Zpracování digitalizovaného obrazu (ZDO) - Popisy II

Další přístupy popisů oblastí

Ing. Zdeněk Krňoul, Ph.D.

Katedra Kybernetiky
Fakulta aplikovaných věd
Západočeská univerzita v Plzni

Podpořeno: ESF projekt Západočeské univerzity v Plzni
reg. č. CZ.02.2.69/0.0/0.0/16_015/0002287

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS

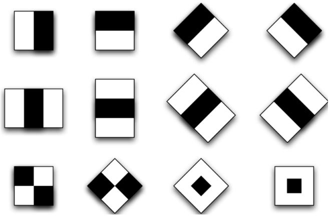
Obsah:

- POPISY II
  - Haar-like
  - LBP
  - HoG

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
DEPARTMENT OF CYBERNETICS

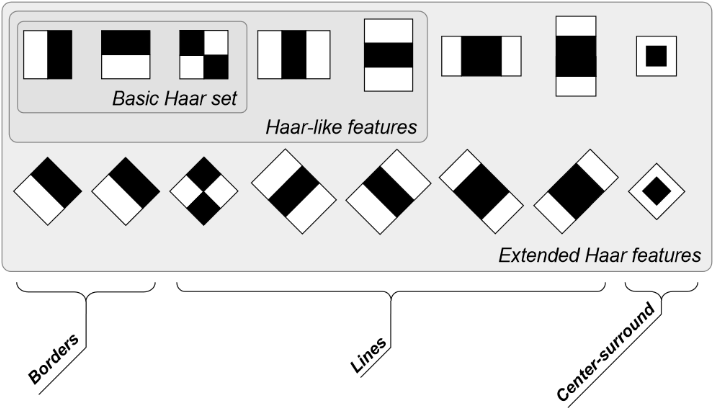
Haar-like popis - **Haar-like features**

- Vycházejí z Haar walelets
- Byly použity v prvním **real-time** face detektoru (Viola–Jones, r. 2001)
- Princip spočívá v přilehlých obdélníkových oblastech ve specifické pozici detekčního okénka

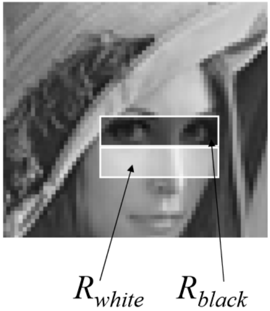


Haar-like features

- Haar-like features are used for image description similar to cosine transform



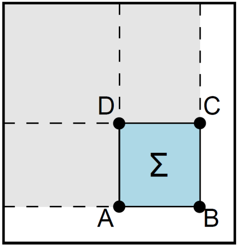
- the black regions are subtracted from the white regions
- the Haar-like filters can be computed efficiently by using integral image



$$F_{Haar} = E(R_{white}) - E(R_{black})$$

- integral image computation

$$I_{\Sigma}(x,y) = \sum_{i=0}^{i\leq x} \sum_{j=0}^{j\leq y} I(i,j) \tag{1}$$



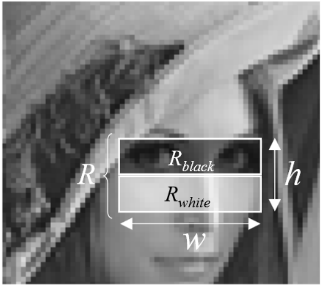
Σ = B - A - C + D

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS

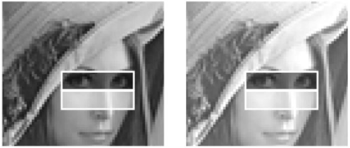
Normalization (monotonic illumination changes)

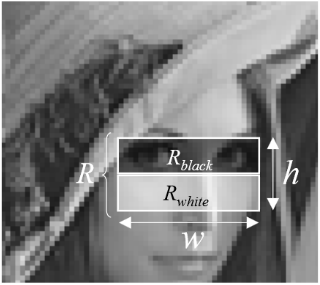


$$F_{Haar} = \frac{E(R_{black}) - E(R_{white})}{\underbrace{\sqrt{\left|E(R_{\mu})^2 - E(R_{\mu}^{\ 2})\right|}}_{\textit{Standard Deviation}}}$$

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
DEPARTMENT OF CYBERNETICS

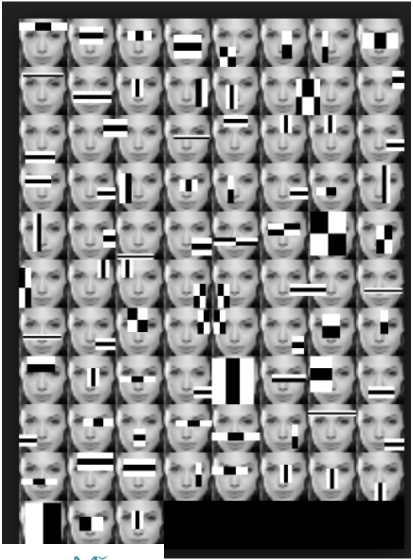
Normalization (monotonic illumination changes and scale)





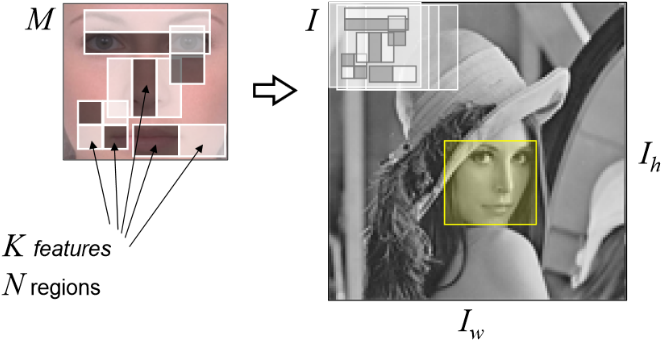
$$F_{Haar} = \frac{E(R_{black}) - E(R_{white})}{w \cdot h \cdot \sqrt{\left| E(R_{\mu})^2 - E(R_{\mu}{}^2) \right|}}$$

- Pro přesný a robustní popis nějaké textury/objektu se používá velký počet Haar-like features najednou



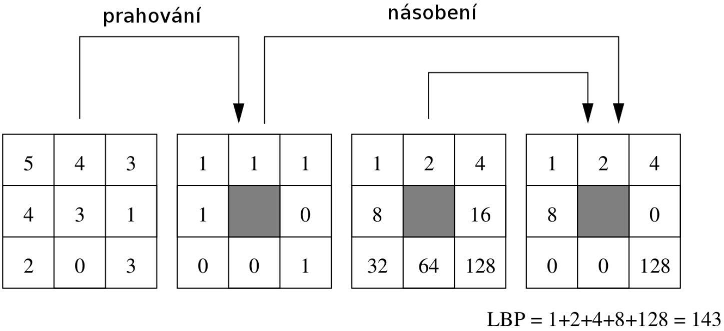
Face detection

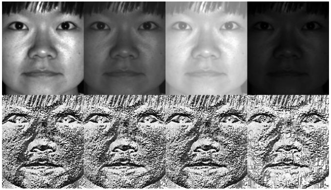
- a boosted classifier is used to train the right responses to certain (most informative) filters
- a sliding window in different scales is used to compute responses on different Haar filters



Local Binary Patterns - LBP

- Ojala 1996: *How to describe texture around a pixel with one scalar?*
- the basic version uses the 8-neighborhood of a pixel
- from this neighborhood a binary representation is build





for a given image patch a **histogram of LBP** codes is constructed and used as a feature:

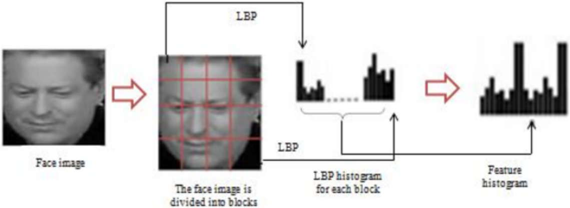


Fig 3: Face description using LBP operator

Rajeshwari et al. "Video Based Face Recognition Using Gabor Features and LBP under Varying Illumination or Pose." (2015)

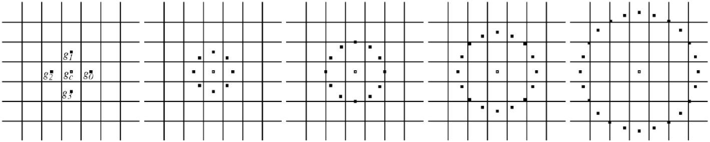
Extensions of LBP

- the basic LBP is invariant to brightness and contrast changes
- they are **variant with scale and rotation** - this is an issue
- texture description:

$$T = t(g_c, g_0, \ldots, g_{P-1}) \tag{2}$$

- the position of pixels in the neighborhood is defined as:

$$g_p = (-R\sin(2\pi p/P), R\cos(2\pi p/P)) \tag{3}$$



(P=4,R=1.0) (P=8,R=1.0) (P=12,R=1.5) (P=16,R=2.0) (P=24,R=3.0)

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \tag{4}$$

Rotation invariance and uniformity

- Rotation invariance is achieved by rotating the local neighborhood

$$LBP_{P,R}^{ri} = \min\left(ROR(LBP_{P,R}, i) | i = 0, 1, \ldots, P-1\right) \quad (5)$$

- ROR is a bitwise rotation operator to find minimal number
- **uniform patterns** - are patterns with at most 2 changes between 0 and 1
- there are a total of 58 uniform patterns, while the rest are put into 59^{th} bin

Histogram of Oriented Gradients - HoG

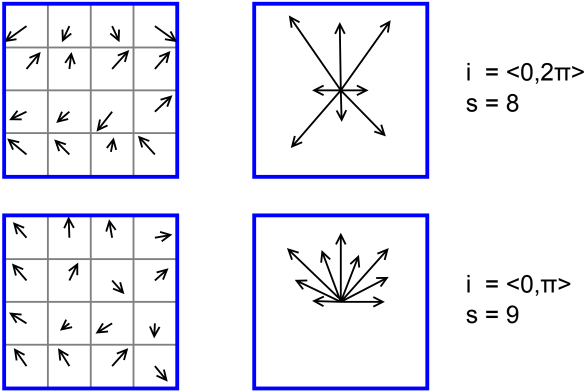
- method for describing images via histogram analysis
- the results of the method are directly dependent on the gradient operator, many had been tested
- the best results were obtained for simple gradient approximation

$$I_x = I * [-1, 0, 1]$$
$$I_y = I * [-1, 0, 1]^T \quad (6)$$

- for every pixel the size and orientation of the gradient is computed
- a histogram is constructed from these values

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
DEPARTMENT OF CYBERNETICS

- the histogram is parametrized by interval i and number of sectors s
- the interval i is mostly $i =< 0, \pi >, i =< 0, 2\pi >$
- the magnitude of the histogram is added to each bin and moreover bilinearly distributed into neighboring bins



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
DEPARTMENT OF CYBERNETICS

Histogram Normalization

- the normalization is useful to cope with brightness transformations
- the most used normalizations are

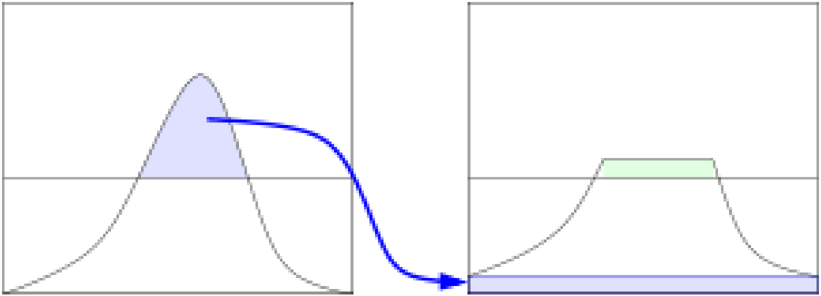
$$L^1 - norm = \frac{v}{(\|v\|_1 + e)}, \tag{7}$$

$$L^1 - sqrt = \sqrt{\frac{v}{(\|v\|_1 + e)}}, \tag{8}$$

$$L^2 - norm = \frac{v}{\sqrt{(\|v\|_2^2 + e^2)}}, \tag{9}$$

- v is the histogram to be normalized and e is a small constant
- a special case of normalization $L_2 - Hys$ - the normalized vector is clipped as in CLAHE (Contrast Limited Adaptive Histogram Equalization)[1]

[1] více v předmětu MPV

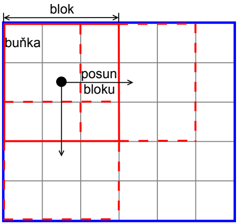


- the values (after normalization) above a given threshold are distributed into all the bins
- the process is repeated until no value is above the threshold

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
DEPARTMENT OF CYBERNETICS

HoG descriptor

- the image is divided into blocks of size (k, k)
- individual blocks are divided into cells of size (l, l)



- for each cell the normalized histograms of gradients are computed
- for each block the cell histograms are averaged
- then the block shifts by some pixels and the process is repeated
- the averaged block histograms are concatenated to obtain the

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
DEPARTMENT OF CYBERNETICS

A co dál?

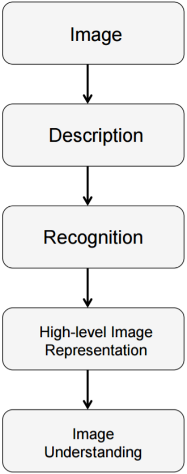
- Existují další popisovače založené na jasových vlastnostech oblastí
- často se prolínají s technikou segmentace, tedy používám detektor + descriptor např. pro key point matching
- BRIEF, DAISY, ORB, BRISK, SIFT, SURF, KAZE, AKAZE[2]
- CNN a end-to-end techniky (příště)

[2] více v předmětu MPV

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
DEPARTMENT OF CYBERNETICS

Shrnutí:

Obvyklý postup operací v úlohách počítačového vidění:



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
DEPARTMENT OF CYBERNETICS

Shrnutí:

Víme, že existují dva základní přístupy reprezentace oblastí tj. popis projekce skutečného předmětu ve 2D obraze:

- externí charakteristiky (hranice) - popis zaměřen pouze na tvar
- interní charakteristiky (vnitřní oblast) - popis zaměřen na výplň oblasti ... barvu, texturu apod.

Některé popisy mohou být invariantní k posunu rotaci a změně měřítka a z hlediska klasifikace musíme rozhodnout zda:

- je tato vlastnost vyžadována
- naopak je potřeba tyto vlastnosti v popisu zachytit

Často se popisy vzájemně kombinují, včetně kombinace popisu tvaru a oblasti (zmíníme v MPV)

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
DEPARTMENT OF CYBERNETICS