Image Fusion: Principles, Methods, and Applications

Tutorial EUSIPCO 2007

Lecture Notes

Jan Flusser, Filip Šroubek, and Barbara Zitová

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic
E-mail: {flusser,sroubekf,zitova}@utia.cas.cz

Introduction

The term *fusion* means in general an approach to extraction of information acquired in several domains. The goal of *image fusion* (IF) is to integrate complementary multisensor, multitemporal and/or multiview information into one new image containing information the quality of which cannot be achieved otherwise. The term quality, its meaning and measurement depend on the particular application.

Image fusion has been used in many application areas. In remote sensing and in astronomy, multisensor fusion is used to achieve high spatial and spectral resolutions by combining images from two sensors, one of which has high spatial resolution and the other one high spectral resolution. Numerous fusion applications have appeared in medical imaging like simultaneous evaluation of CT, MRI, and/or PET images. Plenty of applications which use multisensor fusion of visible and infrared images have appeared in military, security, and surveillance areas. In the case of multiview fusion, a set of images of the same scene taken by the same sensor but from different viewpoints is fused to obtain an image with higher resolution than the sensor normally provides or to recover the 3D representation of the scene. The multitemporal approach recognizes two different aims. Images of the same scene are acquired at different times either to find and evaluate changes in the scene or to obtain a less degraded image of the scene. The former aim is common in medical imaging, especially in change detection of organs and tumors, and in remote sensing for monitoring land or forest exploitation. The acquisition period is usually months or years. The latter aim requires the different measurements to be much closer to each other, typically in the scale of seconds, and possibly under different conditions.

The list of applications mentioned above illustrates the diversity of problems we face when fusing images. It is impossible to design a universal method applicable to all image fusion tasks. Every method should take into account not only the fusion purpose and the characteristics of individual sensors, but also particular imaging conditions, imaging geometry, noise corruption, required accuracy and application-dependent data properties.

Tutorial structure

In this tutorial we categorize the IF methods according to the data entering the fusion and according to the fusion purpose. We distinguish the following categories.

- *Multiview fusion* of images from the same modality and taken at the same time but from different viewpoints.
- *Multimodal fusion* of images coming from different sensors (visible and infrared, CT and NMR, or panchromatic and multispectral satellite images).
- *Multitemporal fusion* of images taken at different times in order to detect changes between them or to synthesize realistic images of objects which were not photographed in a desired time.
- *Multifocus fusion* of images of a 3D scene taken repeatedly with various focal length.
- *Fusion for image restoration*. Fusion two or more images of the same scene and modality, each of them blurred and noisy, may lead to a deblurred and denoised image. Multichannel deconvolution is a typical representative of this category. This approach can be extended to superresolution fusion, where input blurred images of low spatial resolution are fused to provide us a high-resolution image.

In each category, the fusion consists of two basic stages: image registration, which brings the input images to spatial alignment, and combining the image functions (intensities, colors, etc) in the area of frame overlap. Image registration works usually in four steps.

- *Feature detection.* Salient and distinctive objects (corners, line intersections, edges, contours, closed-boundary regions, etc.) are manually or, preferably, automatically detected. For further processing, these features can be represented by their point representatives (distinctive points, line endings, centers of gravity), called in the literature *control points*.
- *Feature matching.* In this step, the correspondence between the features detected in the sensed image and those detected in the reference image is established. Various feature descriptors and similarity measures along with spatial relationships among the features are used for that purpose.

- *Transform model estimation.* The type and parameters of the so-called *mapping functions*, aligning the sensed image with the reference image, are estimated. The parameters of the mapping functions are computed by means of the established feature correspondence.
- *Image resampling and transformation.* The sensed image is transformed by means of the mapping functions. Image values in non-integer coordinates are estimated by an appropriate interpolation technique.

We present a survey of traditional and up-to-date registration and fusion methods and demonstrate their performance by practical experiments from various application areas.

Special attention is paid to fusion for image restoration, because this group is extremely important for producers and users of low-resolution imaging devices such as mobile phones, camcorders, web cameras, and security and surveillance cameras.

Supplementary reading

Šroubek F., Flusser J., and Cristobal G., ”Multiframe Blind Deconvolution Coupled with Frame Registration and Resolution Enhancement”, in: *Blind Image Deconvolution: Theory and Applications*, Campisi P. and Egiazarian K. eds., CRC Press, 2007.

Šroubek F., Flusser J., and Zitová B., ”Image Fusion: A Powerful Tool for Object Identification”, in: *Imaging for Detection and Identification*, (Byrnes J. ed.), pp. 107-128, Springer, 2006

Šroubek F. and Flusser J., ”Fusion of Blurred Images”, in: *Multi-Sensor Image Fusion and Its Applications*, Blum R. and Liu Z. eds., CRC Press, Signal Processing and Communications Series, vol. 25, pp. 423-449, 2005

Zitová B. and Flusser J., ”Image Registration Methods: A Survey”, *Image and Vision Computing*, vol. 21, pp. 977-1000, 2003,

Handouts

Image Fusion

Principles, Methods, and Applications

Jan Flusser, Filip Šroubek, and Barbara Zitová

Institute of Information Theory and Automation
Prague, Czech Republic

Empirical observation

- One image is not enough
- We need
  - more images
  - the techniques how to combine them

Image Fusion

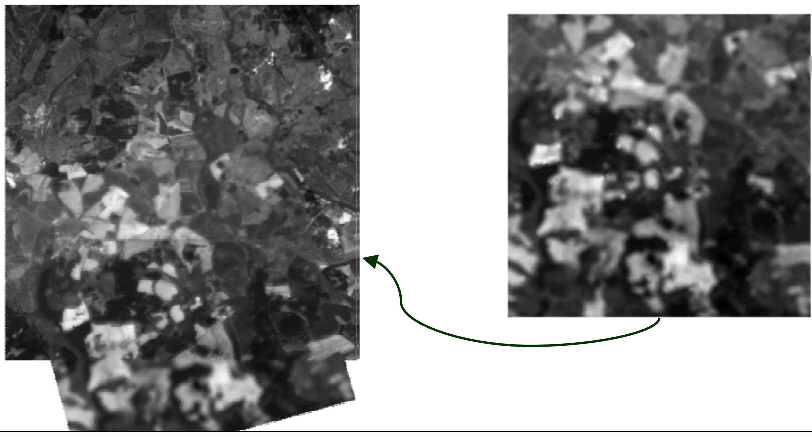
Input: Several images of the same scene

Output: One image of higher quality

The definition of "quality" depends on the particular application area

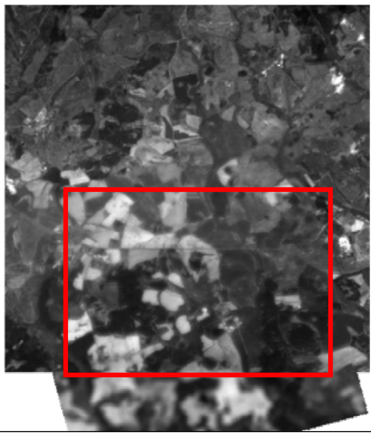
Basic fusion strategy

- Acquisition of different images
- Image-to-image registration



Basic fusion strategy

- Acquisition of different images
- Image-to-image registration
- The fusion itself (combining the images together)



The outline of the talk

- Fusion categories and methods (J. Flusser)
- Fusion for image restoration (F. Šroubek)
- Image registration methods (B. Zitová)

Fusion categories

- **Multiview fusion**
- **Multimodal fusion**
- **Multitemporal fusion**
- **Multifocus fusion**
- **Fusion for image restoration**

Multiview Fusion

- Images of the same modality, taken at the same time but from different places or under different conditions
- Goal: to supply complementary information from different views

Multiview fusion

Foreground

Background

Reprinted from R. Redondo et al.

Fusion categories

- Multiview fusion
- Multimodal fusion
- Multitemporal fusion
- Multifocus fusion
- Fusion for image restoration

Multimodal Fusion

- Images of different modalities: PET, CT, MRI, visible, infrared, ultraviolet, etc.

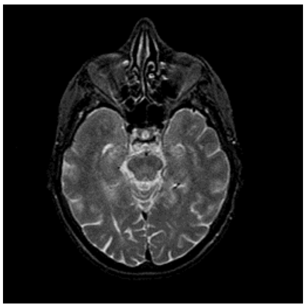
- Goal: to decrease the amount of data, to emphasize band-specific information

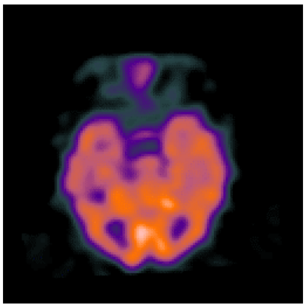
Multimodal Fusion

Common methods

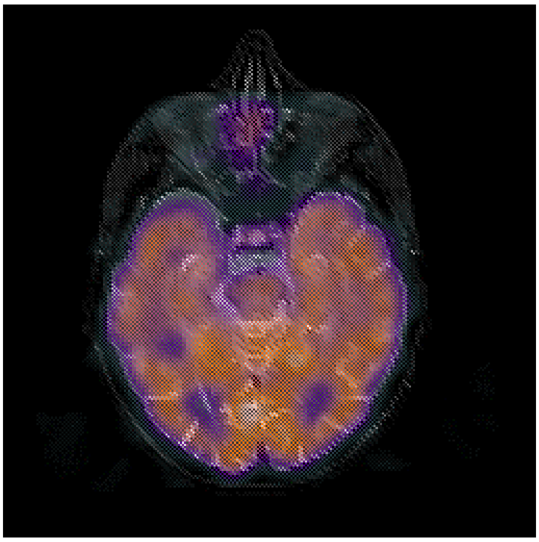
- Weighted averaging pixel-wise
- Fusion in transform domains
- Object-level fusion

Medical imaging – pixel averaging



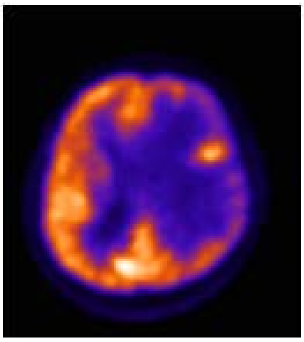


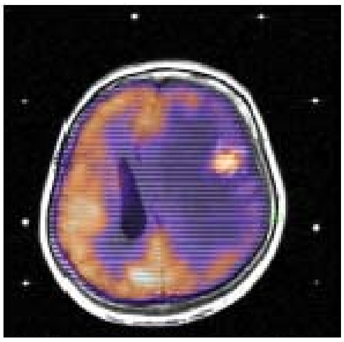
NMR + SPECT



Medical imaging – pixel averaging

PET + NMR





Visible + infrared

different modalities

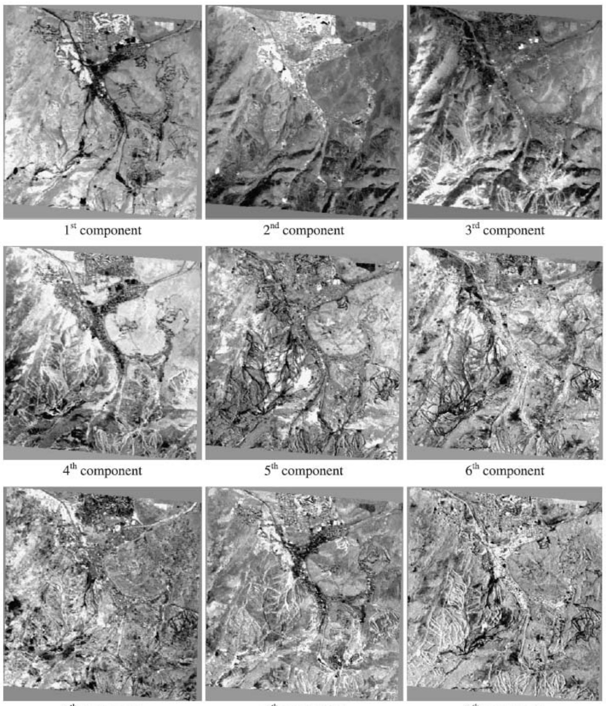
VS

IR

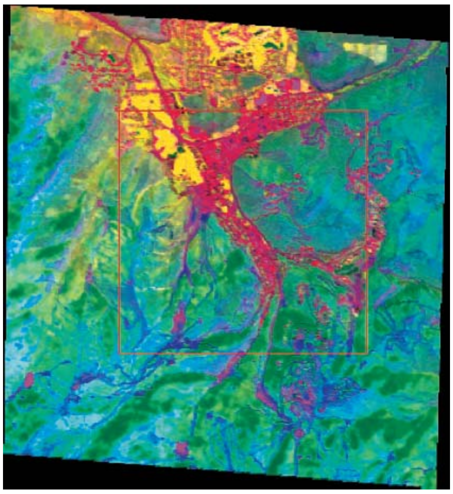
weighted average

Reprinted from R.Blum et al.

Multispectral data – fusion by PCA



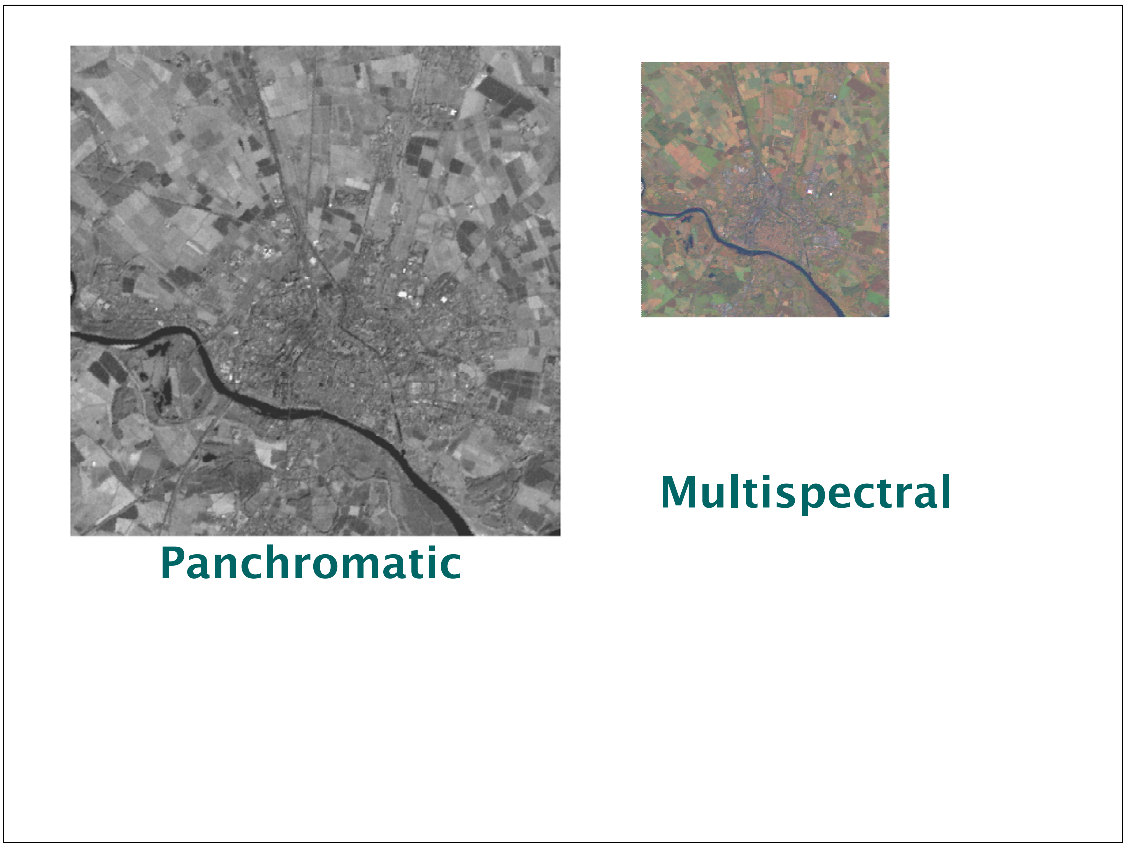
Fused image in pseudocolors

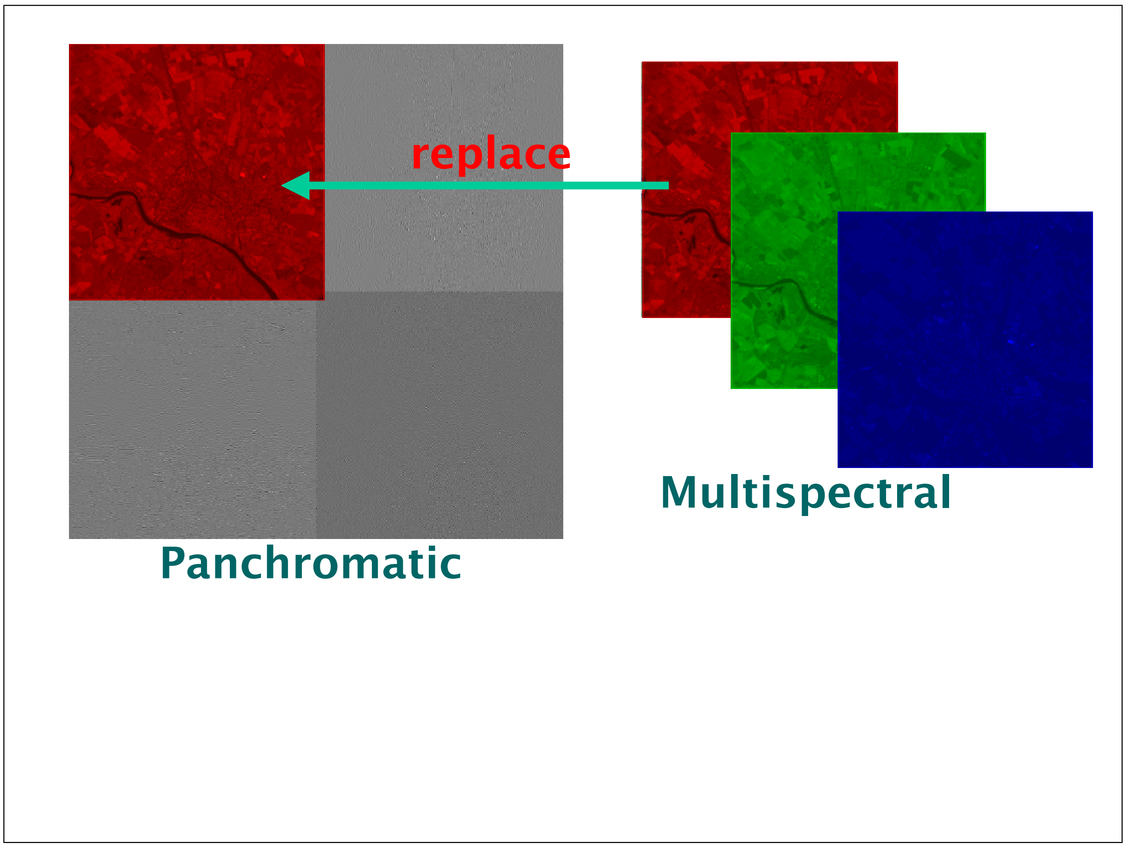


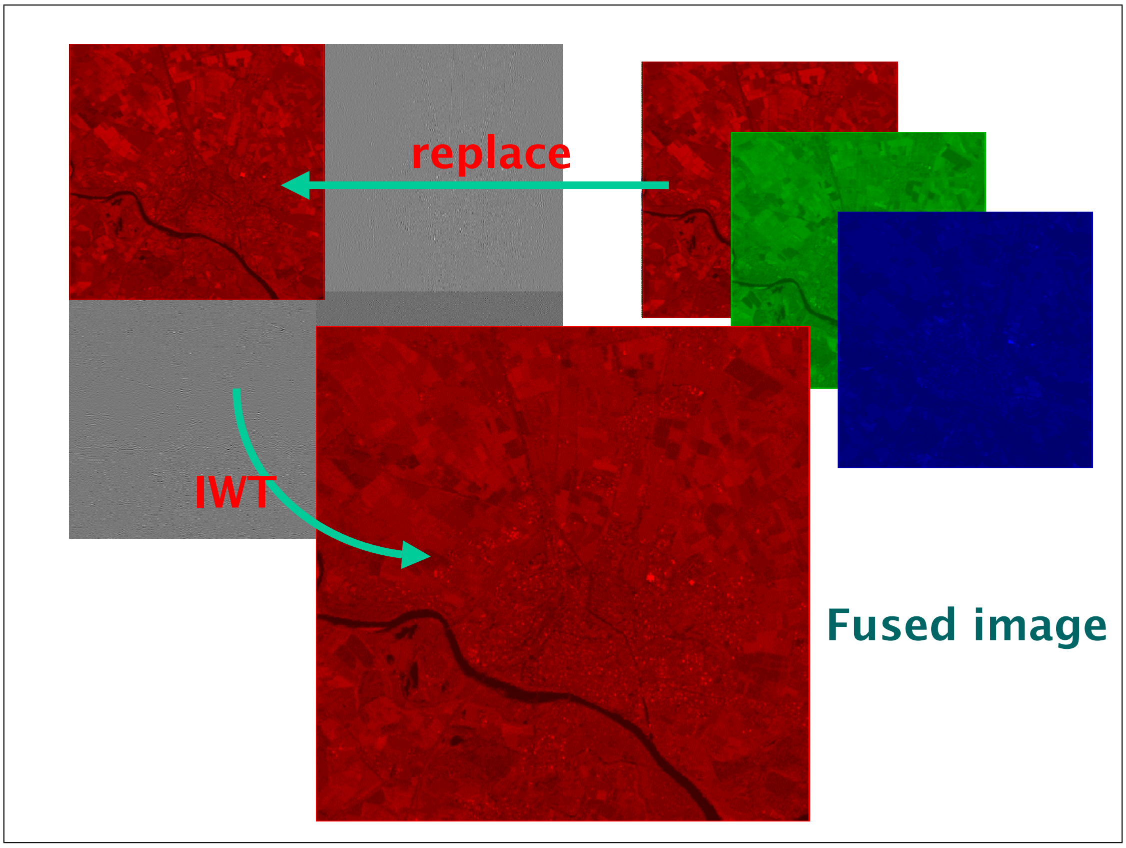
RGB = first 3 components

Multimodal fusion with different resolution

- One image with high spatial resolution, the other one with low spatial but higher spectral resolution.
- Goal: An image with high spatial and spectral resolution
- Method: Replacing bands in DWT

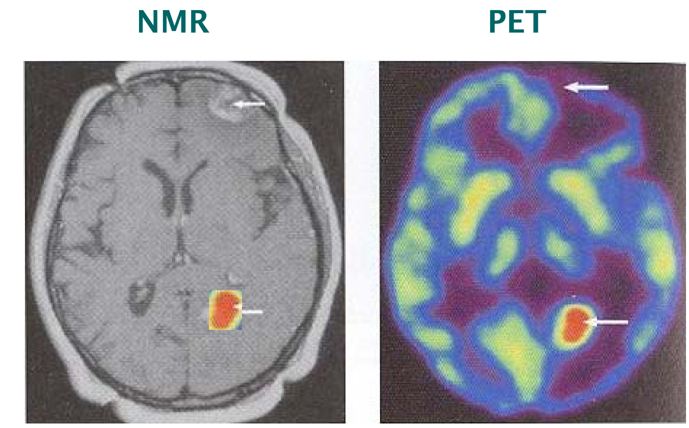

Multispectral
Panchromatic


replace
Multispectral
Panchromatic



FUSED MS + OPT

Challenge for the future: Object-level fusion



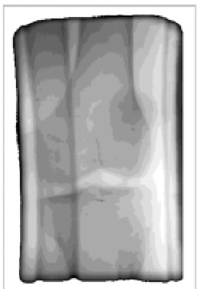
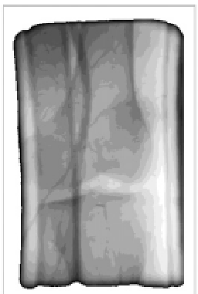
Fusion categories

- Multiview fusion
- Multimodal fusion
- Multitemporal fusion
- Multifocus fusion
- Fusion for image restoration

Multitemporal Fusion

- Images of the same scene taken at different times (usually of the same modality)
- Goal: Detection of changes
- Method: Subtraction

Digital subtraction angiography

A  B 

F 

Reprinted from Y. Bentoutou et al.

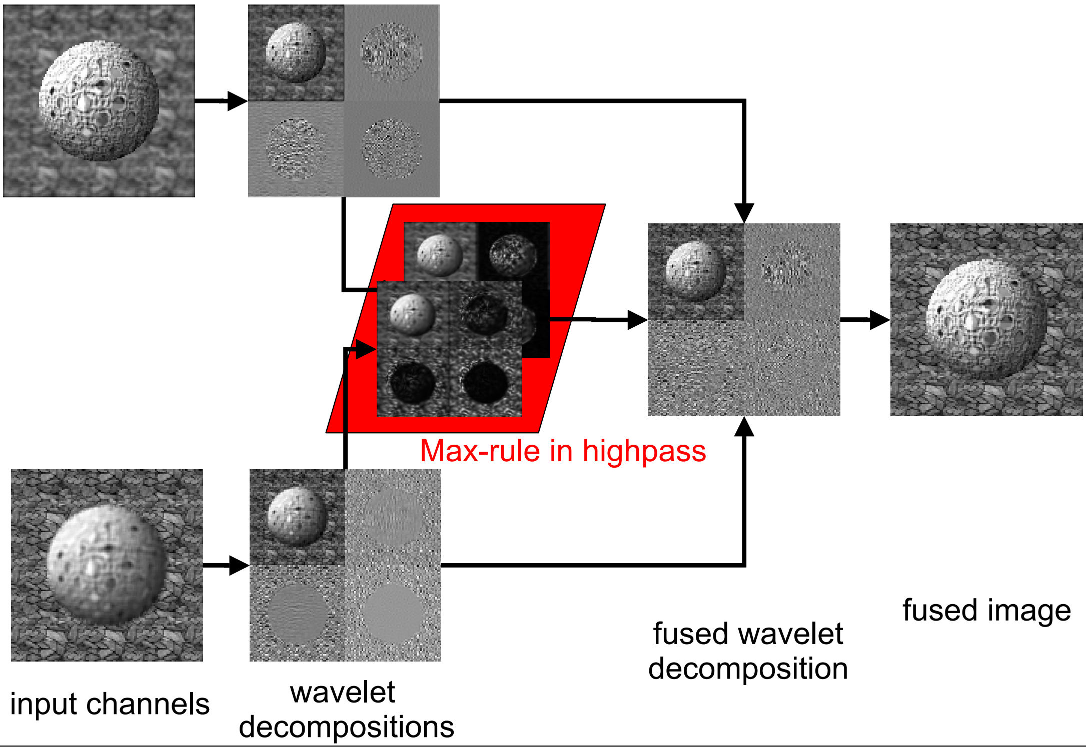
Fusion categories

- **Multiview fusion**
- **Multimodal fusion**
- **Multitemporal fusion**
- **Multifocus fusion**
- **Fusion for image restoration**

Multifocus fusion

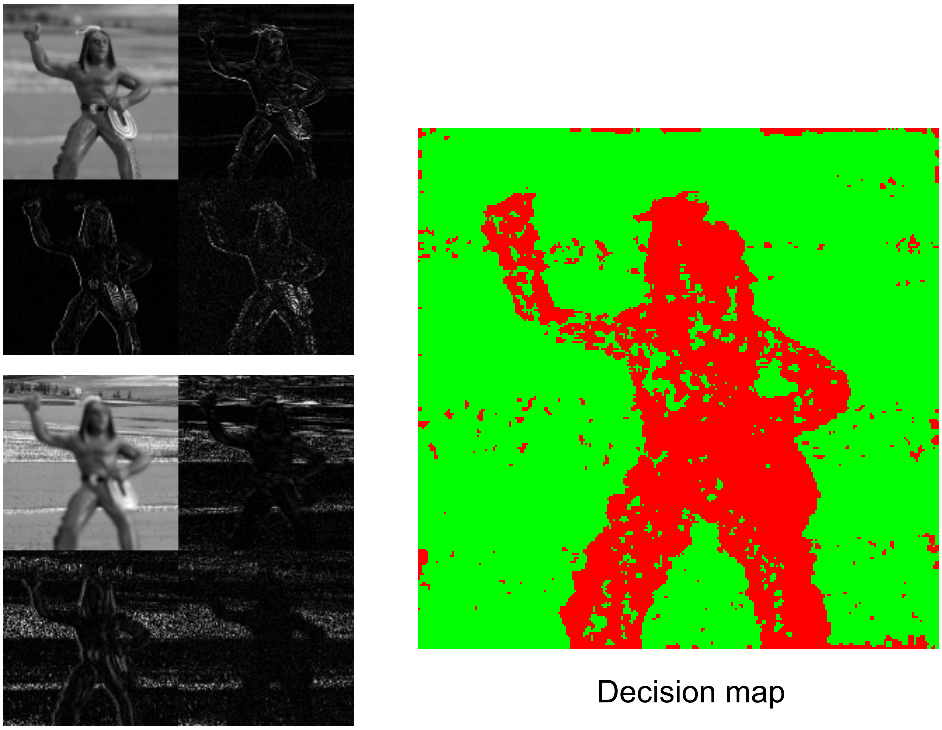
- The original image can be divided into regions such that every region is in focus in at least one channel
- Goal: Image everywhere in focus
- Method: identify the regions in focus and combine them together

Multifocus fusion in wavelet domain

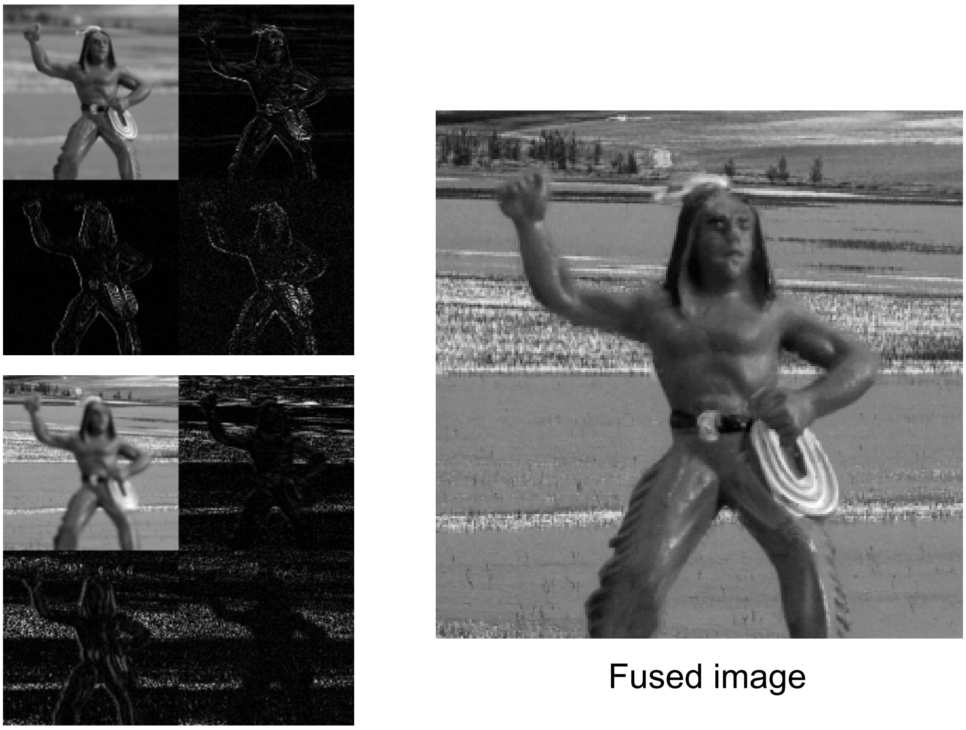


Artificial example

Images with different areas in focus



Decision map



Fused image

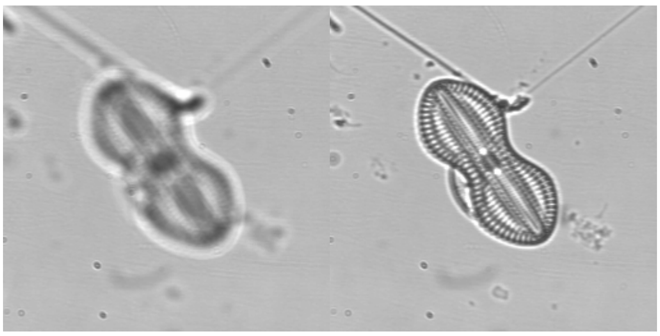
Regularized Decision Map

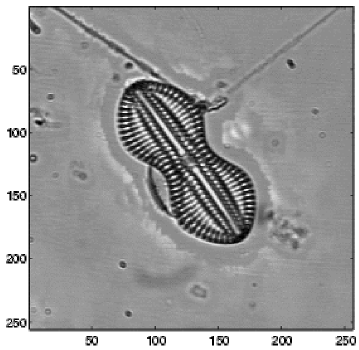
max rule

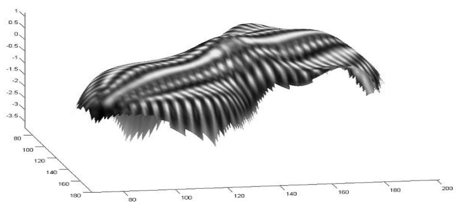
max rule

with regularization

Microscopic images: fusion and 3D reconstruction



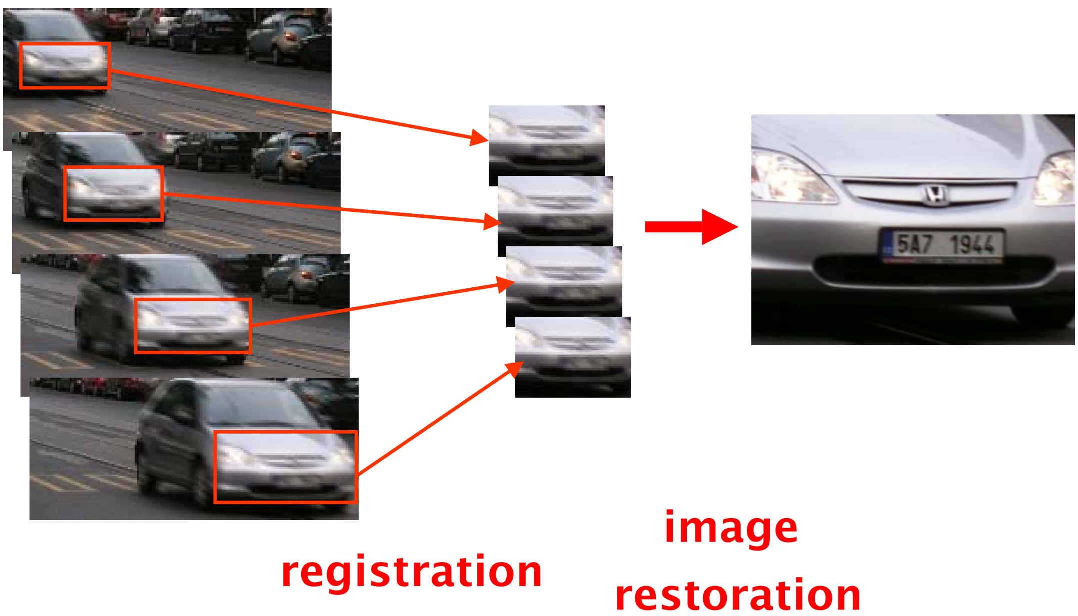




Fusion categories

- Multiview fusion
- Multimodal fusion
- Multitemporal fusion
- Multifocus fusion
- Fusion for image restoration

Realistic imaging



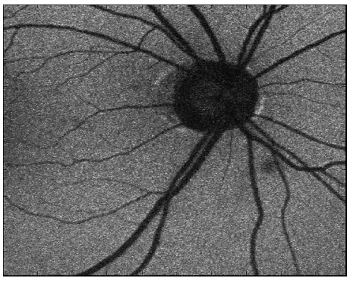
Fusion for image restoration

- Idea: Each image consists of “true” part and “degradation”, which can be removed by fusion
- Types of degradation:
  - additive noise: image denoising
  - convolution: blind deconvolution
  - resolution decimation: superresolution

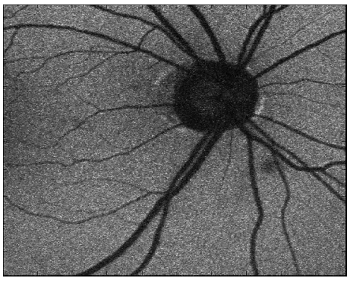
Denoising

- averaging over multiple realizations (averaging in time)

Denoising via time averaging

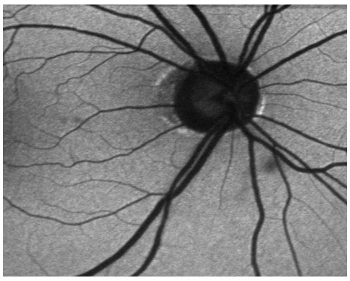


Before registration



After registration

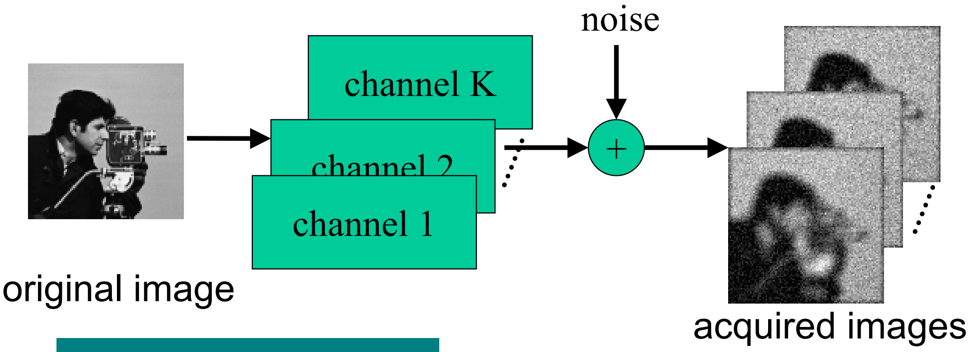
Averaging



Blind deconvolution

- Ill-posed problem for one single image
- Solution:
  - strong prior knowledge of blurs and/or the original image
  - multiple acquisitions of the same object (multichannel blind deconvolution)

Realistic acquisition model (1)



$$[u * h_k](x,y) + n_k(x,y) = z_k(x,y)$$

MC Blind Deconvolution

- System of integral equations
  (ill-posed, underdetermined)

$$z_k(x) = (h_k * u)(x) + n_k(x)$$

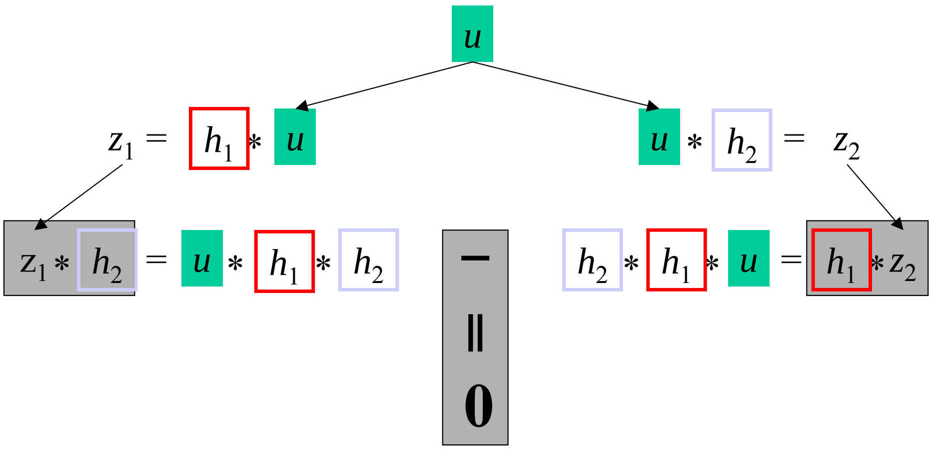
- Energy minimization problem (well-posed)

$$E(u, \{h_i\}) = \frac{1}{2} \sum_{i=1}^{K} \|h_i * u - z_i\|^2 + \lambda Q(u) + \gamma R(\{h_i\}),$$

Image Regularization

- $Q(u)$ captures local characteristics of the image => Markov Random Fields
- Identity: $\int_\Omega |u|^2$
- Tichonov (GMRF): $\int_\Omega |\nabla u|^2$
- Variational integral: $\int_\Omega \phi(|\nabla u|)$
- Huber MRF, bilateral filters, …

PSF Regularization



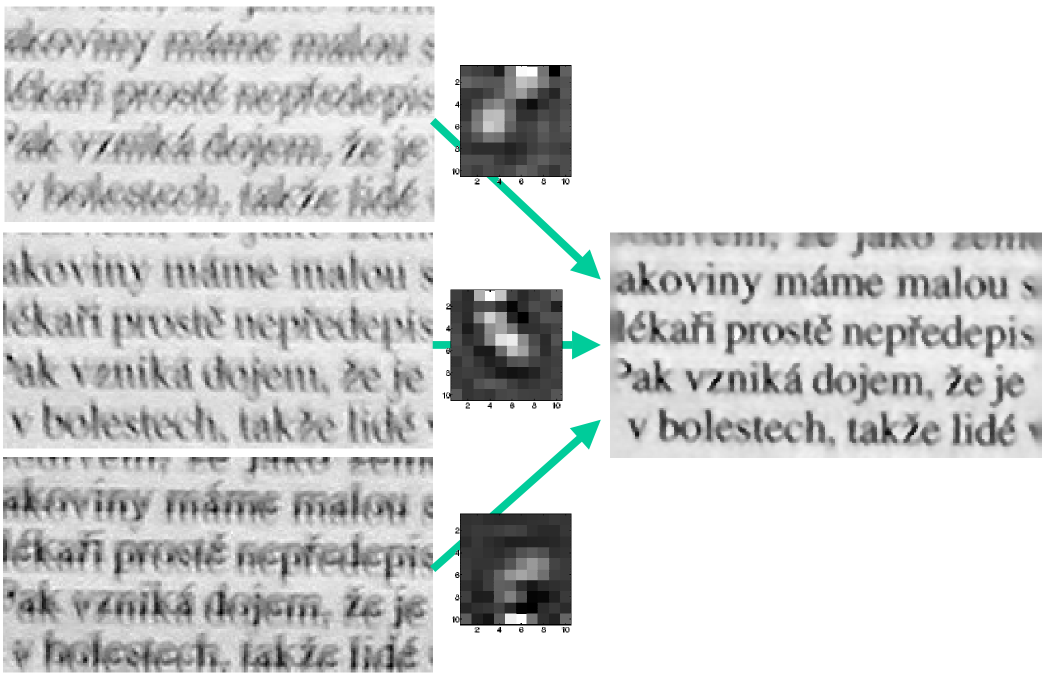
$$R(\{h_i\}) = \frac{1}{2} \sum_{1 \le i,j \le K} \|z_i * h_j - z_j * h_i\|^2$$

with one additional constraint $0 \le h_i(x) \le 1, \quad \forall x, i$

AM Algorithm

- Alternating minimizations of $E(u,\{h_i\})$ over u and h_i
- input: blurred images and estimation of PSF size
- output: reconstructed image and PSFs

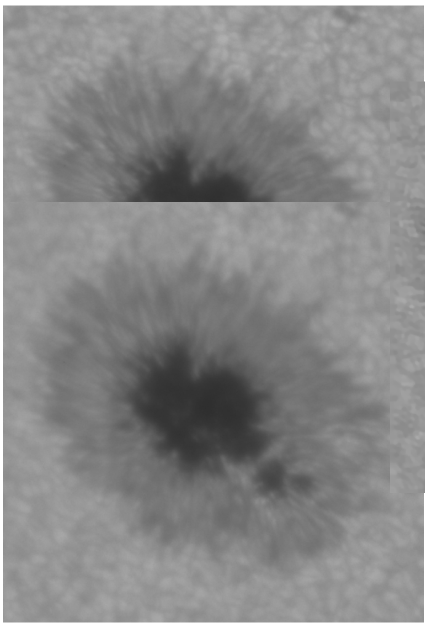
Vibrating Objects



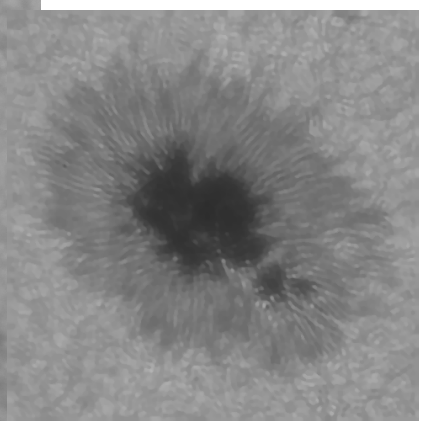
Long-time exposure

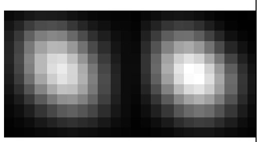
Astronomical Imaging

degraded image



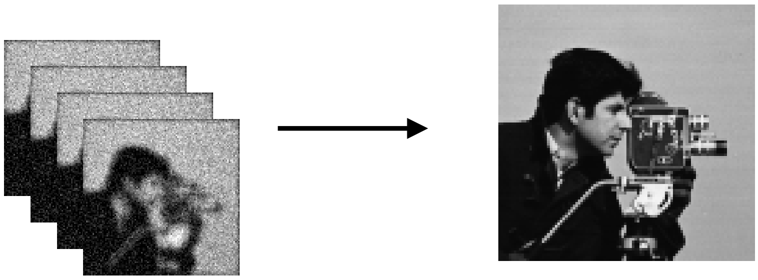
reconstructed image



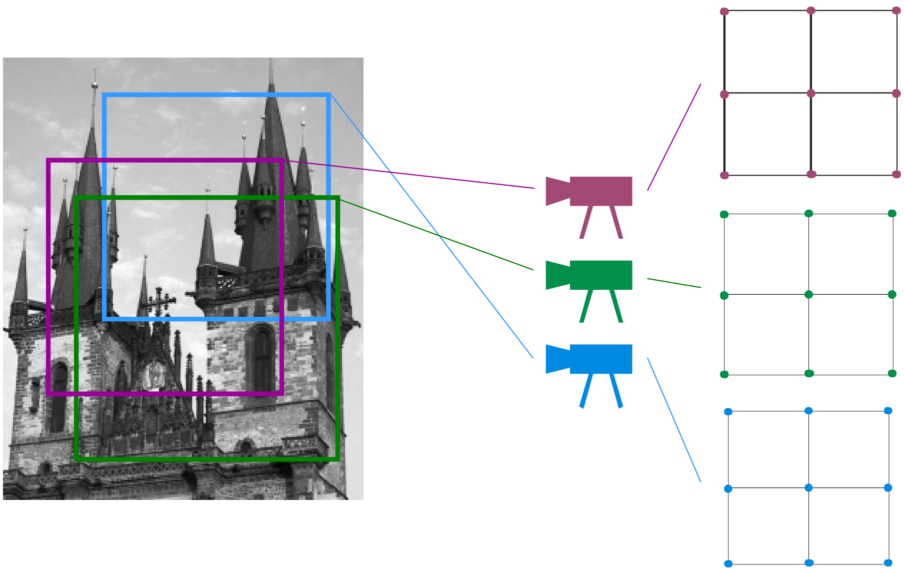


Superresolution

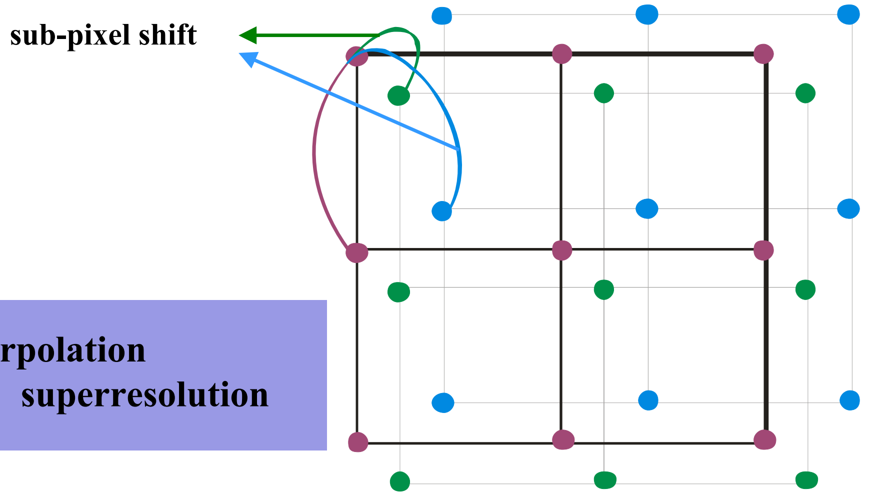
Goal: Obtaining a high-res image from several low-res images



Traditional superresolution

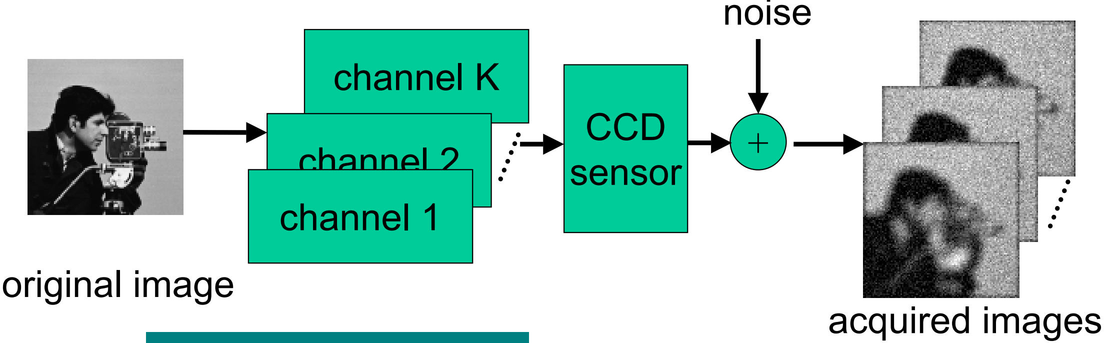


Traditional superresolution



pixel interpolation
→ superresolution

Realistic acquisition model (2)



$$\boldsymbol{D}\left([u \; * \; h_k](x, y) \right) \; + \; n_k(x, y) \; = \; z_k(x, y)$$

SR & MBD

- Incorporating between-image shift

$$[u * h_k](\tau_k(x,y)) + n_k(x,y) = z_k(x,y)$$

$$[u * g_k](x,y) + n_k(x,y) = z_k(x,y)$$

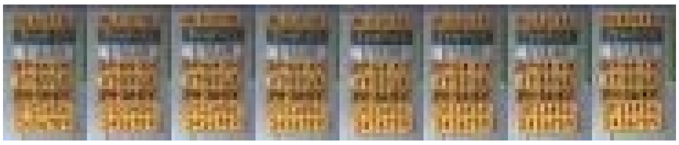
- Incorporating downsampling operator D

$$D[u * g_k](x,y) + n_k(x,y) = z_k(x,y)$$

Superresolution: No blur, SRF = 2x

Superresolution with High Factor

Input LR frames

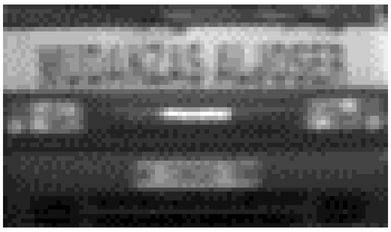


interpolated

SR

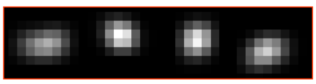
Original frame

Superresolution and MBD



Scaled LR input images

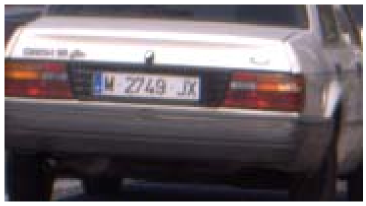
MBD+SR



PSFs

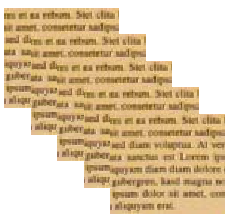
Superresolution and MBD

Superresolved image (2x)

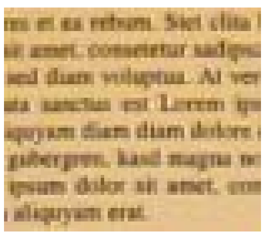


Optical zoom (ground truth)

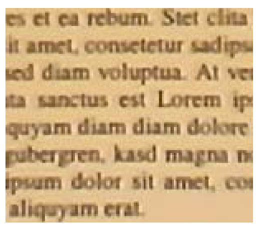
Cell-phone images



LR input images



Scaled input image



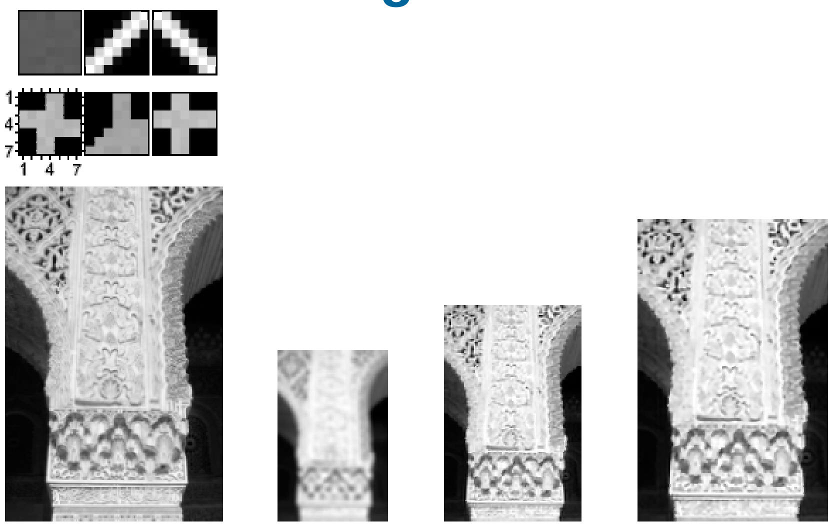
Superresolved image (2x)

Webcam images

LR input frame

Superresolution image (2x)

Superresolution with noninteger factors



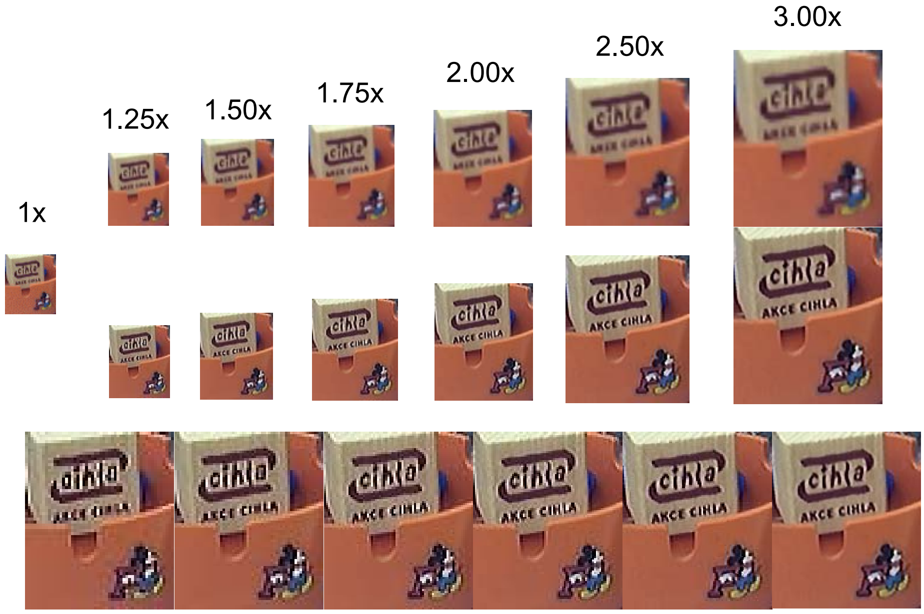
original image & PSFs

LR image

SR=1.25x

SR=1.75x

Noninteger SR factors



Challenges

- 3D scene
- objects with different motion
- improving registration

→

- space-variant deblurring
- motion field
- minimization over registration param.

IMAGE REGISTRATION

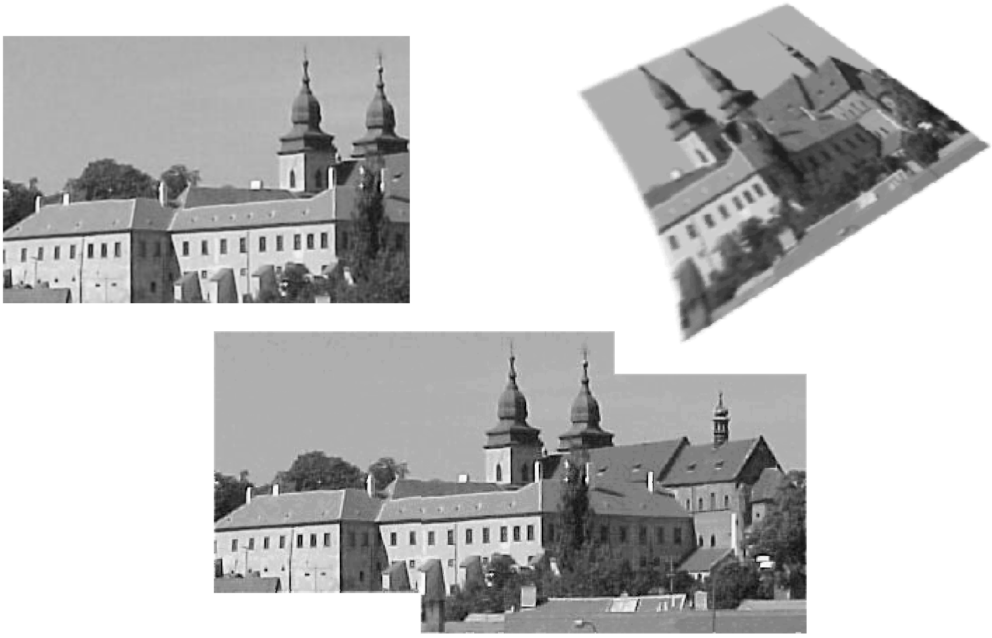
IMAGE REGISTRATION

methodology

feature detection
feature matching
transform model estimation
image resampling and transformation
accuracy evaluation

trends and future

METHODOLOGY: IMAGE REGISTRATION



METHODOLOGY: IMAGE REGISTRATION

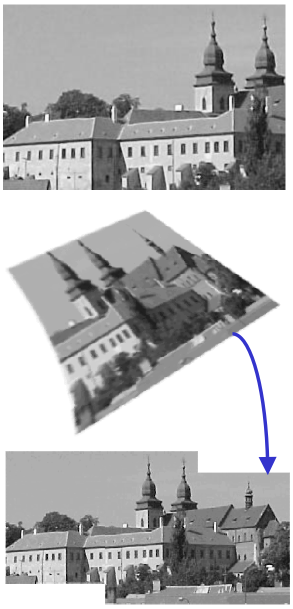
Overlaying two or more images of the same scene

Different imaging conditions

Geometric normalization of the image

Preprocessing of the images entering image analysis systems

METHODOLOGY: IMAGE REGISTRATION - TERMINOLOGY



reference image

sensed image

features

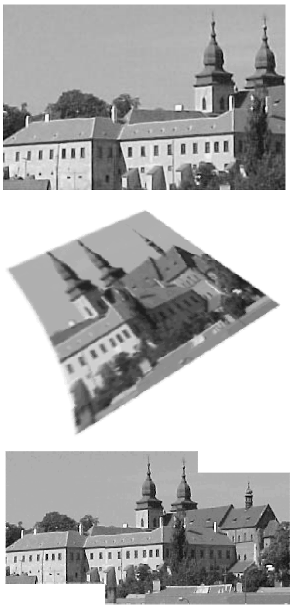
transform function

METHODOLOGY: IMAGE REGISTRATION

Main application categories

1. Different viewpoints - multiview
2. Different times - multitemporal
3. Differet modalities - multimodal
4. Scene to model registration

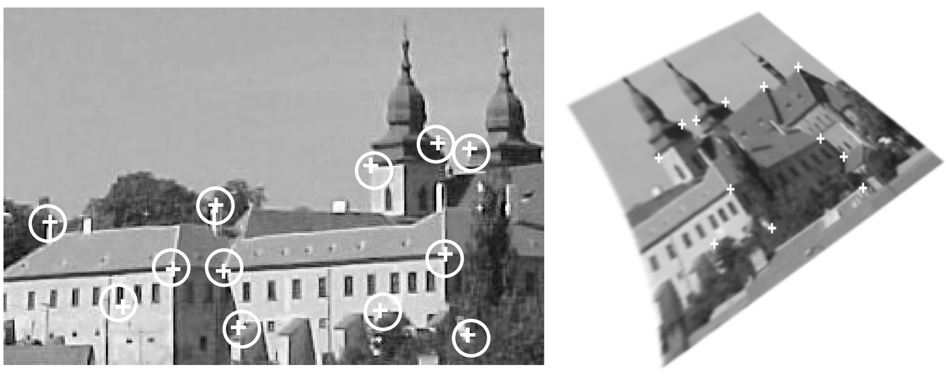
METHODOLOGY: IMAGE REGISTRATION



Four basic steps of image registration

1. **Feature detection**
2. **Feature matching**
3. **Transform model estimation**
4. **Image resampling and transformation**

FEATURE DETECTION



FEATURE DETECTION

Distinctive and detectable objects

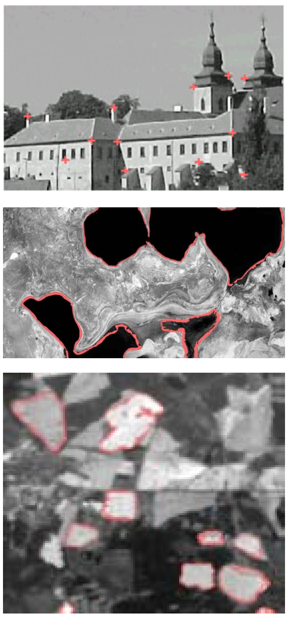
Physical interpretability

Frequently spread over the image

Enough common elements in all images

Robust to degradations

FEATURE DETECTION



Area-based methods - windows

Feature-based methods (higher level info)

- distinctive points
- corners
- lines
- closed-boundary regions
- invariant regions

FEATURE DETECTION **POINTS AND CORNERS**

distinctive points
- **line intersections**
- **max curvature points**
- **inflection points**
- **centers of gravity**
- **local extrema of wavelet transform**

corners
- **image derivatives** **(Kitchen-Rosenfled, Harris)**
- **intuitive approaches** **(Smith-Brady)**

FEATURE DETECTION **LINES AND REGIONS**

lines
- **line segments (roads, anatomic structures)**
- **contours**
- **edge detectors (Canny, Maar, wavelets)**

regions
- **closed- boundary objects (lakes, fields, shadows)**
- **level sets**
- **segmentation methods**

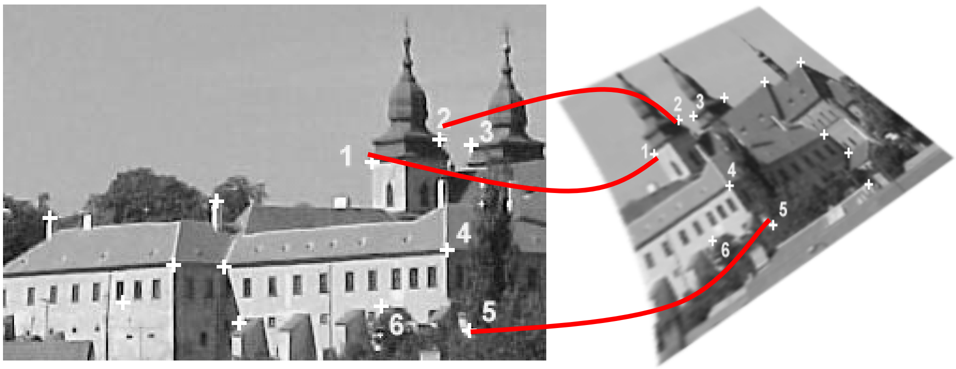
invariant regions with respect to assumed degradation

scale - virtual circles (Alhichri & Kamel)

affine - based on Harris and edges (Tuytelaars&V Gool)

affine - maximally stable extremal regions (Matas et al.)

FEATURE MATCHING



FEATURE MATCHING

Area-based methods

similarity measures calculated directly from the image graylevels

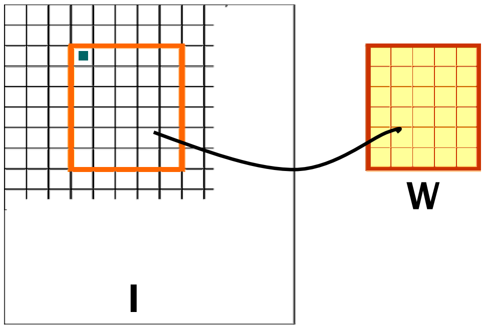
image correlation, image differences
phase correlation, mutual information, …

Feature-based methods

symbolic description of the features matching in the feature space (classification)

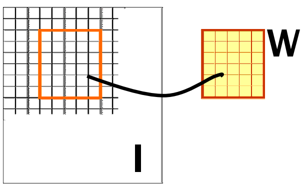
FEATURE MATCHING CROSS-CORRELATION

(k,m)



$$C(k,m) = \frac{\sum (I_{k,m} - \text{mean}(I_{k,m})) \cdot (W - \text{mean}(W))}{\sqrt{\sum (I_{k,m} - \text{mean}(I_{k,m}))^2} \cdot \sqrt{\sum (W - \text{mean}(W))^2}}$$

FEATURE MATCHING CORRELATION-LIKE METODS



edge, vector correlation

extension to complex transformations

hardware correlation

SSDA sequential similarity detection algorithm

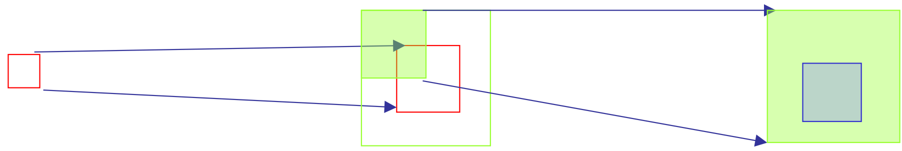
various similarity measures

error functions

subpixel accuracy

FEATURE MATCHING PYRAMIDAL REPRESENTATION

processing from coarse to fine level



wavelet transform

FEATURE MATCHING PHASE CORRELATION

equivalent to standard correlation of "whitened" images

similar to correlation of edges

does not depend on actual image colors

multimodal registration

FEATURE MATCHING PHASE CORRELATION

Fourier shift theorem

if *f(x)* is shifted by *a* to *f(x-a)*

- **FT magnitude stays constant**
- **phase is shifted by $-2\pi a\omega$**

shift parameter – spectral comparison of images

FEATURE MATCHING PHASE CORRELATION

SPOMF symmetric phase - only matched filter

image f window w

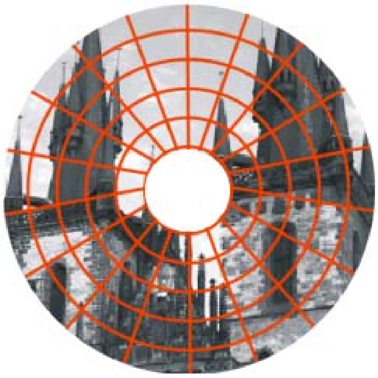
$$\frac{W \cdot F^*}{|W \cdot F|} = e^{-2\pi i(\omega a + \xi b)}$$

$$\text{IFT}\left(e^{-2\pi i(\omega a + \xi b)}\right) = \delta(x-a, y-b)$$

FEATURE MATCHING PHASE CORRELATION

shift solved, what about rotation and change of scale ?

log-polar transform



polar

$$r = [\ (x-x_c)^2 + (y-y_c)^2]^{1/2}$$

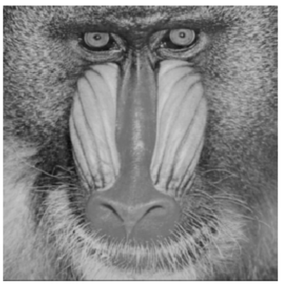
$$\theta = \tan^{-1}((y-y_c) / (x-x_c))$$

log

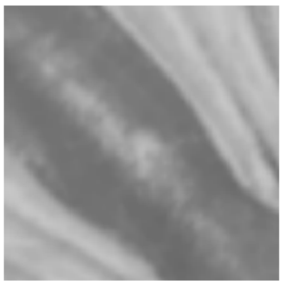
$$R = \frac{(n_r-1)\log(r/r_{min})}{\log(r_{max}/r_{min})}$$

$$W = n_w \theta / (2\pi)$$

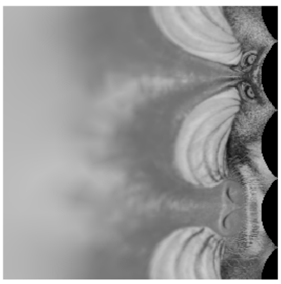
FEATURE MATCHING LOG-POLAR TRANSFORM



(a) input image



(b) scale=4; rotation=45°



FEATURE MATCHING RTS PHASE CORRELATION

Rotation, translation, change of scale

$$FT[f(x-a)](\omega) = \exp(-2\pi i a\omega) FT[f(x)](\omega)$$
$$FT[f_{rotated}](\omega) = FT[f]_{rotated}(\omega)$$
$$FT[f(ax)](\omega) = |a|^{-1} FT[f(x)](\omega/a)$$

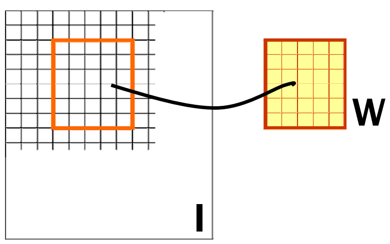
FT → | | → log-polar → FT → phase correlation

π - amplitude periodicity - > 2 angles
dynamics - log(abs(FT)+1)
discrete problems

FEATURE MATCHING MUTUAL INFORMATION

statistical measure of the dependence between two images

often used for multimodal registration



popular in medical imaging

FEATURE MATCHING MUTUAL INFORMATION

Entropy function

$$H(X) = -\sum_{x} p(x)\log p(x)$$

Joint entropy

$$H(X,Y) = -\sum_{x}\sum_{y} p(x,y)\log p(x,y)$$

Mutual infomation

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

FEATURE MATCHING MUTUAL INFORMATION

Entropy **measure of uncertainty**

Mutual information **reduction in the uncertainty of X due to the knowledge of Y**

Maximization of MI **measure *mutual agreement* between object models**

FEATURE MATCHING — FEATURE-BASED METHODS

Combinatorial matching

no feature description, global information

- graph matching
- parameter clustering
- ICP (3D)

Matching in the feature space

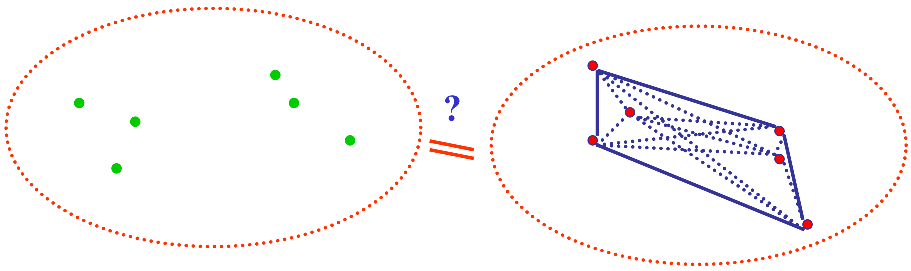
pattern classification, local information

- invariance
- feature descriptors

Hybrid matching

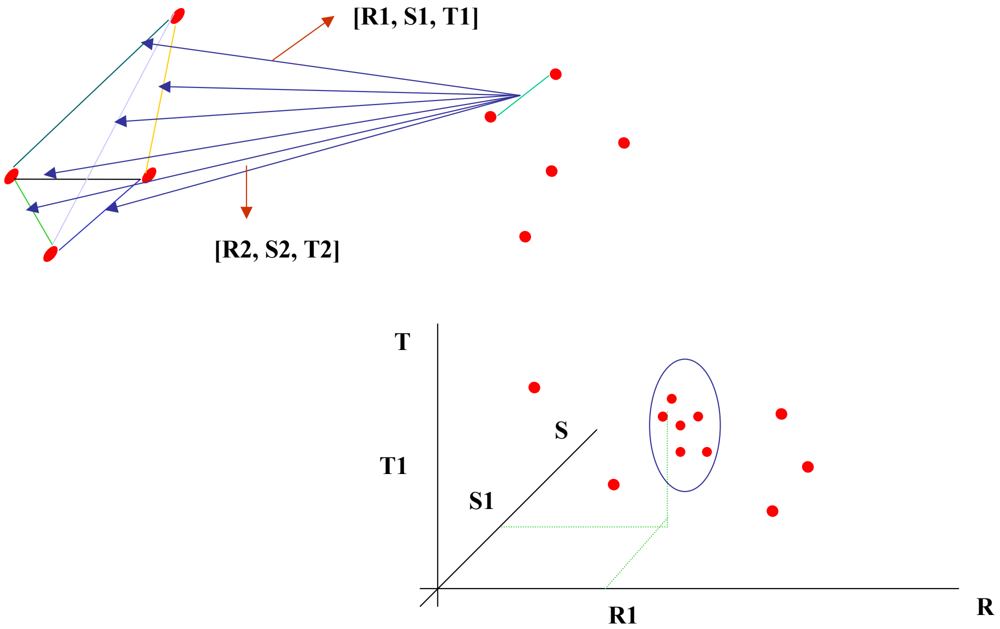
combination, higher robustness

FEATURE MATCHING — COMBINATORIAL - GRAPH



transformation parameters with highest score

FEATURE MATCHING **COMBINATORIAL - CLUSTER**



FEATURE MATCHING **FEATURE SPACE MATCHING**

Detected features **- points, lines, regions**

Invariants description

- intensity of close neighborhood

- geometrical descriptors (MBR, etc.)

- spatial distribution of other features

- angles of intersecting lines

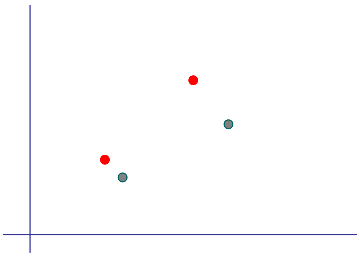
- shape vectors

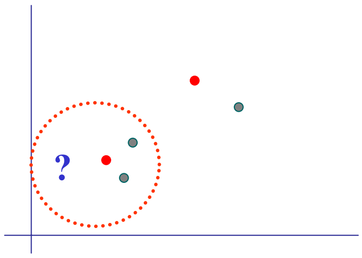
- moment invariants

- …

Combination of descriptors

FEATURE MATCHING FEATURE SPACE MATCHING





FEATURE MATCHING FEATURE SPACE MATCHING

maximum likelihood coefficients

	W1	W2	W3	W4
V1	Dist			
V2				
V3				
V4				

…

⋮

min (best / 2^{nd} best)

FEATURE MATCHING FEATURE SPACE MATCHING

relaxation methods – consistent labeling problem solution

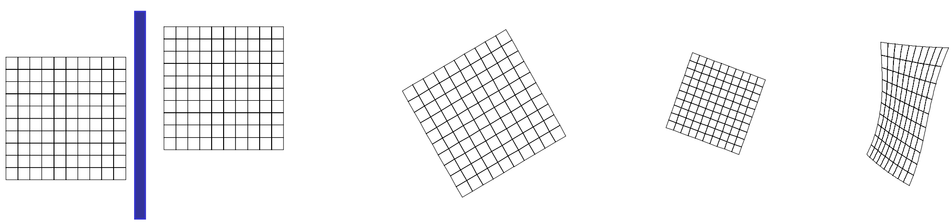
iterative recomputation of matching score

based on
- **match quality**
- **agreement with neighbors**
- **descriptors can be included**

RANSAC
- **random sample consensus algorithm**
- **robust fitting of models, many data outliers**
- **follows simpler distance matching**
- **refinement of correspondences**

TRANSFORM MODEL ESTIMATION

$$x' = f(x,y)$$
$$y' = g(x,y)$$



incorporation of *a priory* known information

removal of differences

TRANSFORM MODEL ESTIMATION

Global functions

similarity, affine, projective transform

low-order polynomials

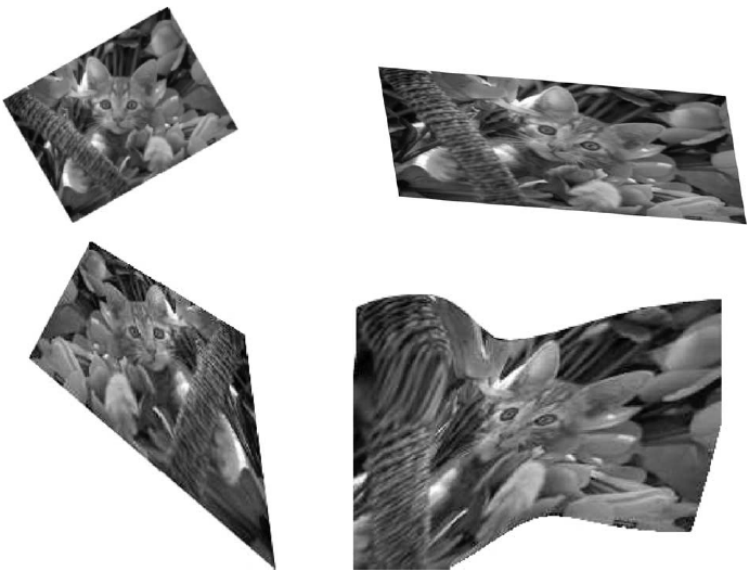
Local functions

piecewise affine, piecewise cubic

thin-plate splines

radial basis functions

TRANSFORM MODEL ESTIMATION



TRANSFORM MODEL ESTIMATION

Affine transform

$$x' = a_0 + a_1 x + a_2 y$$
$$y' = b_0 + b_1 x + b_2 y$$

Projective transform

$$x' = (a_0 + a_1 x + a_2 y) / (1 + c_1 x + c_2 y)$$
$$y' = (b_0 + b_1 x + b_2 y) / (1 + c_1 x + c_2 y)$$

TRANSFORM MODEL ESTIMATION - SIMILARITY TRANSFORM

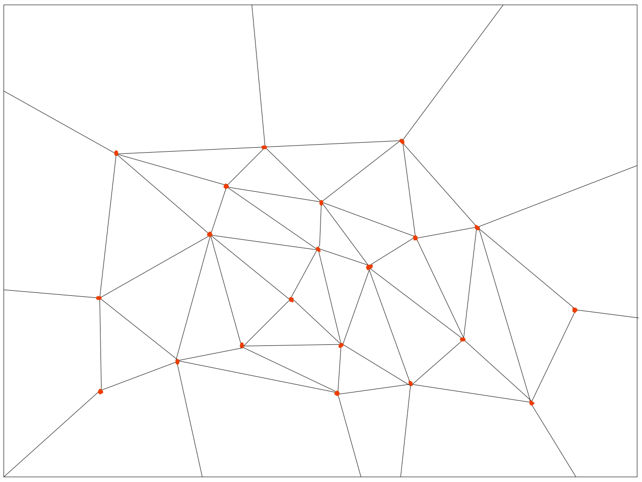
translation $[\Delta x, \Delta y]$ **rotation** φ **uniform scaling** s

$$x' = s(x * \cos\varphi - y * \sin\varphi) + \Delta x$$
$$y' = s(x * \sin\varphi + y * \cos\varphi) + \Delta y$$
$$s \cos\varphi = a, \quad s \sin\varphi = b$$

$$\min \left(\Sigma_{i=1} \{ [x_i' - (ax_i - by_i) - \Delta x]^2 + [y_i' - (bx_i + ay_i) - \Delta y]^2 \} \right)$$

$$\begin{vmatrix} \Sigma(x_i^2 + y_i^2) & 0 & \Sigma x_i & \Sigma y_i \\ 0 & \Sigma(x_i^2 + y_i^2) & -\Sigma y_i & \Sigma x_i \\ \Sigma x_i & -\Sigma y_i & N & 0 \\ \Sigma y_i & \Sigma x_i & 0 & N \end{vmatrix} \bullet \begin{vmatrix} a \\ b \\ \Delta x \\ \Delta y \end{vmatrix} = \begin{vmatrix} \Sigma(x_i' x_i - y_i' y_i) \\ \Sigma(y_i' x_i - x_i' y_i) \\ \Sigma x_i' \\ \Sigma y_i' \end{vmatrix}$$

TRANSFORM MODEL ESTIMATION - PIECEWISE TRANSFORM



TRANSFORM MODEL ESTIMATION UNIFIED APPROACH

Pure interpolation – ill posed

Regularized approximation – well posed

$$\min \ J(f) = a\,E(f) + b\,R(f)$$

E(f) **error term**

R(f) **regularization term**

a,b **weights**

TRANSFORM MODEL ESTIMATION UNIFIED APPROACH

Choices for **min** $J(f) = a\,E(f) + b\,R(f)$

$E(f) = \Sigma\,(x_i' - f(x_i, y_i))^2$

$R(f)$ $>=0$

$||L(f)||$

$a << b$ **least-square fit,**

f from the null-space of L

$a >> b$ **"smooth" interpolation**

TRANSFORM MODEL ESTIMATION UNIFIED APPROACH

Choices for **min** $J(f) = a\,E(f) + b\,R(f)$

$$R(f) = \iint \left(\frac{\partial^2 f}{\partial x \partial x}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y \partial y}\right)^2 dx\,dy$$

$$f(x, y) = \alpha_1 + \alpha_2 x + \alpha_3 y + \sum_{i=1}^{N} a_i g_i(\|x - x_i, y - y_i\|),$$

TPS $g_i(t) = t^2 \log t$

another choice ***G-RBF*** $g_i(t) = \exp\left(\frac{-t^2}{\sigma^2}\right)$

TRANSFORM MODEL ESTIMATION OTHER REGISTRATIONS

Elastic registration

- not parametric models
- "rubber sheet" approach

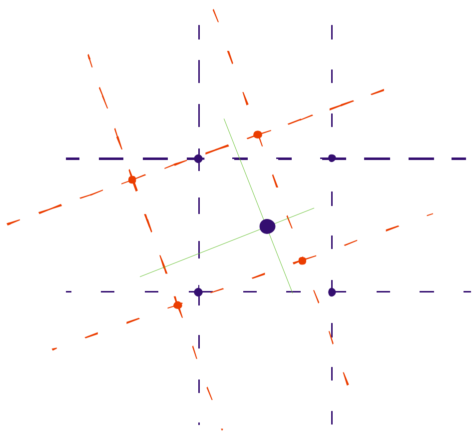
Fluid registration

- viscous fluid model to control transformation
- reference image – thick fluid flowing to match

Diffusion-based registration

Optical flow registration

IMAGE RESAMPLING AND TRANSFORMATION



trade-off between accuracy and computational complexity

IMAGE RESAMPLING AND TRANSFORMATION

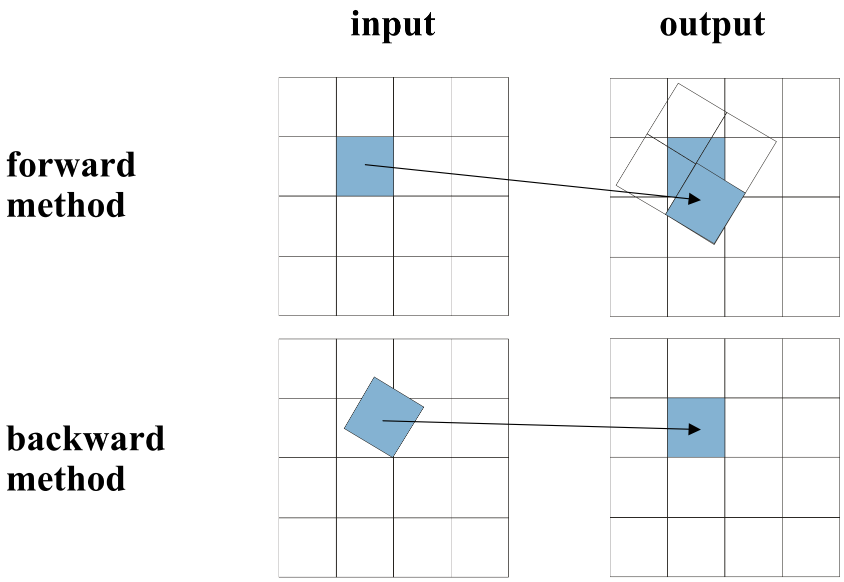


IMAGE RESAMPLING AND TRANSFORMATION

Interpolation nearest neighbor

bilinear

bicubic

Implementation 1-D convolution

$$f(x_o,k) \quad = \Sigma\, d(I,k).c(i-x_0)$$

$$f(x_0, y_0) \; = \Sigma\, f(x_0, j).c(j-y_0)$$

ideal $c(x) \; = k.sinc(kx)$

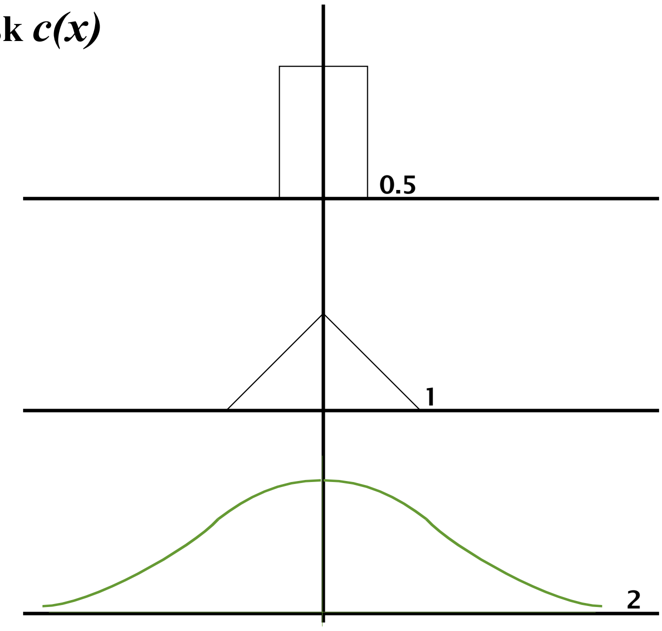
IMAGE RESAMPLING AND TRANSFORMATION

Interpolation mask $c(x)$

closest neighbour

linear

smooth cubic



ACCURACY EVALUATION

Localization error - displacement of features
- due to detection method

Matching error - false matches
- ensured by robust matching (hybrid)
- consistency check, cross-validation

Alignment error - difference between model and reality
- mean square error
- test point error (excluded points)
- comparison ("gold standard")

TRENDS AND FUTURE

complex local transforms

multimodal data

robust systems, based on combination of approaches

3D data sets

expert systems

APPLICATIONS

Different viewpoints

Different times (change detection)

Different sensors/modalities

Scene to model registration

PUBLICATIONS

Papers

- **L. G. Brown, A survey of image registration techniques, ACM Computing Surveys, 24:326-376, 1992**

- **B. Zitová and J. Flusser, Image registration methods: a survey, Image and vision computing, 21(11): 977-1000, 2003**

Books

- **A. Goshtasby, 2-D and 3-D Image Registration, Wiley Publishers, New York, April 2005**

- **J. Hajnal, D.Hawkes, and D. Hill, Medical Image registration, CRC Press, 2001**

- **J. Modersitzki , Numerical Methods for Image Registration, Oxford University Press, 2004**

- **K. Rohr, Landmark-Based Image Analysis - Using Geometric and Intensity Models, Kluwer Academic Publishers, Dordrecht Boston, London 2001**