Fundamentals of Image Processing

▶ Lecture 1: introduction and basics ◀

Master of Computer Science
Sorbonne University

September 2022

Outline

Semester organization

Brief history

Acquisition

Definitions

Applications

Important points

Organization

- Group 1 (in French)
    - Lecture on Tuesday 13h45-15h45: Dominique Béréziat
    - Tutorial/practical works on Thursday 13h45-18h: Dominique Béréziat
- Group 2 (in English)
    - Lecture on Monday 13h45-15h45: Isabelle Bloch
    - Tutorial/practical works on Tuesday 8h30-12h30: Garance Martin and Théo Archambault

Important points

- Materials:
  - lecture material (these slides) and bibliography available online on BIMA web site `https://www-master.ufr-info-p6.jussieu.fr/parcours/ima/bima/`
    Moodle: `https://moodle-sciences-22.sorbonne-universite.fr/course/view.php?id=2631`
  - tutorial and practical works material available online on BIMA web site, printed version available at ALIAS office (see https://alias-asso.fr/ for location and schedule)
- Course evaluation based on practical works and final written exam.

Lectures overview

1. Fundamentals tools:
    - introduction and basics of image processing
    - basic operations and image enhancement
    - continuous Fourier transform
    - sampling, discrete Fourier transform.
    - spatial filtering and filtering in frequency domain
    - edge detection
2. Applications:
    - extraction of image primitives
    - segmentation
    - characterization of image primitives
    - introduction to pattern recognition

Tutorial and practical works overview

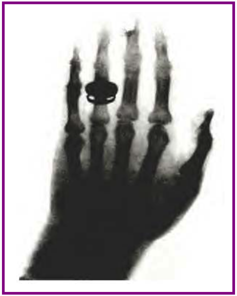
1. Basic tools for image processing:
    - basic operations, image enhancement
    - continuous Fourier transform
    - image digitization, aliasing
    - spatial filtering, filtering in frequency domain
    - edge detection
2. Applications:
    - extraction of image primitives: Harris detector
    - segmentation
    - indexation by color descriptors
    - eigenfaces
3. Official scheduling:
   `https://cal.ufr-info-p6.jussieu.fr/master/`

Outline

The Origins of Image Processing

Some important dates

- 1826: first photograph in history (Joseph Nièpce)
- 1839: photography (Louis Jacques Mandé Daguerre)
- 1895: cinematograph (frères Lumière)
- 1885: X-rays (Röntgen)



The origins of image processing

Some important dates

- 1920: the press exchange images between London and New-York by radio transmission
  $\hookrightarrow$ Bartlane system for image coding (5 levels, then 15 levels in 1929)
- 1960s: conquest of space and diffusion of images
- Late 60s, early 70s: medical imaging (tomography)
- Since the 1970s: constant increase in the use of images
  $\hookrightarrow$ geography, biology, medicine, nuclear, internet, television, satellite, microscopy, video monitoring...

Outline

Semester organization

Brief history

Acquisition

Definitions

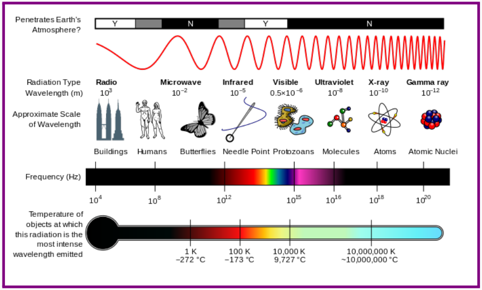
Applications

Image acquisition

- Acquisition = Image formation
- Visible or non-visible spectrum
- Color
- Acquisition *vs* perception

The different types of images and their use

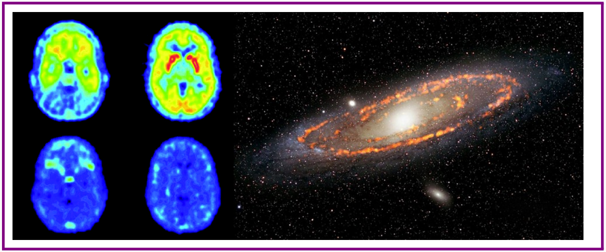
- The electromagnetic spectrum is the main source of energy for the images.
- There are many other sources: acoustics, electronics, ultrasound...
- Electromagnetic waves: sinusoids at different wavelengths containing a quantity of energy, or photons.



(from Wikipedia)

Gamma-ray imaging

- Mainly used in:
  - Nuclear medicine: detecting metabolic, functional... phenomena (infections, tumors)
    $\hookrightarrow$ PET (positron emission tomography)
  - Astronomy: measurement of the force of radiation (novae, super novae)

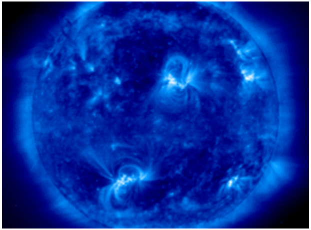
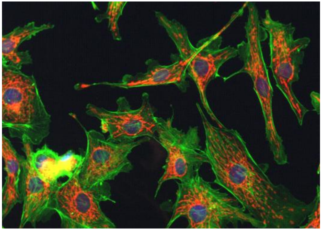


X-ray imaging

- The source of the electromagnetic spectrum in use for the longest period of time
- Rays penetrating "soft matter" (skin, muscles, organs...)
- Mainly used in:
    - Medicine: for example, to detect and localize pathologies $\hookrightarrow$ radiography, angiography, CT scan
    - Industry, astronomy

Imaging in ultraviolet band

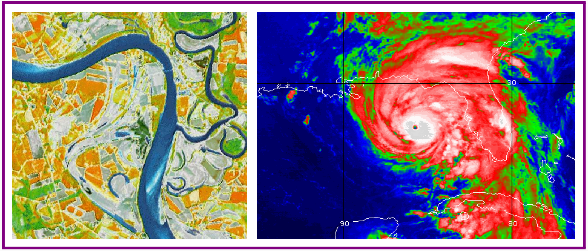
- Ultraviolet is not visible.
- Mainly used in:
  - Microscopy: various fluorescence to highlight different areas of the same image
  - Astronomy, lithography, biology...

The non-visible electromagnetic spectrum

Imaging in infrared band

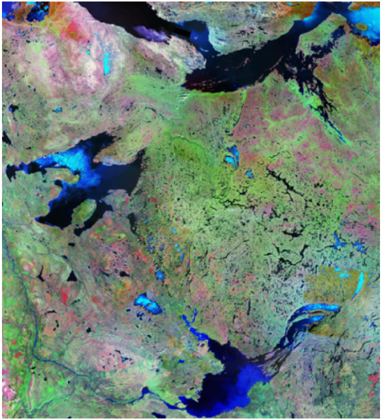
- Often used in conjunction with the visible spectrum
- Mainly used in:
    - Remote sensing: multi-spectral images
      $\hookrightarrow$ weather forecasting, ocean observation, cartography, agriculture, geology ...
    - Microscopy, industry...
    - Photography, video...
    - Astronomy, video monitoring



The non-visible electromagnetic spectrum

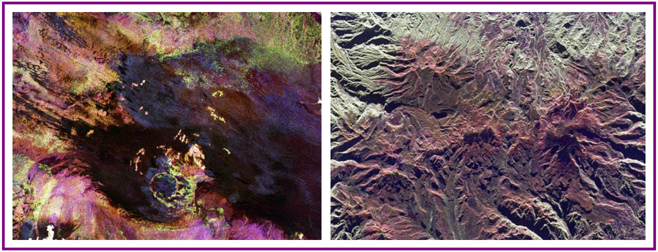
Example of multispectral images: LANDSAT

- 1 (blue): differentiation soil/vegetation, coastal zones
- 2 (green): vegetation
- 3 (red): differentiation of plant species
- 4 (infrared): biomass
- 5 (infrared): snow/cloud differentiation
- 6 (infrared): heat
- 7 (infrared): lithological (rocks)
- hyperspectral imaging: $>$ 200 spectra!



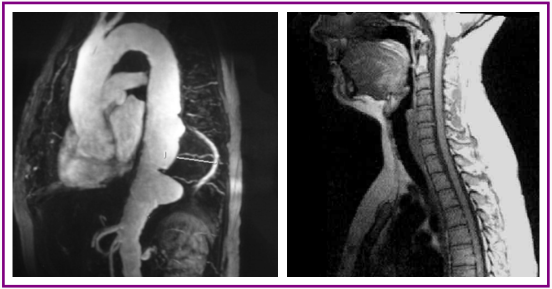
Imaging in microwave band

- Used mainly in radar imaging to see inaccessible areas (clouds)
  $\hookrightarrow$ cartography, agriculture...



Imaging in radio band

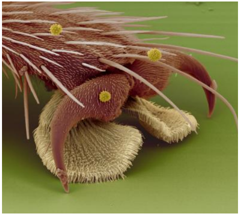
- Mainly used in:
    - Medicine: magnetic resonance imaging (MRI)
    - Astronomy...



Non electromagnetic waves

Other imaging modalities

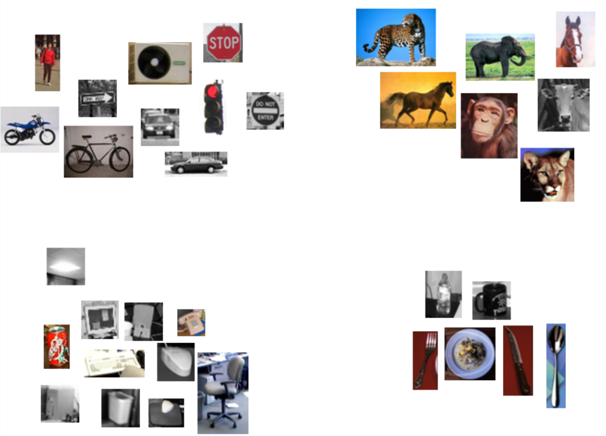
- Ultrasound imaging (acoustic waves): medicine, geology
- Transmission Electron Microscope (TEM): biology, medicine, nanotechnology, molecules, even (big) atoms!




The visible electromagnetic spectrum

The visible spectrum

- The images of everyday life: photography, natural images.

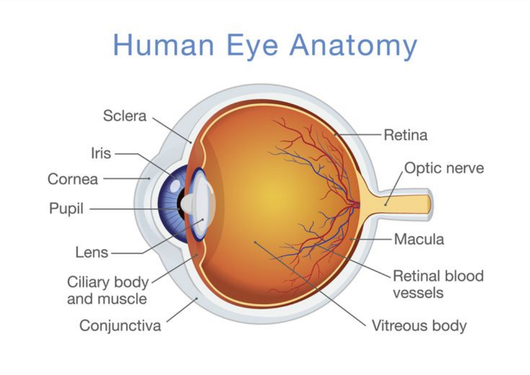


The visible electromagnetic spectrum

The visible spectrum

- The images of everyday life: photography, natural images.
- Challenges from a Digital Image Processing point of view: Analyzing data content with often uncontrolled acquisition conditions.

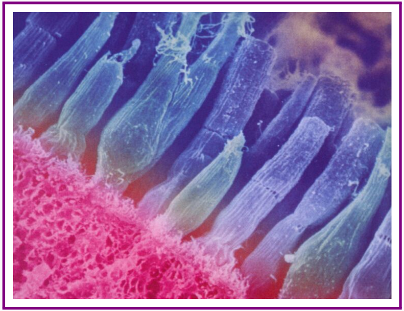
The human visual perception



- The fovea is the region where vision is most accurate and sensitive.
- The eye is moving to align the fovea, the optical axis and the target object.

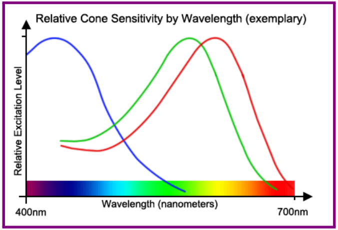
The human eye

- Two types of photoreceptors in the retina (with different distributions):
  - Cones cells: three different kinds of pigments response to short (blue), medium (green) and long (red) wavelengths of light
    $\hookrightarrow$ provide the eye's color sensitivity.
  - Rod cells: not sensitive to color but responsible for our dark-adapted vision (scotopic vision).



Role of cones in visual perception

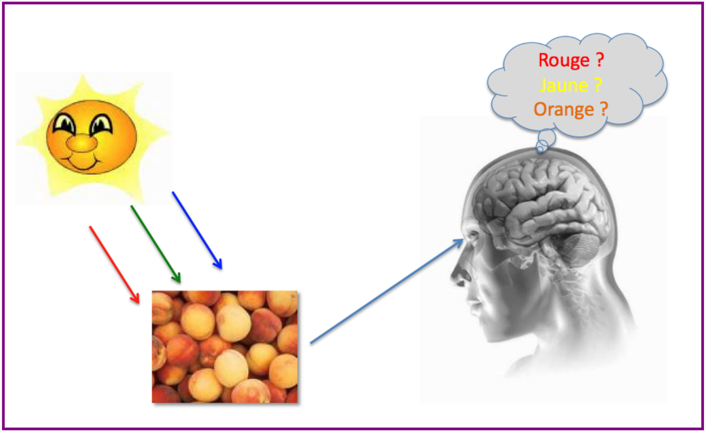
- Provide color sensitive vision
- 3 kinds of cone cells: 64% red sensitive, 32% green sensitive and 2% blue sensitive.
  - Blue cones: short wavelength, $\lambda \approx 420$ nm.
  - Green cones: medium wavelength, $\lambda \approx 530$ nm.
  - Red cones: long wavelength, $\lambda \approx 660$ nm.
- Around 6 to 7 millions of cone cells, mainly located in the Fovea.
- Allow for high resolution vision, daylight vision (photopic vision).



Role rods in visual perception

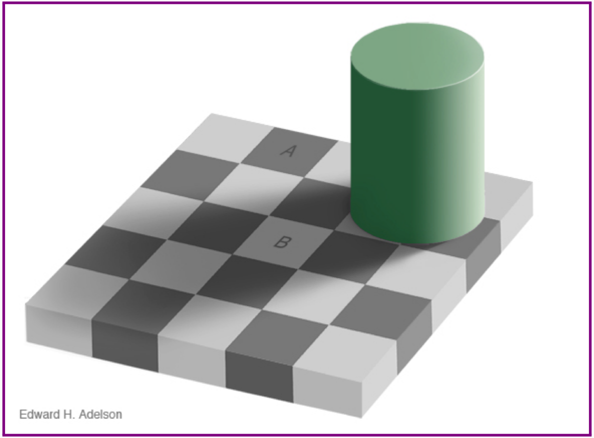
- Much more numerous ($\simeq$ 120 millions) and sensitive to light than cone cells (1000 times more sensitive).
- Very sensitive to low illumination, responsible for dark vision but need a long period of darkness ($\simeq$ 30 minutes) for an optimal vision.
- If only one type of pigment: monochromatic vision, cannot discriminate colors.
- Very sensitive to motion (change of brightness).
- Spread over a large area of the retina and outside of the fovea: responsible for peripheral vision, provide a visual field of 160 degrees.

Human perception

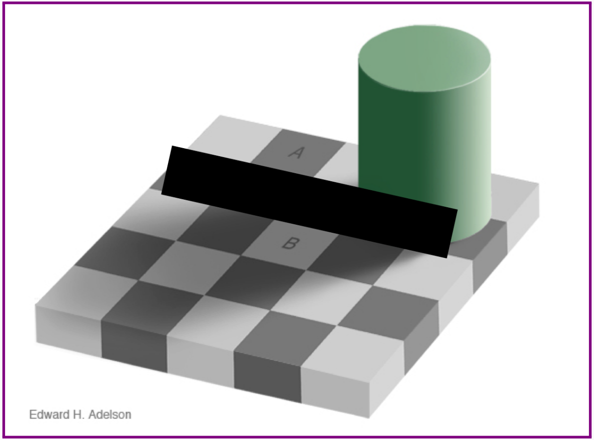


Role of brain: a simple experiment

- What is the brightest square? A or B?

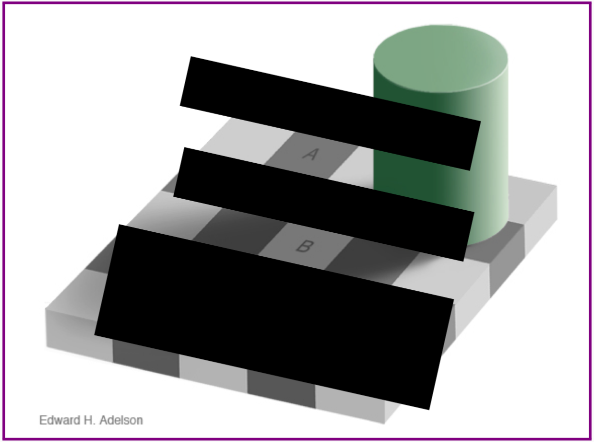


- What is the brightest square? A or B?



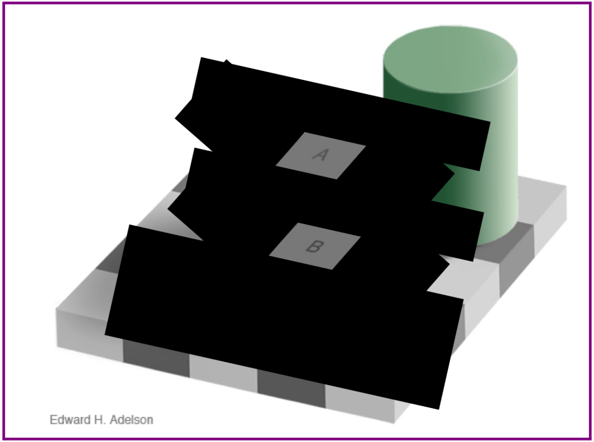
Role of brain: a simple experiment

- What is the brightest square? A or B?



Role of brain: a simple experiment

- What is the brightest square? A or B?

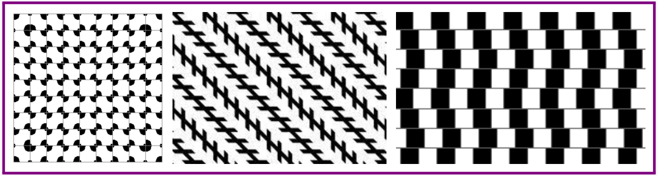


Role of the brain: a simple experiment

- Squares have the same color!
- Made without faking!
- The brain forces itself to picture the squares as they should be: one dark and one bright.
  $\hookrightarrow$ Absolute color is not important
  $\hookrightarrow$ since the contiguous squares are supposed to have different colors.
  $\hookrightarrow$ The brightness perception is not a simple function of intensity, it also depends on other factors, here square boundaries and spatial configurations.
- The human visual perception may be much more different from the reality.

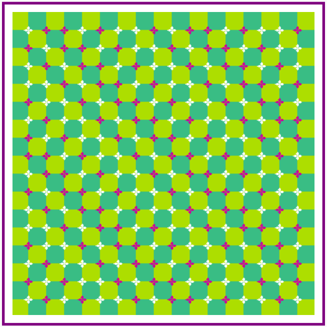
Role of the brain: more examples

Optical illusions



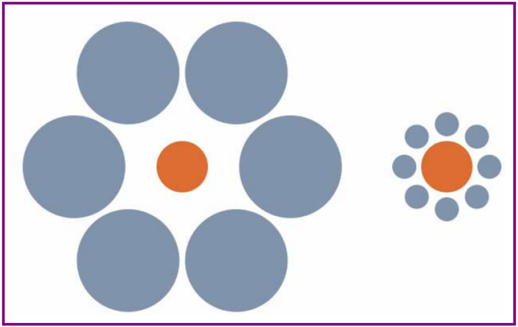
Role of the brain: more examples

Motion illusion



Role of the brain: more examples

Titchener illusion



Role of the brain: more examples

Color illusion

Outline

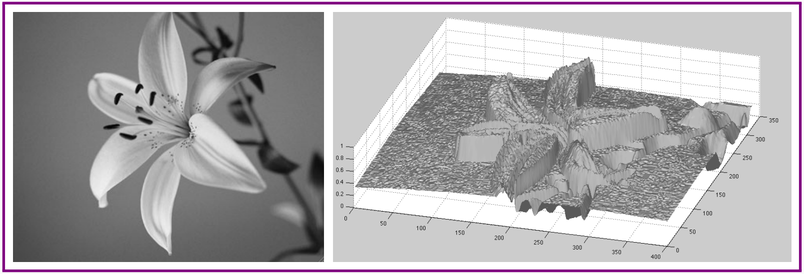
What is an image?

Discrete approach

179	198	204	208	208
182	202	206	208	208
183	203	206	207	208
184	202	205	207	207
184	201	203	205	206
182	199	200	203	205
180	195	197	200	203
177	192	193	198	202
175	189	191	196	201
170	190	192	189	192
167	185	189	188	191

What is an image?

Image as the graph of a 2D function



What is an image?

Definition and notations

- An image may be represented by:
  - a continuous function $f(x, y)$, $x, y \in \mathbb{R}$;
  - a discrete numerical function $f(i, j)$ $i, j \in \mathbb{N}$ and $f(i, j) \in \mathbb{N}$, after digitization.
- Analogical image $\rightarrow$ digital image: Digitization in two steps:
  1. spatial sampling: digitization of the coordinates of the real image,
  2. quantization: digitization of the intensities of the real image.
- A digital image is a finite set of elements named pixels (picture elements) for a 2D image, or voxels (volume elements) for a 3D image.

What is a digital image?

After spatial sampling (notations)

- N: number of rows,
- M: number of columns,
- (i, j): spatial coordinates of a pixel (row i, column j),
- Image (or spatial) domain: the portion of the real plane spanned by the spatial coordinates.

What is a digital image?

After quantization (notations)

- $f(i, j)$ discrete intensity value at pixel (i, j)
- k an intensity value (also called gray scale value), given by f
- m the number of bits required to encode a gray level value
- L the number of intensity values, also called informally the image dynamic range
  $\hookrightarrow L = 2^m$, and $k \in [0, \ldots, 2^m - 1]$

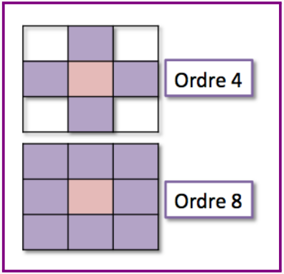
Representing digital images

- 2D-array representation:
  - Cartesian coordinates in $[0, \ldots, N-1] \times [0, \ldots, M-1]$, 2D-array (matrix) f.
  - Image width given by the number of columns M of f, image height given by the number of rows N.
  - Pixel value at row i and column j is given by element (i, j) of matrix f.
- 1D-array (or vector) representation:
  - Image rows are juxtaposed in order to form a vector $v = [0, \ldots, M \times N - 1]^t$
  - The pixel (i, j) is represented by the element of index $(i \times M + j)$ in v.

Neighbors of a pixel

Connectivity

- Fundamental notion in image processing
  - topology
  - contours of objects, regions growing...
- 4-connectivity: the four closest neighboring pixels: N, S, W, E
- 8-connectivity: the eight closest neighboring pixels: N, S, W, E, NE, NW, SE, SW.



Digitization: a crucial step

A subjective vision

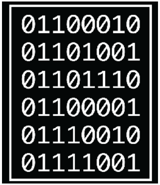
- Spatial sampling defines the size of the smallest element (pixel) of an image.
- Intensity quantization defines the smallest intensity change in an image (always perceptible by human vision?).
- Sampling and quantization define the image size (in bits): $s_b = N \times M \times m$.

What is a digital image?

Types of images

- $m = 1$, $k \in \{0, 1\}$: binary image
- $m = 8$, $k \in [0, \ldots, 255]$: grayscale image
  - usually encoded on 1 byte ($m = 8$)
  - conventionally: black = 0 (no signal, no photons), white = 255 or 1 if normalization

- $m = 24$, $k \in [0, \ldots, 16777215]$: color image (3 bytes)



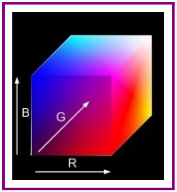
Definition

- Any color can be represented as a linear combination of three primary colors c_1, c_2, c_3.
- Two types of models: additive model, subtractive model.

- Many color spaces: RGB, CMYK ($\sim$ RGB in a subtractive formulation), HSV, YUV, Lab, CIE...

RGB color model

- Additive model, 3 primary colors: red, green and blue, based on a Cartesian coordinate system.



- Luminance: $L = 0.3R + 0.59G + 0.11B$
  - Separation and visualization of each color component, and alternative combinations $\Rightarrow$ Practical work #4

Color image

R

G

B

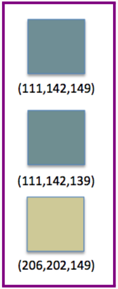
$R \leftrightarrow B$

$R \leftrightarrow G$

RGB color model: limitations

Limitations

- Cannot describe all colors perceptible by the human vision.
- Components are often highly correlated.
- The perceptual difference between two colors is not related to their Euclidean distance in RGB space.



Extensions

- Alternative representation (change of basis) in which components are uncorrelated (PCA), or statistically independent (ICA).
- Use models closer to human perception of colors: HSV, XYZ, YUV, Lab...

HSV color model

Definition

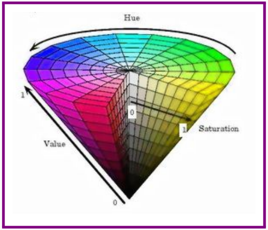
- Three components: Hue, Saturation, Value.
- A polar representation of colors.
- Conversion RBG $\leftrightarrow$ HSV:

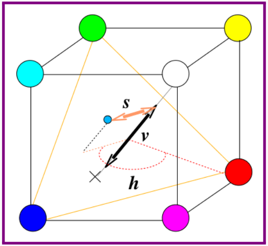
$$V = \frac{R + G + B}{3}$$

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 2\pi - \theta, & \text{otherwise} \end{cases}$$

$$\theta = \arccos\left(\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right)$$

$$S = 1 - \frac{3\min(R, G, B)}{R + G + B}$$





HSV color model: example

- HSV color model: see practical work #7 “visual descriptors”.

Digital image formats

- Raw data
- Standard (free or not): `gif`, `bitmap`, `tiff`, `ppm`, `eps`...
- Standard in medical domain: `DICOM`
- Images may be embedded in a more general data format: `NetCDF`, `HDF5`...
- Some examples:
  - `BMP` (*Bitmap*): array of bits (up to 24 bits per pixel for color representation)
  - `GIF`: 8 bits per pixel, compressed, deprecated by `PNG`
  - `JPG` (*jpeg*): compression standard for photographic images

Some examples of digital images

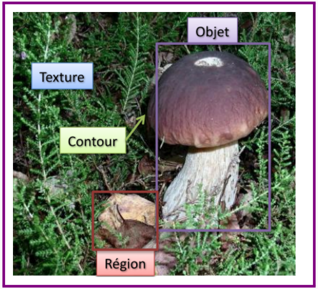
- 2D Image: static scene representing 2D objects by a bi-dimensional array of elementary surfaces (pixels)
- Video: dynamic scene presenting 2D objects evolving in time (moving objects)
  $\hookrightarrow$ set of 2D images, the temporal axis is seen as a third dimension
- 3D image: objects represented by a tri-dimensional array of elementary volumes (voxels)
  $\hookrightarrow$ set of 2D images (for instance, slices of a CT-scan), the third dimension is a spatial dimension
- 3D video: dynamical scene describing 3D moving objects

Image quality

- Sharpness: determines the amount of details an image can convey
- Dynamic range: the range of light levels captured in the image
- Contrast: quality of the dynamic range
- Noise: a random variation of image values, generally coming from the sensor (if the signal is low), depends on the sensor nature, whose distribution law is often unknown
- Geometric distortion: various defaults occurring during image acquisition and spatially deforming the image
- Compression artifacts: visual deformation induced by a high destructive compression (JPEG) or transmission losses

Structures within an image

- Texture: statistical or geometrical distribution of image intensities
- Region: connected set of pixels having the same characteristics (similar intensity or texture, same motion...)
- Edge point: a pixel at the location of a significant local intensity change in the image.
- Contour: connected set of edge pixels
- Object: a region (or group of regions), fully delimited by a contour, and independent of the rest of image
  $\hookrightarrow$ semantic description: we can name the object



Brightness

- The mean of image values:

$$B = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} f(i,j)$$

Important image properties

Contrast

- Definition 1: maximal variation between the lowest and highest values in the image:

$$C_1 = \frac{\max_{i,j}[f(i,j)] - \min_{i,j}[f(i,j)]}{\max_{i,j}[f(i,j)] + \min_{i,j}[f(i,j)]}$$

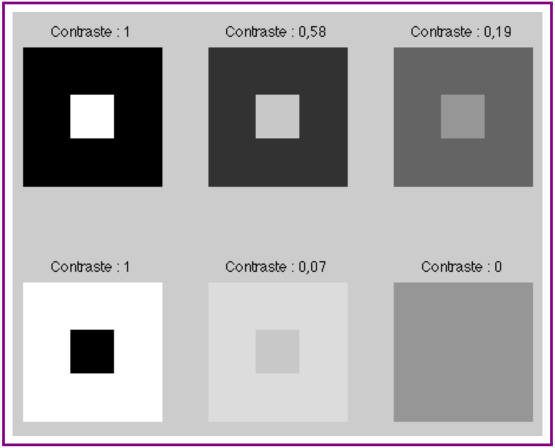
- Definition 2: standard deviation of image values:

$$C_2 = \sqrt{\frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (f(i,j) - \mathrm{B})^2}$$

  Standard deviation is a second moment, C_1 is a first moment.

- Two different images may have the same contrast.

Examples of contrast



Outline

Semester organization

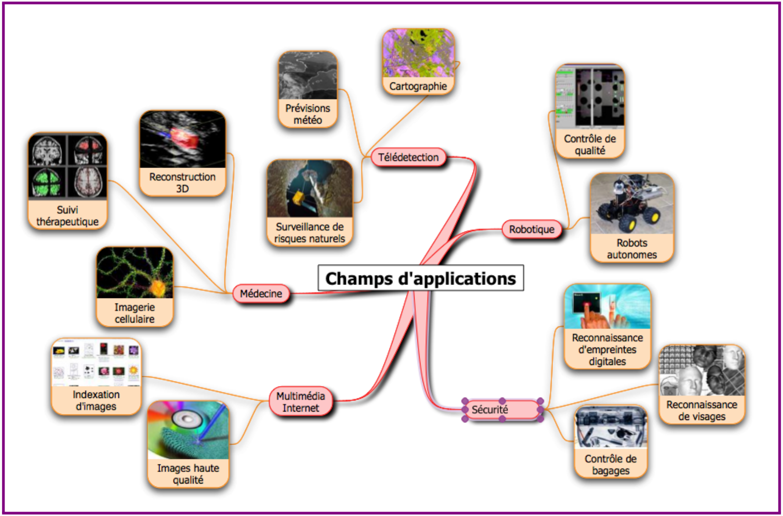
Brief history

Acquisition

Definitions

Applications

Applications

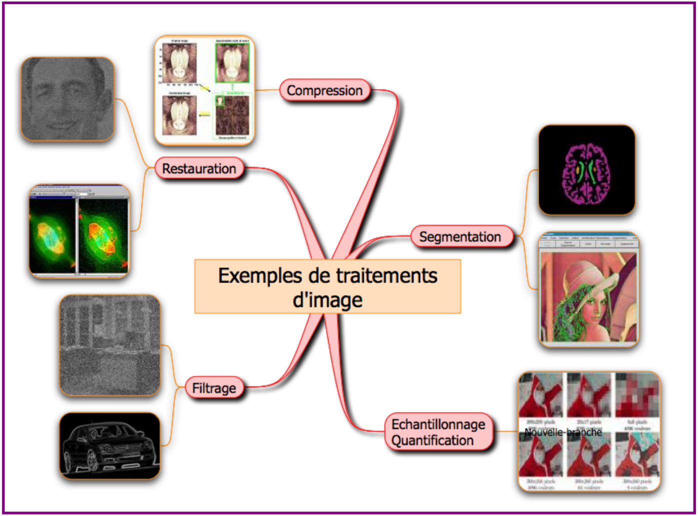


$\Rightarrow$ thousand of potential applications!

Image processing, analysis and understanding: three levels

- Low level:
  - input: image
  - output: image
- Mid level:
  - input: image or attributes
  - output: new attributes
- High level:
  - input: image or attributes
  - output: understanding (semantic description)

Example of low level image processing



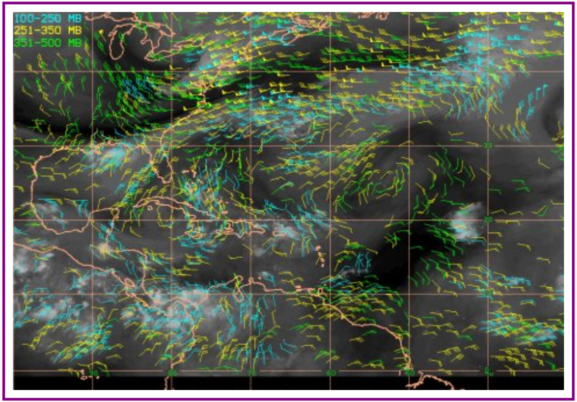
Examples of low level image processing

- Image enhancement: improve the images aspect (for human visualization or further processing)
  $\hookrightarrow$ image restoration (practical work #1), noise reduction, filtering (practical work #4)
- Frequency analysis: change of space representation (change of basis) $\hookrightarrow$ Fourier transform (practical work #2)
- Image acquisition: obtain a digital image
  $\hookrightarrow$ sampling, quantization (practical work #3)
- Image compression : reduction of the amount of data required to represent an image
  $\hookrightarrow$ coding, transmission
- Image segmentation: subdivision of an image into homogeneous regions or objects
  $\hookrightarrow$ edges detection (practical work #5), split & merge (practical work #10)

Examples of mid and high level image processing

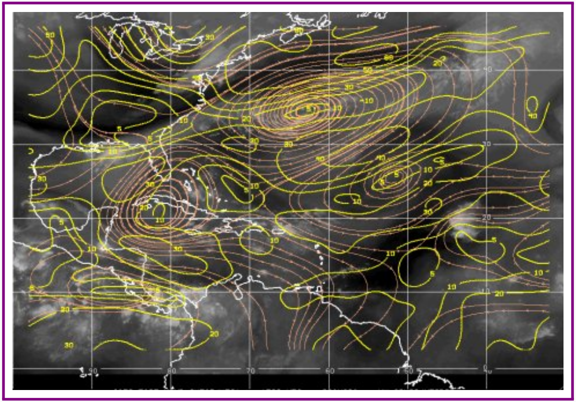
- Representation and description of images: the image is transformed in a set of "concepts" usable by the computer for more advanced algorithms $\hookrightarrow$ Detection of characteristics, Harris corner detector for instance (practical work #6), models (graphs...)
- Image indexing: images are sorted using criteria based on image descriptors
  $\hookrightarrow$ query in image databases (practical work #7)
- Pattern recognition: Assigning a label to an object based on its descriptors $\hookrightarrow$ Face recognition (practical work #8-9)

A first example: meteorological forecasts



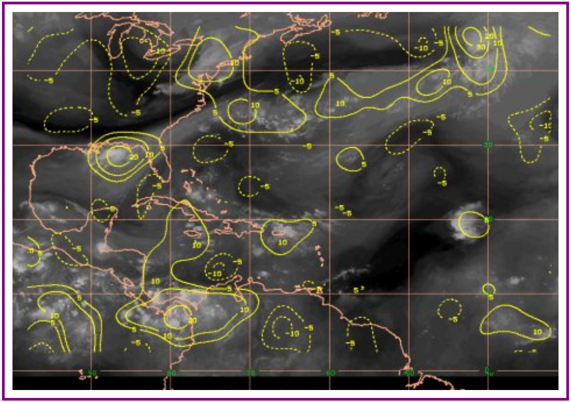
Low level processing: map of vectors velocity

A first example: meteorological forecasts



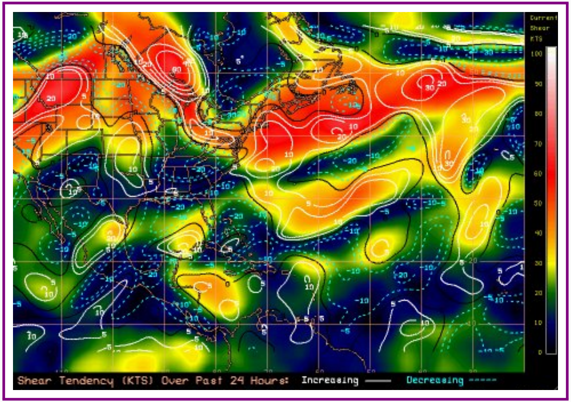
Mid level processing: velocity map $\Rightarrow$ pressure map

A first example: meteorological forecasts



Mid level processing: velocity map $\Rightarrow$ pressure map

A first example: meteorological forecasts



High level processing: pressure map $\Rightarrow$ weather report

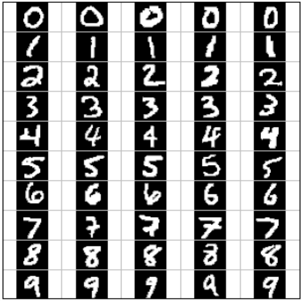
A second example: image understanding

- Image understanding: a broad field in image processing
  - Classification
  - Detection
  - Segmentation
  - Combination: Classification + Detection + Segmentation
  - Global scene understanding
- Image processing and computer vision methods coupled with statistical learning tools and prior knowledge representation

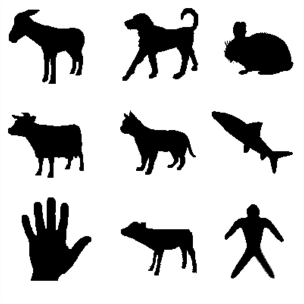
Image classification

Basic cases: $\sim$ resolved

- Basic geometrical objects, centered in the image and segmented
- Annotated databases, extraction of descriptors $\Rightarrow$ prediction of object class



MNIST



Binary shapes

More complex cases

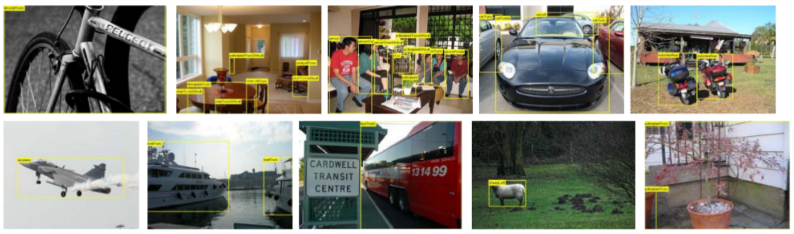
- Several objects, several categories
- Natural images (photography), Flickr $\Rightarrow$ some major issues

PASCAL VOC

Image classification

More complex cases: main challenges

1. Variability: illumination, scale, orientation, view point
2. Object occultation, background clutter
3. Visual variation of objects in a same class



PASCAL VOC: hard cases

Image classification

Challenges

Illumination change

Viewpoint change

Challenges

Scale change

Occultations

Challenges

Deformable object

Background clutter

Image classification

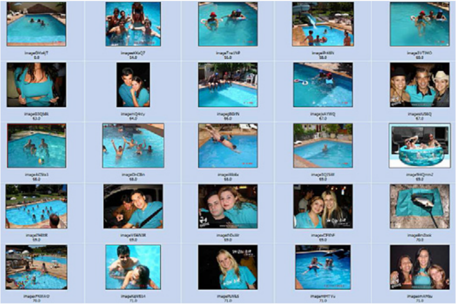
Challenges: high variability of intra-class visual aspect

Semantic query

- Image descriptors: low level (color, texture, form, gradient)
- Semantic descriptors: high level (objects, scenes, abstract concepts)
- Challenge: bridging the semantic gap

Query :

Result of a query based on color descriptors :



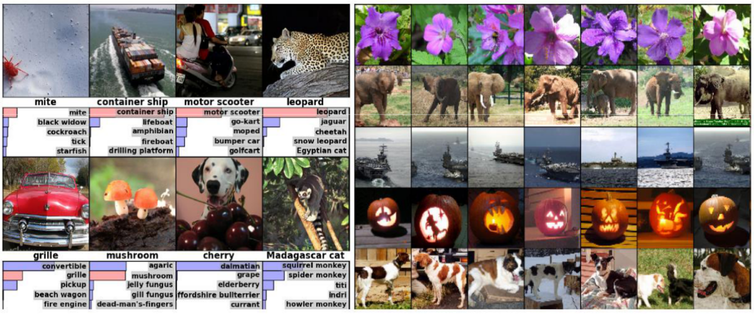
Semantic query

- Search for images by abstract concepts:

- Joy, love, hate...
- Extreme semantic gap

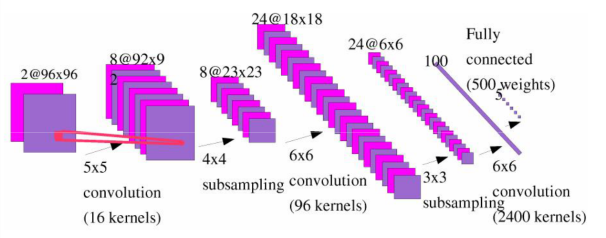
Semantic query

- ImageNet competition during ECCV 2012
  - Classification task: 1000 classes, 1 million images for training
  - Predict the right class: complex



Semantic query

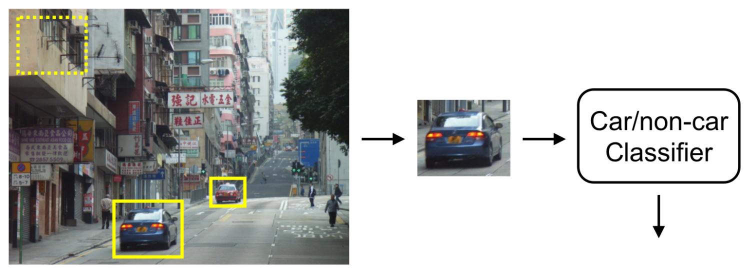
- ImageNet competition during ECCV 2012
  - CNN (Convolutional Neural Network)
  - Very important gain compared to traditional methods
  - Main feature: the semantic gap is bridged by learning internal representations from data



Object detection

Classification *vs* detection: where are the objects of a given semantic category located in the image?

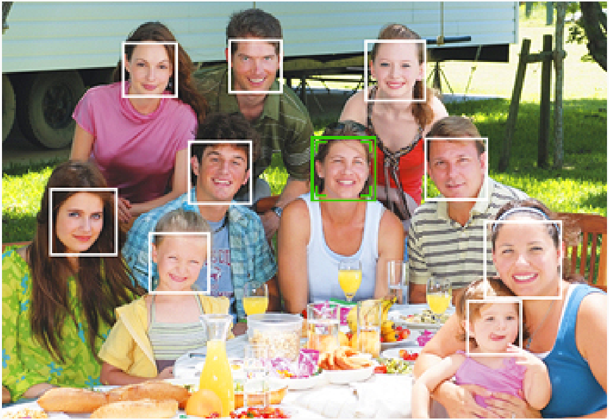
Principle: sliding window



Object detection

Face detection: success!

- Embedded in most of commercial cameras

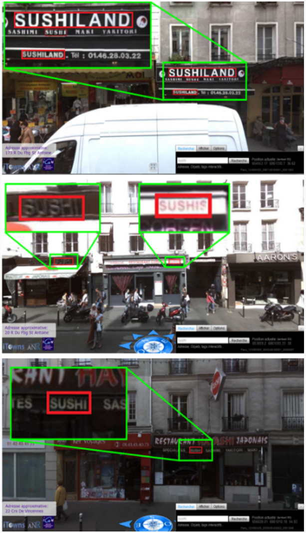


Another example: text detection

- Detect text areas in an image
- Urban context
- Numerous applications

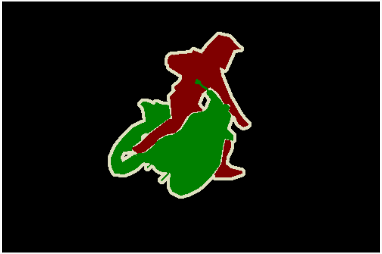
Object detection

Another example: text detection



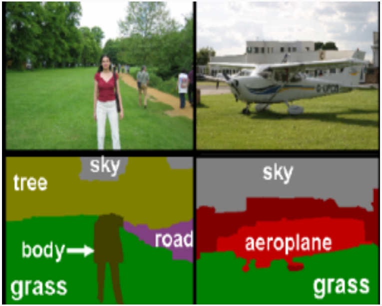
- Combined with OCR: tool for semantic query
- Combining visual information with other sources of information: geolocalization...
- Applications for smartphones

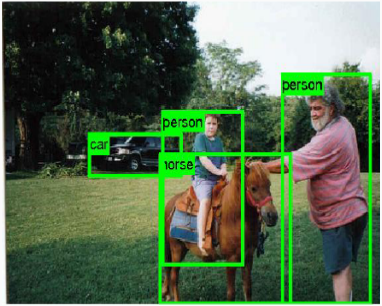
Another example: segmentation
Classification *vs* detection *vs* segmentation: identify pixels belonging to semantic categories in the image



The Holy Quest

- Combine global classification, detection and segmentation





Last example: ID photo compliance

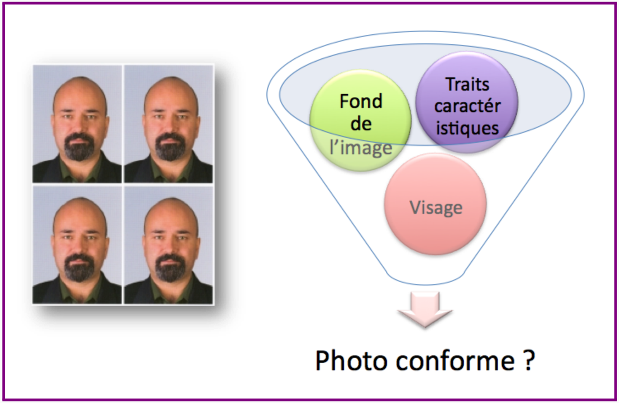
- Linked to the problem of face recognition
- Face detection: identify image regions belonging to the class "face"
- Face identification: identify a particular instance in the class "face"
- Hard problem in the general case

Last example: ID photo compliance

- The ISO/IEC norm (2005)

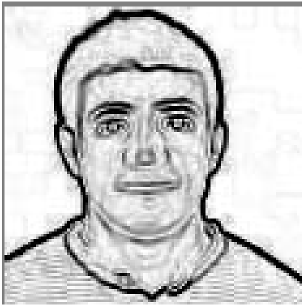
Last example: ID photo compliance

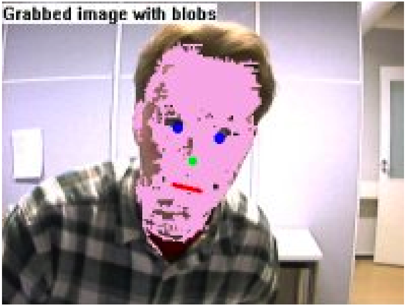
- Many issues to handle



Last example: ID photo compliance

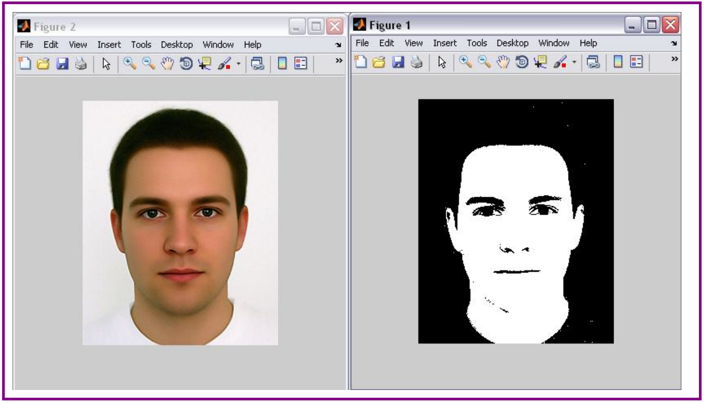
- Face detection and skin threshold





Last example: ID photo compliance

- Skin threshold



Last example: ID photo compliance

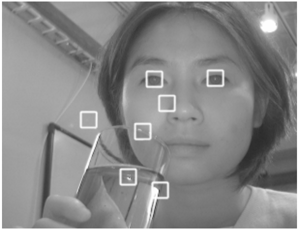
- Detection of characteristic features (eyes, mouth...)

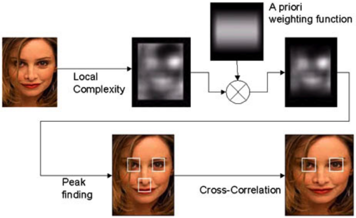
Last example: ID photo compliance

- Detection of characteristic features (eyes, mouth...): easy ?

Last example: ID photo compliance

- Detection of characteristic features (eyes, mouth...): many methods!



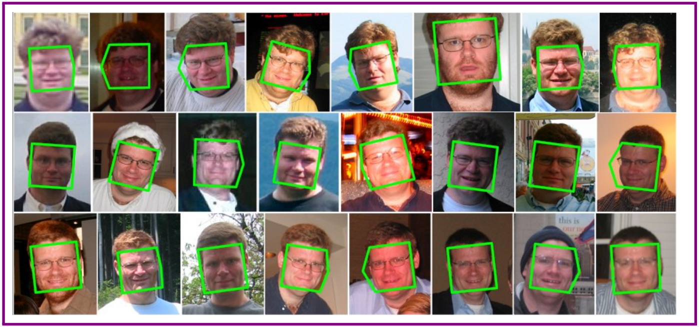


Last example: ID photo compliance

- Issue: multiple representations

Last example: ID photo compliance

- Issue: face orientation



Last example: ID photo compliance

- Background?

Last example: ID photo compliance

Other important issues:

- Potential biases
- Image integrity
- Ethics
- Usage control
- ...