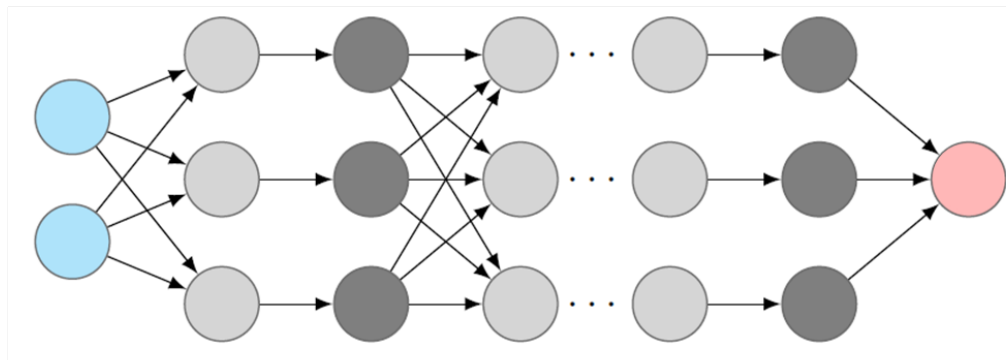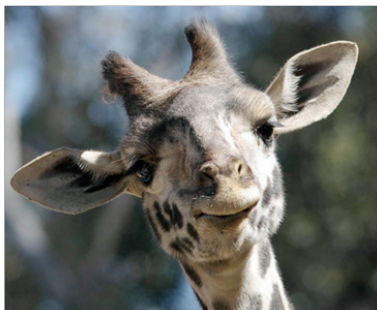# Working with visual data

Lecturer: Michael Möller – michael.moeller@uni-siegen.de
Exercises: Hartmut Bauermeister – hartmut.bauermeister@uni-siegen.de

What is a grayscale image?

Continuous point of view: A function

$$\tilde{f} : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$$

where $\tilde{f}(x)$ describes how bright the point at x is.
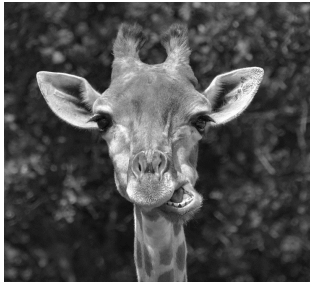
Discrete point of view: A matrix

$$f \in \mathbb{R}^{n_y \times n_x}$$

$n_y n_x$ is the total number of *pixels*.

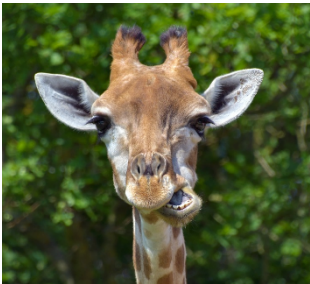The value in $f_{i,j}$ indicates how bright the (i,j)-th pixel is.

We will always work with float values in [0,1] = [black, white]. Other common formats work with uint8 on a scale [0,255].

Other visual data?

$$f \in \mathbb{R}^{n_y \times n_x}$$

Grayscale image

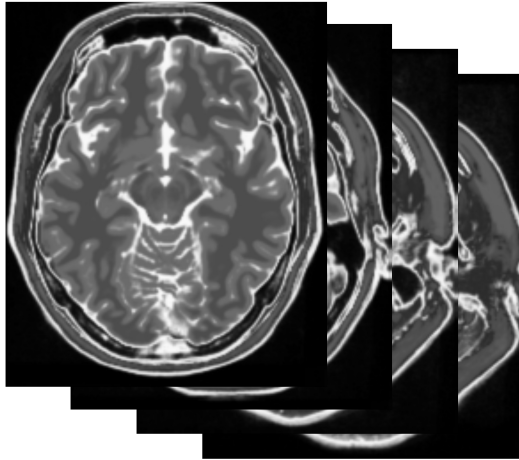$$f \in \mathbb{R}^{n_y \times n_x \times 3}$$

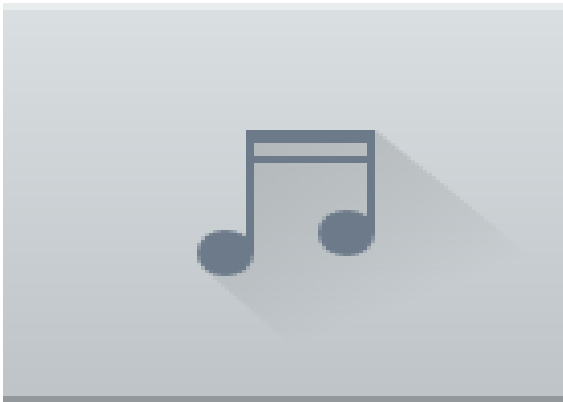Color image

RGB image *channels*

$$f \in \mathbb{R}^{n_y \times n_x \times 4}$$

RGB-D image

$$f \in \mathbb{R}^{n_y \times n_x \times n_z}$$

Image volume

$$f \in \mathbb{R}^{n_y \times n_x \times n_c \times n_f}$$

Video with n_c many channels and n_f many *frames*

But this is how the dog "looks" like in the computer! (red channel only)

178	189	190	187	182	174	120	139	172	175	141	124	149	168	183	176	184	182	184	192	192	195	185
170	174	181	162	144	178	174	183	178	180	183	181	180	186	183	181	185	189	190	190	170	172	208
186	187	185	185	188	187	186	185	181	182	170	176	188	189	189	185	187	194	192	182	182	180	193
182	188	185	178	164	150	159	170	180	184	173	153	128	114	119	153	182	137	145	195	188	191	191
186	137	75	45	33	20	25	32	50	63	41	22	2	0	0	18	46	29	38	154	186	191	175
119	26	2	2	2	0	0	0	0	0	0	0	0	3	9	2	0	0	3	17	52	91	120
19	9	7	5	5	3	6	8	8	4	3	2	2	3	6	7	4	6	4	0	0	0	1
3	0	4	6	7	4	6	9	8	1	4	5	6	4	5	6	6	8	11	10	10	6	0
118	60	1	4	6	7	6	6	6	5	8	6	9	10	10	7	3	8	27	29	23	11	5
204	98	2	7	8	13	13	12	12	14	11	4	4	6	8	9	6	8	13	30	38	18	16
113	104	26	9	11	11	13	13	13	15	12	7	6	5	7	7	9	21	44	68	50	23	13
38	102	56	14	16	13	26	31	20	22	11	7	9	8	8	10	42	101	179	165	113	29	4
52	23	40	29	14	4	78	160	82	46	25	6	7	7	9	40	121	166	126	92	127	85	26
37	41	51	25	16	9	77	204	180	134	107	60	14	16	36	114	182	161	131	120	168	160	134
76	174	144	31	12	9	87	187	194	198	198	186	59	36	39	122	197	189	204	207	200	203	212
156	201	202	98	4	7	124	197	191	191	194	206	79	92	65	67	185	194	202	212	200	145	125
209	201	206	136	6	4	143	196	182	193	195	186	69	169	156	39	137	219	202	201	190	137	137
140	113	93	93	22	44	174	179	188	199	206	186	79	172	206	97	70	159	122	106	169	190	194
107	134	155	148	31	116	203	199	194	192	210	173	77	135	160	175	60	80	141	130	167	190	194
132	157	195	134	20	156	197	172	151	176	181	94	96	168	163	177	114	49	172	161	132	139	137
137	172	212	105	25	164	188	171	153	88	74	53	117	164	143	130	147	101	93	134	155	143	111
183	186	177	106	31	157	196	188	95	63	139	144	152	166	158	175	175	188	123	69	67	118	132

Generate realistic images, e.g. in computer graphics, for games, movies, etc.

Reconstruct images from data that contains implicit information about them, e.g. in medical imaging
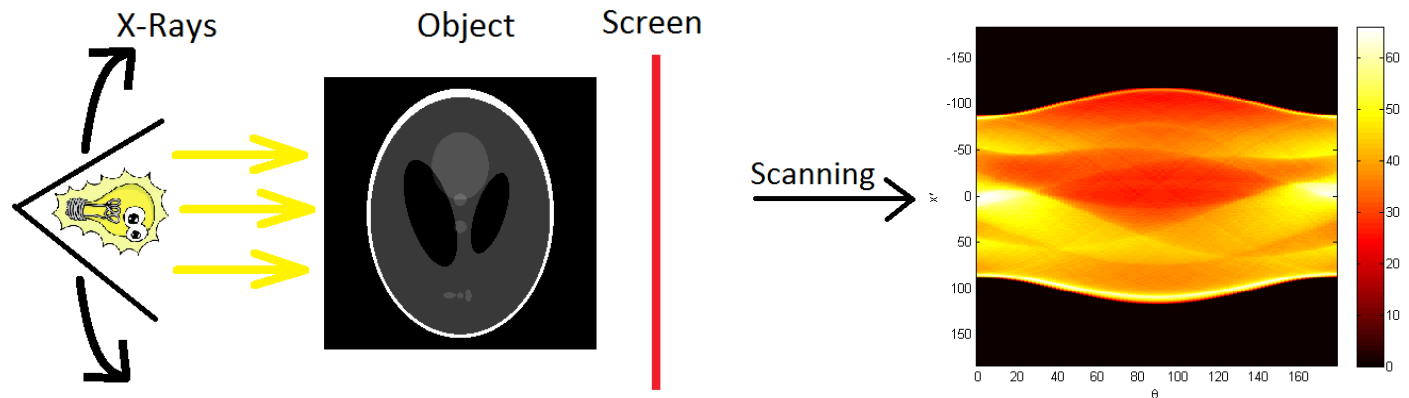
Improve image quality, e.g. denoising, super resolution, or inpainting.

Analyze visual data, e.g. in computer vision, including image classification, semantic segmentation, or activity recognition.

Michael Möller – michael.moeller@uni-siegen.de