Branch-and-Mincut: Global Optimization for Image Segmentation with High-Level Priors

Marcel Cordoví Solés

IN2107 - Image Segmentation and Shape Analysis

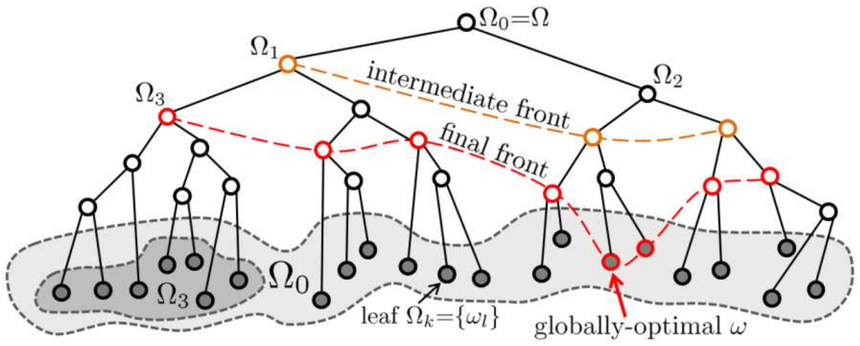
Problem Statement

GOAL

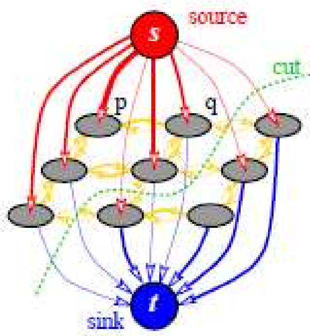
Applying st-mincut with localized lowlevel cues

- A large discrete set of non-local parameters as the space representing the prior knowledge Ω is defined.

- An energy based on both, the graph partition and the non-local parameters is defined.



- The Branch and Bound algorithm guaranties to find the energy's global minima.

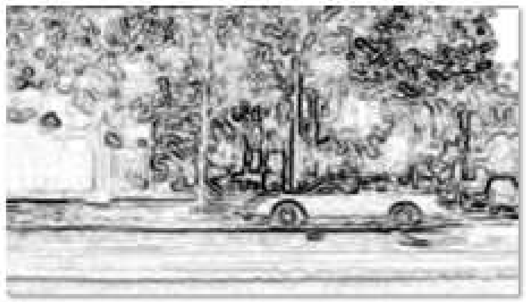


- The st-mincut algorithm is used within Branch-and-Bound search to efficiently compute lower bounds over the tree branches.

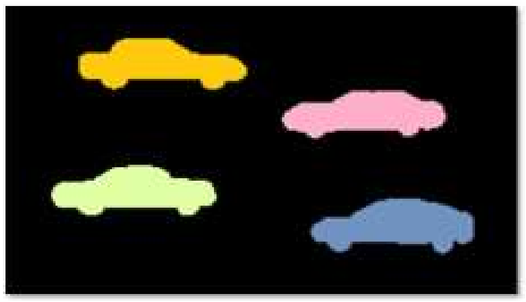
Branch-and-Mincut

Application to the UIUC car dataset

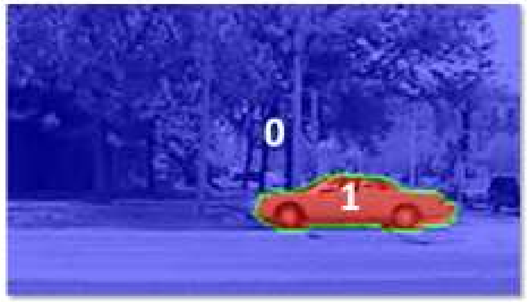
Input image



Local cues encoded on a graph.



Shape prior. Parameterized by (x, y, shape)



Global optimum

General Energy Function

$$E(\boldsymbol{x},\omega) = C(\omega) + \sum_{p\in\mathcal{V}} F^p(\omega)\cdot x_p + \sum_{p\in\mathcal{V}} B^p(\omega)\cdot(1-x_p) + \sum_{(p,q)\in\varepsilon} P^{pq}(\omega)\cdot|x_p - x_q|$$

Constant: $C(\omega)$

Data term: $\sum_{p\in\mathcal{V}} F^p(\omega)\cdot x_p + \sum_{p\in\mathcal{V}} B^p(\omega)\cdot(1-x_p)$

Pairwise potential defining the cost of assigning different labels to adjacent pixels: $\sum_{(p,q)\in\varepsilon} P^{pq}(\omega)\cdot|x_p - x_q|$

where,

$\omega \in \Omega$:
- Shape prior case: product space of various poses and deformations
- Color prior case: set of parametric color distributions

$\boldsymbol{x} \in 2^{\mathcal{V}}$: Is a segmentation, a labeling assignment to the pixels in $\mathcal{V}$

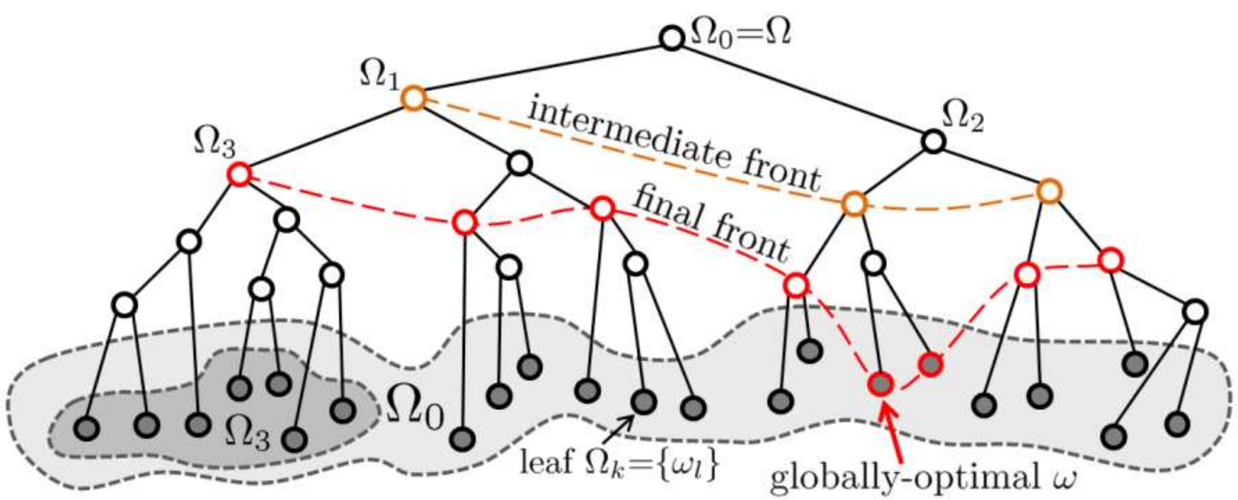
$\varepsilon \subset \mathcal{V}\times\mathcal{V}$: Defines a vecinity of the pixel p.

The domain Ω is hierarchically clustered in a tree of subregions: $T_\Omega = \{\Omega_0, \Omega_1, \dots, \Omega_N\}$

Intermediate nodes correspond to subregions of the domain: $\Omega_{child(k)} \subset \Omega_k$

Leafs corresponds to the singleton subsets: $\Omega_{leaf} = \{\omega_t\}$

A best-first branch and bound search is performed on T_Ω by mantaining an active front that moves towards the leafs decreasing the lowest lower bound at each step.



For applying this framework we need to:

- Define an energy function on $2^{\mathcal{V}} \times \Omega$.
  $(F^p(\omega), B^p(\omega), P^{pq}(\omega))$
- A lower bound function for the energy.
- Build the Branch and Bound tree on Ω .

Let $L(\Omega)$ be the lower bound of $E(\boldsymbol{x}, \omega)$ in the domain $2^{\mathcal{V}} \times \Omega$. We can derive it as follows:

$$\min_{x \in 2^{\mathcal{V}}, \omega \in \Omega} E(\boldsymbol{x}, \omega) = \min_{x \in 2^{\mathcal{V}}} \min_{\omega \in \Omega} E(\boldsymbol{x}, \omega) \geq$$

$$\min_{x \in 2^{\mathcal{V}}} \left[\min_{\omega \in \Omega} C(\omega) + \sum_{p \in \mathcal{V}} \min_{\omega \in \Omega} F^p(\omega) \cdot x_p + \sum_{p \in \mathcal{V}} \min_{\omega \in \Omega} B^p(\omega) \cdot (1 - x_p) + \sum_{p,q \in \varepsilon} \min_{\omega \in \Omega} P^{pq}(\omega) \cdot |x_p - x_q| \right]$$

Which have three esential properties that makes it suitable for the Branch-and-Mincut framework:

- **Monotonicity:** for nested domains of non-local parameters $\Omega_2 \subset \Omega_1$ it holds that:

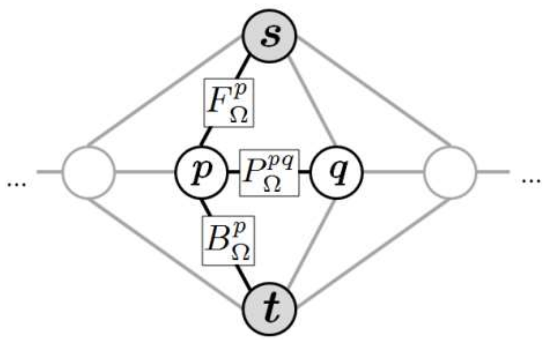
$$L(\Omega_2) \leq L(\Omega_1)$$

- **Computability:** The bound equals the minimum of a quadratic submodular boolean function which can be realized on a network graph such that each configuration of the binary variables is in one-to-one correspondence with an st-cut of the graph having weight equal to the value of the function. $L(\Omega)$ can then be calculated in polynomial time on $|\boldsymbol{\mathcal{V}}|$.

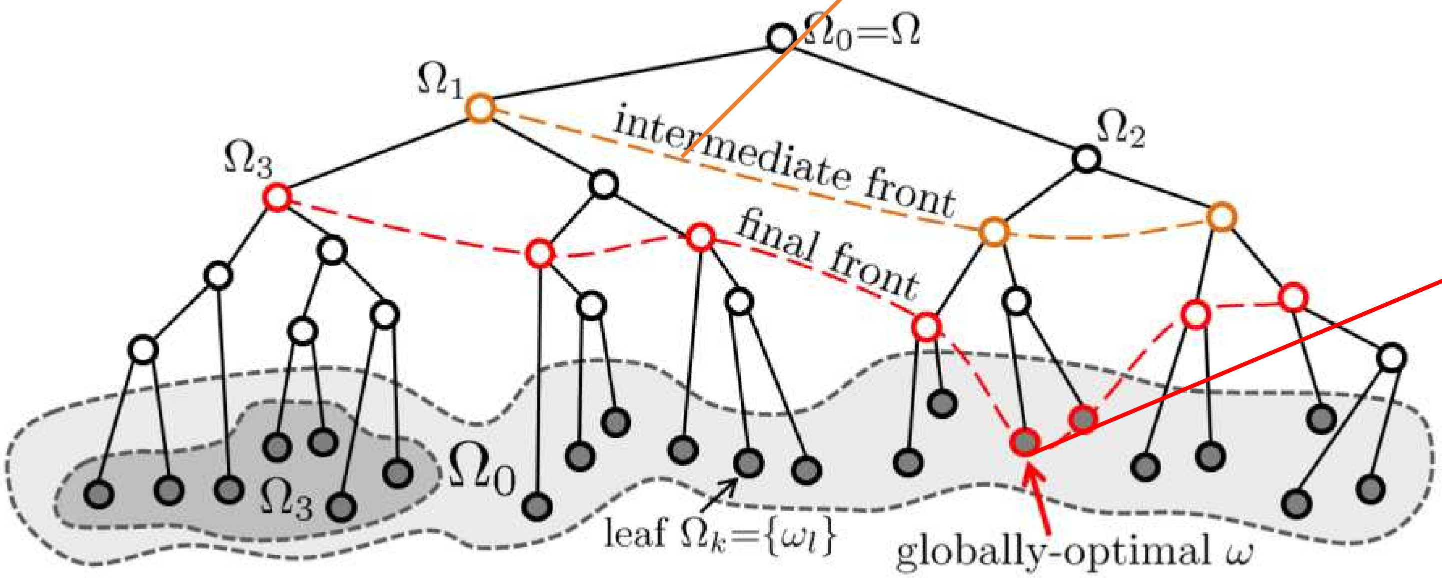
- **Tightness:** for each singleton Ω, it holds that:

$$L(\{\omega\}) = \min_{x \in 2^{\mathcal{V}}} E(\boldsymbol{x}, \omega)$$

In which case the the minimal st-mincut yields the optimal segmentation for ω.



Due to the monotonicity: $L(\Omega_{child(k)}) \leq L(\Omega_k)$
And the lowest lower bound of the active front constitutes a lower bound on the global optimum over the entire domain.



The first leaf $\{\omega'\}$ reached has the lowest bound in the front and hence of the whole domain. Due to the tightness property, $E(\boldsymbol{x}', \omega') = L(\omega')$ which is consequently a global minimum of the energy function.

The prior is defined as the set of exemplar binary segmentations

$$\Omega = \{\boldsymbol{y}^{\omega} \in 2^{\mathcal{V}}\}$$

Were Ω is the discrete set of exemplar segmentations.

The following functional is introduced:

$$E_{prior}(\boldsymbol{x}, \omega) = \rho(\boldsymbol{x}, \boldsymbol{y}^{\omega}) = \underbrace{\sum_{p \in \mathcal{V}} \overbrace{(1 - y_p^{\omega})}^{F^p(\omega)} \cdot x_p + \sum_{p \in \mathcal{V}} \overbrace{y_p^{\omega}}^{B^p(\omega)} \cdot (1 - x_p)}_{\text{Joint prior over the segmentation and the non-local parameter}} + \underbrace{\sum_{p,q \in \varepsilon} \overbrace{\lambda \frac{e^{-\frac{\|\boldsymbol{\kappa}_p - \boldsymbol{\kappa}_q\|}{\sigma}}}{|p - q|}}^{P^{pq}(\omega)} \cdot |x_p - x_q|}_{\text{Contrast sensitive edge term}}$$

ρ denotes the hamming distance Between segmentations.

Joint prior over the segmentation and the non-local parameter

Contrast sensitive edge term

The domain Ω is factorized into the cartesian product of two sets:

$$\Omega_{shape} = \Delta \times \Theta, \qquad \omega = \delta \times \theta$$

Where Δ indexes the set of all segmentations centered at the origin corresponding to variations in scale, orientation and non-rigid deformations.

And Θ corresponds to shift transformations and ensures the translation invariance of the prior.

The branch and bound prior tree $T_\Omega = \{\Omega_0, \Omega_1, \dots, \Omega_N\}$ is generated by combining the two factor trees:

$$T_\Delta = \{\Delta_0, \Delta_1, \dots, \Delta_N\} \qquad T_\Theta = \{\Theta_0, \Theta_1, \dots, \Theta_N\}$$

for which the concept of *loseness* is introduced.

The loseness of a nodeset Ω_t is defined as the number of pixels that change their mask value under different shapes in Ω_t.

$$\Lambda(\Omega_t) = \left| \left\{ p \mid \exists \delta_1, \theta_1 \colon y_p^{\delta_1 \times \theta_1} = 0 \;\; \wedge \;\; \exists \delta_2, \theta_2 \colon y_p^{\delta_2 \times \theta_2} = 1 \right\} \right|$$

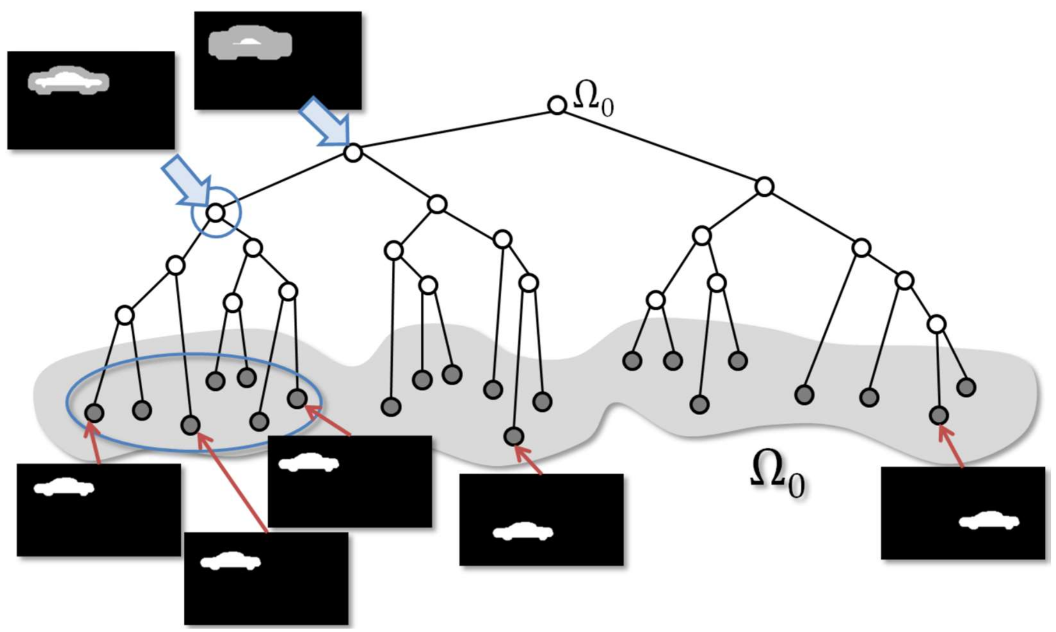
Given the two possible node splits along each of the dimensions:

Split along the shape dimension:
$\Delta_{ch1(p(t))} \times \Theta_{q(t)}$ and $\Delta_{ch2(p(t))} \times \Theta_{q(t)}$

Split along the shape dimension:
$\Delta_{p(t)} \times \Theta_{ch1(q(t))}$ and $\Delta_{p(t)} \times \Theta_{ch2(p(t))}$

The one that minimizes the sum of loseness is prefered.

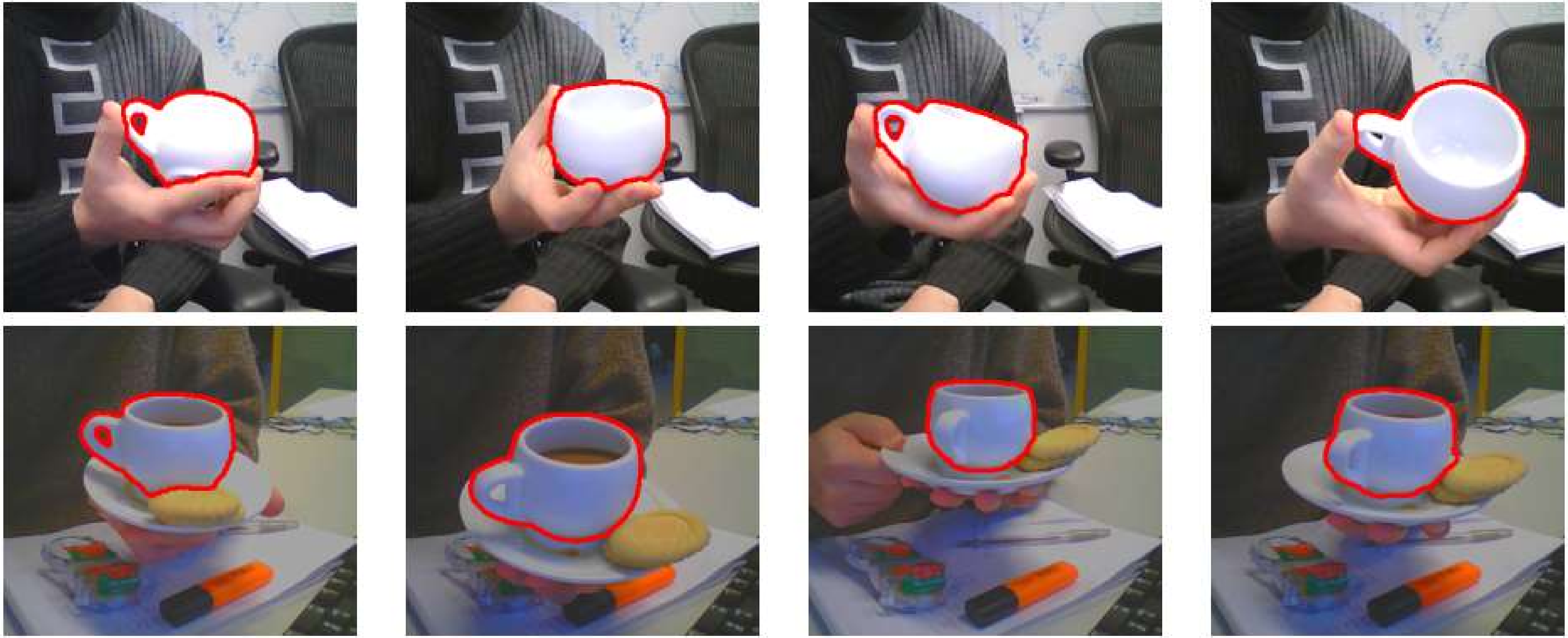
Hierarchical clustering of similar segmentations

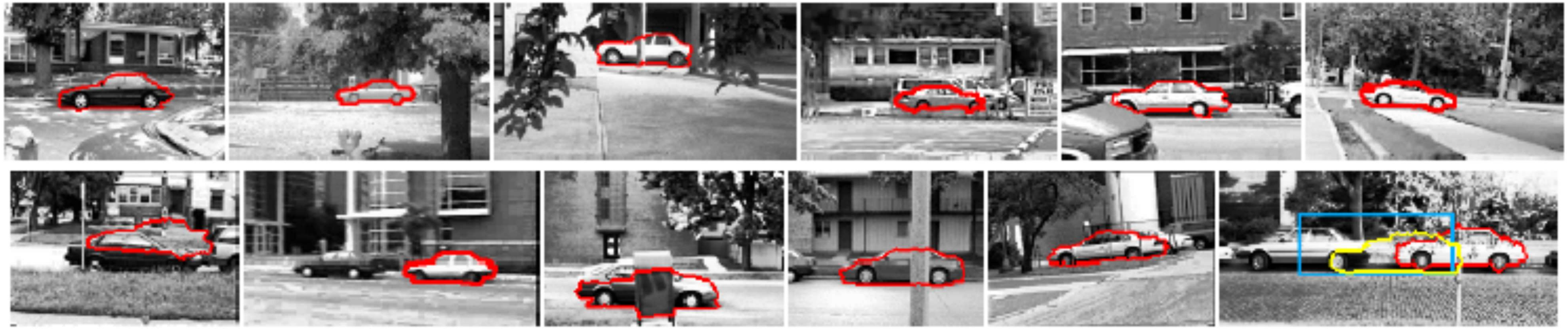


- Each Node correspond to a set of shapes.
- The set of leafs corresponds to the set of binary masks defining the shape prior.
- The image shows the potentials B_{Ω} and F_{Ω} of a node.
- The gray pixels corresponds to the amount of loseness of the mask.

Branch-and-Mincut

Experimental results





Questions?