The part and the whole:

Simple rules on complex image structures

Tamás Szirányi

MTA SZTAKI

Budapest

SSIP 2016

- Robot vision around us

- Axiomatic rules in Computer Vision
- Scale space and anisotropic diffusion
- 3D modelling
- Markov Random Fields Models
- Stochastic image alignment

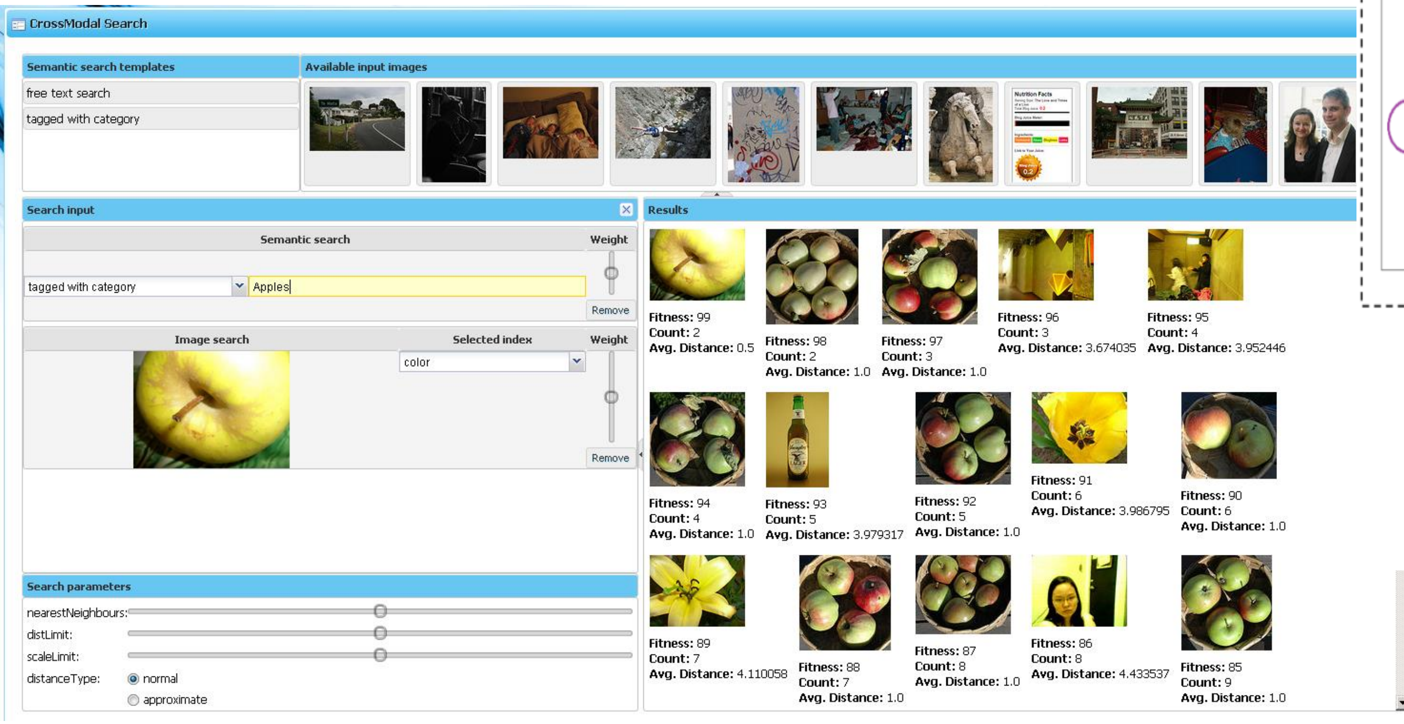
The robotic world on the street: the near future

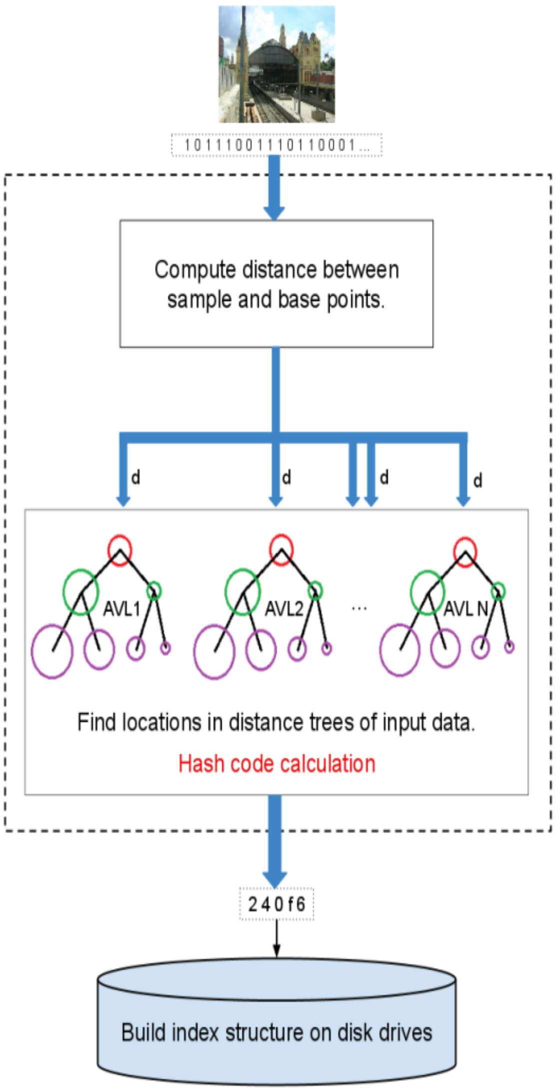
MTA SZTAKI

Recognition and Clustering in Multimedia Databases:

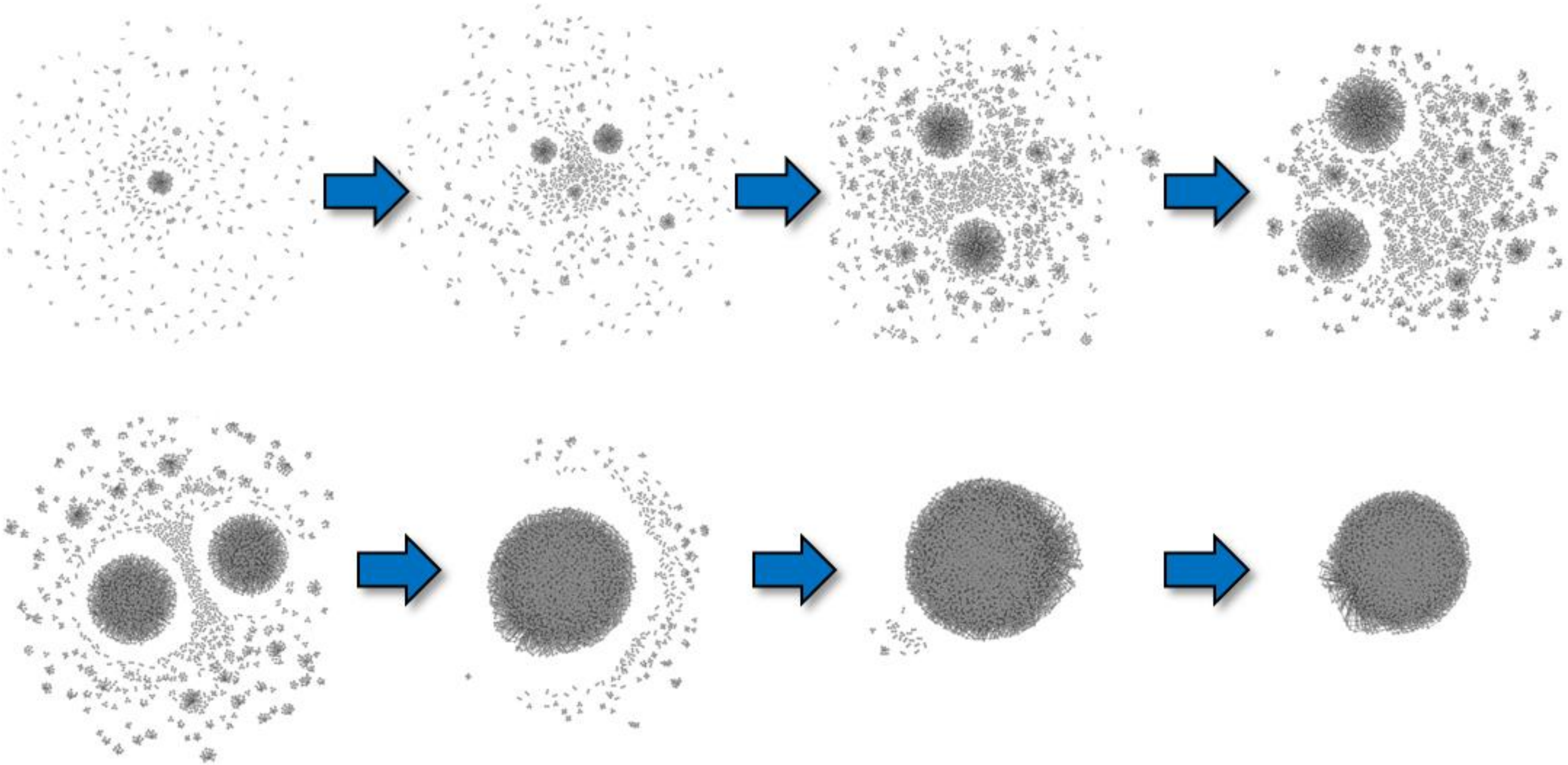
- 1000-100.000 parameter,
- 10^5 - 10^7 elements

What is the graph-structure of elements?
Which parameters can be best clustered?





Clustering behavior of big databases follows the Giant Components (Erdős-Rényi) for random geometric graphs *CLEF, 2012*



Keszler A. & Kovacs L. & Sziranyi T.

Identification vs Recognition

Identification ⟺ Recognition

- N cardinality of the community (e.g. 200)
- ε false accept (e.g. 1%)
- $1 - \varepsilon$ probability of right rejection (e.g. 99%): **identification**
- $(1 - \varepsilon)^N$ probability of right rejection for N persons
- $1 - (1 - \varepsilon)^N$ probability that person N+1. (intruder) could go through
- If $\varepsilon = 1/N$, then the probability of false accept

$$1 - (1 - 1/N)^N \Rightarrow 1 - 1/e = 63\%$$

prob. of false passing

”Scale-space” axioms

If zooming or rotating, the richness of details may change, but the essence not; give us rules:

- *Linearity:* $T_t(af + bh) = aT_tf + bT_th$
- *Shift invariance*
- *Semi-group:* $g(x,y,t_1) * g(x,y,t_2) = g(x,y,t_1 + t_2)$ *(cascade filter)*
- *Local max/min relations do not change during scalig*
- *No new features (zero-crossings)*
- *rotation symmetry*

- Lindeberg & Haar Romeny: ”Linear Scale-Space”, Kluwers, 1994
- Babaud et al: ”Uniqueness of the Gaussian kernel for scale-space filtering”, IEEE PAMI, 1986

A Scale-Space axioms to a new paradigm of image description

- Scale-Space theory: *scaling and* Gauss function
- *Anisotropic diffusion* → de-noising
- *Wavelet* description and uncertainty principle
- Image structure from scales: Lindeberg' iterations
- Finding the most featuring scale
- Descriptors around the key-points in 128 dimesions: SIFT

A Scale-Space axioms

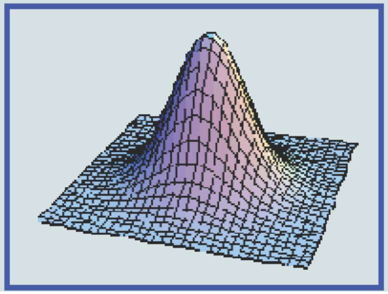
- Lineariy
- Shift invariance
- Isotropy
- Causality
- Separability

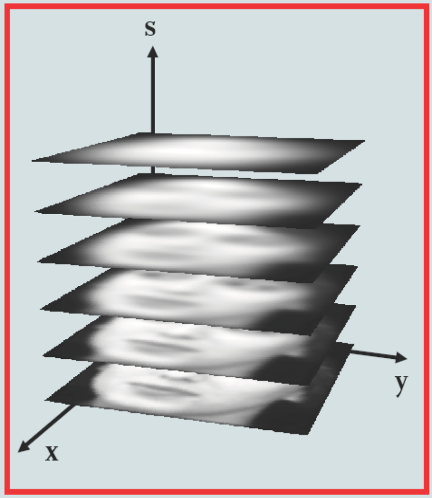
lead to the Gaussian smoothing.

Gaussian Scale Space

$$\boxed{u(\mathbf{x};s)} = (f * \varphi_s)(\mathbf{x})$$

$$\boxed{\varphi_s(\mathbf{x})} = \frac{1}{\sqrt{4\pi s}^{\,n}} e^{-\frac{\|\mathbf{x}\|^2}{4s}}$$





From fine details to overall features: Gaussian smoothing

Original

Gradient (scale 1 pixel)

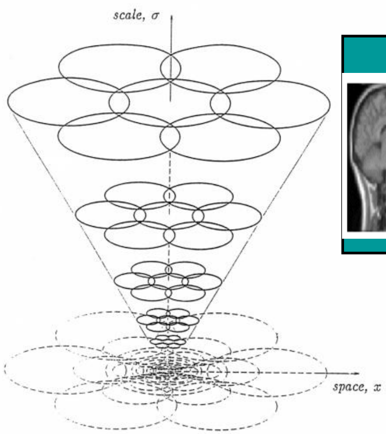
Gradient (scale 4 pixels)

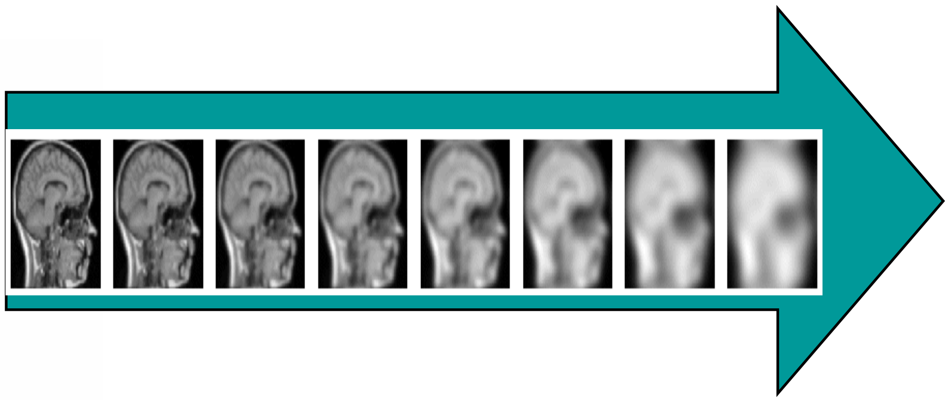
Scale induces an image hierarchy.

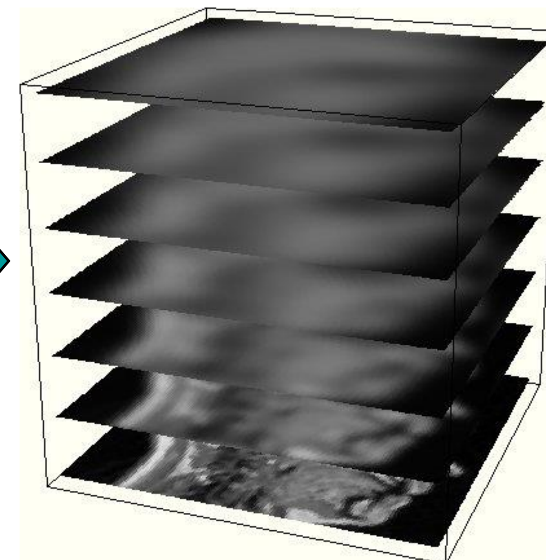
Looking for the most discriminative scale

The position and the detailness fulfils the uncertainty discipline

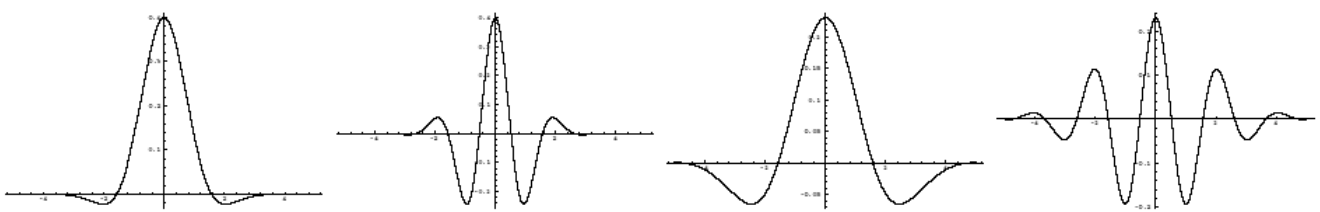
Vision processes several scales as well





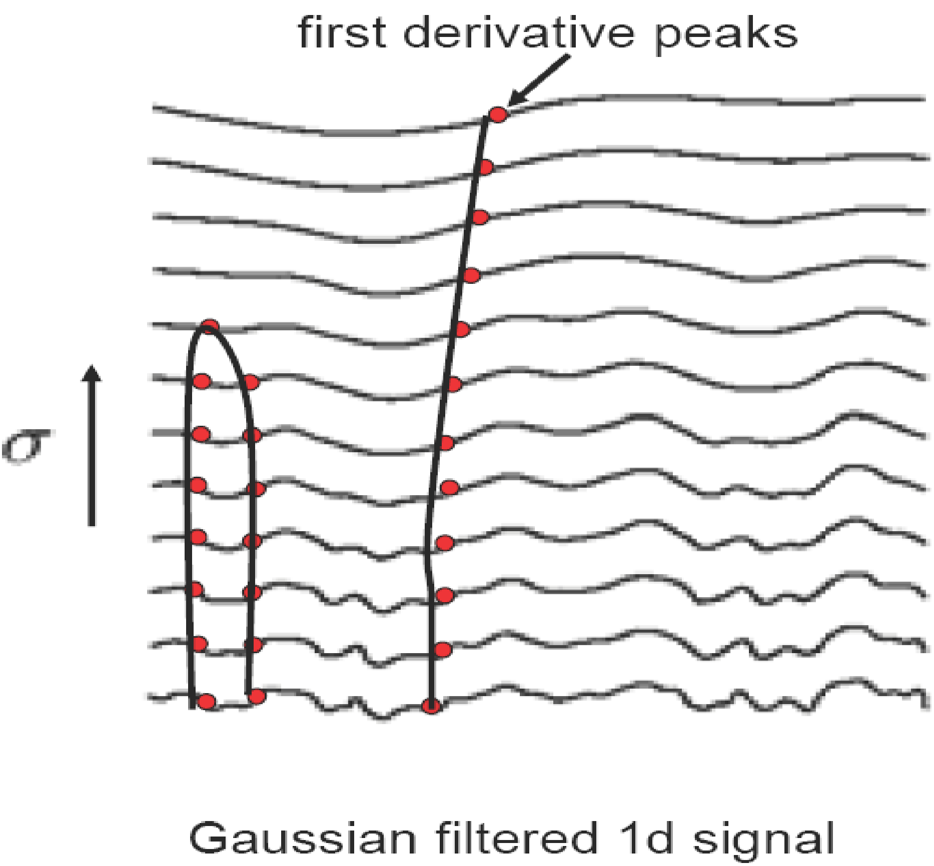


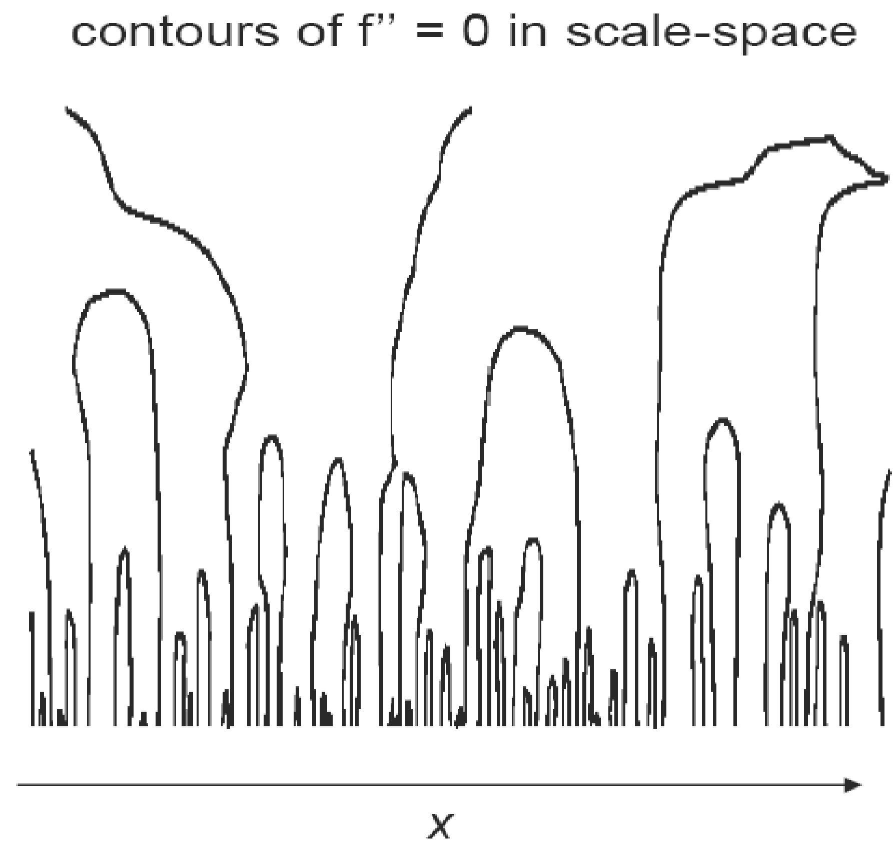
scale-space



$$< \Delta x, \Delta \omega > \geq \frac{1}{2} \qquad \Delta x \Delta \nu \geq \frac{n}{4\pi}$$

The edge transitions follow the axioms



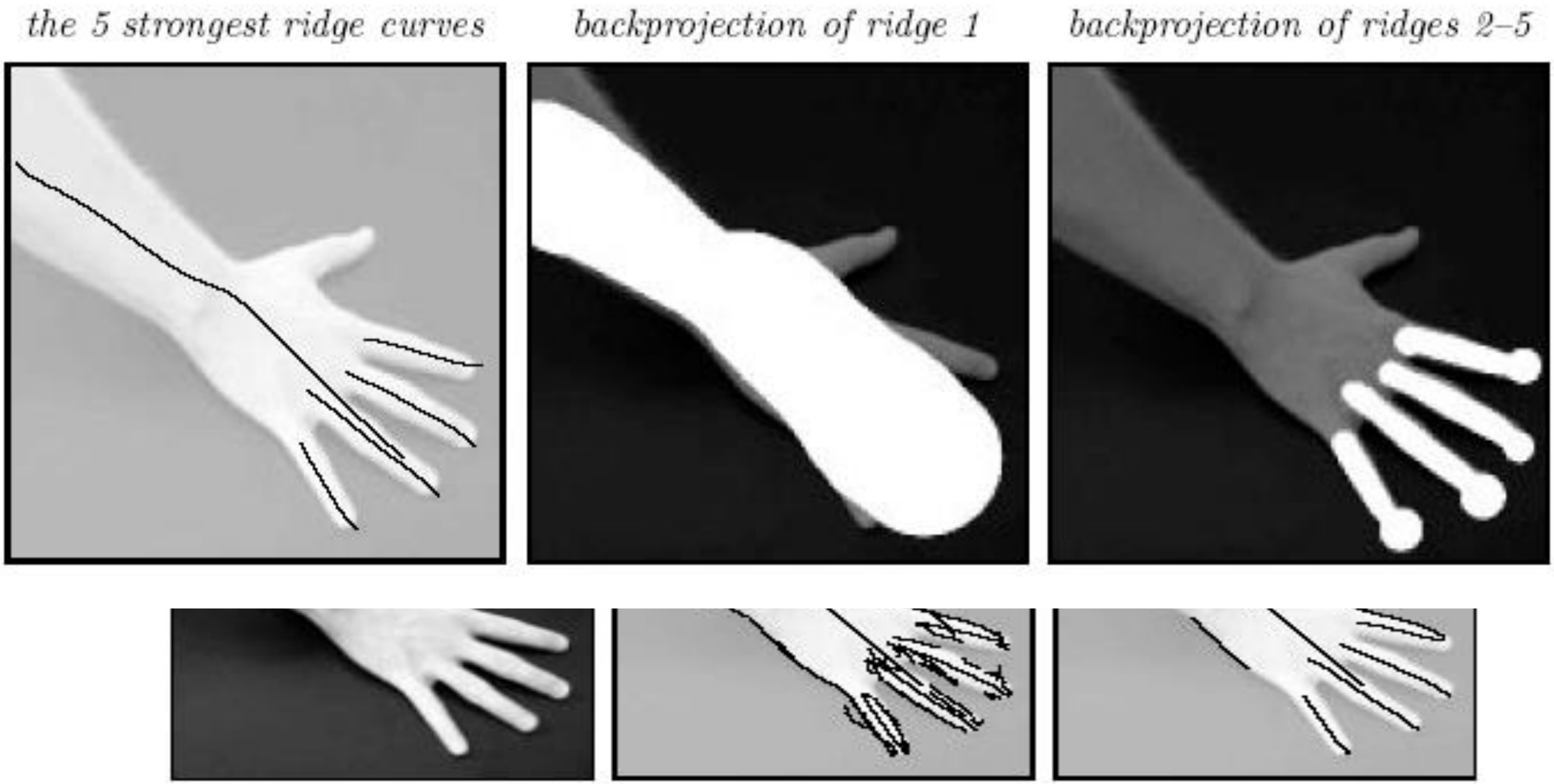


Gradually smoothed signal

The positions of gradient-changes do not cross each other through scales

Scale space (Witkin 1983)

Defining edge, ridge and featuring scale



Tony Lindeberg: ”Edge detection and ridge detection with automatic scale selection”, IJCV, 1998

Anisotropic diffusion (de-noising)

$$\partial_t u = w_K \left(\left| \nabla G_\sigma * u \right| \right) \left((1\text{-}\alpha)\, u_{\parallel} + \alpha\, u_{\perp} \right)$$

- Perona & Malik: "Scale-space and edge detection using anisotropic diffusion“, 1987
- Alvarez, Guichard, Lions, Morel: "Axioms and fundam. eqs. of image proc.“, 1993

$\partial_t u = p\left(|\nabla G_\sigma * u| \right) u_{\parallel} + n\left(|\nabla G_\sigma * u| \right) u_{\perp}$

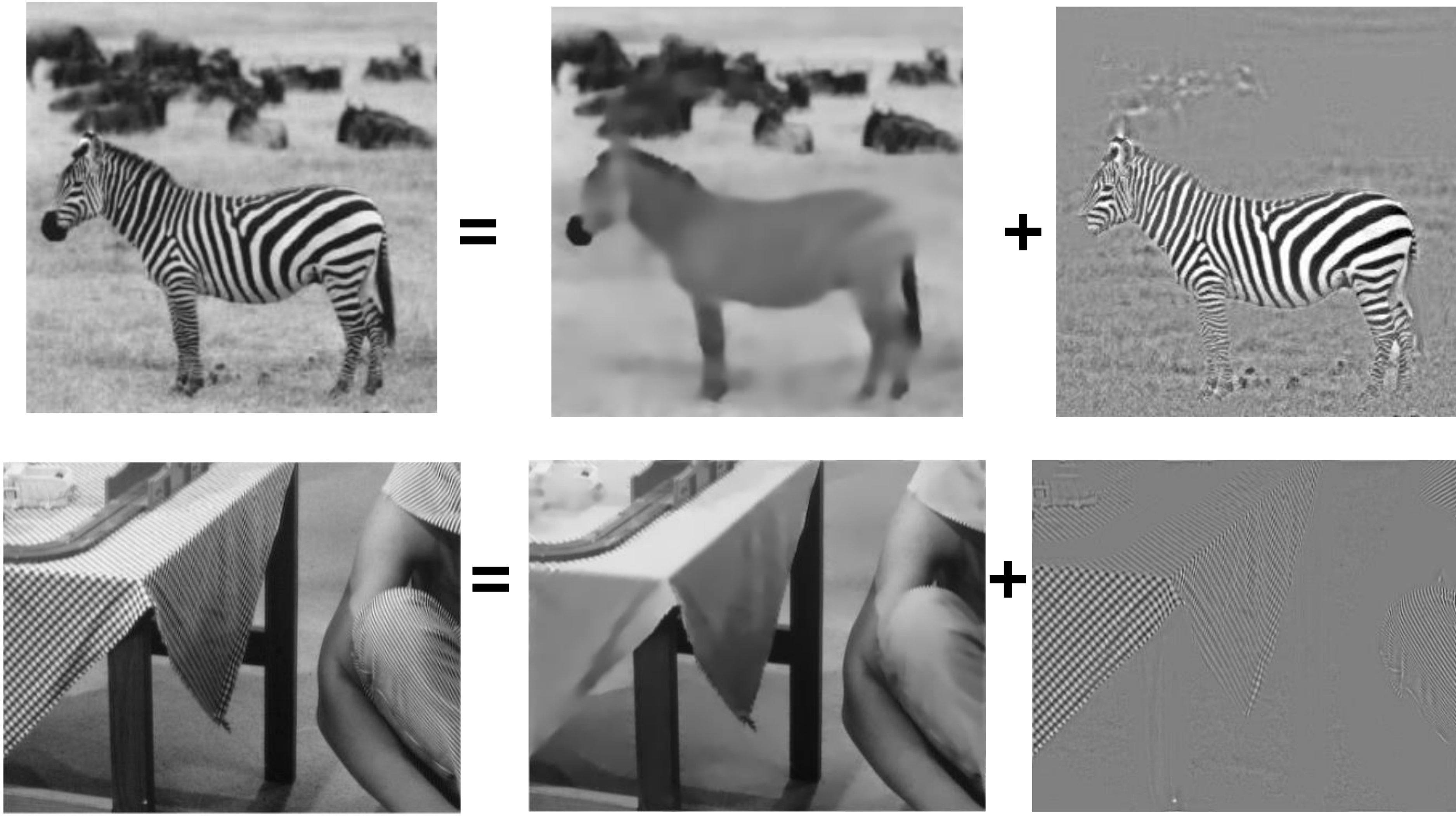
Anisotropic diffusion in preporocessing for JPEG compression 1998 / 2005

Original

JPEG

AD prefiltered JPEG

Kopilovic & Szirányi; Optical Engineering, 2005



Cartoon/Texture separation with TotVar and adaptive anisotropic diffusion -*2012*

Scale Invariant Feature Transform

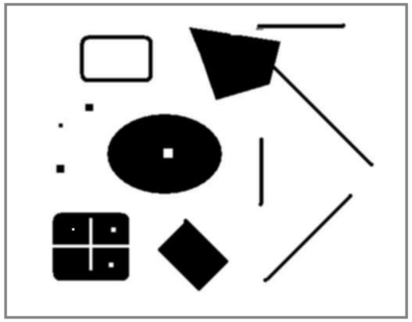
- Local features are specific for shape and less information loss in case of occlusion
- Main steps:
  - Keypont definitions,
  - Featuring areas,
  - Descriptors.

Hessian detector *(Beaudet, 1988)*

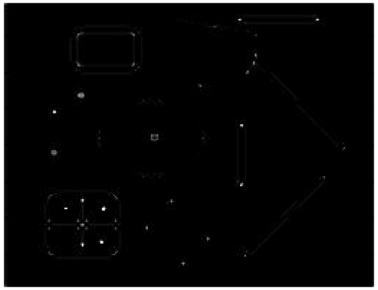
- Search for strong derivatives in two orthogonal directions:

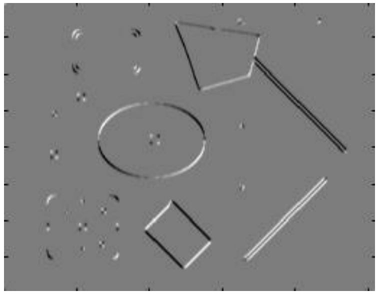
$$Hessian(I) = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix}$$

$$\det(Hessian(I)) = I_{xx} I_{yy} - I_{xy}^2$$



I_{xx}





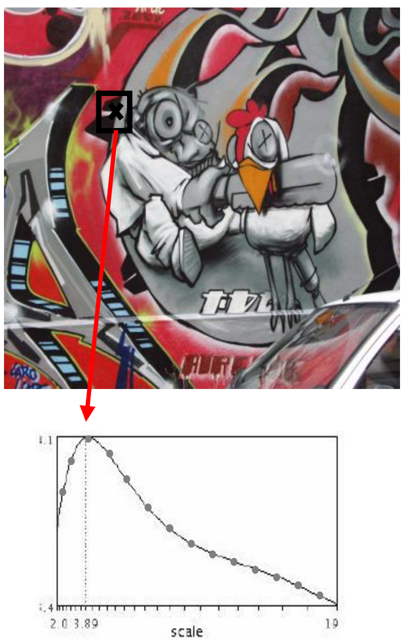
Responses mainly on corners and strongly textured areas.

I_{xy} I_{yy}

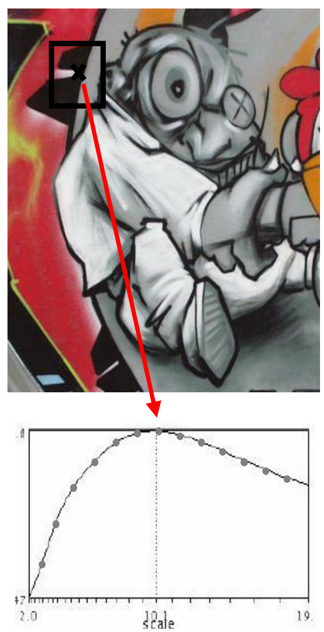
MTA SZTAKI

Automatic scale selection: *LOG*

- Function response of scale signature

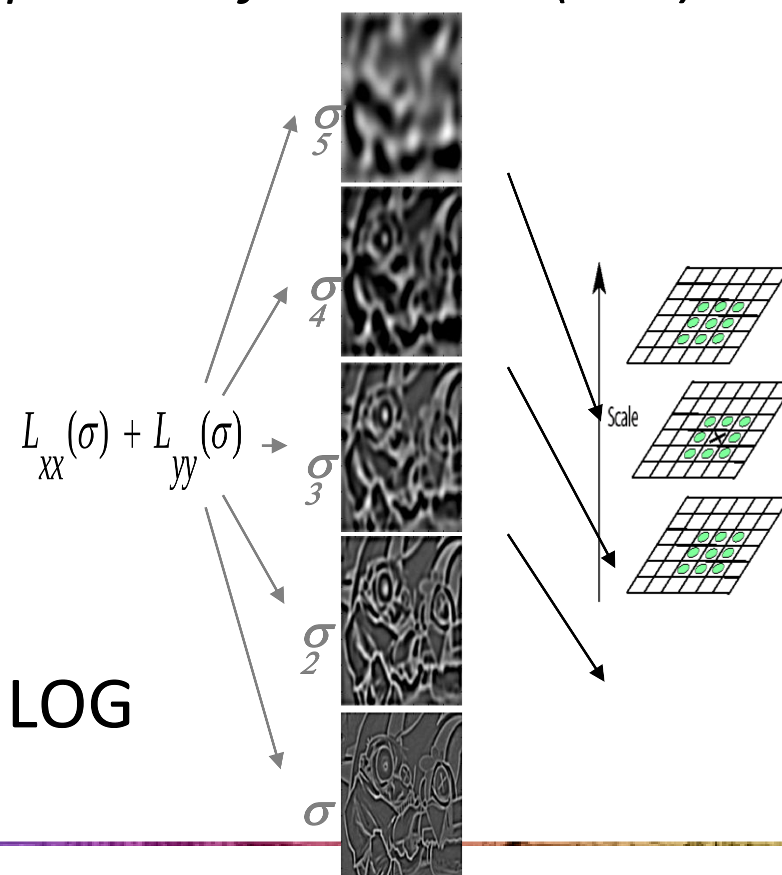


$f(I_{i_1 \dots i_m}(x,\sigma))$



$f(I_{i_1 \dots i_m}(x',\sigma'))$

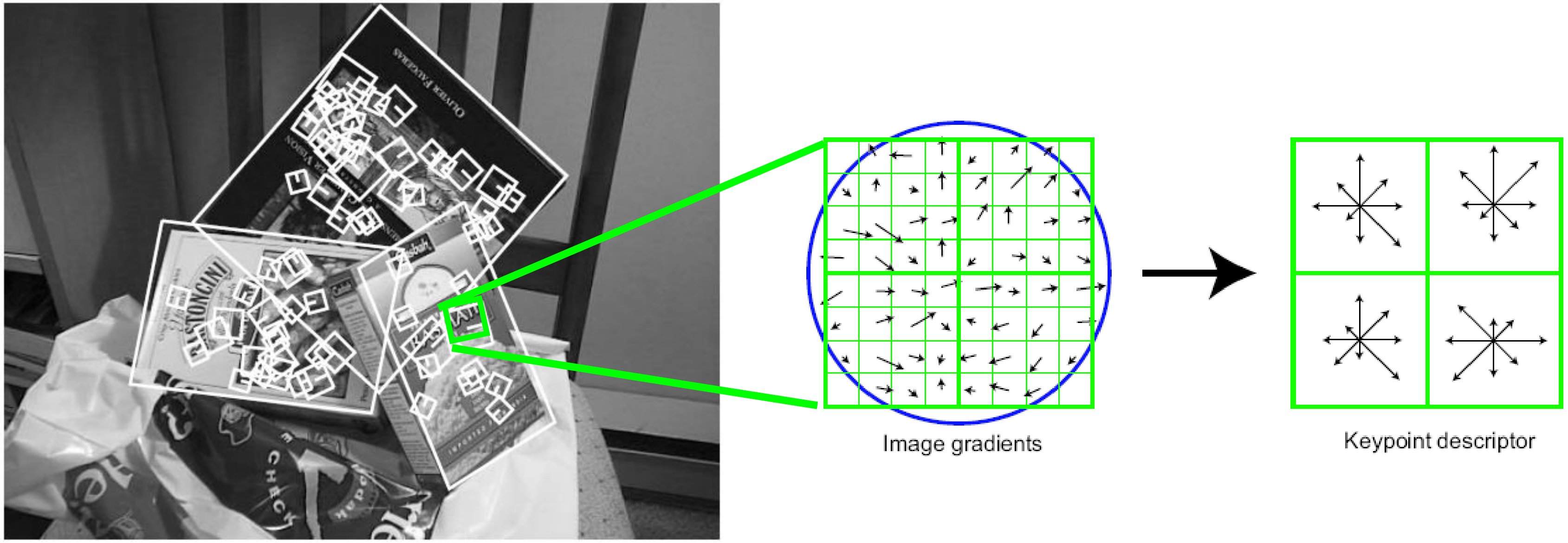
Laplacian of Gaussians (LOG)



- Local maxima in scale-space of LOG

SIFT descriptor *(Lowe, ICCV, 1999)*

- Histogram of oriented gradients
- Captures important texture information
- Robust to small translations and affine deformations



Remote sensing:

Fusion MRF model for multi-time image segmentation and change detection

Tamás Szirányi and Maha Shadaydeh

IEEE GRSL 2014

7/13/2016

1984

2000

2005

2007

Fused segmentation of the different subclasses

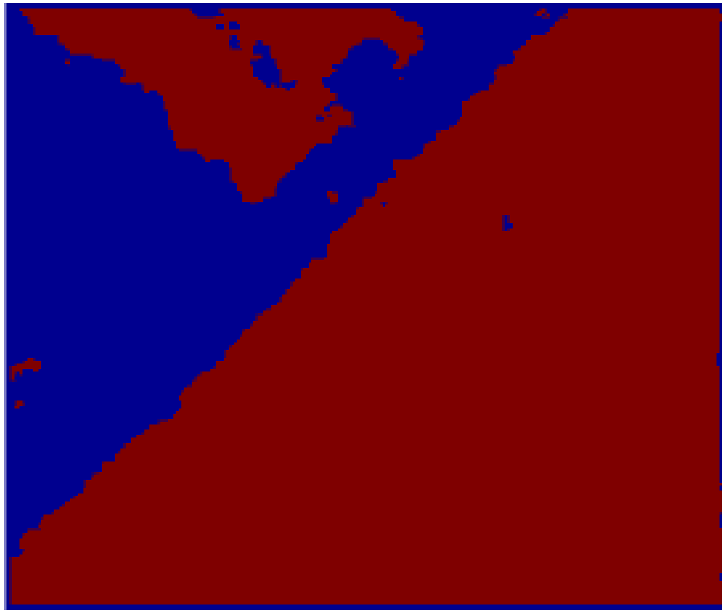
More images

↓

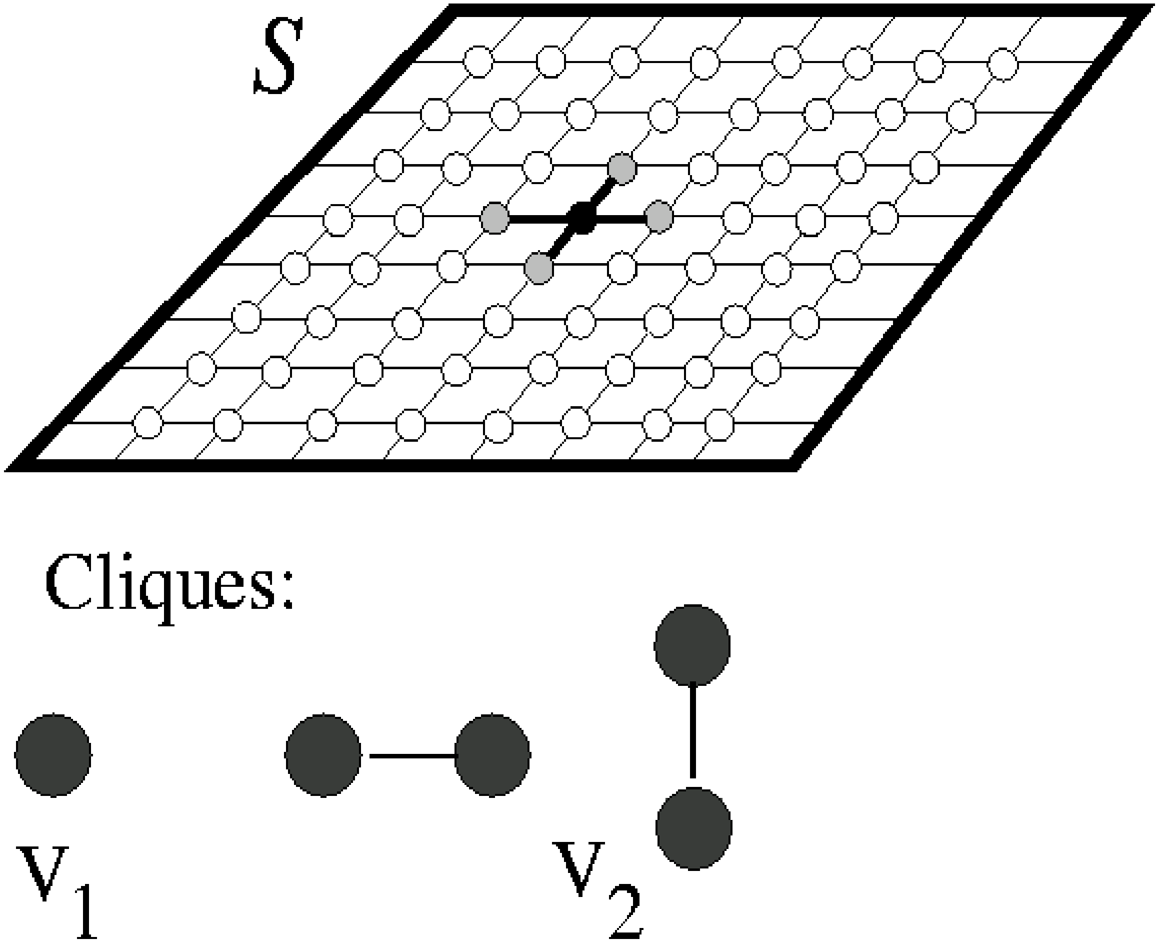
More information

↓

More complex details



Markov Random Field Optimization
Local steps – global optimum



MRF models and segmentation levels:

- *Single layer*

  single year - some supervision is needed

- *Multiple layers* (stack of years' layers):

  the source of supervision for single layer step

$$\widehat{\Omega} = \operatorname{argmin}_{\Omega} \sum_{s \in S} \underbrace{-\log P(\overline{x}_s | \omega_s)}_{\epsilon_{\omega_s}(s)} + \sum_{r,s \in S} \Theta(\omega_r, \omega_s)$$

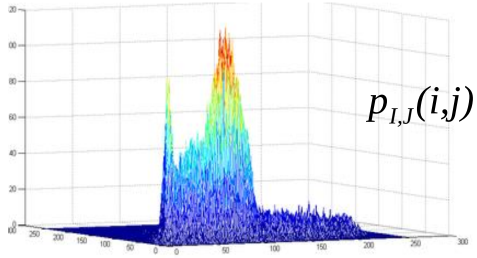
$$\Theta(\omega_r, \omega_s) = \begin{cases} 0 & \text{if } \omega_r = \omega_s \\ +\beta & \text{if } \omega_r \neq \omega_s \end{cases}$$

MTA SZTAKI

Local Similarity Measure Estimation

Cluster Reward Algorithm (CRA) Similarity Measure

$$CRA(I,J,s) = \frac{\sum_{i,j} p_{IJ}^2(i,j,w_s) - \sum_i p_I^2(i,w_s)\sum_j p_J^2(j,w_s)}{\sqrt{\sum_i p_I^2(i,w_s)\sum_j p_J^2(j,w_s)} - \sum_i p_I^2(i,w_s)\sum_j p_J^2(j,w_s)}.$$



$$p_{I,J}(i,j) = \frac{h(i,j)}{\sum_{i,j} h(i,j)}$$

Joint local histogram h(i, j) of the two corresponding windows in images I and J

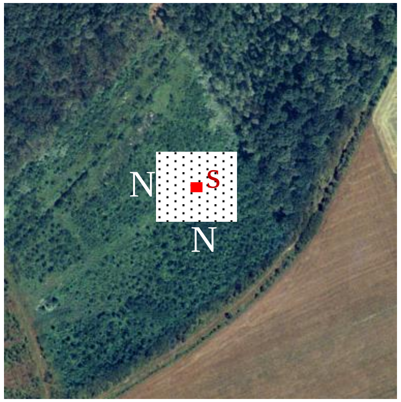


Image I

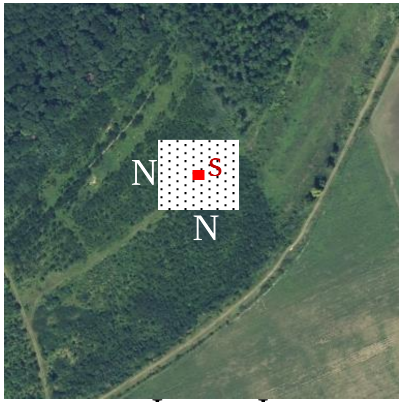
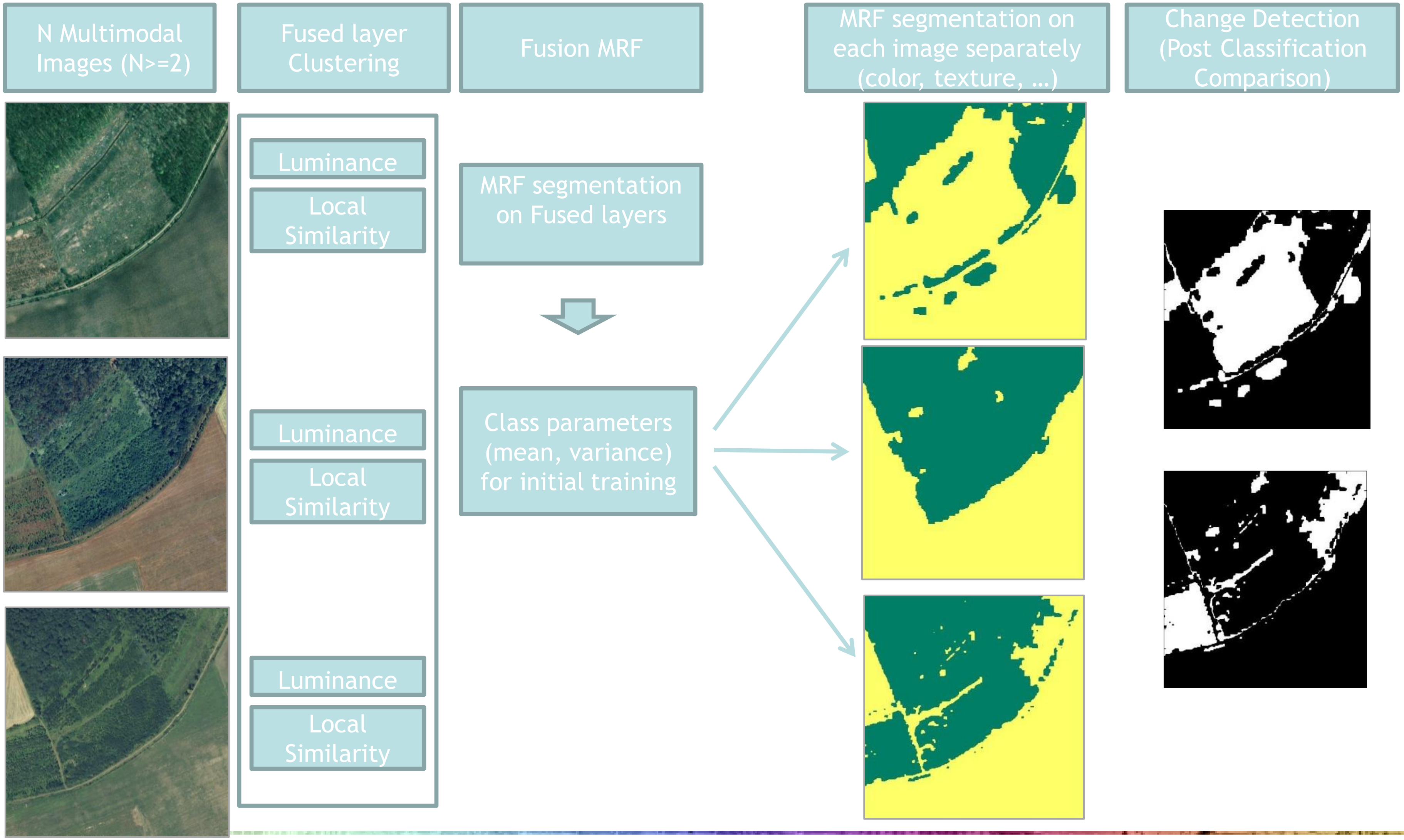


Image J

Problem: Choice of window size N.

- Large window provide better estimation for the pdfs, but can not detect smaller changes.
- Small window gives large estimation error.

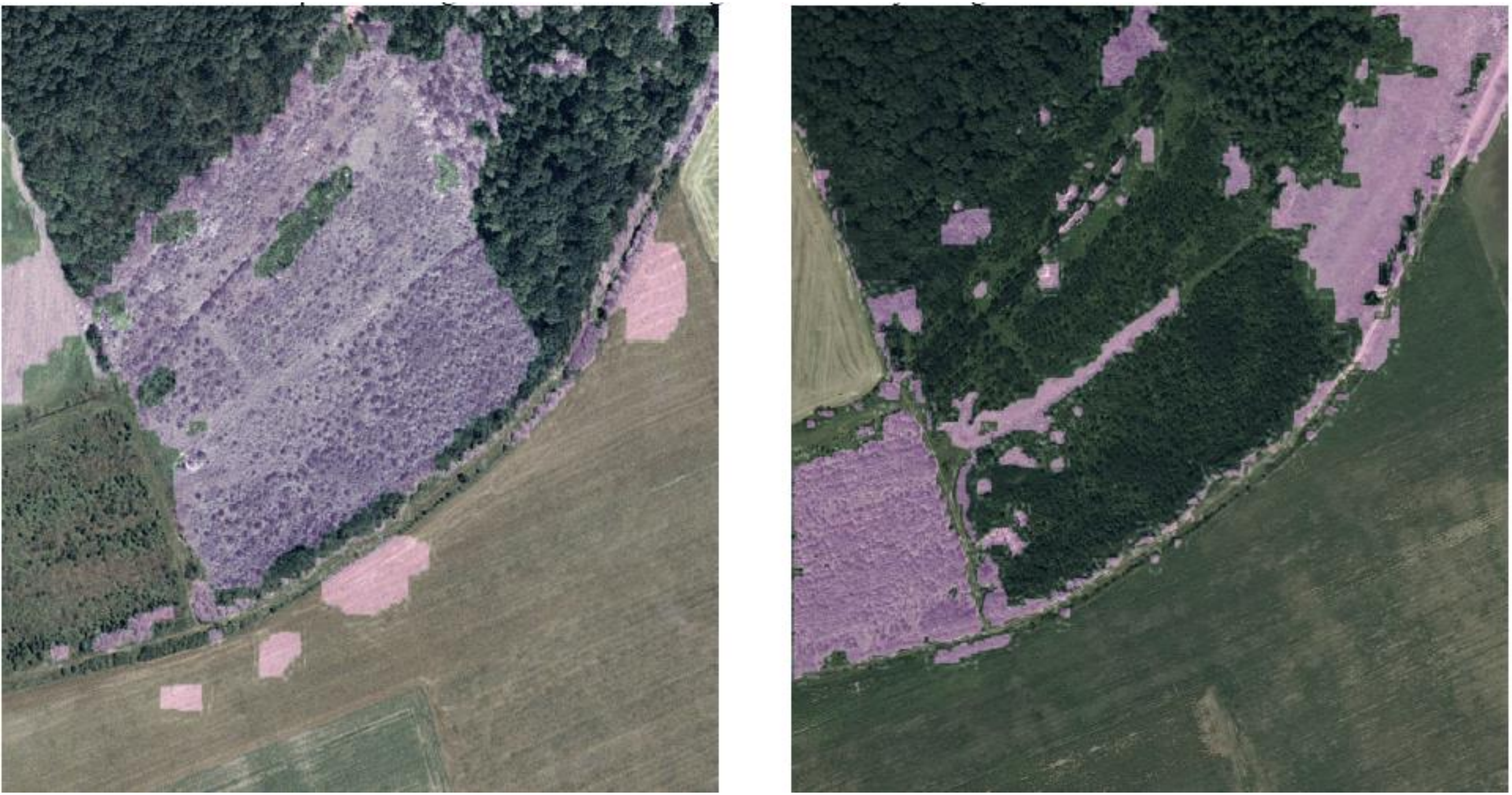
Multi-Layer Fusion MRF Model[1]



[1] *Sziranyi & Shadaydeh, IEEE GSRS Letters, 2014*

Forest / Meadow changes: 2000 - 2005 - 2007

Unsupervised segmentation and change detection by using CRA cross-layer measure and color info



What is in the background?

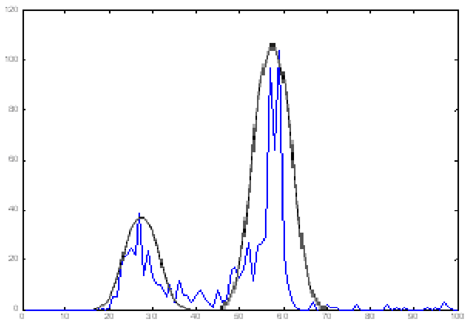
Stochastic and physical models for background discrimination:

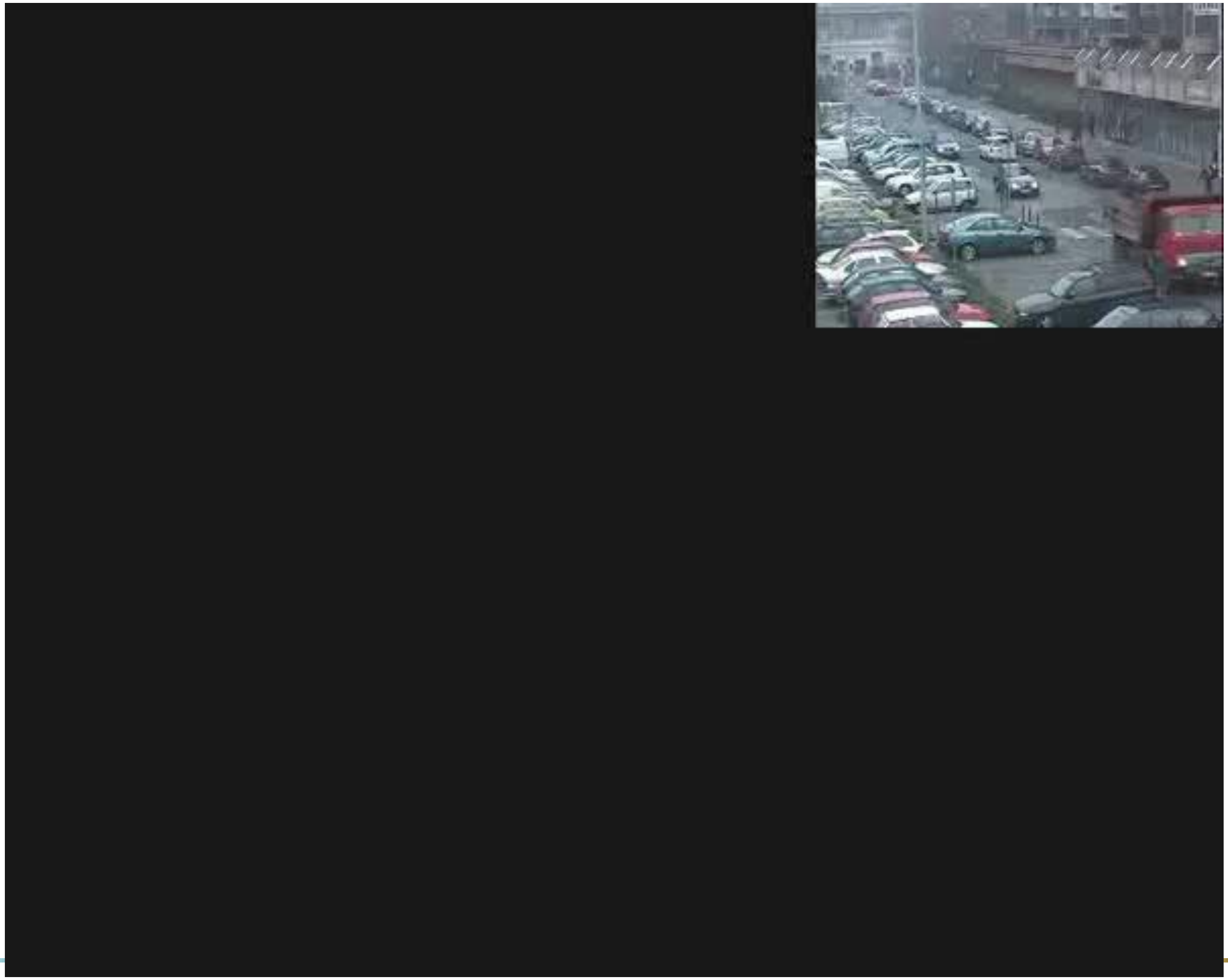
- Gaussian models
- Doppler measurement for passive radars

Background filters

Stauffer and Grimson
IEEE Tr. PAMI, 2000
Benedek and Szirányi,
IEEE Tr. IP, 2008

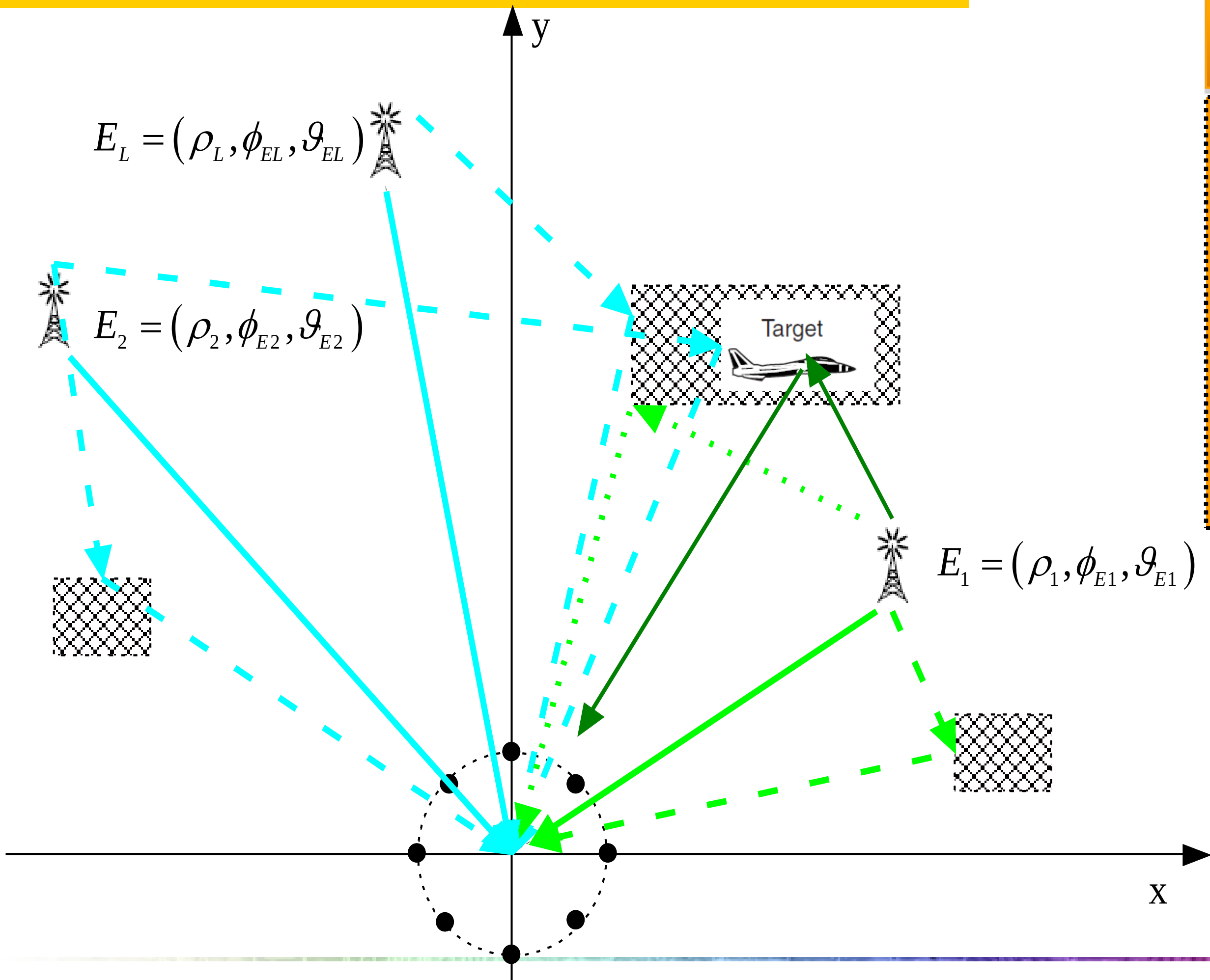
- Mixture of Gaussians
- On-line k-means

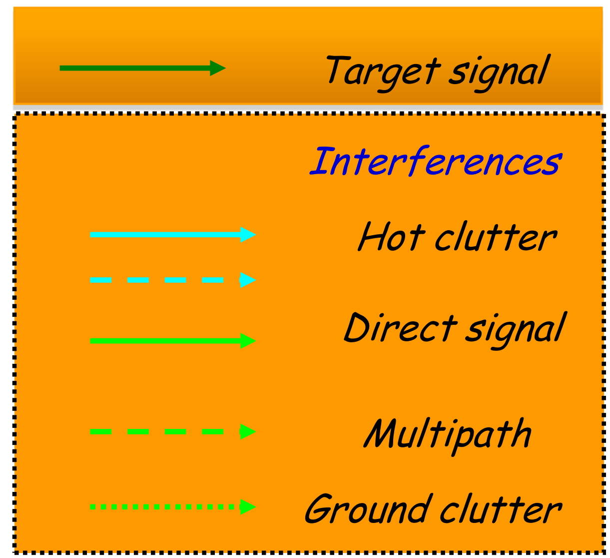




Background filter from „radar” signatures

Passiv reflections
Doppler effect helps to find motion





Array received signal vector

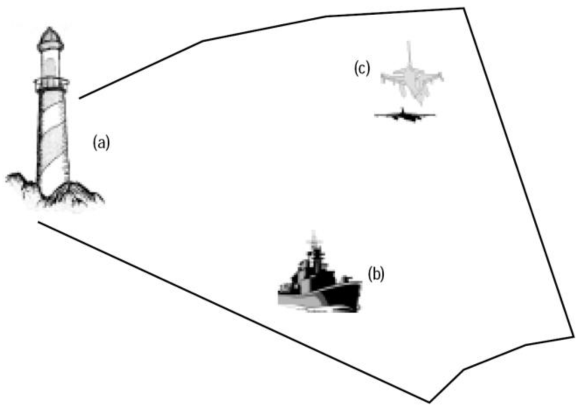
$$s(\mathbf{r}_k, t) = s_r(t) e^{i2\pi f_0 (t-\tau_k)}$$

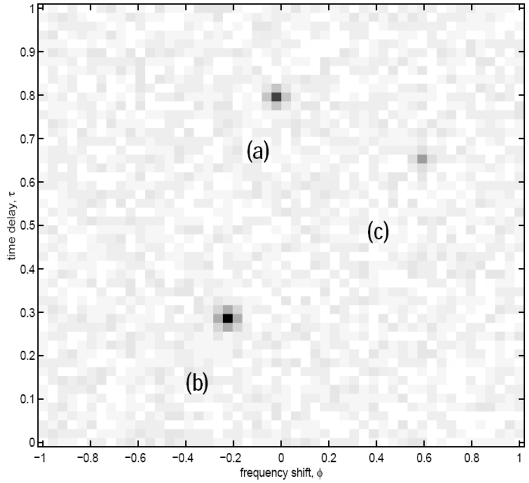
$$\tau_k = \frac{\mathbf{k}(\phi,\vartheta)\mathbf{r}_k}{c} = \frac{\cos(\phi)\sin(\vartheta)x_k + \sin(\phi)\sin(\vartheta)y_k}{c}$$

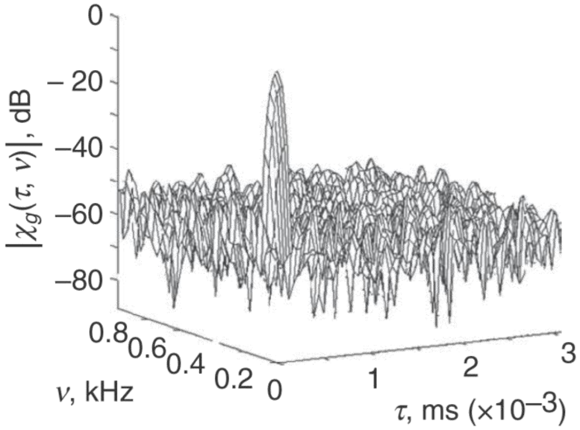
Ambiguity function describes how correlated a signal is with itself delayed in time by τ and shifted in Doppler (Sinsky and Wang, 1974).

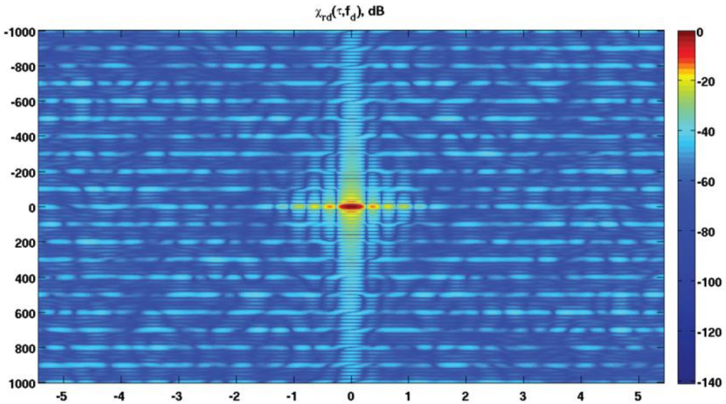
$$|\chi(\tau,\omega_d)|^2 = \left|\int_{-\infty}^{\infty} u(l)e^{j\omega_d l}u^*(l+\tau)\,dl\right|^2$$

Saini, Cherniakov, “DTV signal ambiguity function analysis for radar application”, IEE Proc.-Radar Sonar Navigation 2005









Stochastic methods, correlation models

Bayes iterations

- *Lucy-Richardson algorithm for blind deconvolution*
- Imsage registration through correlation calculus

Lucy-Richardson algorithm for blind deconvolution:

- Given observation *g*, give an estimation of the original image *f* and the blurring function (PSF) *h* :

$$g = f * h$$

- Starting from Richardson's original formula based on Bayesians:

$$P(f_i|g) = [P(g|f)P(f_i)] \,/\, \textstyle\sum_j [P(g|f_j)P(f_j)] \qquad P(f|g) = P(fg)/P(g)$$

$$P(f_i) = \sum_l P(f_i g_l) = \sum_l P(f_i|g_l)P(g_l) = \sum_l \frac{P(g_l|f_i)P(f_i)P(g_l)}{\sum_j P(g_l|f_j)P(f_j)}$$

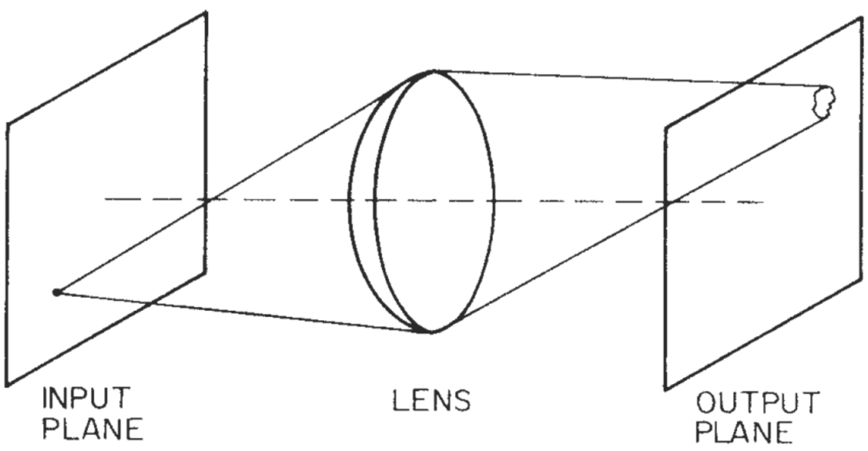
$$P_{k+1}(f_i) = P_k(f_i) \sum_l \frac{P(g_l|f_i)P(g_l)}{\sum_j P(g_l|f_j)P_k(f_j)}$$

Lucy-Richardson algorithm for blind deconvolution:

- Probability is proportional to the luminance
- Conditional probability is prop. to the PSF

$$f_{i,k+1} = f_{i,k} \sum_l \frac{h_{i,l} g_l}{\sum_j h_{j,l} f_{j,k}} = f_{i,k} \sum_l h_{i,l} \frac{g_l}{\sum_j f_{j,k} h_{j,l}}$$

$$f_{k+1} = f_k \left(h_k \otimes \frac{g}{f_k \otimes h_k} \right)$$

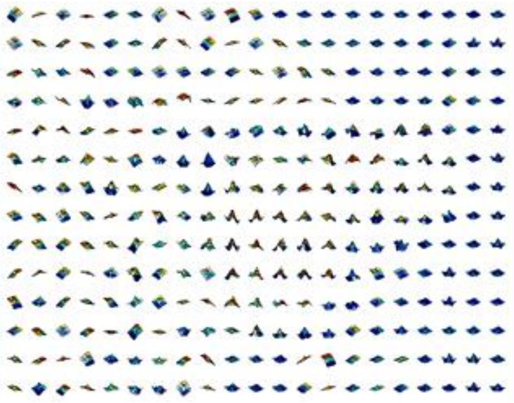


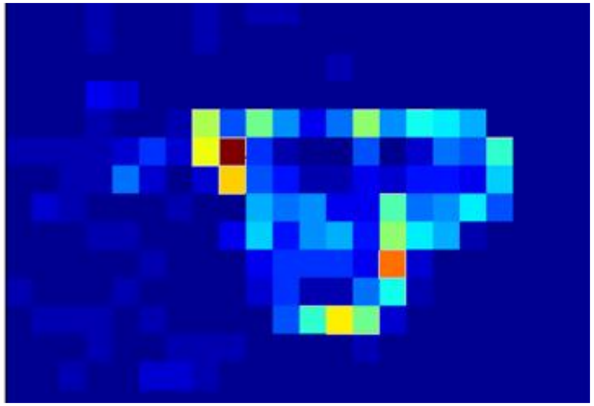
The double iteration

$$\begin{cases} f_{k+1}(r) = f_k(r)\left[h_k(r) * \frac{g}{g_k}(r)\right] \\ h_{k+1}(r) = \frac{h_k(r)}{\gamma}\left[f_k(r) * \frac{g}{g_k}(r)\right] \end{cases}$$

$$g_k = f_k * h_k$$

- f and the PSF vary locally according to the amount of blur (distortion) present on the image locally
- Stop the double iteration at a finite step (here #5) and check the **error between the measured and the estimated blurred image** blocks: $\| g - g_k \|$
- Is MSE usable for comparison the BD residual errors of different blocks?





Comparison of in-focus measures through convergence error

- Comparison of blocks is based on the $\|g - g_k\|$ error distance between the estimated g_k and measured g average values over a sample area
- The greater the error distance, the better in the focus: more effort is needed for convergence from the stopped iteration
- Problem: The residual error makes the comparison unbalanced
- A measure of zero theoretical residual is needed

The independence of signal and noise makes the orthogonality criterion usable as an error measure itself

ADE : angle deviation error
Orthogonality criterion: signal and noise are independent

$$\left| \arcsin \frac{\langle g, g - g_k \rangle}{|g| \cdot |g - g_k|} \right|$$

In case of

$$g - g_k = [\ +1, -1, -1, +1, -1, +1 \ \ldots\ldots\ -1, +1\ ]$$

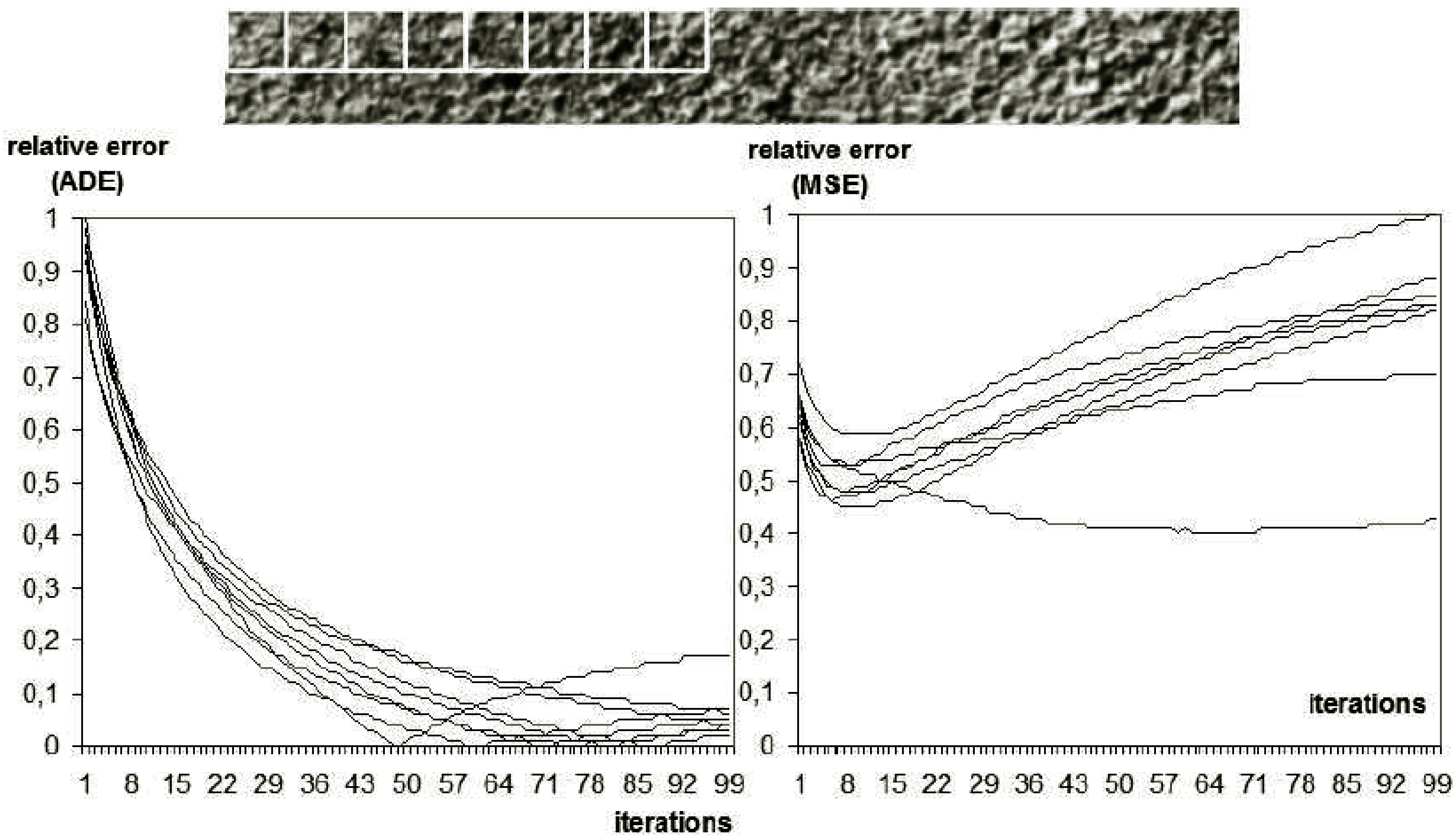
$$g = [10, 10, 10, 10, 10, 10\ \ldots\ldots\ 10, 10]$$

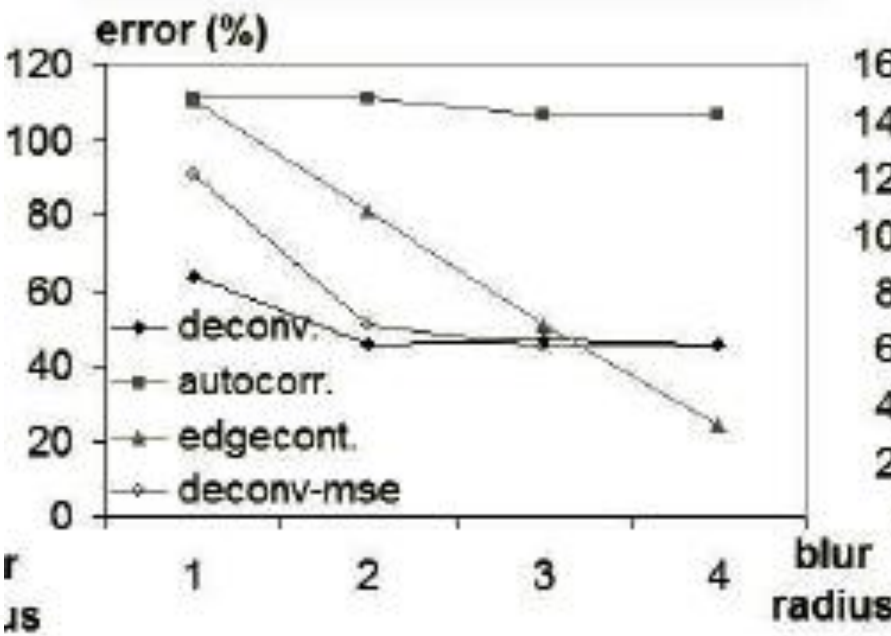
$\| g - g_k \|$ is high, while

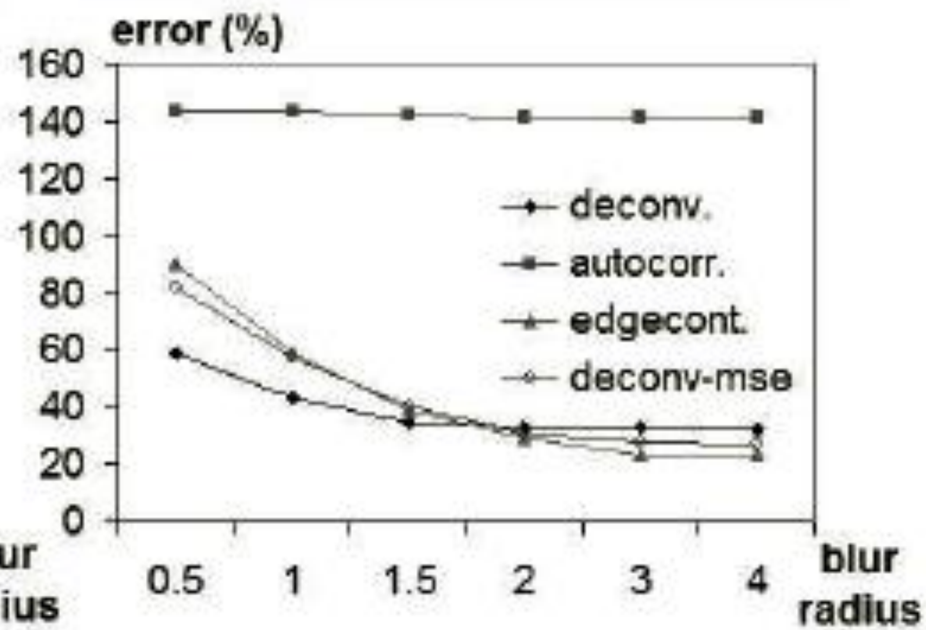
$< g, g - g_k > \ \rightarrow$ zero

Error curves for 8 neighboring blocks (each curve stands for one block) on a blurred texture sample (top) for the same blur with **ADE** (left), and **MSE** (right) .

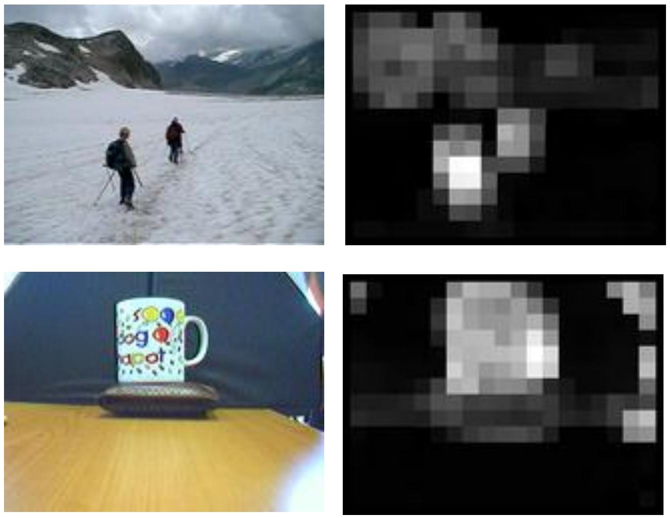
Ideally, curves of the same measure should remain close to each other.

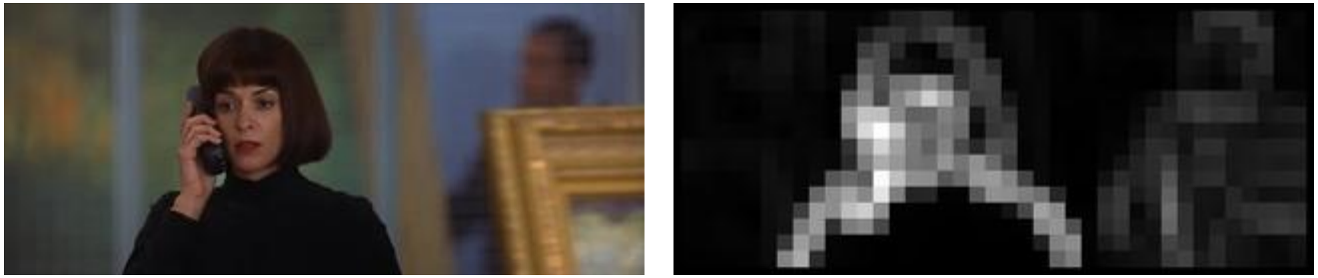



error (%)
120
100
80
60
40
20
0
deconv.
autocorr.
edgecont.
deconv-mse
1
2
3
4
blur radius


error (%)
160
140
120
100
80
60
40
20
0
deconv.
autocorr.
edgecont.
deconv-mse
0.5
1
1.5
2
3
4
blur radius

Find relative focus map:





Stochastic methods, correlation models

Bayes iterations

- Lucy-Richardson algorithm for blind deconvolution
- *Image registration through correlation calculus*

Stereo registration through random motion

Szlavik, Sziranyi, Havasi - IEEE Tr. Image Processing, 2007

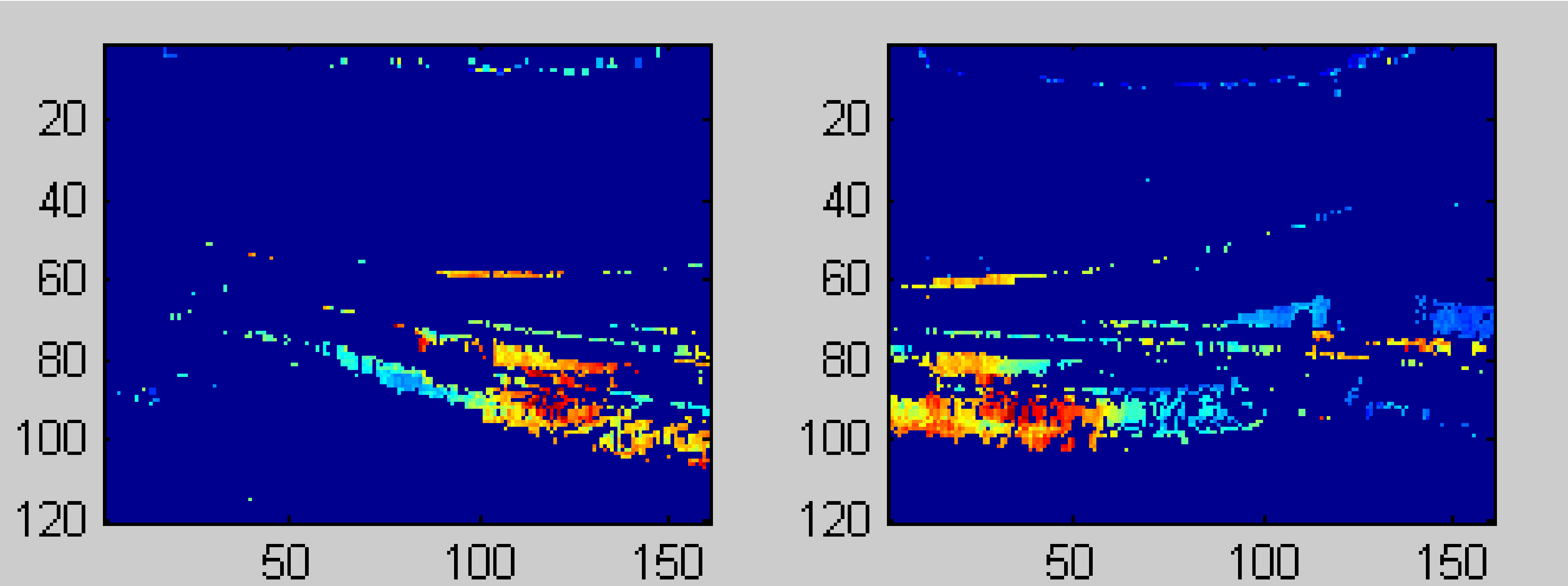
Ergodic regular Markov chain has a unique stationary distribution

$$\left(\underline{p}_1 \quad \underline{p}_2\right) = \left(\underline{p}_1 \quad \underline{p}_2\right)\underline{\underline{\Pi}}$$

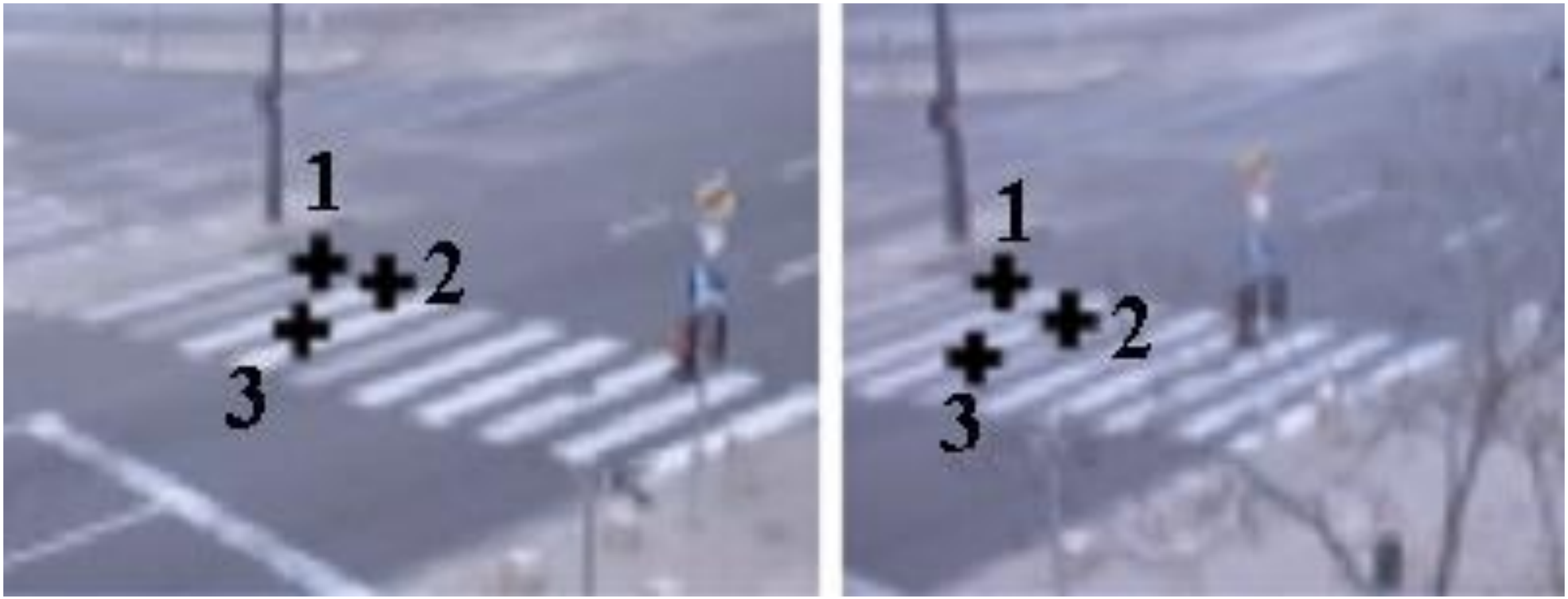
$$P(m_{1i})_{r+1} = P(m_{1i})_r \sum_k \frac{P(m_{2k} \mid m_{1i}) P(m_{2k})_r}{\sum_j P(m_{2k} \mid m_{1j}) P(m_{1j})_r}$$

$$P(m_{2k})_{r+1} = P(m_{2k})_r \sum_i \frac{P(m_{1i} \mid m_{2k}) P(m_{1i})_r}{\sum_j P(m_{1i} \mid m_{2j}) P(m_{2j})_r}$$

Bayesian iterations of Ergodic regular Markov chain with a unique stationary distribution



Point pairs



Multi-view scene reconstruction

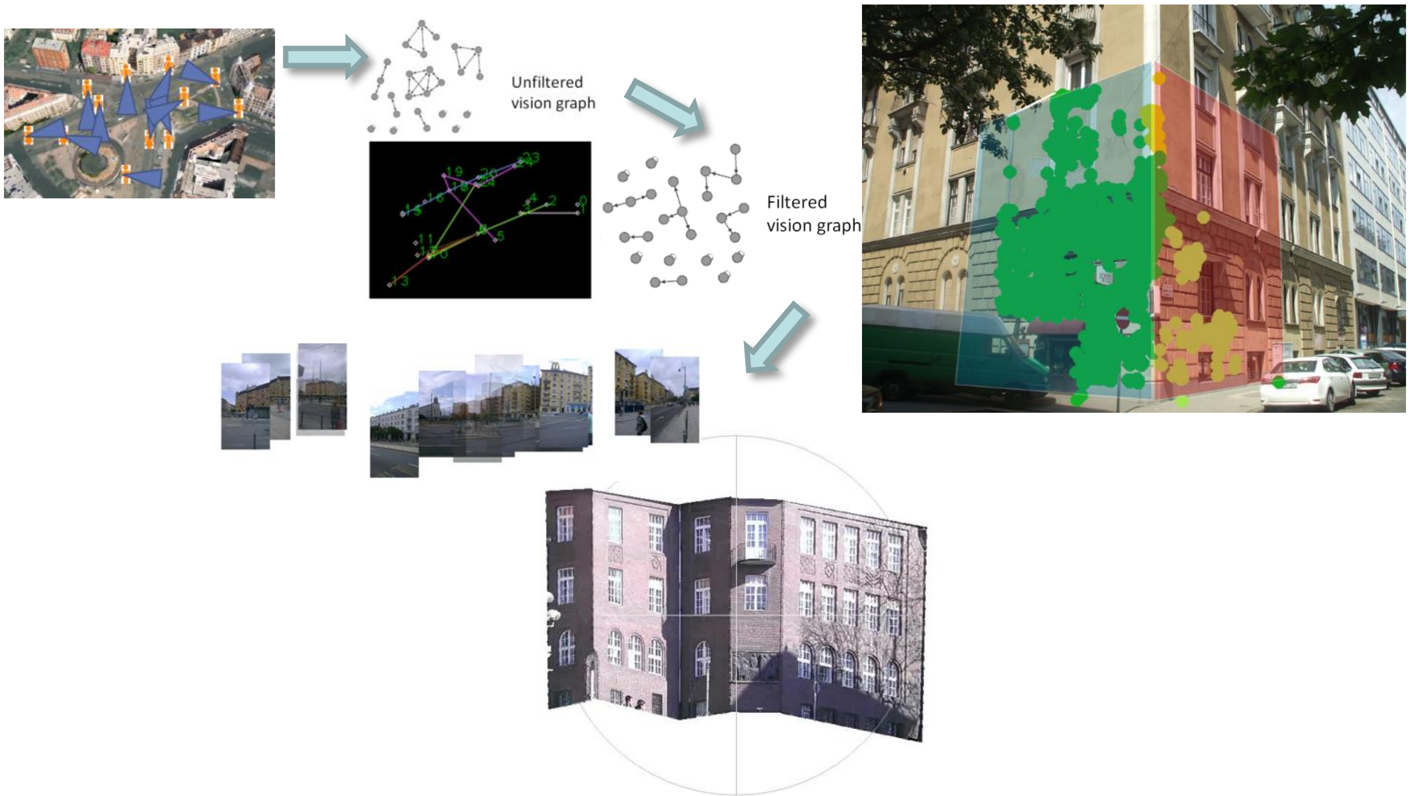
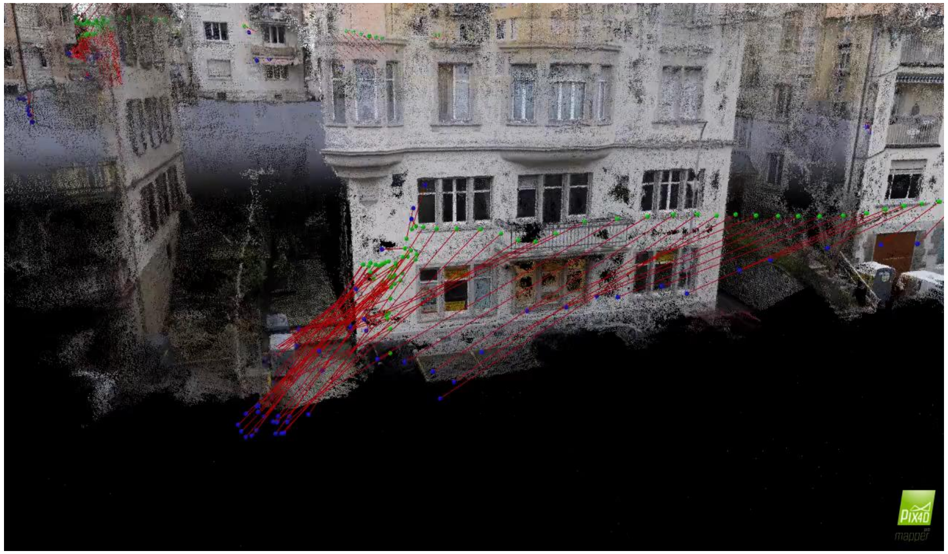


Image based localization in the city

MAV image based 3D
city reconstruction

Mono-camera SLAM
Budapesti Régi Közvágóhíd



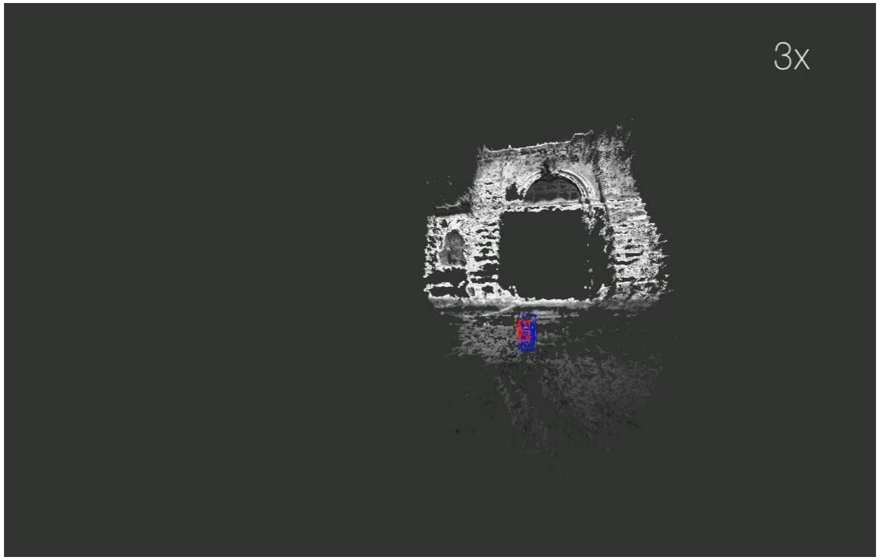


Image based track
GPS position
Error

MAV camera position
Key position
Transition

Thank you for the attention !