Lesson 04

KAZE, Non-linear diffusion filtering, ORB, MSER

Ing. Marek Hrúz, Ph.D.

Katedra Kybernetiky
Fakulta aplikovaných věd
Západočeská univerzita v Plzni

DEPARTMENT OF CYBERNETICS

KAZE

ORB: an efficient alternative to SIFT and SURF

MSER - Maximally stable extremal regions

DEPARTMENT OF CYBERNETICS

KAZE - "wind" in Japanese

- Classical Gaussian scale spaces used in SIFT et.al. have undesirable property of blurring the edges of images
- This lowers the localization precision of key-points
- KAZE introduces a new scheme for creating the scale-space using the **Nonlinear Diffusion Filtering**
- The NDF has a nice property of preserving the image edges
- "NDF describes the evolution of the luminance of an image through increasing scale levels as the divergence of a certain *flow* function that controls the diffusion process." (WTF?!?)

Diffusion

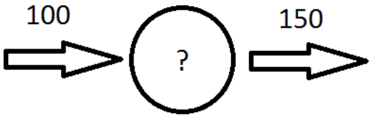
- Diffusion is the net movement of molecules or atoms from a region of high concentration (or high chemical potential) to a region of low concentration (or low chemical potential).
- NDF is normally described by nonlinear partial differential equations

$$\frac{\partial L}{\partial t} = \mathrm{div}\left(c\left(x, y, t\right) \cdot \nabla L\right) \tag{1}$$

- L is the luminance function (the image)
- div is the divergence operator
- $c(x, y, t)$ is a conductivity function

Divergence

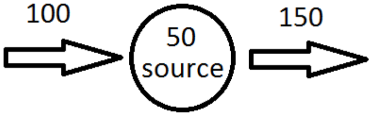
- In vector calculus, divergence is a vector operator that produces a signed scalar field giving the quantity of a vector field's source at each point



DEPARTMENT OF CYBERNETICS

Divergence

- In vector calculus, divergence is a vector operator that produces a signed scalar field giving the quantity of a vector field's source at each point

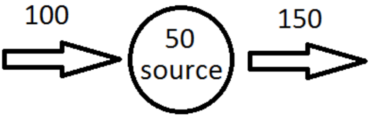


$$\operatorname{div} \mathbf{F}(p) = \lim_{V \to \{p\}} \iint_{S(V)} \frac{\mathbf{F} \cdot \mathbf{n}}{|V|} dS \tag{2}$$

DEPARTMENT OF CYBERNETICS

Divergence

$$\operatorname{div} \mathbf{F}(p) = \lim_{V \to \{p\}} \iint_{S(V)} \frac{\mathbf{F} \cdot \mathbf{n}}{|V|} dS \tag{3}$$



$$\operatorname{div} \mathbf{F}(p) = \nabla \cdot \mathbf{F} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \cdot (U, V, W) = \frac{\partial U}{\partial x} + \frac{\partial V}{\partial y} + \frac{\partial W}{\partial z} \tag{4}$$

DEPARTMENT OF CYBERNETICS

Perona and Malik Diffusion Eq.

- Nonlinear diffusion filtering was introduced in 1990
- $\frac{\partial L}{\partial t} = \mathrm{div}\left(c\left(x, y, t\right) \cdot \nabla L\right)$
- Perona and Malik proposed to make the conductivity function c dependent on the gradient magnitude
- $c(x, y, t) = g\left(\nabla L_\sigma(x, y, t)\right)$
- σ is the blurring parameter - variance of Gaussian
- More functions have been proposed, KAZE uses:
- $g = \exp\left(-\frac{|\nabla L_\sigma|^2}{k^2}\right)$
- k is the contrast parameter – empirical or estimated

DEPARTMENT OF CYBERNETICS

Additive Operator Splitting

- No analytical solution for PDEs - we need a numerical approciamtion = AOS
- Lets assume a 1-D case

$$\frac{\partial u}{\partial t} = \mathrm{div}\left(g\left(\left|\frac{\partial u_\sigma}{\partial x}\right|\right) \cdot \frac{\partial u}{\partial x}\right) = \frac{\partial\left(g\left(\left|\frac{\partial u_\sigma}{\partial x}\right|\right) \cdot \frac{\partial u}{\partial x}\right)}{\partial x} \qquad (5)$$

- The simplest approximation is

$$\frac{u_i^{k+1} - u_i^k}{\tau} = \sum_{j \in N(i)} \frac{g_j^k + g_i^k}{2h^2}\left(u_j^k - u_i^k\right) \qquad (6)$$

- Where u is the image i, j are the locations, τ is the time difference, h is the grid size, k is the time, $N(i)$ is the neighbourhood of location i

Notation

- Gradient is approximated by central differences

$$g_i^k = g\left(\frac{1}{2}\sum_{p,q\in N(i)}\left(\frac{u_p^k - u_q^k}{2h}\right)^2\right) \tag{7}$$

- The matrix notation: $\frac{u^{k+1}-u^k}{\tau} = A(u^k)u^k$
- And matrix A has entries:

$$a_{ij}(u^k) := \begin{cases} \frac{g_i^k + g_j^k}{2h^2} & (j \in \mathcal{N}(i)), \\ -\sum_{n\in\mathcal{N}(i)} \frac{g_i^k + g_n^k}{2h^2} & (j = i), \\ 0 & (\text{else}). \end{cases}$$

- This yields a set of linear equations which can be solved

DEPARTMENT OF CYBERNETICS

Scheme

- By arranging the expression we obtain
- $u^{k+1} = \left(I + \tau A(u^k)\right) u^k$
- This is the explicit scheme (restrictions on step size $=>$ slow)
- KAZE uses so-called semi-explicit scheme which is
- $\frac{u_i^{k+1} - u_i^k}{\tau} = A(u^k) u^{k+1}$
- $\left(I - \tau A(u^k)\right) u^{k+1} = u^k$ which has no explicit solution
- We have to solve as:

$$u^{k+1} = \left(I - \tau A(u^k)\right)^{-1} u^k \tag{8}$$

- In the KAZE paper notation: $L^{i+1} = \left(I - \tau \sum_{l=1}^{m} A_l(L^i)\right)^{-1} L^i$
- Thu sum in the expression reflects that there are more direction in an image

DEPARTMENT OF CYBERNETICS

KAZE scale space, detector, descriptor

- Contrast parameter k: 70% percentile of the gradient histogram of a smoothed version of the original image
- Scale σ: has octaves and sub-levels as SIFT
- $\sigma_i(o, s) = \sigma_0 2^{o+s/2}$
- Time step t: is a mapping from scale
- $t_i = \frac{1}{2}\sigma_i^2$
- Detector as in SIFT
- Descriptor as in SURF

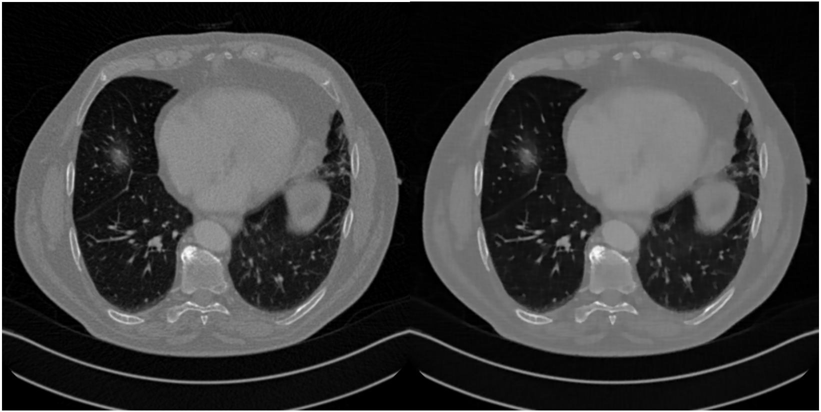
Comparison of scale-spaces

DEPARTMENT OF CYBERNETICS

Comparison of scale-spaces

DEPARTMENT OF
CYBERNETICS

Example of filtered image



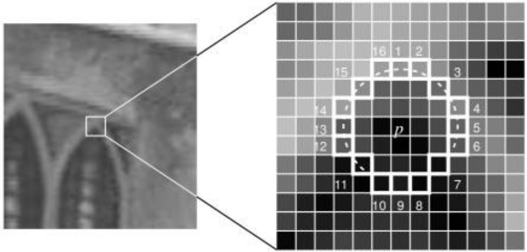
ORB: an efficient alternative to SIFT and SURF

- SIFT and SURF - impose large computational burden (speed and energy)
- ORB - Oriented FAST and Rotated BRIEF
- FAST - Features from Accelerated Segment Test - paper "Machine learning for high-speed corner detection" in 2006, revisited in 2010
- BRIEF - Binary Robust Independent Elementary Features - paper "BRIEF: Binary Robust Independent Elementary Features" in 2010

DEPARTMENT OF CYBERNETICS

FAST

- Algorithm for keypoint detection:
- Select a pixel p with intensity I_p
- Select appropriate threshold t
- Consider a circle of 16 pixels around the pixel under test



- A point p is an interest point (in this case a corner) if there are n contiguous pixels brighter than $I_p + t$ or darker than $I_p - t$
- High-speed test - check only 4 pixels (1, 9, 5, 13), if at least 3 are brighter/darker

DEPARTMENT OF CYBERNETICS

BRIEF

- It is a binary descriptor of interest points
- It selects pairs of pixels of a smoothed image around an interest point and compares them binary
- If $I(p) < I(q)$ then the result is 1, else it is 0
- The selection of the pixel pairs will be explained later (details also in the paper)
- Based on the number of pixel pairs an n-dimensional binary feature vector is obtained which serves as the descriptor of the interest point
- Distance of the descriptors is calculated as a Hamming distance (which is a XOR operator and bit count)

Oriented FAST

- The location of keypoints is detected using FAST-9 algorithm (circular radius of 9 px)
- Since FAST has large response on edges the oFAST computes a Harris corner measure for each keypoint
- The keypoints are ordered according to this measure of cornerness
- Furthermore a pyramid of the image is built and thus multi-scale FAST keypoints are detected
- N best keypoints are considered (N is defined by the user)

Oriented FAST - Orientation computation

- Orientation is computed as *intensity centroid*

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y) \tag{9}$$

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \tag{10}$$

- A vector from corner's center O to the intensity centroid C is computed
- the orientation is simply:

$$\theta = \mathrm{atan2}(m_{01}, m_{10}) \tag{11}$$

- To obtain better results the C is computed only from pixels in a circular region with radius r

Classical BRIEF - details

- A test is defined over a smoothed image patch p of size $S \times S$ as:

$$\tau(p; x, y) = \begin{cases} 1 : p(x) < p(y) \\ 0 : p(x) \geq p(y) \end{cases} \tag{12}$$

- Binary feature vector (descriptor) is constructed as

$$f_n(p) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i, y_i) \tag{13}$$

- How to choose the pairs of pixel for the test τ?
- The pixels x and y are sampled independently from a Gaussian distribution centered on the analyzed patch center with a variance of $\frac{1}{25} S^2$
- The smoothing is achieved using an integral image, where each test point is a 5×5 sub-window of a 31×31 pixel patch

DEPARTMENT OF CYBERNETICS

Steered BRIEF

- Steered BRIEF respect the found orientation θ
- We define a matrix S to represent the binary tests

$$S = \begin{pmatrix} x_1, \dots, x_n \\ y_1, \dots, y_n \end{pmatrix} \tag{14}$$

- Using the patch orientation θ we construct a rotation matrix R_θ

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{15}$$

- New test locations are computed as

$$S_\theta = R_\theta S \tag{16}$$

- In praxis the test locations are pre-generated
- Orientations are discretized into 12 degree regions
- The pre-generated test locations are pre-computed into the $360/12 = 30$ different orientations and stored into a look-up-table

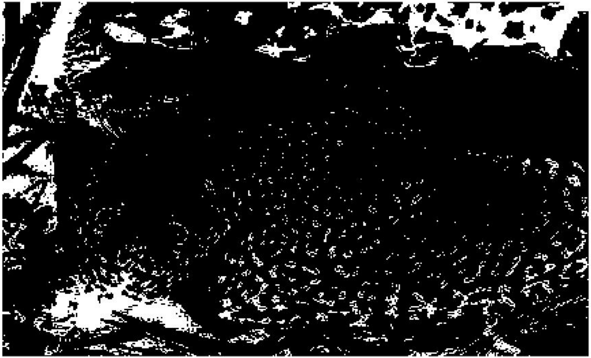
rBRIEF - rotation aware BRIEF

- Experiments show that Steered BRIEF generate descriptors that are highly correlated (unwanted!)
- To recover from this an algorithm which yields rBRIEF is presented:
- Enumerate all possible binary tests (in their case it's 205590 tests)
- Take a lot of images, detect keypoints and run each test against all training patches
- Order the tests by their distance from a mean of 0.5, forming a vector T
- Greedy search:
  - Put the first test into the result vector R and remove it from T
  - Take the next test from T and compare it against all tests in R. If its absolute correlation is greater than a threshold, discard it; else add it to R
  - Repeat the previous step until there are 256 tests in R. If there are fewer than 256, raise the threshold and try again

DEPARTMENT OF CYBERNETICS

MSER - Maximally stable extremal regions

- Image I is a mapping $I : \mathrm{D} \subset \mathrm{Z}^2 \rightarrow \mathrm{S}$. External region can be defined if:
  1. S is fully ordered
  2. adjacency exists $\mathrm{A} \subset \mathrm{D} \times \mathrm{D}$
- Region Q is a connected subset from D - for every $p, q \in \mathrm{Q}$ there exists a sequence $p, a_1, a_2, \ldots, a_n, q$ which fulfilsl $pAa_1, a_1Aa_2, \ldots, a_nAq$
- (Outer) boundary of a region is defined $\partial\mathrm{Q} = \{q \in \mathrm{D}\backslash\mathrm{Q} : \exists p \in \mathrm{Q} : qAp\}$
- Extremal region $\mathrm{Q} \subset \mathrm{D}$ is a region which fulfills that for every $p \in \mathrm{Q}, q \in \partial\mathrm{Q} : I(p) > I(q)$ or $I(p) < I(q)$
- **Maximally stable extremal region (MSER)**. Let $\mathrm{Q}_1, \ldots, \mathrm{Q}_{i-1}, \mathrm{Q}_i, \ldots$ be a sequence of nested extremal regions (ie. $\mathrm{Q}_i \subset \mathrm{Q}_{i+1}$). Extremal region Q_{i*} is maximally stable if:
  - $q(i) = \frac{|\mathrm{Q}_{i+\Delta}\backslash\mathrm{Q}_{i-\Delta}|}{|\mathrm{Q}_i|}$ has a local minimum in $i*$ ($|.|$ means cardinality). $\Delta \in \mathrm{S}$ is the parameter of the method.



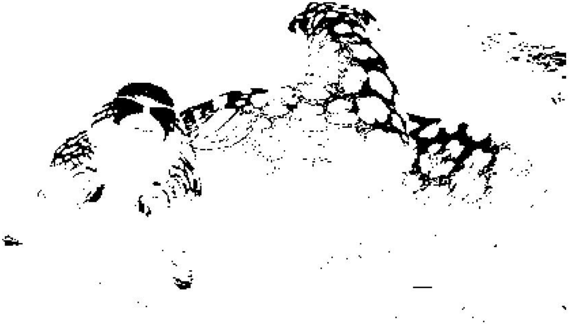
DEPARTMENT OF CYBERNETICS

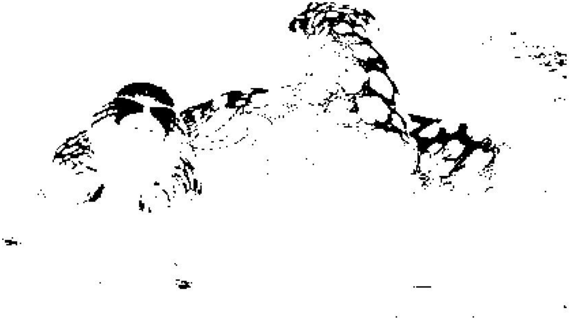
DEPARTMENT OF
CYBERNETICS

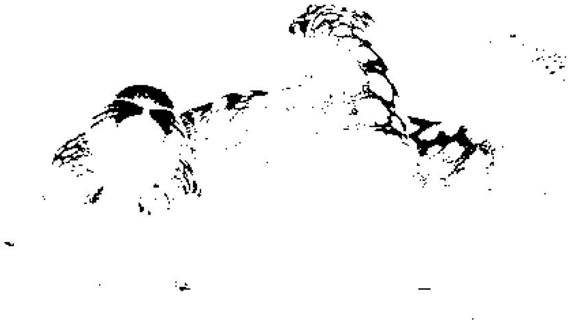
DEPARTMENT OF CYBERNETICS

DEPARTMENT OF
CYBERNETICS

DEPARTMENT OF
CYBERNETICS







DEPARTMENT OF CYBERNETICS