EE 417 Introduction to Computer Vision

Images, Windows and Histograms

Instructor: Associate Prof. Mehmet Keskinöz

Communication Theory &Technologies (CTT) Group,
Electronics Engineering Program,
Computer Science and Engineering,
Cyber-Security Program
Faculty of Engineering and Natural Sciences

Sabancı Üniversitesi

Email: keskinoz@sabanciuniv.edu

An Array of 50 $\times$ 50 Numbers

46	58	70	76	76	69	60	67	71	54	80	103	100	87	77	55	30	22	23	22	23	26	22	22	20	20	22
44	61	71	73	72	65	52	49	54	46	39	64	65	57	40	21	17	18	21	16	16	21	20	19	18	19	20
48	61	70	71	64	59	49	41	43	48	24	34	30	26	17	13	17	20	16	9	9	15	16	16	16	17	18
43	58	62	61	63	60	44	39	41	43	26	16	15	14	13	16	17	16	11	11	14	17	16	11	8	14	17
38	48	58	56	52	58	47	41	43	39	27	14	19	18	16	14	12	14	19	30	39	42	41	34	27	19	11
36	37	47	46	48	60	44	43	46	45	25	19	20	18	16	13	14	26	43	59	71	78	80	76	73	50	25
37	38	35	32	66	87	47	47	58	51	21	19	20	18	14	13	20	36	52	61	68	71	70	73	82	78	66
39	36	35	27	67	108	67	52	67	52	22	17	20	18	15	15	23	34	40	40	43	44	45	56	58	72	91
35	37	39	32	50	87	87	55	56	53	29	19	22	14	13	19	26	40	46	49	55	57	59	64	61	63	73
37	34	33	38	46	61	69	44	52	45	27	24	24	13	9	24	44	50	56	63	67	72	77	75	76	76	69
37	38	43	37	48	53	42	38	45	35	24	27	27	16	9	31	59	63	67	74	79	79	84	88	87	79	67
35	46	47	38	58	53	38	43	41	30	22	25	24	19	10	33	67	76	74	74	80	82	89	93	75	49	39
39	49	52	45	51	48	44	39	38	26	19	20	17	17	10	30	69	83	82	77	85	92	89	64	45	45	49
46	53	60	60	50	42	43	36	33	21	15	16	13	14	11	27	66	74	63	59	81	99	94	60	57	50	43
62	75	70	81	72	45	43	41	30	18	13	15	13	13	13	21	37	45	38	42	72	96	102	80	51	26	21
61	81	81	75	72	63	42	39	28	16	12	14	14	13	15	15	38	50	55	58	67	83	96	72	24	23	45
43	58	65	55	42	41	40	33	23	13	13	13	14	12	11	28	57	48	28	32	35	54	76	53	41	52	62
47	50	51	47	40	32	29	30	23	12	11	12	13	12	12	24	32	12	13	23	28	35	53	58	59	69	72
41	50	51	46	44	40	33	30	22	12	9	11	11	10	10	16	15	28	48	57	55	46	62	67	61	68	80

This is actually a digital image when ...

Interpreting Numbers by Intensities

0 is "Black", 255 is "White", numbers in between are levels of gray

Each number is shown as a shaded small square, called a *pixel*

41	22	11	20	19	28	62	93	78	29	53	66	59	64	57	57	52	48
48	27	14	20	21	28	49	60	105	130	109	106	82	89	81	58	49	49
50	37	19	22	24	30	24	86	244	254	171	118	97	89	69	63	55	54
58	49	28	21	21	9	56	255	255	232	155	124	102	80	78	91	90	88
74	60	37	21	19	13	54	203	193	147	139	141	121	116	111	123	111	98
82	74	36	15	22	31	28	71	99	100	141	156	142	132	121	136	114	83
91	79	38	6	22	37	40	49	74	108	121	112	103	102	104	95	73	58
93	81	31	6	35	29	41	68	84	92	98	93	86	86	87	90	91	112
91	79	25	13	26	35	64	85	97	91	94	101	99	89	92	119	127	94
92	71	16	10	21	45	87	87	95	95	94	100	97	101	94	87	75	31
83	47	40	70	49	46	88	93	98	95	102	110	97	95	93	81	71	59
81	34	30	69	47	39	85	95	93	98	106	113	105	98	103	100	88	69
68	46	15	26	20	30	78	88	94	103	102	101	102	95	100	99	73	56
71	54	19	28	26	21	54	79	95	99	97	96	96	91	90	71	48	42
79	57	21	17	20	21	44	76	90	98	101	102	98	88	76	45	33	38
82	56	21	16	20	27	50	86	89	99	102	103	98	85	65	35	38	64
78	52	23	24	27	40	72	95	98	101	99	101	93	82	62	33	39	51
74	47	25	24	28	47	87	93	87	111	100	100	93	77	53	30	36	49
79	50	34	32	36	58	86	77	84	104	111	97	91	76	60	61	60	59
80	50	36	39	52	78	98	87	85	103	108	100	92	73	68	74	65	48
75	51	37	34	31	34	59	79	91	107	105	101	95	71	39	45	42	33
77	60	36	27	26	24	34	42	61	81	83	83	90	87	56	32	39	43
77	69	44	25	28	42	53	46	47	68	66	57	72	77	70	43	3	0
78	73	54	35	29	39	47	46	46	65	71	49	63	60	45	67	131	138
78	76	61	37	30	40	39	34	34	60	78	61	61	67	50	65	214	255
71	71	63	44	25	37	53	39	43	67	80	69	61	73	72	56	93	209
63	58	53	45	28	40	66	59	50	66	79	76	69	77	79	77	61	138

These are regions in the shown image with intensities ≥ 180; the pixels showing the teeth are still below threshold 180

Arrays of Numbers

... allow us to do numerical calculations, for example with the intention to improve the contrast in the digital image

Image Carrier

An image I is a rectangular array of *pixels* (x, y, u)

A pixel combines location $p = (x, y) \in \mathbb{Z}^2$ and sample u at p

$\mathbb{Z}$ is the set of all integers; points $(x, y) \in \mathbb{Z}^2$ form a *regular grid*

An image I is defined on a *carrier*

$$\Omega = \{(x, y) : 1 \leq x \leq N_{cols} \wedge 1 \leq y \leq N_{rows}\} \subset \mathbb{Z}^2$$

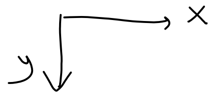
of N_{cols} times N_{rows} pixel locations (grid points)

$Z^2 = Z \times Z$

cartesian product

$\{a, b\} \times \{a, b\}$

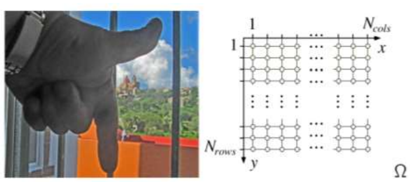
$= \{ (a, a), (a, b), (b, a), (b$



$|\Omega| =$ # elements in Ω

cardinality of set Ω

Image Coordinate System



Assuming a left-hand coordinate system, the thumb defines the x-axis and the pointer the y-axis while looking into the palm of the hand

---

Image Rows and Columns

We assume a left-hand coordinate system in Ω

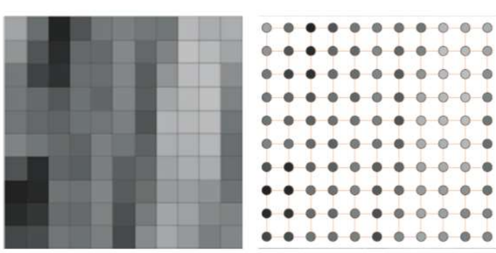
Row y

contains grid points $\{(1, y), (2, y), \ldots, (N_{cols}, y)\}$, for $1 \leq y \leq N_{rows}$

Column x

contains grid points $\{(x, 1), (x, 2), \ldots, (x, N_{rows})\}$, for $1 \leq x \leq N_{cols}$

Pixels



Left: Image values as shades in *grid squares* (grid cells)
Right: Image values as labels at *grid points* (centers of grid squares)

Grid Cells, Grid Points, and Adjacency

Grid cell model

A pixel is a homogeneously shaded square cell

Grid point model

A pixel is a labelled grid point

Pixel adjacency

Not defined by the pixels themselves; needs to be defined by us

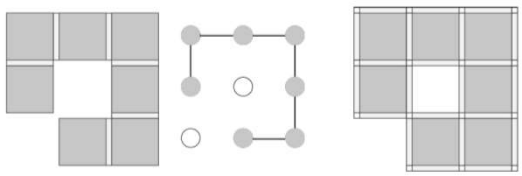
Two Examples

Two pixel locations p and q in grid cell model are adjacent *iff* $p \neq q$ and

(1) their tiny shaded squares share an edge ($\equiv$ 4-adjacency)
(2) their tiny shaded squares share an edge or corner ($\equiv$ 8-adjacency)

Two Examples

Adjacency defines *connectedness* and *connected sets* of pixels



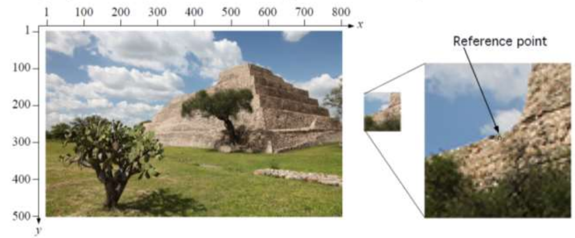
Left and middle: Connectedness of gray pixels due to 4-adjacency, either in grid cell or in grid point model; the two white pixels are not 4-adjacent

Right: Connectedness of gray pixels due to 8-adjacency in grid cell model forms a loop; the 8-adjacency of the two white pixels is "crossing" this loop

(Edges or corners of grid cells are shown as elongated rectangles or squares)

An Image Window

Image processing is often local (just addressing data in a window)



73×77 window in image `SanMiguel`; marked pixel location $p = (453, 134)$

The *reference point* of a window is usually at the window's centre

Image Windows

Unique identifier for an image window:

A *window* $W_p^{m,n}(I)$

is a subimage of image I of size $m \times n$

positioned with respect to a *reference point* p

Default: odd number $m = n$ and p at the center of the window

Example (on page before): $W_{(453,134)}^{73,77}(\texttt{SanMiguel})$

Usually: Simplify notation to W_p

(because image and size of window are known by given context)

Integer Values and $G_{\max}$

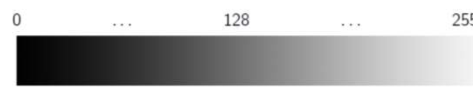
Scalar image: Values are integers $u \in \{0, 1, \ldots, 2^a - 1\}$

Common: $a = 8$ (i.e. one byte) or $a = 16$

Let $G_{\max} = 2^a - 1$ be the general maximum image value

Example

A *gray-level image* is a scalar image where scalar values represent gray levels, with $0 =$ black and $2^a - 1 =$ white; all other gray-levels are linearly interpolated between black and white



Binary and Vector-Valued Images

A *binary image* (e.g. produced by *thresholding*) has two values, traditionally $0 =$ white and $1 =$ black (black *objects* on white *background*)

A *vector-valued image* has more than one *channel* or *band*

Image values $[u_1, \ldots, u_{N_{channels}}]^\top$ are vectors

Example: Color images in the *RGB color model* have channels for the red, green, and blue component; values u_i in each channel are in the set $\{0, 1, \ldots, G_{\max}\}$ (like a gray-value image)

Example of an RGB Color Image

RGB colour image `Fountain` and its three channels for Red (*upper right*), Green (*lower left*), and Blue (*lower right*)

Basic Statistics

Mean

Given: $N_{cols} \times N_{rows}$ scalar image I

Mean (i.e., the "average gray level") of image I

$$\mu_I = \frac{1}{N_{cols} \cdot N_{rows}} \sum_{x=1}^{N_{cols}} \sum_{y=1}^{N_{rows}} I(x,y)$$
$$= \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y)$$
$$= \frac{1}{|\Omega|} \sum_{p\in\Omega} I(p)$$

$|\Omega| = N_{cols} \cdot N_{rows}$ is the cardinality of the carrier Ω

Variance and Standard Deviation

Variance of image I

$$\sigma_I^2 = \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} [I(x,y) - \mu_I]^2$$

Root σ_I is the *standard deviation* of image I

Well-known formula in statistics:

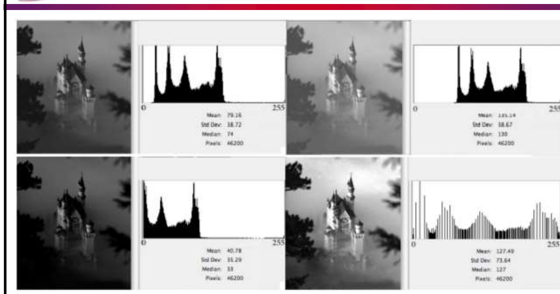
$$\sigma_I^2 = \left[\frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} I(x,y)^2 \right] - \mu_I^2$$

Thus: Mean and variance calculated by running through image I only once

Why?

Standard deviation $= \sqrt{\text{Variance}} \geq 0$

Four Histograms



Histograms for a 200×231 image `Neuschwanstein`

Upper left: Original image. *Upper right*: Brighter version. *Lower left*: Darker version. *Lower right*: After histogram equalization (defined later)

Definition of Histograms

A *histogram* represents tabulated frequencies, typically by using bars in a graphical diagram

Given: Scalar image I with pixels (x,y,u), where $0 \leq u \leq G_{max}$

Absolute frequencies (count of appearances of u in Ω)

$$H_I(u) = |\{(x,y) \in \Omega : I(x,y) = u\}|$$

$|\ldots|$ denotes the cardinality of a set

$H_I(0), H_I(1), \ldots, H_I(G_{max})$ define the (absolute) *gray-level histogram* of I

Relative frequencies (between 0 and 1) define a *relative histogram*

$$h_I(u) = \frac{H_I(u)}{|\Omega|}$$

More on Histograms

We have:

$$\mu_I = \sum_{u=0}^{G_{max}} u \cdot h_I(u) \quad \text{or} \quad \sigma_I^2 = \sum_{u=0}^{G_{max}} [u - \mu_I]^2 \cdot h_I(u)$$

Absolute and relative *cumulative frequencies* define *cumulative histograms*

$$C_I(u) = \sum_{v=0}^{u} H_I(v) \quad \text{and} \quad c_I(u) = \sum_{v=0}^{u} h_I(v)$$

Observation

Relative frequencies are comparable to the *probability density function* $\Pr[I(p) = u]$ of discrete random numbers $I(p)$, relative cummulative frequencies are comparable to the *probability function* $\Pr[I(p) \le u]$

Histograms for Two Image Windows

Two 104×98 windows in image `Yan` and corresponding histograms

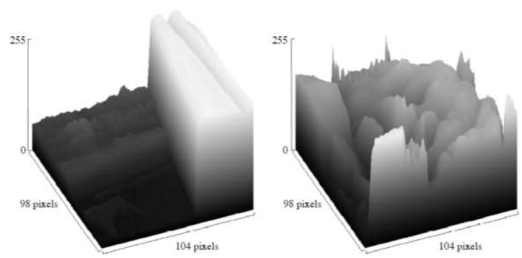
Upper window: $\mu_{W_1} = 133.7$ and $\sigma_{W_1} = 55.4$

Lower window: $\mu_{W_2} = 104.6$ and $\sigma_{W_2} = 89.9$

Draw a sketch of the cumulative histograms of both windows!

3D Views of Gray-Level Images

3-dimensional (3D) views illustrate different "degrees of homogeneity" in an image window; here for the two windows from Page 32



Left: Steep slope from lower plateau to higher plateau illustrates an "edge"

Value Statistics in a Window

Given: Window $W = W_p^{n,n}(I)$, with $n = 2k + 1$ and $p = (x, y)$

We have in window coordinates

$$\mu_W = \frac{1}{n^2} \sum_{i=-k}^{+k} \sum_{j=-k}^{+k} I(x + i, y + j)$$

Formulas for variance, histograms, and so forth, can be adapted analogously