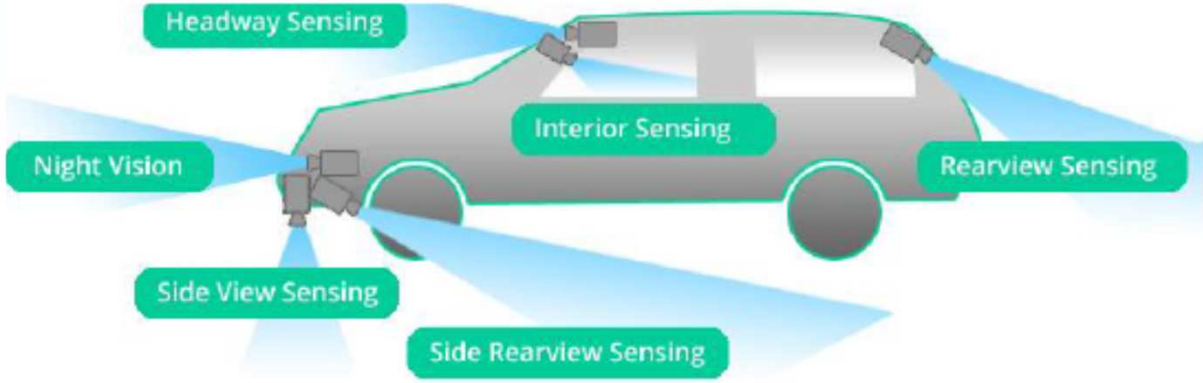
- **A lot of personal information!**
- **Allows to make link between different sources of information**

Context – Market Sizes

- **Video surveillance**
  - 150M cameras/year

- **Mobile phones**
  - 1B cameras/year

- **Automotive industry**
  - 100M cameras/year in 2020



Privacy Protection in Visual Data

Definition

- **Privacy is linked to personal information**
  - Identifiable individuals
  - Gender, race, age, color of clothes, facial features, etc.

- **Privacy protection**
  - Limit access to personal information in recorded or streamed video
  - Require to identify regions with privacy-sensitive information
    - Predefined static zones
    - Automatic and dynamic using video analytics
    - Active with RFID tags
  - May depend on the context, external knowledge, and other linkable sources of information

Categories of Privacy Protection Solutions

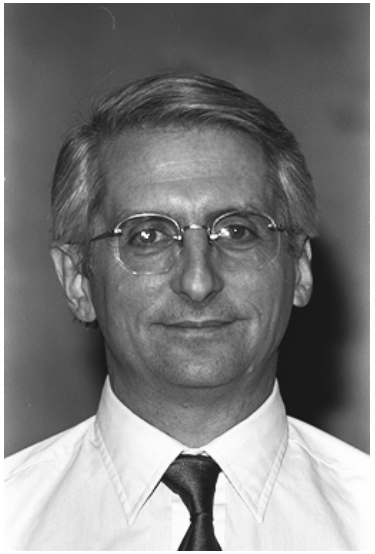
- **Visual privacy filters**
  - Distort, remove or hide visual information in regions containing privacy-sensitive information
- **Smart cameras**
  - Cameras which embedded video analytics tools and only output alerts or metadata descriptors
- **System-level security**
  - Access management and policies at the system-level
  - Private information is only accessible by those users granted such rights

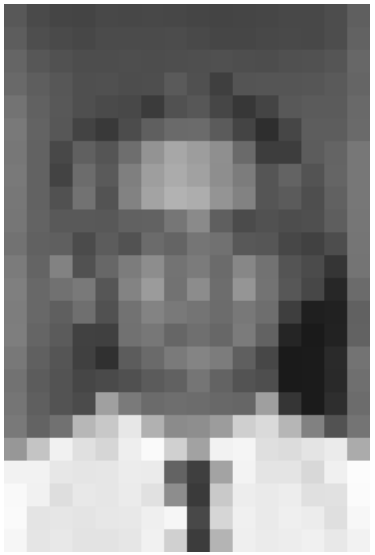
Pixelization

- **Naïve approach for privacy protection**
  - Sometimes used in TV, Internet, social networks, etc. in order to obscure faces for anonymity
- **Notable reduction of resolution in ROI**
  - Substitute a square block of pixels with its average
  - Very easy to implement!!

$$I_{\text{pixelization}}(x, y) = \frac{1}{b^2} \sum_{i=0,\ldots b-1} \sum_{j=0,\ldots b-1} I\left(\left\lfloor \frac{x}{b} \right\rfloor + i, \left\lfloor \frac{y}{b} \right\rfloor + j\right)$$

- **Drawback**
  - Irreversible
  - Not efficient at concealing information!





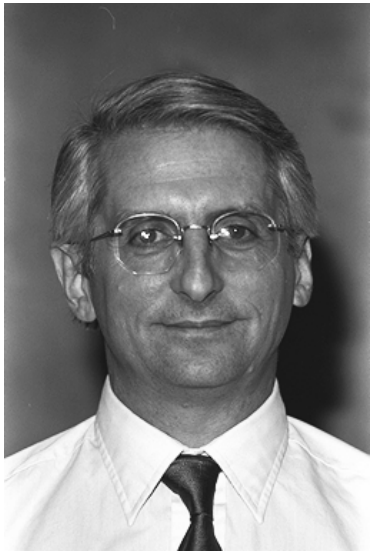
Blur

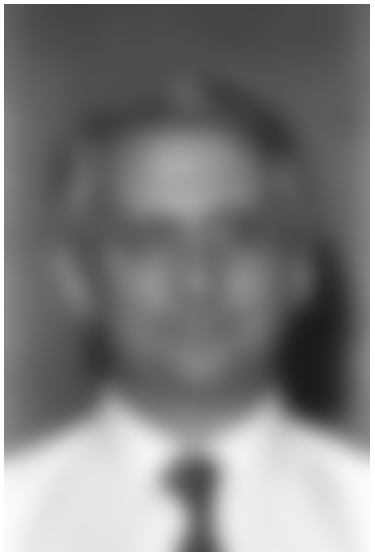
- **Naïve approach for privacy protection**
  - Sometimes used in TV, Internet, social networks, etc. in order to obscure faces for anonymity
- **Removes details in ROI by applying a Gaussian low pass filter**
  - Image is convolved with a 2D Gaussian function
  - Very easy to implement!!

$$I_{\text{Gaussian blur}}(x, y) = I(x, y) * G(x, y)$$

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

- **Drawback**
  - Irreversible
  - Not efficient at concealing information!





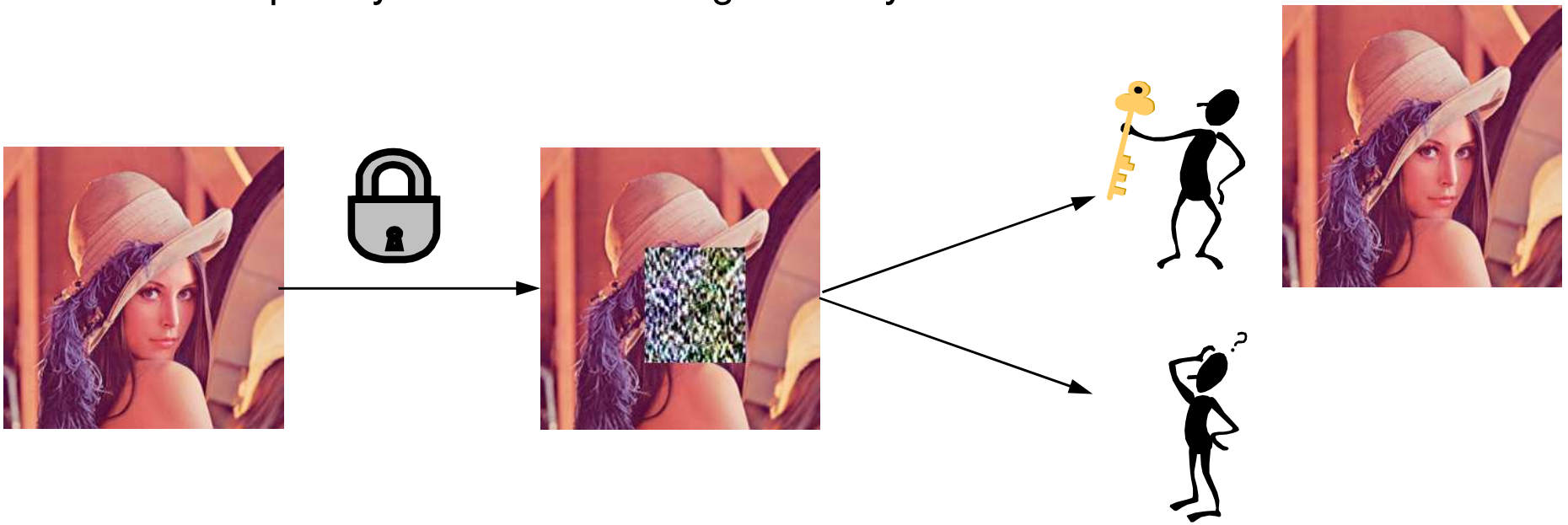
Masking

- **Naïve approach for privacy protection**
  - Sometimes used in TV, Internet, social networks, etc. in order to obscure faces for anonymity
- **Replace ROI by a plain form**
  - Silhouette of privacy-sensitive regions
  - Very easy to implement!!
- **Drawback**
  - Irreversible

Scrambling / Encryption

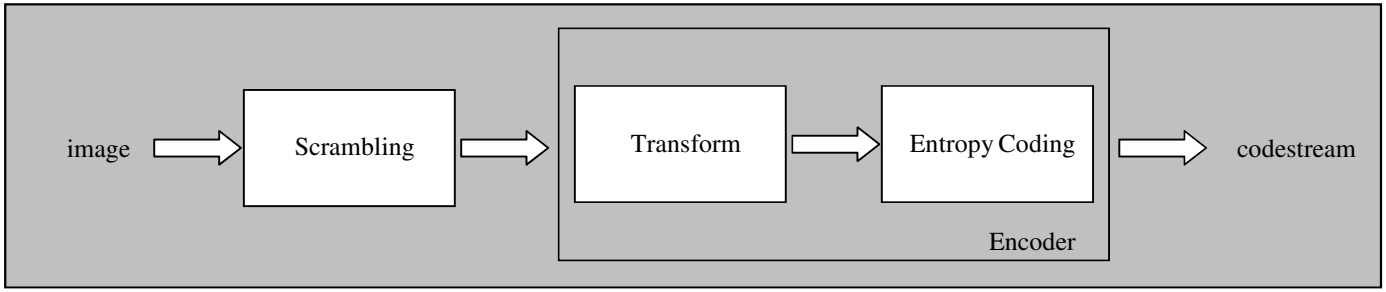
- **Scrambling is applied to the ROIs**
  - Sensitive information is concealed, e.g. people, license plate, …
  - Process is reversible with secret encryption key, kept by trusted third party
  - Cryptographically secure

  - Standard-compliant codestream / standard decoder
  - Coding performance should not be adversely affected
  - Complexity should not be significantly increased



Scrambling / Encryption

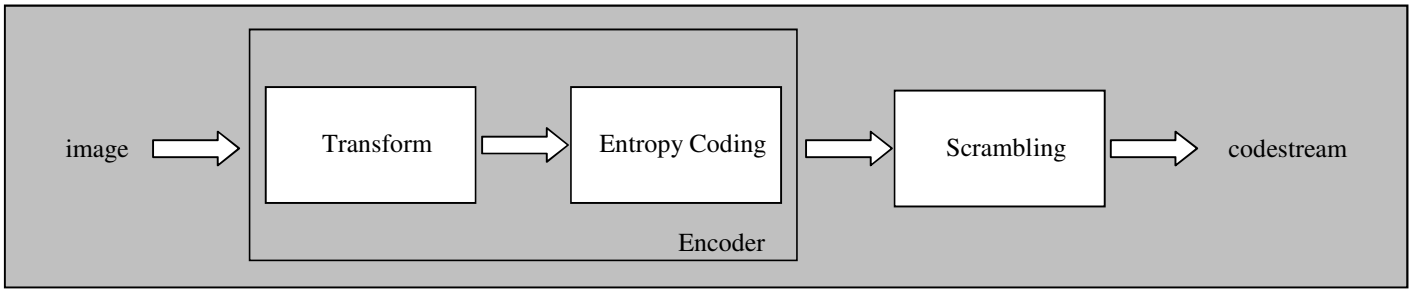
- **Image-domain**
  - Randomly flip bits in one or more bit planes



- **Pros**
  - Very simple
  - Independent from the subsequent encoding scheme
  - Does not affect the codestream syntax → standard compliance
- **Cons**
  - Significantly alter statistics of video signal
  - Ensuing compression less efficient

Scrambling / Encryption

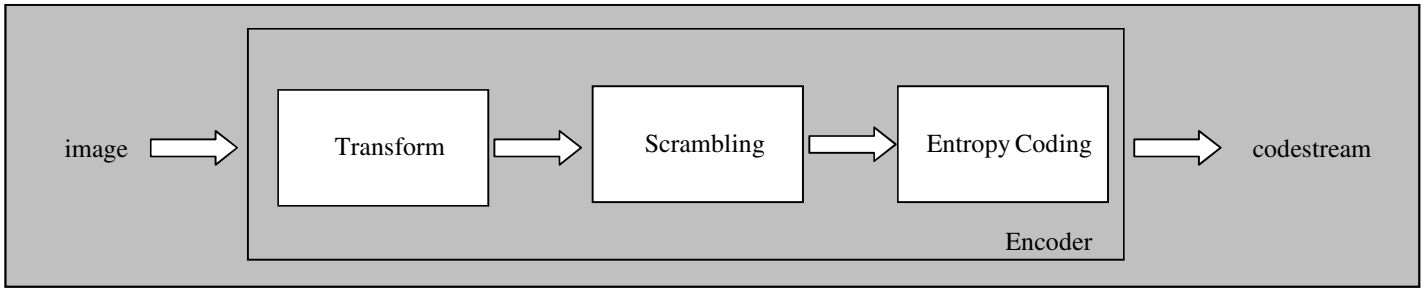
- **Codestream-domain**
  - Randomly flip bits in codestream



- **Pros**
  - Applied on codestream after encoding
- **Cons**
  - Require parsing of codestream
  - Difficult to guarantee syntax remains standard compliant and will not crash a decoder

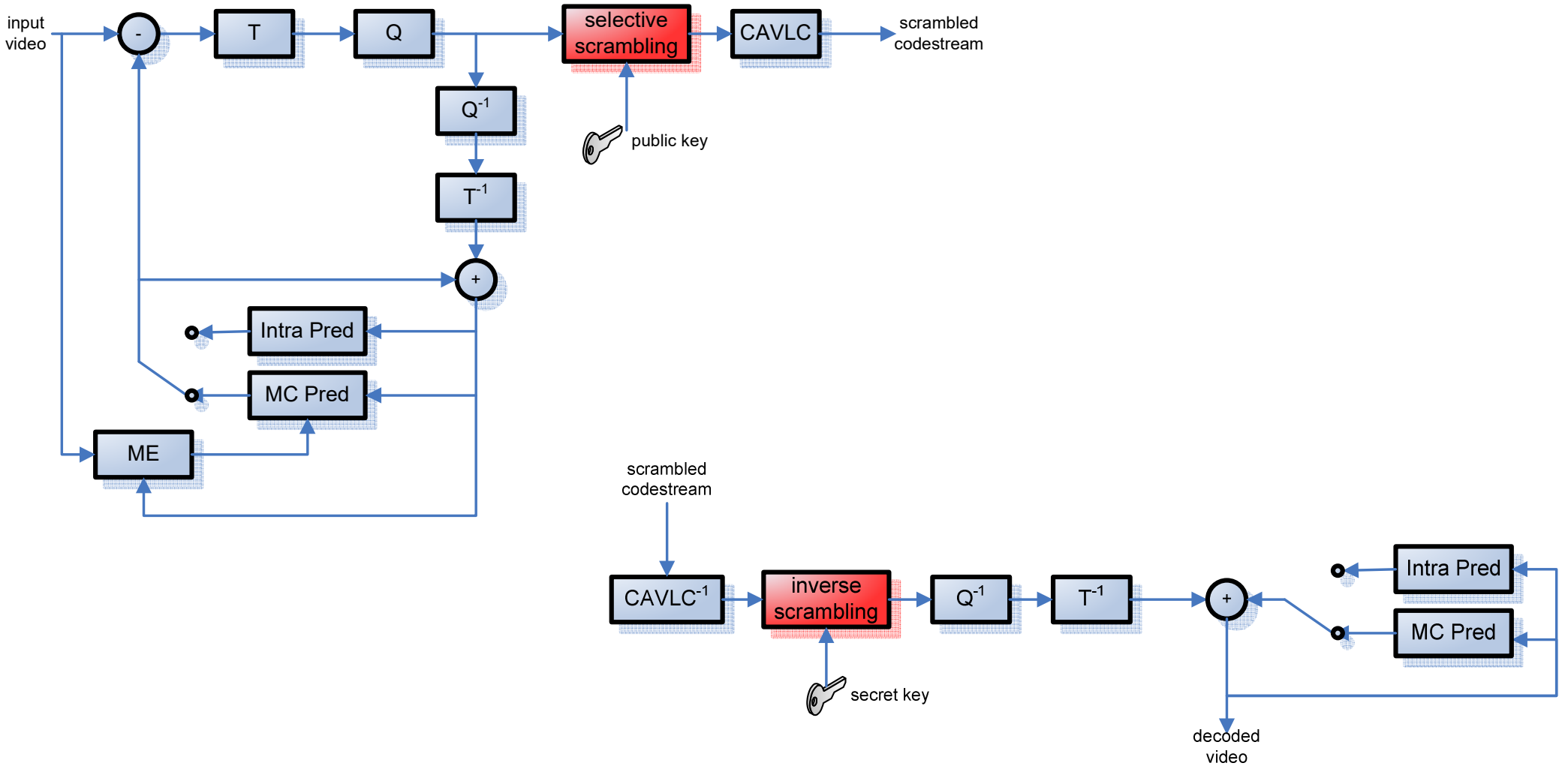
Scrambling / Encryption

- **Transform-domain**
  - Randomly flip transform coefficients



- **Pros**
  - Does not adversely affect subsequent entropy coding
  - Strength of scrambling can be controlled
  - Does not affect the codestream syntax → standard compliance
- **Cons**
  - Must be integrated inside the encoder

Scrambling in H.264/AVC



Scrambling in H.264/AVC

- **Flexible Macroblock Ordering (FMO)**
  - Two slice groups for foreground and background MBs
  - Slice Group Map Type 6
  - Explicit assignment of each MB to one of the slice groups
  - ROI segmentation mask is restricted to 16x16 MB boundaries
  - Background MB will not use scrambled foreground MB for spatial Intra prediction

- **Temporal Inter prediction**
  - Modify the mode selection to force some MB to be coded in Intra mode to prevent use of scrambled data for Inter prediction

Scrambling in H.264/AVC

- **Each 4x4 block in foreground MB**

- **Random sign inversion**
  - Pseudo-randomly flip the sign of quantized coefficients (weakly correlated)

$$qcoeff[i] = \begin{cases} -qcoeff[i] & \text{if } random\_bit = 1 \\ +qcoeff[i] & \text{otherwise} \end{cases}$$

- **Random permutation**
  - Rearrange the order of coefficients
  - Knuth shuffle to generate a permutation of *n* items with uniform random distribution

$$\begin{pmatrix} 0 & 1 & \dots & 14 & 15 \\ x_0 & x_1 & \dots & x_{14} & x_{15} \end{pmatrix}$$

Scrambling in H.264/AVC

- **Advantages**
  - Fully reversible
  - Same scrambled stream is transmitted to all users
  - Small impact in terms of coding efficiency
  - Requires a low computational complexity

Pseudo-random sign inversion

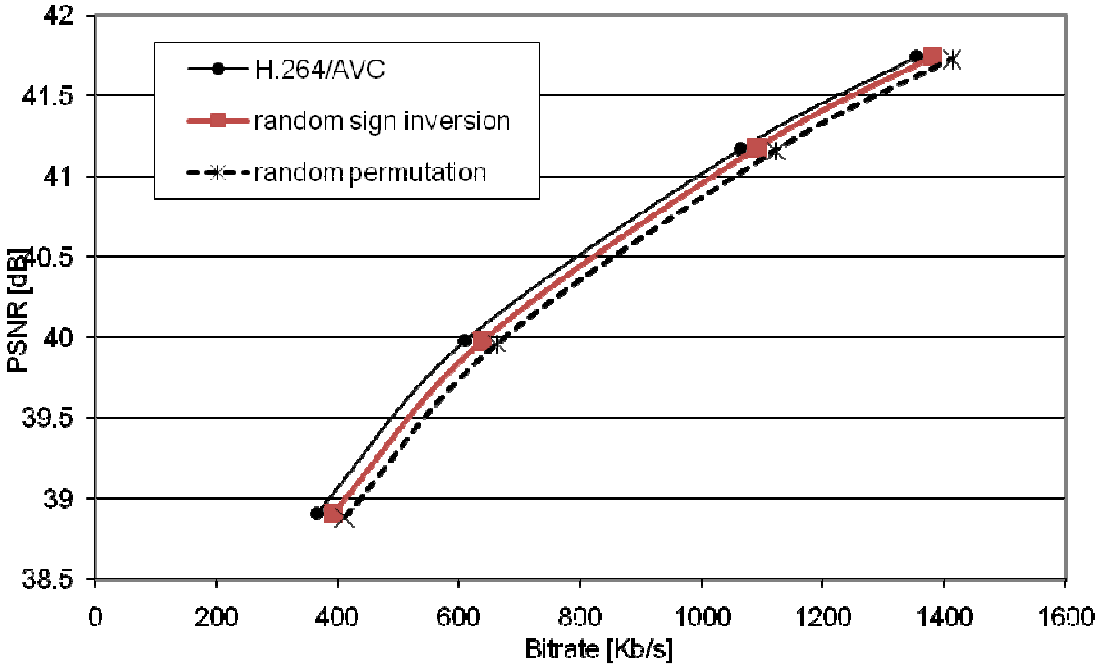
Pseudo-random permutation

Scrambling in H.264/AVC

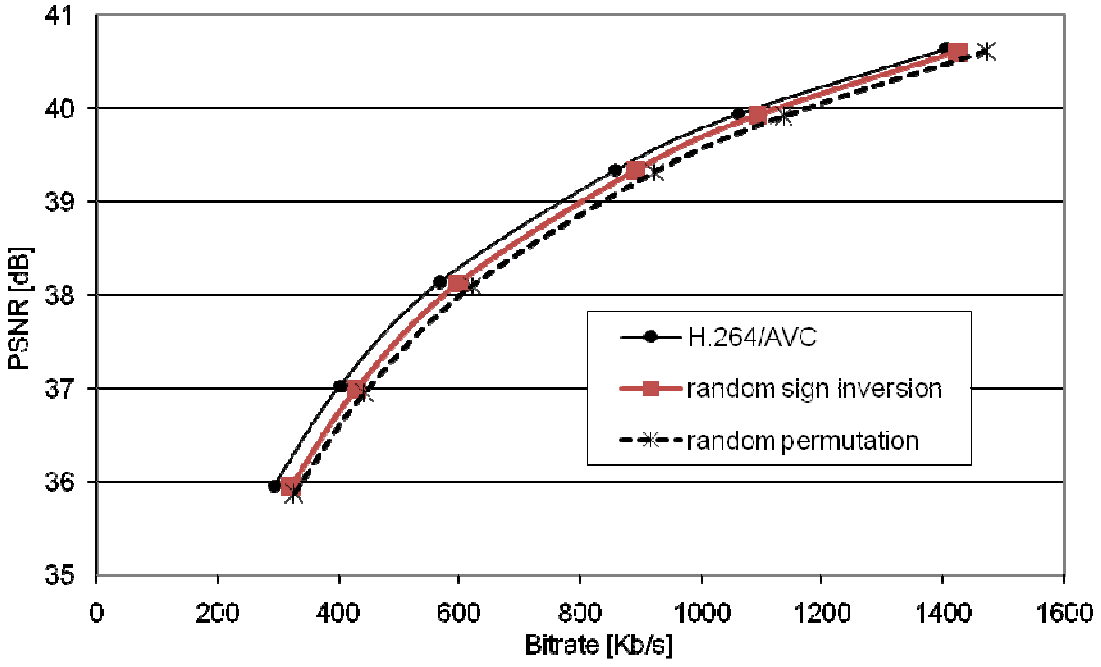
Pseudo-random sign inversion

Pseudo-random permutation

Results: Coding Efficiency



(a) Hall Monitor



(b) Road

- Overhead for FMO Slice Group Map Type 6
  - 1 bit per MB
  - CIF luminance frame: 12 Kb/s

Security: Brute-Force Attack

- **Assumptions**
  - CIF luminance frame
    - *101376 coefficients*
  - ROIs is known and cover 5% of image
    - *5068 coefficients*
    - *316 blocks*

- **Random sign inversion**
  - 5% of coefficients are non-zero → 253 coefficients
  - 2^{253} combinations for each frame
- **Random permutation**
  - $(16!)^{316}$ combinations for each frame

Smart MPEG-7 camera

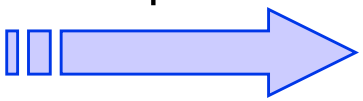
- **The MPEG-7 camera describes a scene in terms of semantic objects and of their properties**

XML scene description

- Image analysis: segmentation, change detection, and tracking implemented on the camera DSP
- Scene description represented using MPEG-7 (XML)

Smart MPEG-7 camera

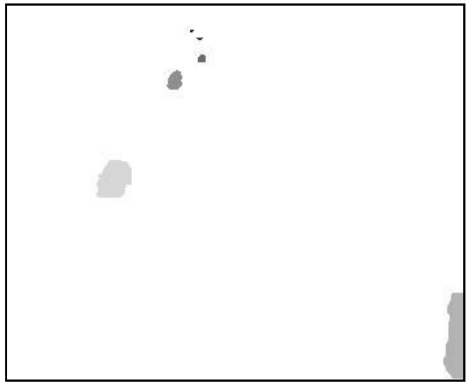
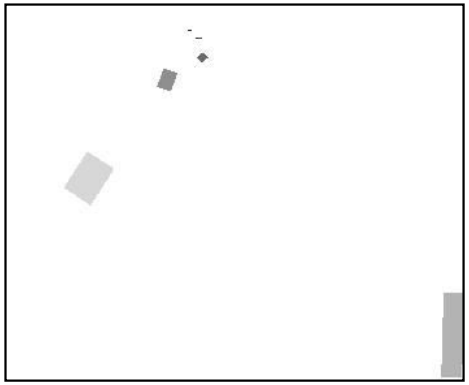
XML scene description

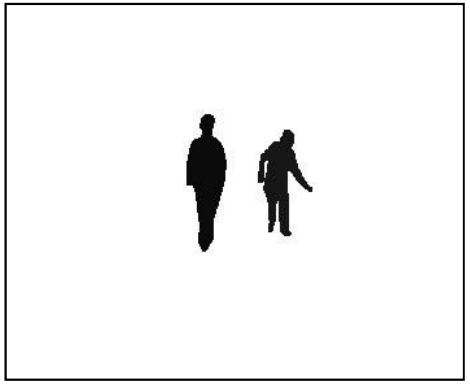
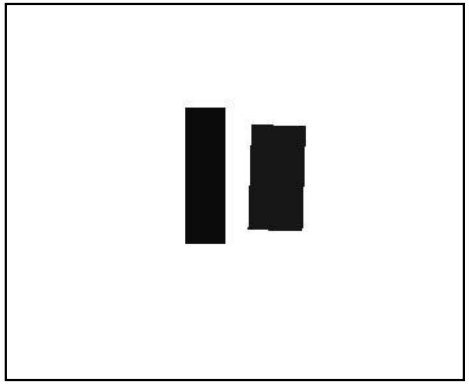


```
<!-- ################################################ --!>
<!-- DDL output for object 4                          --!>
<!-- ################################################ --!>

<Object id="4">

  <RegionLocator>
    <BoxPoly> Poly </BoxPoly>
    <Coords1> 237 222 </Coords1>
    <Coords2> 230 252 </Coords2>
    <Coords3> 240 286 </Coords3>
    <Coords4> 308 287 </Coords4>
    <Coords5> 312 284 </Coords5>
  </RegionLocator>

  <DominantColor>
    <ColorSpace>  YUV  </ColorSpace>
    <ColorValue1> 143.4 </ColorValue1>
    <ColorValue2> 123.3 </ColorValue2>
    <ColorValue3> 128.2 </ColorValue3>
  </DominantColor>

  <HomogeneousTexture>
    <TextureValue> 9.02 </TextureValue>
  </HomogeneousTexture>

  <MotionTrajectory>
    <TemporalInterpolation>
      <KeyFrame>  100  </KeyFrame>
      <KeyPos> 268.6 251.7 </KeyPos>
      <KeyFrame>  101  </KeyFrame>
      <KeyPos> 262.8 241.0 </KeyPos>
                 ...

      <KeyFrame>  138  </KeyFrame>
      <KeyPos> 192.9 79.0 </KeyPos>
    </TemporalInterpolation>
  </MotionTrajectory>

</Object>
```

Smart MPEG-7 camera

Various statistics and simple decisions can be derived without revealing identity of people

original frame | segmentation mask | bounding box

Benchmarking of Privacy Protection Solutions

Objectives

- **Performance analysis of privacy protection solutions is still lacking**
- **It is paramount to validate proposed privacy protection solutions against user and system requirements for privacy**
  - It is unclear to which extend current privacy protection approaches can be efficiently integrated into existing architectures and deployed in large scale systems

Face Recognition Algorithms

- **Principal Components Analysis (PCA)**
  - Also known as eigenfaces
  - A linear transformation is applied to rotate feature vectors from the initially large and highly correlated subspace to a smaller and uncorrelated subspace

- **Linear Discriminant Analysis (LDA)**
  - LDA aims at finding a linear transformation which stresses differences between classes while lessening differences within classes (a class corresponds to all images of a given individual)

Face Identification and Evaluation

- **Preprocessing to reduce variations between images**
  - Face alignment aligned using eye coordinates
  - Pixel values equalization, contrast and brightness normalization
- **Training**
  - Create the subspace into which test images are subsequently projected and matched
- **Testing**
  - A distance matrix is computed in the transformed subspace for all test images
  - Two image sets are defined:
    - gallery set is made of known faces
    - probe set corresponds to faces to be recognized.
- **Performance analysis**
  - For each probe image, the recognition rank is computed
    - rank 0 means that the best match is of the same subject
    - rank 1 means that the second best match is of the same subject, etc.
  - The cumulative match curve is obtained by summing correct matches for each rank
- **Standard training, gallery and probe sets from the FERET test**

Face Identification and Evaluation

- **Principal Components Analysis (PCA) (aka eigenfaces)**
- **Linear Discriminant Analysis (LDA)**

- **Preprocessing to reduce variations between images**
- **Training**
  - Create the subspace into which test images are subsequently projected and matched
- **Testing**
  - A distance matrix is computed in the transformed subspace for all test images
  - Two image sets are defined:
    - gallery set is made of known faces
    - probe set corresponds to faces to be recognized.
- **Performance analysis**
  - Cumulative match curve based on the recognition rank
- **Standard training, gallery and probe sets from the FERET test**

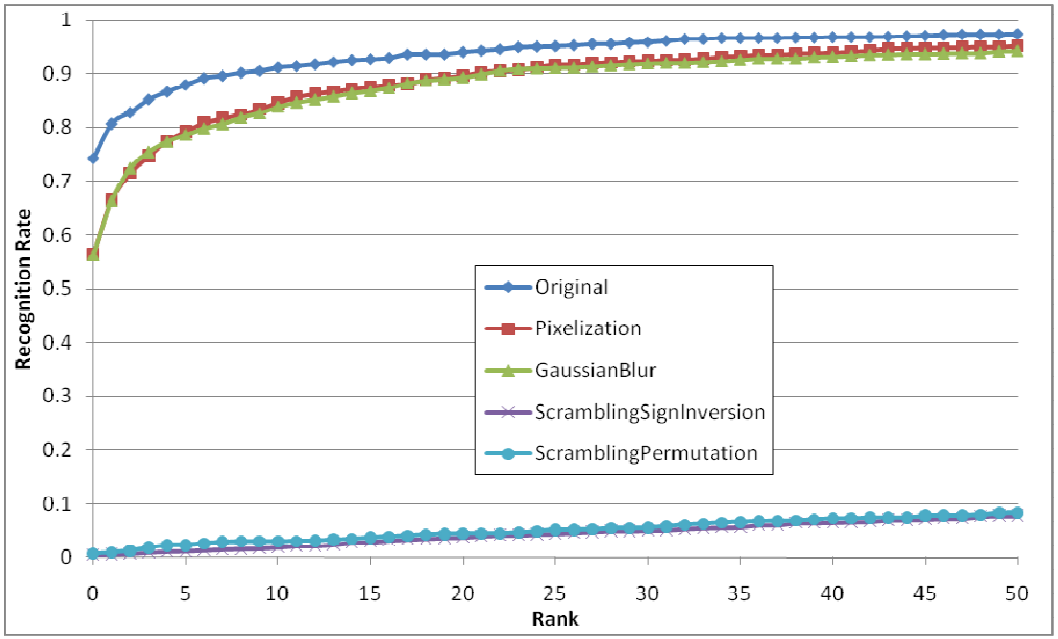
Test Data

- **Grayscale Facial Recognition Technology (FERET)**
  - Although it is not representative of typical video surveillance footage, this database is widely used for face recognition research
  - We consider a subset of 3368 images of frontal faces for which eye coordinates are available
  - Images have 256 by 384 pixels with eight-bit per pixel
  - We further consider two series of images denoted by ‘fa’ and ‘fb’
    - ‘fa’ indicates a regular frontal image
    - ‘fb’ indicates an alternative frontal image, taken within seconds of the corresponding ‘fa’ image, where a different facial expression was requested from the subject.

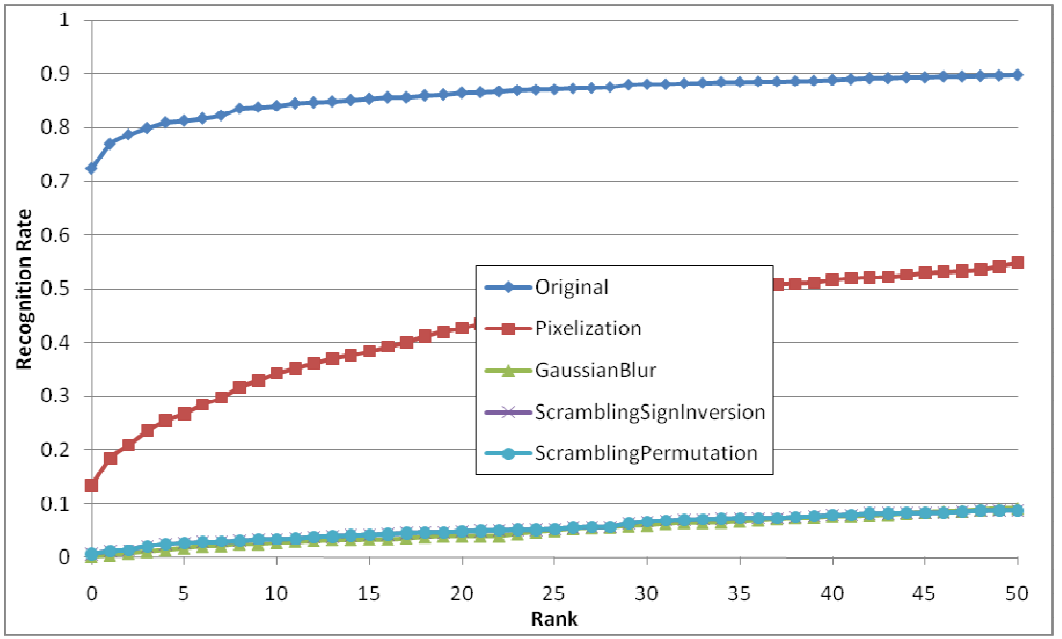
- **Standard training, gallery and probe sets from the FERET test**
  - Training set: 501 images from the ‘fa’ series
  - Gallery set: 1196 images from the ‘fa’ series
  - Probe set: 1195 images from the ‘fb’ series

Performance Analysis – Attack #1

- **Simple attack**
  - Training and gallery sets are made of unaltered images
  - Probe set corresponds to images with privacy protection
  - In other words, altered images are merely processed by the face recognition algorithms without taking into account the fact that privacy protection tools have been applied.



PCA



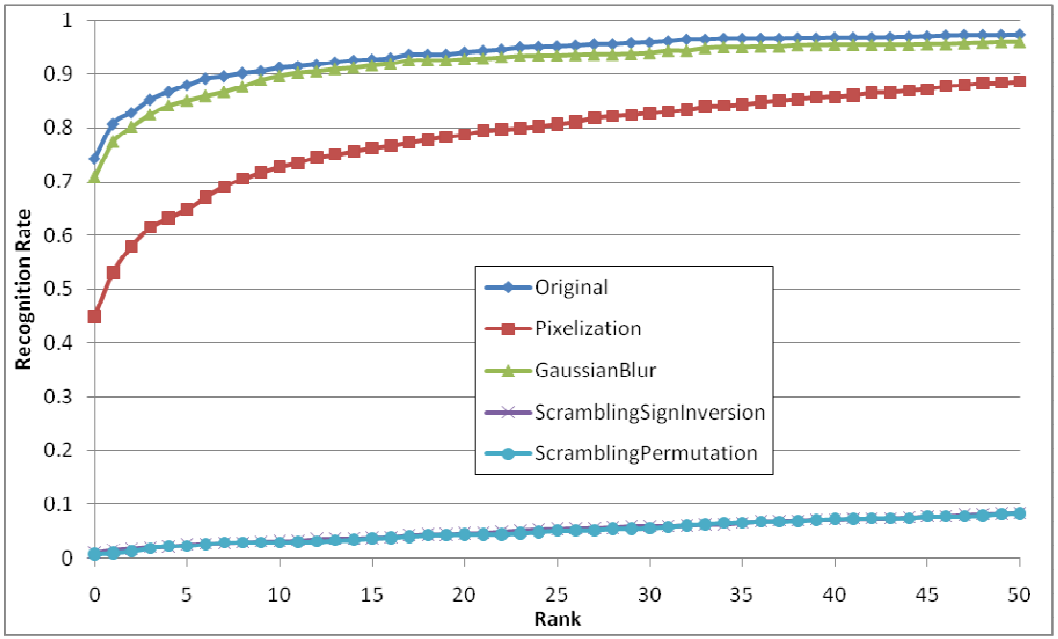
LDA

Performance Analysis – Attack #1

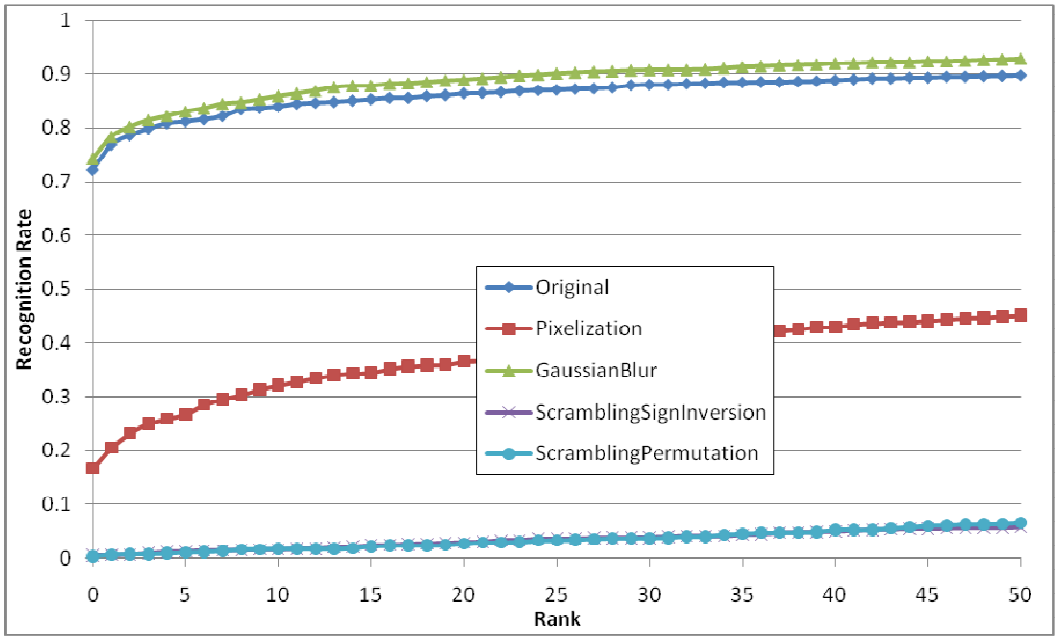
- For both PCA and LDA schemes applied on original images, recognition rate is superior to 70% at rank 0 (i.e. the best match is of the same subject as the probe), and superior to 90% at rank 50
- When applying a Gaussian blur, the performance drops radically for LDA. However, recognition rate remains high for PCA with 56% success at rank 0
- Pixelization fares worse. The recognition rate is 56% and 13% at rank 0 for PCA and LDA respectively
- Results clearly show that both region-based transform-domain scrambling approaches are successful at hiding identity. The recognition rate is nearly 0% at rank 0, and remains below 10% at rank 50, for both PCA and LDA algorithms. In addition, it can be observed that both random sign inversion and random permutation schemes achieve nearly the same performance

Performance Analysis – Attack #2

- **More sophisticated attack**
  - Privacy protection tools are now applied to all images in the training, gallery and probe sets
  - This corresponds to an attacker which gets access to protected data
  - Alternatively, an attacker may attempt replicating the alteration due to privacy protection techniques on his own training and gallery sets



PCA



LDA

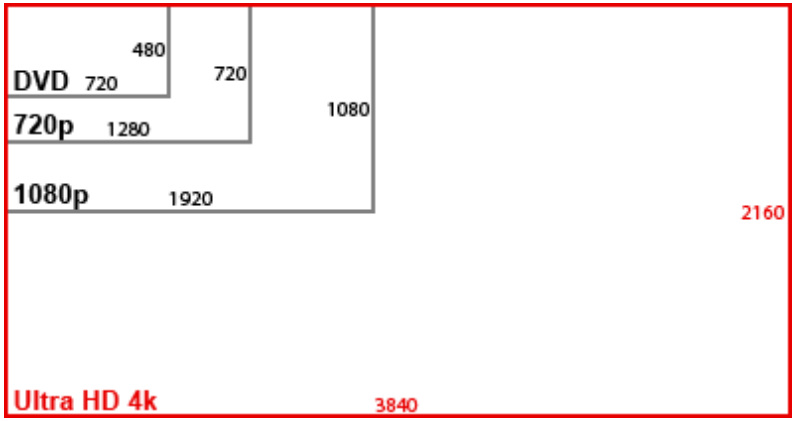
Performance Analysis – Attack #2

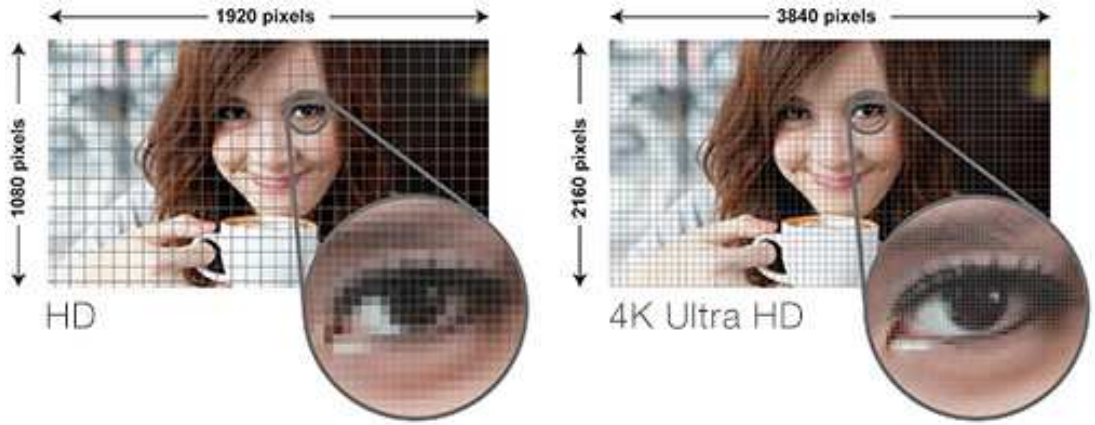
- With Gaussian blur, the performance remains nearly identical. It even improves slightly for LDA
- Pixelization is not much better at hiding facial information. The recognition rate is still 45% and 17% at rank 0 for PCA and LDA respectively
- Finally, both region-based transform-domain scrambling approaches are again successful at hiding identity. The recognition rate is nearly 0% at rank 0 for both PCA and LDA algorithms.

New Trends in Imaging Technology

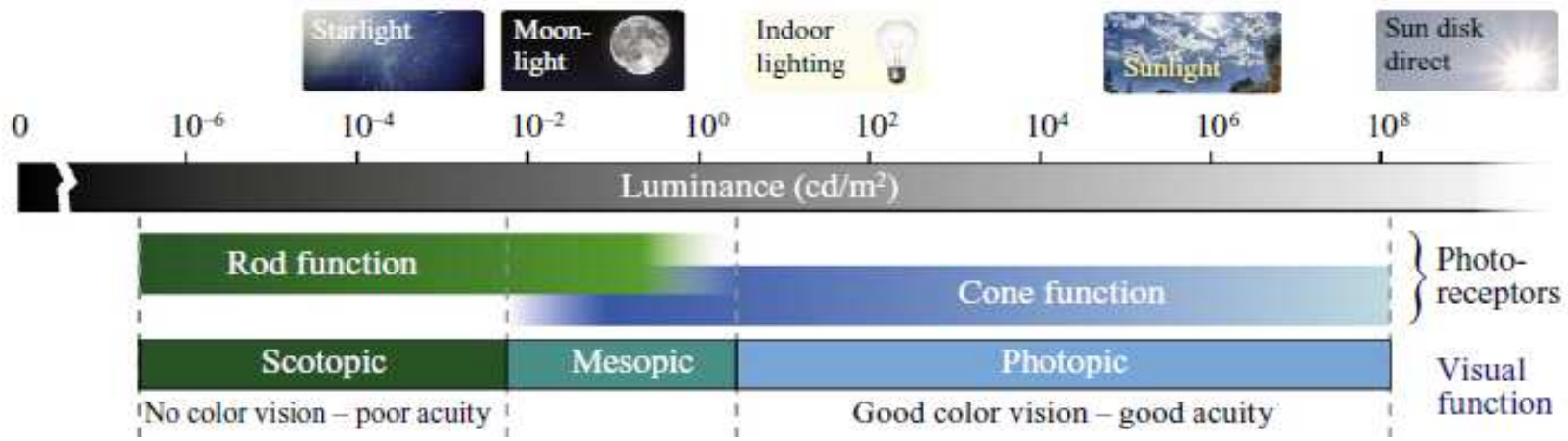
Trends

- **Always more and more…**
- **Higher spatial and temporal resolutions**
  - Ultra High Definition (UHD), 4K, 8K
  - High Frame Rate (HFR)
- **Higher pixel depth**
  - High Dynamic Range (HDR)
- **More views**
  - 3D, multi-view, free viewpoint
  - Lightfield image representation





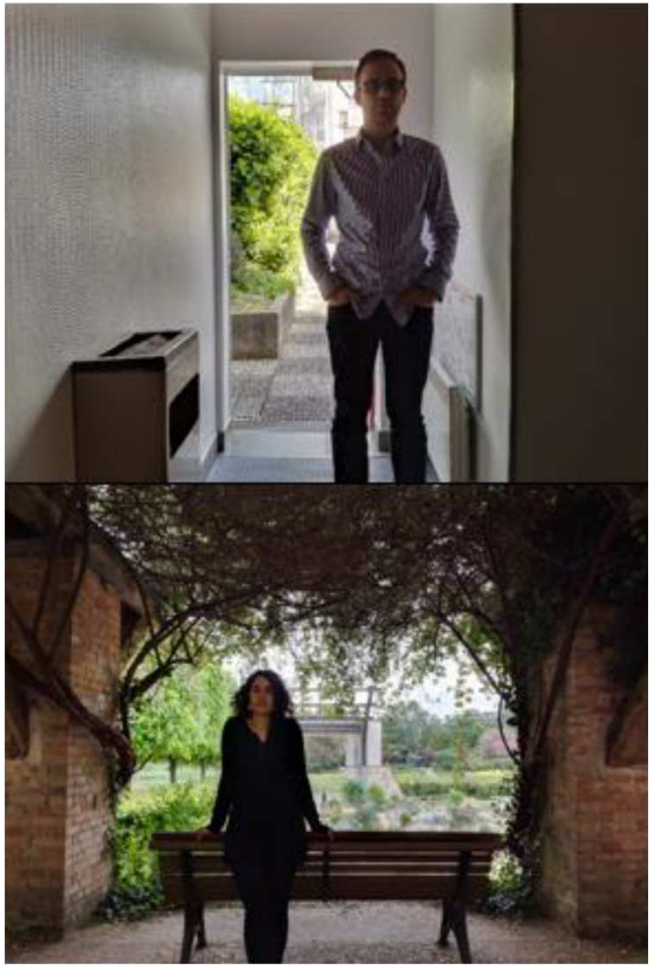
High Dynamic Range and Human Visual System



- **Human Visual System can adapt to a very large range of light intensities**
  - At a given time: 4-5 orders of magnitude
  - With adaptation: 14 orders of magnitude
    - *Mechanical, photochemical and neuronal adaptive processes*

High Dynamic Range

without HDR



with HDR

Enhanced contrast with ability to capture details in both dark and bright regions

High Dynamic Range

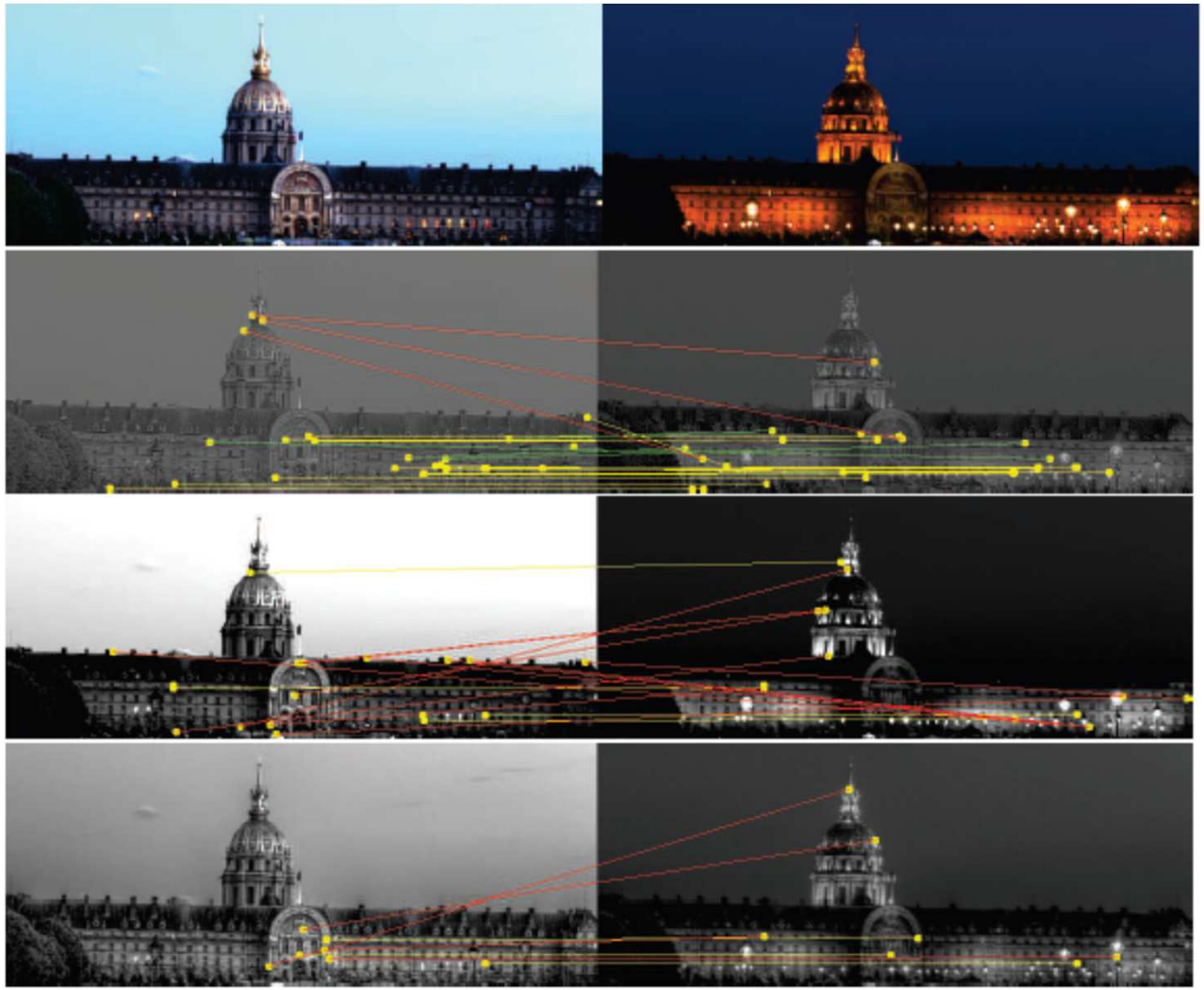


Image matching using SURF

2 HDR images (log scaling)

Proposed descriptor-optimal tone mapping operator (DoTMO)
(11 correct and 3 incorrect matches).

Reinhard TMO
(3 correct and 11 incorrect matches).

MantiukTMO
(4 incorrect and 3 correct matches).

Correct and incorrect matches are shown with yellow and red lines respectively. Green lines represent the special case of mismatch due to repetitive structure.

Potential to greatly improve computer vision algorithms performance

New devices

- **Drones**
  - Defense & security
  - Increasing civilian use

- **Mini-drones**
  - Cheap and easy to deploy
  - Mostly unregulated

Wrap-up

Conclusions

- **Privacy protection in visual data**
  - Difficult and complex problem!

- **Challenges**
  - Application- and context-dependent
  - User and system requirements have to be better understood
  - Lack of thorough performance analysis
  - Systematic security analysis
  - Effective integration into existing large-scale systems
  - New imaging technologies
  - Lack of business incentives

The Last Word

Thank you for your attention !!

Any questions ?

frederic.dufaux@l2s.centralesupelec.fr

References

- F. Dufaux and T. Ebrahimi, Scrambling for Privacy Protection in Video Surveillance Systems, IEEE Trans. on Circ. Syst. for Video Tech., vol. 18, no. 8, pp. 1168-1174, Aug. 2008.
- F. Dufaux and T. Ebrahimi, H.264/AVC Video Scrambling for Privacy Protection, in Proc. IEEE International Conference on Image Processing (ICIP'2008), San Diego, CA, Oct. 2008.
- F. Dufaux and T. Ebrahimi, Recent Advances in MPEG-7 Cameras, in SPIE Proc. Applications of Digital Image Processing XXIX, San Diego, CA, August 2006.
- F. Dufaux and T. Ebrahimi, A Framework for the Validation of Privacy Protection Solutions in Video Surveillance, in Proc. IEEE International Conference on Multimedia & Expo (ICME 2010), Singapore, July 2010.

- F. Dufaux, P. Le Callet, R. Mantiuk, M. Mrak, High Dynamic Range Video - From Acquisition, to Display and Applications, Academic Press, 2016.
- A. Rana, G. Valenzise, F. Dufaux, Learning-Based Tone Mapping Operator for Image Matching, in Proc. IEEE International Conference on Image Processing (ICIP'2017), Beijing, China, Sept. 2017.
- A. Rana, G. Valenzise, F. Dufaux, Learning-based Adaptive Tone Mapping for Keypoint Detection, in Proc. IEEE International Conference on Multimedia & Expo (ICME'2017), Hong Kong, July 2017.