Image Fusion

Tutorial

Dr.-Ing. Michael Heizmann

2006 IEEE International Conference on
Multisensor Fusion and Integration
for Intelligent Systems (MFI 2006)

Heidelberg, September 03, 2006

Contact Information:
Dr.-Ing. Michael Heizmann
Fraunhofer IITB Institut für Informations- und Datenverarbeitung
Fraunhoferstraße 1, D-76131 Karlsruhe
Michael.Heizmann@iitb.fraunhofer.de

Fraunhofer Institut Informations- und Datenverarbeitung

Overview

- Why image fusion?
- Acquisition of data foundation:
  - Image series
  - Information content of image data
- Abstraction levels in image fusion
- General structure of image fusion
- Application examples

Fraunhofer Institut Informations- und Datenverarbeitung

Definition: Image fusion

Image fusion can be defined as

- combination of **images** (input quantity)
- from **different sources** (image sensors, cameras)
- with the aim to obtain **new or more precise knowledge** (output quantities: image, features, decisions) about the scene (objects, events or situations)

Objective: Generate a result which describes the scene »better« than any single image with respect to some relevant properties

Condition: Information sources used for fusion refer to a **common underlying event**

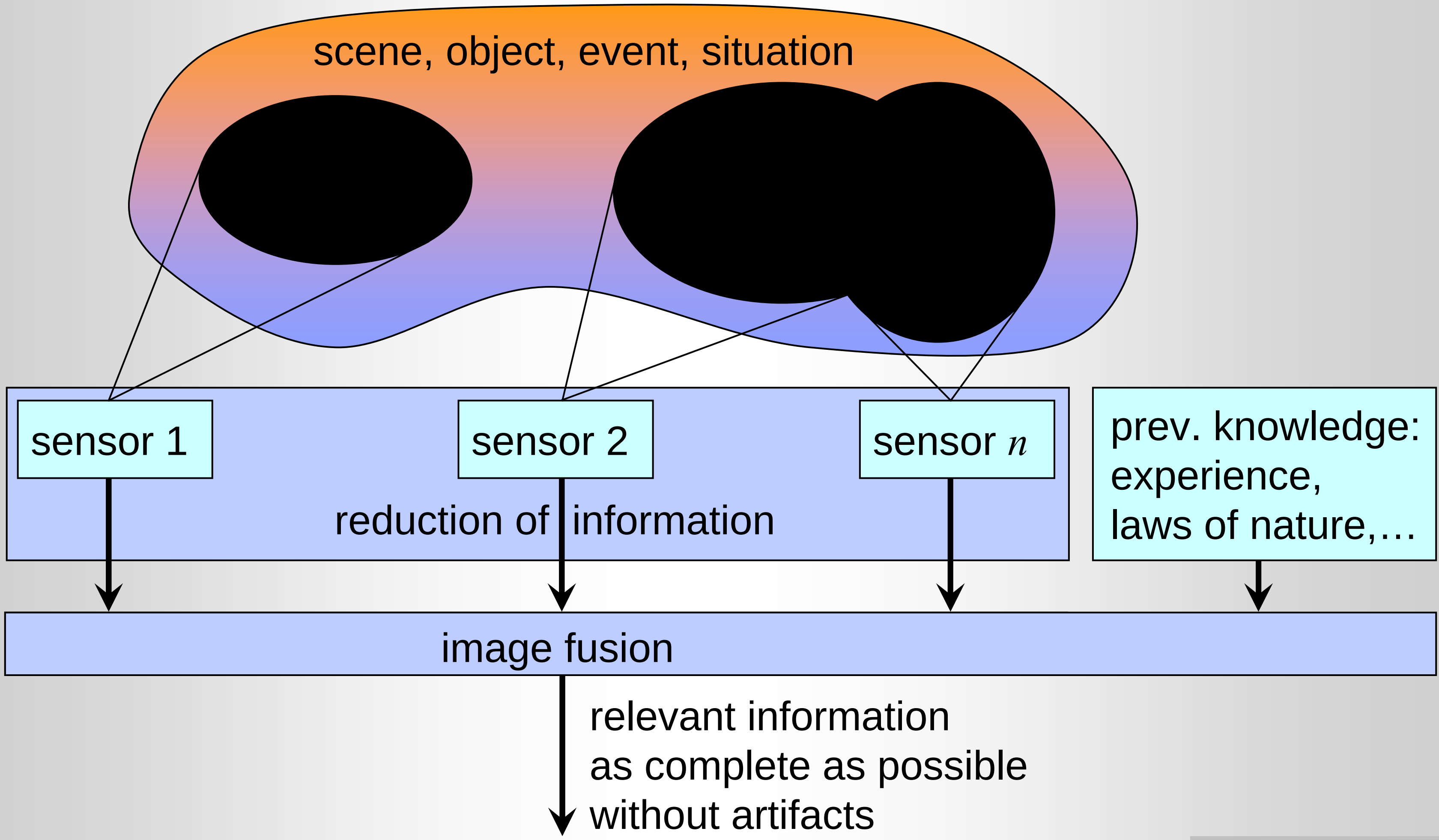
IITB
Fraunhofer Institut Informations- und Datenverarbeitung

Benefits of image fusion

- **Higher accuracy and reliability**
  by evaluation of redundant information,
  Requires sensors receiving identical properties of the scene/object
- **Feature vector with higher dimensionality**
  by evaluation of complementary information,
  Requires sensors receiving different properties of the scene/object
- **Faster acquisition of information**
  Simultaneous data acquisition using multiple sensors
- **Cost-effective acquisition of information**
  Substitution of expensive special sensors with several low-cost sensors

Fraunhofer Institut Informations- und Datenverarbeitung

Why image fusion?

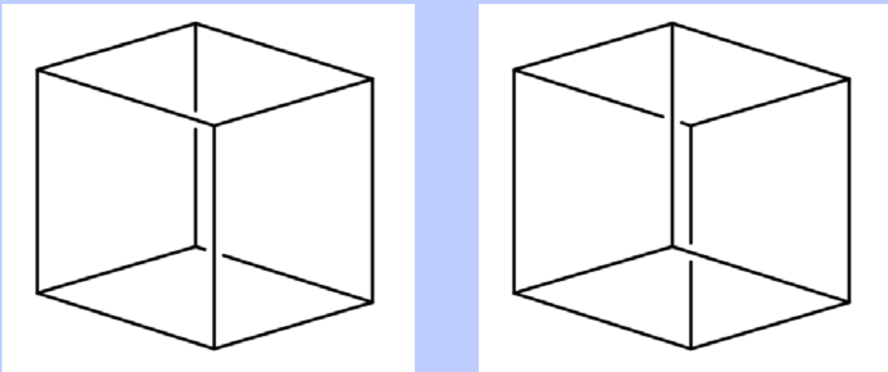


Reduction of information by the imaging sensor

- **Windowing**: Observation of a part of the scene limited in time and space
- **Projection**: Reduction of dimensions (geometrical, spectral, temporal dimension)

$\Rightarrow$ Mapping of the world in images is not invertible

Identical objects? flat objects? …?



Reduction of information by the imaging sensor

- **Sampling**:
  - **Discretization of space**:
    - Convolution of the spatial irradiance on the imaging sensor with the windowing function of the pixels
    - Sampling with the sensor grid
  - **Discretization of time**:
    - Convolution of the temporal irradiance with the temporal exposure function
    - Sampling with the shutter frequency
  - **Quantization of values**:
    Discrete values, e.g. gray-value image with 8 bits: 256 values
- **Infliction of disturbances**:
  Caused e.g. by thermal noise of the imaging sensor

Fraunhofer Institut Informations- und Datenverarbeitung

Data foundation: Image series

Variation of illumination and/or observation parameters

- **Illumination**
  - Direction
  - Spatial distribution of intensity (structured illumination)
  - Spectrum (»color«)
  - …
- **Observation**
  - Camera position and orientation (object pose)
  - Spectral response
  - Focus
  - …

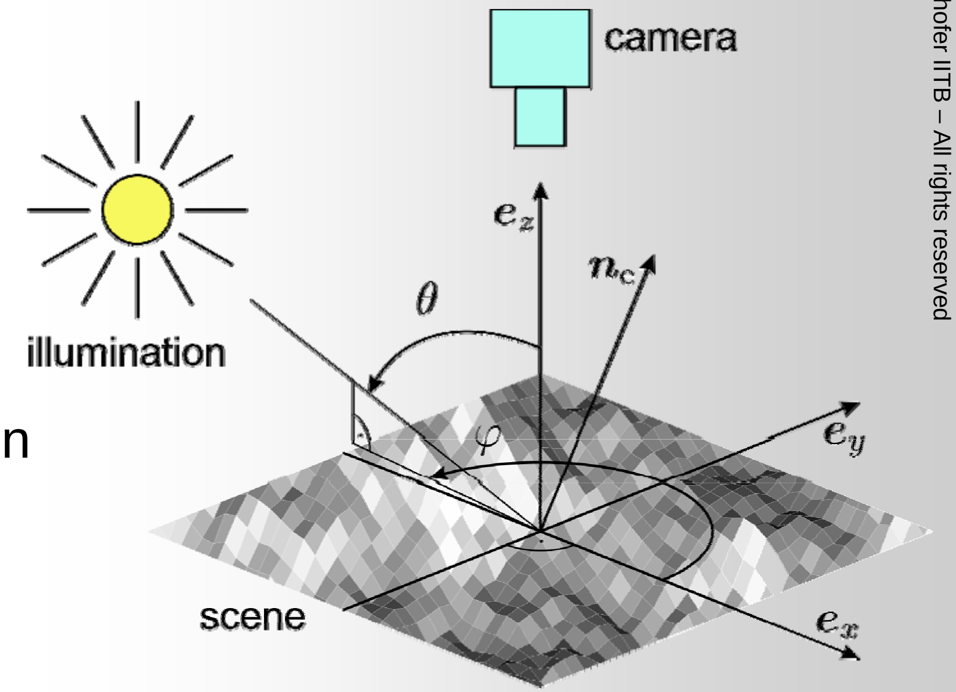


Image series: **Multidimensional data objects**
Each varied parameter represents its own dimension

Fraunhofer Institut Informations- und Datenverarbeitung

Data foundation: Image series

Example: Image series with varied azimuth of a directional illumination

→ Irradiance on the imaging sensor: $E(x, y, \phi): D \to V$,

$D = [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}] \times [0, 2\pi)$, V: Range of values $\left[\frac{W}{m^2}\right]$

Discretization and quantization:

→ Discrete image series

$$d(x, y, \phi): D \to V, \quad D = \{0, 1, \ldots, x_{\max}\} \times \{0, 1, \ldots, y_{\max}\} \times \{0, \Delta\phi, \ldots, 2\pi - \Delta\phi\},$$

$$V = \{0, \ldots, d_{\max}\}$$

here: $D = \{0, 1, \ldots, 511\} \times \{0, 1, \ldots, 511\} \times \{0, 120°, 240°\}$,

$$V = \{0, \ldots, 255\}$$

Fraunhofer Institut Informations- und Datenverarbeitung

Characteristics of imaging sensors

- **Commensurability**
  (data with equal number of dimensions, here: two dimensions)
  Imaging sensors: Principally commensurable data

- **Virtual sensors**
  Image acquisition with the same sensor, but with at least one varied acquisition parameter

  Examples:
  - **Gray-value camera**: At least different acquisition times
  - **RGB-camera**: At least different spectral response of the R-, G-, B- subpixels

Fraunhofer Institut Informations- und Datenverarbeitung

Characteristics of imaging sensors

- **Homogeneity**
  Sensors recording the same/comparable physical quantity
  Examples:
  Similar cameras with different observation parameters

  For homogeneous sensors:
  - Processing of data usually practicable without costly preprocessing
  - Example: Gray-value cameras with identical spectral response and comparable linearity

  For inhomogeneous sensors:
  - Information not directly linkable
  - Usually preprocessing necessary (e.g. feature extraction, classification)
  - Example: Fusion of gray-value images and 3D-data

Characteristics of imaging sensors

- **Collocated sensors**
  Image acquisition with constant relative sensor position and orientation and constant projection scale:
  Imaging of **identical parts** of the scene

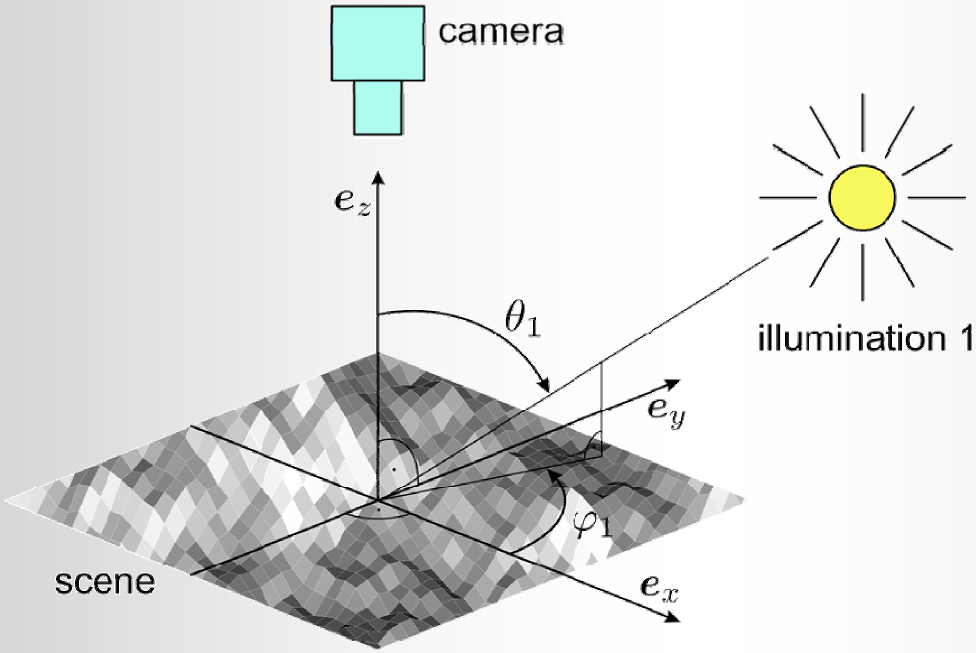
  - Example: Stationary imaging sensor used for several shots
  - If sensors are not collocated: Image registration required
    - Identification of image features (edges, corners)
    - Alignment of the images by means of geometrical transforms (translation, rotation, scaling, projective transform)

Fraunhofer Institut Informations- und Datenverarbeitung

Homogeneous collocated sensors

Example:

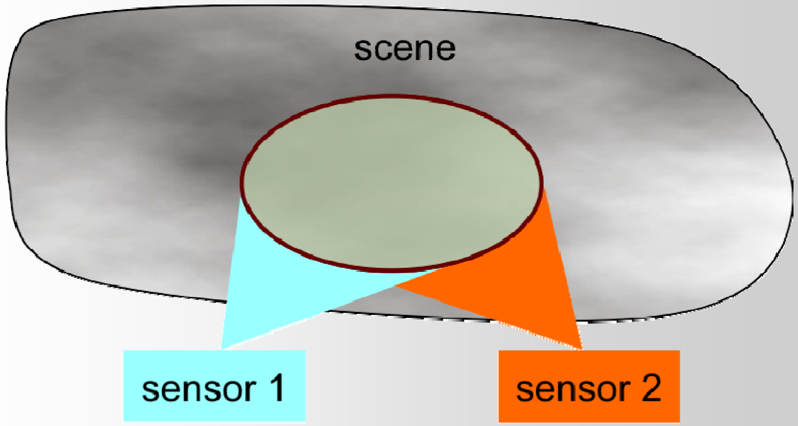
- Stationary camera, constant observation parameters
- Varied illumination



Information content of image data

Imaging sensors record only **parts of the entire information** available about the scene:

- **Redundant (concurrent) information**
  - Useful information is contained in the same way in all images of the series
  - Homogeneous sensors
  - Concurrent fusion reasonable, e.g. by averaging
  - All images of the series contribute equivalently to the fusion result
  - Examples: Reduction of noise, increase of reliability



Redundant information

Example:

- Simulation: Stationary camera, heavy additive Gaussian noise: $\sigma_{\text{sensor}} = 20$

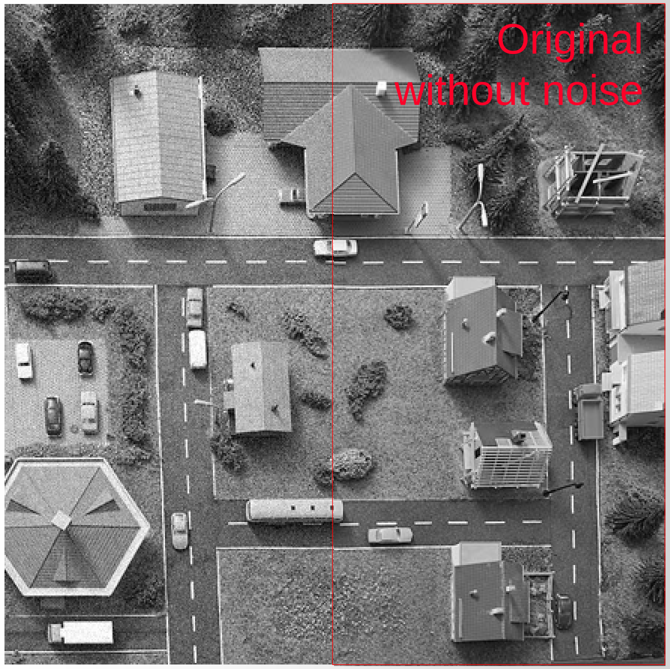
Image without noise

Image with heavy noise

Redundant information

- Suppression of noise by means of multiple recordings and averaging
- Assumption: Noise in different recordings is uncorrelated

Image series $N=4,\ \sigma_{\text{sensor}}=20$



Fusion result $\sigma_{\text{result}}=\dfrac{\sigma_{\text{sensor}}}{\sqrt{N}}=10$

Fraunhofer Institut Informations- und Datenverarbeitung

Redundant information

- Modeling of image acquisition:
  Measured image intensity: $g_i(\boldsymbol{x}) = d(\boldsymbol{x}) + r_i(\boldsymbol{x}),\ i = 1,\ldots,N$
  - Useful information: $d(\boldsymbol{x})$ deterministic signal
    represents scene properties
  - Sensor noise (e.g. thermal): $r_i(\boldsymbol{x}),\quad \mathrm{E}\{r_i(\boldsymbol{x})\} = 0,$
    $\mathrm{E}\{r_i(\boldsymbol{x}) r_j(\boldsymbol{x})\} = \delta_i^j \sigma_R^{\ 2}$ uncorrelated

- Statistics of one recording: $\mathrm{E}\{g_i(\boldsymbol{x})\} = d(\boldsymbol{x}),$

$$\sigma_G^{\ 2} = \mathrm{var}\{g_i(\boldsymbol{x})\} = \mathrm{E}\left\{\left(g_i(\boldsymbol{x}) - \mathrm{E}\{g_i(\boldsymbol{x})\}\right)^2\right\}$$
$$= \mathrm{E}\left\{\left(r_i(\boldsymbol{x})\right)^2\right\} = \sigma_R^{\ 2}$$

- Averaging of N recordings:

$$m(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^{N} g_i(\boldsymbol{x}) = d(\boldsymbol{x}) + \frac{1}{N}\sum_{i=1}^{N} r_i(\boldsymbol{x})$$

Fraunhofer Institut Informations- und Datenverarbeitung

Redundant information

- Statistics of the average $m(\boldsymbol{x})$ of N recordings:

$$\mathrm{E}\{m(\boldsymbol{x})\} = d(\boldsymbol{x}) + \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}\{r_i(\boldsymbol{x})\} = d(\boldsymbol{x}) \qquad \text{(as before)}$$

$$\sigma_M{}^2 = \mathrm{var}\{m(\boldsymbol{x})\} = \mathrm{E}\left\{\left(m(\boldsymbol{x}) - \mathrm{E}\{m(\boldsymbol{x})\}\right)^2\right\}$$

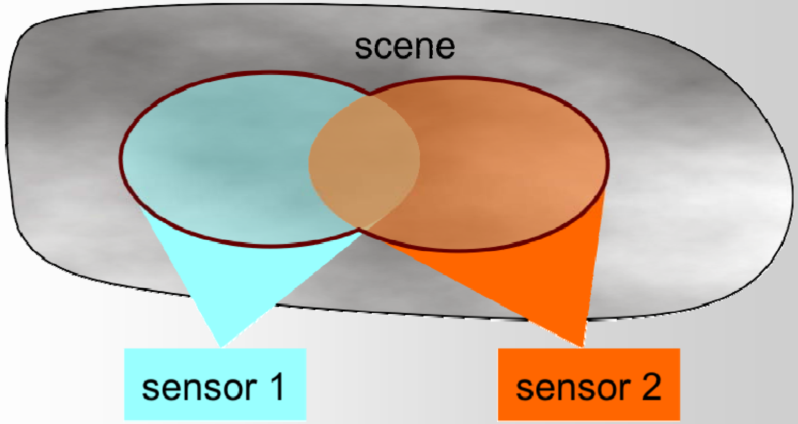
$$= \mathrm{E}\left\{\left(\frac{1}{N}\sum_{i=1}^{N} r_i(\boldsymbol{x})\right)^2\right\} = \frac{1}{N^2}\,\mathrm{E}\left\{\sum_{i=1}^{N} r_i{}^2(\boldsymbol{x}) + 2\underbrace{\sum_{i=1}^{N}\sum_{\substack{j=1\\ i\neq j}}^{N} r_i(\boldsymbol{x})\, r_j(\boldsymbol{x})}_{=0}\right\}$$

$$= \frac{1}{N}\sigma_R{}^2$$

$\rightarrow$ Variance of average $m(\boldsymbol{x})$ is reduced by factor N compared to single recording $g_i(\boldsymbol{x})$

Information content of image data

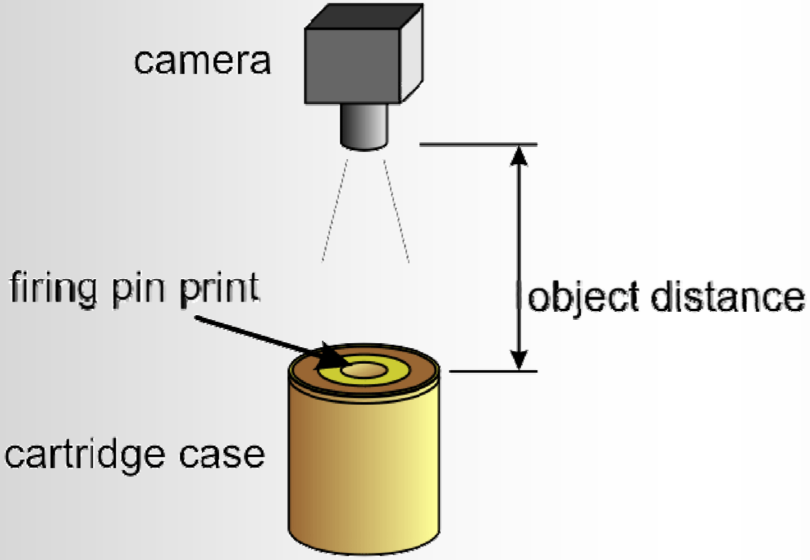
- **Complementary information**
  - Useful information for a certain image location is concentrated in one or few images of the series
  - Complementary fusion:
    Selection of the useful information
  - Concurrent fusion (e.g. averaging) usually useless:
    Destructive combination
  - Examples:
    - Synthetically enhanced depth of focus
    - High-contrast images by fusion of illumination series
    - Generation of panoramic images
    - Multispectral images

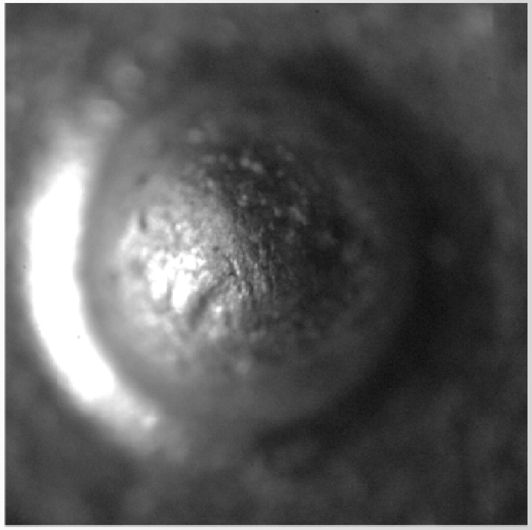


Fraunhofer Institut Informations- und Datenverarbeitung

Example: Focus series

- Variation of the object distance
- Constant projection scale in the focal plane (»Virtually telecentric projection«)

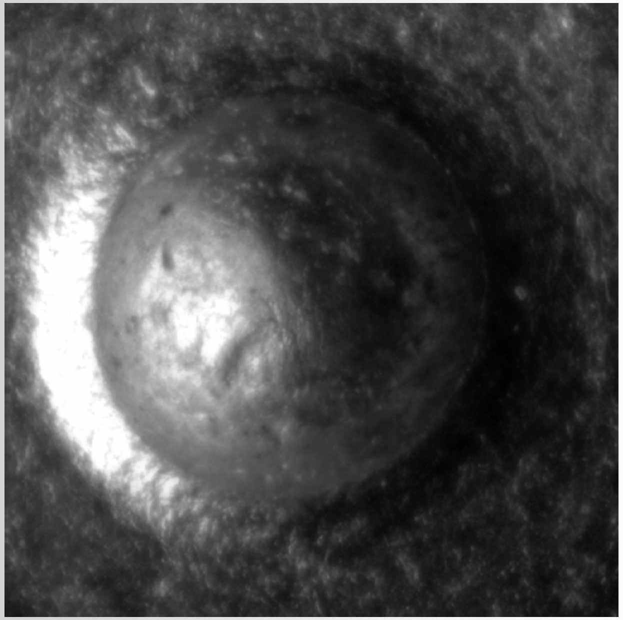




Fraunhofer Institut Informations- und Datenverarbeitung

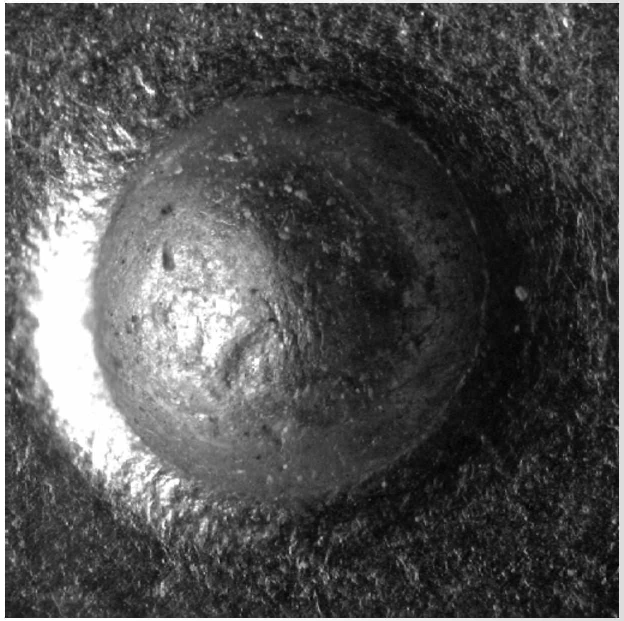
Complementary information

- Combination of the images of the series



Averaging

→ Low-contrast image



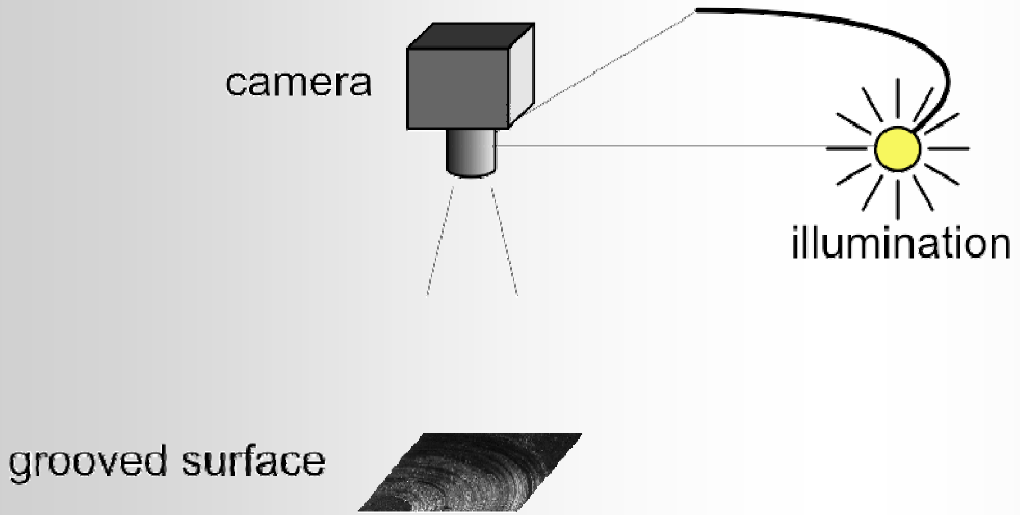
Cutting and combining of areas with satisfactory focus

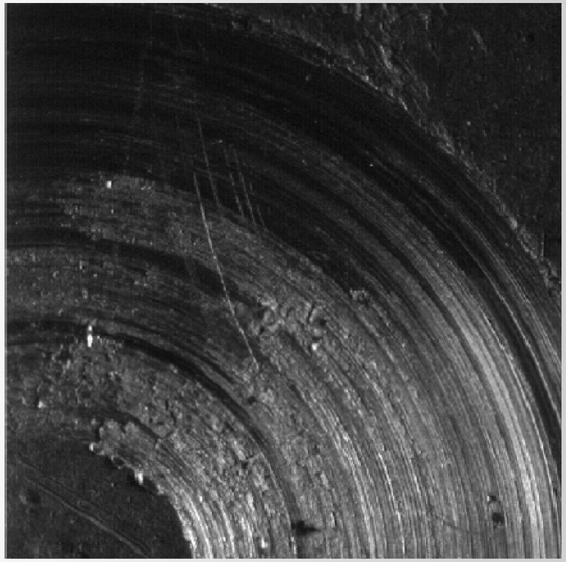
→ Synthetically enhanced depth of focus

Complementary information

Example: Illumination series

- Variation of the illumination direction: Azimuth of a point light source
- Optimal contrast for directional illumination perpendicular to the grooves





Complementary information

- Optimization of the local contrast

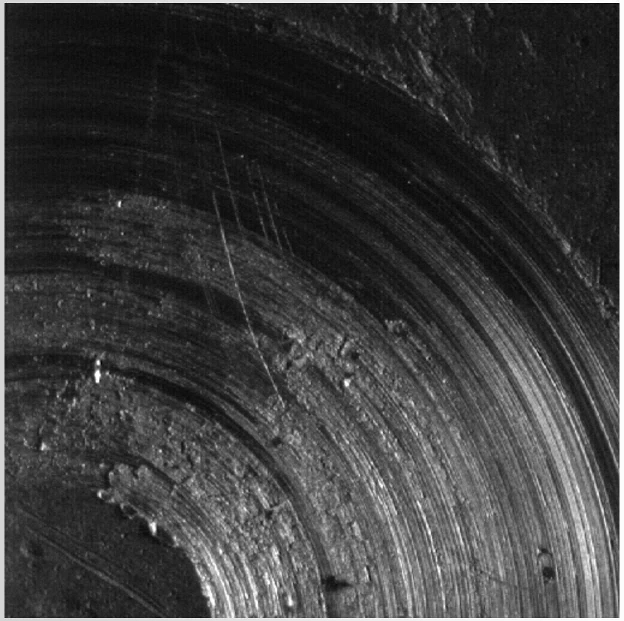
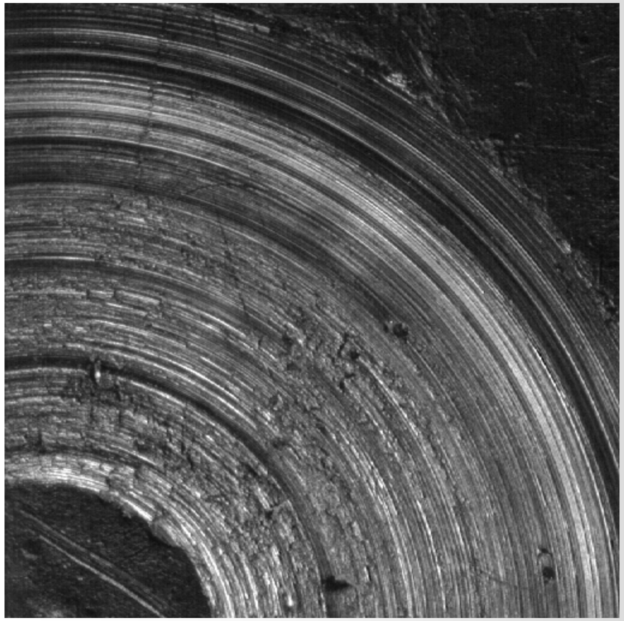


Image series



Cutting and combining of areas with high contrast
→ Enhanced visibility of grooves

Complementary information

Example: Moving camera

- Generation of a panoramic image

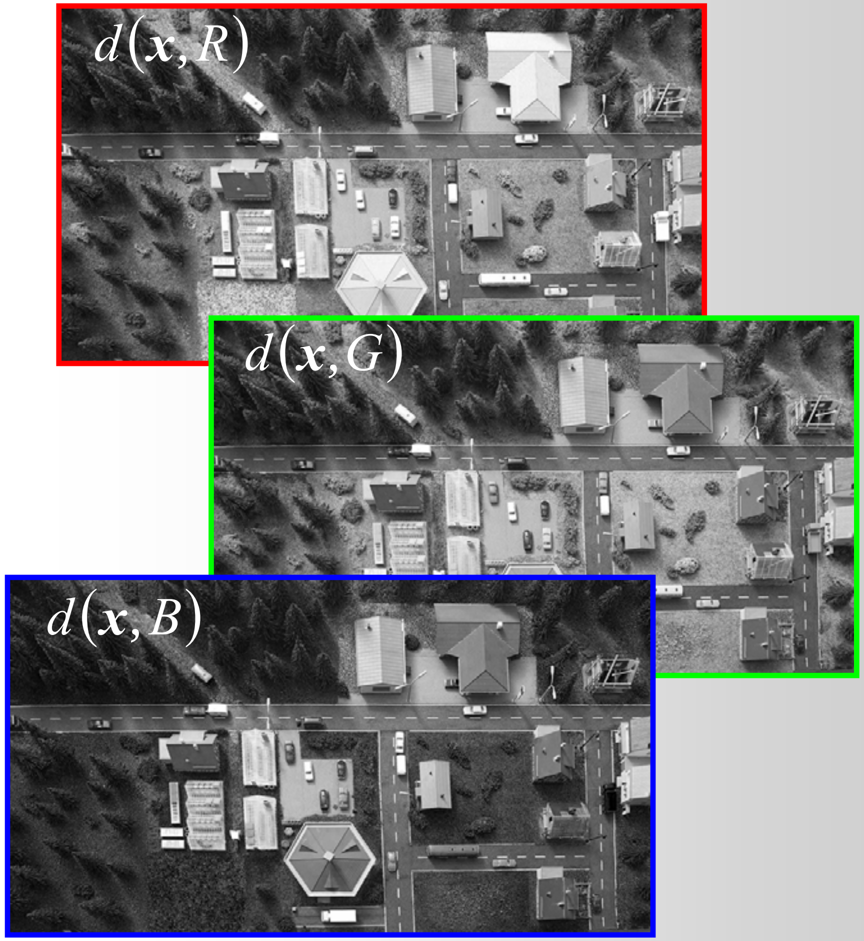
Fraunhofer Institut Informations- und Datenverarbeitung

Complementary information

Example: Stationary RGB-camera

- (Red/Green/Blue)

$d(\boldsymbol{x}, p),\ p \in \{R, G, B\}$



IITB
Fraunhofer Institut Informations- und Datenverarbeitung

Complementary information

- Fusion of RGB-images

Brightness
(here: Intensity in HSV color space)

$$I(\boldsymbol{x}) = \max\left(d(\boldsymbol{x}, R), d(\boldsymbol{x}, G), d(\boldsymbol{x}, B)\right)$$

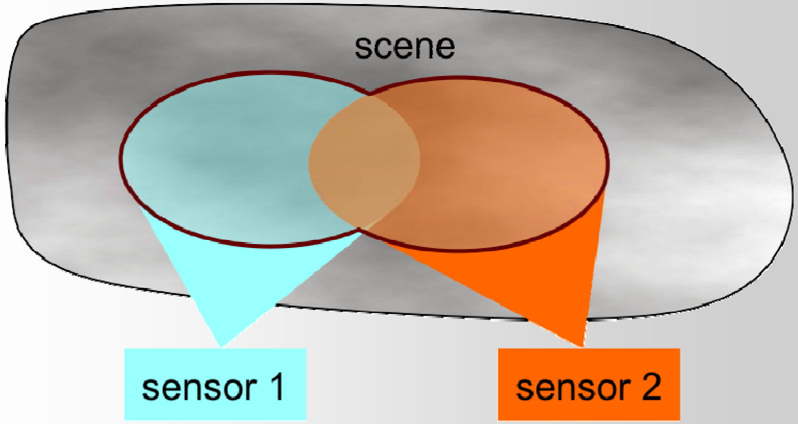
Saturation
(in HSV color space)

$$S(\boldsymbol{x}) = \frac{I - \min\limits_{p \in \{R,G,B\}} \left(d(\boldsymbol{x}, p)\right)}{I}$$

Fraunhofer Institut Informations- und Datenverarbeitung

Information content of image data

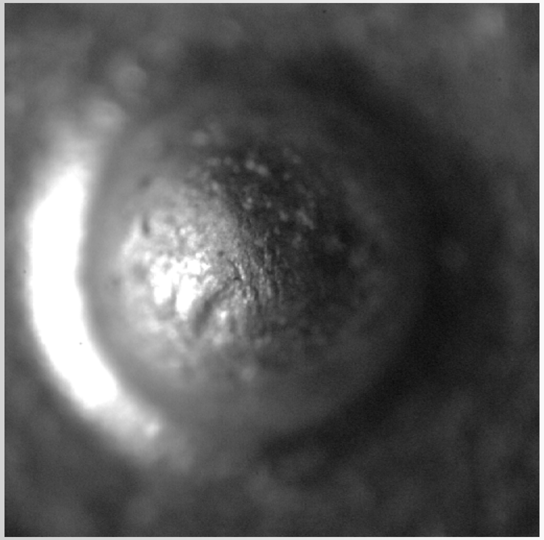
- **Distributed information**
  - Useful information is distributed over the whole image series
  - Only evaluation of all images allows statements on the desired properties
  - Complementary fusion of features extracted from the images
  - Examples:
    - Depth maps based on stereo series or focus series (*depth from focus*)
    - Photometric stereo



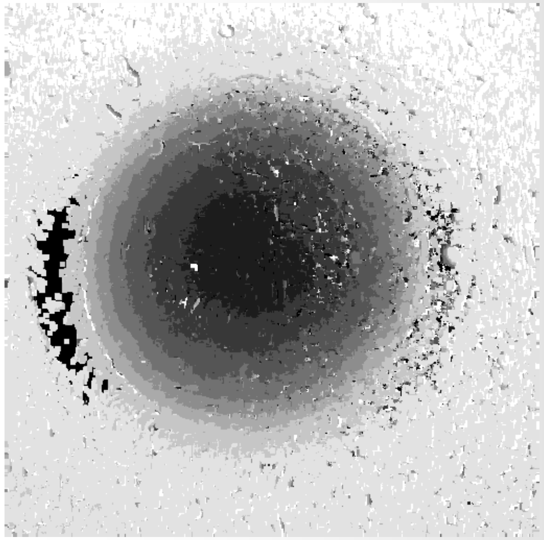
Distributed information

Example: Depth map from focus series

- Image number represents vertical position of focus
- For each image position: Transfer of the »number« of the image with maximum contrast into the fusion result



Focus series

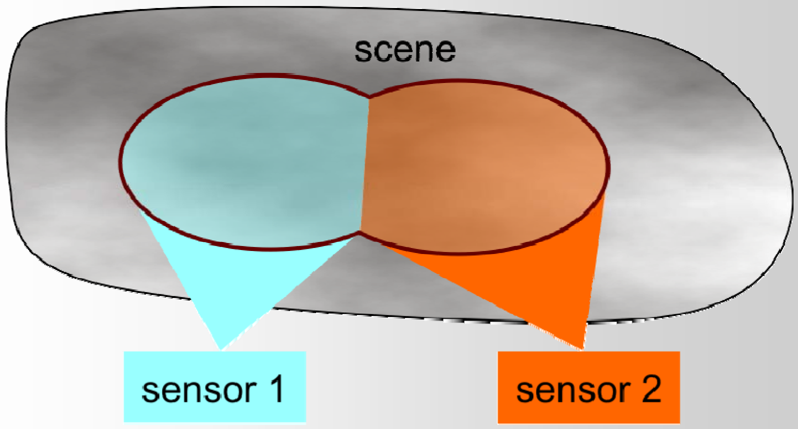


Depth map
dark: »deep«, light: »high«

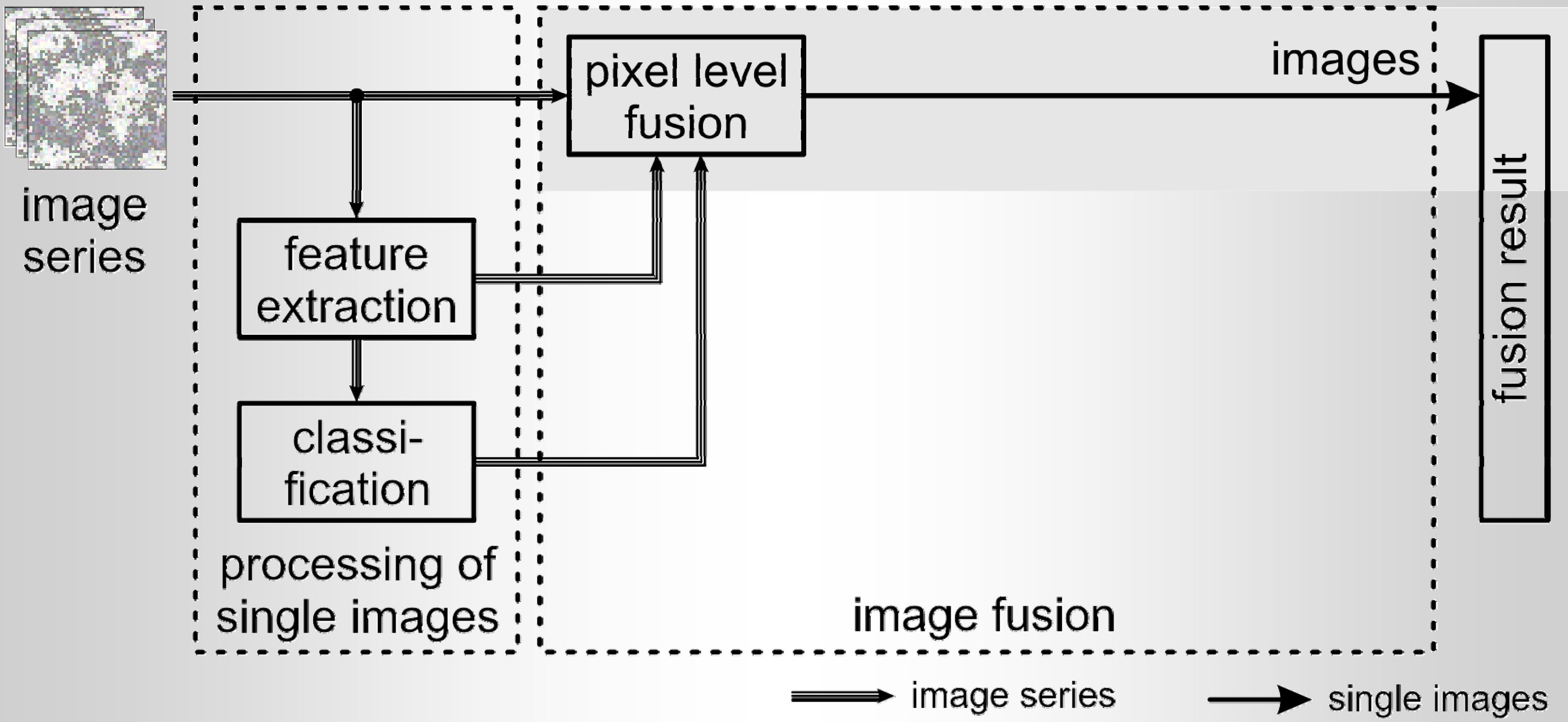
Fraunhofer Institut Informations- und Datenverarbeitung

Information content of image data

- **Orthogonal information**
  - Data used in fusion contain useful information on disjoint properties of the scene, e.g. different physical quantities
  - Combination of features or classification results (decisions)
  - Example:
    Combination of gray-value images (reflectance of the scene) and depth maps (3D-shape of the scene) for vehicle detection



Abstraction levels in image fusion

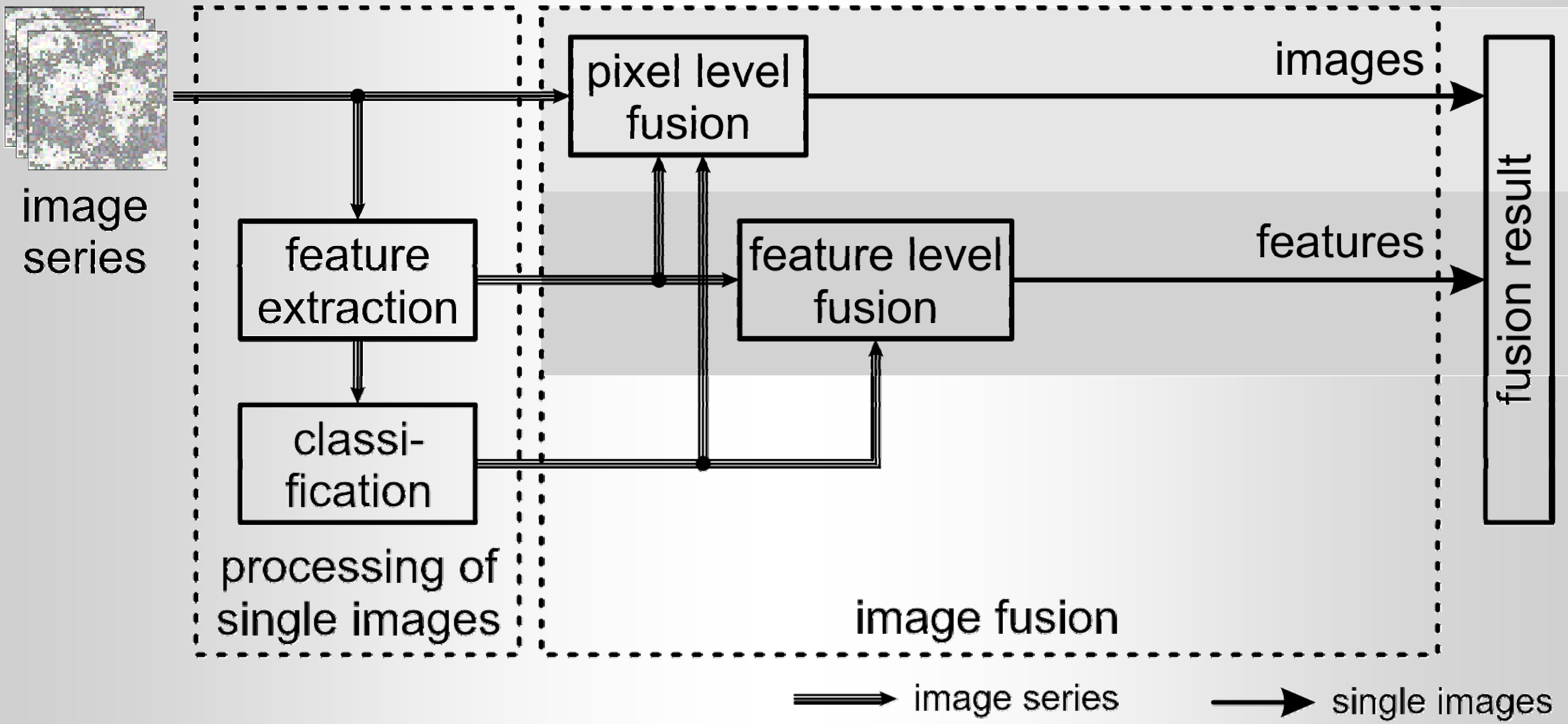


Fraunhofer Institut Informations- und Datenverarbeitung

Pixel-level fusion

- Combination on the **abstraction level of the image (pixel) data**
  Image intensities: Gray values or RGB-tripels
- Precondition: **Homogeneous sensors**
  Identical or comparable physical properties of the scene
- Objective: Fusion result with **better image quality** or **enhanced definition area**:
  - Concurrent fusion, e.g. suppression of sensor noise
  - Complementary fusion, e.g. enhanced depth of focus
- Information from **higher abstraction levels** may be necessary or helpful

Abstraction levels in image fusion



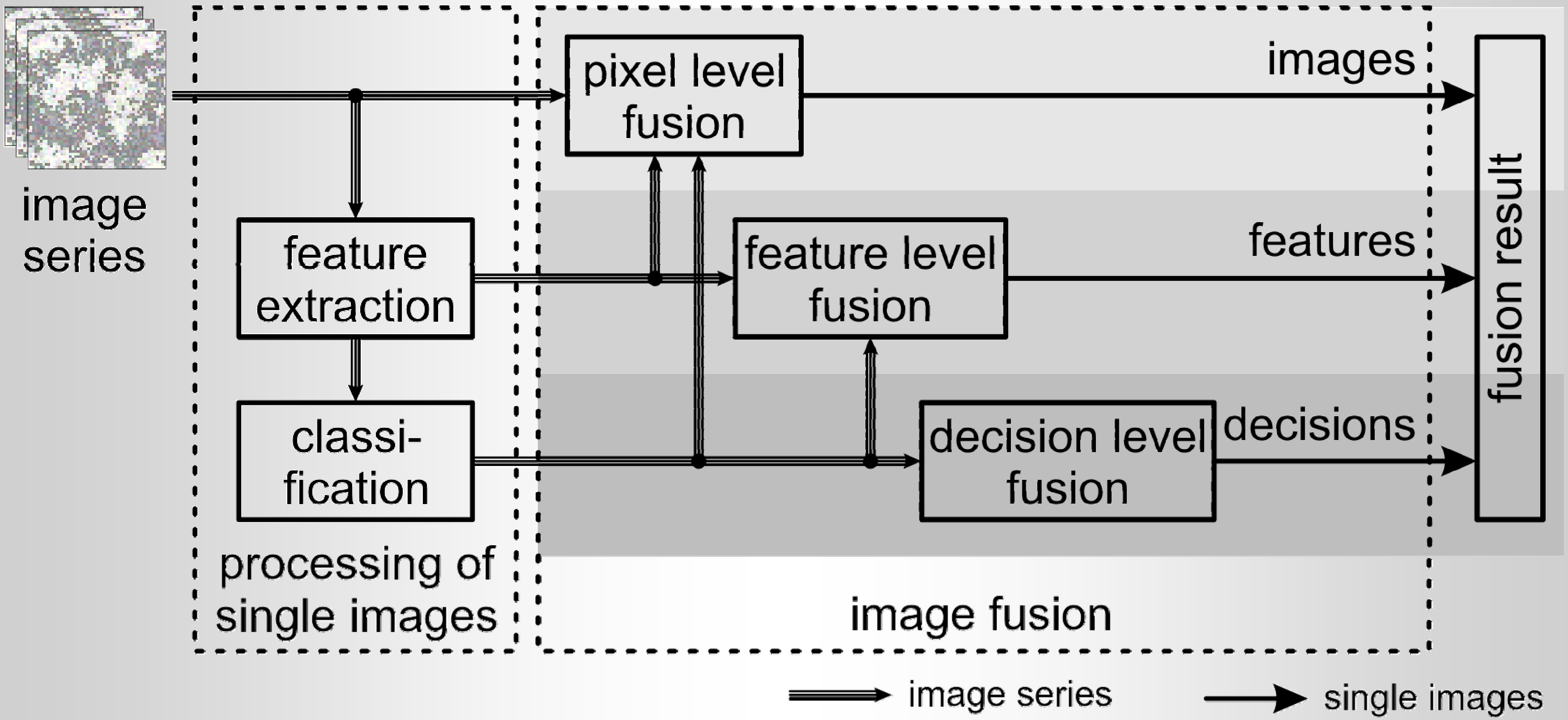
Fraunhofer Institut Informations- und Datenverarbeitung

Feature-level fusion

- Combination on the **abstraction level of features**
- Requires preceding **extraction of relevant features**:
  a) From single images (e.g. local texture descriptors)
  b) By simultaneous evaluation of several or all images of the series (e.g. mean local intensity within an illumination series)
- **Preconditions**:
  a) Sensors allowing the extraction of equal/comparable features
  b) Homogeneous sensors
- **Objectives**:
  - Improved extraction of image features (accuracy/reliability)
  - Access to features that are distributed over the series
- Integration of **information from higher abstraction levels**, if necessary or helpful

Fraunhofer Institut Informations- und Datenverarbeitung

Abstraction levels in image fusion



Fraunhofer Institut Informations- und Datenverarbeitung

Decision-level fusion

- Combination on the **abstraction level of decisions/symbols**
- Preceding **making of decisions required**:
  classification results obtained from single images
- **Precondition**:
  Sensors allowing the extraction of equal/comparable symbolic information
- **Objective**:
  Better reliability of classification,
  e.g. for defect detection of object recognition

Choice of the abstraction level

High abstraction level

+ Usually **relatively simple processing**,
since standard procedures for feature extraction and evaluation for single images are available

– Often **lower quality** of fusion result,
since potentially undesired reduction of information in the feature extraction and evaluation is necessary

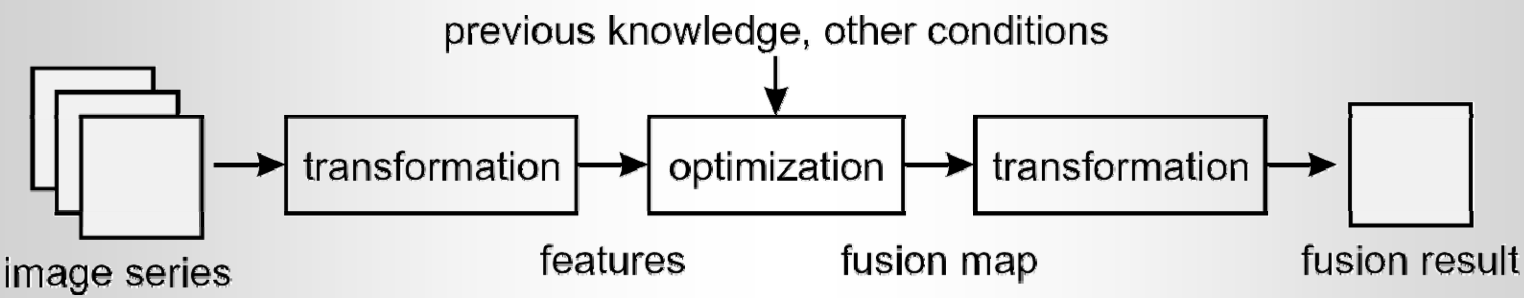
Low abstraction level

+ Frequently **higher quality of fusion result**,
since information content of single images can be entirely preserved until fusion

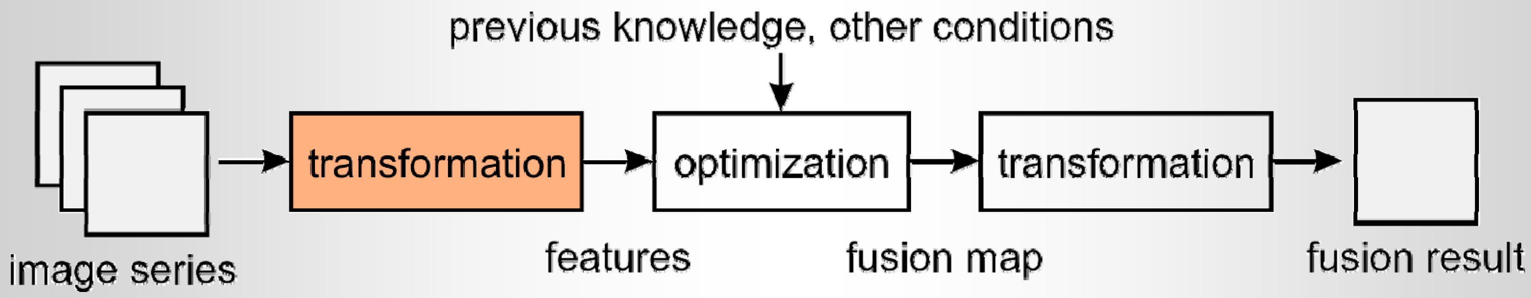
– Usually **appropriate processing methods** for image series must be individually developed

General structure of image fusion

- Many frameworks for image fusion show **comparable processing structure**
- Processing steps:
  - **Input transformation** of image series
  - **Optimization** with respect to a suitable representation of information
  - **Output transformation** to obtain result in usable form



General structure of image fusion



- Transformation of the image series into a **representation suitable for the extraction of the relevant information**
- Generation of **suitable descriptors** for the information content
- Generally:
  Transformations for image signals with **two or more dimensions**
  (e.g. local operators may not only refer to spatial dimensions)

Fraunhofer Institut Informations- und Datenverarbeitung

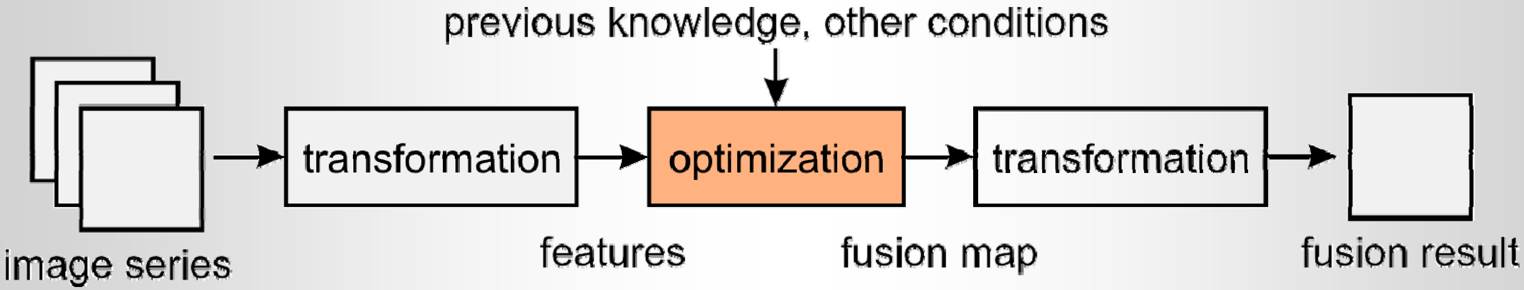
General structure of image fusion

Input transformation

- Examples for methods and fields of application:

Geometric transforms	Image alignment, correction of geom. distortions
Intensity transforms	Adaptation of sensor characteristics, illumination correction
Principal Component Analysis	Decorrelation of signal components, compression
Cross correlation	Detection of similarities or redundancies
Local operators	Determination of local image descriptors (e.g. local mean, contrast, edges, structure tensor, texture features)
Fourier transform	Harmonic analysis, global selection of spatial frequencies (for stationary processes)
Wavelet transform, Pyramid transform	Locally resolved spectrum analysis (for non-stationary processes)
Morphological transforms	Locally resolved analysis of textures/structures

General structure of image fusion



- **Selection of useful information** from the transformed image series
- Incorporation of **previous knowledge (e.g. experience, laws of nature)** with respect to the transformed image series and/or the fusion result
- Fusion in **appropriate domain**:
  E.g. spatial domain, domain of spatial frequencies, parameter domain, domain of parameter frequencies

Fraunhofer Institut Informations- und Datenverarbeitung

Optimization

- Examples for methods and fields of application:

Linear operators	Averaging used for concurrent fusion of image data or features (weighted averaging, if necessary)
Nonlinear operators	Maximization/minimization of feature values or quality measures, hierarchical operators for robust feature selection (e.g. median)
Energy minimization	Combination of several optimization objectives (e.g. maximization of a quality criterion, compliance with smoothness or other constraints)
Bayesian statistics	Interpretation of observations and previous knowledge as Degree-of-Belief
Neural Networks	Classification w/o explicit formulation of fusion instructions
Support-Vector-Machines	Classification in features spaces with high dimensionality
Kalman-Filtering	State estimation of dynamical systems

Optimization: Energy minimization

- Objective: **Universal approach** to formalize fusion tasks
- Introduction of »**energy terms**« E_k to model relevant information:
  - Available data (a priori knowledge)
  - Desired properties of the fusion result
  - Constraints referring to the data to be fused, to intermediate results or to fusion results
  - Mutual relations of the available data
- Modeling of the energy terms E_k such that the relevant information is reflected in **monotonous functions**:
  More desirable property $\leftrightarrow$ lower functional value
- Total energy: **Sum of energy terms**
- Fusion result: **Minimization of total energy** leads to desired result

Optimization: Energy minimization

- **Example**:

$$E(r) = E_d(d,r) + \lambda E_c(r), \quad \lambda > 0$$

$d(\boldsymbol{x},\boldsymbol{\omega})$: image data (observations)

: fusion result

- $E_d(.,.)$: Models the connection between given image data $d(\boldsymbol{x},\boldsymbol{\omega})$ and fusion result $r(\boldsymbol{x})$
- $E_c(.)$: Models desired or previously known properties of $r(\boldsymbol{x})$
- λ : Regularization parameter for weighting the energy components
- E.g. smoothing of a one-dimensional signal $d(x)$:

$$E(r(x)) = \underbrace{\int (r(x) - d(x))^2 \,\mathrm{d}\,x}_{E_d(d,r)} + \lambda \underbrace{\int \left(\frac{\mathrm{d}}{\mathrm{d}\,x} r(x) \right)^2 \mathrm{d}\,x}_{E_c(r)}$$

Fraunhofer Institut Informations- und Datenverarbeitung

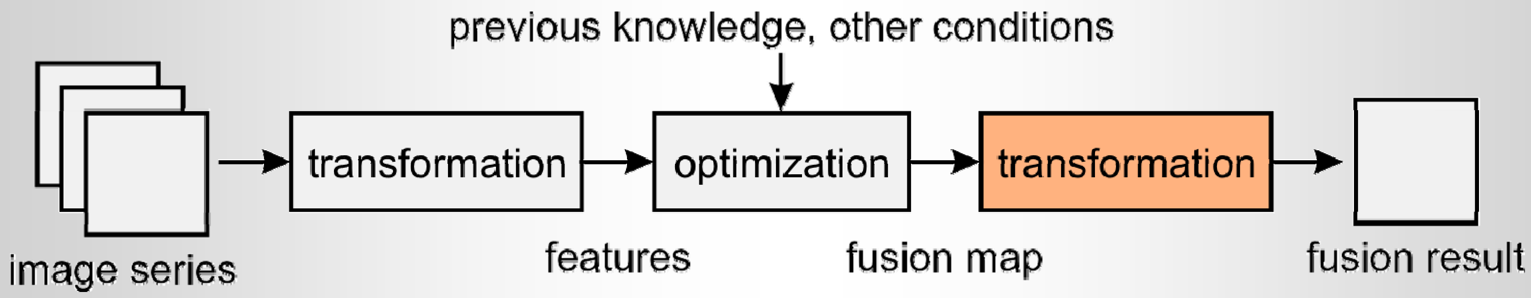
Optimization: Energy minimization

- Representation of the fusion task as **energy functional**:

$$E := \sum_k \lambda_k E_k, \quad \lambda_k > 0$$

- **Implicit and compact representation** of fusion task
- Additional information and constraints can be easily included by introduction of **suitable energy terms**
- Different relevance of data and information can be considered by suitable **coefficients** λ_k
- Fusion is performed by **simultaneous minimization** of energy functional with respect to fusion result (optimization problem)
- Problem: **No universally applicable optimization method** for minimization of total Energy E

General structure of image fusion



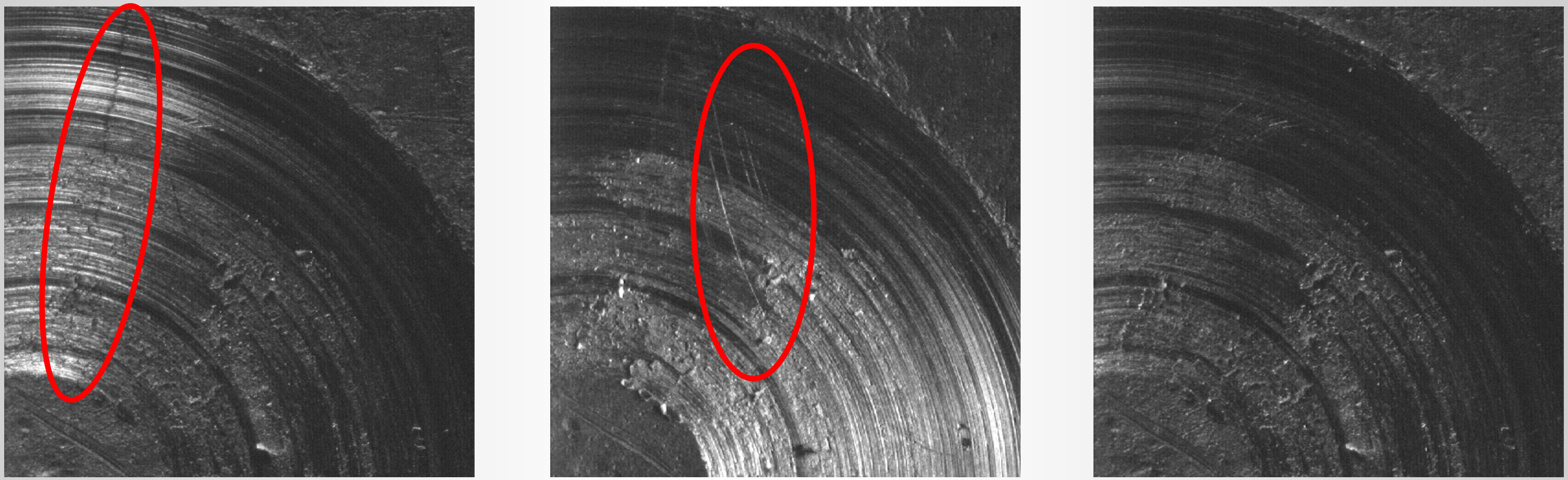
- Result of optimization: Fusion map represents »**blueprint**« for design of fusion result
- Final transformation of fusion map to obtain **fusion result on the desired abstraction level**
- Examples:

Direct adoption of the fusion map	Construction of depth maps from focus or stereo series
Look-up-table	Synthetically enhanced depth of focus from focus series

Fraunhofer Institut Informations- und Datenverarbeitung

Example: Generation of synthetical high-quality images

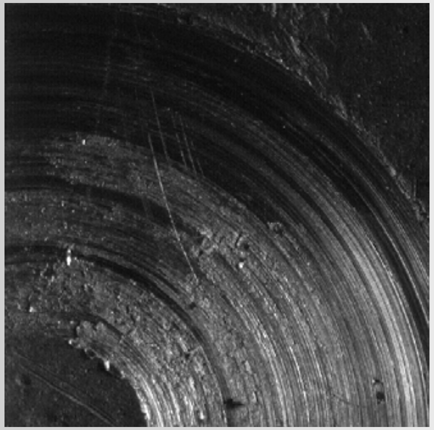
- Objective: **Reproducible acquisition of high-contrast images** for visual inspection of tool marks in forensic science
- Single images unsuitable: Locally insufficient contrast
- **Illumination series**: Variation of illumination direction, stationary camera
- Virtual, homogeneous, collocated sensors



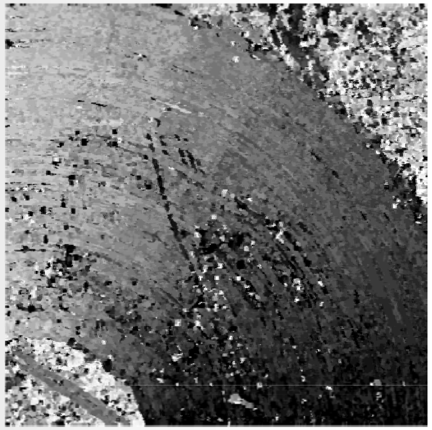
exemplary images from illumination series

Example: Generation of synthetical high-quality images

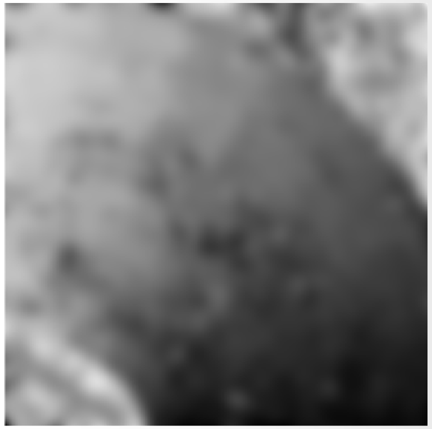
- Input transformation: Calculation of the **local contrast** in the single images
- Optimization: **Maximization of the local contrast** → fusion map
- Constraint: »Smoothness« of the fusion map → smoothed fusion map
- Output transformation: Smoothed fusion map as **blueprint** for fusion result: Patchwork from input image series



illumination series



fusion map



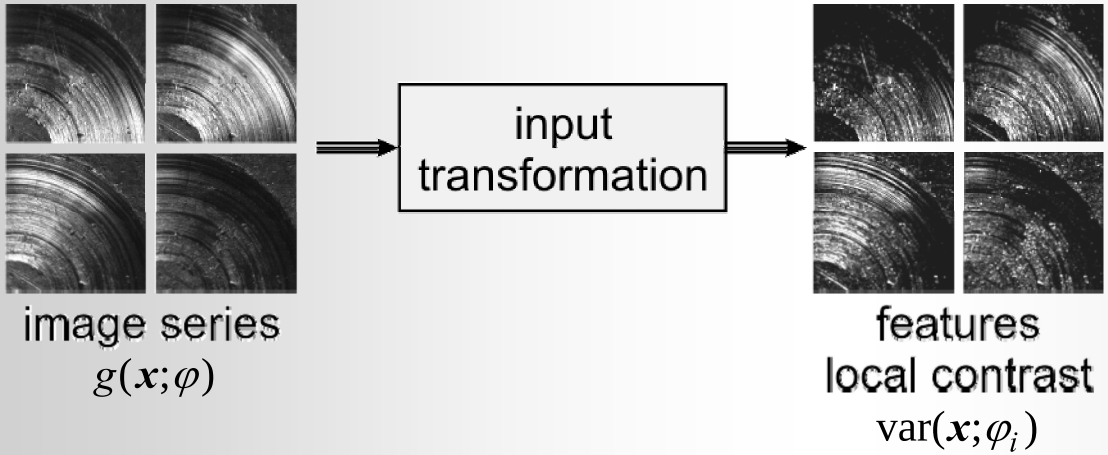
smoothed fusion map

fusion result

Example: Generation of synthetical high-quality images

- **Input transformation**: Transformation of the image series into a suitable representation
- Input data: Illumination series $g(\boldsymbol{x};\varphi)$
- Relevant information: **Local contrast** as quality measure
- Estimation of local contrast: **Local variance**

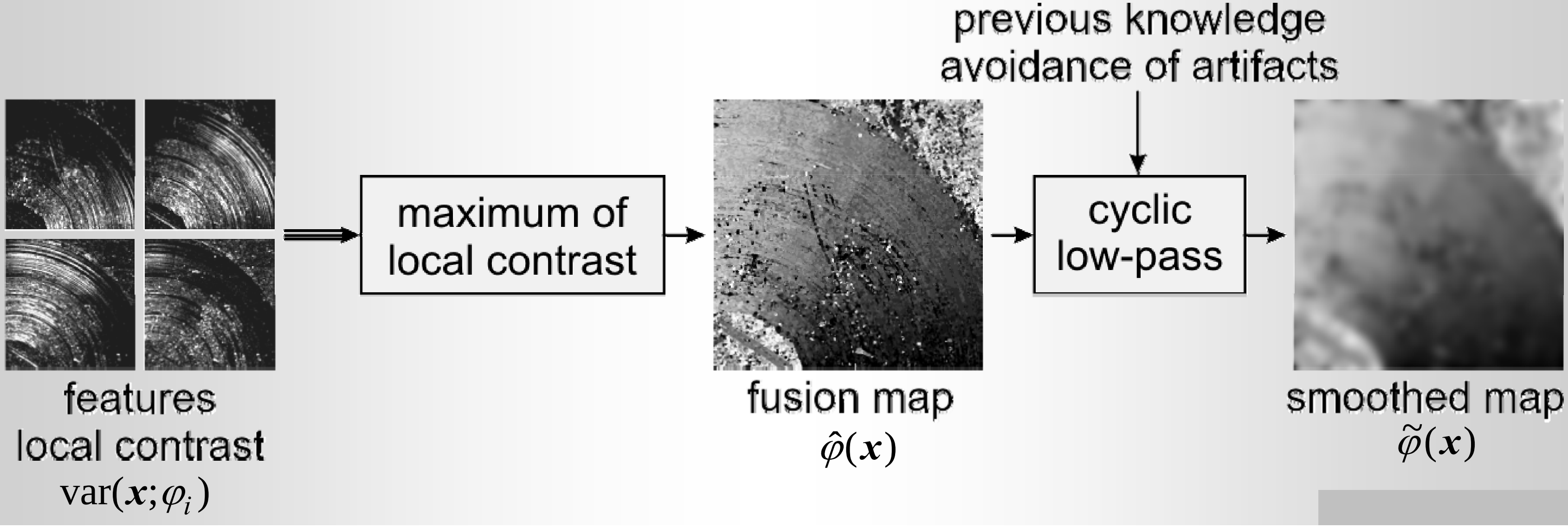
$$\operatorname{var}(\boldsymbol{x};\varphi_i) = \frac{1}{|U|}\sum_{U(\boldsymbol{x})}\big(g(\boldsymbol{x};\varphi_i) - m(\boldsymbol{x};\varphi_i)\big)^2, \quad m(\boldsymbol{x};\varphi_i) = \frac{1}{|U|}\sum_{U(\boldsymbol{x})} g(\boldsymbol{x};\varphi_i)$$



Fraunhofer Institut Informations- und Datenverarbeitung

Example: Generation of synthetical high-quality images

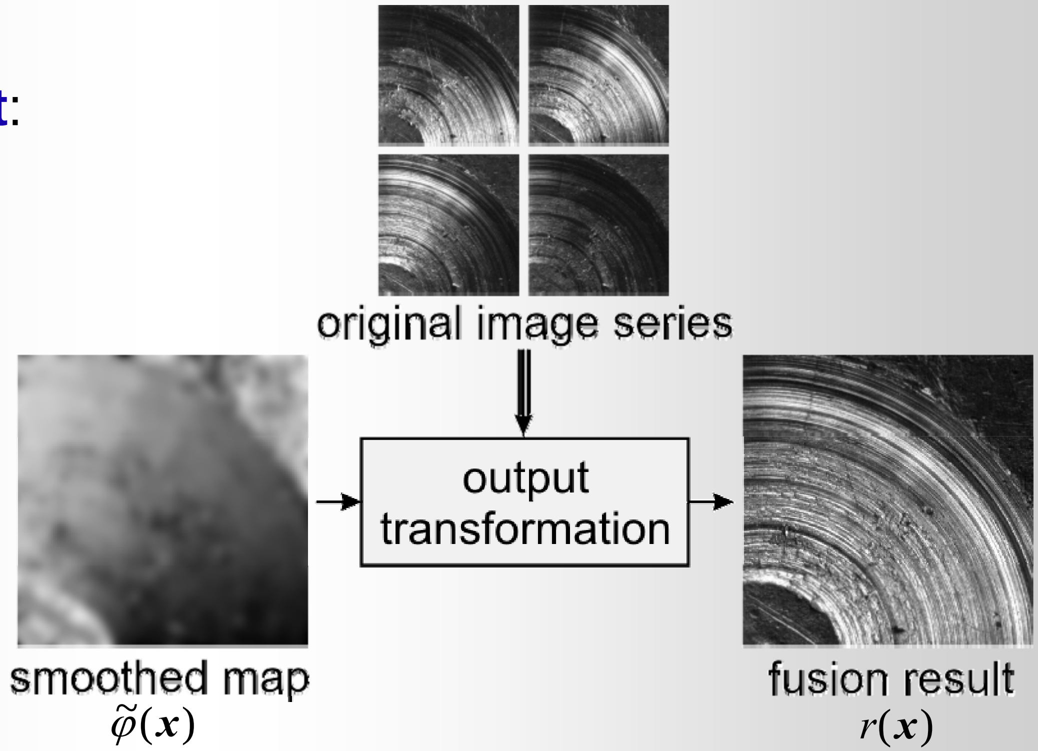
- **Optimization**: Selection of useful information
- Optimization domain: Parameter space of illumination azimuth
- Criteria:
  - Maximum local contrast for each location: $\hat{\varphi}(\boldsymbol{x}) = \arg\max_{\varphi_i}\{\mathrm{var}(\boldsymbol{x};\varphi_i)\}$
  - Smoothing of $\hat{\varphi}(\boldsymbol{x})$ to avoid artifacts (constraint)
    Cyclic low-pass: $\tilde{\varphi}(\boldsymbol{x}) = \angle \mathrm{LP}\{\exp(\mathrm{j}\hat{\varphi}(\boldsymbol{x}))\}$



Example: Generation of synthetical high-quality images

- **Output transformation**:
  Conversion into desired abstraction level: Image
- Smoothed map: **Look-up-table** to locally select optimal input image

- **Properties of fusion result**:
  - Result locally resembles best image of the series
  - Artifacts are avoided by means of smoothing



Fraunhofer Institut Informations- und Datenverarbeitung

Example: Generation of synthetical high-quality images

- Compact problem statement as **energy minimization task** (simplified):

$$E(r) = \lambda_d \overbrace{\underbrace{\sum_{\boldsymbol{x}} \sum_{\varphi_i} \delta_{\varphi_i}^{\tilde{\varphi}(\boldsymbol{x})} \left(r(\boldsymbol{x}) - d(\boldsymbol{x}, \varphi_i)\right)^2}}^{E_d(d,r,\tilde{\varphi})} + \lambda_s \overbrace{\underbrace{\sum_{\boldsymbol{x}} \left(\mathrm{HP}\{\tilde{\varphi}(\boldsymbol{x})\}\right)^2}}^{E_s(\tilde{\varphi})} + \lambda_c \overbrace{\underbrace{\left(-\sum_{\boldsymbol{x}} \left(\mathrm{K}\{r(\boldsymbol{x})\}\right)^2\right)}}^{E_c(r)}$$

- connection of image data and fusion result
- squared intensity distance to the »nearest« image in the series (with respect to $\tilde{\varphi}(\boldsymbol{x})$)

- smoothness constraint
- applied to fusion map

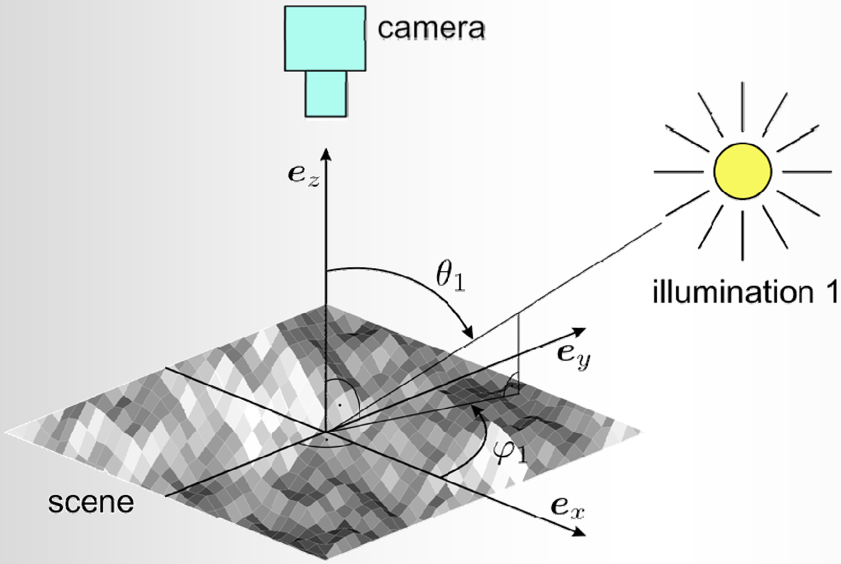
- objective: high contrast
- contrast measure $\mathrm{K}(.)$, e.g. local variance

- Solution of fusion task by **successive minimization**:
  1. Minimization of $E_c(r)$: $\hat{\varphi}(\boldsymbol{x}) = \arg\max_{\varphi_i} \{\mathrm{var}(\boldsymbol{x};\varphi_i)\}$
  2. Minimization of $E_s(\tilde{\varphi})$: $\tilde{\varphi}(\boldsymbol{x}) = \angle\mathrm{LP}\{\exp(\mathrm{j}\hat{\varphi}(\boldsymbol{x}))\}$
  3. Fusion by applying smoothed fusion map $\tilde{\varphi}(\boldsymbol{x})$

Fraunhofer Institut Informations- und Datenverarbeitung

Example: Photometric stereo

- Reconstruction of **2½D shape and reflectance** of a surface
- Illumination series: Variation of illumination direction, stationary camera (virtual, homogeneous, collocated sensors)
- Optimization: **Modeling of the surface reflection**
- Previous knowledge: Reflection properties of the surface (e.g. material property)



Fraunhofer Institut Informations- und Datenverarbeitung

Example: Photometric stereo

- Evaluation by modeling the observed intensity $g(\boldsymbol{x})$
  e.g. with **assumption of ideal diffuse reflection**: $g(\boldsymbol{x}) = \rho(\boldsymbol{x}) \cdot \boldsymbol{b}_{\mathrm{e}}{}^{\mathrm{T}} \boldsymbol{n}_{\mathrm{e}}(\boldsymbol{x})$
- Acquisition of illumination series:

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{pmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_n(\boldsymbol{x}) \end{pmatrix} = \rho(\boldsymbol{x}) \cdot \boldsymbol{B}^{\mathrm{T}} \boldsymbol{n}_{\mathrm{e}}(\boldsymbol{x})$$

$\boldsymbol{g}(\boldsymbol{x}) \in \mathbb{R}^n$: observation vector

$0 \leq \rho(\boldsymbol{x}) \leq 1$: diffuse reflectance

$\boldsymbol{n}_{\mathrm{e}}(\boldsymbol{x}) \in \mathbb{R}^3$: normal unit vector

$\boldsymbol{b}_{\mathrm{e}} = (\cos\varphi \sin\theta, \sin\varphi \sin\theta, \cos\theta)^{\mathrm{T}} \in \mathbb{R}^3$: illumination direction

$\boldsymbol{B} = \left(\boldsymbol{b}_{\mathrm{e},1}, \ldots, \boldsymbol{b}_{\mathrm{e},n}\right) \in \mathbb{R}^{3 \times n}$: illumination matrix

Example: Photometric stereo

- For three illumination directions $n = 3$:
  Direct inversion of observation equation $\boldsymbol{g}(\boldsymbol{x}) = \rho(\boldsymbol{x}) \cdot \boldsymbol{B}^{\mathrm{T}} \boldsymbol{n}_{\mathrm{e}}(\boldsymbol{x})$

$$\boldsymbol{n}(\boldsymbol{x}) := \rho(\boldsymbol{x}) \cdot \boldsymbol{n}_{\mathrm{e}}(\boldsymbol{x}) = \left(\boldsymbol{B}^{\mathrm{T}}\right)^{-1} \boldsymbol{g}(\boldsymbol{x})$$

$$\rho(\boldsymbol{x}) = \left\|\boldsymbol{n}(\boldsymbol{x})\right\|, \quad \boldsymbol{n}_{\mathrm{e}}(\boldsymbol{x}) = \frac{\boldsymbol{n}(\boldsymbol{x})}{\rho(\boldsymbol{x})}$$

$\boldsymbol{B}^{\mathrm{T}}$ invertible
if illumination directions
not coplanar

Length of $\boldsymbol{n}(\boldsymbol{x})$: Reflectance
Direction of $\boldsymbol{n}(\boldsymbol{x})$: Normal unit vector

- For more than three illumination directions:

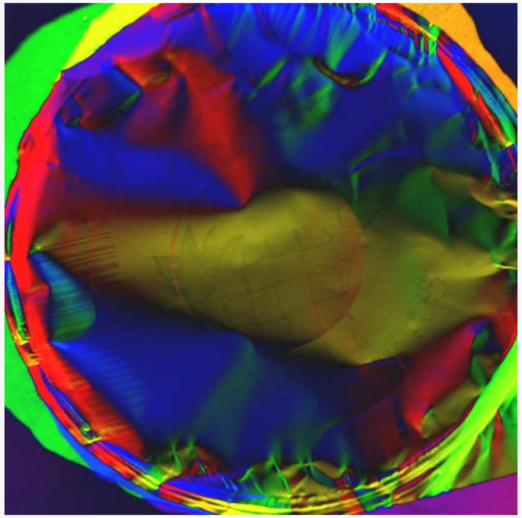
$$\boldsymbol{n}(\boldsymbol{x}) := \rho(\boldsymbol{x}) \cdot \boldsymbol{n}_{\mathrm{e}}(\boldsymbol{x}) = \left(\boldsymbol{B}\,\boldsymbol{B}^{\mathrm{T}}\right)^{-1} \boldsymbol{B}\,\boldsymbol{g}(\boldsymbol{x})$$

$\rightarrow$ Least-squares estimation

Fraunhofer Institut Informations- und Datenverarbeitung

Example: Photometric stereo

illumination series



normal vector $\boldsymbol{n}_\mathrm{e}(\boldsymbol{x})$
(color: azimuth,
intensity: polar angle)

reflectance
$\rho(\boldsymbol{x})$

- Result: Satisfactory **separation of shape and reflectance**
- Obvious deviations of reflection modeling (specular reflections): Still acceptable results

Summary

- Image fusion: Generation of a **result which describes the scene »better«** than any image obtained in a single shot
- Interpretation of image series as **multidimensional data objects**
- Evaluation of redundant, complementary, distributed or orthogonal **useful information**
- Classification of fusion architectures according to the **abstraction level**
- **General fusion structure**:
  Input transformation, optimization, output transformation

Thank you for your attention!

Literature

- J. J. Clark, A. L. Yuille: *Data Fusion for Sensory Information Processing Systems*. Kluwer Academic Publishers, 1990.
- F. Puente León: *Enhanced imaging by fusion of illumination series*. In: Otmar Loffeld (ed.), Sensors, Sensor Systems, and Sensor Data Processing, Proc SPIE, Vol. 3100, pp. 297-308, 1997.
- M. Heizmann: *Techniques for the Segmentation of Striation Patterns*. IEEE Transactions on Image Processing 15 (3), pp. 624-631, 2006.
- M. Heizmann: *Fusion von Bildserien in der automatischen Sichtprüfung*. In: Informatik 2005, A. B. Cremers, R. Manthey, P. Martini, V. Steinhage (eds.), Vol. 2: Proceedings, pp. 322-326, Gesellschaft für Informatik (GI), Bonn, 2005. (in German)