Discrete 2D Signals

Image Representation & Matching

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

Image Representation

- Image space
- Decomposition and reconstruction
- Frequency space
- Fourier transformation

Digitalisation

- Sampling theorem
- Aliasing

Image Matching

- Convolution & Correlation
- Point spread function & Transfer function

Image Analysis

Second Part of the Lecture

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

1. Part: Basics of Image acquisition and Image geometry

- Physics, Geometry and Linear Algebra -

2. Part: Basics of Image Analysis

- Signal Theory, Pattern Recognition and Probability -

Linear algebra & probability are very important to master (deep) learning

Interview Andrew Ng & Ian Goodfellow, 2017.

Image Analysis

Second Part of the Lecture

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

Image Analysis is a special subfield of Pattern Recognition and thus a subfield of Artificial Intelligence. Pattern recognition has its origins in engineering and machine learning originated in computer science. Both research fields deal with the same problems with different emphases. In this course, we focus on the signal-theoretic aspects of generating appropriate representations of images.

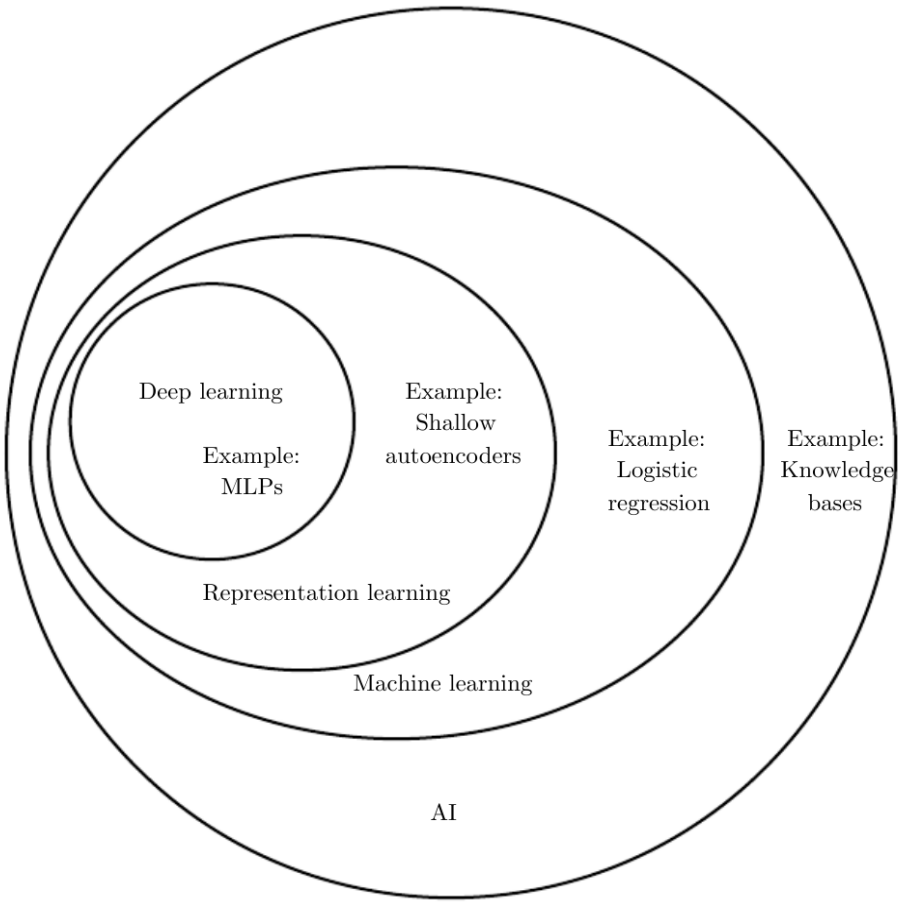


Image Analysis

Second Part of the Lecture

Technical University of Applied Sciences Würzburg-Schweinfurt

Representations of images or parts of images are stored in feature vectors, also called descriptors. These representations can be designed by hand, or learned via a learning architecture using data, in which case the design of the learning architecture and the selection of learning data must be performed by hand. When learning feature vectors, priority is given to unsupervised learning methods.

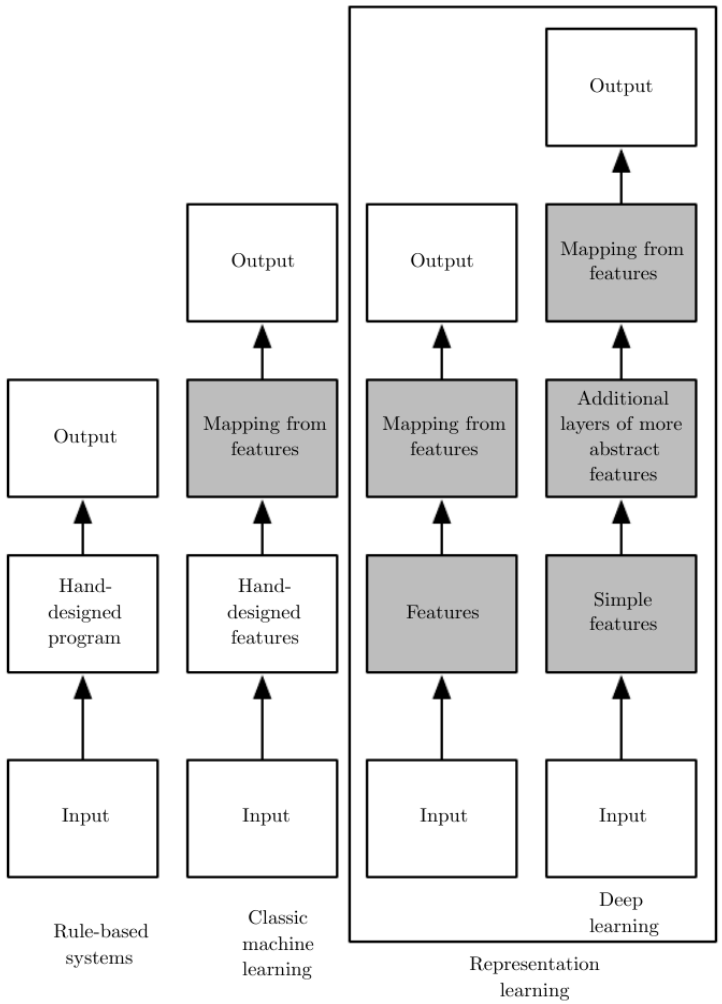


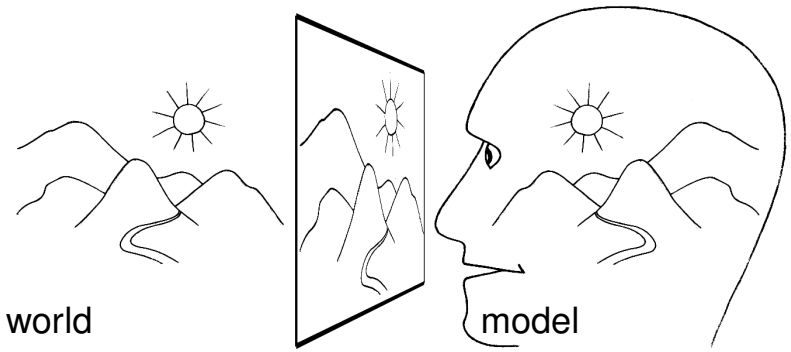
Image Analysis

How to solve image processing problems?

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

Mathematical Models for

- Geometry and Physics of the world ✓
- How to extract
  features from image data ?
- How to model
  spatiotemporal dependencies ?



How does our brain work?

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

Example: Deep Learning Architecture

Example of learned filter masks for feature extraction at different abstraction levels (local - global). These learned filter masks have many properties that hand-designed filters also have. To evaluate learned filter masks, a good understanding of filter design and implementation is also needed.

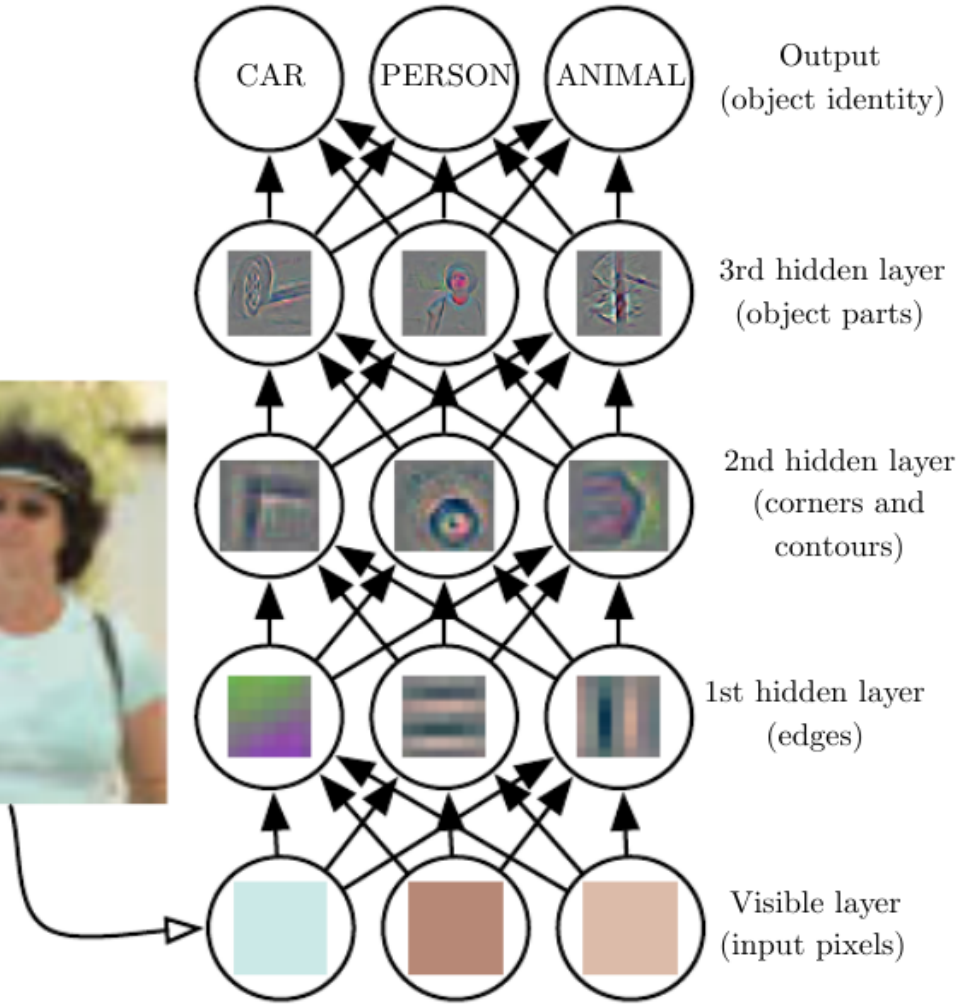


Image Analysis

Motivation

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

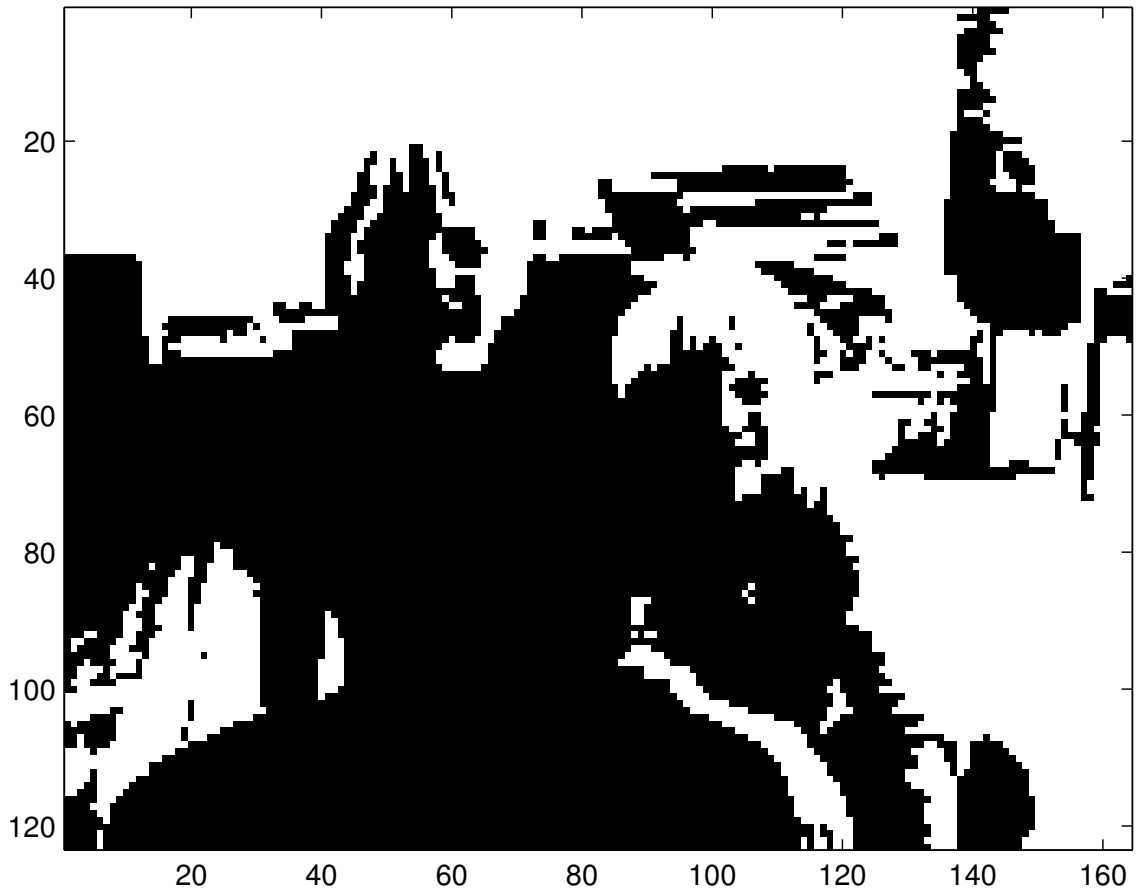
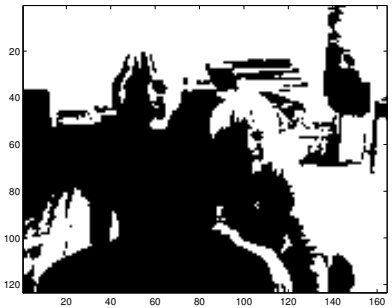


Image Analysis

Motivation



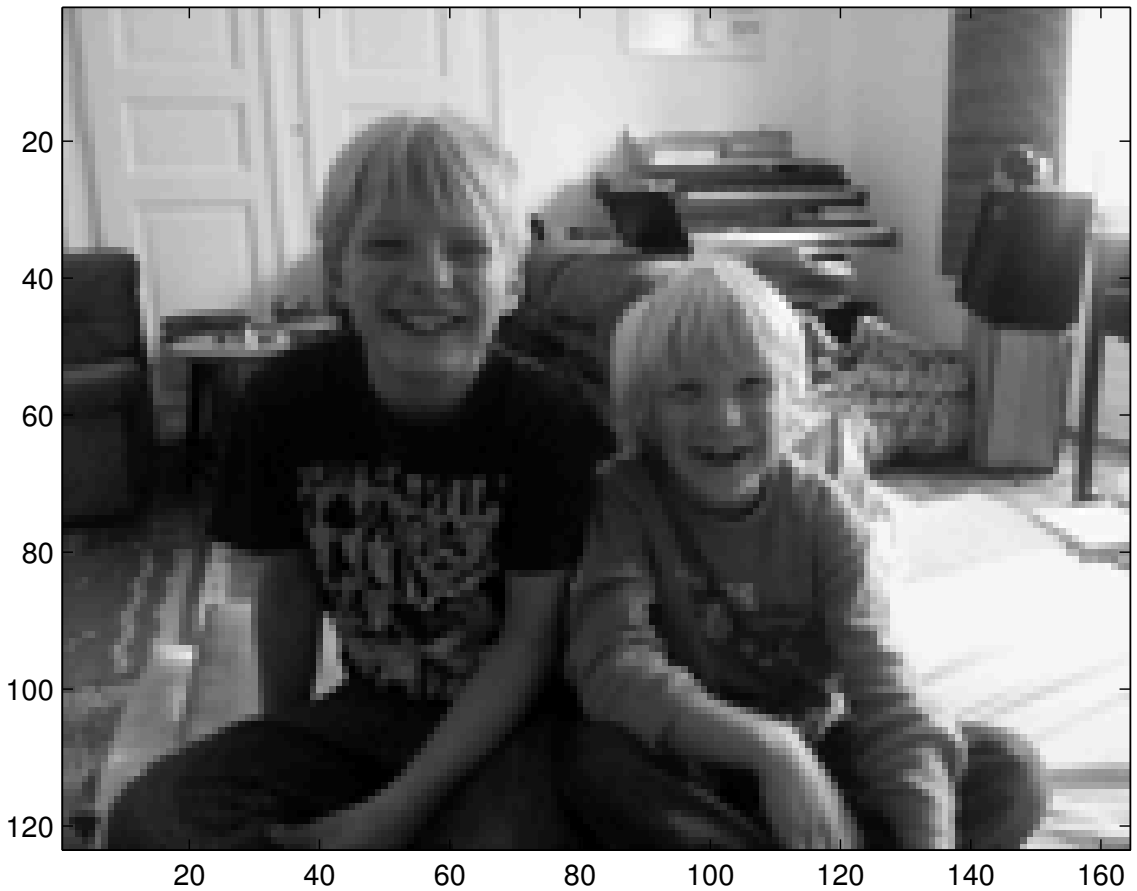
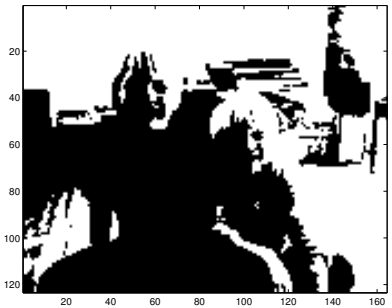


Image Analysis

Motivation



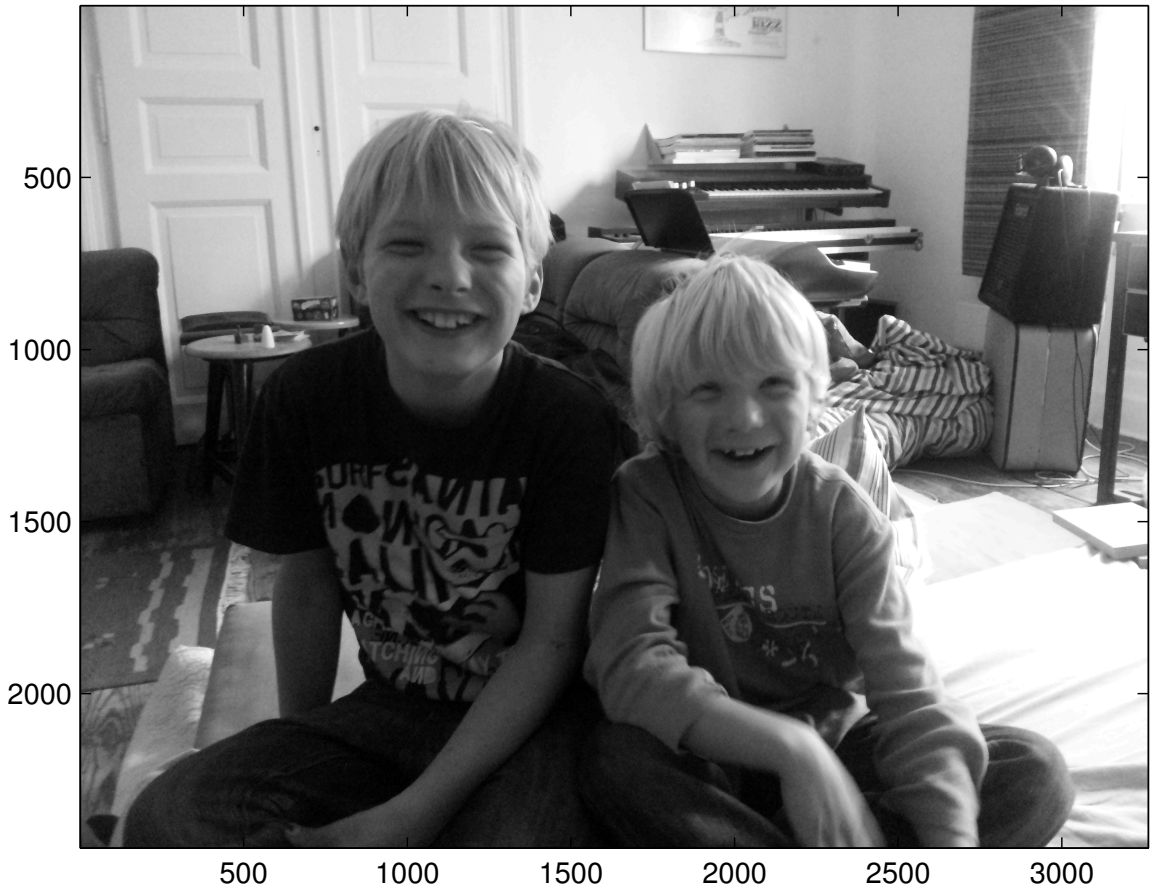
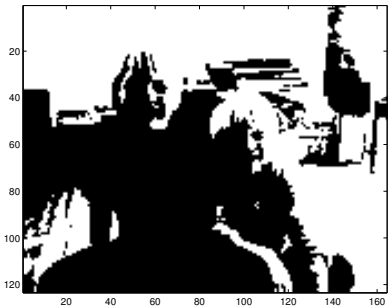


Image Analysis

Motivation



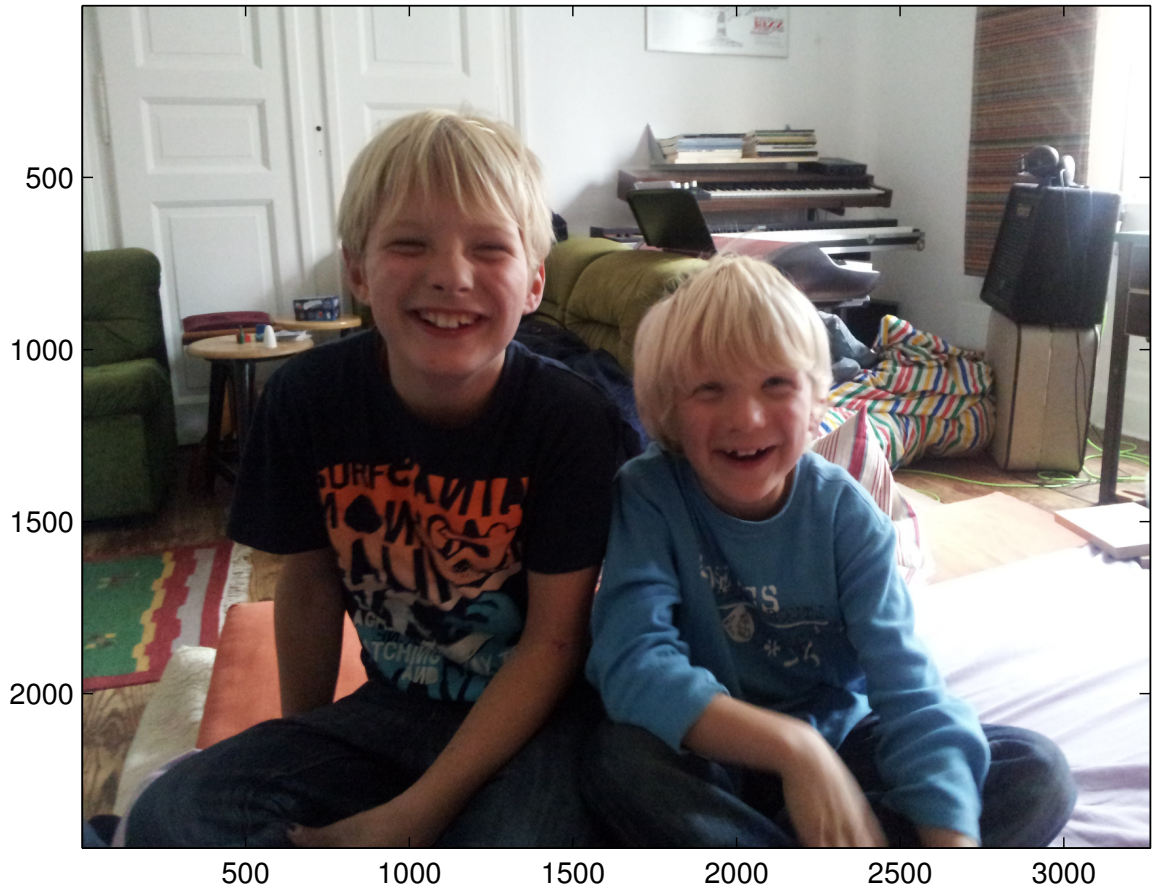
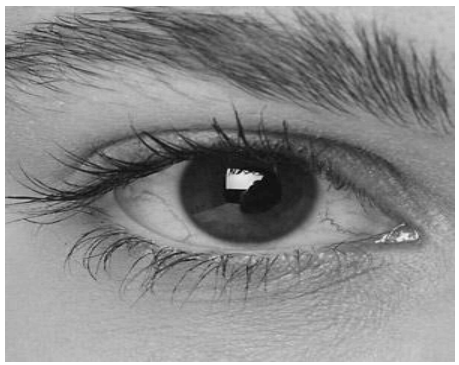


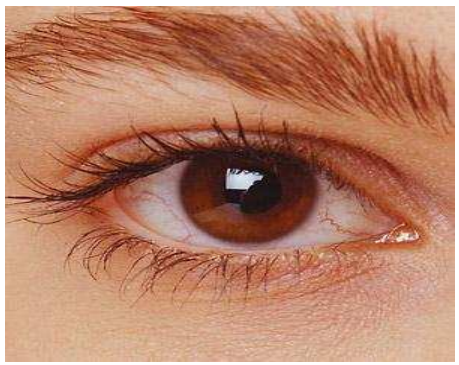
Image Representation

The Digital Image Space

Images represent an areal distribution of irradiance E of a certain section of the environment in a plane (x, y).



gray values



colors

E(x,y)

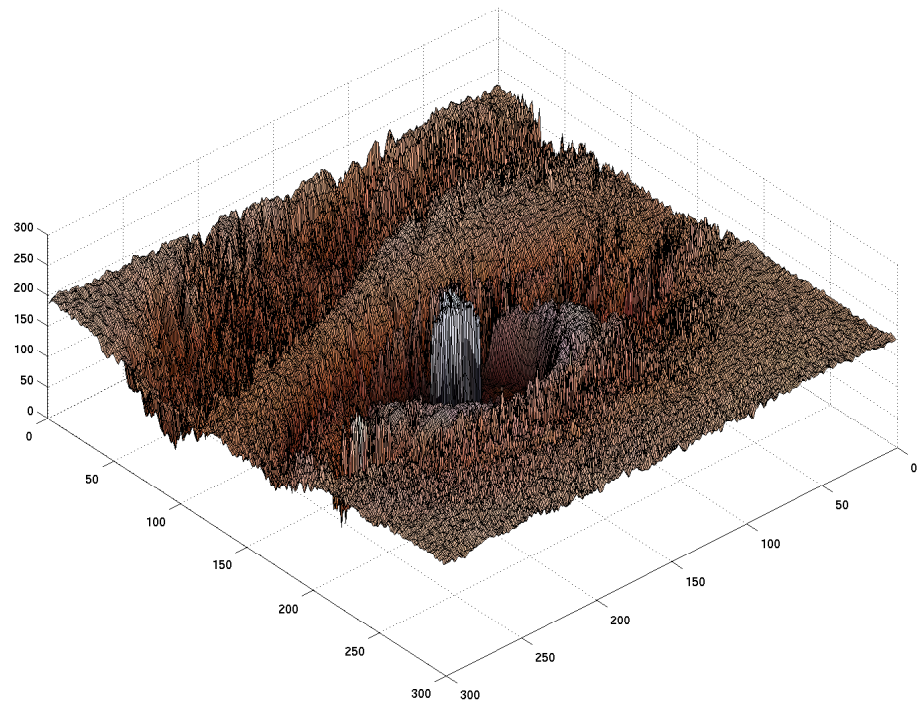


Image Representation

Digital Image Space - Pixel

Digitization is the rasterization of an image $E(\mathbf{x})$ with continuous coordinates $\mathbf{x} = [x, y]^\top$ onto a two-dimensional image matrix $\mathbf{G} = \{G(\mathbf{p})\}_\mathbf{p}$ with the grid points $\mathbf{p} = [m, n]^\top$.

$$E(\mathbf{x}) = E(x, y)\,,\ \ x,y \in \mathbb{R} \quad \rightarrow \quad G(\mathbf{p}) = G_{m,n}\,,\ \ m,n \in \mathbb{N}\,.$$

A matrix element on the grid is called a pixel (picture element) with the rectangular pixel area $\delta x \times \delta y$. The pixel position $\mathbf{p}$ is given by a pair of numbers

$$(m = 0, 1, \ldots, M-1 \text{ and } n = 0, 1, \ldots, N-1)$$

where m is the row index and n is the column index. M is the row number and N is the column number of the image matrix.

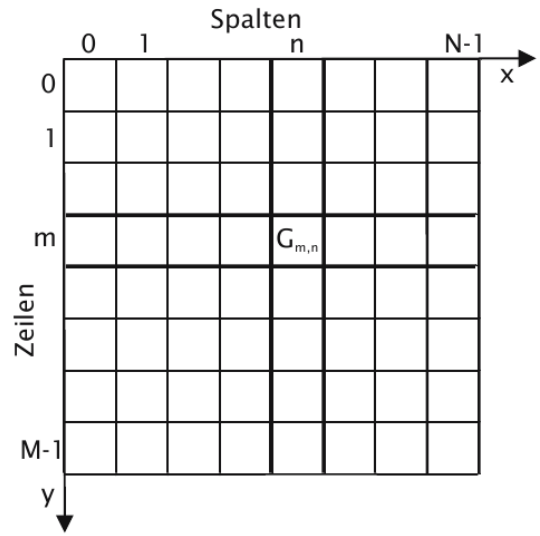
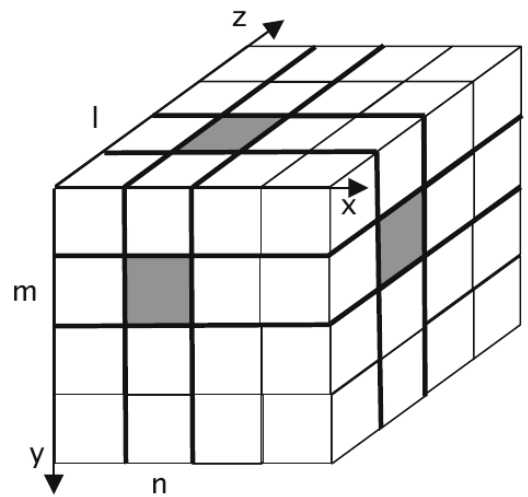


Image Representation

Digital Image Space - Pixel & Voxel

Each pixel represents a rectangular region, a unit cell, of the grid. The value corresponds to the average irradiance.



Three- or higher-dimensional signals also occur in image processing:

- Color images (RGB channel images).
- Series of sectional images (e.g. tomography).
- Video sequence of gray or color images.

3-D images consist of voxels (volume element) whose values $G(m, n, l)$ represent the mean gray value of a cube element at the location (x, y, z).

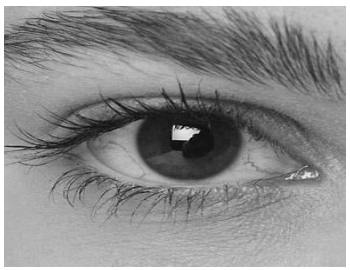
Image Representation

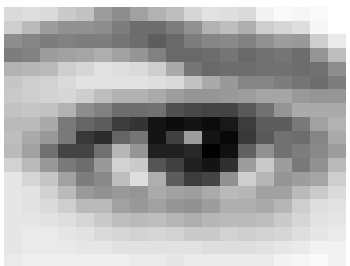
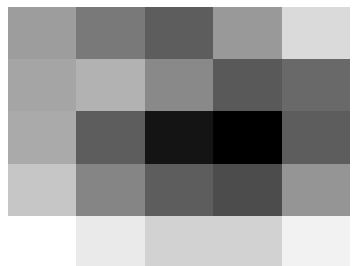
Digital Image Space - Resolution

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

The lower the sampling, the smaller the resolution of the digital image.

Which spatial (and temporal) resolution
is necessary/meaningful?
The question depends on the task:



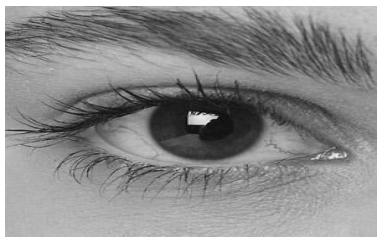



- image viewing: pixel size should be e.g. be smaller than the visual resolution of the human visual system.
- object recognition: pixel size should be e.g. smaller than the smallest objects to be examined.
- image motion: The time discretization must be so fine that, e.g., all pixels to be analyzed remain within one ROI (Region of Interest) despite pixel shifts between two successive images.

Image Representation

Digital Image Space - Quantization

The continuous range of values of the measured irradiance is mapped onto a limited number Q (quantization levels) of discrete gray values.



How many quantization levels are necessary?

This question again depends on the task:

- How high is the contrast in the image areas in which I want to recognize an object?
- From how many quantization levels does the human perceive gray value transitions as continuous?
- Are edge locations shifted?

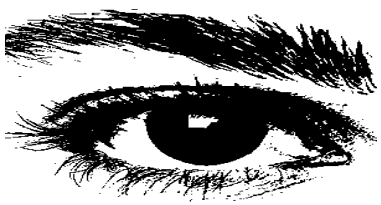


Image Representation

Digital Image Space - Memory

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

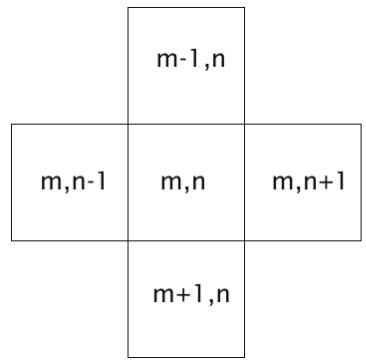
By default, grayscale images or a color channel are quantized with a color bit depth of 8 bit, which corresponds to 256 integer value steps. For a standard resolution of 640 $\times$ 480 pixels (e.g. graphics standard VGA=Video Graphics Array), a color depth of 8 bits (e.g. RGB color space) per color channel and a refresh rate of 25 Hertz (PAL television standard), the following memory requirements result:

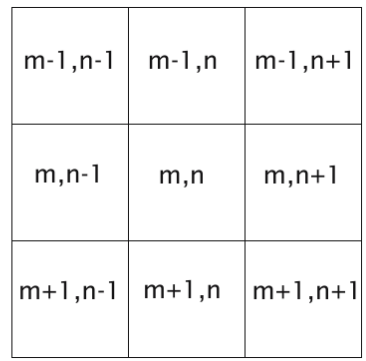
- 1 gray value image: $640 \times 480 Byte = 300 KB$.
- 1 color image: $640 \times 480 Byte \times 3$Kanäle $= 900 KB$.
- 1 channel video signal: $640 \times 480 Byte \times 25 Hz \approx 7,3 \frac{MB}{sec}$.
- Color video sequence: $640 \times 480 Byte \times 3 \times 25 Hz \approx 22 \frac{MB}{sec}$.

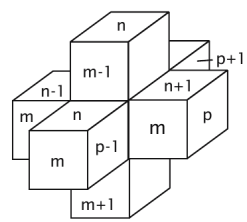
Image Representation

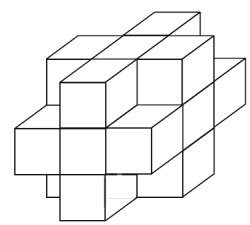
Digital Image Space - Neighborhoods

Objects almost always cover a contiguous region of individual pixels. For the description of an object or a image region, the evaluation of neighborhood relations plays a fundamental role. A direct neighbor pixel must have a common edge or at least one common corner. In principle, however, also larger neighborhoods can be defined for a pixel, whereby not all neighbor pixels are directly adjacent.









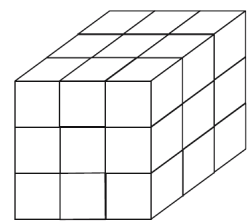


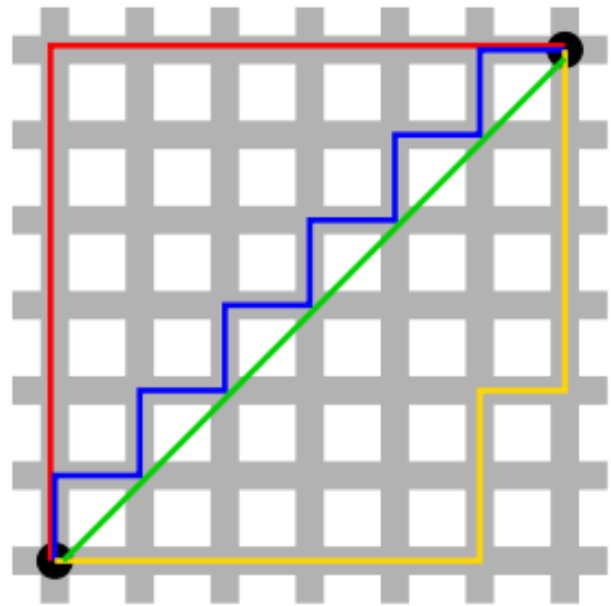
Image Representation

Digital Image Space - Distances

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

Each grid point **p** of a discrete image has a grid vector

$$\mathbf{r}_{m,n} = \begin{bmatrix} n\Delta x \\ m\Delta y \end{bmatrix} .$$



Thus, the Euclidean distance on a discrete grid is obtained as follows:

$$d(\mathbf{r}, \mathbf{r}') = \|\mathbf{r} - \mathbf{r}'\| = \sqrt{(n - n')^2 \Delta x^2 + (m - m')^2 \Delta y^2} .$$

Other distances can be defined on a grid, but they are not relevant for photogrammetry tasks, since they are direction-dependent.

Block distance: $d = |n - n'|\Delta x + |m - m'|\Delta y$, Checkerboard distance: $d = max(|n - n'|\Delta x, |m - m'|\Delta y)$.

Image representation

Images as Vectors

The reordering from an image matrix **G** to an image vector **g** can be done row-wise or column-wise.

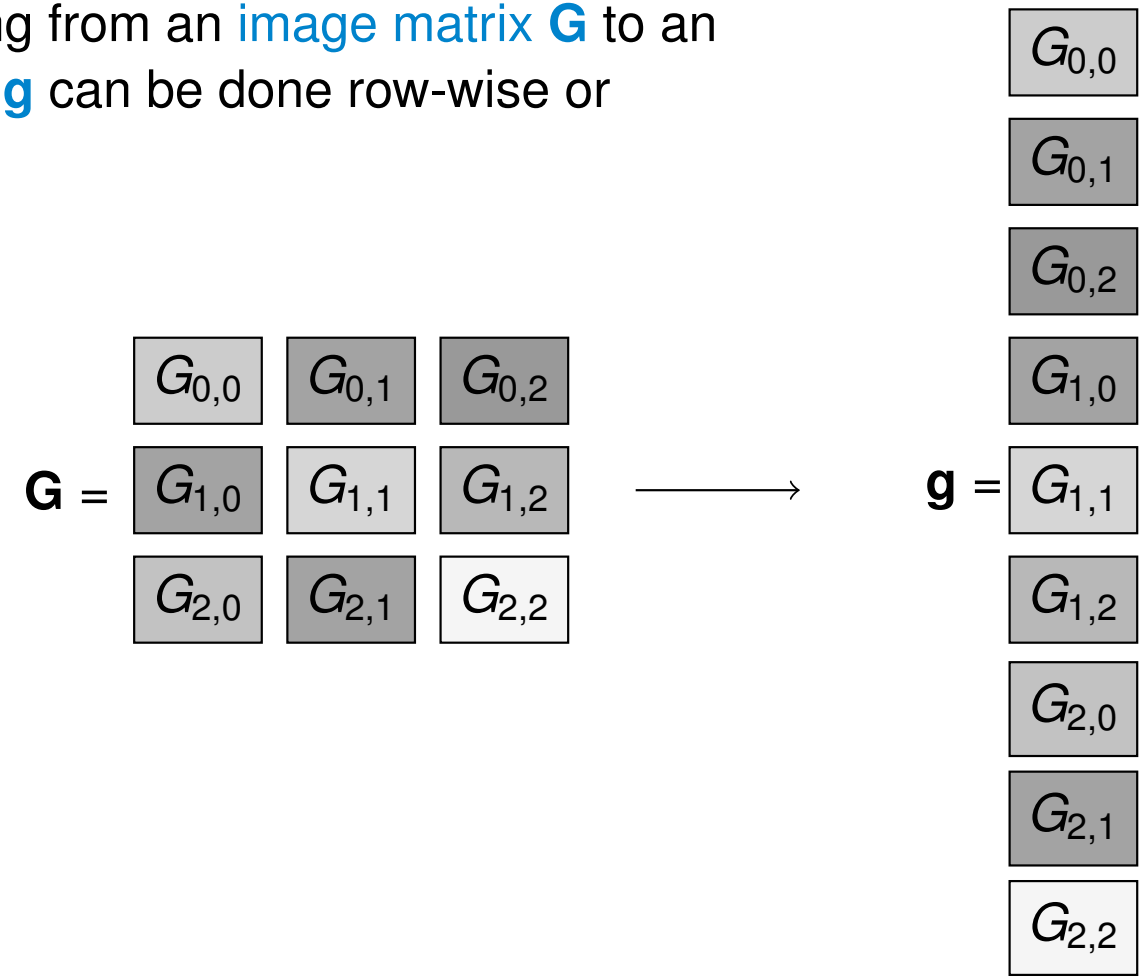


Image representation

Decomposition & Reconstruction - Basis Images

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

One can think of each $M \times N$ image $\mathbf{G}$ as a weighted sum of $M \times N$ different basis images $\mathbf{B}^{m,n}$ with the individual gray values $G_{m,n}$ as weights:

$$\mathbf{G} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} G_{m,n} \mathbf{B}^{m,n} \quad .$$

Each base image $\mathbf{B}^{m,n}$ has the value 1 only at one image point (m, n), at all other image points $\{(m', n')\}_{m' \neq m, n \neq n'}$ the value is zero:

$$\mathbf{B}^{m,n} : B^{m,n}_{m',n'} = \begin{cases} 1 & \text{für } m = m' \wedge n = n', \\ 0 & \text{sonst.} \end{cases}$$

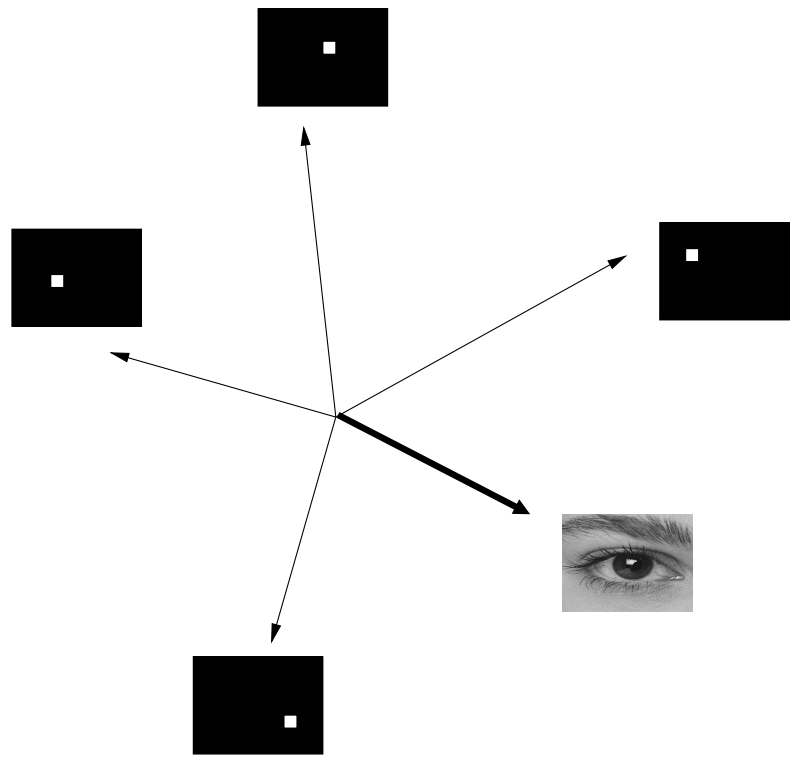


Image representation

Decomposition & Reconstruction - Basis Images

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

These basis images form an orthonormal basis as the inner product (scalar product) of two different basis images (vectors)

$$\left\langle \mathbf{B}^{m',n'} \middle| \mathbf{B}^{m'',n''} \right\rangle = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} B_{m,n}^{m',n'} B_{m,n}^{m'',n''} = \delta_{m'-m''} \delta_{n'-n''} ,$$

always results in zero and the inner product of a base image with itself is always one.

The basis vectors thus span a $M \times N$ - dimensional vector space. Thus, an image corresponds to a $M \times N$ - dimensional vector in this vector space.

If you now change the coordinates, no image information is lost, only the coordinates are changed. This means that the same information is viewed from a different „angle of view“.

Image representation

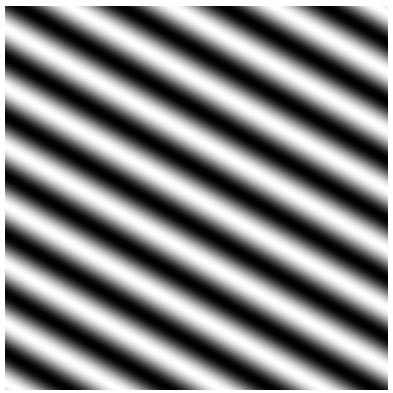
Frequency Space - Periodic Basis Images

A two-dimensional periodic pattern

$$G(\mathbf{x}) = r\cos(2\pi\mathbf{k}^\top\mathbf{x} - \varphi)$$

is given by:

- Wavenumber vector $\mathbf{k} = [k_1, k_2]^\top$
- Wavelength λ
- Magnitude of wavenumber vector $|\mathbf{k}| = 1/\lambda$
- Amplitude r
- Distance of the first maximum from the origin Δx
- Phase angle $\varphi = 2\pi\Delta x/\lambda = 2\pi\Delta x|\mathbf{k}|$



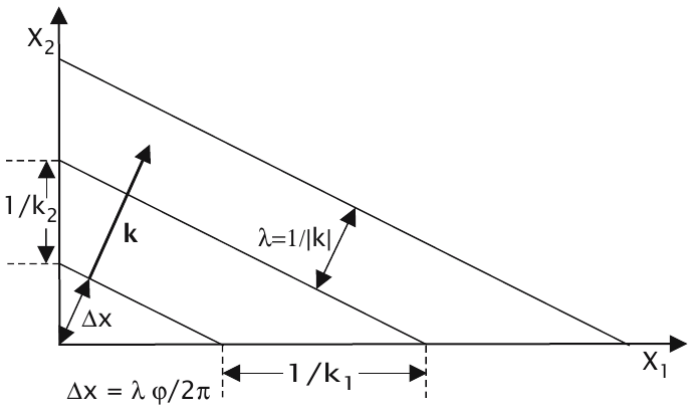


Image representation

Frequency Space - 2D Fourier Transform

The continuous 2D Fourier transform decomposes a continuous 2D image function into periodic structures of different wavelength and direction.

$$\text{Transformation:} \quad G(\mathbf{x}) \circ\!\!-\!\!\bullet \hat{G}(\mathbf{k}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\mathbf{x}) e^{-2\pi j \mathbf{k}^T \mathbf{x}} d\mathbf{x}$$

Expansion of $G(\mathbf{x})$ according to periodic basis functions: $e^{-2\pi j \mathbf{k}^T \mathbf{x}} = \cos(2\pi \mathbf{k}^T \mathbf{x}) - j \sin(2\pi \mathbf{k}^T \mathbf{x})$.

$$\text{Inverse Transformation:} \quad \hat{G}(\mathbf{k}) \bullet\!\!-\!\!\circ G(\mathbf{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{G}(\mathbf{k}) e^{2\pi j \mathbf{k}^T \mathbf{x}} d\mathbf{k}$$

Signal representation via superposition of $\hat{G}(\mathbf{k})$ weighted periodic basis functions:
$e^{2\pi j \mathbf{k}^T \mathbf{x}} = \cos(2\pi \mathbf{k}^T \mathbf{x}) + j \sin(2\pi \mathbf{k}^T \mathbf{x})$.

Image representation

Frequency Space - 2D Discrete Fourier Transform

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

The 2D DFT maps real-valued (and complex-valued) $M \times N$ matrices to complex-valued $M \times N$ matrices:

$$
\begin{aligned}
G_{m,n} \circ\!\!-\!\!\bullet \hat{G}_{u,v} &= \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} G_{m,n} e^{-2\pi jmu/M} e^{-2\pi jnv/N} \\
&= \frac{1}{MN} \sum_{m=0}^{M-1} \left(\sum_{n=0}^{N-1} G_{m,n} w_N^{-nv} \right) w_M^{-mu} \\
&= \frac{1}{MN} \langle \mathbf{W}^{u,v} | \mathbf{G} \rangle = \frac{1}{MN} \left(\mathbf{G} \mathbf{w}_N^{-v} \right)^\top \left(\mathbf{w}_M^{-u} \right),
\end{aligned}
$$

whereas $e^{2\pi j/M} = w_M$, $e^{2\pi j/N} = w_N$ and $\mathbf{W}^{u,v} = (\mathbf{w}_M^u)(\mathbf{w}_N^v)^\top$,

with $\mathbf{w}_M^u = \left[w_M^0, w_M^u, w_M^{2u}, \ldots, w_M^{(M-1)u} \right]^\top$, $\mathbf{w}_N^v = \left[w_N^0, w_N^v, w_N^{2v}, \ldots, w_N^{(N-1)v} \right]^\top$.

Image representation

Frequency Space - 2D Discrete Fourier Transform

The inverse 2D DFT results accordingly to:

$$\begin{aligned}
\hat{G}_{u,v} \bullet\!\!-\!\!\circ G_{m,n} &= \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} \hat{G}_{u,v} e^{2\pi jmu/M} e^{2\pi jnv/N} \\
&= \sum_{u=0}^{M-1}\left(\sum_{v=0}^{N-1} \hat{G}_{u,v} w_N^{nv}\right) w_M^{mu} \\
&= \langle \mathbf{W}^{-u,-v} | \hat{\mathbf{G}} \rangle = \left(\hat{\mathbf{G}}\mathbf{w}_N^v\right)^\top \left(\mathbf{w}_M^u\right).
\end{aligned}$$

The matrices $\mathbf{W}^{u,v}$ can be understood as basis images spanning an $N \times M$-dimensional vector space over the body of complex numbers. Therefore, the DFT can also be viewed as a coordinate transformation to a complex $N \times M$-dimensional vector space.

Image representation

Frequency Space - 2D Discrete Fourier Transform

The 2D DFT computes the projections of the image **G** onto all base images $\mathbf{W}^{u,v}$ of the Fourier space, i.e. the components $\hat{G}_{u,v}$ of **G** in the direction of the base images.

Real and imaginary part of the basic images are sampled periodic patterns with different wavenumber vectors **k** of integer vector components $k_1 = u$, $k_2 = v$.

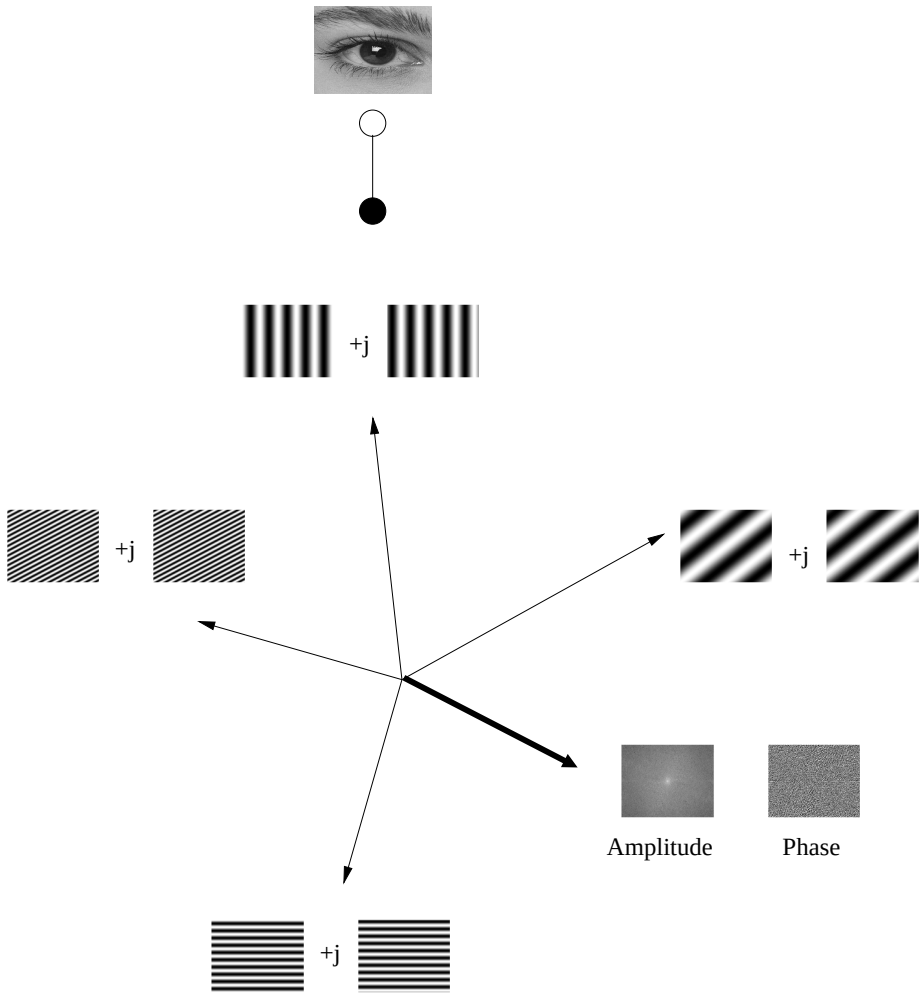
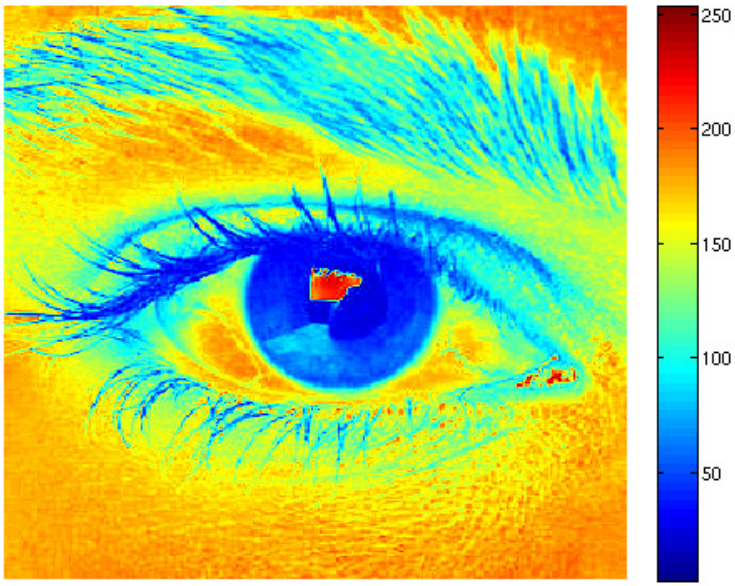


Image representation

Fourier Transform - Example



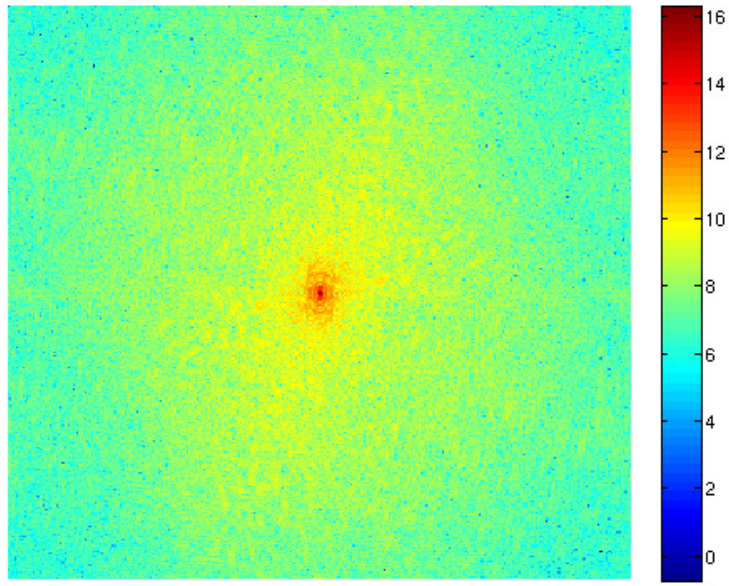


Image representation

Fourier Transform - Separability

Due to the property of separability, problems from the two-dimensional space can be reduced to applying the one-dimensional case two times. In image processing this principle is used a lot, because it permits separate computations along rows and columns of the image matrix.

A multidimensional function is separable if: $f(\mathbf{x}) = \prod_{n=1}^{N} f_n(x_n)$.

Examples:

- $\delta_{m,n} = \delta_m \cdot \delta_n$
- $e^{j\mathbf{k}^\top \mathbf{x}} = \prod_{n=1}^{N} e^{k_n x_n}$

The singular value decomposition (SVD)

$$\mathbf{G} = \sum_i \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_i \mathbf{G}_i$$

generates a matrix decomposition into a weighted sum of separable matrices. This can be used to check whether a matrix is separable ($\sigma_i = 0, \forall i \neq 1$).

Fourier Transform - Separability

Since the kernel of the discrete 2D Fourier transform is separable

$$e^{-2\pi j(mu/M+nv/N)} = e^{-2\pi jmu/M} e^{-2\pi jnv/N} = w_N^{-nv} w_M^{-mu},$$

it can be decomposed as follows:

$$G_{m,n} \quad \overset{\text{1D DFT column-wise}}{\circ\!\!-\!\!\bullet} \quad \acute{G}_{u,n} \quad \overset{\text{1D DFT row-wise}}{\circ\!\!-\!\!\bullet} \quad \hat{G}_{u,v},$$

$$G_{m,n} \circ\!\!-\!\!\bullet \acute{G}_{u,n} = \sum_{m=0}^{M-1} G_{m,n} w_M^{-mu} \circ\!\!-\!\!\bullet \hat{G}_{u,v} = \sum_{n=0}^{N-1} \acute{G}_{u,n} w_N^{-nv},$$

or in reverse order:

$$G_{m,n} \quad \overset{\text{1D DFT row-wise}}{\circ\!\!-\!\!\bullet} \quad \acute{G}_{m,v} \quad \overset{\text{1D DFT column-wise}}{\circ\!\!-\!\!\bullet} \quad \hat{G}_{u,v}.$$

2D Signal Processing

Digitalization

The transformation from a continuous image to a spatially bounded image matrix can be decomposed into three steps:

- the averaging of the intensity,
- the sampling and
- the limitation to a finite window.

From these relations

- the sampling theorem and
- the aliasing effect

can be explained. Furthermore, the sampling theorem provides the condition for a perfect

- reconstruction from sample points.

Result: From a digital image, the original continuous image can no longer be perfectly reconstructed.

2D Signal Processing

2D DFT - Finite Window

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

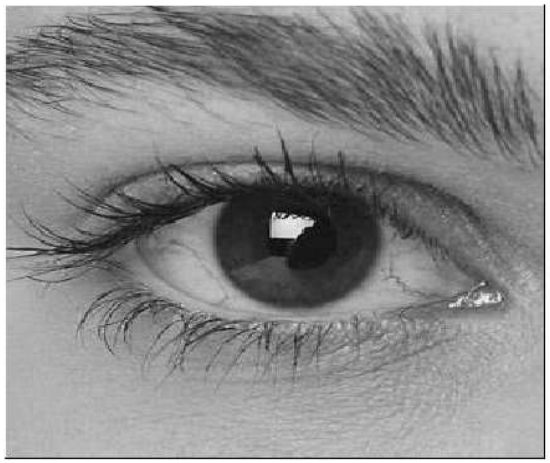
When calculating the DFT of an image, the limitation to a finite window must be taken into account. Since an image is always spatially bounded, discontinuities arise at the horizontal and vertical edges of the image. These discontinuities lead to high spectral densities along the horizontal and vertical axes in frequency space. This effect can be reduced, if a window function that dampens the image towards the image boarders (e.g. Hamming window) is used. Because of the small dynamic range of the gray values a sinusoidal window is sufficient.

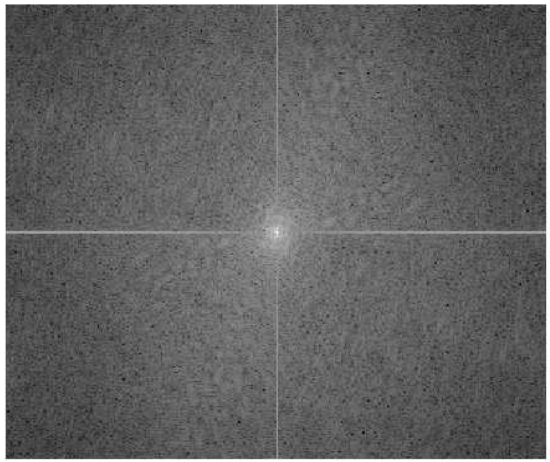
$$W_{m,n} = sin(\pi m/M)sin(\pi n/N)\,, \quad 0 \leq m < M\,,\ 0 \leq n < N\,.$$

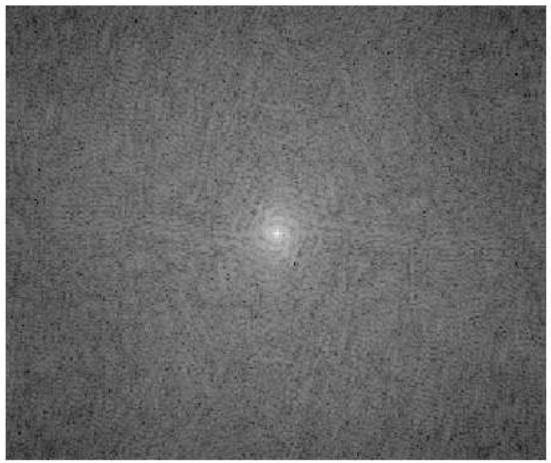
2D Signal Processing

2D DFT - Finite Window

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt







2D Signal Processing

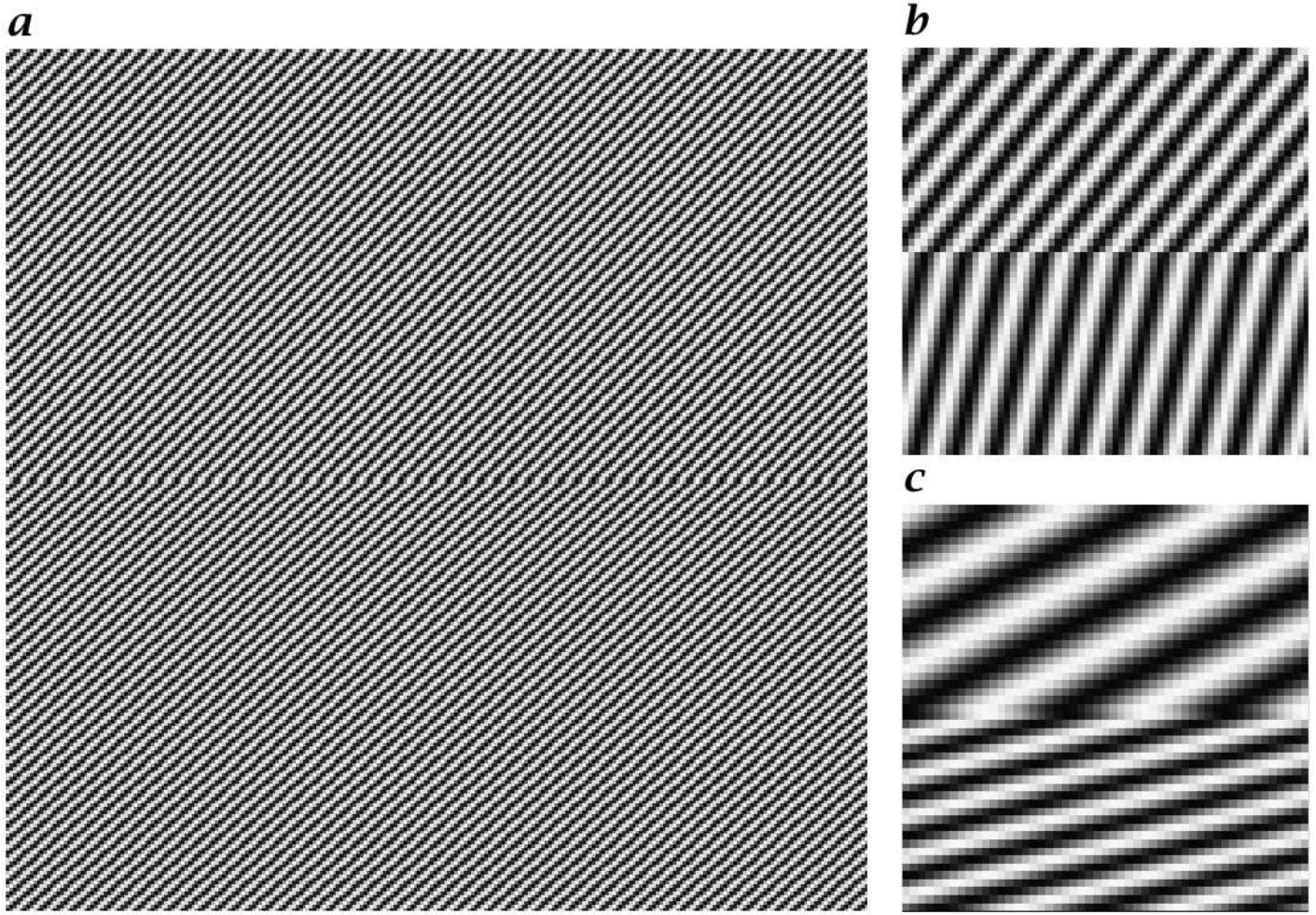
Moiré-Effect - Aliasing

Technical University of Applied Sciences Würzburg-Schweinfurt

a: Two periodic patterns with slightly different wave vectors.

b: Every fourth point sampled.

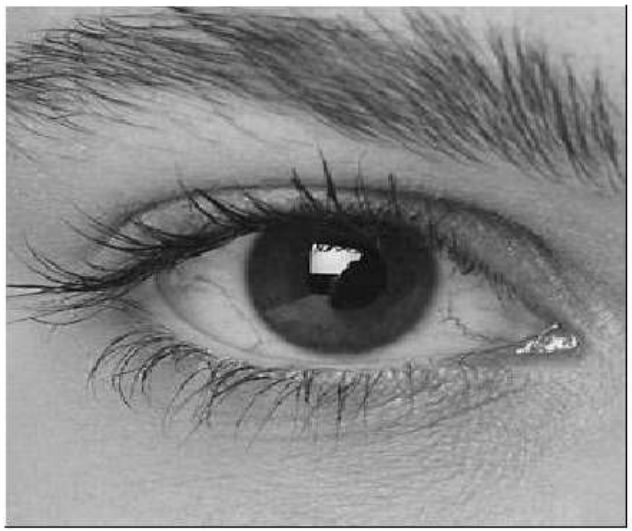
c: Every fifth point sampled.



2D Signal Processing

Moiré-Effect - Aliasing

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt



Sampling: 300 x 300

Sub-sampling: 100 x 100

Sampling Theorem A periodic structure can only be reconstructed from the samples if it is sampled at least twice per wavelength.

Neighborhood Operations - Importance

Technical University of Applied Sciences Würzburg-Schweinfurt

Neighborhood operations are probably the most important and most widely used operations in image processing. This is due to the fact that the essential information of an image is hidden in the spatial configuration of the gray values. Neighborhood operations provide information about certain properties of the spatial configuration within a small image section. Thus they provide the basis for the analysis of images and for more complex pattern and object recognition tasks.

Some applications of neighborhood operators:

- detection of simple local structures, such as edges, corners, and homogeneous regions,
- correction of disturbances,
- texture analysis,
- convolutional layers of neural networks, etc.

Neighborhood Operations - Definition

A neighborhood operator $\mathcal{N}$ concatenates the values in a neighborhood around an image point by a suitable operation and writes the result back to that point. This operation is performed for all points in the image.

$$\text{Continuous:} \quad g'(\mathbf{x}) = \mathcal{N}\left(\{g(\mathbf{x}')\}\,, \forall(\mathbf{x}-\mathbf{x}') \in \mathbb{M}\right).$$

$$\text{Discrete:} \quad G'_{m,n} = \mathcal{N}\left(G_{m'-m,n'-n}\,, \forall\left[m', n'\right]^{\top} \in \mathbb{M}\right).$$

The area $\mathbb{M}$ is called mask or window. The size and shape of $\mathbb{M}$ determine the input values from g and G for the calculation of g' and G', respectively.

Neighborhood Operations - Definition

Different types of operators:

- Nonlinear, shift-variant (location-adaptive).
- Nonlinear, shift invariant (translation invariant)
- Linear, shift-variant
- Linear, shift-invariant = LSI operator (linear, shift-invariant)

Characteristics/properties of a mask:

- Position of the reference point or anchor (usually in the center).
- mask size even or odd
- mask function symmetrical or not

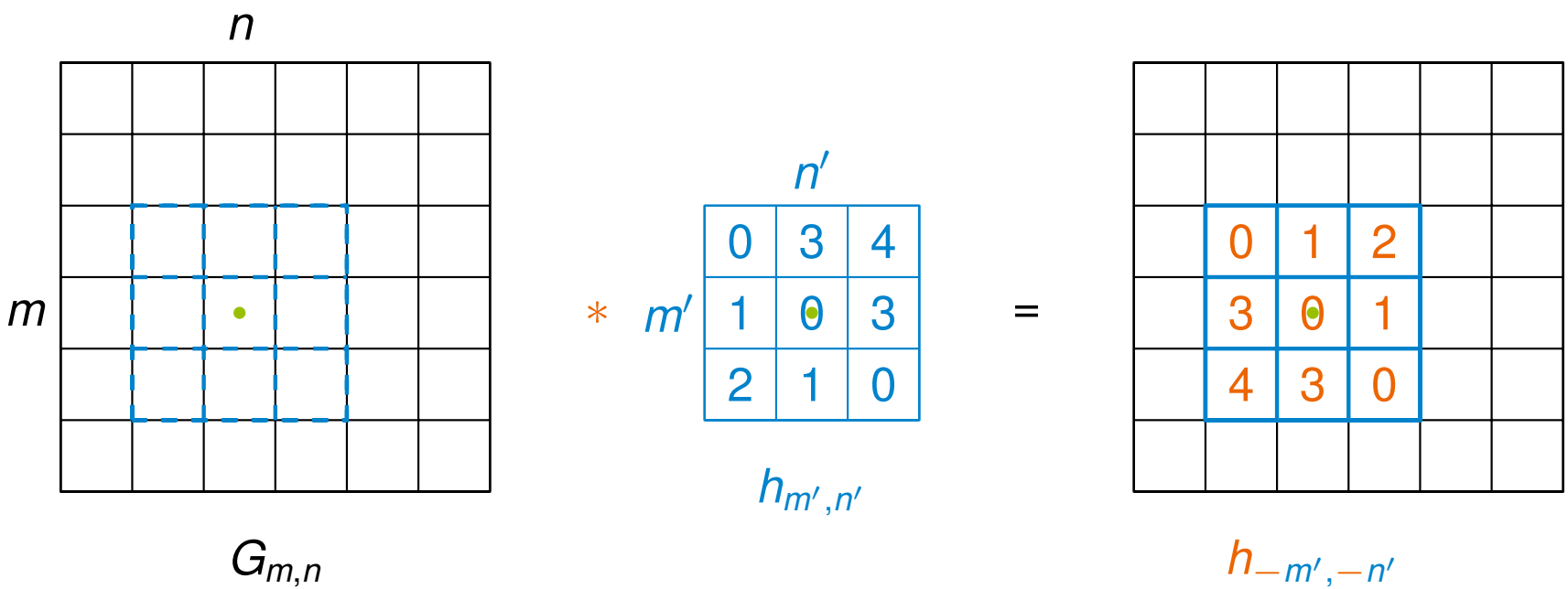
Attention: If the mask size is even, the resulting image will shift by half a pixel. This leads to errors when combining the result image with the source image.

Note: In principle, any shift can be generated with an LSI operator.

2D Signal Processing

Discrete 2D Convolution

In discrete 2D convolution, each weight in the area of the mask is point reflected at the anchor and multiplied by the corresponding pixel. The sum is written to the position of the anchor.

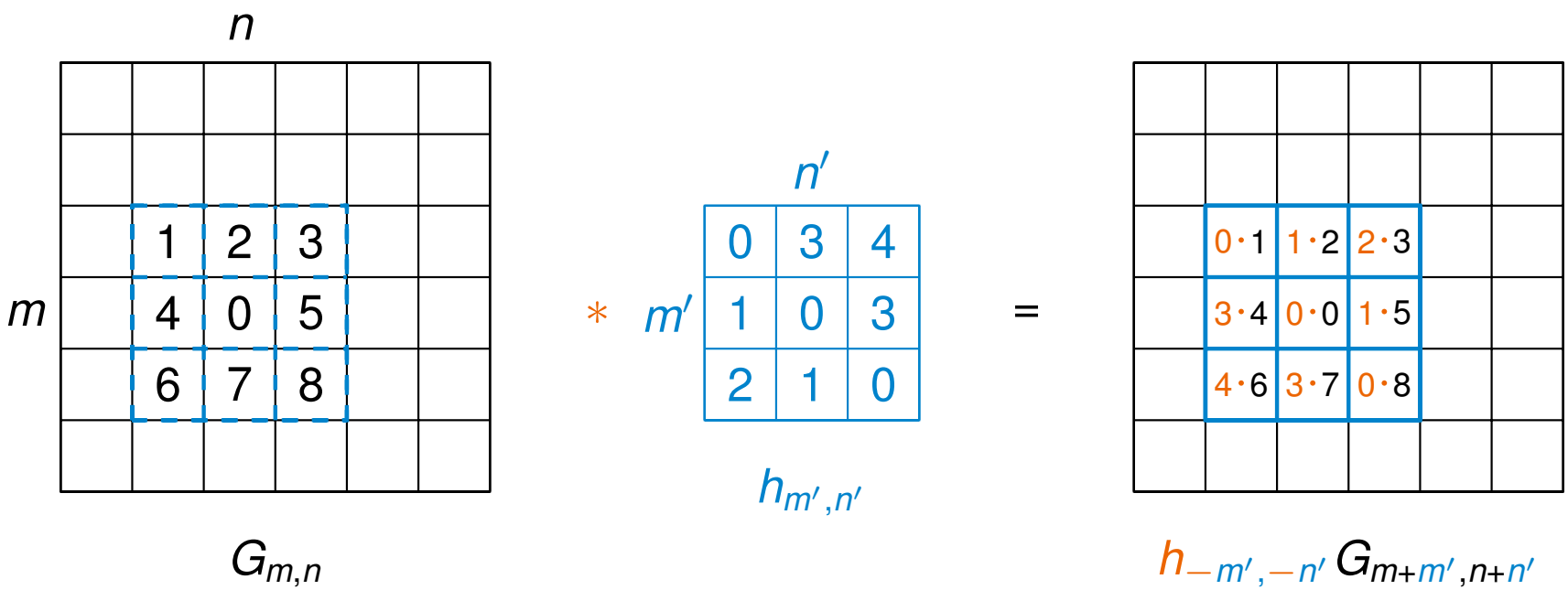


- anchor of the mask

2D Signal Processing

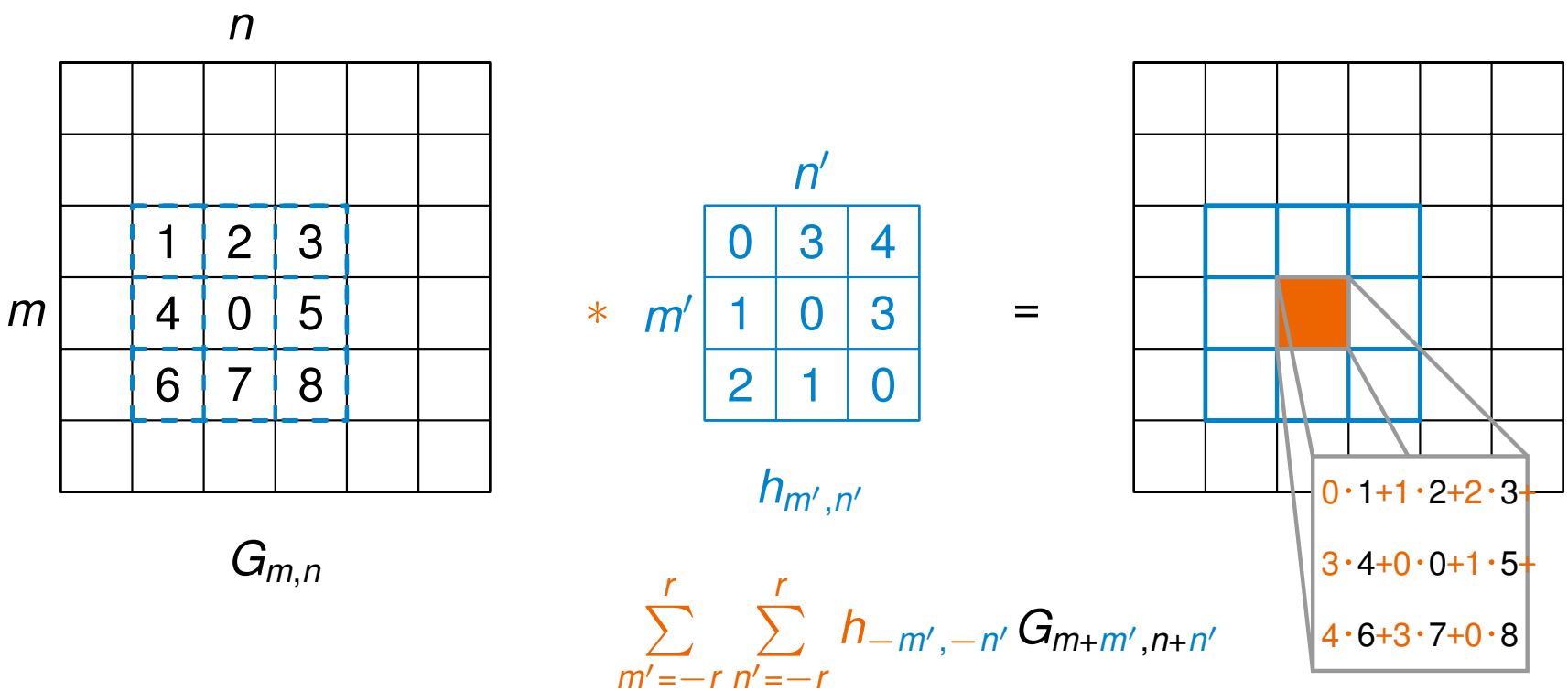
Discrete 2D Convolution

In discrete 2D convolution, each weight in the area of the mask is point reflected at the anchor and multiplied by the corresponding pixel. The sum is written to the position of the anchor.



Discrete 2D Convolution

In discrete 2D convolution, each weight in the area of the mask is point reflected at the anchor and multiplied by the corresponding pixel. The sum is written to the position of the anchor.



$$\sum_{m'=-r}^{r} \sum_{n'=-r}^{r} h_{-m',-n'} G_{m+m',n+n'}$$

2D Signal Processing

2D Convolution and Correlation

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

Discrete 2D convolution with square odd mask and anchor in the center:

$$G'_{m,n} = \mathbf{H} * \mathbf{G} = \sum_{m'=-r}^{r} \sum_{n'=-r}^{r} h_{m',n'} G_{m-m',n-n'}$$

$$= \sum_{m'=-r}^{r} \sum_{n'=-r}^{r} h_{-m',-n'} G_{m+m',n+n'}$$

Discrete 2D correlation with square odd mask and anchor in the center:

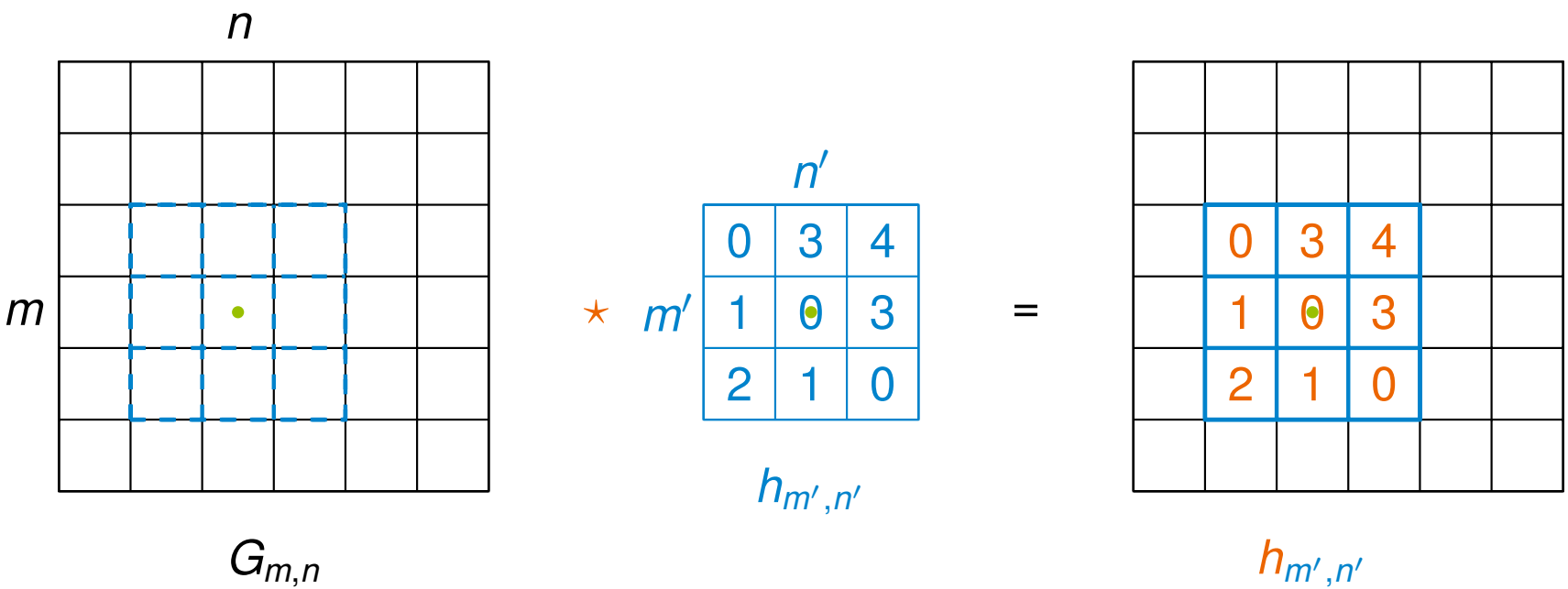
$$G'_{m,n} = \mathbf{H} \star \mathbf{G} = \sum_{m'=-r}^{r} \sum_{n'=-r}^{r} h_{m',n'} G_{m+m',n+n'}$$

If a mask is symmetric $h_{-m,-n} = h_{m,n}$, it is not changed by a reflection.
Thus, convolution and correlation do not differ for symmetric masks.

2D Signal Processing

2D Correlation

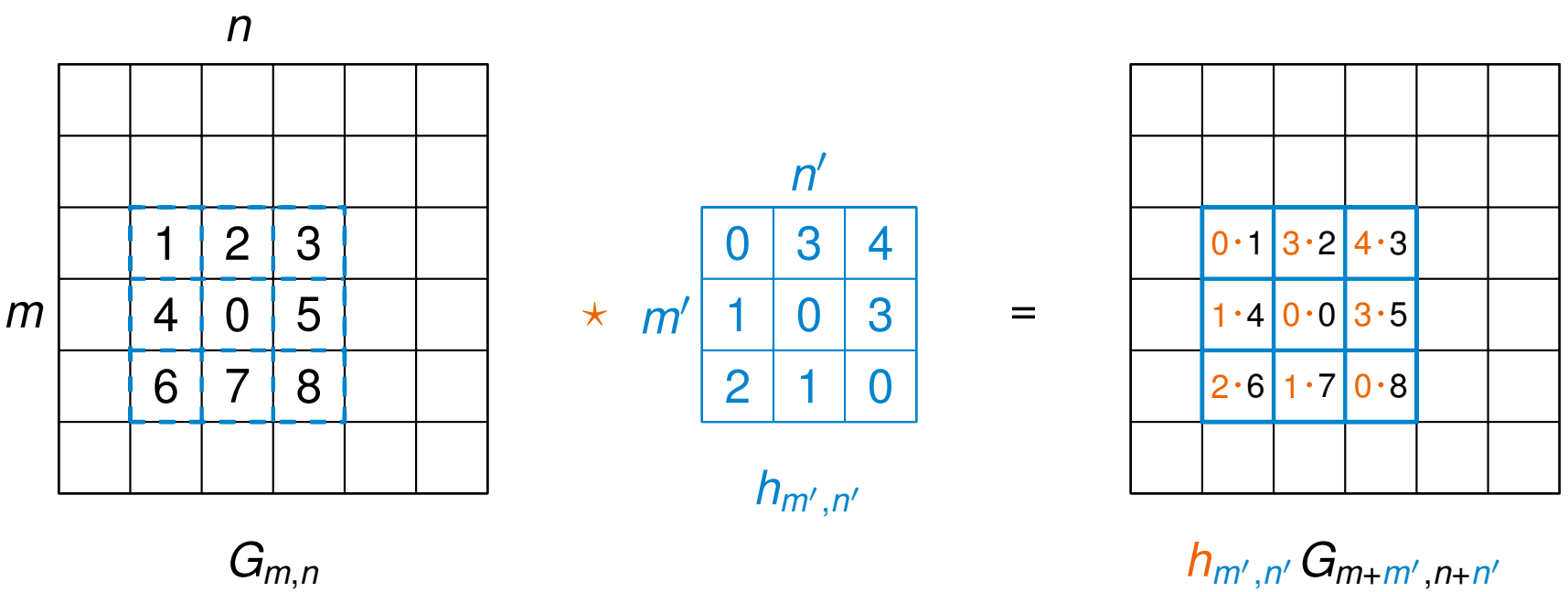
A 2D correlation with a certain mask equals a 2D convolution with the same point reflected mask and vice versa.



- anchor of the mask

2D Correlation

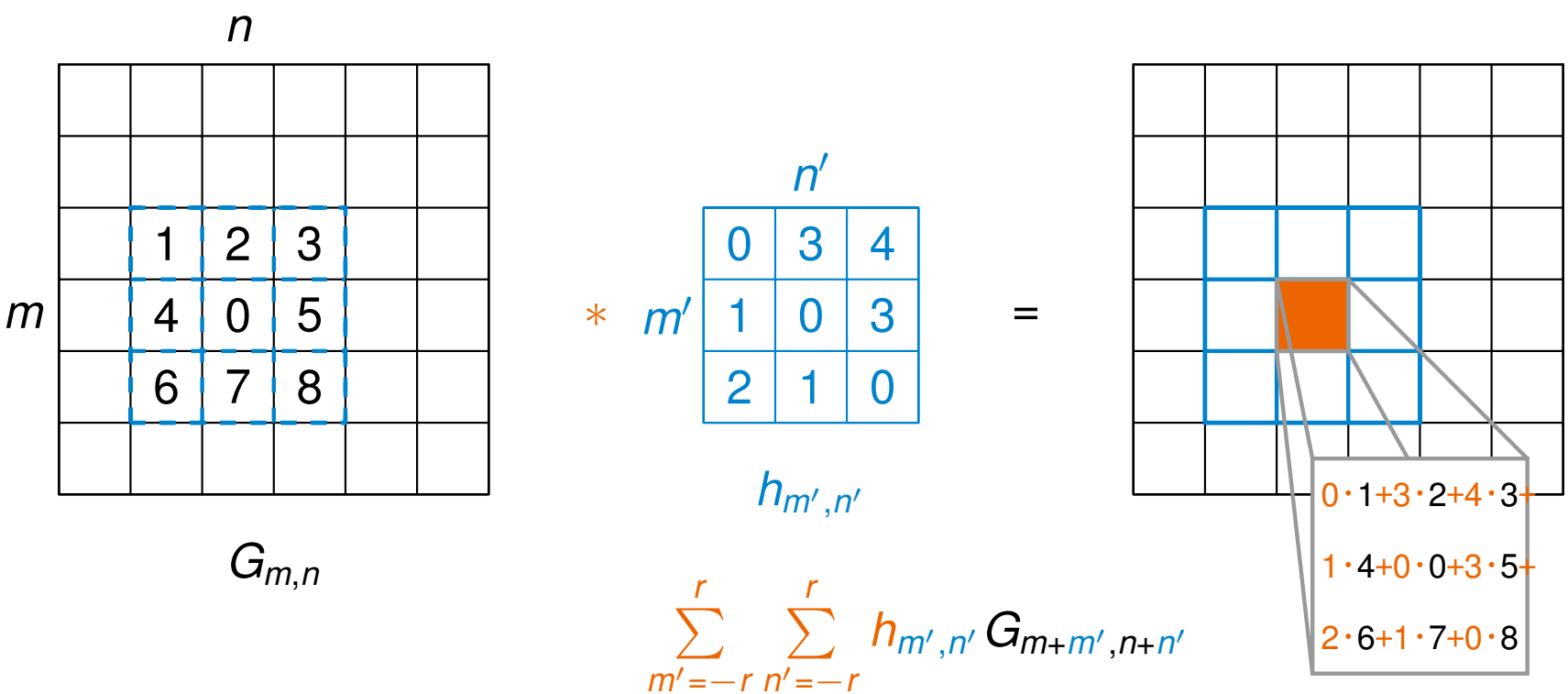
A 2D correlation with a certain mask equals a 2D convolution with the same point reflected mask and vice versa.



2D Signal Processing

2D Correlation

A 2D correlation with a certain mask equals a 2D convolution with the same point reflected mask and vice versa.



$$\sum_{m'=-r}^{r} \sum_{n'=-r}^{r} h_{m',n'} G_{m+m',n+n'}$$

Convolution & Correlation - Image Borders

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

At the border of the image, the area of the mask extends beyond the border of the image. What kind of pixels should be assumed there? If the convolution in the spatial space shall correspond to a multiplication in the Fourier space, then the image matrix must be considered as a periodically repeating structure. If this is assumed, one speaks also of cyclic convolution.

In practice, however, a choice is made between three other methods, with a trade-off between low artifacts and fast calculation.

- Fill border with zeros (fast but lots of artifacts)
- Mirror border
- Extrapolate border (e.g. continue gray value, overemphasize borders)

2D Signal Processing

Convolution & Correlation - Efficient Computation

Technical University of Applied Sciences Würzburg-Schweinfurt

How to calculate a discrete convolution with the smallest possible number of arithmetic operations depends on the one hand on the mask size and on the other hand on the properties of the convolution kernel. Separability and Symmetry are the two properties, whose consistent utilization allow large accelerations.
Above a certain mask size, it is worthwhile to use the additional forward and backward transformation into the Fourier space, in order to carry out a little computationally expensive multiplication in the frequency domain. (Filters can possibly be transformed already before the runtime).
Other ways of speeding up are decomposition of filter masks and lookup tables. If the filter kernels are very small, all possible multiplications can be calculated in advance and stored in lookup tables. Sometimes it is possible to convert the convolution with a large filter mask into a sequential execution of convolutions with smaller masks.

thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

Fourier Transform - Convolution Theorem

Discrete 2D convolution is one of the most important operations of image processing in image space (see also filtering):

$$G'_{m,n} = H_{m,n} * G_{m,n} = \sum_{m'=0}^{M-1} \sum_{n'=0}^{N-1} H_{m',n'} G_{m-m',n-n'} \ .$$

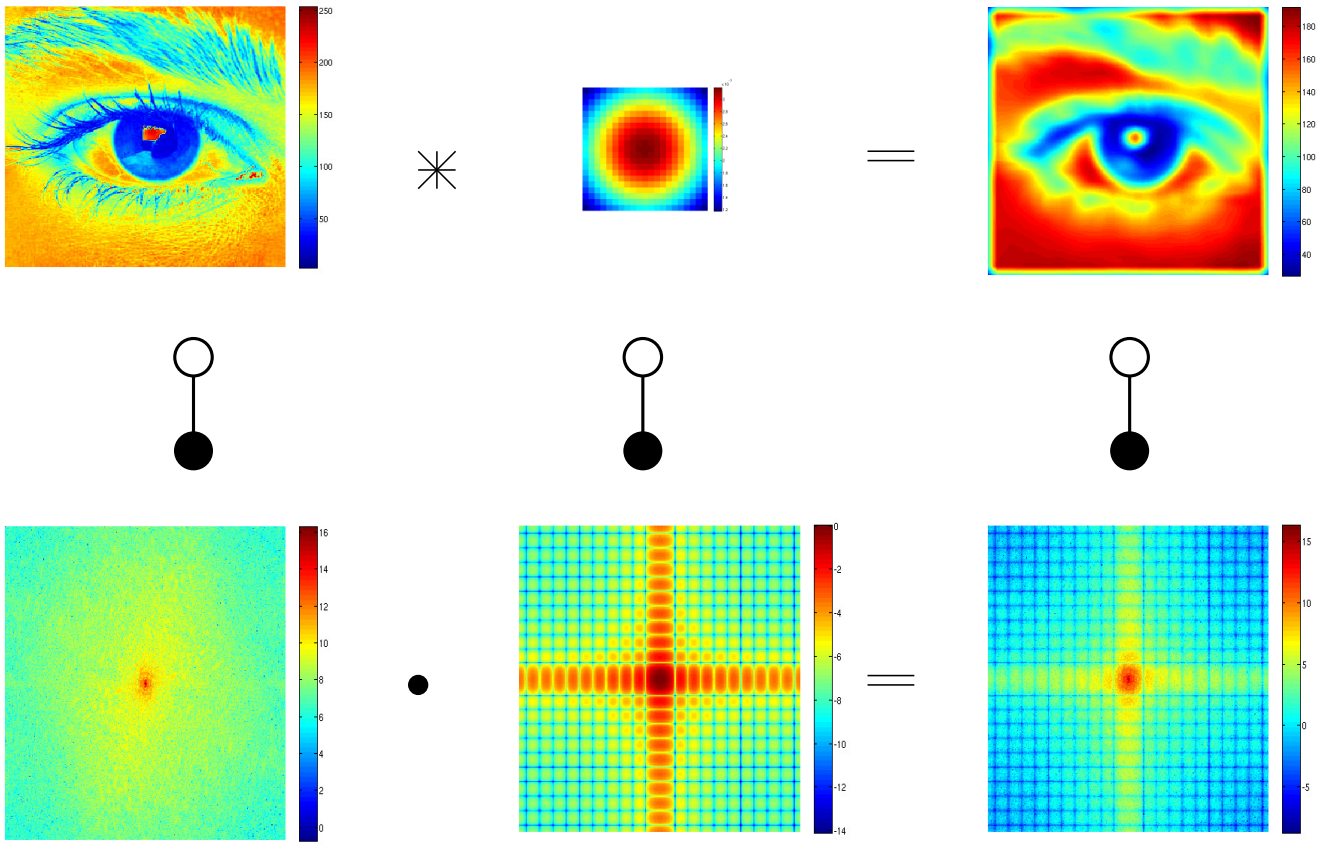
Convolution Theorem:

$$G'_{m,n} = H_{m,n} * G_{m,n} \circ\!\!-\!\!\bullet\, MN \hat{H}_{u,v} \hat{G}_{u,v} \ .$$

This means that convolution in image space is equivalent to complex multiplication in Fourier space. This property is mainly used to quickly compute a set of convolutions of an image with different kernels by transforming all kernels and the image into Fourier space, multiplying each kernel by the image, and then back-transforming all the resulting images into the image space. Additional acceleration is achieved by efficient calculation of the DFT with the FFT (fast Fourier transform).

2D Signal Processing

Fourier Transform - Convolution Theorem

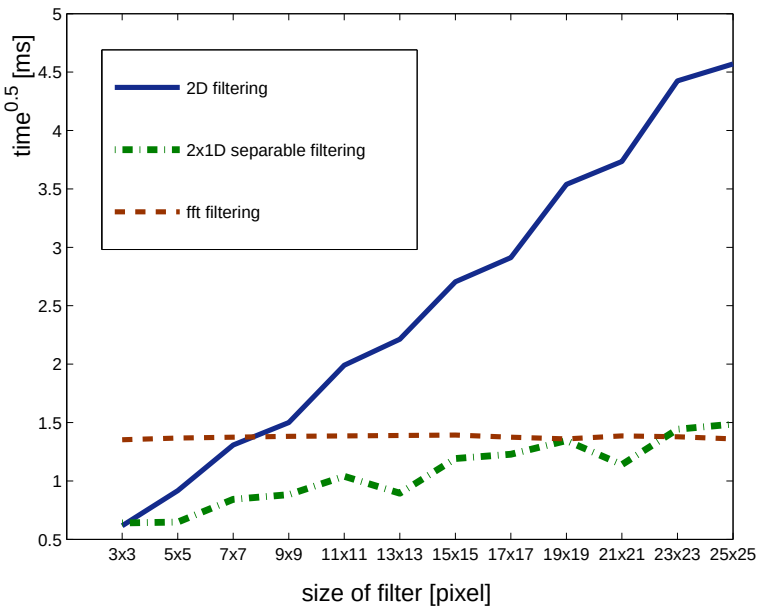


Zero-padding is used to align the filter size to the image size.

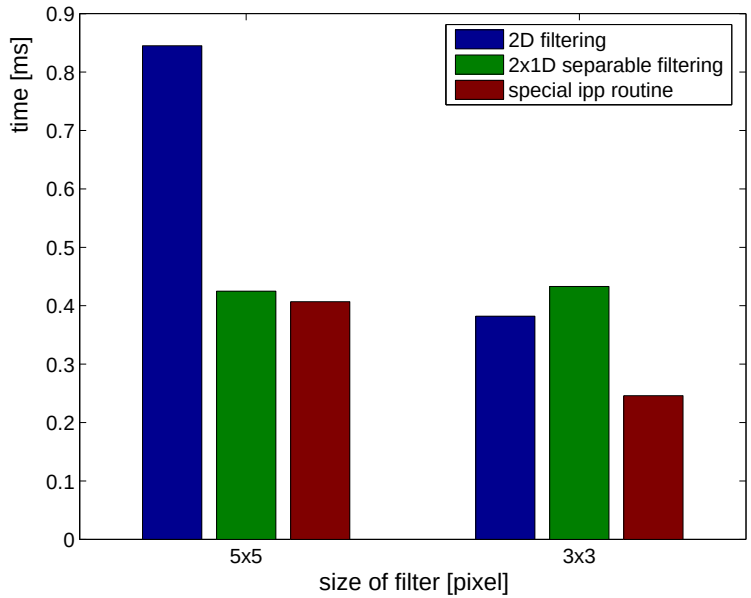
thws
Technical University
of Applied Sciences
Würzburg-Schweinfurt

Efficient Convolutions - IPP Comparison

Comparison of different possibilities of accelerated calculation of convolutions depending on the size of the filter masks. A C library optimized for image processing was used, the Intel Integrated Performance Primitives (Intel® IPP).



FFT vs. separable convolution



Separable convolution vs. Lookup-tables

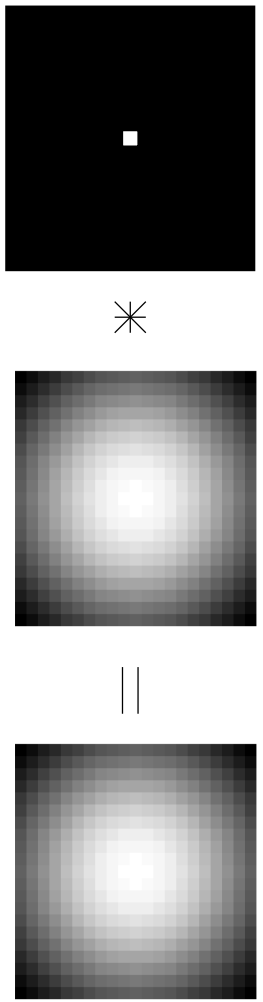
2D Signal Processing

Convolution - Point spread function

The point spread function (PSF) is the response of a convolution operation to a point image (base image) and is identical to the convolution mask:

$$\mathbf{P} = \mathbf{H} * \mathbf{B}^{m,n} = \mathbf{H}\,.$$

The point response shows how each pixel propagates into the neighborhood. Since the convolution is an LSI operator, i.e. the superposition principle applies and it is shift invariant, the result of a convolution is completely determined by the response to any base image.



* Analogy to the impulse response or weight function of an LTI system.

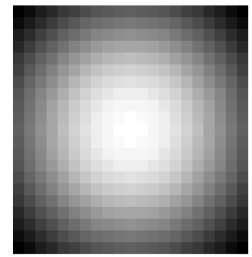
2D Signal Processing

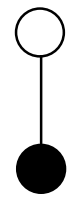
Convolution - Transfer function

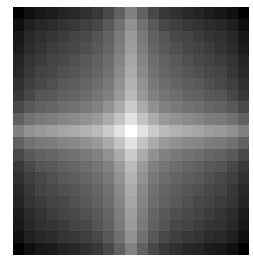
The Fourier transform of a convolution mask or PSF is called transfer function (TF) of a filter*.

TF: $\hat{H}$

For each wavenumber vector, the TF specifies the factor by which the corresponding periodic structure is multiplied by a filter operation. Since this factor is complex, both the amplitude as well as the phase of the periodic structure is changed. So you can see which frequencies in the image are filtered by how much. Thus, the TF is very useful to develop a filter with special properties.







* Analogy to the frequency response of an LTI-system.