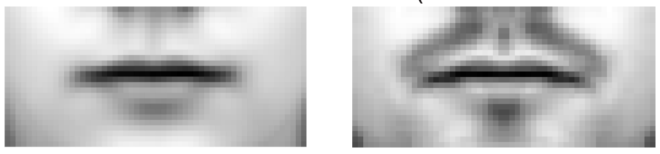
Sparse Versions of Optimized Centroids

Jan Kalina & Petra Vidnerová & Patrik Janáček

Introduction

- Mouth localization in images using centroids

Examples of initial centroids (26 × 56 pixels each)



- 212 gray-scale images of faces
- Centroid-based classification is popular, but lacks sparsity
- We propose several methods for constructing sparse centroids

Motivation for Sparsity

- Computational demands (saving time)
- Energetic demands (saving computational energy)
- Explainability

Centroid-based Object Detection

- Similarity measure - Pearson product-moment correlation coefficient r

$$\arg\min_{x \in \mathbb{E}} r(x, c),$$

c is a centroid and x is an candidate part of the image

- Similarity measure - Euclidean distance [2]

$$\arg\min_{x \in \mathbb{E}} ||x - c||_2$$

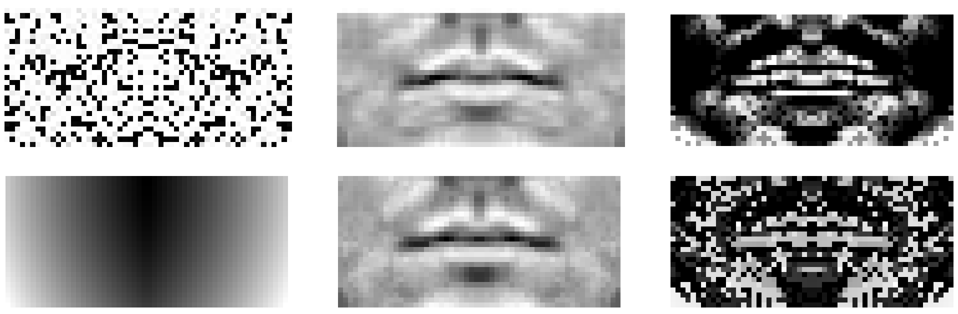
- *Vanilla* approach - centroid is an average of positive examples
- Instead the *vanilla* approach use **optimal centroids** [1]

Methods

- **Weighted approach:** each pixel is assigned a weight $w_i > 0$
- Weights are also optimized
- Several approaches (see the paper):
  - Linear approximation
  - Constrained sparse optimization
  - Evolutionary algorithm
  - Binarized optimal weights
  - Thresholded optimal weights

Illustrative Example

Initial weights: Optimal centroid: Optimal weights:



Experimental Results

Mouth localization accuracy (using test set).

Centroid	Weights	# of used pixels	Localization accuracy
Average	Equal	1456	0.93
Average [2]	Equal [2]	1456	0.90
Optimal	Equal	1456	**1.00**
Optimal	Optimal	932	**1.00**
Optimal	Sparse opt.	1456	**1.00**
Optimal	EA	770	0.88
Optimal	EA	**500**	0.82
Optimal	Binarized optimal	904	0.97
Optimal	Binarized optimal	**500**	0.92
Optimal	Thresholded optimal	904	**1.00**
Optimal	Thresholded optimal	**500**	**0.98**
Average	Binarized equal	536	0.95
Average	Binarized equal	**500**	0.93
Average centroid with r_{LWS}		892	0.96
Viola-Jones [3]	-	1456	**1.00**

Localization accuracy (using test set with added noise).

Method (Section*)	# of pixels	Localization accuracy Noise I	Noise II	Noise III
II-A	1456	0.90	0.88	0.85
II-B	1456	0.87	0.86	0.82
II-C	1456	**1.00**	**0.99**	0.97
III-B	932	**1.00**	0.95	0.96
III-C	500	**1.00**	**0.98**	0.95
III-D	770	0.86	0.81	0.84
III-D	500	0.86	0.81	0.84
III-E	904	0.96	0.92	0.93
III-E	500	0.96	0.92	0.93
III-F	904	**1.00**	**0.98**	**0.99**
III-F	500	**1.00**	**0.98**	**0.99**
IV-A	536	0.93	0.91	0.91
IV-A	500	0.93	0.91	0.91
IV-B	892	0.96	0.89	0.92
Viola-Jones [3]	1456	**0.99**	**0.98**	0.96

* see the paper

Conclusions

- **Sparse** versions of optimal centroids, smaller numbers of pixels
- **Robustness** with respect to noise
- We recommend **thresholded** optimal version

References

J. Kalina and C. Matonoha, "A sparse pair-preserving centroid-based supervised learning method for high-dimensional biomedical data or images", Biocybern. Biomed. Eng., vol. 40, pp. 774–786, 2020.

P. Hall and T. Pham, "Optimal properties of centroid-based classifiers for very high-dimensional data", Ann. Stat., vol. 38, pp. 1071–1093, 2010.

P. Viola and M.J. Jones, "Robust real-time face detection", Int. J. Comput. Vis., vol. 57, pp. 137–154, 2004.

The Czech Academy of Sciences, Prague kalina@cs.cas.cz, petra@cs.cas.cz