# Person Tagging in Still Images by Fusing Face and Full-body Detections

**Vlastislav Dohnal** (dohnal@fi.muni.cz) and Alexander Matecny

DISA Lab – disa.fi.muni.cz

Masaryk University, Brno, Czech Republic

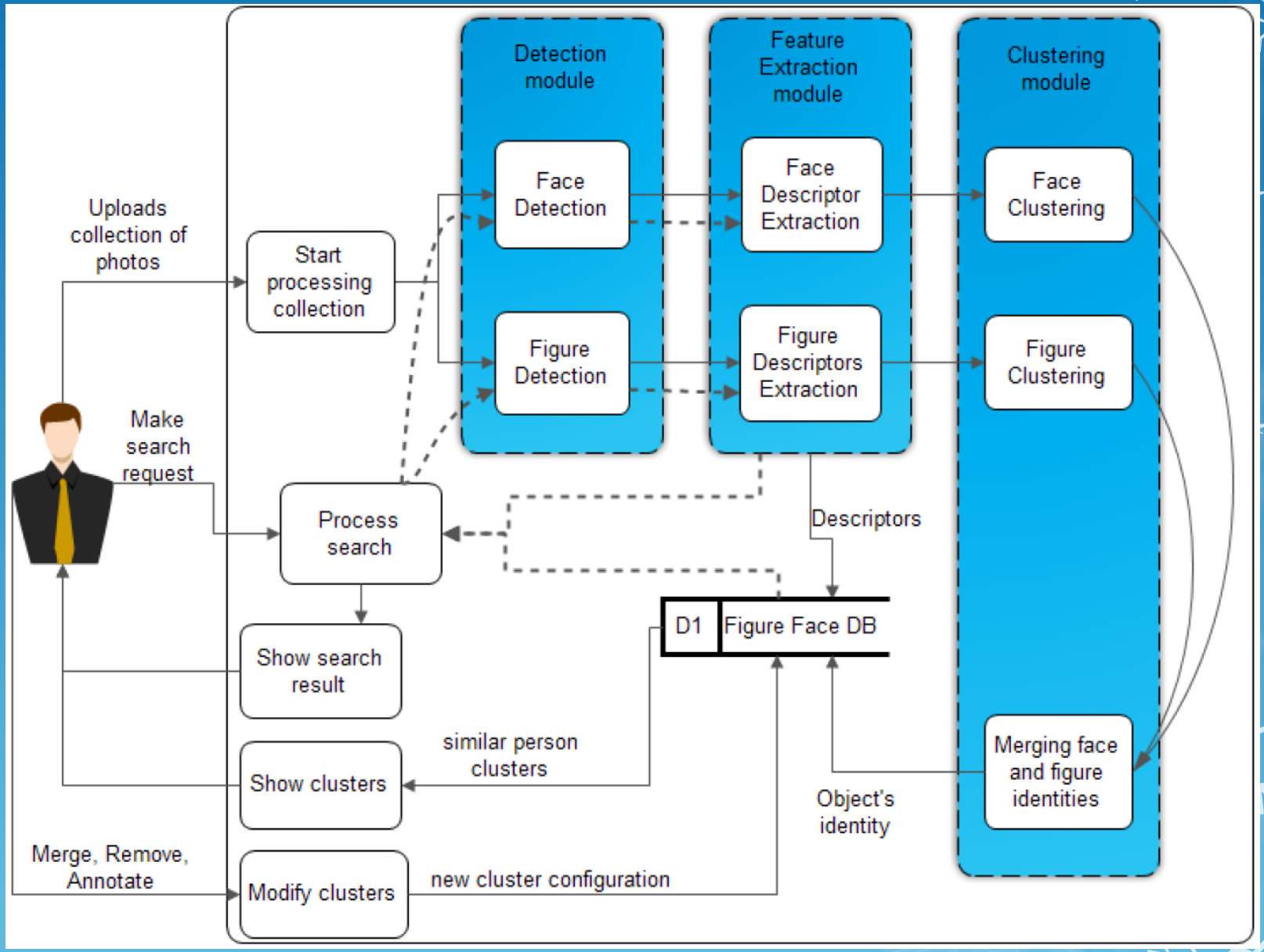
# Outline

- Motivation
- System architecture
  - Figure recognition
- Experimental results
- Conclusions

# Motivation

- Application: person tagging in personal photo albums
  - State of the art: Face recognition is used
- Figure/body detectors exist
  - Based on detecting person head, or
  - full-body/torso detector
    - But typically applied to video signal
- Aim: tag people that do not have their face detected
  - Tag such people using figure recognition, i.e. figure detector applied to still images

# Framework for Person Recognition



# Prototypical Implementation

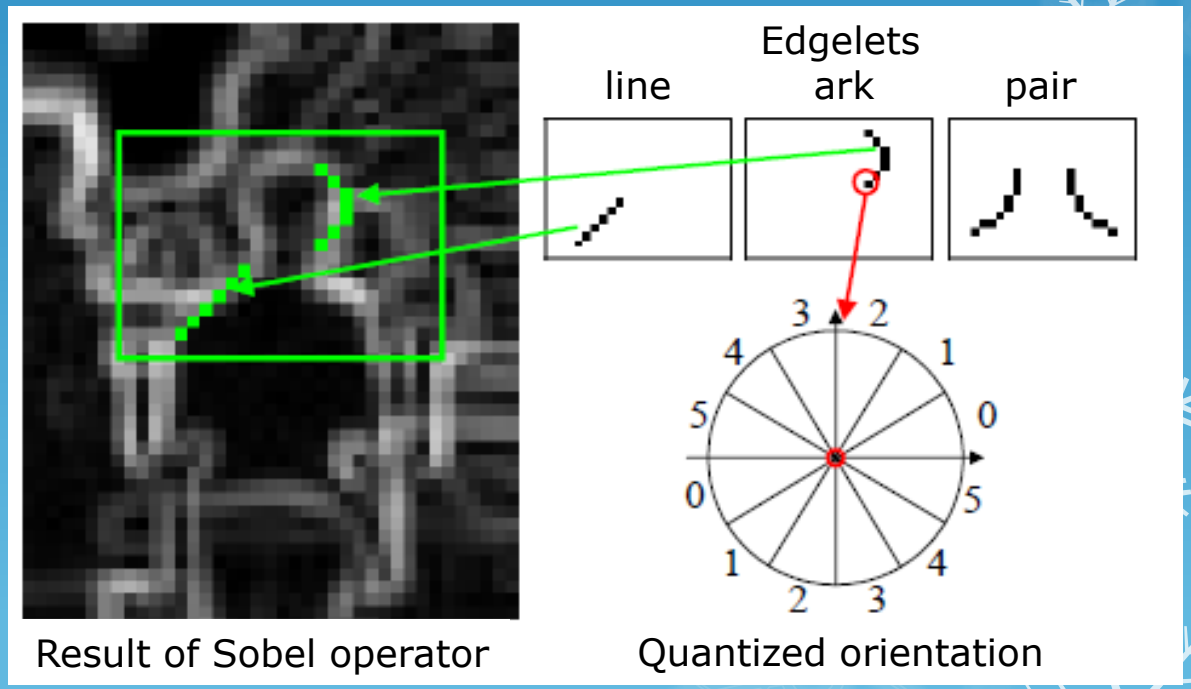
- Detection module
  - Face detector: Luxand SDK
  - Figure detector: Edgelets [1]
- Extraction module
  - Face descriptor: Luxand SDK
  - Figure descriptor:
    - Obtain clothing patch region
    - MPEG-7 visual descriptor
- Clustering module
  - Objective function based on maximum distance within cluster members



[1] B. Wu and R. Nevatia: *Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors*. ICCV 2005.

# Figure Detection: Edgelets

- Detect edges
- Match edgelets (lines, arks, symmetrical pairs)
- Group edgelets to identify a body part
  - By AdaBoost

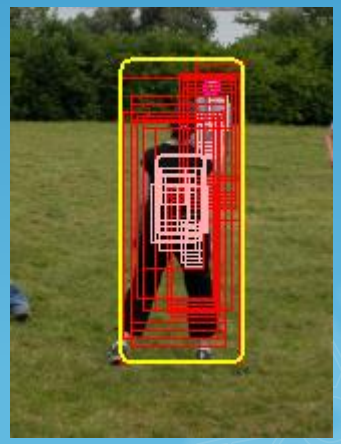
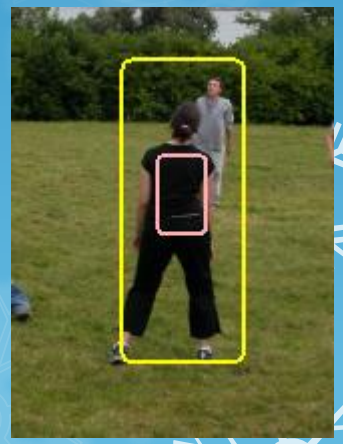


# Figure Detection: Merging & Patch

- Merge detections

$$\frac{area\left(R_{body1} \cap R_{body2}\right)}{min\left(area\left(R_{body1}\right), area\left(R_{body2}\right)\right)} \geq 0.72$$

- Clothing patch
  - Defined by shrinking the figure detection region
    - Width [0.30; 0.70], height [0.32; 0.58]

# Figure Detection: Quality

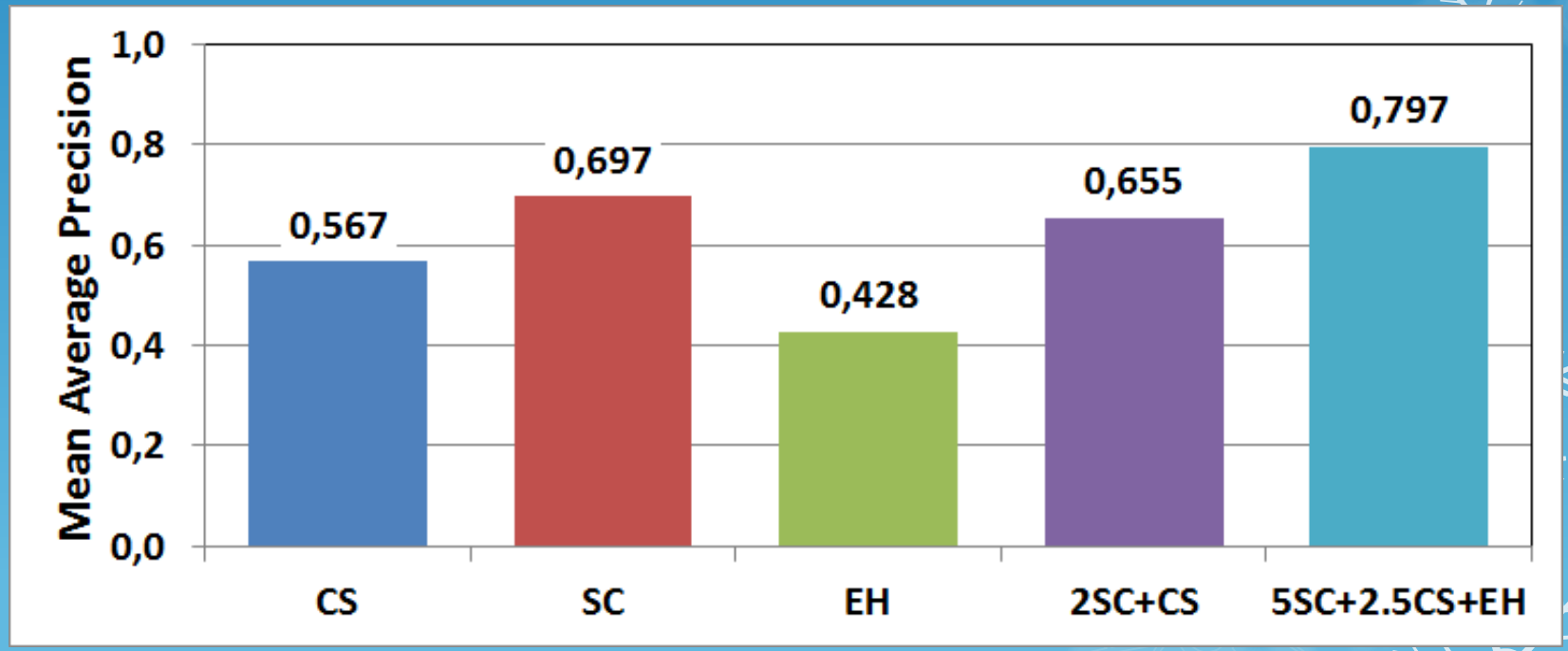
- Figure detector training
  - Positives: 914 person images from MIT dataset
  - Negatives: 1,886 images from INRIA Holidays dataset
- Figure detector quality
  - Tested on ETH dataset
    - 1,201 person figures in 196 images, i.e. ground truth is defined
  - 817 people correctly detected (68%)
  - 1,943 false detections (i.e. 29.6% precision)

# Figure Descriptor

- MPEG7 standard descriptors
  - Extracted from clothing patch region
  - Color structure (CS), scalable color (SC), edge histogram (EH), and their combination

# Figure Descriptor: Recognition

- Figure recognition
  - Tested on ETH dataset
  - Descriptors: Color structure (CS), scalable color (SC), edge histogram (EH), and their combinations



# Clustering Detections

- Face clustering
  - Similar faces form a cluster (distance < 0.14)
- Figure clustering
  - Similar clothing patches (distance < 1.28)
- Face and figure correspondence
  - $\frac{area(R_{face} \cap R_{top\_body})}{min(area(R_{face}), area(R_{top\_body}))} \geq 0.10$
  - $R_{top\_body}$ = top third of the figure region
- Image file IDs are respected
  - Two different detections in the same file cannot be merged

# Experimental Evaluation

- Whole system tested on a small personal album
  - 24 photos, 76 figures, 54 faces, 26 distinct person identities (11 of them repeated)
- Detection results:

  7 correct faces (13%)
  54 correct figures (71%)
  24 false-posit. figs
- Clustering results:

6 - ok (14%)
3 - different people (7%)
7 - mix of false/true positives
27 - only one detection (63%)


More results at http://disa.fi.muni.cz/mmedia2014/

# Conclusions

- Framework for combining face and figure detection for person identification
  - A preliminary implementation
  - Identification using body clothing patches
- Future plans
  - Prepare independent head and upper-body detectors
  - Improve merging detection regions
  - Use personal photos only to train the detectors
  - Test other figure detectors
  - Test better clothing patch extractor (e.g. based on segmentation)

# Experimental Evaluation

- One cluster of different people
  - Clothing patches are very similar

# Experimental Evaluation

- Different clusters of ungrouped detections

# Experimental Evaluation

- A cluster with false and true positive detections
  - Clothing patches are similar