Introduction to Applied Computer Vision

By: Achraf Hsain and Lahoucine Chikry

Human vision vs Computer vision

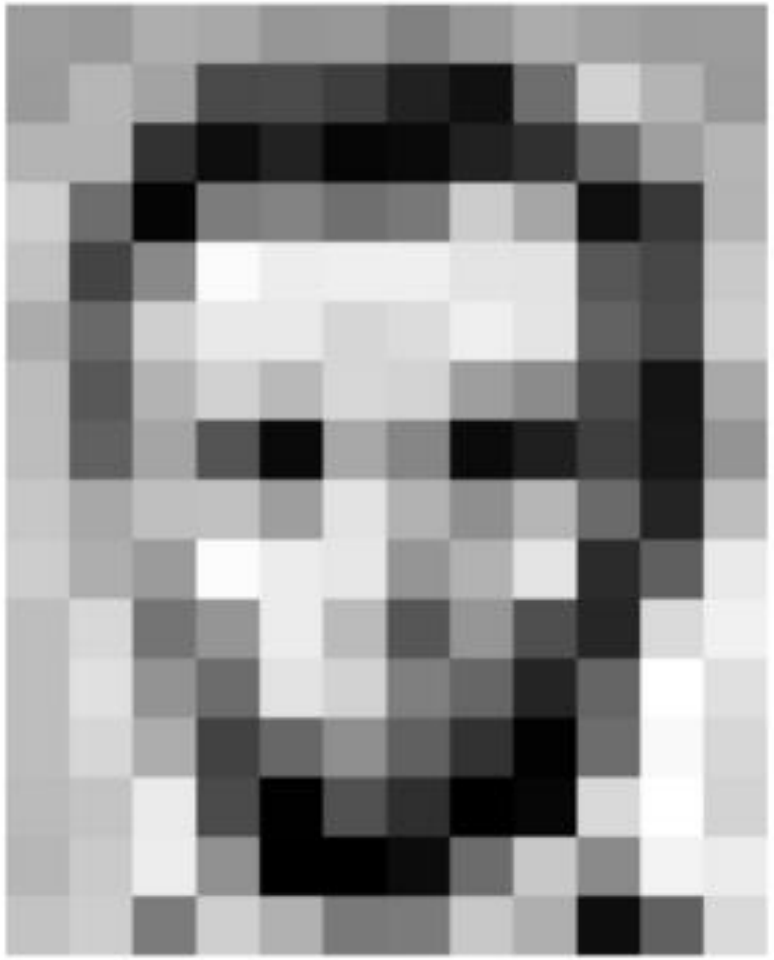
What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

What computers ‘see’: Images as Numbers

What you see



Input Image

What you both see

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Input Image + values

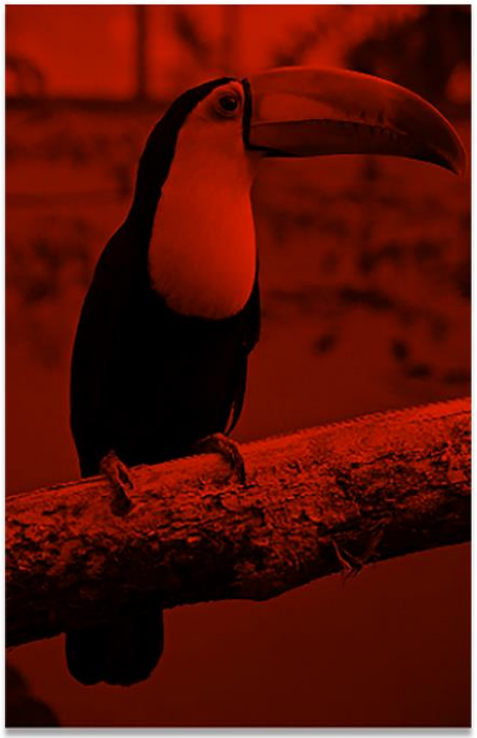
What the computer "sees"

157	153	174	168	150	152	129	151	172	161	155	15
155	182	163	74	75	62	33	17	110	210	180	15
180	180	50	14	34	6	10	33	48	106	159	18
206	109	5	124	131	111	120	204	166	15	56	18
194	68	137	251	237	239	239	228	227	87	71	20
172	105	207	233	233	214	220	239	228	98	74	20
188	88	179	209	185	215	211	158	139	75	20	16
189	97	165	84	10	168	134	11	31	62	22	14
199	168	191	193	158	227	178	143	182	106	36	19
205	174	155	252	236	231	149	178	228	43	95	23
190	216	116	149	236	187	86	150	79	38	218	24
190	224	147	108	227	210	127	102	36	101	255	22
190	214	173	66	103	143	96	50	2	109	249	21
187	196	235	75	1	81	47	0	6	217	255	21
183	202	237	145	0	0	12	108	200	138	243	23
195	206	123	207	177	121	123	200	175	13	96	21

Pixel intensity values
(“pix-el”=picture-element)

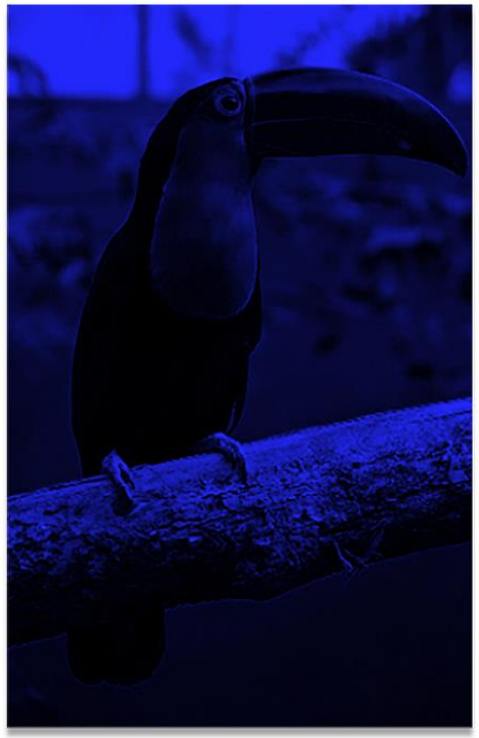
What about Colored Images?

Original Image



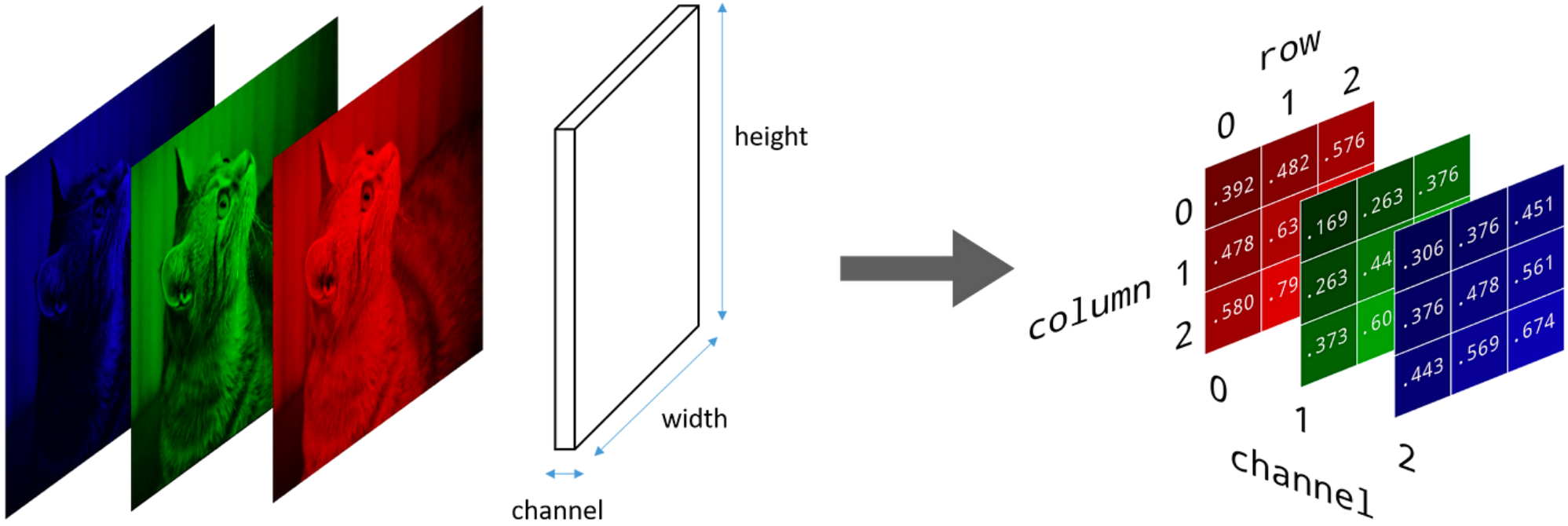
Red Channel

Green Channel



Blue Channel

Colored Images Representation



Computer Vision

- The field of AI that enables machines to interpret and understand the content of digital images or videos.
- Extract meaningful information (e.g., object detection, recognition, and scene understanding) for decision-making.

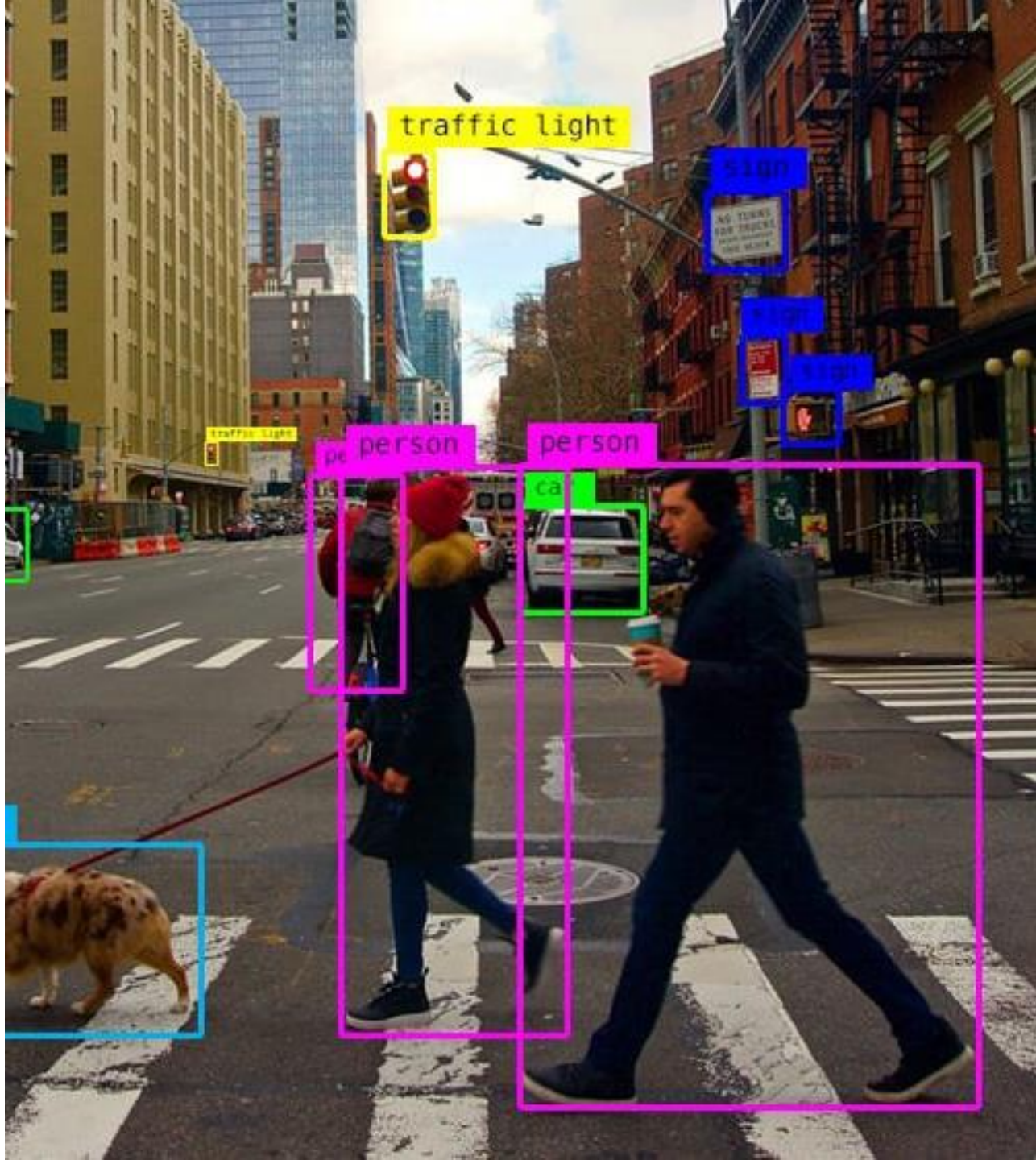
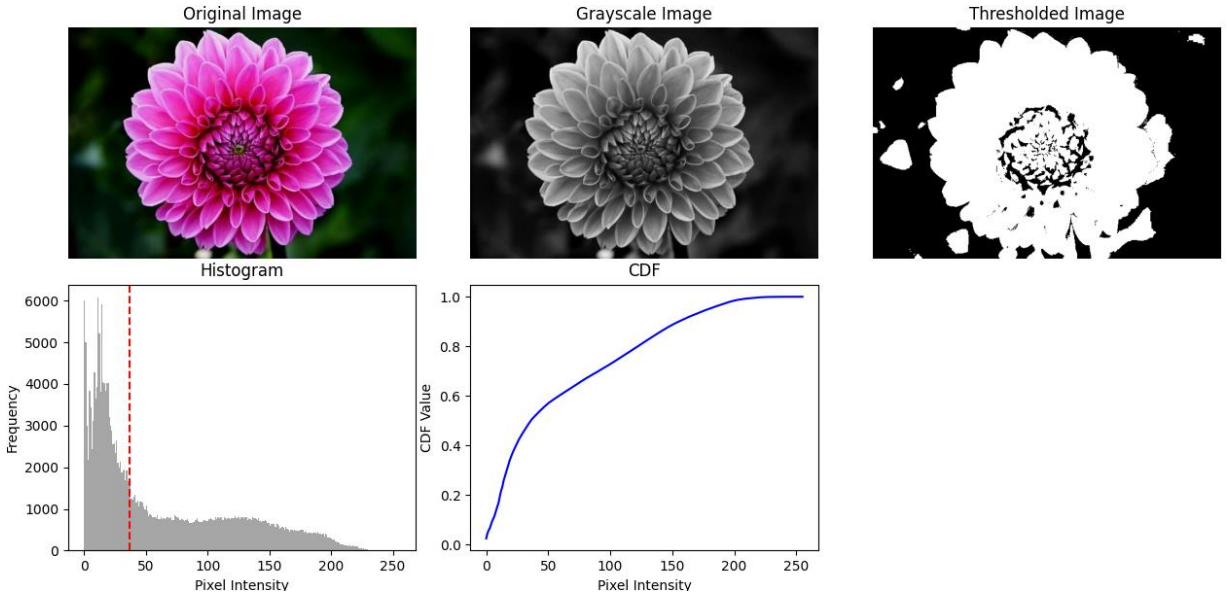


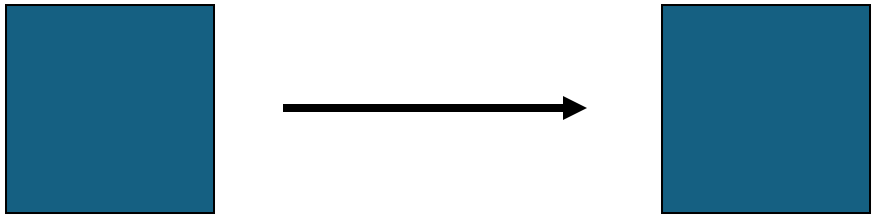
Image Processing



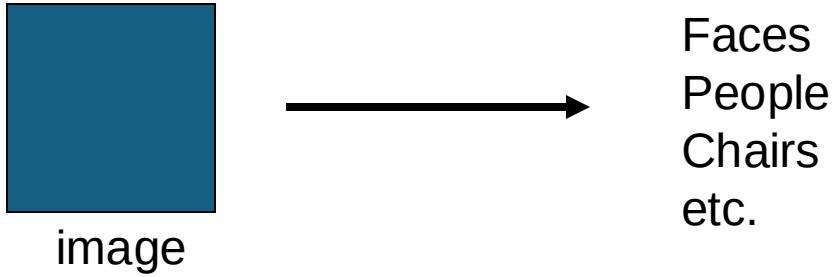
- A subset of signal processing focused on performing operations on images to enhance or manipulate them.
- Improve image quality, extract features, or transform the image for specific purposes.

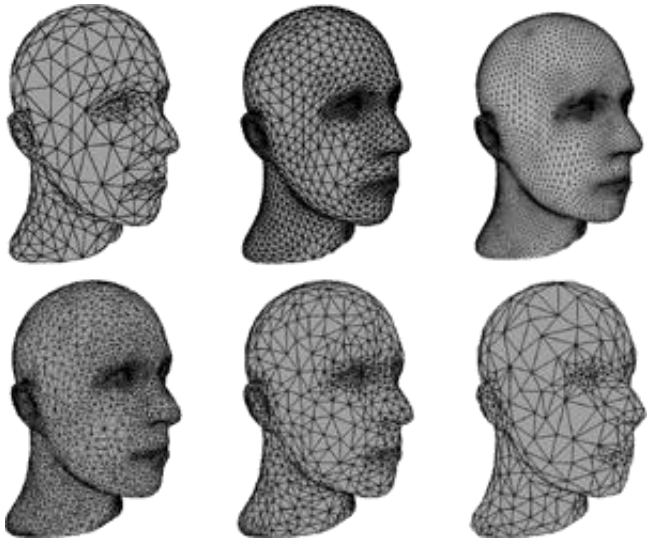
Image Processing vs. Computer Vision

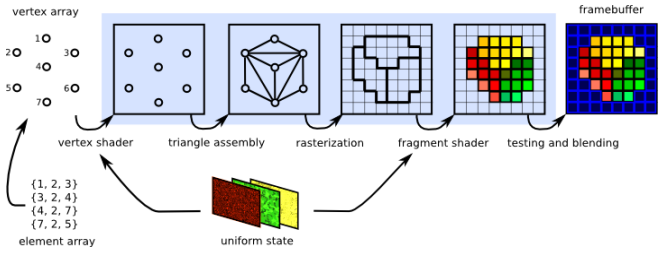
- Image Processing
  - Research area within electrical engineering/signal processing
  - Focus on syntax,
    low level features



- Computer Vision
  - Research area within computer science/artificial intelligence
  - Focus on semantics,
    symbolic or geometric
    descriptions







Computer Graphics

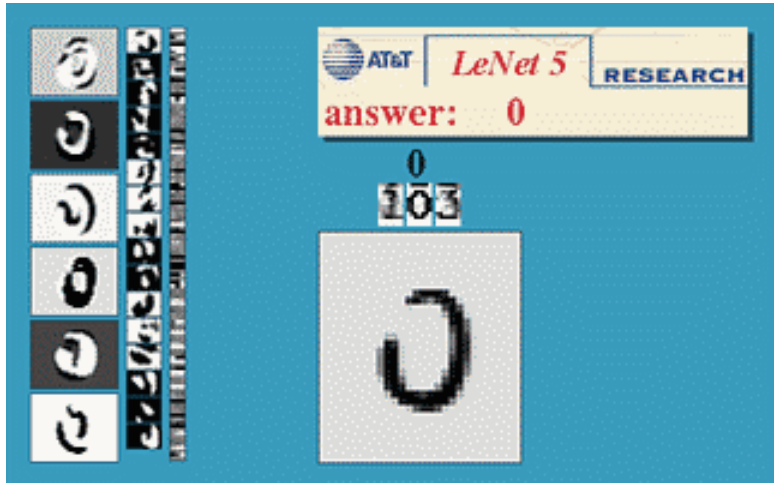
- The creation and manipulation of visual content (images, animations, 3D models) through computational methods.
- Generate realistic or stylized visuals, often for entertainment or design purposes.

Machine Vision

- The industrial application of computer vision technologies to automate visual inspection and quality control tasks.
- Enable machines to inspect, measure, and analyze objects during production or operation.

Applications

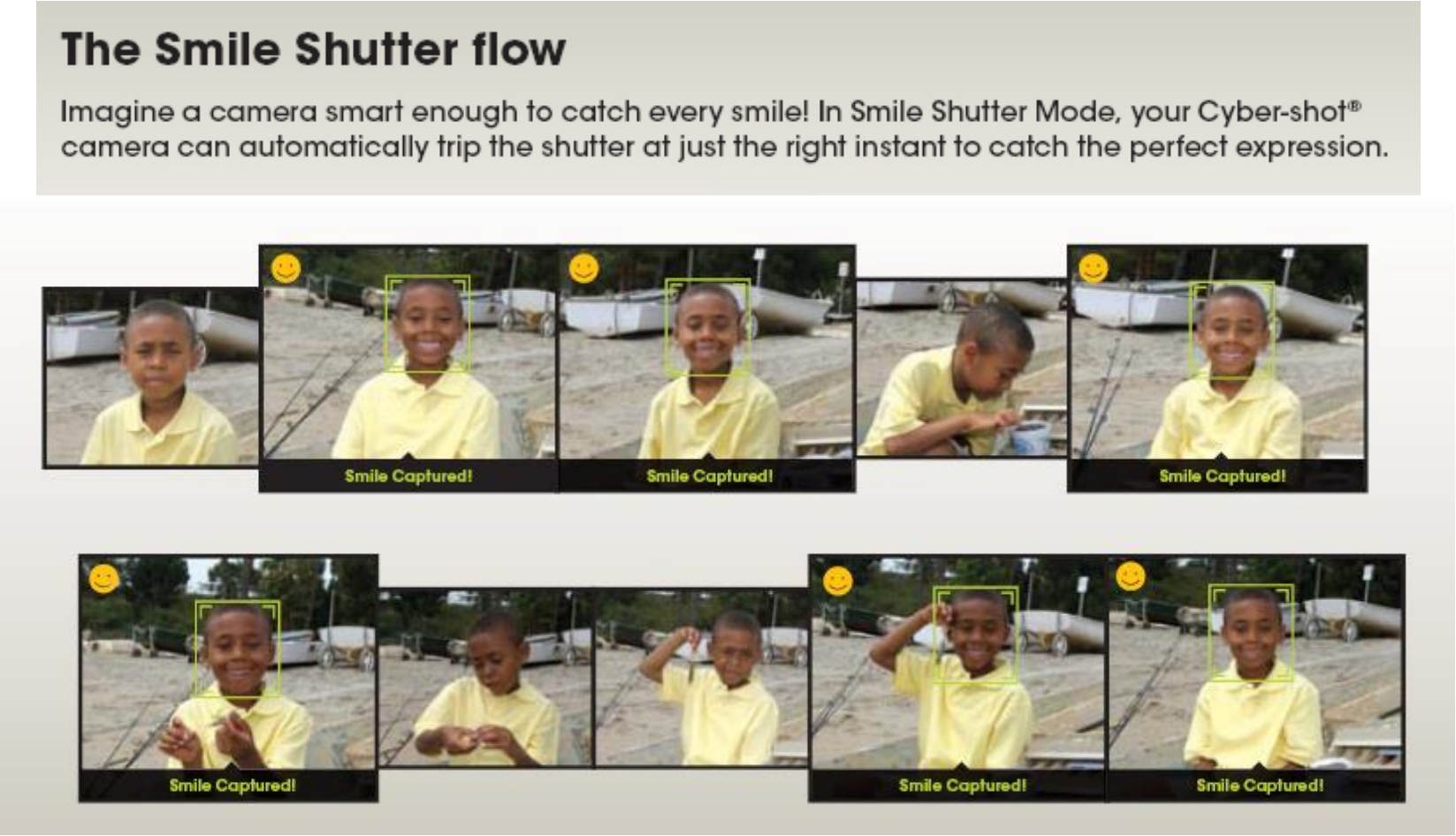
Optical Character Recognition (OCR)



Digit recognition, AT&T labs
http://www.research.att.com/~yann/

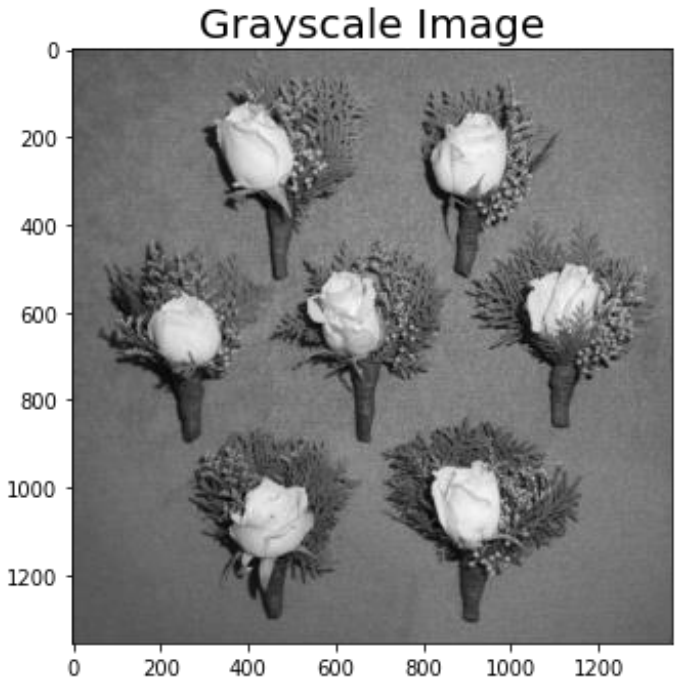
License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

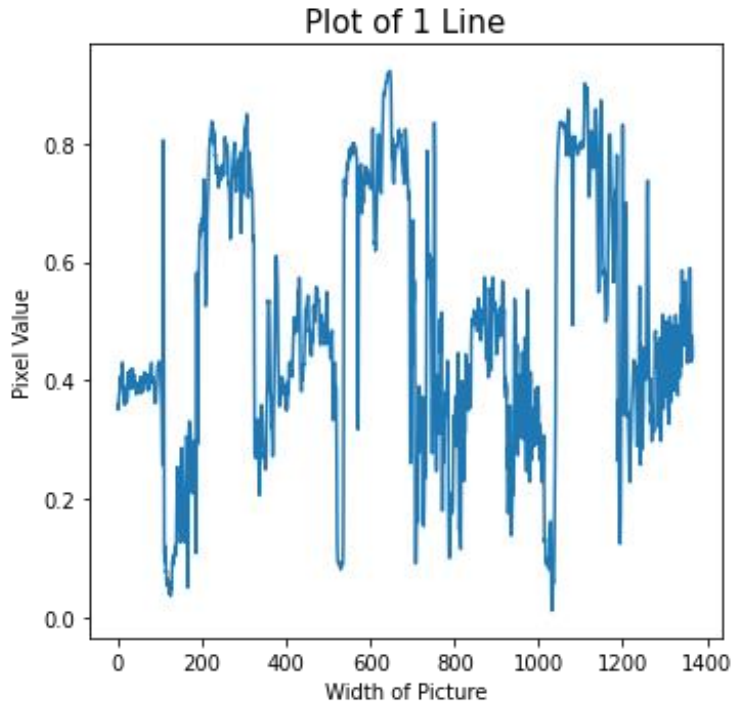
Face Detection with Expressions

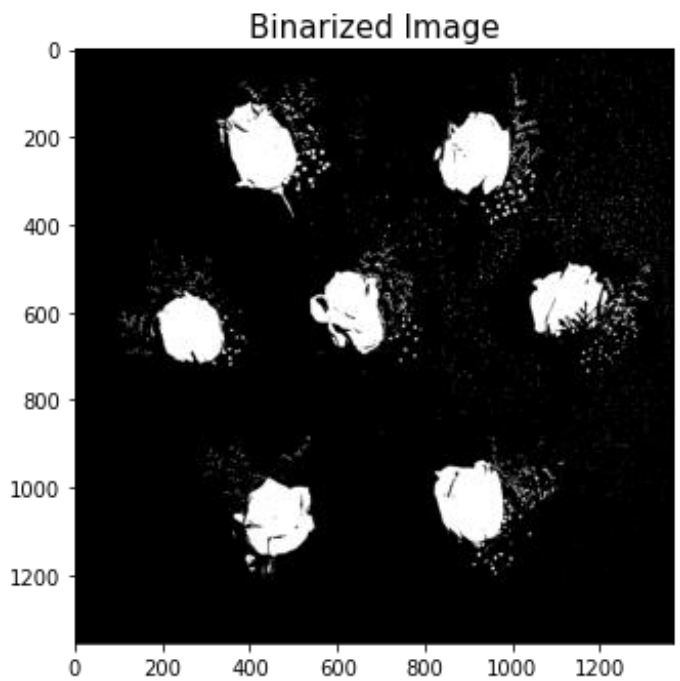


Hand Detection

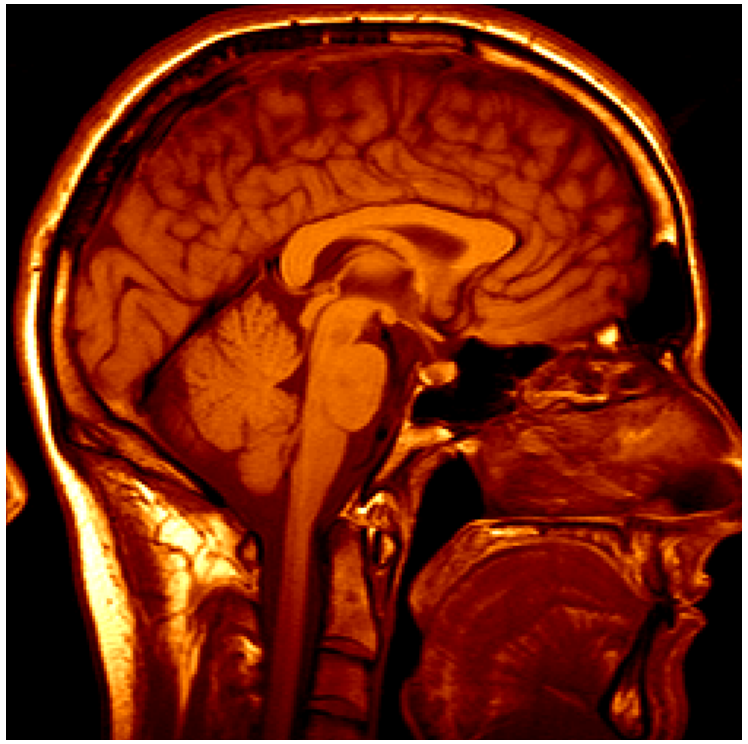
Blob Detection







Medical imaging



3D imaging
MRI, CT

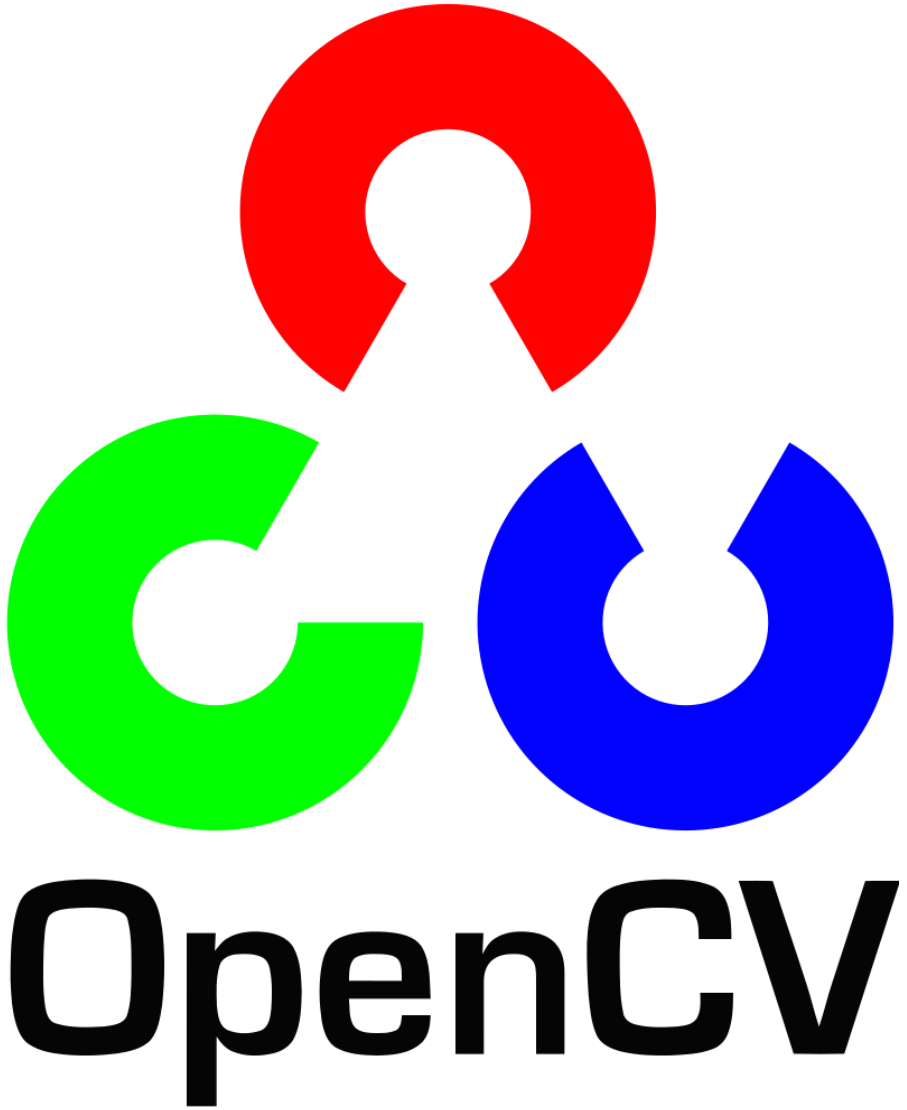
Image guided surgery
Grimson et al., MIT

What normal people see when they walk on street

What Computer Vision folks see

OpenCV

- Open source Computer Vision library: http://opencv.org/
- Originally developed by Intel and released in June 2000
- Has more than 2500 optimized algorithms
- C/C++/Python API
- it is written natively in C++
- Cross-platform also available for Android and iOS
- Released under a BSD license (it's free)



MediaPipe

- An open-source framework developed by Google for building multimodal, cross-platform machine learning pipelines.
- Cross-platform support (works on Android, iOS, web, and desktop).
- Pre-built machine learning models optimized for speed and accuracy.
- Modular and customizable architecture for creating advanced applications.

Workshop Plan

- Open CV basics
- Hand Detection
- Face Detection
- Object Detection
- Contour Segmentation
- QR Code Scanning
- Blob Detection

If time allows:

- Edge Detection
- Image classification
- Image compression
- Image processing
- Machine Learning?
- Image Generation?

Time to get your hand dirty!!!