

# Detector of Facial Landmarks Learned by the Structured Output SVM

UCU Summer school

July 13, 2017

## 1 Introduction

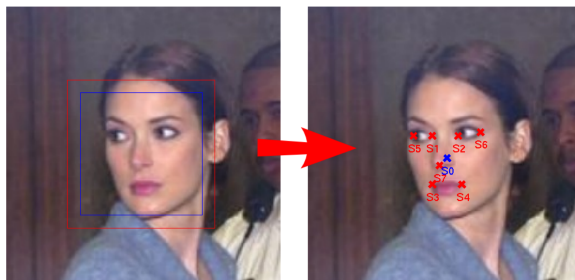


Figure 1: Example of detected landmarks

We will consider a problem of detecting basic set of landmarks from a given normalized image of human face (Fig 2). The inputs are a binary images with given sets of possible hypothesis per each landmark. In Section 2 we will describe structured output classifier for the landmark detection. Your task will be to learn parameters of these classifiers by Bundle Risk Minimization method and to evaluate their performance using data described in Section 3.

## 2 Structured output classifier

Structured output classifier estimates landmark positions from the grey scale image given structure of a landmark's graph and set of possible hypothesis per each landmark by maximizing cost function:

$$f(\mathcal{I}, \mathbf{s}) = \sum_{i=0}^{M-1} q_i(\mathcal{I}, \mathbf{s}_i) + \sum_{i,j \in \mathbb{E}} g_{i,j}(\mathbf{s}_i, \mathbf{s}_j) \quad (1)$$

Graph is shown on the Figure 2.

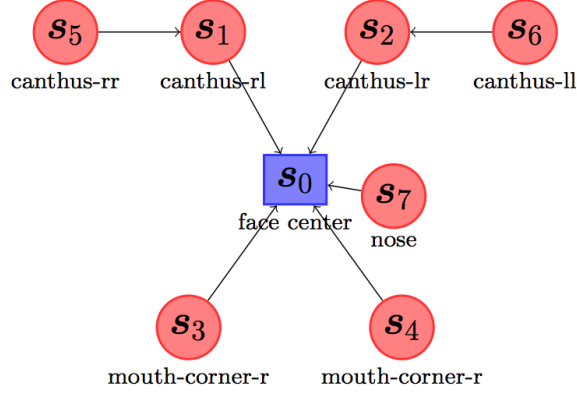


Figure 2: Configuration of  $M$  landmarks described by graph  $G = (V, E)$

The cost function consists of two terms:

- appearance model

$$q_i(\mathcal{I}, \mathbf{s}_i) = \langle \mathbf{w}_i^q, \Psi_i^q(\mathcal{I}, \mathbf{s}_i) \rangle ,$$

where  $\Psi_i^q(\mathcal{I}, \mathbf{s}_i)$  local binary pattern (LBP) features.

- deformation cost

$$g_{i,j}(\mathbf{s}_i, \mathbf{s}_j) = \langle \mathbf{w}_{i,j}^g, \Psi_{i,j}^g(\mathbf{s}_i, \mathbf{s}_j) \rangle ,$$

where  $\Psi_{i,j}^g(\mathbf{s}_i, \mathbf{s}_j)$  is a displacement vector.

$$\Psi_{i,j}^g(\mathbf{s}_i, \mathbf{s}_j) = (dx, dy, dx^2, dy^2) ,$$

$$(dx, dy) = (x_i - x_j, y_i - y_j) .$$

Function  $\Psi_i^q(\mathcal{I}, \mathbf{s}_i)$  crops a square region  $Q_i \times Q_i$  with the center  $\mathbf{s}_i = (x_i, y_i)$  and outputs sparse vector of pyramid LBP features (we provide code for computing LBP features). The set of all possible positions of landmark (hypothesis)  $\mathbf{s}_i$  is given by search window  $S_i \times S_i$ . In this task we consider the following parameters

TODO

To learn proposed structured output classifier you have to implement minimization of the following surrogate loss

$$R(\mathbf{w}) = \frac{1}{m} \sum_{k=1}^m \max_{\mathbf{s} \in \mathcal{A}_0 \times \dots \times \mathcal{A}_{M-1}} \{L(\mathbf{s}, \mathbf{s}^*) + \langle \mathbf{w}, \Psi(\mathcal{I}^k, \mathbf{s}_i) \rangle\} - \langle \mathbf{w}, \Psi(\mathcal{I}^k, \mathbf{s}_i^*) \rangle , \quad (2)$$

where

$$L(\mathbf{s}, \mathbf{s}^*) = \frac{1}{M} \sum_{i=0}^{M-1} \|\mathbf{s}_i - \mathbf{s}_i^*\|^2 .$$

Its input is normalized grey scale image  $40 \times 40$ , M components  $\mathbf{s} = (s_0, \dots, s_{M-1})$  and search space  $\mathcal{A} = (\mathcal{A}_0, \dots, \mathcal{A}_{M-1})$ .

### 3 Dataset

TODO: