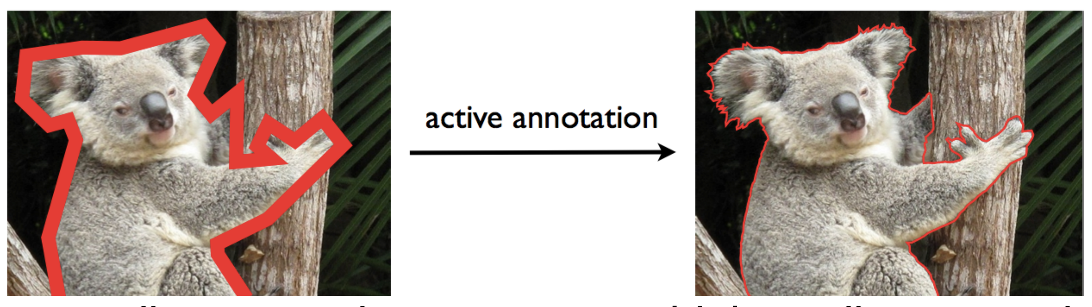
Active Boundary Annotation via Random MAP Perturbations

Subhransu Maji — TTI Chicago
Tamir Hazan — University of Haifa, Israel
Tommi Jaakkola — MIT

1. Overview

Goal: Obtain high quality image annotation with low cost (annotation effort)



low quality annotation — high quality annotation

Approach: Bayesian active learning

- Minimize uncertainty in the boundary of MAP prediction
- Tradeoff *uncertainty* reduction and *cost* of annotation

Contributions

- Entropy bounds that measure the expected perturbation that change MAP prediction.
- Coarse to fine approach for pixel-accurate annotation that saves 33% in cost.

2. Active learning in structured spaces

Traditional Active learning

- Active learner picks which data points to label. Typically assume data is i.i.d.

Bayesian active learning in structured spaces

- Deals with correlated labels, e.g. labels of a single image (non i.i.d. setting)
- **Basic idea:** Construct a probability function over the label space and reduce its uncertainty with minimal annotation cost (clicks)

3. Active annotation framework

Approach

- Let, $y = (y_1, ..., y_n)$ be the set of labels for image x for n pixels
- Let, $A_t = \{a_1, ..., a_t\}$ be the set of *annotations* obtained till time *t*
- Let, p(y) be the joint probability of the labels given the data x and *annotations* till time *t*

Bayesian experimental design

- Given:
  - a function that measures the uncertainty of the labels given the annotation, *U(A)*
  - a function that measures the cost of annotation, *C(a)*
- Pick the annotation task the provides the highest uncertainty reduction/unit cost, i.e.,:

$$a_t = \arg\max_{a \in \mathcal{A}} \frac{U(A_{t-1}) - U(A_{t-1} \cup a)}{C(a)}$$

- Uncertainty, *U(A) = H (p),* is defined as the entropy

$$H(p) = -\sum_y p(y) \log p(y)$$

- Computing entropy is exponential in the size of the patch. for many useful cases, however MAP estimation is tractable for some of these (e.g., via Graph-cuts, MPLP)

$$(MAP) \qquad \arg\max_{y_1,...,y_n} \theta(y_1, ..., y_n; x, A_t)$$

4. Markov Random Fields (MRFs) for image labeling

- Popular for image segmentation (e.g. Grabcut model, *Blake et al., 2004*)
- Let an annotation of an n pixel image be described as a n-tuple $y = (y_1, ..., y_n)$, $\{-1, 1\}$
- The overall score of the pixel label is given by:

$$\theta(y_1, ..., y_n; x) = \sum_{i=1}^n \theta_i(y_i; x) + \sum_{(i,j) \in E} \theta_{i,j}(y_i, y_j; x)$$

$\log P(y_i|x)$ — **Local potentials**: Gaussian mixture models on LAB space

$\theta_{i,j}(y_i, y_j; x) = \exp(-(\bar{x}_i - \bar{x}_j)^2) y_i y_j$ — **Pairwise potentials**: smoothness contraints on adjacent pixels

- The MAP estimate can be obtained via. Graph cuts (*Boykov et al., 2001*)

5. MAP perturbations

The Perturb MAX model (*Papandreou and Yuille, 2011, Tarlow 2012, Gane 2014*)

- Random functions

$$\gamma_i : Y_i \to R \text{ for every pixel } i$$

$$p(\hat{y}) = P_\gamma\left(\hat{y} = \arg\max_y \{\theta(y) + \sum_{i=1}^n \gamma_i(y_i)\}\right)$$

Image + initial boundary — Per-pixel entropy bound

MAP perturbations upper bound the partition function (Hazan & Jaakkola 2012)

- Let $\{\gamma_i(y_i)\}$ be i.i.d. Gumbel random variables with zero mean

$$\log\left(\sum_y \exp(\theta(y))\right) \le E_\gamma\left[\max_y \{\theta(y) + \sum_{i=1}^n \gamma_i(y_i)\}\right]$$

6. Measuring uncertainty in the boundary of MAP prediction

For Perturb MAX models with Gumbel random variables

$$H(p) \le E_\gamma\left[\sum_{i=1}^n \gamma_i(y_i^*)\right]$$

- Where,

$$y^* = \arg\max_y \{\theta(y) + \sum_{i=1}^n \gamma_i(y_i)\}$$

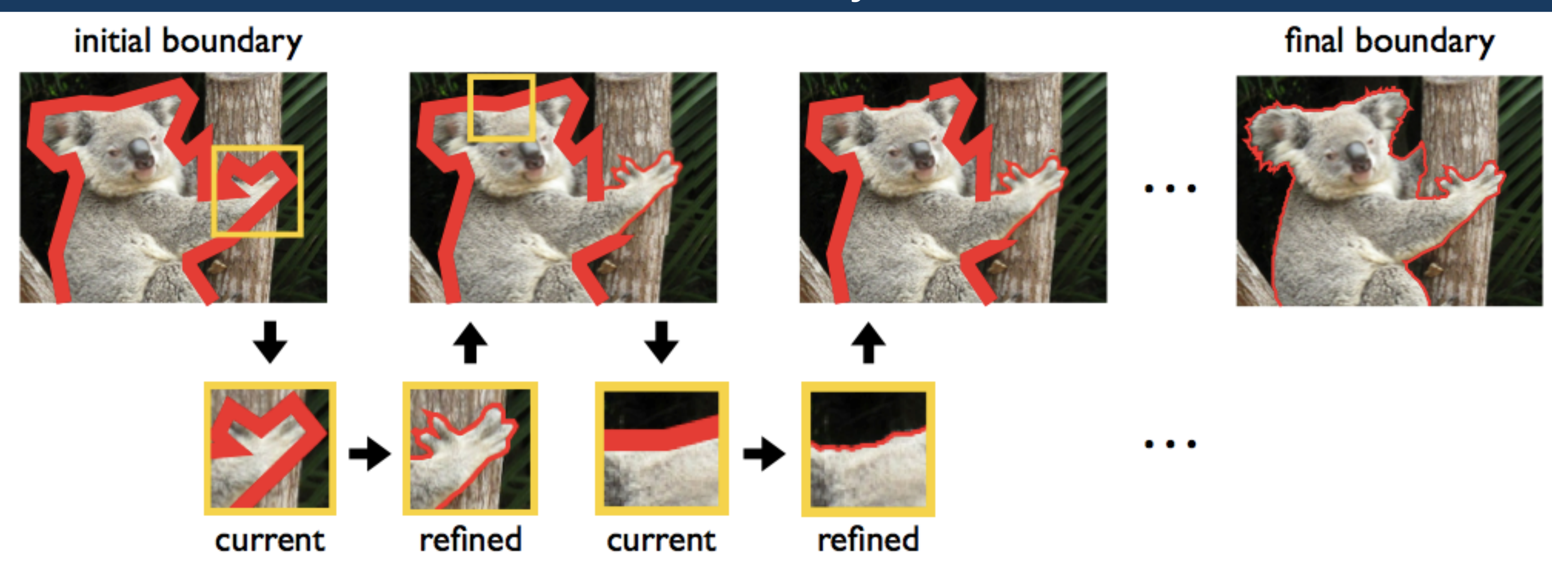
Proof idea:

- Conjugate duality: $H(p) = \min_\theta \left\{ \log\left(\sum_y \exp(\theta(y))\right) - \sum_y p(y)\theta(y) \right\}$
- Use MAP perturb. upper bounds. $H(p) \le \min_\theta \{E_\gamma[\max_y\{\theta(y) + \sum_{i=1}^n \gamma_i(y_i)] - \sum_y p(y)\theta(y)\}$
- The optimal theta attains the perturb-max model p(y).
- The linear term cancels out.

Uncertainty measure $E_\gamma\left[\sum_{i=1}^n \gamma_i(y_i^*)\right]$

- Nonnegative (upper bounds the entropy).
- Attains its minimal value for the zero-one distribution (zero mean perturbations).
- Attains its maximal value for the uniform distribution (symmetry).

7. Active boundary annotation

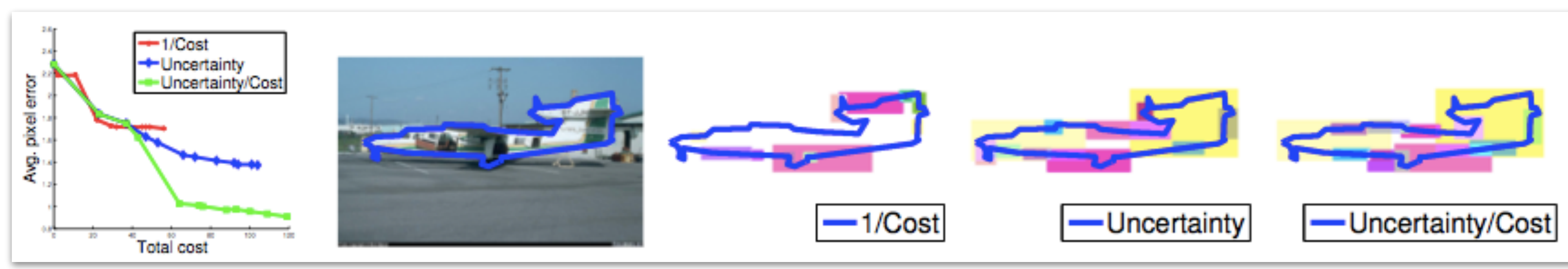


Coarse-to-fine boundary refinement

- We start from a coarse boundary and repeatedly the
  - *regions* are picked by the algorithm, *refinement* is done by the user
  - Cost of refinement = number of points in the polygons (boundary complexity)
- We don't know the truth, so we can compute expectations of *cost* and *uncertainty*

8. Experimental evaluation

An example coarse-to-fine refinement (sampled regions for various strategies)



Active annotation results

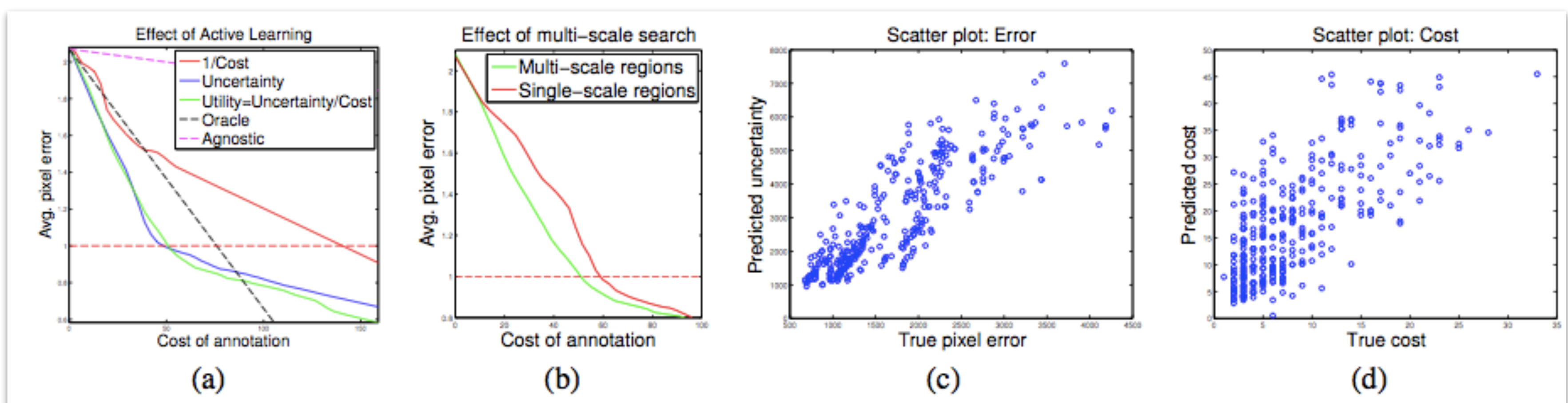


Figure 5: (a) *Error vs. cost tradeoffs for various annotation methods.* Using active learning we are able to annotate the boundaries within an avg. error of 1 pixel at about 66% of the cost required for annotating using the oracle. (b) *Effect of multi-scale.* Multi-scale selection of regions provide a saving of about 20% in cost over single-scale regions. (c,d) *Quality of error and cost estimation.* Scatter plot of predicted and true error (c), and same for the cost (d).

9. Conclusions and future work

We proposed a new uncertainty measure

- Avoids expensive MCMC sampling by randomly perturbing the model and using a MAP solver as a black box tool.
- Applications for parameter estimation and active learning in a number of areas such as matchings, parse trees, and other combinatorial structures.

Active learning in structured spaces

- Sampling based approach allows us to consider non-decomposable cost functions. For the boundary annotation task we used boundary complexity, which is not possible to compute with marginal estimates.
- This led to 33% savings in annotation time for pixel-accurate boundary annotations.

Challenges

- MAP perturbation based entropy bounds for higher dimensional perturbations.
- Beyond super-modular functions in the context of active learning.