

Facial landmark detection

Michal Uříčář and Vojtěch Franc

[uricamic, xfrancv]@cmp.felk.cvut.cz

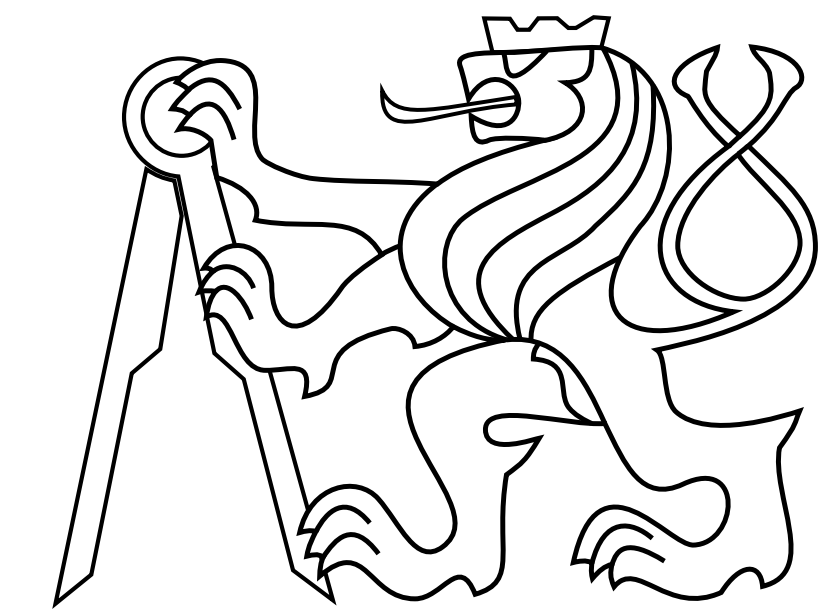
Department of Cybernetics

Faculty of Electrical Engineering, Czech Technical University

<http://cmp.felk.cvut.cz>



CENTER FOR MACHINE
PERCEPTION



Motivation & Overview

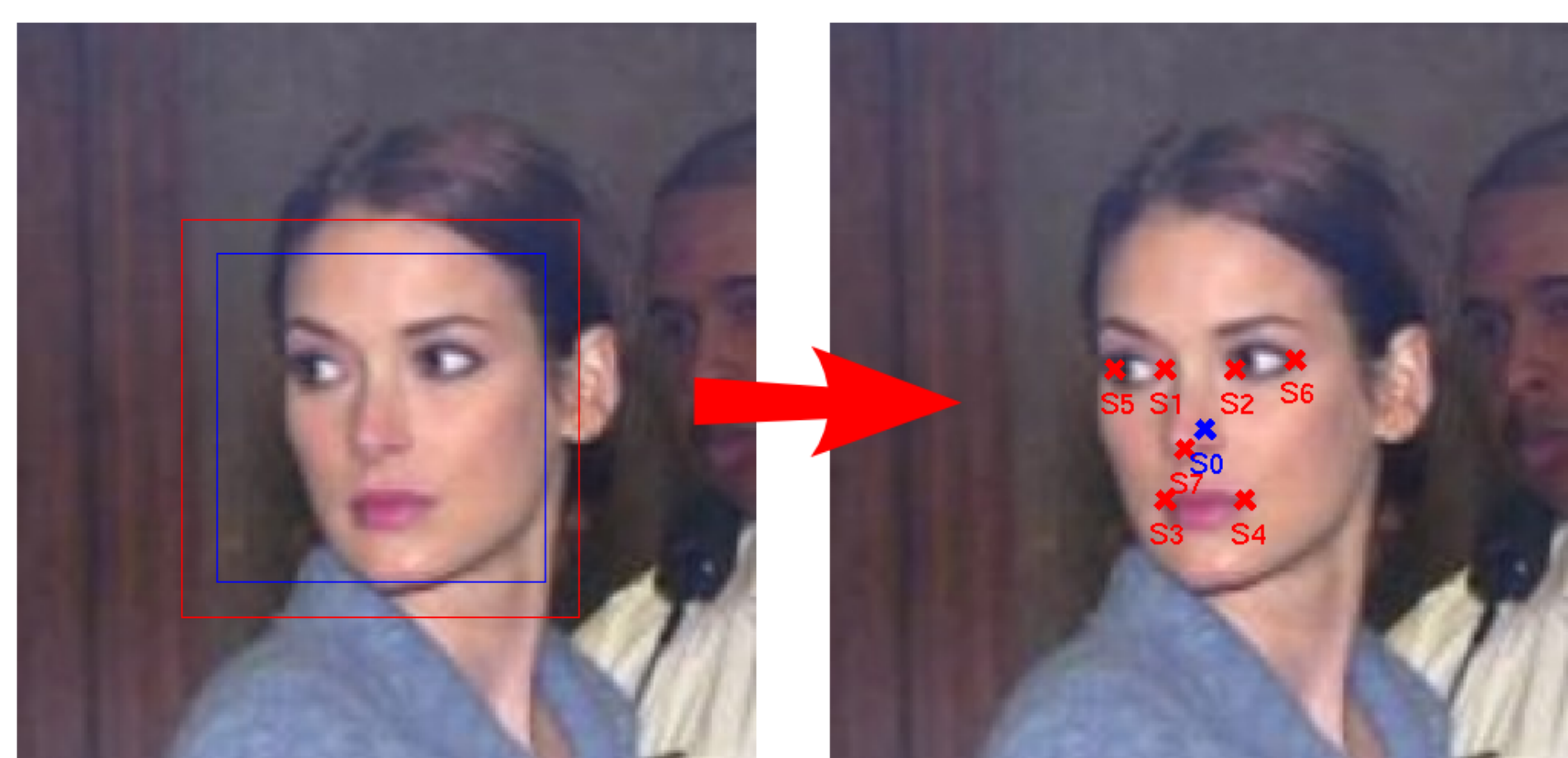


Figure 1: Functionality of detector

- Detection of facial landmarks for enhancement of face localization.
- Essential part of face recognition and further processing.
- Accuracy of face localization strongly influences the quality of recognition.
- High performance open source implementation for academic use
<http://cmp.felk.cvut.cz/~uricamic/flandmark>

Conclusions

- Landmark detector based on the Deformable Part Models (DPM) [1].
- Instance of the structured output classifier.
- Structured Output Support Vector Machines used for learning parameters of the detector.
- Objective function of the learning algorithm directly related to the performance of the resulting detector.
- Outperforms public domain implementations based on Active Appearance Models (AAM) [2] and DPM [3].

Structured Output Classifier

- Quality of a landmark configuration $\mathbf{s} = (s_0, \dots, s_{M-1}) \in \mathcal{S}$ for a given image I is measured by a scoring function $f: \mathcal{I} \times \mathcal{S} \rightarrow \mathbb{R}$.

- Scoring function is defined as a sum of the appearance model and the deformation cost.

- Exact formulation of f is derived from the graph (Fig. 5)

$$f(I, \mathbf{s}) = \sum_{i=0}^{M-1} q_i(I, \mathbf{s}_i) + \sum_{i=1}^{M-3} g_i(\mathbf{s}_0, \mathbf{s}_i) + g_5(\mathbf{s}_1, \mathbf{s}_5) + g_6(\mathbf{s}_2, \mathbf{s}_6) + g_7(\mathbf{s}_0, \mathbf{s}_7)$$

- Maximization of f solved by Dynamic Programming (DP), thanks to the form of graph constraints (directed acyclic graph).

- Loss function contains normalization constant κ which is inversely proportional to the length measured on each face individually.

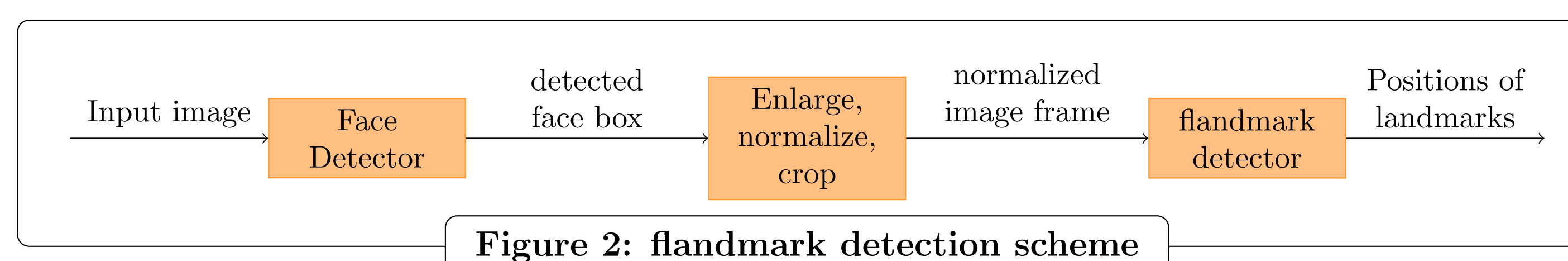


Figure 2: landmark detection scheme

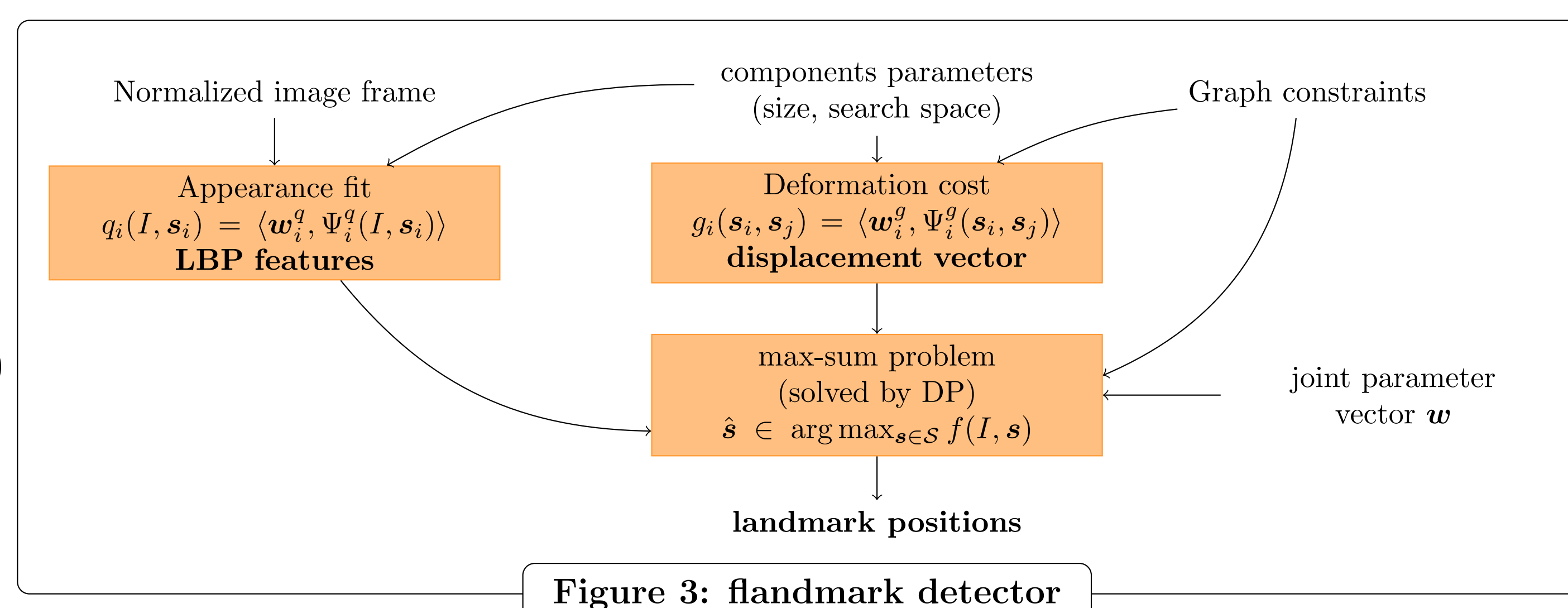


Figure 3: landmark detector

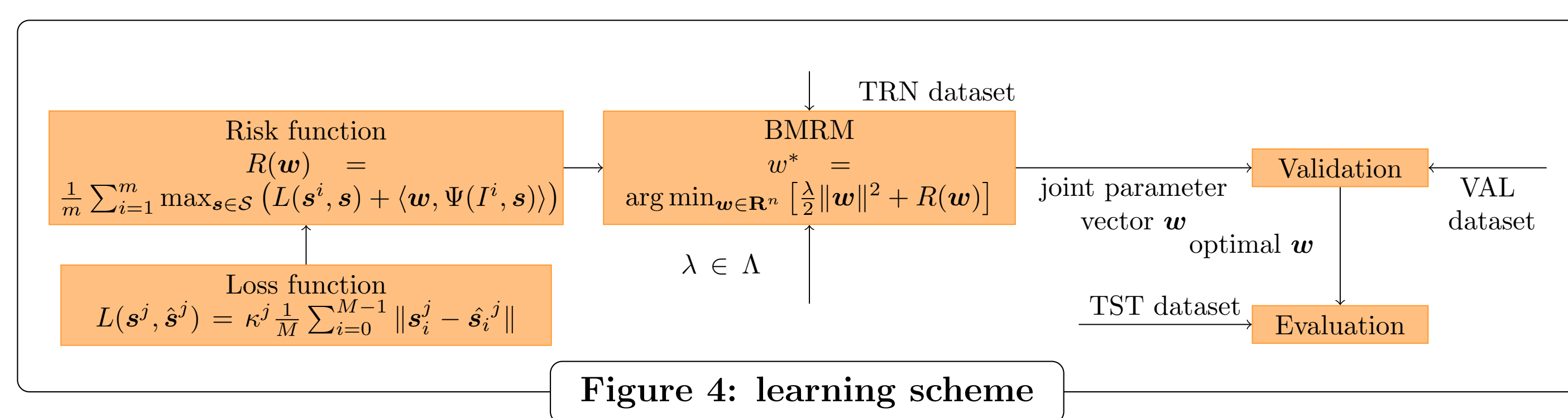


Figure 4: learning scheme

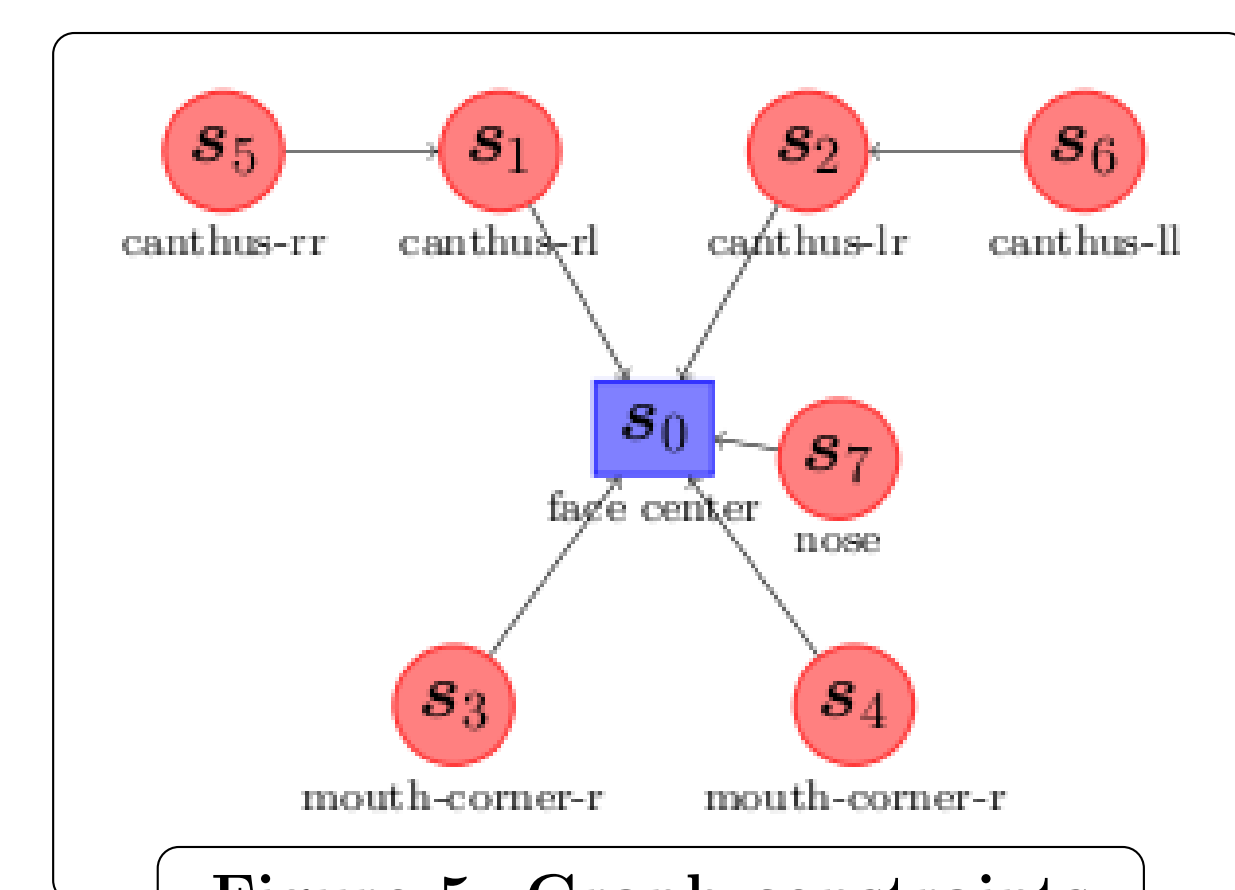


Figure 5: Graph constraints

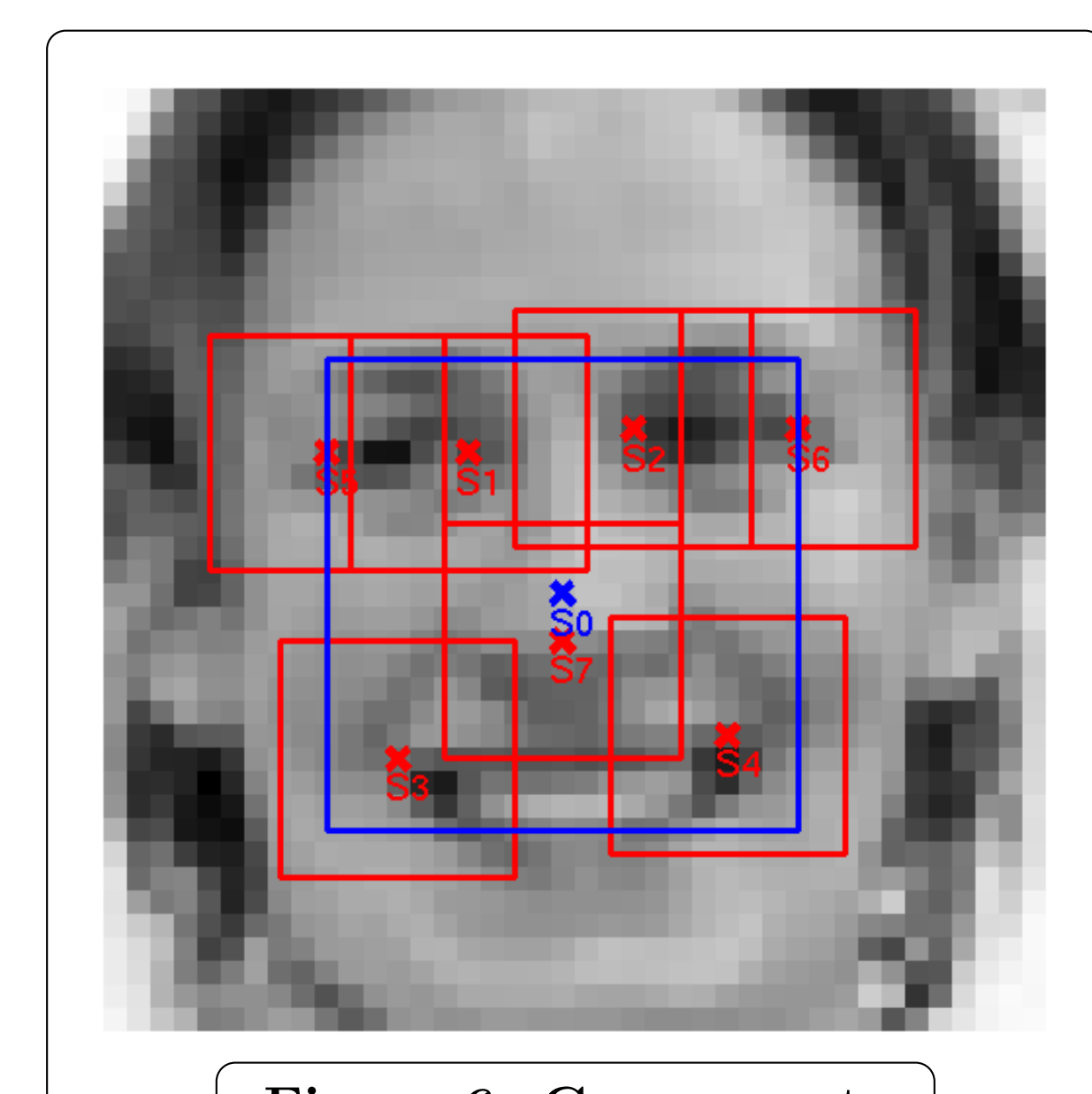


Figure 6: Components

Results

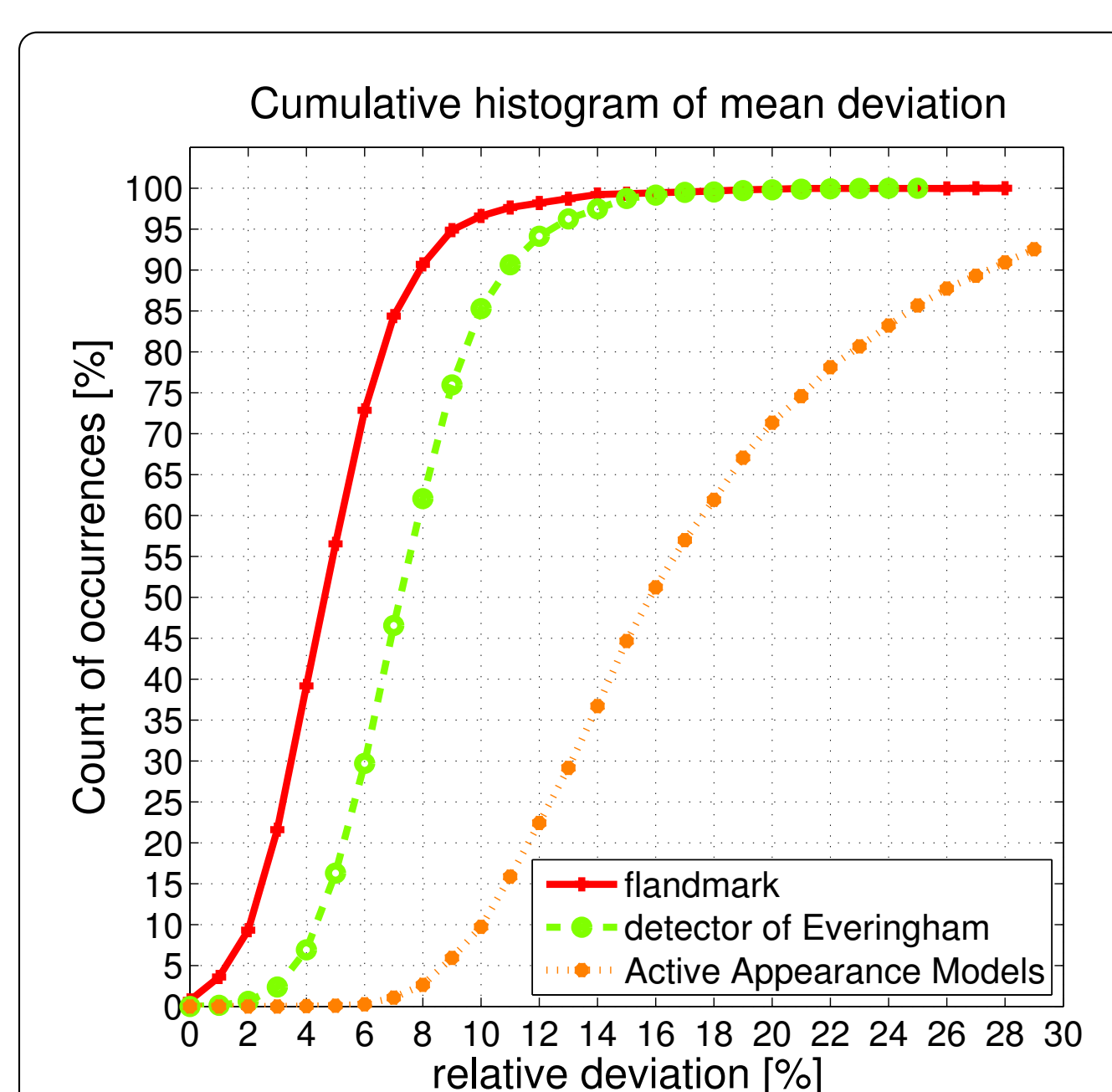


Figure 7: Mean deviation

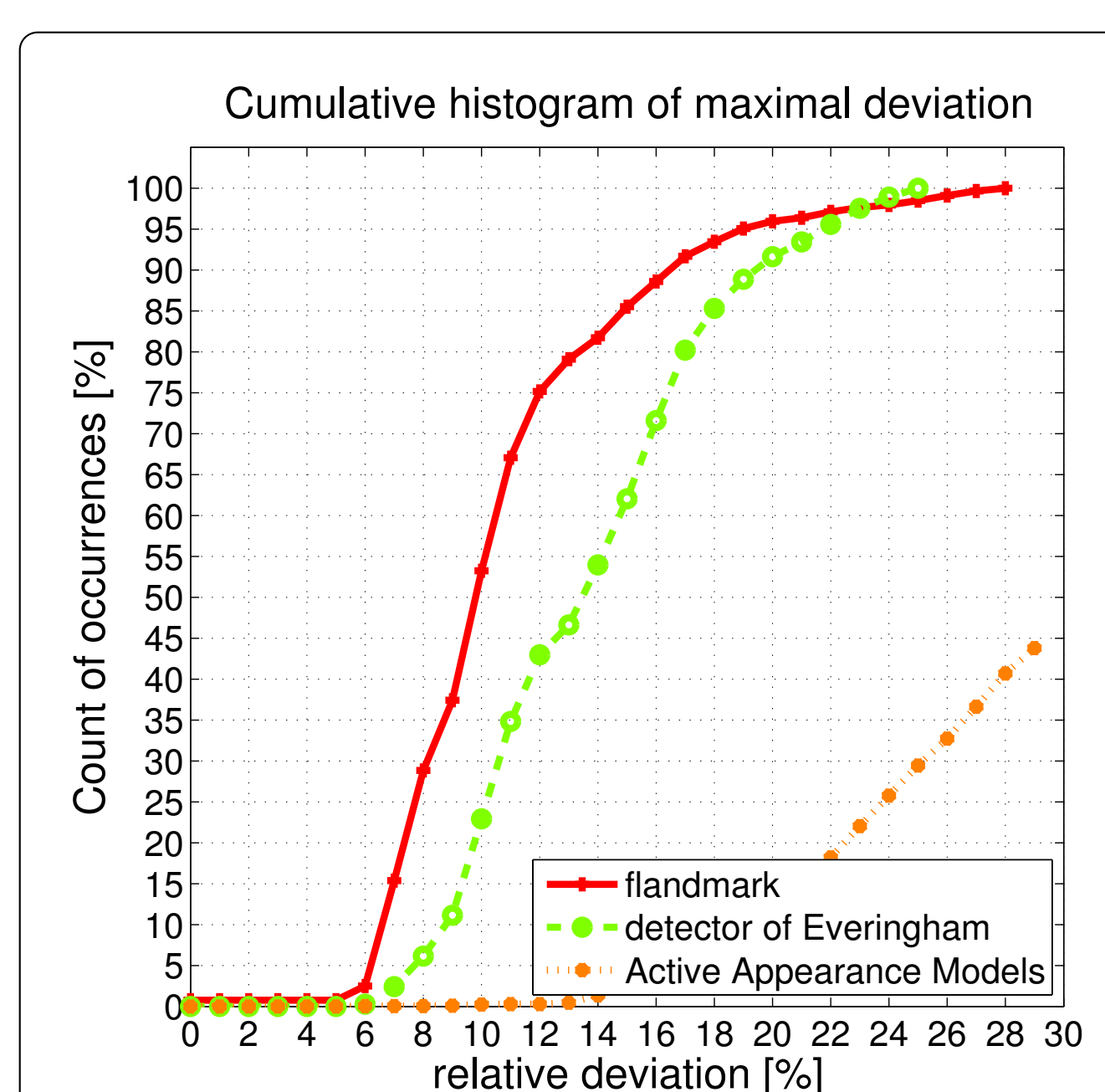


Figure 8: Maximal deviation

Detail 10 % of relative deviation

	mean	maximal
AAM	9.72%	0.22%
Everingham [3]	85.28%	22.93%
flandmark	96.59%	53.23%

- All detectors evaluated on the same TST set.
- Deviation is measured relatively to the distance of the center of eyes and mouth.
- Average time spent on detection — 7ms on a standard PC.

Acknowledgements

The authors were supported by EC projects FP7-ICT-247525 HUMAVIPS and PERG04-GA-2008-239455 SEMISOL.

References

- [1] Fischler, M. A. and Elschlager, R.A., *The representation and matching of pictorial structures*, IEEE Trans. on Computers, 1973.
- [2] Cootes, T., Edwards, G. J. and Taylor, C. J., *Active Appearance Models*, IEEE Trans. Pattern Analysis and Machine Intelligence, 2001.
- [3] Sivic, J., Everingham, M. and Zisserman, A., “Who are you” — learning person specific classifiers from video, In Proc. of IEEE Conference on Computer Vision and Pattern Recognition, 2009.
- [4] Michal Uříčář, *Detector of facial landmarks*, Master’s thesis. CTU in Prague, May 2011.
- [5] M. Uříčář and V. Franc, *Detector of landmarks on human face*, 16th Computer Vision Winter Workshop, February 2011.