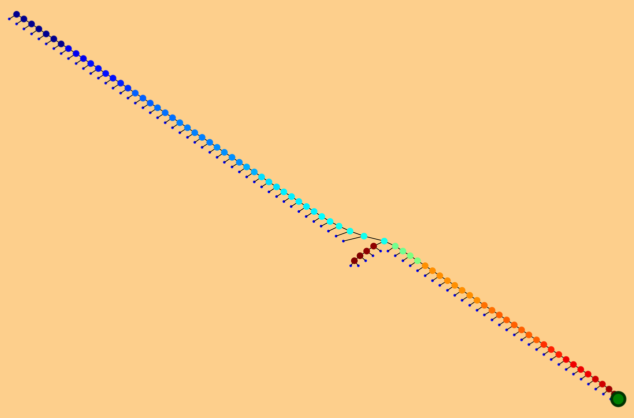


Optimal Landmark Detection using Shape Models and Branch and Bound

{ Brian.Amberg and Thomas.Vetter }@unibas.ch

UNI BASEL

Problem

Fitting statistical 2D and 3D shape models to images is necessary for a variety of tasks, such as video editing and face recognition. Much progress has been made on local fitting from an initial guess, but determining a close enough initial guess is still an open problem. We propose a method to locate fiducial points, which can then be used to initialize the fitting.

Contributions

We overcome the inherent ambiguity in landmark detection by using global shape information. We solve the combinatorial problem of selecting out of a large number of candidate landmark detections the configuration which is best supported by a shape model. Our method, as opposed to previous approaches, always finds the globally optimal configuration.

The algorithm can be applied to a very general class of shape models and is independent of the underlying feature point detector. Its theoretic optimality is shown, and it is evaluated on a large face dataset.

Formulation

The solution is constrained by a shape model

$$M(\mathbf{\Theta}) = (m_1(\mathbf{\Theta}), \ldots, m_N(\mathbf{\Theta})) \qquad (1)$$
$$m_i : \mathbb{R}^{N_\Theta} \to \mathbb{R}^2$$

mapping model parameters $\mathbf{\Theta}$ to image positions $m_i(\mathbf{\Theta})$. For each fiducial point m_i a set of candidate positions

$$L_i = \{l_i^1, l_i^2, \ldots\} \qquad l_i^j \in \mathbb{R}^2 \qquad (2)$$

is detected in the image. The task is to assign to every model vertex one of the candidate positions such that the shape model can be best fit to the selection $\mathbf{S}$, written as a tuple

$$\mathbf{S} = (j_1, j_2, \ldots, j_N) \qquad j_i \in \mathbb{N}, \qquad (3)$$

where j_i is the index of a candidate of landmark i.

So we minimize the distance between the shape model and the image landmarks:

$$\mathbf{S}^* = \operatorname*{arg\,min}_{\mathbf{S}=(j_1,\ldots,j_N)} f(\mathbf{S})$$
$$f(\mathbf{S}) = \min_{\mathbf{\Theta}} \sum_i \rho\left(\left\| m_i(\mathbf{\Theta}) - l_i^{j_i} \right\|\right) \quad . \qquad (4)$$

Where $\rho : \mathbb{R} \to \mathbb{R}$ is a robust function, allowing us to handle missing detections, and points which are invisible due to occlusion.

References

[1] B. Amberg, T. Vetter. Optimal Landmark Detection using Shape Models and Branch and Bound In *ICCV '11*

Results

Some randomly chosen images from the color feret database for each pose, and the detected landmark positions. The first two rows are success cases, the last row shows a failure case.

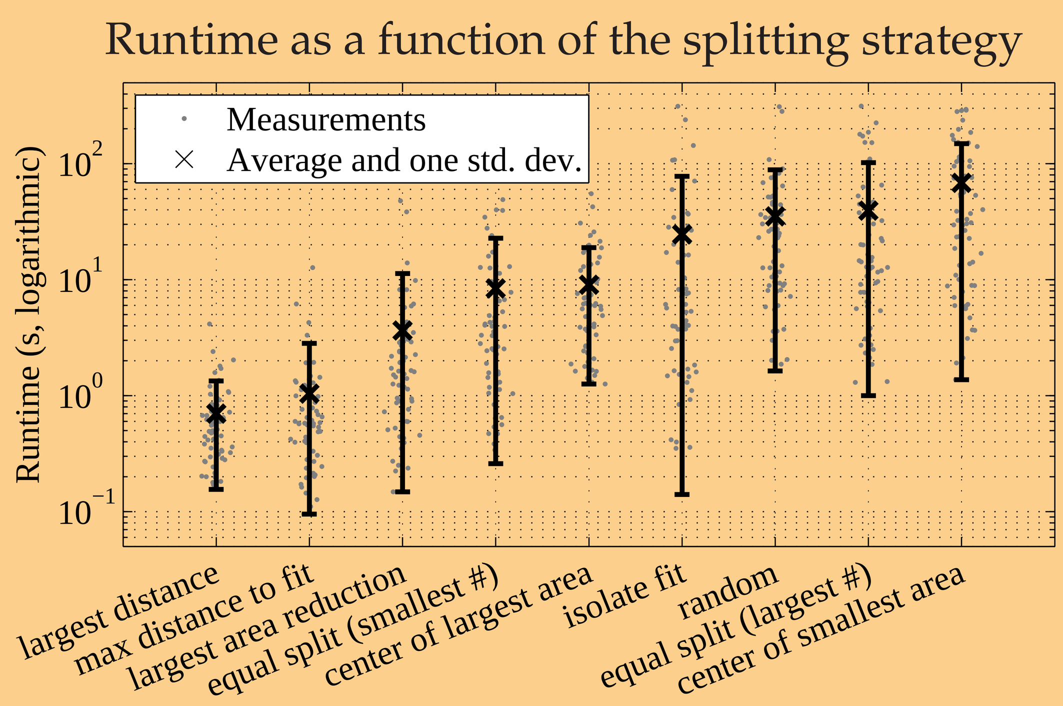
Solution

This discrete optimization is solved by Branch and Bound, which is a method to minimize a function over a set. It requires us to (1) efficiently specify solution subsets, (2) determine a lower bound on the minimal cost of the solutions within a subset, and (3) specify a strategy to split a solution subset into two new subsets.

The ingredients in our case are:

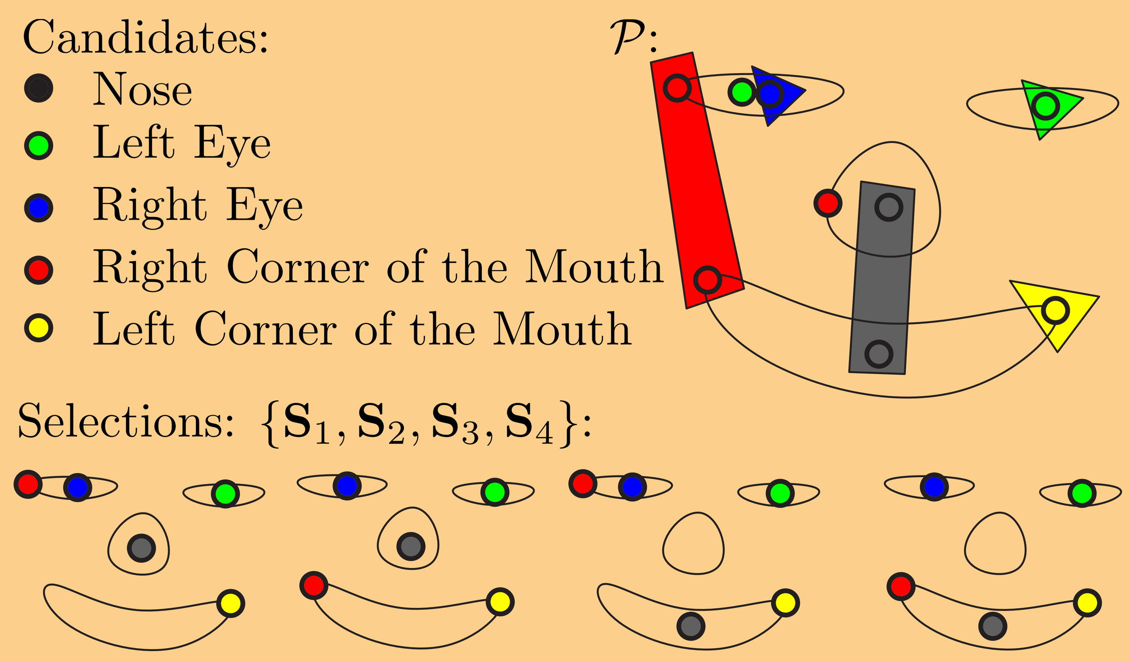
1. Solution subsets are created by taking subsets of landmark candidates, and considering the Kartesian product of all selected landmark candidates
2. We bound the cost for such a solution set by taking for each landmark the minimal distance to the convex hull of the selected candidates
3. We found that splitting landmark candidates such that the convex hull of the resulting two landmark candidates are as distant as possible is most effective.

Splitting Strategy



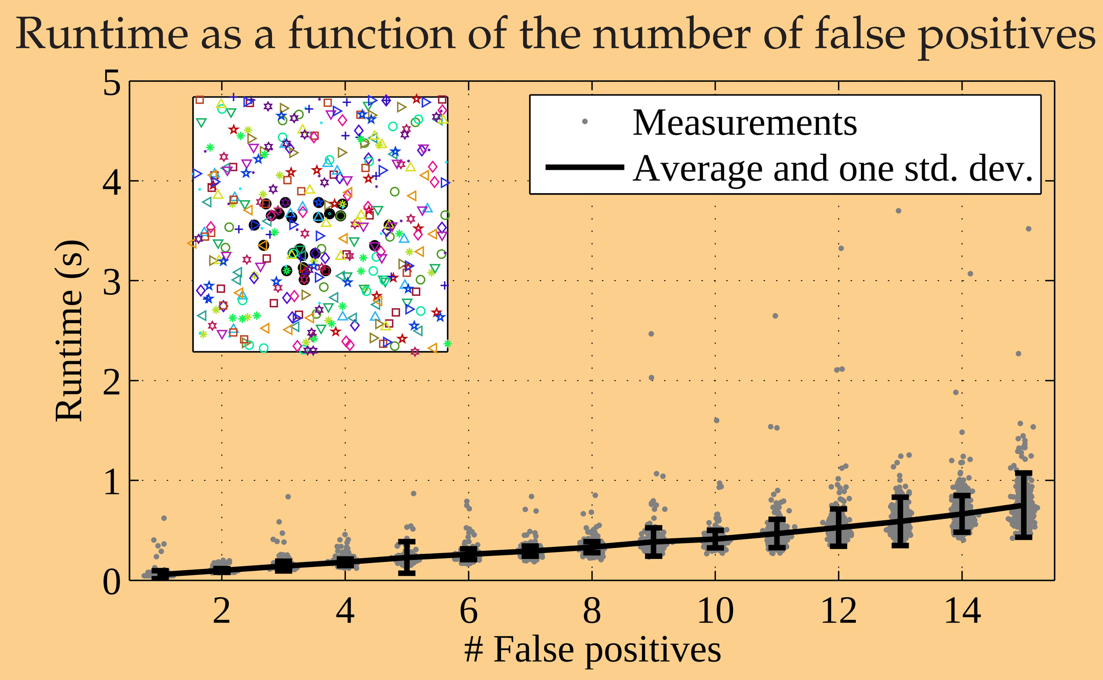
Different splitting strategies result in vastly different performance. Note that 'split into equal sized problems' is one of the worst strategies for branch and bound.

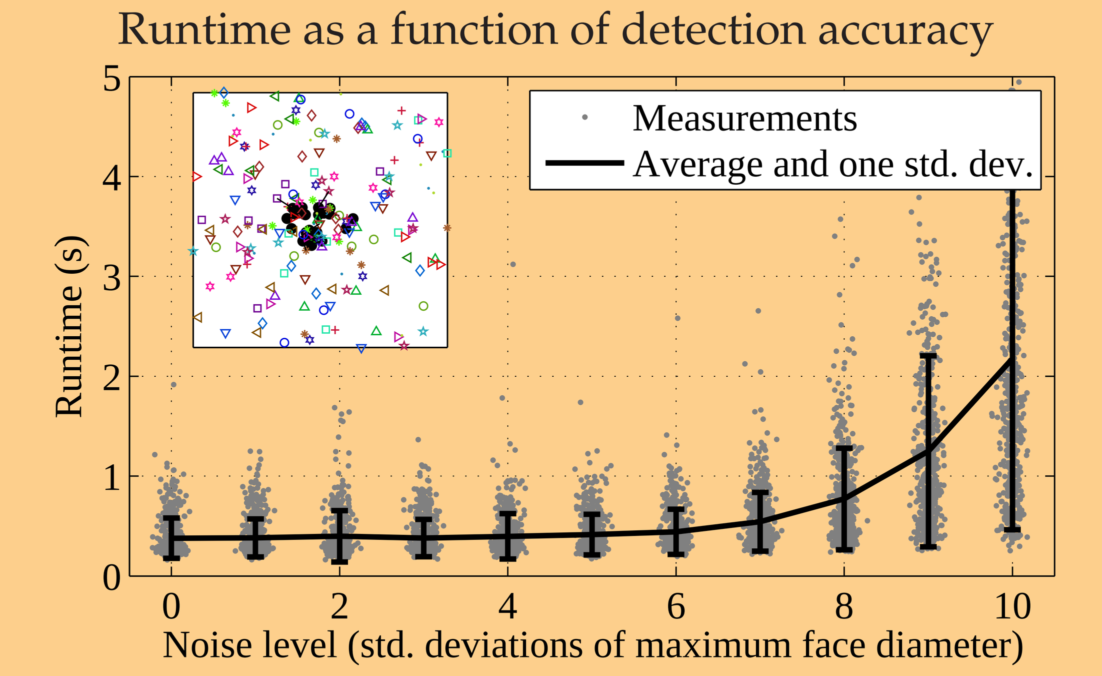
Representation



Subsets of solutions are encoded as the Kartesian product of subsets of landmark candidates per fiducial point.

Scaling Behaviour





Source Code

The source code is available at `http://www.cs.unibas.ch/personen/amberg_brian/bnb/`